



US 20160298151A1

(19) **United States**

(12) **Patent Application Publication**

Butt et al.

(10) **Pub. No.: US 2016/0298151 A1**

(43) **Pub. Date: Oct. 13, 2016**

(54) **NOVEL METHOD FOR THE CHEAP, EFFICIENT, AND EFFECTIVE PRODUCTION OF PHARMACEUTICAL AND THERAPEUTIC API'S INTERMEDIATES, AND FINAL PRODUCTS**

C12P 13/00 (2006.01)

C12P 17/06 (2006.01)

(52) **U.S. Cl.**

CPC *C12P 13/02* (2013.01); *C12P 17/06* (2013.01); *C12P 7/22* (2013.01); *C12P 13/001* (2013.01)

(71) Applicants: **Sher Ali Butt**, San Diego, CA (US);
Jacob Michael Vogan, Oakland, CA (US)

(57) **ABSTRACT**

(72) Inventors: **Sher Ali Butt**, San Diego, CA (US);
Jacob Michael Vogan, Oakland, CA (US)

The present invention is a method for the biosynthesis of hundreds of compounds, mainly found in the cannabis plant. The starting material for these compounds can be any biological compound that is used/produced in a biological organism from the sugar family starting materials or other low cost raw materials processed via enzymes or within organisms to give final products. These final products include, but are not limited to: cannabinoids, terpenoids, stilbenoids, flavonoids, phenolic amides, lignanamides, spermidine alkaloids, and phenylpropanoids. Specifically, the present invention relates to the regular/modified/synthetic gene(s) of select enzymes are processed and inserted into an expression system (vector, cosmid, BAC, YAC, phage, etc.) to produce modified hosts. The modified host is then optimized for efficient production and yield via manipulation, silencing, and amplifying inserted or other genes in the host, leading to an efficient system for product.

(21) Appl. No.: **15/096,164**

(22) Filed: **Apr. 11, 2016**

Related U.S. Application Data

(60) Provisional application No. 62/145,430, filed on Apr. 9, 2015.

Publication Classification

(51) **Int. Cl.**
C12P 13/02 (2006.01)
C12P 7/22 (2006.01)

Figure 1 (General Pathway Scheme)

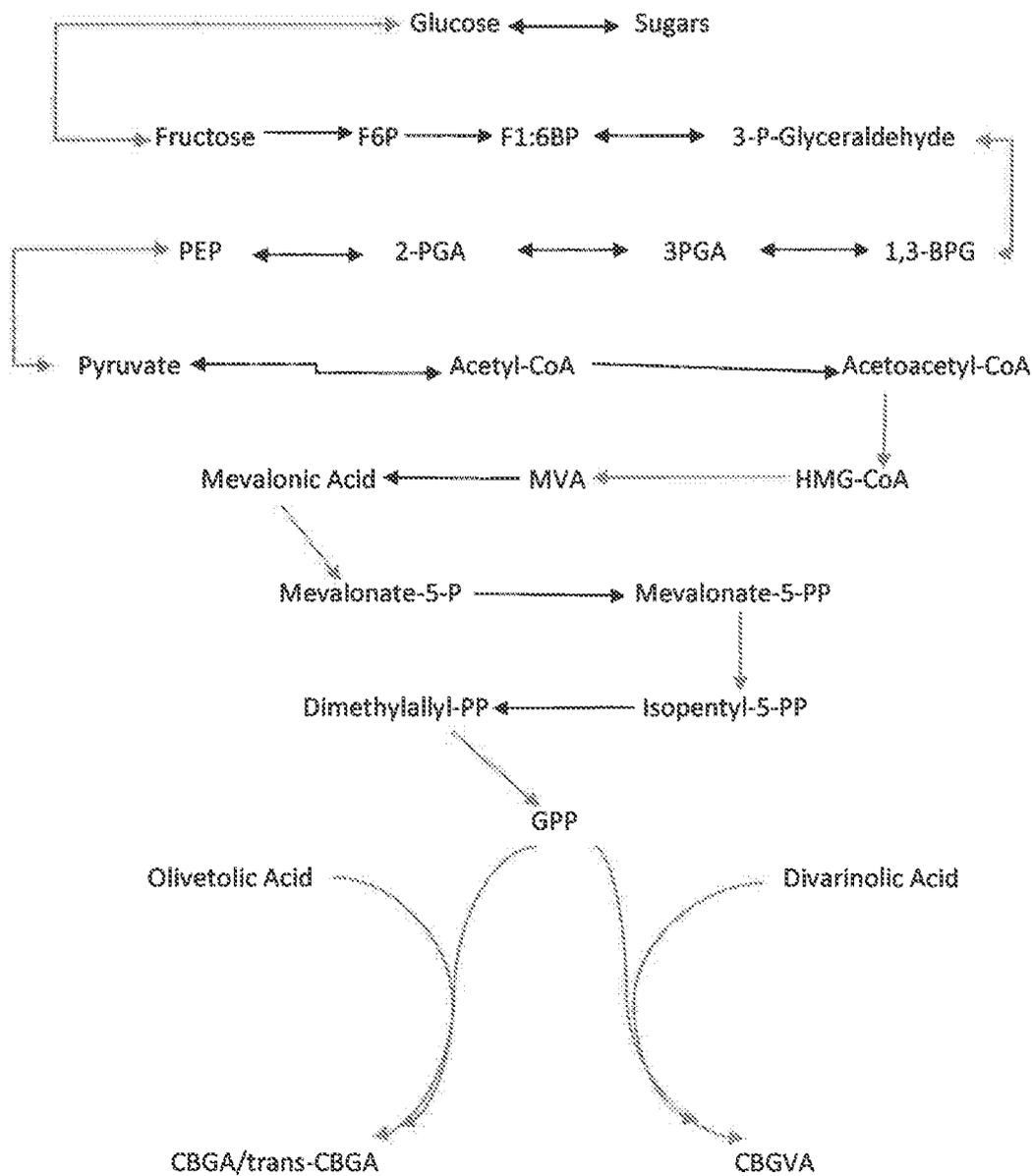


Figure 2 (27 Cannabinoids)

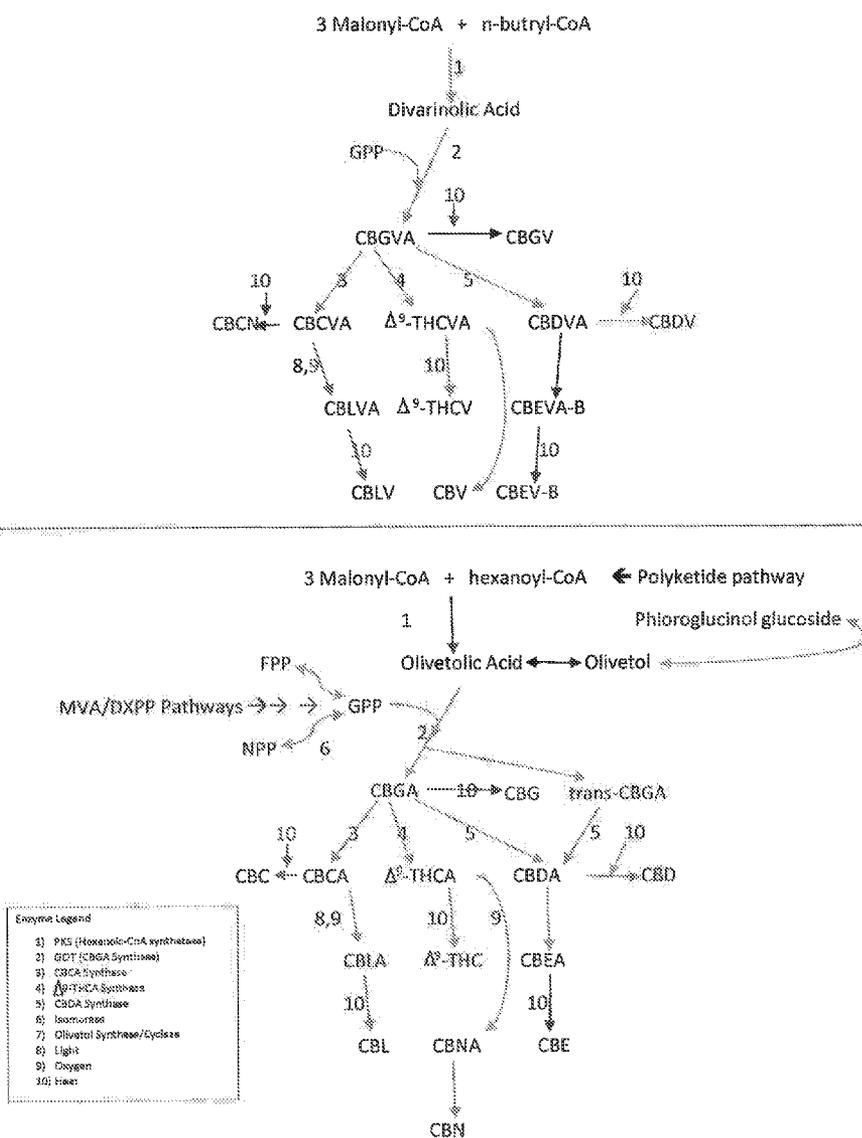
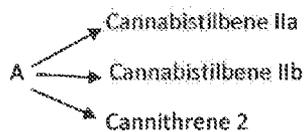
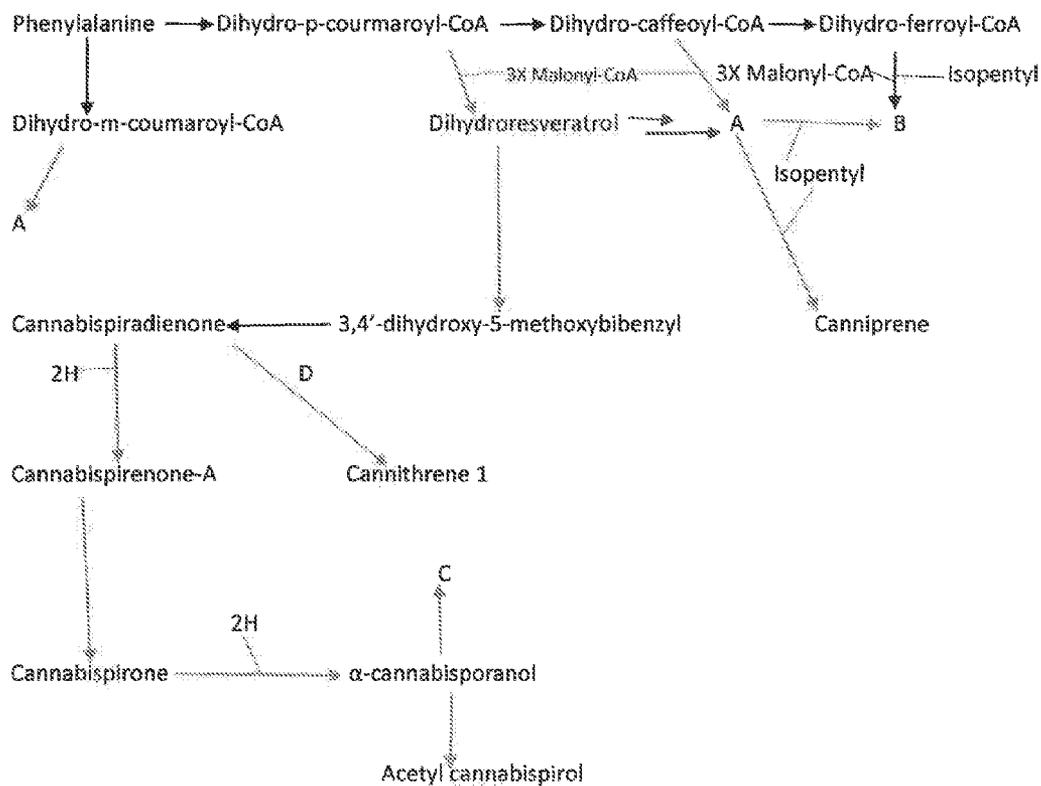


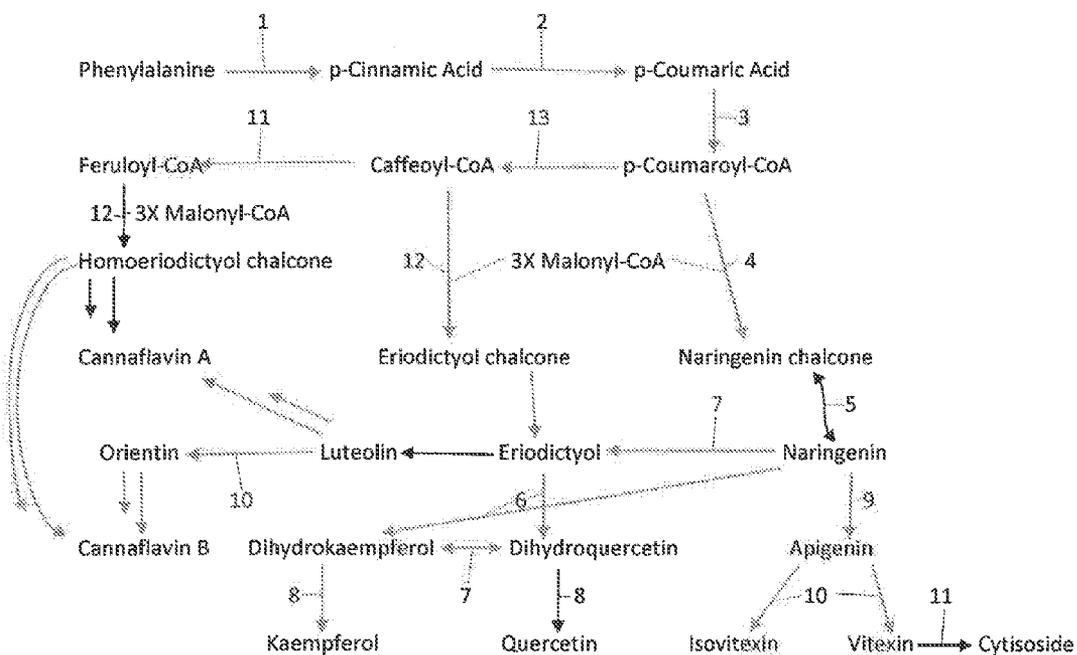
Figure 3 (Stilbenoids)



Legend

- A- 3,3'-dihydroxy-5,4'-dimethoxybibenzyl
- B- 3,4'-dihydroxy-5,3'-dimethoxy-5'-isoprenylbibenzyl
- C- 5-methoxyindan-1-spiro-cyclohexane
- D- Dienone-phenol [in vitro rearrangement by using heat and acidic pH]

Figure 4 (Phenylpropanoids & Flavonoids)



Enzyme Legend	
1)	PAL
2)	C4H
3)	4CL
4)	CHS
5)	CHI
6)	F3H
7)	F3'H
8)	FLS
9)	FNS1/FNS2
10)	UGT
11)	OMT
12)	HEDS/HvCHS
13)	C3H

Figure 5 (Phenolic Amides, Lignanamides)

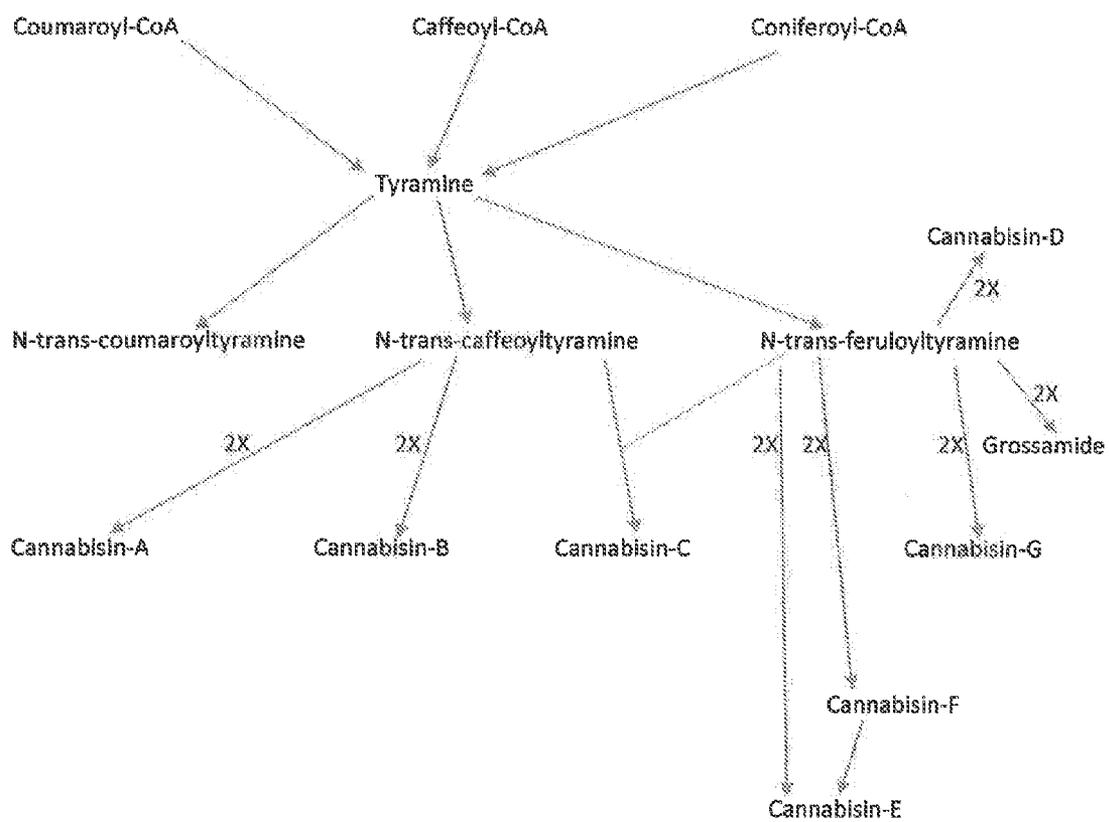
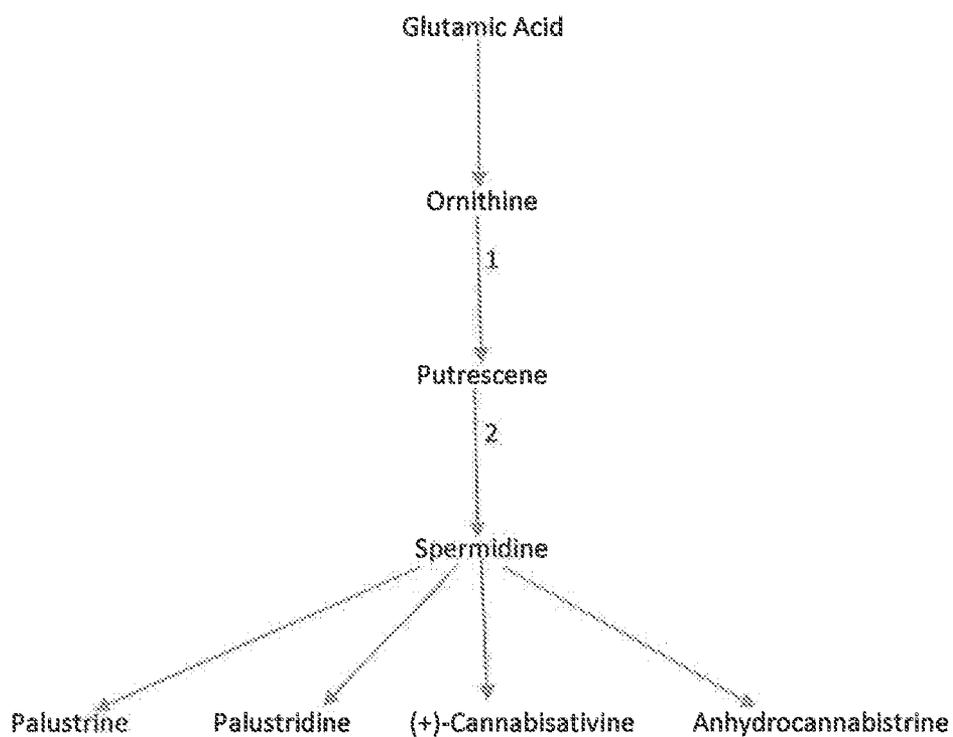


