

FLAVONOID AND ANTHOCYANIN BIOPRODUCTION USING MICROORGANISM HOSTS

I. RELATED APPLICATIONS

5 This application claims priority to U.S. Provisional Application No. 63/174,403, filed on April 13, 2021. The content of U.S. Provisional Application No. 63/174,403 is hereby incorporated by reference in its entirety.

II. SEQUENCE LISTING

10 This application contains a sequence listing filed in electronic form as an ASCII.txt file entitled DEBU-009-01-WO-Sequence-Listing.txt, created on March 21, 2022, last modified April 12, 2022, and having a size of 448 KB. The content of the sequence listing is incorporated herein its entirety.

III. FIELD OF THE INVENTION

15 The invention related to materials (including engineered cells and cell lines) and methods involved in the production of flavonoids, anthocyanins and other organic compounds.

IV. BACKGROUND OF THE INVENTION

20 Flavonoids and anthocyanins are natural products produced in plants that find a variety of roles such as antioxidants, ultraviolet (UV) defense mechanisms, and colors. Over the past several years, the health benefits of flavonoids and anthocyanins have been widely demonstrated. These compounds are capable of scavenging radicals and can act as enzyme inhibitors and anti-inflammatory agents. With these recognized health and color benefits, much research has gone into understanding how these compounds are made in nature. Flavonoids and anthocyanins are synthesized from phenylpropanoid starter units and malonyl-Cofactor-A (malonyl-CoA) extender units that then undergo modifications to create many polyphenol compounds such as taxifolin, naringenin, and (+)-catechin. However, in most cases, these compounds are extracted
25 or chemically manufactured.

V. SUMMARY OF THE INVENTION

To move away from agriculture and chemically derived products, we have created engineered cells for the bioproduction of flavonoids and anthocyanins. This approach provides a feasible route for the rapid, safe, economical, and sustainable production of a wide variety of important flavonoids.

Herein, a range of flavonoids and anthocyanins including naringenin, eriodictyol, taxifolin, dihydrokaempferol, (+)-catechin, cyanidin, and cyaninidin-3-glucoside are biomanufactured using a modified microbial host. Herein, the engineered cells include one or more genetic modifications that increase(s) flavonoid and anthocyanin bioproduction by increasing metabolic flux to flavonoid precursors and/or reducing carbon losses resulting from the production of byproducts.

Provided herein are cells engineered for the production of flavonoids, anthocyanins and other organic compounds, where the engineered cells include one or more genetic modifications that increase flavonoid production by increasing metabolic flux to flavonoid precursors and/or reducing carbon losses resulting from the production of byproducts. As nonlimiting examples, a genetic modification can be a modification for over-expressing or under-expressing one or more endogenous genes in the engineered host cell or can be a modification for expressing one or more non-native genes in the engineered host cell. Engineered cells as provided herein can include multiple genetic modifications.

Also provided are cell cultures for producing one or more flavonoids or anthocyanins. The cell cultures include engineered cells as disclosed herein in a culture medium that includes a carbon source that can also be an energy source, such as glycerol, sugar, or an organic acid. In various embodiments, the culture medium can include at least one feed molecule such as but not limited to one or more organic acids or amino acids that can be converted into a flavonoid precursor (such as tyrosine, p-coumaroyl-CoA or malonyl-CoA). Examples of feed molecules include, but are not limited to, acetate, malonate, tyrosine, phenylalanine, pantothenate, coumarate, etc. In some embodiments, the feed molecules may be of reduced or low purity. For example, glycerol as a feed molecule may be crude glycerol, including a biomass comprising glycerol, for example, glycerol obtained as a byproduct of biodiesel processing. Alternatively, or in addition, the culture medium can include a supplemental compound that can be a cofactor or a

precursor of a cofactor used by an enzyme that functions in a flavonoid pathway, such as, for examples, bicarbonate, biotin, thiamine, pantothenate, alpha-ketoglutarate, ascorbate, or 5-aminolevulinic acid.

Further provided are methods for producing flavonoids and anthocyanins that include
5 culturing a cell engineered for the production of flavonoids or anthocyanins as provided herein under conditions in which the cell produces flavonoids or anthocyanins. In some examples, the methods include culturing the engineered cells in a culture medium that includes at least one feed molecule or supplement such as but not limited to: tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate,
10 and 5-aminolevulinic acid. The methods can further include recovering at least one of the flavonoids from culture medium, whole culture, or the cells.