Figure 6 (Spermidine Alkaloids)



Enzyme Legend

- 1) Ornithine decarboxylase
- 2) Spermidine synthase

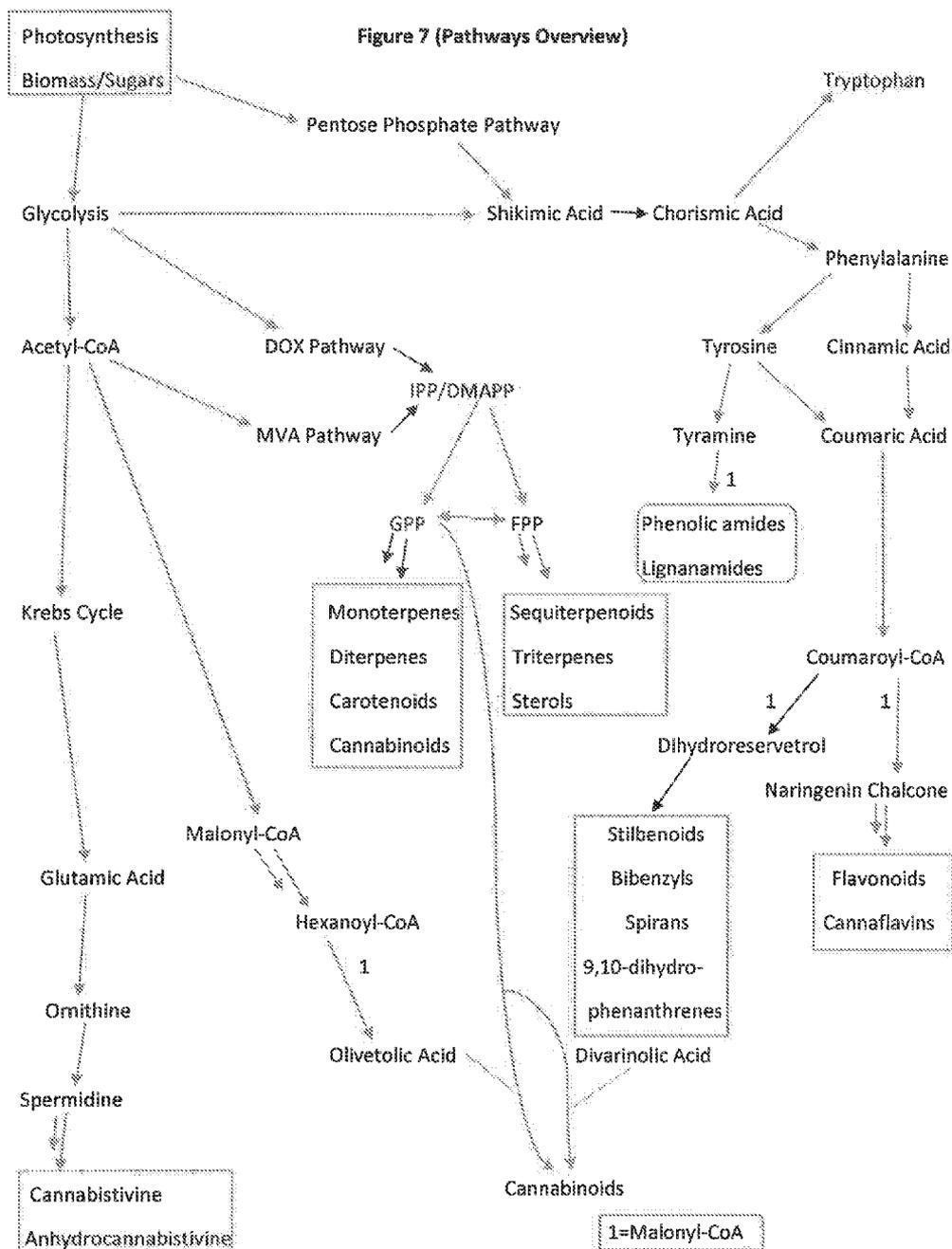
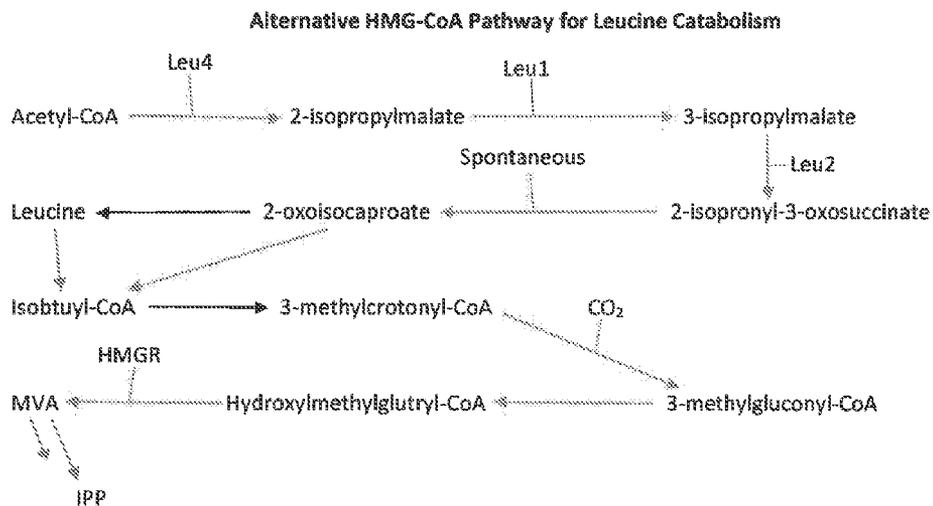
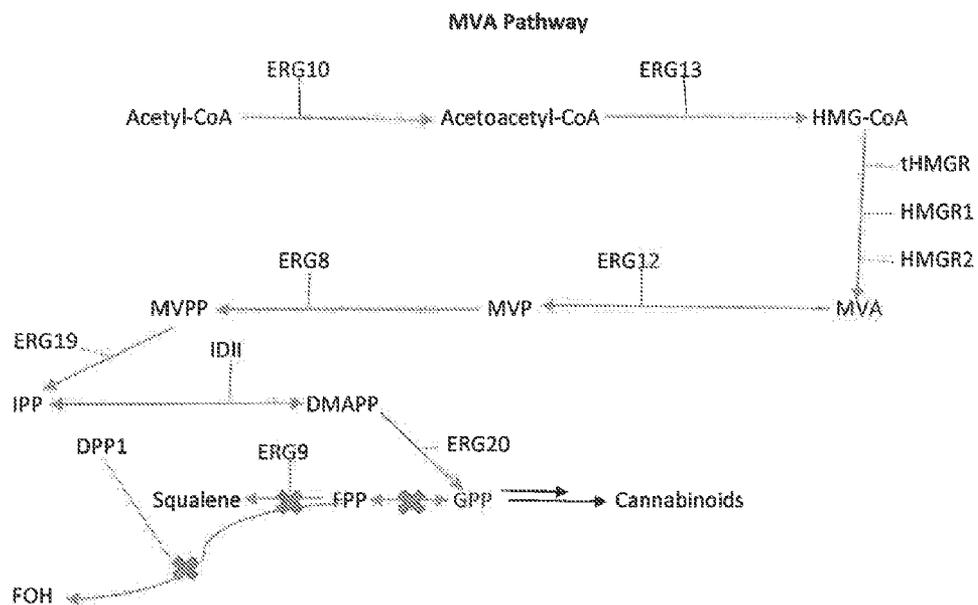


Figure 8 (Genetic Modifications of genes in our pathway for accelerated high yield development)



**NOVEL METHOD FOR THE CHEAP,
EFFICIENT, AND EFFECTIVE PRODUCTION
OF PHARMACEUTICAL AND
THERAPEUTIC API'S INTERMEDIATES,
AND FINAL PRODUCTS**

RELATED APPLICATIONS

[0001] The present application claims the benefit of priority to U.S. Provisional Patent Application Ser. No. 62/145,430, entitled "A Novel Method for the Cheap, Efficient, and Effective Production of Pharmaceutical and Therapeutic API's, Intermediate, and Final Products", filed Apr. 9, 2015, and currently co-pending.

FIELD OF THE INVENTION

[0002] The present invention is in the technical field of large scale production of pharmaceutical and supplemental products for various common illnesses, medical conditions, and general industrial use. More particularly, the present invention is in the technical field of bio-synthesis of cannabinoids, terpenoids, stilbenoids, flavonoids, phenolic amides, lignanamides, spermidine alkaloids, and phenylpropanoids; compounds found in cannabis sativa, along with various combinations and specialized formulations which are beneficial in ailments ranging from cancer to glaucoma. The final product(s) can be an intermediate or a compound of interest. The core concept of the invention is based on the idea of cheaper and more efficient production, along with novel products and applications.

BACKGROUND OF THE INVENTION

[0003] Cannabinoids from cannabis have been used for thousands of years for treatment of various ailments and conditions in many different cultures around the world. However, most of various types of cannabinoids in cannabis are at a very low concentration in the plant. Therefore, most patients/users never get a threshold dosage for any kind of relief from anything other than tetrahydrocannabinolic acid (THC/A), cannabinolic-acid (CBD/A), and cannabinol (CBN). There are a few strains or concentrates available that have a rare cannabinoid, but are usually very highly concentrated in tetrahydrocannabinol (THC) or cannabidiol (CBD) to have any pronounced effect by the rare cannabinoid.

[0004] In other words, the pharmaceutical industry has not tapped into the real potential of the cannabis plant. With time, more research is being conducted into the different kinds of cannabinoids and their medicinal applications. Researchers are finding that many of the other cannabinoids also have unique medicinal properties.

SUMMARY OF THE INVENTION

[0005] Biosynthesis of important molecules can be used for therapeutic applications, bulk substance production, intermediate API biosynthesis, and various other novel formulations and applications for such substances, as known to those skilled in the art. Many biological molecules can be changed/converted into molecules of importance by using enzymes and other processes. This process can be utilized by employing methods for transforming a range of starting materials into final products to be used in pharmaceuticals and supplements as active ingredients, or donating a significant portion of their structure to the final active ingredients.

The final products can also be used in other industries and applications, such as food, beverage, and other goods production. For example, table sugar, starch, and cellulose can be converted to glucose, creating a molecule that can readily be utilized by any organism as an energy source. Therefore, depending on the specific compound(s) being manufactured, and the kind(s) of starting materials available, along with the host and production technique(s) any kind of host engineering, various expression systems and methods, and varying materials, a spectrum of different methods and products is possible.

[0006] The advantages of the present invention include, without limitation, creation of hundreds of compounds from readily available biological molecules that can be produced and harvested from virtually all known sources of plants and other energy producing organisms. Since sugar producing plants and organisms, biomass, and carbon based industrial waste products are very abundant, our "raw material" will be very cheap and easy to obtain anywhere in the world. After scaling up the given methods, hundreds of compounds with medicinal properties can be produced at a very low cost, allowing the widespread distribution and aiding of millions of people.

[0007] Another advantage is that there is no need or use of growing any illegal plants. For example, no marijuana, poppy, or other plant production is necessary. This is advantageous as it will lead to drastically cutting down the production, consumption, and trafficking of many unregulated substances.

[0008] The most important advantage of the present invention is that we can make and use many compounds that are virtually so low in concentration in the cannabis plant, that there is no effect in using cannabis if we are only after the therapeutic effects of these compounds. For example, patients using marijuana can only benefit from tetrahydrocannabinolic acid (THCA), THC, cannabidiolic acid (CBDA), CBD, CBN, and a few other compound class families, as the concentrations of the other compounds is so low that it has no effect. This invention will allow the production of hundreds of compounds in pure form, leading to many new medical discoveries and applications.

BRIEF DESCRIPTION OF THE FIGURES

[0009] The nature, objects, and advantages of the present invention will become more apparent to those skilled in the art after considering the following detailed description in connection with the accompanying figures, in which like reference numerals designate like parts throughout, and wherein:

[0010] FIG. 1 is a diagram of the pathway for the biosynthesis of all molecules of interest via the conversion of starting materials to glucose and then to final products;

[0011] FIG. 2 is a diagram of the pathway for the biosynthesis of cannabinoids;

[0012] FIG. 3 is a diagram of the pathway for the biosynthesis of stilbenoids;

[0013] FIG. 4 is a diagram of the pathway for the biosynthesis of phenylpropanoids and flavonoids;

[0014] FIG. 5 is a diagram of the pathway for the biosynthesis of phenolic amides and lignanamides;

[0015] FIG. 6 is a diagram of the pathway for the biosynthesis of spermidine alkaloids;

[0016] FIG. 7 is a diagram of the combined biosynthetic pathways of FIGS. 1-6; and

[0017] FIG. 8 is diagram of the genetic modification of certain genes for higher product yield in *Saccharomyces cerevisiae* yeast.

DETAILED DESCRIPTION OF THE INVENTION

[0018] The present invention is a method for the biosynthesis of hundreds of compounds, mainly found in the cannabis plant. The starting material for these compounds can be any biological compound that is used/produced in a biological organism from the sugar family starting materials or other low cost raw materials processed via enzymes or within organisms to give final products. These final products include, but are not limited to: cannabinoids, terpenoids, stilbenoids, flavonoids, phenolic amides, lignanamides, spermidine alkaloids, and phenylpropanoids (collectively, "final products").

DEFINITIONS, TERMS, ELEMENTS

[0019] The following are a list of terms and their definitions:

- [0020] Genetic engineering: targeted manipulation of a cell's genetic information
- [0021] Rational Metabolic Engineering: engineering of enzymes, transporters, or regulatory proteins based on available information about enzymes, pathways, and their regulation.
- [0022] Evolutionary engineering: encompasses all methods for empirical strain improvement (mutagenesis [natural or induced] and recombination and/or shuffling of genes, pathways, and even whole cells; usually performed in cycles or sequentially
- [0023] Cannabinoids: compounds that are terpenophenolic with 22 carbons (21 carbons for neutral forms), found in cannabis
- [0024] Terpenoids: also known as isoprenoids, class of organic compounds
- [0025] Stilbenoids: hydroxylated derivatives of stilbene
- [0026] Flavonoids/phenylpropanoids: compounds derived from or using phenylalanine as a precursor
- [0027] Lignanamides/phenolic amides: compounds produced through tyramine pathways
- [0028] Spermidine alkaloids: compounds produced through glutamic acid pathways
- [0029] Starting material/reactant/excipient: compounds used for the initial step of biosynthesis, which are cheap and readily available
- [0030] Intermediate: products that are formed within the biosynthesis pathways, which can further be processed to make final products, or can, themselves, be utilized as a final product
- [0031] Final product/product/end product/compounds of interest: cannabinoids, terpenoids, stilbenoids, flavonoids, phenolic amides, lignanamides, spermidine alkaloids, and phenylpropanoids
- [0032] In-vivo: inside the cell
- [0033] In-vitro: outside the cell
- [0034] BAC: bacterial artificial chromosome, carrier of DNA of interest into host
- [0035] YAC: yeast artificial chromosome, carrier of DNA of interest into host
- [0036] Vector/cosmid/phage: carrier of DNA of interest into host

Starting Materials

[0037] All biological organisms produce organic molecules that are processed in many different processes in the organism. The present invention utilizes starting materials that are either:

- [0038] 1) Readily available and relatively pure
- [0039] 2) Cheap to produce or buy
- [0040] 3) Easily modified (via enzymes, catalysts, or other methods)

[0041] Based on the above criteria, there are multiple groups and families of compounds that would fit one or all three of the above criteria. These groups and families of compounds include, but are not limited to: ligno-cellulosic biomass, forest biomass, energy/food production waste, commonly available sugar-based substrates, food and feed grains.

[0042] Sugars and metabolic intermediates from cellular processes can be used as starting materials. Sugars can be found in abundance in many substances, including, but not limited to the following: rice, soya/rape, cereals (maize), wheat, beans, sugar beet (sugar cane), plant biomass (wood), grasses, and various other sources. Starch, cellulose, fructose, ethanol, and saccharose in the aforementioned substances can be enzymatically converted to glucose, which, after filtration and purification steps, can be used as a raw material for the final products.

[0043] Subsequent steps can also be performed on the lignocellulose, which further makes hemicellulose and cellulose, both which make glucose. An advantage of this method is that there are by-products generated which can be sold as raw material to make hydrocarbons, biogas, and other fuel sources. Whole crops or parts of crops, or waste matter from crop products can be used and incorporated into this system, yielding an "eco-friendly" facility. Products made from these raw materials can use any of the starting materials listed in Table 2.

[0044] Within the realm of readily available non-biomass/crop bulk material, HFCS (high fructose corn syrup) is a cost effective syrup made with fruit sources that contains anywhere from 30-90% fructose, along with some other sugars. Plants that make molasses, HFCS, and other sugars can be genetically modified to enhance the production of sugar, leading to better yields of starting material from the crop. Other products from these plants can also be incorporated into compounds of interest production via slight system modification. Biodiesel, ethanol, glycerol, lactic acid, whey and glucose are a few others. These work due to the fact that any of these products can be converted into starting material for our own purposes using enzymatic or physiochemical tools.

[0045] Plants also have their own innate levels of metabolites that can be harvested into the process from a plant biomass source. Processes can be crafted that utilize most of the metabolites and biomass for API production giving the maximum efficiency and usability per amount of starting material used. (Enzyme combinations or chambers that utilize most intermediates, sugars, oils, etc. in each biomass load).

[0046] Biorefineries can be custom designed that cater to specific raw material (plant biomass for harvesting lignocellulose which is further processed and refined into a simple carbohydrate used in the API manufacturing processes). During certain steps throughout the process, ther-

mochemical and other processing can be used for higher efficiencies which are not possible with biochemical processing alone.

[0047] Another group of cheap starting materials is agricultural residue, grass, aquatic biomass, and water hyacinth. Products such as oils and alcohols can be made with these bulk materials. These materials can be converted enzymatically and chemically into starting materials that can readily be injected into our API production system.

[0048] Specifically, biorefineries can be designed to be extremely efficient, using all parts of the raw material. For example, concerning plant biomass, the biomass can be step-wise processed so we are able to harvest all individual components. The first step can be using solvent to extract terpenes, alkaloids, etc. Other methods can be used to extract steroids, triglycerides, and other valuable metabolites. Finally the biomass can be treated with cellulases to give glucose, which is one of the primary raw materials of choice.

Production Roadmap Summary

[0049] The present invention is a method that covers the bio-synthesis of hundreds of compounds, mainly found in the cannabis plant. The starting material for these compounds can be any biological compound that is used/produced in a biological organism from the sugar family starting materials or other low cost raw materials processed via enzymes or within organisms to give final products. Information related to the starting materials were detailed in the previous section.

[0050] Most sugars and related compounds can be interchanged using various enzyme systems. For example, we can convert glucose to fructose using Fructose 6-Phosphate (F-6-P) as an intermediate.

[0051] Apart from starting materials, we can either:

[0052] 1) Make enzymes via vectors in bacteria (e.g. *E. coli*) or yeast (e.g. *S. cerevisiae*), extract enzymes, and create in vitro models for making cannabinoids.

[0053] 2) Make enzymes via protein synthesizing systems (Protein Synth. Robot, Cell Free Expression Systems, etc.)

[0054] 3) Make final products (compounds of interest) in bacteria or yeast via vectors, plasmids, cosmids, mRNA, various RNA, etc; feed them substrate and purify product.

[0055] 4) Genetically engineer strains of bacteria and yeast that specialize in cannabinoid production, or intermediate production, or substrate production, etc.

[0056] 5) Use organic chemistry for certain parts of the above processes.

[0057] 6) Use various plant starting material for large quantities of substrates or intermediates.

[0058] 7) Genetically engineer various plants to produce cannabinoids. (e.g. Tomatoes or celery that naturally produce cannabinoids, or algae that produces cannabinoids)

[0059] 8) Using bioengineered or unengineered *C. sativa* or any other plant/algae cell lines for enzyme/substrate/intermediates/product(s) production.

[0060] 9) Protein engineering on the various proteins involved in the processes; engineering will enhance the functionality, ruggedness, and efficiency of the enzymes, and altering them into a novel protein, one not found to be covered in any of the above prior art patents.

[0061] 10) Genetically engineer various plant species to produce higher yielding raw material (sugars) to be used in production of the products. A possibility is to have an indoor/grow for different plants to be used as raw material producers.

[0062] After the final product is made, a purification system will filter and concentrate the target molecules. Examples include large scale filtration systems such as chromatography. Once a pure product, we can utilize liquid solutions, caps, sprays, and other delivery systems.

[0063] As many of these final products are made, their applications can be seen from glaucoma to cancer, or general well-being. Certain cofactors can be combined with certain final products for more efficacy against specific medical conditions (e.g. combine certain vitamins or other therapeutic compounds with certain compounds of interest). We can also make final products that have certain combinations of compounds of interest with other cofactors as well (e.g. combine THCA/CBDANitamin C, or CBDVA/CBD). This patent covers all the products above and also ones discovered in the future based on the same principles and methods.

DETAILED DESCRIPTION OF THE FIGURES

[0064] Referring now to the invention in more detail, in FIG. 1 there is shown a family of sugars and other common derivatives. Along each arrow for each reaction, the number denotes a specific enzyme that catalyzes the reaction. Starting with any sugar in Figure 1 (list of starting materials in Table 1), we can convert it to glucose to incorporate it into the reaction using the appropriate enzyme, as known to those skilled in the art. An unlimited number of ways are possible when dealing with any starting material, as described above. Enzymes needed for each kind of substrate can be made in vivo or in vitro just as we will be doing for the enzymes in the final product or intermediate production. The final sugar that enters our mechanism will be either glucose or fructose. Through the glycolysis pathway, the sugar will be converted into Acetyl-CoA with the addition of ATP and CoA (shown in FIG. 1). From this point on, the intermediate can follow a variety of paths that can lead to hundreds of products. There are many alternative ways of doing this. We can use the DOX, MEP or MVA pathways to get IPP and DMAPP, which give us GPP and NPP. For a reaction with Olivetolic Acid or Divarinolic Acid, we get many cannabinoids as final products.

[0065] The generalized pathway for the production of cannabinoids once the starting material is converted to glucose is the following, using appropriate enzymes as known by those skilled in the art:

[0066] Glucose→Fructose→F-6-P→FI:6BP→3-P-Glyceraldehyde→1,3-BPG→3PGA→2-PGA→PEP→Pyruvate→Acetyl-CoA→Acetoacetyl CoA→HMG-CoA→MVA→Mevalonic Acid→Mevalonate-5-P→Mevalonate-5-PP→Isopentyl-5-PP→Dimethylallyl-PP→NPP/GPP→GPP

[0067] This general pathway is outlined in FIG. 1. From this point on, the pathway can utilize Olivetolic Acid or Divarinolic Acid with GPP, yielding CBGA or CBGVA, which can further yield other cannabinoids, as shown in FIG. 2.

[0068] The pathways for stilbenoids, phenylpropanoids, and flavonoids work in a similar fashion. Phenylalanine is

generated from sugars, which is then further processed into other compounds using enzymes to final compounds, as shown in FIG. 3 and FIG. 4.

[0069] Phenolic amides and lignanamide pathways are derived from tyramine molecules reacting with other compounds, as shown in FIG. 5. Tyramine can also be synthesized in our cells of interest as most living organisms contain the pathway to synthesize tyramine on their own. Same is the case for spermidine alkaloid production, as most cells already produce glutamic acid, which can be further processed by enzymes into the final components, as shown in FIG. 6.

[0070] FIG. 7 is the total pathway overview, showing how all the different classes of compounds can be made, and the general paths they take for being biosynthesized in the cell.

Overview of Procedure

[0071] A general scheme of the work flow is as follows:

[0072] 1) Regular/modified/synthetic gene(s) of select enzymes are processed and inserted into an expression system (vector, cosmid, BAC, YAC, phage, etc.) to produce modified hosts.

[0073] 2) Mod host is then optimized for efficient production and yield via manipulation, silencing, and amplifying inserted or other genes in the host, leading to an efficient system for product. It is important to remember that every organism is different, and to get a specific compound each optimization will also be different.

[0074] 3) Mod host can produce enzymes and final products/intermediates, or be further modified using host engineering techniques. (Host engineering Can also be performed before insertion of exp. System)

[0075] 4) Mod and engineering hosts produce products and intermediates.

[0076] 5) Product is purified and can be further modified/processed.

[0077] In Table 1, different final products are listed along with possible uses. This list is by no means exhaustive, and as such this patent covers any molecules that are made this way. Table 2 lists all possible starting materials that can be utilized for a cheap and efficient biosynthesis.

[0078] In more detail, referring to the inter-conversion of sugars, we employ enzymes readily available in the market. Pure enzyme stock can be diluted and added to a solution with the substrates. Once the reaction is complete, we can filter out the enzyme via dialysis tubing, by precipitation out of the solution, chromatography, or other industrial methods for filtration and purification. Each step in FIGS. 1 to 7 will give work with this strategy, leading us up to the final products or key intermediate molecules. Certain steps in the process can be worked on by using chemical and physical methods as well. For example, prenylation of certain compounds can be done outside the cell, as it may be advantageous to do so since unprenylated compounds are also high value compounds. Small batches can be prenylated accordingly to demand via a chemical process.

[0079] There are also commercially available cell free expression systems, which are able to produce proteins without the need of any host. With appropriate optimization steps, it is possible to get a cheap and efficient process for production of these compounds using identified starting molecules.

Application Techniques

[0080] Referring to bacterial, yeast, plant, and algae incorporation of genes, there are a number of strategies that can be applied to achieve this. We can:

[0081] 1) Add genes for 1-10 enzymes in various commercially available vectors, cosmids, plasmids, etc. Only need 1-10 enzymes added, as others are already built in most living organisms. For example, glycolysis pathway and related enzymes are already present in most hosts.

[0082] 2) Bioengineer genes for better yield and suitability in the host.

[0083] 3) Bioengineer strains of bacteria and yeast that are specialized in producing important molecules. Many metabolic strategies exist, with identification by appropriate screening methods:

[0084] 1) Rational metabolic engineering: engineering pathways using available information

[0085] 2) Evolutionary engineering: using random genetic perturbations and/or mutations (via random mutagenesis in whole genome and/or parts)

[0086] 3) Transposon mutagenesis & gene overexpression libraries: overexpression and/or deletion of single or multiple genes;

[0087] 4) Global transcription machinery engineering: basal transcription factors mutagenesis causing a global reprogramming of gene transcription and/or translation

[0088] One strategy is to suppress any pathway that is not essential to our goals or the survival of the host.

[0089] Another is to enhance our key pathways, or mixing and matching the two methods. The second strategy is through rapid directed evolution, possible by producing many generations so eventually we get a generation of host that has evolved with our genes/functions of interest.

[0090] 4) Bioengineer custom basic life forms that are specifically making our products, using another organism or using synthetic/modifications. Components from other hosts and system to make a custom organism.

[0091] 5) Bioengineer bacteria and yeast to have enzyme genes in their chromosomes, and make intermediates or final products inside the host. The product of this process can further be modified.

[0092] 6) Propagate various colonies of organisms which co-exist symbiotically, with the first making our starting material after utilizing a precursor, and the other colonies making our final product. This process can also be incorporated into an ecosystem type setup of different chambers, each holding different organisms that use specific parts of the raw material to produce intermediates or final products that can be modified post-manufacturing.

[0093] Referring to the extraction of enzymes once they have been produced in the host, there are many ways to isolate and purify our enzymes. Many organisms have the ability to excrete proteins, which can be collected much easier than cell lysis, as known by those skilled in the art. This technique is the preferred method.

[0094] Another method is to lyse the host culture and purify with traditional biochemistry methods (gels, centrifugation, ammonium sulfate precipitation, etc.), use a specialized nickel column with a prep HPLC (need to add a HIS tag to our proteins; remove HIS tag after purification), etc.

Example

Bacterial

[0095] Bacteria (*E. Coli*, etc.) are inserted with exp. system giving us a modified host. The mod host can either be further processed or it can generate products. Products/intermediates are made in the host, and may be either enzymes that are further extracted and used in vitro, or we add substrates into the bacterial culture so they use the enzymes produced in them to make the substrate. Either way (protein or prod production), purification is carried out to get final products, or intermediates that can be further processed in vitro to give final products. Throughout this procedure, host engineering can be carried out at any step of any process to get better yields.

Example 2

Plants

[0096] Plant tissue can be used as a starting material to get a tissue culture going. Appropriate expression vectors/systems carry our interest genes into the cells. Alternatively, cell engineering can lead to many combinations that may have similar or different outcomes. The culture can be grown into full plants, and products are ingested by consuming the plants (e.g. tomatoes with certain cannabinoids produced within, etc.). The second way uses the cell culture in a synthetic environment to produce final products/intermediates. Finally, product is purified and used.

Example 3

Algae

[0097] Algae are modified with the usual techniques used for host engineering. Once completed, the mod host can be embedded into a system similar to biofuel production from algae. Using sunlight and some nutrients, the algae produces final products/intermediates, which is appropriately filtered from the bulk. Other products generated can be further processed to get biofuels or other important compounds that can readily be sold in the market.

Example 4

Fungi

[0098] Fungi modified with the techniques can:

- [0099] 1) Use plastic to produce final products/intermediates. Plastic needs to be processed and broken down into components before being used in this process via chemical and biological processes, known by those skilled in the art.
- [0100] 2) Clean up waste, whilst producing final products/intermediates at the same time.
- [0101] 3) Produce beer and wine with fungi that also makes final prod/intermediates. Beer and wine will contain our compounds of interest.
- [0102] 4) Use fungi cultures to produce compounds of interest.
- [0103] 5) Genes for *s. cerevisiae* strains to be modified for better yields of final products:
 - [0104] tHMGR
 - [0105] upc2-1 (allows higher uptake of exogenous sterol five-fold from medium)

[0106] ERG genes (ERG6, ERG2, ERG3, ERG1, ERG11, ERG24, ERG25, ERG9, ERG10, ERG13, ERG12, ERGS, ERG19, ERG20)

[0107] HMGR1 and HMGR2

[0108] IDI genes

[0109] Gal80p

[0110] DPP1, ADH2, and ALD6 genes

[0111] FPP/GPP synthase (chose avian FPP synthase as it exhibits higher catalytic turnover rates and lower Kms for substrates than other prenyltransferases)

[0112] Manipulation, deletion, overexpression, and other modifications to the genes listed above will produce strains that are highly efficient for the production of our compounds of interest. These strains have an exogenous sterol uptake, as the internal sterol pathway has been disabled by manipulations so that all the carbon flux can be directed toward the production of our compounds of interest. Example of genetic pathway regulation in yeast is shown in FIG. 8.

[0113] Our initial strategy in *S. Cerevisiae* was to increase the carbon flux of our pathways of interest, while decreasing or eliminating pathways that led carbon flux away from our pathways as well. We also focused on exogenous sterol uptake for higher production and secretion levels, cell permeability for more efficient and cheaper production, along with focusing the pathways on utilizing the cheapest sugars. Dynamic control over ergosterol regulation can increase yields as well. Overall result is a strain that is has increased yield many fold, while making the overall production more stable and cheaper.

[0114] 1) Perform EMS mutagenesis on yeast strains (BY4741, BY4742, CEN.PK, CEN.PK2, EPY300) to get colonies with a SUE (sterol uptake exogenous) mutation. This enables us to provide exogenous sterol to the yeast while cancelling out the gene that diverts carbon flux towards ergosterol, thereby increasing total carbon flux. Without the SUE mutation, the cell diverts lots of carbon flux toward manufacturing sterols, thereby diverting the pools of intermediates away from our compounds and interest leading to very low yields.

[0115] 2) Perform ERG1 and ERG9 gene knockouts. ERG1 knockout stops the activity of conversion of squalene to squalene epoxide, thereby complementing the SUE mutation and allowing higher uptake of exogenous ergosterol, while ERG9 knockout takes out the cells ability to divert carbon flux towards other metabolites.

[0116] 3) On some lines, we can perform a DPP1 knockout. DPP1 knockout ensures that FPP/GPP are not converted to FOH, thereby blocking the pathway towards FOH products in the cell.

[0117] 4) Perform ERG2, ERG3, or ERG6 mutations in different cell lines, while performing upregulation mutation on upc2-1 gene (general transcription factor) on all three lines. This helps increase cell membrane permeability for better excretion of our compounds without the need for cell lysis and having the ability to use two-phase or continuous fermentation. This also allows the cells to uptake more fatty acids, thereby increasing the yield many fold.

[0118] 5) Overexpression of ERG10, ERG13, HMGR1/2 or tHMGR, ERG12, ERG8, IDI11, HFA1 genes in yeast inserted via vectors. By overexpression of these genes, we are amplifying the enzymes of the

MVA pathway from the sugars to our compounds, thereby amplifying the intermediates and final products.

- [0119] 6) Modification of avian and/or salmonella ERG20 gene encoded FPP synthase (ERG20p). Some cells lines can also be modified using the Erg20p (F96C) mutations. This allows for higher Kms and increased catalytic turnover compared to endogenous GPP synthase, while the engineering itself allows for production of GPP.
- [0120] 7) Gal80p gene deletion so we do not need to use galactose sugar when inducing promoter expression. This is important since others have used galactose promoters, which need expensive galactose sugars for production. By deleting this gene, the cells bypass the need for galactose to express enzymes, leading to cheaper and more efficient biosynthesis.
- [0121] 8) Adding ADH2p promoter to induce strong transcription under conditions with low glucose. This promoter is more efficient than the GAL promoter, and has best results while using non-glucose sugars (ethanol, fructose, etc.) which are cheaper.
- [0122] 9) On some lines, we also overexpress ADH2 and ALD6 genes, along with overexpression of an acetyl-CoA C-acetyltransferase to increase efficiency of the system, while also gaining the ability to convert ethanol to acetate efficiently.
- [0123] 10) Adding and overexpressing enzymes for the production of CBDA (OS-OAC fusion enzyme, CsPtI, CBDA Synthase), constructed in a single vector. These enzymes are codon optimized.
- [0124] 11) Grow colonies while adding free fatty acids, and hexanoic acid (for THCA, CBDA, CBGA, CBCA) or butyric acid (for THCVA, CBDVA, CBGVA, CBCVA).
- [0125] 12) For production of THCA/THCVA, use THCA synthase in step 10 instead of CBDA synthase. For production of CBGA/CBGVA, follow step 10 but don't use CBDA synthase in vector construct. For production of CBCA/CBCVA, use CBC synthase in step 10 instead.
- [0126] Our strategy for *Pichia pastoris* (*Pichia* Pink 1, 2, 3 from Invitrogen) yeast was similar to *S. Cerevisiae*, except for the following differences:
- [0127] 1) Each enzyme, vector, and primer were optimized for insertion into *pichia* cells instead of *s. cerevisiae*.
- [0128] 2) Methanol is used to supplement cells in addition to free fatty acid, hexanoic acid, and butyric acid, thereby reducing the total cost of production many fold, while eliminating any contamination issues from other species.
- [0129] 3) No EMS mutagenesis is performed.
- [0130] 4) Knockouts of *pep4* (encoding Proteinase A), *prb1* (encoding Proteinase B), and *YPS1* genes are also introduced. These knockouts allow for the integration of high copy plasmids leading to higher yields.
- [0131] 5) Steps 7, 8, and 9 from the *S. cerevisiae* strategy above are not to be performed in *pichia* cells.

Example 5

Cell Free Expression Systems

[0132] Vectors are introduced into cell free expression systems, and make either enzymes or intermediate/final products. Further processing or steps are needed to get purified final products.

Procedures

EMS Mutagenesis (*S. Cere.*; BY4741, BY4742, CEN.PK, CEN.Pk2, BY300)

- [0133] 1) Cells incubated overnight @30 C in 5 mL TPD medium while shaking @200 rpm to establish 200 mL YPD shake flask culture.
- [0134] 2) When OD600 of yeast culture reaches 1.0, cells are spun down by centrifugation (12 mins at 4,000 g), washed twice with 20 mL 0.1M sodium phosphate buffer, pH7.0.
- [0135] 3) Cells concentrated by centrifugation again, re-suspended in 1 mL 0.1M sodium phosphate buffer, transferred to 30 mL FALCON tubes, treated with 300 uL EMS (1.2 g/mL).
- [0136] 4) Cells are incubated at 30 C for 1 hr while shaking.
- [0137] 5) Stop mutagenesis by adding 8 mL of sterile 5% sodium thiosulfate to yeast cells.
- [0138] 6) Cells are pelleted, washed with 8 mL sterile water, concentrated by centrifugation, re-suspended in 1 mL sterile water and 100 uL aliquots plated into YPD-NCS agar plate (YPD+50 mg/L each of cholesterol, nystatin, squalastatin, and 2% Bacto-agar).
- [0139] 7) In some instances, washed cells were re-suspended in 1 mL YPDE liquid media for overnight recovery before plating to YPD-NCS agar medium.
- [0140] 8) Incubate cultures for up to two weeks at 30 C until distinct colonies are visible.