In a first aspect, provided herein are cells engineered to produce one or more flavonoids or anthocyanins, wherein the cells include, in addition to nucleic acid sequences encoding either tyrosine ammonia lyase activity and/or phenylalanine ammonia lyase activity and cinnamate-4-
15 hydroxylase activity, 4-coumarate-CoA ligase activity, chalcone synthase activity, chalcone isomerase activity, flavanone-3-hydroxylase activity, flavonoid 3'-hydroxylase activity or flavonoid 3'5'-hydroxylase activity, cytochrome P450 reductase activity, leucoanthocyanidin reductase activity, and dihydroflavonol-4- reductase activity, one or more genetic modifications for improving production of the flavonoids or anthocyanins. As set forth herein, a cell that is
20 engineered to produce one or more of the flavonoids is engineered to include an exogenous nucleic acid sequence encoding tyrosine ammonia lyase activity that can form 4-coumaric acid using tyrosine as substrate (e.g., tyrosine ammonia lyase TAL, EC: 4.3.1.25) or, alternatively or in addition, an exogenous nucleic acid sequence encoding phenylalanine ammonia lyase activity that can convert phenylalanine to trans-cinnamic acid and an exogenous nucleic acid sequence
25 encoding cinnamate-4-hydroxylase activity that forms 4-coumaric acid from trans-cinnamic acid, an exogenous nucleic acid sequence encoding CoA ligase activity that forms p-coumaroyl-CoA from coumaric acid (e.g., 4-coumarate--CoA ligase, 4CL, EC:6.2.1.12), an exogenous nucleic acid sequence encoding polyketide synthase activity that forms naringenin chalcone using malonyl-CoA and p-coumaroyl-CoA as substrates (e.g., chalcone synthase, CHS, EC:2.3.1.74),
30 an exogenous nucleic acid sequence encoding chalcone isomerase activity that forms naringenin

from naringenin chalcone via its cyclase activity (e.g., chalcone-flavonone isomerase, CHI, EC:5.5.1.6), an exogenous nucleic acid sequence encoding flavanone-3-hydroxylase activity that forms dihydrokaempferol from naringenin or forms taxifolin from eriodictyol (e.g., naringenin 3-dioxygenase, F3H, EC: 1.14.11.9), an exogenous nucleic acid sequence encoding flavonoid 3'-hydroxylase or flavonoid 3'5'-hydroxylase activity coupled with an exogenous nucleic acid sequence encoding cytochrome P450 reductase activity to form taxifolin or dihydromyricetin from dihydrokaempferol or to form eriodictyol or pentahydroxyflavone from naringenin (e.g., flavonoid 3'-monooxygenase, F3'H, EC: 1.14.13.21, EC: 1.14.14.82; cytochrome P450/NADPH-P450 reductase, EC:1.14.14.1; F3'5'H, EC:1.14.14.81), an exogenous nucleic acid sequence encoding dihydroflavonol-4-reductase activity that forms leucocyanidin from taxifolin, leucodelphinidin from dihydromyricetin, or leucopelargonidin from dihydrokaempferol (e.g., dihydroflavonol 4-reductase, EC:1.1.1.1), and an exogenous nucleic acid sequence encoding leucoanthocyanidin reductase activity that forms catechin from leucocyanidin (e.g., leucoanthocyanidin reductase, LAR, EC:1.17.1.3). Optionally, a cell that is engineered to produce anthocyanins is further engineered to include an exogenous nucleic acid sequence encoding anthocyanin synthase activity that forms cyanidin from catechin or leucocyanidin, forms delphinidin from leucodelphinidin, or forms pelargonidin from leucopelargonidin (e.g., anthocyanin synthase, ANS, EC:1.14.20.4) and to include an exogenous nucleic acid sequence encoding glucosyltransferase activity that forms cyanidin-3-O-beta-D-glucoside from cyanidin, delphinidin-3-O-beta-D-glucoside from delphinidin, or pelargonidin-3-O-beta-D-glucoside from pelargonidin (e.g., anthocyanidin 3-O-glucosyltransferase, 3GT, EC:2.4.1.115). The cells provided herein that are engineered to produce flavonoids or anthocyanins are further engineered to increase the production of flavonoids or anthocyanins product, for example by increasing metabolic flux to a flavonoid or anthocyanin pathway, or by decreasing byproduct formation.