Bacteria & Yeast Culturing

- [0141] 1) Grown using standard culture practices.
- [0142] 2) YPD media without selection consisted of 1% Bacto-yeast extract, 2% Bacto-peptone, and 2% glucose.
- [0143] 3) Add 40 mg/L ergosterol to YPD media to get YPDE media.
- [0144] 4) Add 40 mg/L each of nystatin, cholesterol, and squalastatin to YPD media to get TPDNCS media.
- [0145] 5) Add 40 mg/L each of ergosterol and squalastatin to YPD media to get YPDSE media.
- [0146] 6) Prepare minimal media, SCE (pH5.3), by adding 0.67% Bacto-yeast nitrogen base (without amino acids), 2% dextrose, 0.6% succinic acid, 0.14% Sigma yeast dropout soln (-his, -leu, -ura, -trp), uracil (300 mg/L), L-tryptophan (150 mg/L), L-histidine (250 mg/L), L-methionine (200 mg/L), L-leucine (1 g/L), and 40 mg/L of ergosterol.
- [0147] 7) Cholesterol and ergosterol stocks are 10 mg/mL in 50% Triton X-100, 50% ethanol and kept at -20 C.
- [0148] 8) Selection media prepared similarly except without supplementation of media with indicated reagent based on the yeast auxotrophic markers.

- [0149] 9) All solid media plates are prepared with 2% Bacto-agar.

Yeast Transformation & Culture Performance

- [0150] 1) Used FROZEN-EZ Yeast Transformation II Kit from Zymo Research, Orange, CA, according to manufacturer's recommendations.
- [0151] 2) μ g of plasmid was used per transformation, followed by selection on agar plates of SCE medium lacking specified amino acids for auxotrophic markers, or YPDE containing 300 mg/L hygromycin B for screening *erg9* knockout at 30 C.
- [0152] 3) Colonies are picked and used to start 3 mL cultures in minimal media to characterize their terpene production capabilities. (6 days incubation at 30 C while shaking)
- [0153] 4) Best cultures are chosen to move further, using 30 mL shake flask cultures.
- [0154] 5) Cultures are grown to saturation in minimal media, inoculated into 30 mL SCE media and 1 mL aliquots are taken out daily for 15 days.
- [0155] 6) Cell growth is monitored via change in optical density at 600 nm every two days using dilutions at later stages of growth.
- [0156] 7) Production of terpenes is determined via testing.

ERG9 Knockout Mutations

- [0157] 1) Primers ERG9PS1 and ERG9-250downS2 used to amplify hygromycin resistance gene, *hphNT1*, from the pFA6-*hph-NT1* vector.
- [0158] 2) Simultaneously add 42 bp nucleotide sequences homologous to regions surrounding ERG9 gene in yeast genome.
- [0159] 3) Purified PCR fragment is transformed into various cell lines identified in phase 2 with the ability to accumulate farnesol and selected on YPDE plates containing 300 mg/L hygromycin.
- [0160] 4) Independent single colonies are picked for ergosterol dependent test, PCR confirmation of recombination with *hphF* and ERG9 450DWR primer.
- [0161] 5) Farnesol production analysis done by GC-MS/LC-MS.

ERG1 Knockout Mutations

- [0162] 1) Primers ERG1F and ERG1R used to amplify the squalene epoxidase synthase ERG1 gene by using Takara high fidelity Primerstar taq polymerase.
- [0163] 2) Obtained PCR fragment is gel purified, A tailed and ligated into the pGEM-Teasy vector.
- [0164] 3) Obtained vector is used as template to run second PCR with primers *Erg1-splitF* and *EGR1-splitR* to obtain PCR fragment with deletion of 8916 bp CDS in the middle, yet containing 310 bp at 5' end region and 291 bp at 3' end region of ERG1 gene which are the target homologous recombination sequence for ERG1 knockout.
- [0165] 4) After digestion with BamHI, self-ligation, and transformation to DH5alpha competent cells, resulting vector is pGEM-ERG1-split.
- [0166] 5) *Padh-Kanmx4-Tcyc-LoxP* antibiotic selection marker cassette is constructed by assembly PCR of three fragments.
- [0167] 6) *Padh* promoter is PCR amplified with *Padh-loxP-ManHIF* and *Padh-Kanmx4R* primers using *Yep352* vector as a template.
- [0168] 7) *Kanmx4* selection gene is PCR amplified using *Padh-kanmx4F* and *Tcyc-kanmx4R* primers using *PYM-N14* plasmid as a template.
- [0169] 8) *Tcyc* terminator was PCR amplified with *Padh-loxP-BamHIF* and *Padh-Kanmx4R* primers using *Pesc* vector as a template.
- [0170] 9) 3 PCR fragments containing homologous regions with each other were gel purified and 250 ng of each fragment were mixed together to serve as template for the secondary assembly PCR reaction to yield *pAdh-Kanmx4-Tcyc-LoxP* cassette.
- [0171] 10) Cassette is digested and inserted into pGEM-ERG1-split vector, and used as template to run PCR with ERG1F and ERG1R to get PCR fragment used to generate cell lines.
- [0172] 11) *Pgpd-tHMGR-Tadh* fragment was amplified from *Pesc-Gpd-leu-tHMGR* vector with primers *GPD-BamHIP* and *Tadh-XhoIIR*.
- [0173] 12) Insert fragment into pGEM-ERG1-split vector containing *kanmx4* cassette.
- [0174] 13) Use construct as template to amplify with ERG1F and EGR1R primers to gain the fragment for building slightly different cell lines, which include integration of one copy of *tHMGR* into the ERG1 gene.

Primer name	Primer Sequence
ERG9pS1 (SEQ ID NO 1)	GTACATTTTCATAGCCCATCTTCAACAACAATACCGACTTACCCGTACGCTGCAGGTCGAC
ERG9 250dws2 (SEQ ID NO 2)	CAGATTGACGGAGAGAGGGCCACATTGTTTGTTCGGAATAAATCGATGAATTCGAGCTCG
Hph F (SEQ ID NO 3)	ATGGGTA AAAAGCCTGAACTCA
Hph R (SEQ ID NO 4)	TTATTCCTTTGCCCTCGGACGAG
ERG9 450dWR (SEQ ID NO 5)	AGATGCTAGTCAATGGCAGAAG
WEG9p300upF (SEQ ID NO 6)	TGCTTACACAGAGTGAACCTGC

-continued

Primer name	Primer Sequence
ERG9 300R (SEQ ID NO 7)	CTCGTGGAAAGTGACGCAAC
pGPD-BAMHI F (SEQ ID NO 8)	cgGGATCCagttttatcattatcaatactcgcc
pGPD-NotIR (SEQ ID NO 9)	gggGCGGCCGcgagctcagttttatcattatc
tHMGR-NotIF (SEQ ID NO 10)	GGGGCGGCCGCAAACAATGTTGTCACGACTTTTCCGTATGC
tHMGR-SpeIR (SEQ ID NO 11)	GACTAGTTCAAGCTGACTTCTTGGTGCACGTTTCCTTG
ERG1F (SEQ ID NO 12)	ATGTCGTCTGTTAACGTTGCACCTG
ERG1R (SEQ ID NO 13)	TTAACCAATCAACTCACCAAAC
ERG1-split F (SEQ ID NO 14)	CGGGATCCCTCGAG TTGTTGCTGCTGACAGCGATAAC
ERG1-splitR (SEQ ID NO 15)	CGGGATCCGCTAGCGGTACCACATGGGTCCTTTATATGACACG
ERG1 90up F (SEQ ID NO 16)	ATCAGAACAATTGTCCAGTATTG
ERG1100dwR (SEQ ID NO 17)	AATGTACTATACAAGCCTTCC
bSQS-NotIF (SEQ ID NO 18)	GGGGCGGCCGCAAACAATGGGGATGCTTCGCTGGGGAGT
bSQS-SpeIR (SEQ ID NO 19)	GACTAGTTTAGCTCCTCAATTCGTCAAAGGT
Cre-NotIF (SEQ ID NO 20)	GGGGCGGCCGCAAACAATGGACATGTTTCAGGGATCGCCAGG
Cre-SpeIR (SEQ ID NO 21)	GACTAGTCTAATCGCCATCTTCCAGCAGGCG
Padb-Loxp- BamHIF (SEQ ID NO 22)	CGGGATCCATAACTTCGTATAGCATAACATTATACGAAGTTATGTGGAA TATTTCCGATAT
Padh-Kanmx4F (SEQ ID NO 23)	GCATACAATCAACTAAGCTAAGCTAAAAAATGGGTAAGGAAAAGACT CACGTTC
Padh-Hanmx4R (SEQ ID NO 24)	GAAACGTGAGTCTTTTCCTTACCATTGTTTAGCTTAGCTTAGTTGA TTGTATGC
Kanmx4-TcycF (SEQ ID NO 25)	CATTTGATGCTCGATGAGTTTTTCTAAATCCGCTCTAACCGAAAAGGA AGGAG
Kanmx4-TcycR (SEQ ID NO 26)	CTCCTTCCTTTTCGGTTAGAGCGGATTTAGAAAAACTCATCGAGCATC AAATG
Tcyc-LoxP-NheIR (SEQ ID NO 27)	GGGGCTAGCATAACTTCGTATAATGTATGCTATACGAAGTTATCTTCG AGCGTCCCAAAA
Gpd-GamHIF (SEQ ID NO 28)	CGGGATCCAGTTTATCATTATCAATACTCG
Tadh-XohI (SEQ ID NO 29)	GGGCTCGAG GAGCGACCTCATGCTATACCTG
Kanmx4R (SEQ ID NO 30)	TTAGAAAAACTCATCGAGCATC

Expression of Enzymes for Cannabinoid Production

All REV

OLS 5' FWD

SEQ ID NO 33

SEQ ID NO 31

Length: 55

Length: 58

Type: DNA

Type: DNA

Organism: Artificial Sequence

Organism: Artificial Sequence

Notes: Primer

Notes: Primer

[0175]**[0177]**Gcatagcaatctaataagtttaaa atgaatcatttgagagcagaa
gggcctgccataactaattacatgatttaaccTAAACATCAGATTCAATAGAGCCGCC
TCCACTG

CB 5' FWD

SEQ ID NO 32

Length: 56

Backbone|CBGA synthase|Flexible spacer|CBD syn-
thase|target peptide

Type: DNA

SEQ ID NO 34

Organism: Artificial Sequence

Notes: Primer

Length:

[0176]

Type: DNA

caccagaacttagtttcgacggataaa atggaaaccggtttgcctcgg
tttgac**[0178]** Organism: artificial sequence

Notes: Codon optimized

```

1  ggtaaataca tgtaattagt tatgtcacgc ttacattcac gccctcccc cacaaccgct
61  ctaaccgaaa aggaaggagt tagacaacct gaagtctagg tccctattta tttttttata
121 gttatgtagg tattaagaac gttatttata tttcaaattt ttcttttttt tctgtacaga
181 cgogtgtacg catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg
241 aaggctttaa tttgcggccc ctcaacctgca cgcaaaatag gataattata ctctatttct
301 caacaagtaa ttggttggtt ggccgagcgg tctaaggcgc ctgattcaag aaatatcttg
361 accgcagtta actgtgggaa tactcaggta tcgtaagatg caagagttag aatctcttag
421 caaccattat ttttttcctc aacataacga gaacacacag gggcgctatc gcacagaatc
481 aaattcgatg actggaaatt ttttgtaat ttcagaggtc gectgacgca tatacctttt
541 tcaactgaaa aattgggaga aaaaggaaag gtgagagcgc cggaaccggc ttttcatata
601 gaatagagaa gcgttcatga ctaaagctt gcacacacat acttgaagtt gacaatatta
661 ttaaggacc tattgttttt tccaataggt ggtagcaat cgtcttactt ttaactttt
721 cttacctttt acatttcagc aatatatata tatatatttc aaggatatac cattctaagt
781 tctgccctta agaagatcgt cgttttgcca ggtgaccacg ttggtcaaga aatcacagcc
841 gaagcatta aggttcttaa agctatttct gatgttcggt ccaatgtcaa gttcgatttc
901 gaaaatcatt taattggtgg tgctgctatc gatgctacag gtgttccact tccagatgag

```

-continued

961 gcgctggaag cctccaagaa ggctgatgcc gttttgttag gtgctgtggg tggctcctaaa
1021 tggggtaccg gtagtgttag acctgaacaa ggtttactaa aaatccgtaa agaacttcaa
1081 ttgtacgcca acttaagacc atgtaacttt gcatccgact ctcttttaga cttatctcca
1141 atcaagccac aatttgcctaa aggtactgac ttcgttgttg tcagagaatt agtgggaggt
1201 atttactttg gtaagagaaa ggaagatgat ggtgatggg tgcgttggga tagtgaacaa
1261 tacaccgttc cagaagtgca aagaatcaca agaatggccg ctttcatggc cctacaacat
1321 gagccaccat tgcctatttg gtccttggat aaagctaag ttttggcctc ttcaagatta
1381 tggagaaaaa ctgtggagga aaccatcaag aacgaattcc ctacattgaa ggttcaacat
1441 caattgattg attctgcccg catgatccta gttaagaacc caaccacct aaatggtatt
1501 ataatcacca gcaacatggt tggatgatc atctccgatg aagcctccgt tatcccaggt
1561 tccttggggt tgttgccatc tgcgtccttg gcctctttgc cagacaagaa caocgcattt
1621 ggtttgtacg aaccatgcca cggttctgct ccagatttgc caaagaataa ggtcaaccct
1681 atcgccacta tcttgtctgc tgcaatgatg ttgaaattgt cattgaactt gcctgaagaa
1741 ggtaaggcca ttgaagatgc agttaaaaag gttttgatg caggcatcag aactggtgat
1801 ttaggtggtt ccaacagtag caccgaagtc ggtgatgctg tcgccgaaga agttaagaaa
1861 atccttgctt aaaaagattc tcttttttta tgatatttgt acataaactt tataaatgaa
1921 attcataata gaaacgacac gaaattacaa aatggaatat gttcataggg taacgctatg
1981 atccaatc aaaggaaatg atagcattga aggatgagac taatccaatt gaggagtggc
2041 agcatataga acagctaaaag ggtagtgtg aaggaagcat acgatacccc gcatggaatg
2101 ggataatc acaggaggta ctgactacc tttcatccta cataaataga cgcataataag
2161 tacgcattta agcataaaca cgcactatgc cgttcttctc atgtatata atatacaggc
2221 aacacgcaga tataggtgcy acgtgaacag tgagctgtat gtgcgcagct cgcgttgcat
2281 tttcggaagc gctcgtttc gaaacgctt tgaagttcct attccgaagt tcctattctc
2341 tagaaagtat aggaacttca gagcgctttt gaaaaccaa agcgcctcga agtcgcactt
2401 tcaaaaaacc aaaaacgcac cggactgtaa cgagctacta aaatattgcy aataccgctt
2461 ccacaacat tgctcaaaaag tatctctttg ctatatact ctgtgtata tcctatata
2521 acctaccct ccaaccttgc ctccctgaac ttgcatctaa actcgacctc tacatttttt
2581 atgtttatct ctagtattac tctttagaca aaaaaattgt agtaagaact attcatagag
2641 tgaatcgaac acaatcgaac aatgtaaca tttcctatc gtagtatata gagacaaat
2701 agaagaaacc gttcataatt ttctgacca tgaagaatca tcaacgctat cactttctgt
2761 tcacaaagta tgcccaatcc acatcggat agaataaat cggggatgcc tttatcttga
2821 aaaaatgcac ccgagcttc gctagtaatc agtaaacgcy ggaagtggag tcaggctttt
2881 tttatggaag agaaaaataga caccaaatga gccttcttct aaccttaacy gacctacgt
2941 gcaaaaagt atcaagagac tgcattatag agcgcacaaa ggagaaaaaa agtaatctaa
3001 gatgctttgt tagaaaaata gcgctctcgg gatgcatttt tgtagaacaa aaaagaagta
3061 tagattcttt gttggtaaaa tagcgtctc gcgttgcaat tctgttctgt aaaaatgcag
3121 ctcagattct ttgtttgaaa aattagcgt ctcgcttgc atttttgttt taaaaaatg
3181 aagcacagat tcttcgttgg taaaatagcy ctttcgctt gcatttctgt tctgtaaaaa
3241 tgcagctcag attctttgtt tgaaaaatta gcgctctcgc gttgcatttt tgttctacaa

-continued

3301 aatgaagcac agatgcttcg ttcaggtggc acttttcggg gaaatgtgcg cggaaccct
3361 atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca ataacctga
3421 tattggtcag aattgggtaa ttggttgtaa cactgacccc tatttgttta ttttctaaa
3481 tacattcaaa tatgtatccg ctcatgagac aataaccctg ataatgctt caataatatt
3541 gaaaaaggaa gaatatgagc catattcaac gggaaacgtc gaggcgcgca ttaaatcca
3601 acatggatgc tgatttatat gggataaaat gggctcgcga taatgtcggg caatcagggtg
3661 cgacaatcta tcgcttgat ggggaagccc atgcgccaga gttgtttctg aaacatggca
3721 aaggtagcgt tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat
3781 ttatgccact tccgaccatc aagcatttta tccgtactcc tgatgatgca tggttactca
3841 ccaactgcat ccccgaaaa acagcgttcc aggtattaga agaatatcct gattcagggtg
3901 aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt gcaactcgatt cctgtttgta
3961 attgtccttt taacagcgat cgcgtatttc gcctcgcgca ggcgcaatca cgaatgaata
4021 acggtttggt tgatgcgagt gattttgatg acgagcgtaa tggctggcct gttgaacaag
4081 tctgaaaaga aatgcataaa cttttgcat tctcaccgga ttcagtcgct actcatggtg
4141 atttctcact tgataacctt attttgacg aggggaaatt aataggttgt attgatggtg
4201 gacgagtcgg aatcgcagac cgataccagg atcttgccat cctatggaac tgctcgggtg
4261 agttttctcc ttcattacag aaacggcttt tcaaaaata tggatattgat aatcctgata
4321 tgaataaatt gcaatttcat ttgatgctcg atgagttttt ctaactcatg accaaaaatc
4381 cttaacgtga gttacgcgcy cgtcgttcca ctgagcgtca gaccccgtag aaaagatcaa
4441 aggatcttct tgagatcctt ttttctgcy cgtaactctgc tgcttgcaaa caaaaaacc
4501 accgctacca gcggtgggtt gtttgccgga tcaagagcta ccaactcttt tccgaaaggt
4561 aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc cgtagttagc
4621 ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc
4681 agtggctgct gccagtgcyg ataagtcgty tcttaccggg ttggactcaa gacgatagtt
4741 accggataag gcgcagcgyt cgggctgaa cgggggttcg tgcacacagc ccagcttgga
4801 gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct
4861 tcccgaaggy agaaagcgc acaggtatcc ggtaagcggc agggtcggaa caggagagcy
4921 cacgaggyg cttccaggyg gaaacgcctg gtatctttat agtcctgctg ggtttcgcca
4981 cctctgactt gagcgtcgat ttttgatgag ctctcaggy gggcggagcc tatggaaaa
5041 cgccagcaac gcggtctttt tacggttctt ggccttttgc tggccttttg ctcaatggt
5101 ctttctgcy ttatccctg attctgtgga taaccgtatt accgcctttg agtgagctga
5161 taccgctcgy ggtcgtgcyg gtagtttatc attatcaata ctccgcttt caaagaatac
5221 gtaataaatt aatagtagty attttctaa ctttatttag tcaaaaaatt agccttttaa
5281 ttctgctgta acccgtacat gcccataata gggggcgggt tacacagaat atataacatc
5341 gtaggtgtct ggggtgaacg tttattctg gcacccacta aatataatgg agcccgtct
5401 ttaagctgcy atccagaaaa aaaaagaatc ccagcaccia aatattgttt tcttcaccia
5461 ccatcagttc ataggtccat tctcttagcy caactacaga gaacaggggc acaaacaggy
5521 aaaaacggg cacaacctca atggagtgat gcaaccagcc tggagtaaat gatgacacia
5581 ggcaattgac ccacgcatgt atctatctca ttttcttaca ctttctatta ctttctgctc

-continued

5641 tctctgattt ggaaaaagct gaaaaaaaaag gttgaaacca gttccctgaa attattcccc
5701 tacttgacta ataagtatat aaagacggta ggtattgatt gtaattctgt aaatctattt
5761 cttaaacttc ttaaattcta cttttatagt tagtcttttt tttagtttta aaacaccaga
5821 acttagtttc gacggataaa atggaaaccg gtttgctctc ggtttgcaact ttctccttcc
5881 aaacaaacta tcatacactc ctgaaccgc acaataacaa tcccaaaact tcctgctgtg
5941 gttataggca cccaagaca ccaatcaaat actcctacaa taactttcca tctaagcatt
6001 gtagcacaaa aagtttccat ttgcaaaata agtgttccga atctctgtcc atcgcaaaa
6061 attccattag ggctgccact actaatcaaa ctgaaccacc agagtctgat aatcattctg
6121 tcgccacaaa gattctgaat tttgggaagg cttgttgga gttacaaaga ccatatacaa
6181 ttattgcctt tacctcttgt gcctgtggtt tatttggtaa ggaactgttg cataatacaa
6241 atttaatatc ttggtcattg atggaaacgt tcaaagcatt tttttctta gtogctatcc
6301 tttgtattgc ttctttcacc accactatca accagattta cgacttacat attgacagaa
6361 ttaacaagcc agatttgcca ctggcttcgg gcgagatttc cgtcaatact gcctggatca
6421 tggaaacttc tattattgtt gccttgtttg gattgataat caccataaaa atggaaacta
6481 aggggtggtc attgtatatt ttcggttact gttttggtat cttcgggggc atcgctact
6541 ctgttcctcc attcagatgg aaacaaaatc cttccacagc attccttttg aacttcctgg
6601 cgcacattat aaccaacttt actttttatt atgcctccag agccgcctg gggctgcct
6661 ttgaattacg cccctccttt acatttttac tggccttcat ggagaccaag tccatggaga
6721 ctggttctgc tctcgcgttg atcaaagatg cttccgatgt ggaaggtgac accaaatttg
6781 gtatatccac tttggccagc aagtatggtt ccaggaattt gaccctattt tgttctggt
6841 tcgtgctgct gtcttatgtt gcagccatct tggctggcat catttggcca caggctttca
6901 attcaaatgt tatggagacg ctgctctcgc atgctatttt ggcatttttg ttgattctac
6961 agacaagaga ttttgcttta accaattatg acccagaagc tggtagaaga ttttcaaat
7021 ttatggaaac atggaaatta tactatgctg aatatttagt gtacgttttc attgggggcg
7081 gctccagcgc cggcgggcgc tcttctcgg gcggttggtc tcatccacaa tttgagaaag
7141 gtgggtcgtc tggcggcggc agcggggcg ggtccggcg ggggagcggc ggtatgaaat
7201 gttcgacctt ctctttttgg tttgtctgta aaataatttt tttttcttc agctttaaca
7261 ttcaaacagc cattgcaaat ccaagagaaa atttcttgaa atgcttttca caatatatcc
7321 ccaataatgc tactaacttg aagctagttt atactcaaaa caacccttg tacatgtccg
7381 tgctcaactc caccattcac aacctaatg tcacttcaga cactaccca aaaccattag
7441 ttattgtgac accttctcac gtttcacata tccaaggtac tattttatgc tccaagaagg
7501 tcggcctgca aattagaact agatctggag gtcattgattc agaaggaatg tcttacatct
7561 ctcaagttcc atttgtgatt gtcgatttaa gaaatagag gagcattaag atcgatgttc
7621 actcccaaac ggcattgggt gaagccggtg ccaccttggg cgaagtttac tactgggtca
7631 acgagaagaa tgaaaaacta tcaactagccg caggttattg tccaactggt tgtgctggtg
7741 gccatttcgg agggcggcgc tacggctctc taatgagaaa ctacggctta gctgctgaca
7801 atatcatcga cgctcacttg gttaacgttc atggtaaagt tttagataga aaatctatgg
7861 gtgaggatct tttctgggct ttgagaggtg gcggcgcaga atcatttggc attatcggtg
7921 cttggaagat cagattggtg gctgtcccca agtctacaat gttttctgtg aagaaaatta

- continued

```

7981  tggaaatcca tgaattggct aaactgggtga ataaatggca aaacatagct tacaagtacg
8041  ataaagactt gctgttaatg acacatttta ttaccaggaa catcactgat aaccaaggca
8101  agaacaagac tgcaattcat acttattttt cctccgtttt tttgggtggg gtcgactccc
8161  tcgtggatct gatgaataaa tcattccctg aactagggtat taaaaaaacc gattgtagac
8221  aattgagttg gattgatacc atcatattct acagtgggtg tgtaattat gatactgaca
8281  acttcaacaa agaaatactg ctggaccggt ccgcccggcca gaatgggtgct tttaaaatca
8341  agttggatta tgtgaaaaag cctattccag aatccgtatt tgttcaata ttggaaaagc
8401  tgtatgaaga agacattggg gcaggcatgt acgctcttta tccttatggc ggcataatgg
8461  atgaaatttc tgaaagtgcc attcctttcc cacatagggc cgggatcctg tacgagttat
8521  ggtacatttg ttcatgggaa aagcaagaag ataataaaaa acatttaaat tggataagaa
8581  atatttataa ttttatgact ccatacgtct ccaaaaaacc acgcttgcca tatttgaatt
8641  acagagacct ggatattggc atcaatgatc ctaaaaaacc aaataattac actcaggcaa
8701  gaatatgggg tgaaaaatat ttccggcaaaa attttgatag gctggccaag gttaaaacac
8761  tggttgatcc aaacaatttc ttagaaaacg aacaatctat cccacctctg cctagacata
8821  gacacggcgg tggaagcagt ggaggcggct ctattgaatc tgatgtttaa tga

```

[0179] Backbone|OLS|Flexible spacer|OAC|target peptide

SEQ ID NO 35

Length:

Type: DNA

[0180] Organism: artificial sequence

Notes: Codon optimized

```

1   ggtaaataca tgtaattagt tatgtcacgc ttacattcac gcctcctccc cacatccgct
61  ctaaccgaaa aggaaggagt tagacaacct gaagtctagg tcctatttta tttttttata
121 gttatgtagg tattaagaac gttatttata tttcaaattt ttcttttttt tctgtacaga
181 cgcggtgtacg catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg
241 aaggctttaa tttgcggccc ctcacctgca cgcaaaaagc ttttcaattc aattcatcat
301 ttttttttta ttcttttttt tgatttcggg ttctttgaaa tttttttgat tcggtaatct
361 ccgaacagaa ggaagaacga aggaaggagc acagacttag attggtatat atacgcatat
421 gtagtggtga agaaacatga aattgccagc tattcttaac ccaactgcac agaacaaaaa
481 ccagcaggaa acgaagataa atcatgtcga aagctacata taaggaacgt gctgctactc
541 atcctagtcc tggtgctgcc aagctattta atatcatgca cgaaaagcaa acaaacttgt
601 gtgcttcatt ggatgttcgt accaccaagg aattactgga gttagttgaa gcattaggtc
661 caaaaatttg tttactaaaa acacatgtgg atatcttgac tgatttttcc atggaggggca
721 cagttaagcc gctaaaggca ttatccgcca agtacaattt tttactcttc gaagatagaa
781 aatttgctga cattggtaat acagtcaaat tgcagtactc tgcgggtgta tacagaatag
841 cagaatgggc agacattacg aatgcacacg gtgtgggtggg cccagggtatt gttagcgggt
901 tgaagcaggc ggcagaagaa gtaacaaagg aacctagagg ccttttgatg ttagcagaat
961 tgtcatgcaa gggctcceta tctactggag aatatactaa gggtactggt gacattgcca

```

-continued

1021 aaagcgacaa agattttggt atcggcttta ttgctcaaag agacatgggt ggaagagatg
1081 aaggttacga ttggttgatt atgacacccg gtgtgggttt agatgacaag ggagatgcat
1141 tgggtcaaca gtatagaacc gtggatgatg ttgtctctac aggatctgac attattattg
1201 ttggaagagg actatattgca aagggaaagg atgctaaggt agagggtgaa cgttacagaa
1261 aagcaggctg ggaagcatat ttgagaagat gcgccagca aactaaaaa actgtattat
1321 aagtaaatgc atgtatacta aactcacaaa ttagagcttc aatttaatta tatcagttat
1381 taccacgct atgatccaat atcaaaggaa atgatagcat tgaaggatga gactaatcca
1441 attgaggagt ggcagcatat agaacagcta aagggtagt ctgaaggaag catacgatac
1501 cccgcatgga atgggataat atcacaggag gtactagact acctttcatc ctacataaat
1561 agacgcatat aagtacgcat ttaagcataa acacgcacta tgccgttctt ctcatgtata
1621 tatatataca ggcaacacgc agatataggt ggcagctgaa cagttagctg tatgtgcgca
1681 gctcgcgttg cattttcggga agcgcctcgt ttcggaaaacg ctttgaagtt cctattccga
1741 agttcctatt ctctagaaaag tataggaact tcagagcgt tttgaaaacc aaaagcgtc
1801 tgaagtcgca ctttcaaaaa accaaaaacg caccggactg taacgagcta ctaaaatatt
1861 gcgaataccg cttccacaaa cattgctcaa aagtatctct ttgctatata tctctgtgct
1921 atatccctat ataacctacc catccacctt tcgctccttg aacttgcatc taaactcgac
1981 ctctacattt tttatgttta tctctagtat tactctttag acaaaaaaat tgtagtaaga
2041 actattcata gagtgaatcg aaaacaatac gaaaatgtaa acatttccta tacgtagtat
2101 atagagacaa aatagaagaa accgttcata atttctgac caatgaagaa tcatcaacgc
2161 tatcactttc tgttcacaaa gtatgcgcaa tccacatcgg tatagaatat aatcggggat
2221 gcctttatct tgaaaaaatg caccgcagc ttcgctagta atcagtaaac gcgggaagtg
2281 gagtcaggct tttttatgag aagagaaaaat agacacccaaa gtagccttct tctaacctta
2341 acggacctac agtgcaaaaa gttatcaaga gactgcatta tagagcgac aaaggagaaa
2401 aaaagtaatc taagatgctt ttttagaaaa atagogctct cgggatgcat tttttagaaa
2461 caaaaaagaa gtatagattc tttgttggtg aaatagcgt ctgcggttgc atttctgttc
2521 tgtaaaaatg cagctcagat tctttgtttg aaaaattagc gctctcgcgt tgcatttttg
2581 ttttacaaaa atgaagcaca gattctcgt tggtaaaata gcgcttctgc gttgcatttc
2641 tgttctgtaa aaatgcagct cagattcttt gttgaaaaa ttagcgtctc cgcggtgcat
2701 ttttgttcta caaaatgaag cacagatgct tcgctcaggt ggcacttttc ggggaaatgt
2761 gcgcggaacc cctattttgt tatttttcta aatacattca aatagtatc cgctcatgag
2821 acaataaacc tgatattggt cagaattggt taattggtg taacactgac ccctatttgt
2881 ttatttttct aaatacattc aaatagtat ccgctcatga gacaataacc ctgataaatg
2941 cttcaataat attgaaaaag gaagaatag agtattcaac atttccgtgt cgccttatt
3001 cccttttttg cggcattttg ccttctgtt tttgctcacc cagaaacgct ggtgaaagta
3061 aaagatgctg aagatcagtt ggggtgcacga gtgggttaca tcgaactgga tctcaacagc
3121 ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa
3131 gttctgctat gtggcgcggt attatccgt attgacgccc ggcaagagca actcggctgc
3241 cgcatacact attctcagaa tgacttggtt gactactcac cagtacagaa aaagcatctt
3301 acggatggca tgacagtaag agaattatgc agtgctgcca taacatgag tgataaacct

-continued

3361 gcggccaaact tactttctgac aacgatcggg ggaccgaagg agctaaccgc ttttttgca
3421 aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata
3481 ccaaacgacg agcgtgacac cacgatgcct gttagcgtgg caacaacggt gcgcaaaacta
3541 ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg
3601 gataaagtgt caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat
3661 aaatccggag ccggtgagcg tggttctcgc ggtatcatcg cagcgtggg gccagatggt
3721 aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga
3781 aatagacaga tcgctgagat aggtgcctca ctgattaage attggttaact catgacaaa
3841 atcccctaac gtgagttacg cgcgcgtcgt tccactgagc gtcagacccc gtagaaaaga
3901 tcaaaggatc tccttgagat ccttttttcc tgcgcgtaat ctgctgcttg caaacaaaa
3961 aaccaccgct accagcgggt gtttgtttgc cggatcaaga gctaccaact ctttttccga
4021 aggtaactgg cttcagcaga gcgcagatac caaatactgt tcttctagt tagccgtagt
4081 tagcccacca cttcaagaac tctgtagcac cgcctacata cctcgtctg ctaatcctgt
4141 taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttgac tcaagacgat
4201 agttaccgga taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct
4261 tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca
4321 cgcttcccga agggagaaa ggcgacaggt atccggtaag cggcagggtc ggaacaggag
4381 agcgcacgag ggagcttcca ggggaaacg cctggatatc ttatagtcct gcgggttcc
4441 gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcg agcctatgga
4501 aaaacgccag caacgcggcc tttttacggt tcttggcctt ttgctggcct tttgctcaca
4561 tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag
4621 ctgataccgc tcggggctgt gcaggtatag cttcaaaatg tttctactcc ttttttactc
4681 ttccagattt tctcggactc cgcgcacatc cgtaccactt caaacacccc aagcacagca
4741 tactaaattt cccctcttcc tctctcagg gtgtcgttaa ttaccgtagc taaaggttg
4801 gaaaagaaaa aagtgaccgc ctcgtttctt tttcttcgtc gaaaaaggca ataaaaattt
4861 ttatcagctt tctttttctt gaaaattttt ttttttgatt tttttctctt tcgatgacct
4921 cccattgata ttaagttaa taaacgact tcaatctctc aagtttcagt ttcatttttc
4981 ttgttctatt acaacttttt ttacttcttg ctcattagaa agaaagcata gcaatcta
5041 ctaagttaa aatgaatcat ttgagagcag aagggcctgc tccgctgctg gctattggta
5101 ccgccaatcc agaaaaatc ctgctgcagg acgaattccc agattactat tttagggtca
5161 ccaaatctga acatatgaca caattgaaag agaaattcag aaagatttgt gacaagtcca
5221 tgattaggaa aagaaattgt tttttgaaat aagaacactt gaagcaaat cctcgcctgg
5281 tggagcatga aatgcaaac ttggatgcta gacaagacat gttggtgggt gaagttccaa
5341 agctggggaa ggatgcctgt gcccaaggcca ttaaagaatg gggccaacca aaatccaaaa
5401 ttcccacct gattttcacc tccgcctcca cactgatat gccaggtgca gactatcatt
5461 gtgctaatt gttgggtttg tccccctccg tgaagagagt tatgatgat caattaggtt
5521 gttatggcgg cggcaccggt ctgagaattg ccaagacat tgctgaaac aataaagggtg
5581 cgcgcgtttt ggctgtttgt tgtgatatta tggcatgttt atttagaggt ccaagtgaaa
5641 gtgacttggg attgctagtg ggcagggcca tatttgggtg tgggtgcgct gctgtgatcg

-continued

5701 ttgggtgctga gectgatgaa tctgtcggtg aaagaccaat tttggaactg gtttccactg
5761 gtcaaaccat tttgccaaat tcagaaggta ctattggcgg ccatatcaga gaagctgggt
5821 taatctttga tttgcacaag gatgtcccaa tgtaatttc caataatatt gaaaaatggt
5881 tgatcgaagc atttaccccc atcgggtatt ctgattggaa ttccatcttc tggattacac
5941 atcctggcgg taaagctatc ttagataaag ttgaggagaa gttgcattta aagtctgaca
6001 aatttggtga ttcaagacat gtccctgtctg agcacggtaa tatgtcttcc tgcaccgtct
6061 tgtttgcat ggatgagttg aggaagaggt ccctggaaga aggcaagagc accaccgggtg
6121 acggttttga gtggggggct ctctttggat ttggggccagg cctgaccgta gaaaggggtg
6181 ttgtccgctc ggtgccaatc aaatatggtg ggggggtccag cgcgggtggc gggagctccg
6241 cgggcgggtg gtctcaccca caatttgaaa aggggtggcag cagcggcggc ggctctggcg
6301 gaggtctcgg cgggggctcg ggggggtatg ctgtcaagca tctgatcgtg ctgaagttca
6361 aagatgaaat tactgaagcc caaaaggagg aatttttcaa gacatatggt aatttggtta
6421 acatcattcc agcaatgaaa gatgtttatt ggggtaagga cgttactcaa aaaaataagg
6481 aagagggtta cactcatatt gttgaagtca ctttcgaatc cgtcgaaca atcaagatt
6541 atattattca tccagctcat gttgggtttg gcgatgtgta cagatcattt tgggaaaaat
6601 tattgatttt tgactacaca ccaagaaaag gcggtggaag cagtggaggc ggctctattg
6661 aatctgatgt ttaatag

Overexpression of ERG8, HFA1, ERG 10, ERG13, tHMGR, HMGR, ERG12, ERG8, IDI Genes (for Higher Levels of Intermediates)

Same process as expression of Synthase expression, but with 3 copies expressed in yeast cells.