A cell engineered to produce a flavonoid is further engineered to increase the supply of precursor malonyl-CoA. One strategy for increasing malonyl-CoA includes increasing acetyl-CoA carboxylase (ACC) activity. In various embodiments, the ACC enzyme, which in most eukaryotes, including fungi, is a large single chain polypeptide, and in plant and bacteria such as *E. coli* is a multi-subunit enzyme, is overexpressed in the host strain. Examples of acetyl-CoA carboxylase that can be expressed in a host cell engineered to produce a flavonoid or anthocyanin include, without limitation, the ACC genes of *Mucor circinelloides*, *Rhodotorula*

toruloides, *Lipomyces starkeyi*, *Ustilago maydis*, and orthologs of these ACCs in other species having at least 50% amino acid identity to these ACCs.

Additional strategies for increasing malonyl-CoA include increasing acetyl-CoA, which is converted to malonyl-CoA by acetyl-CoA carboxylase (ACC). In some embodiments, acetyl-CoA synthase (ACS) that converts acetate and CoA to acetyl-CoA is over-expressed in the host cells. Cultures of engineered host cells that include overexpressed nucleic acid sequence encoding ACS can optionally include acetate in the culture medium. Examples of acetyl-CoA synthase that can be expressed in a host cell engineered to produce a flavonoid or anthocyanin include, without limitation, the ACS gene of *E. coli*, the ACS of *Salmonella typhimurium*, orthologs of these ACSs in other species having at least 50% amino acid identity to these ACSs.

Also considered, in further embodiments, is an engineered host cell that overexpresses a gene encoding pyruvate dehydrogenase (PDH), which converts pyruvate to acetyl-CoA. Further, in *E. coli*, a variant of the Lpd subunit of PDH can be expressed that includes a mutation (E354K) that reduces inhibition of PDH by NADH.

Alternatively, or in addition to strategies for increasing ACC activity and strategies for increasing acetyl-CoA, strategies for increasing malonyl-CoA by mechanisms that do not rely on the activity of an ACC can be employed. In some embodiments, a cell engineered to produce a flavonoid, or an anthocyanin, is further engineered to increase the cell's supply of malonyl-CoA includes an exogenous nucleic acid sequence encoding a malonyl-CoA synthetase that generates malonyl-CoA from malonate. Examples of malonyl-CoA synthetases include the malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a malonyl-CoA synthetase having at least 50% identity to any of these or other naturally occurring malonyl-CoA synthetases. Malonate can optionally be added to the culture medium of a culture that includes a cell engineered to express a malonyl-CoA synthetase. An engineered cell that includes an exogenous gene encoding a malonyl-CoA synthetase can also include an exogenous nucleic acid sequence encoding a malonate transporter, such as a malonate transporter encoded by a *matC* gene, for example, of *Streptomyces coelicolor*, or a malonate transporter encoded by *DctPQM* of *Sinorhizobium medicae*.

In additional embodiments, a cell engineered to produce a flavonoid or an anthocyanin is further engineered to include an exogenous nucleic acid sequence encoding malonate CoA-

transferase that makes malonyl-CoA by direct transfer of the CoA from acetyl-CoA. Examples of malonate CoA-transferase that can be expressed in an engineered cell as provided herein include, without limitation, the alpha subunit (mdcA) of malonate decarboxylase from *Acinetobacter calcoaceticus*, *Geobacillus sp.*, or a transferase having at least 50% identity to any of these or
5 other naturally occurring malonate CoA-transferases.

In some embodiments, a cell engineered to produce flavonoids or anthocyanins is further engineered to increase the supply of coenzyme A (CoA) to increase its availability for producing acetyl-CoA, malonyl-CoA, and/or p-coumaroyl-CoA. Strategies for increasing CoA supply include upregulating endogenous pantothenate kinase (PanK) (EC:2.7.1.33) that produces CoA
10 from pantothenate. Alternatively, or in addition, a host cell can be engineered to include a nucleic acid sequence encoding type III pantothenate kinase that is not feedback inhibited by coenzyme A, such as CoaX gene of *pseudomonas aeruginosa* (EC:2.7.1.33). Cultures of cells engineered for the production of flavonoids or anthocyanins can in some embodiments include a medium that includes pantothenate, a precursor of CoA biosynthesis, and can optionally also
15 include cysteine, used in the CoA biosynthesis.