Backbone|GGPS1|2a protease|HMG-CoA reductase|flexible spacer|IDI1

SEQ ID NO 36

Length:

Type: DNA

[0181] Organism: artificial sequence

Notes: Codon optimized

1 atggagaaga ctcaagaaac agtccaaaga attcttctag aacctataa atacttactt
61 cagttaccag gtaaacaagt gagaacaaa ctttcacagg catttaatca ttggctgaaa
121 gttccagagg acaagctaca gattattatt gaagtgcagc aaatggtgca taatgccagt
181 ttactcatcg atgatattga agacaactca aaactccgac gtggcttctc agtggccacc
241 agcatctatg gaatcccatc tgtcatcaat tctgccaatt acgtgtattt ccttggcctg
301 gagaaagtct taacccttga tcaccagat gcagtgaagc tttttaccgg ccagcttttg
361 gaactccatc agggacaagg cctagatatt tactggaggg ataattacac ttgtcccact
421 gaagaagaat ataaagctat ggtgctgcag aaaacagggt gactgtttgg attagcagta
481 ggtctcatgc agttgttctc tgattacaaa gaagatttaa aaccgctact taatacactt
541 gggctctttt tccaaattag ggatgattat gctaactctac actccaaaga ataatagtaa
601 aacaaaagt tttgtgaaga tctgacagag ggaaagtctc catttctctac tattcatgct
661 atttggtcaa ggccgtgaaag caccaggtg cagaatatct tgcgccagag aacagaaaa
721 atagatataa aaaaataactg tgtacattat cttgaggatg taggttcttt tgaatacact
761 cgtaatacce ttaaagagct tgaagctaaa gcctataaac agattgatgc acgtgggtggg
841 aaccctgagc tagtagcctt agtaaacac ttaagtaaga tgttcaaaga agaaaatgaa
901 ggcggttctg gcagcggaga gggcagagga agtcttctaa catgcgggtg cgtggaggag

-continued

961 aatccccggcc ctaggctctgg cagcgggagag ggcagaggaa gtcttctaac atgcggtgac
1021 gtggaggaga atccccggccc taggacacaa aagaaagtcc cagacaattg ttgtagacgt
1081 gaacctatgc tggtcagaaa taaccagaaa tgtgattcag tagaggaaga gacagggata
1141 aaccgagaaa gaaaagttag gggtataaaa cccttagtgg ctgaaacaga taccocaaac
1201 agagctacat ttgtggttgg taactcctcc ttactcgata ctcatcagt actggtgaca
1261 caggaacctg aaattgaact toccagggaa cctcggccta atgaagaatg tctacagata
1321 cttgggaatg cagagaaaag tgcaaaatcc cttagtgatg ctgagatcat ccagttagtc
1351 aatgctaagc atatcccagc ctacaagtgg gaaactctga tggaaactca tgagcgtggt
1441 gtatctatcc gccgacagtt actttccaag aagctttcag aacctcttc tctccagtac
1501 ctacctaca gggattataa ttactccttg gtgatgggag cttggttga gaatgttatt
1561 ggatatatgc ccatccctgt tggagtggca ggacccttt gcttagatga aaaagaattt
1621 caggttccaa tggcaacaac agaaggttgt cttgtggcca gcaccaatag aggctgcaga
1681 gcaataggtc ttggtggagg tgccagcagc cgagtccttg cagaCgggat gactcgtggc
1741 ccagttgtgc gtcttccacg tgcttgtgac tctgcagaag tgaagcctg gctcgaaca
1801 tctgaagggt tcgcagtgat aaaggaggca tttgacagca ctagcagatt tgcacgtcta
1861 cagaaacttc atacaagtat agctggacgc aacctttata tccgtttcca gtccaggtea
1921 ggggatgcca tgggatgaa catgatttca aagggtacag agaaagcact ttcaaaactt
1981 cagcagtatt tccctgaaat gcagattcta gccgttagtg gtaactattg tactgacaag
2041 aacctgctg ctataaattg gatagagga agaggaaat ctggtgttg tgaagctgtc
2101 attccagcca aggttctcag agaagtatta aagactacca cagaggctat gattgaggtc
2161 aacattaaca agaatttagt gggctctgcc atggctggga gcataggagg ctacaacgcc
2221 catgcagcaa acattgtcac cgccatctac attgcctgtg gacaggatgc agcacagaat
2281 gttggtagt caaactgtat tactttaatg gaagcaagtg gtcccacaaa tgaagattta
2341 tatatcagc gcacctgccc atctatagag ataggaacgg tgggtggtgg gaccaaccta
2401 ctacctcagc aagcctgttt gcagatgcta ggtgttcaag gagcatgcaa agataatcct
2461 ggggaaaatg cccgcagct tgcccgaatt gtgtgtggga ccgtaatggc tggggaattg
2521 tcacttatgg cagcattggc agcaggacat cttgtcaaaa gtcacatgat tcacaacagg
2581 tcgaagatca atttacaaga cctccaagga gcttgcacca agaagacagc cggctcagga
2641 ggttcttcag gactggaagt gctgtttcag ggcccgggtg gatctggcat gatgcctgaa
2701 ataaacacta accacctcga caagcaacag gttcaactcc tggcagagat gtgtatcctt
2761 attgatgaaa atgacaataa aattggagct gagaccaaga agaattgtca cctgaacgag
2821 aacattgaga aaggattatt gcacagagct tttagtgtct tcttatcaaa caccgaaaat
2881 aagcttctgc tacagcaaaag atcagatgct aagattacct ttccagggtg ttttacgaat
2941 acgtgttga gtcatccatt aagcaatcca gccgagcttg aggaaagtga cggccttggg
3001 gtgaggcgag cagcacagag acggctgaaa gctgagctag gaattccctt ggaagaggtt
3061 cctccagaag aaattaatta tttaacacga attcactaca aagctcagtc tgatggtatc
3121 tggggtgaac atgaaaattg ttacattttg ttggtgagga agaattgtaac tttgaatcca
3181 gatcccaatg agattaanaag ctattgttat gtgtcaaagg aagaactaaa agaacttctg
3241 aaaaaagcag ccagtggtga aattaagata acgccatggt ttaaaattat tgcagcgact

-continued

3301 tttctcttta aatggtggga taacttaaat catttgaatc agtttggtga ccatgagaaa
 3361 atatacagaa tg

ERG2, ERG3, and ERG6 mutations/deletions for increased membrane permeability

[0182] Same process as ERG9 knockout, but targeting ERG2, ERG3, and ERG6 genes.

ERG20p Modification

[0183] We experimented with a few types of ERG20 genes, (avian, salmon entrica, and human). Currently we are still trying to see which is the best by engineering the ERG20p gene into a FPP synthase, thereby creating a new enzyme that can create GPP instead at high rates.

Gal80p Deletion/Mutation for not Needing Expensive Galactose to Induce Promoter

[0184] Induce mutation in Gal80 gene by site directed mutagenesis.

Using ADH2p (Glucose Repressible Promoter) Induces Strong Transcription with No Glucose, Better than GAL Promoter

[0185] Same process as Gal80p deletion.

Overexpression of ADH2 and ALD6 (Ethanol to Acetate), as Well as Overexpression of an Acetyl-CoA C-Acetyltransferase

[0186] Same process as IDI and HMGR overexpression, but with genes for ADH2 and ALD6.

Tables

[0187] Below is a table of various cannabinoids, along with their structure and variants and main pharmacological characteristics as well as a table listing potential starting materials.

TABLE 1

Compounds	Pharmacological Characteristics
Cannabinoids (FIG. 1 and 2)	
Cannabigerolic acid (CBGA)	Antibiotic (1)
Cannabigerolic acid monomethylether (CBGAM)	
Cannabigerol (CBG)	Antibiotic, antifungal, anti-inflammatory, analgesic (1) Partial agonist at CB1/CB2 receptors (2)
Cannabigerovarinic acid (CBGVA)	
Cannabigerovarin (CBGV)	
Cannabichromenic acid (CBCA)	
Cannabichromene (CBC)	Anti-inflammatory, antibiotic, antifungal, analgesic (1)
Cannabichromevarinic acid (CBCVA)	
Cannabichromevarin (CBCV)	
Cannabidiolic acid (CBDA)	Antibiotic
Cannabidiol (CBD)	Anxiolytic, antipsychotic, analgesic, anti-inflammatory, antioxidant, antispasmodic (1) Ant schizophrenic, antiepileptic, sleep-promoting, anti-oxidizing, anti-inflammatory, immunomodulation properties (2)

TABLE 1-continued

Compounds	Pharmacological Characteristics
Cannabidiol monomethylether (CBDM)	
Cannabidiol-C4 (CBD-C4)	
Cannabidivarinic acid (CBDVA)	
Cannabidivarin (CBDV)	
Cannabidioreol (CBD-C1)	
Tetrahydrocannabinolic acid A (THCA-A)	
Tetrahydrocannabinolic acid B (THCA-B)	
Delta-9-tetrahydrocannabinol (THC)	Euphoriant, analgesic, anti-inflammatory, antioxidant, antiemetic (1)
Delta-9-tetrahydrocannabinolic acid-C4 (THCA-C4)	
Delta-9-tetrahydrocannabinol-C4 (THC-C4)	
Delta-8-tetrahydrocannabivarin (D8-THCV)	Exhibit in vitro pharma properties similar to THCV, and both can antagonize THC; behave as agonists or antagonists in dose dependent manner (2)
Delta-9-tetrahydrocannabivarinic acid (THCVA)	
Delta-9-tetrahydrocannabivarin (THCV)	Analgesic, euphoriant (1) Strong antagonist of anandamide (due to interactions with non-CB1/2 receptors), neuromodulator (in animal and human organs), some affects due to interaction with non CB1/CB2 receptors (2)
Delta-9-tetrahydrocannabiorcolic acid (THCA-C1)	
Delta-9-tetrahydrocannabiorcol (THC-C1)	
Delta-7-cis-iso-tetrahydrocannabivarin (D7-THCV)	
Delta-8-tetrahydrocannabinolic acid (D8-THCA)	
Delta-8-tetrahydrocannabinol (D8-THC)	Similar to THC (1) Several 1-O-methyl- and 1-deoxy-delta-8-THC analogs have high CB2 receptor affinity [JWH133, JWH359, trans-(6aR,10aR)-3-(1,1-dimethylhexyl)-1-O-methyl-delta-8-THC]; antiemetic effects similar to THC (2)
Cannabicyclolic acid (CBLA)	
Cannabicyclol (CBL)	
Cannabicyclovarin (CBLV)	
Cannabielsoic acid A (CBEA-A)	
Cannabielsoic acid B (CBEA-B)	
Cannabielsoin (CBE)	
Cannabinolic acid (CBNA)	
Cannabinol (CBN)	Sedative, antibiotic, anticonvulsant, anti-inflammatory (1)

TABLE 1-continued

Compounds	Pharmacological Characteristics
Cannabinol methylether (CBNM)	
Cannabinol-C4 (CBN-C4)	
Cannabivarin (CBV)	
Cannabinol-C2 (CBN-C2)	
Cannabinol-C1 (CBN-C1)	
Cannabinodiol (CBND)	
Cannabinodivarin (CBVD)	
Cannabitriol (CBT)	
10-Ethoxy-9-hydroxy-delta-6a-tetrahydrocannabinol	
8,9-Dihydroxy-delta-6a-tetrahydrocannabinol	
Cannabitriolvarin (CBTV)	
Ethoxy-cannabitriolvarin (CBTVE)	
Dehydrocannabifuran (DCBF)	
Cannabifuran (CBF)	
Cannabichromanon (CBCN)	
Cannabicitran (CBT)	
10-oxo-delta-6a-tetrahydrocannabinol (OTHC)	
Delta-9-cis-tetrahydrocannabinol (Cis-THC)	
3,4,5,6-Tetrahydro-7-hydroxy-alpha-alpha-2-trimethyl-9-n-propyl-2,6-methano-2H-1-benzoxocin-5-methanol (OH-iso-HHCV)	
Cannabiripsol (CBR)	
Trihydroxy-delta-9-tetrahydrocannabinol (triOH-THC)	
Terpenes/Terpenoids	
Beta-Myrcene	Analgesic, anti-inflammatory, antibiotic, antimutagenic
d-Limonene	Immune potentiator, antidepressant, antimutagenic
Linalool	Sedative, antidepressant, anxiolytic, immune potentiator
Trans-Ocimene	
Beta-Pinene	
Alpha-Pinene	Anti-inflammatory, bronchodilator, stimulant, antibiotic, antineoplastic, AChE inhibitor
Beta-Caryophyllene	Anti-inflammatory, cytoprotective, antimalarial, CB2 agonist
Delta-3-Carene	
Pulegone	AChE inhibitor, sedative, antipyretic
Trans-gamma-Bisabolene	
Trans-alpha-Farnesene	
Beta-Fenchol	
Beta-Phellandrene	
Alpha-Humulene	
Guajol	
Alpha-Gualene	
Alpha-Eudesmol	
Terpinolene	
Alpha-Selinene	
Alpha-Terpineol	Sedative, antibiotic, AChE inhibitor, antioxidant, antimalarial
Fenchone	
Camphene	
Cis-Sabinene hydrate	
Cis-Ocimene	
Beta-Eudesmol	
Beta-Selinene	
Alpha-trans-Bergamolene	
Gamma-Eudesmol	
Borneol	
Cis-beta-Farnesene	
Gamma-Curcumene	
Cis-gamma-Bisabolene	
Alpha-Thujene	
Epi-alpha-Bisabolol	
Ipsdienol	
Alpha-Yiangene	

TABLE 1-continued

Compounds	Pharmacological Characteristics
Beta-Elemene	
Alpha-cis-Bergamontene	
Gamma-Murolene	
Alpha-Cadinene	
Alpha-Longipinene	
Caryophyllene oxide	
Spermidine Alkaloids (FIG. 6)	
(+)-Cannabisativine	
Palustridine	
Palustrine	
Spermidine	
Anhydrocannabisativine	
Phenolic Amides and Lignanamides (FIG. 5)	
N-trans-Feruloyltyramine	
N-p-Coumaroyltyramine	
N-trans-Caffeoyltyramine	
Grossamide	
Cannabisin-A	
Cannabisin-B	
Cannabisin-C	
Cannabisin-D	
Cannabisin-E	
Cannabisin-F	
Cannabisin-G	
Phenylpropanoids and Flavonoids (FIG. 4)	
Apigenin	
Luteolin	
Kaempferol	
Quercetin	
Orientin	
Vitexin	
Cannflavin A	Inhibit prostaglandin E2 in human rheumatoid synovial cells
Cannflavin B	Inhibit prostaglandin E2 in human rheumatoid synovial cells
Stilbenoids (FIG. 3)	
Cannabispiran	
Isocannabispiran	
Cannabistilbene-IIa	
Cannabistilbene-IIb	
Cannithrene-1	
Cannithrene-2	
Acetyl cannabispireol	
Alpha-cannabisporeanol	
Canniprene	
Cannabispiron	

TABLE 2

(Starting Materials)		
Sugar based concentrates (High Fructose Corn Syrup, Molasses)	Hemicellulose	Glycerol
Glucose	Xylose	Whey
Sucrose	Methanol	Biodiesel
Cellulose	Lactic Acid	Citrate
Ethanol	Lignin	Fructose
Succinic Acid	Arabinose	Biofuels
Biomass	Saccharose	Starch based products
Agricultural residue	Water hyacinth	Aquatic biomass

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 36

<210> SEQ ID NO 1
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 1
gtacatttca tagcccatct tcaacaacaa taccgactta cccgtacgct gcaggtcgac 60

<210> SEQ ID NO 2
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 2
cagattgacg gagagagggc cacattgttt gtcggcaata aatcgatgaa ttcgagctcg 60

<210> SEQ ID NO 3
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 3
atgggtaaaa agcctgaact ca 22

<210> SEQ ID NO 4
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 4
ttattccttt gcctcggac gag 23

<210> SEQ ID NO 5
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 5
agatgctagt caatggcaga ag 22

<210> SEQ ID NO 6
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 6
tgcttacaca gagtgaacct gc 22

<210> SEQ ID NO 7

-continued

<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 7

ctcgtggaag tgacgcaac 19

<210> SEQ ID NO 8
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 8

cgggatccag tttatcatta tcaatactcg cc 32

<210> SEQ ID NO 9
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 9

ggggcggccg cgagctcagt ttatcattat c 31

<210> SEQ ID NO 10
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 10

ggggcggccg caaaacaatg ttgtcacgac tttccgatat gc 42

<210> SEQ ID NO 11
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 11

gactagtcca agctgacttc ttggtgcacg ttccttg 37

<210> SEQ ID NO 12
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 12

atgtctgctg ttaacgttgc acctg 25

<210> SEQ ID NO 13
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

-continued

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 13

ttaaccaatc aactcaccaa ac

22

<210> SEQ ID NO 14

<211> LENGTH: 38

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 14

cgggatccct cgagttgttc gctgctgaca gcgataac

38

<210> SEQ ID NO 15

<211> LENGTH: 44

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 15

cgggatccgc tagcgggtacc acatgggtcc tttatattga cacg

44

<210> SEQ ID NO 16

<211> LENGTH: 23

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 16

atcagaacaa ttgtccagta ttg

23

<210> SEQ ID NO 17

<211> LENGTH: 21

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 17

aatgtactat acaagccttc c

21

<210> SEQ ID NO 18

<211> LENGTH: 40

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 18

ggggcggccg caaaacaatg gggatgcttc gctggggagt

40

<210> SEQ ID NO 19

<211> LENGTH: 31

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 19

-continued

gactagttta gctcctcaat tcgtcaaagg t 31

<210> SEQ ID NO 20
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 20

ggggcggccg caaaacaatg gacatgttca gggatcgcca gg 42

<210> SEQ ID NO 21
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 21

gactagtcta atgccatct tccagcaggc g 31

<210> SEQ ID NO 22
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 22

cgggatccat aacttcgat agcatacatt atacgaagtt atgtggaata ttcggatat 60

<210> SEQ ID NO 23
<211> LENGTH: 56
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 23

gcatacaatc aactaagcta agctaaaaca atgggtaagg aaaagactca cgtttc 56

<210> SEQ ID NO 24
<211> LENGTH: 56
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 24

gaaacgtgag tcttttcctt acccattggt ttagcttagc ttagttgatt gtagtc 56

<210> SEQ ID NO 25
<211> LENGTH: 53
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 25

catttgatgc tcgatgagtt tttctaaatc cgctctaacc gaaaaggaag gag 53

<210> SEQ ID NO 26

-continued

<211> LENGTH: 53
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 26

ctccttcctt ttcggttaga gcggatttag aaaaactcat cgagcatcaa atg 53

<210> SEQ ID NO 27
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 27

ggggctagca taacttcgta taatgtatgc tatacgaagt tatcttcgag cgtcccaaaa 60

<210> SEQ ID NO 28
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 28

cgggatccag tttatcatta tcaatactcg 30

<210> SEQ ID NO 29
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 29

gggctcgagg agcgacctca tgctatact g 31

<210> SEQ ID NO 30
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 30

ttagaaaaac tcatcgagca tc 22

<210> SEQ ID NO 31
<211> LENGTH: 55
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer, OLS 5' FWD

<400> SEQUENCE: 31

gcatagcaat ctaatctaag tttaaatga atcatttgag agcagaaggg cctgc 55

<210> SEQ ID NO 32
<211> LENGTH: 56
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

-continued

<223> OTHER INFORMATION: Primer, CB 5' FWD

<400> SEQUENCE: 32

caccagaact tagtttcgac ggataaaatg gaaaccgggt tgcctcgggt ttgcac 56

<210> SEQ ID NO 33

<211> LENGTH: 58

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer, ALL REV

<400> SEQUENCE: 33

cataactaat tacatgattt aaccttaaac atcagattca atagagccgc ctccactg 58

<210> SEQ ID NO 34

<211> LENGTH: 8933

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Codon Optimized

<400> SEQUENCE: 34

ggttaaatca tgtaattagt tatgtcacgc ttacattcac gccctcccc cacatccgct 60

ggttaaatca tgtaattagt tatgtcacgc ttacattcac gccctcccc cacatccgct 120

ctaaccgaaa aggaaggagt tagacaacct gaagtctagg tccctattta tttttttata 180

gttatgttag tattaagaac gttatttata tttcaaattt ttcttttttt tctgtacaga 240

cgcgtgtacg catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg 300

aaggctttaa tttgcggccc ctcaactgca cgcaaatag gataattata ctctatttct 360

caacaagtaa ttggttggtt ggccgagcgg tctaaggcgc ctgattcaag aaatatcttg 420

accgcagtta actgtgggaa tactcaggta tcgtaagatg caagagttcg aatctcttag 480

caaccattat ttttttcctc aacataacga gaacacacag gggcgctatc gcacagaatc 540

aaattcgatg actggaaatt ttttgtaaat ttcagaggtc gccctgacgca tatacctttt 600

tcaactgaaa aattgggaga aaaaggaaag gtgagagcgc cggaaaccggc tttcatata 660

gaatagagaa gcgttcatga ctaaaatgctt gcatcacaat acttgaagtt gacaatatta 720

tttaaggacc tattgttttt tccaataggt ggtagcaat cgtcttactt tctaactttt 780

cttacctttt acatttcage aatatatata tatatattc aaggatatac cattctaatg 840

tctgccctta agaagatcgt cgttttgcca ggtgaccacg ttggtcaaga aatcacagcc 900

gaagccatta aggttcttaa agctatttct gatgttcggt ccaatgtcaa gttcgatttc 960

gaaaatcatt taattgggag tgctgctatc gatgctacag gtgttccact tccagatgag 1020

gcgctggaag cctccaagaa ggctgatgcc gttttgtag gtgctgtggg tggtcctaaa 1080

tggggtaccg gtagtgtagt acctgaacaa ggtttactaa aaatccgtaa agaacttcaa 1140

ttgtacgcca acttaagacc atgtaacttt gcatccgact ctctttttaga cttatctcca 1200

atcaagccac aatttgctaa aggtactgac ttcggtgttg tcagagaatt agtgggaggt 1260

atttactttg gtaagagaaa ggaagatgat ggtgatggtg tcgcttggga tagtgaacaa 1320

tacaccgttc cagaagtgca aagaatcaca agaatggcgg ctttcatggc cctacaacat 1380

gagccaccat tgccattttg gtccttgat aaagctaagt ttttgccctc ttaagatta 1440

-continued

tggagaaaa ctgtggagga aaccatcaag aacgaattcc ctacattgaa ggttcaacat	1500
caattgattg attctgccgc catgatccta gttaagaacc caaccacct aatggtatt	1560
ataatcacca gcaacatggt tgggtgatatc atctccgatg aagcctcogt tatcccaggt	1620
tccttgggtt tgttgccatc tgcgtccttg gectctttgc cagacaagaa cacgcattt	1680
ggtttgtacg aacctgcca cggttctgct ccagatttgc caaagaataa ggtcaaccct	1740
atcgccacta tcttgtctgc tgcaatgatg ttgaaattgt cattgaaact gectgaagaa	1800
ggtaaggcca ttgaagatgc agttaaaaag gttttggatg caggcatcag aactggtgat	1860
ttaggtggtt ccaacagtac caccgaagtc ggtgatgctg tcgccgaaga agttaagaaa	1920
atccttgctt aaaaagattc tcttttttta tgatatttgt acataaaact tataaatgaa	1980
atcataata gaaacgacac gaaattacaa aatggaatat gttcataggg taacgctatg	2040
atccaatc atcaagaaatg atagcattga aggatgagac taatccaatt gaggagtggc	2100
agcatataga acagctaaaag ggtagtgtctg aaggaagcat acgatacccc gcatggaatg	2160
ggataatc acaggaggta ctagaactacc tttcatccta cataaataga cgcatataag	2220
tacgcattta agcataaaca cgcactatgc cgttcttctc atgtatatat atatacaggc	2280
aacacgcaga tataggtgcg acgtgaacag tgagctgtat gtgcgcagct cgcggttgc	2340
ttcgggaagc gctcgttttc ggaaaacgctt tgaagttcct attccgaagt tcctattctc	2400
tagaaagtat aggaacttca gagcgtttt gaaaaccaa agcgcctctga agtcgcactt	2460
tcaaaaaac aaaaacgcac cggactgtaa cgagctacta aaatattgcg aataccgctt	2520
ccacaaacat tgctcaaaag tatctctttg ctatatatct ctgtgctata tcctatata	2580
acctacccat ccacctttcg ctccctgaaac ttgcatctaa actcgacctc tacattttt	2640
atgtttatct ctagtattac tcttttagaca aaaaaattgt agtaagaact attcataag	2700
tgaatcgaaa acaatacgaa aatgtaaaaa tttcctatac gtagtatata gagacaaaat	2760
agaagaaaac gttcataatt ttctgaccaa tgaagaatca tcaacgctat cactttctgt	2820
tcacaaaagta tgcgcaatcc acatcgggat agaataaat cgggggatgcc tttatcttga	2880
aaaaatgcac ccgcagcttc gctagtatc agtaaacgcg ggaagtggag tcaggctttt	2940
tttatggaag agaaaataga caccaaagta gccttcttct aaccttaacg gacctacagt	3000
gcaaaaagt atcaagagac tgcattatag agcgcacaaa ggagaaaaaa agtaaatctaa	3060
gatgctttgt tagaaaaata gcgctctcgg gatgcatttt tgtagaacaa aaaagaagta	3120
tagattcttt gttggtaaaa tagcgcctctc gcggtgcatt tctggtctgt aaaaatgcag	3180
ctcagattct ttgtttgaaa aattagcgtc ctgcggttgc atttttgttt tacaaaaatg	3240
aagcacagat tcttcggttg taaaatagcg ctttcgctgt gcatttctgt tctgtaaaaa	3300
tgagctcag attctttggt tgaaaaatta gcgctctcgc gttgcatttt tgttctacaa	3360
aatgaagcac agatgcttcg ttcagggtggc acttttcggg gaaatgtgcg cggaaaccct	3420
atgtgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga	3480
tattggtcag aattggttaa ttggtgtaa cactgacccc tattgtttta ttttctaaa	3540
tacattcaaa tatgatccg ctcatgagac aataaccctg ataaatgctt caataatatt	3600
gaaaaaggaa gaatatgagc catattcaac gggaaacgtc gaggccgcga ttaaatcca	3660
acatggatgc tgatttatat gggataaat gggctcgcga taatgtcggg caatcagggtg	3720

-continued

cgacaatcta	tcgcttgat	gggaagcccg	atgcgccaga	gttgtttctg	aaacatggca	3780
aaggtagcgt	tgccaatgat	gttacagatg	agatggtcag	actaaactgg	ctgacggaat	3840
ttatgccact	tccgaccatc	aagcatttta	tccgtactcc	tgatgatgca	tggttactca	3900
ccactgcgat	ccccgaaaa	acagcgttcc	aggtattaga	agaatcct	gattcaggtg	3960
aaaaattgt	tgatgcgctg	gcagtgttcc	tgcgccggtt	gactcgcatt	cctgtttcta	4020
attgtccttt	taacagcgat	cgcgtatttc	gcctcgctca	ggcgcaatca	cgaatgaata	4080
acggtttggg	tgatgcgagt	gattttgatg	acgagcgtaa	tggctggcct	ggtgaacaag	4140
tctggaaga	aatgcataaa	cttttgccat	tctcaccgga	ttcagtcgtc	actcatggtg	4200
atcttcaact	tgataacctt	atctttgacg	aggggaaatt	aataggttgt	attgatgttg	4260
gacgagtcgg	aatcgcagac	cgataaccag	atcttgccat	cctatggaac	tgctcggtg	4320
agttttctcc	ttcattacag	aaacggcttt	ttcaaaaata	tggtattgat	aatcctgata	4380
tgaaaaaatt	gcaatttcat	ttgatgctcg	atgagttttt	ctaactcatg	acaaaaatcc	4440
cttaacgtga	gttacgcgcg	cgctgttcca	ctgagcgtca	gaccccgtag	aaaagatcaa	4500
aggatcttct	tgagatcctt	ttttctcgcg	cgtaatctgc	tgcttgcaaa	caaaaaaacc	4560
accgctacca	gcggtggttt	gtttgocgga	tcaagagcta	ccaactcttt	ttccgaaggt	4620
aaactggcttc	agcagagcgc	agataccaaa	tactgttctt	ctagtgtagc	cgtagttagc	4680
ccaccacttc	aagaactctg	tagcaccgcc	tacatacctc	gctctgctaa	tcctgttacc	4740
agtggctgct	gccagtggcg	ataagtcgtg	tcttaccggg	ttggactcaa	gacgatagtt	4800
accggataag	gcgagcgggt	cgggctgaac	ggggggttcg	tgcacacagc	ccagcttggg	4860
gcgaacgacc	tacaccgaac	tgagatacct	acagcgtgag	ctatgagaaa	gcccacgct	4920
tcccgaaggg	agaaaaggcg	acaggtatcc	ggtaagcggc	agggtcggaa	caggagagcg	4980
cacgagggag	cttccagggg	gaaacgcctg	gtatctttat	agtccctgctg	ggtttcgcca	5040
cctctgactt	gagcctcgat	ttttgtgatg	ctcgtcaggg	ggcgggagcc	tatggaaaaa	5100
cgccagcaac	gcggcctttt	tacggttcct	ggccttttgc	tggccttttg	ctcacatggt	5160
ctttcctgcg	ttatcccctg	attctgtgga	taaccgtatt	accgcctttg	agtgagctga	5220
taccgctcgg	ggtcgtgcag	gtagtttato	attatcaata	ctcgccattt	caaagaatac	5280
gtaaaaaatt	aatagtagtg	atcttcttaa	ctttatttag	tcaaaaaatt	agccttttaa	5340
ttctgctgta	accctgacat	gccccaaaata	ggggcggggt	tacacagaat	atataacatc	5400
gtaggtgtct	gggtgaacag	tttatctctg	gcataccta	aatataatgg	agcccgcctt	5460
ttaagctggc	atccagaaaa	aaaaagaatc	ccagcaccaa	aatattgttt	tcttcaccaa	5520
ccatcagttc	ataggtccat	tctcttagcg	caactacaga	gaacaggggc	acaaacaggc	5580
aaaaaacggg	cacaacctca	atggagtgat	gcaaccagcc	tggagtaaat	gatgacacaa	5640
ggcaattgac	ccacgcatgt	atctatctca	ttttcttaca	ccttctatta	ccttctgctc	5700
tctctgattt	ggaaaaagct	gaaaaaaaag	gttgaaacca	gttccctgaa	attattcccc	5760
tacttgacta	ataagtatat	aaagacggta	ggtattgatt	gtaattctgt	aaatctattt	5820
cttaaacctc	ttaaattcta	cttttatagt	tagtcttttt	tttagtttta	aaacaccaga	5880
acttagtttc	gacggataaa	atggaaaccg	gtttgtcctc	ggtttgcaact	ttctccttcc	5940
aaacaaacta	tcatacactc	ctgaacccgc	acaataacaa	tccccaaact	tcctgctgt	6000

-continued

gttataggca cccaaagaca ccaatcaaat actcctacaa taactttcca tctaagcatt	6060
gtagcacaaa aagtttccat ttgcaaaaata agtgttccga atctctgtcc atcgccaaaa	6120
attocattag ggctgccact actaatcaaa ctgaaccacc agagtctgat aatcattctg	6180
tcgcccacaaa gattctgaat tttgggaagg cttgttgaa gttacaaaga ccatatacaa	6240
ttattgcctt tacctcttgt gcctgtggtt tatttggtaa ggaactgttg cataatacaa	6300
atttaatatc ttggtcattg atggaaacgt tcaaagcatt tttttctta gtcgctatcc	6360
tttgattgc tctttcacc accactatca accagattta cgacttacat attgacagaa	6420
ttaacaagcc agatttgcca ctggcttcgg gcgagatttc cgtcaatact gcctggatca	6480
tggaacttc tattattggt gccttgttg gattgataat caccataaaa atggaaacta	6540
agggtggtcc attgtatatt ttcggttact gttttggtat cttcgggggc atcgtctact	6600
ctgttctccc attcagatgg aaacaaaac cttccacagc attccttttg aacttctctg	6660
cgacattat aaccaacttt actttttatt atgcctccag agccgcccctg gggctgcct	6720
ttgaattacg cccctccttt acatttttac tggccttcat ggagaccaag tccatggaga	6780
ctggttctgc tctcgcgttg atcaaagatg cttccgatgt ggaaggtgac accaaaattg	6840
gtatatccac tttggccagc aagtatggtt ccaggaattt gacctattt tgttctggta	6900
tcgtgctgct gtcttatggt gcagccatct tggctggcat catttgcca caggcttca	6960
attcaaatgt tatggagacg ctgctctcgc atgctatatt ggcattttg ttgattctac	7020
agacaagaga ttttgcctta accaattatg acccagaagc tggtagaaga ttttacgaat	7080
ttatgaaac atgaaaatta tactatgctg aatatttagt gtacgttttc attggggcg	7140
gctccagcgc cggcggcggc tcttctcgg gcggttggtc tcaccacaa tttgagaaag	7200
gtgggtcgtc tggcggcggc agcggggcg ggtccggcg ggggagcggc ggtatgaaat	7260
gttogacctt ctcttttttg tttgtctgta aaataatatt tttttcttc agctttaaca	7320
ttcaaacag cattgcaaat ccaagagaaa atttctttaa atgctttca caatataccc	7380
ccaataatgc tactaacttg aagctagttt atactcaaaa caacccttg tacatgtccg	7440
tgctcaactc caccattcac aacctaatg tcaactcaga cactaccca aaaccattag	7500
ttattgtgac accttctcac gtttcacata tccaaggtac tattttatgc tccaagaagg	7560
tcggcctgca aattagaact agatctggag gtcattgattc agaaggaatg tcttacatct	7620
ctcaagtcc atttgtgatt gtcgatttaa gaaatagag gagcattaag atcgatgttc	7680
actcccaaac ggcattgggt gaagcgggtg ccaccttggg cgaagtttac tactgggtca	7740
acgagaagaa tgaaaactta tcaactagcc caggttattg tccaactgtt tgtgctggtg	7800
gccatttcgg aggcggcggc tacggtctc taatgagaaa ctacggctta gctgctgaca	7860
atatcatcga cgtcacttg gttaacgttc atggtaaagt tttagataga aaatctatgg	7920
gtgaggatct tttctgggct ttgagaggtg gcggcgcaga atcatttggc attatcgttg	7980
cttgaagat cagattgggt gctgtccca agtctacaat gttttctgtg aagaaaatta	8040
tggaatcca tgaattggtc aaactggtga ataaatggca aaacatagct tacaagtacg	8100
ataaagactt gctgttaatg acacatttta ttaccaggaa catcactgat aaccaaggca	8160
agaacaagac tgcaattcat acttattttt cctccgtttt tttgggtggt gtcgactccc	8220
tcgtggatct gatgaataaa tcattccctg aactaggtat taaaaaac gattgtagac	8280

-continued

```

aattgagttg gattgatacc atcatattct acagtgggtg tgттаattat gatactgaca 8340
acttcaacaa agaaatactg ctggaccgtt cgcgcggcca gaatgggtgct tttaaaatca 8400
agttggatta tgtgaaaaag cctattccag aatccgtatt tgttcaata ttggaaaagc 8460
tgtatgaaga agacattggt gcaggcatgt acgctcttta tccttatggc ggcataatgg 8520
atgaaatttc tgaaagtgc attcctttcc cacatagggc cgggatcctg tacgagttat 8580
ggtacatttg ttcattggaa aagcaagaag ataatgaaaa acatttaaat tggataagaa 8640
atatttataa ttttatgact ccatacgtct ccaaaaacc acgcctggca tatttgaatt 8700
acagagacct ggatattgac atcaatgatc ctaaaaacc aaataattac actcaggcaa 8760
gaatatgggg tgaaaaatat ttcggcaaaa atttgatag gctggcaag gttaaaacac 8820
tggttgatcc aaacaatttc tttagaaacg aacaatctat cccacctctg cctagacata 8880
gacacggcgg tggaagcagt ggaggcggct ctattgaatc tgatgttaa tga 8933

```

<210> SEQ ID NO 35

<211> LENGTH: 6677

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Condon Optimized

<400> SEQUENCE: 35

```

ggttaaatca tgtaattagt tatgtcacgc ttacattcac gccctcccc cacatccgct 60
ctaaccgaaa aggaaggagt tagacaacct gaagtctagg tcctattta ttttttata 120
gttatgtag tattaagaac gttatttata tttcaaattt ttctttttt tctgtacaga 180
cgcgtgtacg catgtaacat tatactgaaa acctgcttg agaaggttt gggacgctcg 240
aaggcttaa tttgcggccc ctcacctgca cgaaaaagc tttcaattc aattcatcat 300
tttttttta ttctttttt tgatttcggt ttctttgaaa ttttttgat tcggtaatct 360
ccgaacagaa ggaagaacga aggaaggagc acagacttag attggtatat atacgcatat 420
gtagtgtaga agaaacatga aattgccag tattcttaac ccaactgcac agaacaaaaa 480
ccagcaggaa acgaagataa atcatgtcga aagctacata taaggaaact gctgctactc 540
atcctagtcc tgttgcgtcc aagctattta atatcatgca cgaaaagcaa acaaaactgt 600
gtgcttcatt ggatgttcgt accaccaagg aattactgga gttagttgaa gcattaggtc 660
ccaaaatttg tttactaaaa acacatgtgg atatcttgac tgatttttcc atggagggca 720
cagttaagcc gctaaaggca ttatccgcca agtacaattt tttactcttc gaagatagaa 780
aatttgctga cattgtaaat acagtcaaat tgcagtactc tgcgggtgta tacagaatag 840
cagaatgggc agacattacg aatgcacacg gtgtgggtggg cccaggtatt gttagcggtt 900
tgaagcaggc ggcagaagaa gtaacaaaag aacctagagg ccttttgatg ttagcagaat 960
tgtcatgcaa gggctcccta tctactggag aatatactaa gggtactggt gacattgcga 1020
aaagcgacaa agattttggt atcggcttta ttgctcaaag agacatgggt ggaagagatg 1080
aaggttacga ttggttgatt atgacaccgc gtgtgggttt agatgacaag ggagatgcat 1140
tgggtcaaca gtatagaacc gtggatgatg ttgtctctac aggatctgac attattattg 1200
ttggaagagg actatttgca aaggaagggt atgctaagggt agagggtgaa cgttacagaa 1260
aagcaggctg ggaagcatat ttgagaagat gcggccagca aaactaaaaa actgtattat 1320

```

-continued

aagtaaatgc atgtatacta aactcacaaa ttagagcttc aatttaatta tatcagttat	1380
taccacgct atgatccaat atcaaaggaa atgatagcat tgaaggatga gactaatcca	1440
attgaggagt ggcagcatat agaacagcta aagggtagtg ctgaaggaag catacgatac	1500
cccgcagtgga atgggataat atcacaggag gtactagact acctttcatc ctacataaat	1560
agacgcatat aagtacgcat ttaagcataa acacgcacta tgccgttctt ctcagtata	1620
tatatataca ggcaacacgc agatataggt gcgacgtgaa cagttagctg tatgtgcgca	1680
gctcgcgttg cattttcgga agcgcctggt ttcggaaaacg ctttgaagtt cctattccga	1740
agttcctatt ctctagaag tataggaact tcagagcgtc tttgaaaacc aaaagcgtc	1800
tgaagtcgca ctttcaaaaa accaaaaacg caccggactg taacgagcta ctaaaatatt	1860
gcgaataccg cttccacaaa cattgctcaa aagtatctct ttgctatata tctctgtgct	1920
atatccctat ataacctacc catccacctt tcgctccttg aacttgcac taaactcgac	1980
ctctacattt tttatgttta tctctagat tactctttag acaaaaaaat tgtagtaaga	2040
actattcata gagtgaatcg aaaacaatac gaaaatgtaa acatttccta tacgtagtat	2100
atagagacaa aatagaagaa accgttcata attttctgac caatgaagaa tcatcaacgc	2160
tatcactttc tgttcacaaa gtatgcgcaa tccacatcgg tatagaatat aatcggggat	2220
gcctttatct tgaaaaatg cccccgcgc ttcgctagta atcagtaaac gcgggaagtg	2280
gagtcaggct tttttatgg aagagaaaat agacaccaa gtagccttct tctaacctta	2340
acggacctac agtgcaaaaa gttatcaaga gactgcatta tagagcgac aaaggagaaa	2400
aaaagtaatc taagatgctt tgtagaaaa atagcgtctc cgggatgcat tttttagaa	2460
caaaaaagaa gtatagatc tttgttgta aatatgcgtc ctcgcgttgc atttctgttc	2520
tgtaaaaatg cagctcagat tctttgttg aaaaattagc gctctcgcgt tgcattttg	2580
ttttacaaa atgaagcaca gattctctgt tggtaaaata gcgctttcgc gttgcatttc	2640
tgttctgtaa aaatgcagct cagattcttt gttgaaaaa tttagcgtct cgcgttgc	2700
ttttgttcta caaaatgaag cacagatgct tcggtcaggt ggcacttttc ggggaaatgt	2760
gcgccaacc cctatttgtt tatttttcta aatacattca aatatgtatc cgctcatgag	2820
acaataacc tgatattggt cagaattggt taattggtg taacactgac ccctatttgt	2880
ttatttttct aaatacttc aaatatgtat ccgctcatga gacaataacc ctgataaatg	2940
cttcaataat attgaaaaag gaagaatag agtattcaac atttccgtgt cgccttatt	3000
ccctttttg cggcattttg ccttctggt tttgctcacc cagaaacgct ggtgaaagta	3060
aaagatgctg aagatcagtt ggggtgcaga gtgggttaca tcgaaactgga tctcaacagc	3120
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa	3180
gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc	3240
cgcatacact atttctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt	3300
acggatggca tgacagtaag agaattatgc agtgctgcca taacctgag tgataaacct	3360
gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc ttttttgac	3420
aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata	3480
caaacgcgc agcgtgacac cacgatgcct gtagegatgg caacaacgtt gcgcaacta	3540
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg	3600

-continued

gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctgggt tattgctgat 3660
aaatccggag cgggtgagcg tggttctcgc ggtatcatcg cagcgctggg gccagatggt 3720
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga 3780
aatagacaga tcgctgagat aggtgctca ctgattaagc attggtaact catgacccaaa 3840
atccctaac gtgagttacg cgcgcgctcg tccactgagc gtcagacccc gtagaaaaga 3900
tcaaaggatc ttcttgagat ccttttttct tgcgcgtaat ctgctgcttg caaacaaaaa 3960
aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga 4020
aggtaactgg cttcagcaga gcgcagatc caaatactgt tcttctagtg tagccctagt 4080
tagcccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt 4140
taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttgac tcaagacgat 4200
agttaccgga taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct 4260
tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcga 4320
cgcttcccga agggagaaa ggcgacaggt atccggtaag cggcagggtc ggaacaggag 4380
agcgcacgag ggagcttcca ggggaaaacg cctggtatct ttatagtcct gtcgggttct 4440
gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcg agcctatgga 4500
aaaaagccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca 4560
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag 4620
ctgataccgc tggggctcgt gcaggatatag cttcaaaatg tttctactcc ttttttactc 4680
ttccagattt tctcggactc cgcgcacatc cgtaccactt caaaacaccc aagcacagca 4740
tactaaattt cccctcttcc ttcctctagg gtgtcgtaa ttaccctac taaaggtttg 4800
gaaaagaaaa aagtgaccgc ctcgtttctt tttctctcgc gaaaaaggca ataaaaattt 4860
ttatcacggt tctttttctt gaaaattttt ttttttgatt tttttctctt tcgatgacct 4920
cccattgata ttaagttaa taaacggact tcaatttctc aagtttcagt ttcatttttc 4980
ttgtctattt acaacttttt ttacttcttg ctcatagaa agaaagcata gcaatctaatt 5040
ctaagttaa aatgaatcat ttgagagcag aagggcctgc tcccgctctg gctattggta 5100
ccgccaatcc agaaaatcct ctgctgcagg acgaattccc agattactat ttaggggtca 5160
ccaaatctga acatagaca caattgaaag agaaattcag aaagatttgt gacaagtcca 5220
tgattaggaa aagaaattgt tttttgaaat aagaacactt gaagcaaat cctcgctctg 5280
tggagcatga aatgcaact ttggatgcta gacaagacat gttgggtggg gaagttccaa 5340
agctggggaa ggatgctctg gcccaaggcca ttaaagaatg gggccaacca aaatccaaaa 5400
ttaccacact gattttcacc tccgcctcca ccaactgatat gccaggtgca gactatcatt 5460
gtgctaaatt gttgggtttg tccccctccg tgaagagagt tatgatgat caattagggt 5520
gttatggcgg cggcaccggt ctgagaattg ccaaagacat tgctgaaaac aataaagggt 5580
cgcgcgtttt ggctgtttgt tgtgatatta tggcatggtt atttagaggt ccaagtgaaa 5640
gtgacttggg attgctagtg ggcacggcca tatttgggtg tgggtccgct gctgtgatcg 5700
ttgggtctga gcctgatgaa tctgtcggtg aaagaccaat ttttgaactg gtttccactg 5760
gtcaaacctt tttgcaaat tcagaaggtg ctattggcgg ccatatcaga gaagctgggt 5820
taatctttga tttgcacaag gatgtcccaa tgtaatttc caataatatt gaaaaatggt 5880

-continued

```

tgatcgaagc atttaccccc atcggtatct ctgattggaa ttccatcttc tggattacac 5940
atcctggcgg taaagctatc ttagataaag ttgaggagaa gttgcattta aagtctgaca 6000
aatttgttga ttcaagacat gtcctgtctg agcacggtaa tatgtcttcc tegaccgtct 6060
tgtttgtcat ggatgagttg aggaagaggt ccoctggaaga aggcaagagc accaccggtg 6120
acggttttga gtggggggtc ctctttggat ttgggccagg cctgaccgta gaaagggttg 6180
ttgtccgctc ggtgccaatc aaatatggtg gggggtccag cgccggtggc gggagctccg 6240
cgggcggttg gtctcaccca caatttgaag aggggtggcag cagcggcggc ggctctggcg 6300
gaggctccgg cgggggctcg gggggtatgg ctgtcaagca tctgatcgtg ctgaagtcca 6360
aagatgaaat tactgaagcc caaaaggagg aatTTTTTcaa gacatatggt aatttggtta 6420
acatcattcc agcaatgaaa gatgtttatt ggggtaagga cgttactcaa aaaaaaagg 6480
aagaggggta cactcatatt gttgaagtca ctttcgaatc cgtcgaaaaca attcaagatt 6540
atattattca tccagctcat gttgggttg gcgatgtgta cagatcattt tgggaaaaat 6600
tattgatttt tgactacaca ccaagaaaag gcggtggaag cagtggaggc ggctctattg 6660
aatctgatgt ttaatatg 6677

```

<210> SEQ ID NO 36

<211> LENGTH: 3372

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Codon Optimized

<400> SEQUENCE: 36

```

atggagaaga ctcaagaaac agtccaaaga attcttctag aaccctataa atacttactt 60
cagttaccag gtaaacaagt gagaacccaa ctttcacagg catttaatca ttggctgaaa 120
gttccagagg acaagctaca gattattatt gaagtgcagc aaatggtgca taatgccagt 180
ttactcatcg atgatattga agacaactca aaactccgac gtggctttcc agtggcccac 240
agcatctatg gaatcccatc tgtcatcaat tctgccaaat acgtgtatct ccttggtctg 300
gagaaagtct taacccttga tcaccagat gcagtgaaagc tttttaccgg ccagcttttg 360
gaactccatc agggacaagg cctagatatt tactggaggg ataattacac ttgtcccact 420
gaagaagaat ataaagctat ggtgctgcag aaaacagggtg gactgttttg attagcagta 480
ggtctcatgc agttgttctc tgattacaaa gaagatttaa aaccgctact taatacactt 540
gggctctttt tccaaattag ggatgattat gctaatctac actccaaaga atatagttaa 600
aacaaaaagt tttgtgaaga tctgacagag ggaagttct catttcctac tattcatgct 660
atgtgtcaa ggctgaaag caccaggtg cagaatatct tgcgccagag aacagaaaaac 720
atagatataa aaaaactactg tgtacattat cttgaggatg taggttcttt tgaatacact 780
cgtaataacc ttaaagagct tgaagctaaa gcctataaac agattgatgc acgtggtggg 840
aacctgagc tagtagcctt agtaaaacac ttaagtaaga tgttcaaaga agaaaatgaa 900
ggcggttctg gcagcggaga gggcagagga agtcttctaa catgcgggtga cgtggaggag 960
aatccccgcc ctaggtctgg cagcggagag ggcagaggaa gtcttctaac atgcgggtgac 1020
gtggaggaga atccccggcc taggcacaaa aagaaagtcc cagacaattg ttgtagacgt 1080
gaacctatgc tggtcagaaa taaccagaaa tgtgattcag tagaggaaga gacagggata 1140

```

-continued

aaccgagaaa	gaaaagtga	ggtataaaa	cccttagtgg	ctgaaacaga	taccccaaac	1200
agagctacat	ttgtggttg	taactcctcc	ttactcgata	cttcatcagt	actggtgaca	1260
caggaacctg	aaattgaact	tcccagggaa	cctcggccta	atgaagaatg	tctacagata	1320
cttggaatg	cagagaaagg	tgcaaaatc	cttagtgatg	ctgagatcat	ccagttagtc	1380
aatgctaagc	atatcccagc	ctacaagttg	gaaactctga	tgaaactca	tgagcgtggt	1440
gtatctatc	gccgacagtt	actttccaag	aagctttcag	aaccttcttc	tctccagtac	1500
ctacctaca	gggattataa	ttactccttg	gtgatgggag	cttgttgatg	gaatgttatt	1560
ggatatatgc	ccatccctgt	tggagtggca	ggaccctttt	gcttagatga	aaaagaattt	1620
caggttccaa	tggcaacaac	agaaggttgt	cttgtggcca	gcaccaatag	aggctgcaga	1680
gcaataggtc	ttggtggagg	tgccagcagc	cgagtccttg	cagatgggat	gactcgtggc	1740
ccagttgtgc	gtcttccacg	tgcttgtgac	tctgcagaag	tgaagcctg	gctcgaaca	1800
tctgaagggt	tcgagtgat	aaaggaggca	tttgacagca	ctagcagatt	tgacagctca	1860
cagaaacttc	atacaagtat	agctggagc	aacctttata	tccgtttcca	gtccaggtca	1920
gggatgcca	tgggatgaa	catgatttca	aagggtacag	agaaagcact	ttcaaaactt	1980
cacgagtatt	tccctgaaat	gcagatttca	gccgttagtg	gtaactattg	tactgacaag	2040
aaacctgctg	ctataaattg	gatagaggga	agaggaaaat	ctgttgtttg	tgaagctgtc	2100
attocagcca	aggttgtcag	agaagtatta	aagactacca	cagaggctat	gattgaggtc	2160
aacattaaca	agaatttagt	gggctctgcc	atggctggga	gcataggagg	ctacaacgcc	2220
catgcagcaa	acattgtcac	cgccatctac	attgctgtg	gacaggatgc	agcacagaat	2280
gttggtagtt	caaatgtgat	tactttaatg	gaagcaagtg	gtcccacaaa	tgaagattta	2340
tatatcagct	gcaccatgcc	atctatagag	ataggaacgg	tgggtggtgg	gaccaacctta	2400
ctacctcagc	aagcctgttt	gcagatgcta	gggtttcaag	gagcatgcaa	agataatcct	2460
ggggaaaatg	cccggcagct	tgcccgaatt	gtgtgtggga	ccgtaatggc	tggggaattg	2520
tcacttatgg	cagcattggc	agcaggacat	cttgtcaaaa	gtcacatgat	tcacaacagg	2580
tcgaagatca	atttacaaga	cctccaagga	gcttgacca	agaagacagc	cggtcagga	2640
ggttcttcag	gactggaagt	gctgtttcag	ggcccgggtg	gatctggcat	gatgcctgaa	2700
ataaacacta	accacctcga	caagcaacag	gttcaactcc	tggcagagat	gtgtatcctt	2760
attgatgaaa	atgacaataa	aattggagct	gagaccaaga	agaattgtca	cctgaacgag	2820
aacattgaga	aaggattatt	gcacagagct	tttagtgtct	tcttattcaa	caccgaaaat	2880
aagcttctgc	tacagcaaa	atcagatgct	aagattacct	ttccaggttg	ttttacgaat	2940
acgtgttgta	gtcatccatt	aagcaatcca	gccgagcttg	aggaaagtga	cgcccttgg	3000
gtgaggcgag	cagcacagag	acggctgaaa	gctgagctag	gaattccctt	ggaagagggt	3060
cctcagaag	aaattaatta	tttaacacga	attcactaca	aagctcagtc	tgatggtatc	3120
tgggtggaac	atgaaattga	ttacattttg	ttggtgagga	agaatgtaac	tttgaatcca	3180
gatcccaatg	agattaaaag	ctattgttat	gtgtcaaagg	aagaactaaa	agaacttctg	3240
aaaaagcag	ccagtggtga	aattaagata	acgccatggt	ttaaaattat	tcagcagact	3300
tttctcttta	aatggtggga	taacttaaat	catttgaatc	agtttgttga	ccatgagaaa	3360
atatacagaa	tg					3372

1) Process of biosynthesis of compounds of interest (Table 1); using a cheap and readily available starting materials in any organism or cell-free expression system (Table 2).

2) Process of genetically engineering micro-organisms that produce compounds of interest in large/industrial-scale quantities.

3) Process of altering genes in organisms to allow better yields for compounds of interest.

4) Process of genetically engineering food plants that are optimized for production of compounds of interest.

5) Using *s. cerevisaie*, *e coli*, *p. pastoris*, *n. crassa*, *s. pombe*, *r. palmatum*, *c. longa*, *o. sativa*, *a. arborescens*, *a. terrus*, *c. sativa*, *s. griseus*, *s. erythera*, *s. coelicolor*, *s. toxytruini*, *s. cellulosum*, *p. floresceus*, *a. orientalis*, *s. lividans*, *a. vinelandii*, *b. subtilis*, *m. tuberculosis*, *m. xanthus*, *a. oryzae*, *a. niger*, *r. palmatum*, *h. serrata*, *r. rubra*, *l. erythrhizon*, *l. acetobutylicum*, *c. utilis*, *a. niger*, *c. glutamicum*, *b. spp.*, and other large-scale biotechnology organisms to biosynthesize a total pathway using cheap and readily available starting materials to make our compounds of interest.

6) Process of claims 1 to 5, wherein the products are those listed in table 1.

7) Process of claims 1 to 6, wherein the starting materials are those listed in table 2.

8) Process of claims 1 to 7, wherein certain steps are performed outside or inside an organism or cell free expression system using various techniques, while other steps may use biosynthesis techniques.

9) Producing compositions (e.g. tinctures, balms, capsules, tablets, concentrates, edibles, topicals) that use claims 1 to 8, whether they are classified as pharmaceuticals or not.

10) Producing compositions that contain combinations of certain compounds for certain therapeutic and non-therapeutic applications, as listed in table 1.

11) To claim 10, compositions that have added cofactors for better efficacy, absorption, etc, for our compounds of interest. These include terpenes, sequiterpenes, vitamins, and other cofactors.

12) cDNA sequences of enzymes and organisms in the pathways and processes listed in FIGS. 1-7 for biosynthesis of compounds of interest using cheap and readily available starting materials, including mRNA's, tRNA's, vectors, and other genetic coding molecules needed for process of claims 1 to 8.

13) A method for the efficient and cheap biosynthesis of CBGA, THCA, CBDA, CBCA, CBDVA, THCVA, CBGVA, and CBCVA utilizing manipulation of any ERG, IDI, gal80p, upc2-1, HMGR, tHMGR, ALD6, DPP1, and ADH2 genes that result in exogenous sterol mutations for the uptake of exogenous sterols, more permeable cell membrane for easy inward flow of exogenous materials while easy outward flow of our compounds of interest, and carbon flux manipulation and blocking of other pathways for many fold higher yields in *S. Cerevisiae* and *P. Pastoris*.

* * * * *