Additional strategies to increase malonyl-CoA flux to the flavonoid pathway include mutation or downregulation of one or more genes that function in fatty acid biosynthesis. Without limiting the embodiments to any particular mechanism, limiting fatty acid biosynthesis can increase the malonyl-CoA supply available for flavonoid biosynthesis. In some
20 embodiments, the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF) can be disrupted to reduce fatty acid biosynthesis. Another example of a fatty acid biosynthesis gene of a host cell that may be mutated or downregulated is a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD). Other fatty acid biosynthesis genes of the engineered host cell that can be downregulated include a beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB) and acyl carrier protein (*E. coli* acpP).

Additional genetic modifications that may be present in a host cell engineered to produce flavonoids or anthocyanins include downregulation, disruption, or deletion of genes encoding alcohol dehydrogenase, lactate dehydrogenase, pyruvate oxidase, acetyl phosphate transferase and acetate kinase. In an *E. coli* host cell, genes that are downregulated, disrupted, or deleted can include aldehyde-alcohol dehydrogenase (adhE), lactate dehydrogenase (ldhA), pyruvate oxidase
30 (poxB), and enzyme acetate kinase phosphate acetyltransferase (ackA-pta).

Further, a cell engineered for the production of flavonoids or anthocyanins can have one or more genes encoding thioesterases downregulated, disrupted, or deleted to prevent hydrolysis of precursors malonyl-CoA, acetyl-CoA, and/or p-coumaroyl-CoA. For example, in an *E. coli* host one or more of the thioesterase genes *tesA*, *tesB*, *yciA*, and *ybgC* can be downregulated, 5 disrupted, or deleted.

Alternatively, or in addition, genes encoding enzymes of the tricarboxylic acid cycle (TCA), such as succinate dehydrogenase, can be disrupted or downregulated to increase alpha-ketoglutarate supply which serves as a cofactor for one or more of the flavonoid and anthocyanin pathway enzymes. Other TCA enzymes that can be downregulated include citrate synthase that 10 converts acetyl-CoA to citrate.

Also considered, in further embodiments, is an engineered host cell for the production of flavonoids or anthocyanins to upregulate the endogenous biosynthesis of amino acid tyrosine. Tyrosine is one of the precursors for the flavonoid biosynthesis and its conversion to coumaric acid is the first committed step of the pathway. L-tyrosine is one of the three aromatic amino 15 acids derived from the shikimate pathway. The initial step of the shikimate pathway is catalyzed by DAHP synthase isozymes and regulated through feedback-inhibition. Strategies to increase tyrosine production can include, without limitation, transcriptional deregulation, removing feedback inhibition, overexpression of rate-limiting enzymes, and/or deletion of the L-phenylalanine branch of the aromatic acid biosynthetic pathway. For example, in an *E. coli* host 20 the *tyrR* gene can be disrupted, feedback-inhibition-resistant versions of the DAHP synthase (*aroG*) and chorismate mutase (*tyrA*) can be introduced, and/or rate-limiting enzymes, shikimate kinase (*aroK* or *aroL*) and quinate (QUIN)/shikimate dehydrogenase (*ydiB*) can be overexpressed. Further, the Phosphoenolpyruvate synthase (*ppsA*) and transketolase (*tktA*) can be exogenously introduced to enhance tyrosine production.

Also considered, in further embodiments, is an engineered host cell for the production of flavonoids or anthocyanins further engineered to upregulate the endogenous biosynthesis of cofactor heme. Cytochrome P450 (CYPs), one of the exogenous genes in the engineered cells provided herein, contain heme as a cofactor. Improving heme supply can be an effective strategy to increase flavonoid biosynthesis. 5-aminolevulinic acid (ALA) is the first committed precursor 30 to the heme pathway. Strategies to increase heme supply include overexpression of the genes that

synthesize the precursor ALA. In an *E. coli* host, ALA is formed from the 5-carbon skeleton of glutamate (the C5 pathway). The three enzymes involved in ALA biosynthesis are glutamyl-tRNA synthetase (gltX), glutamyl-tRNA reductase (hemA), and glutamate-1-semialdehyde aminotransferase (hemL). In an *E. coli* host, the engineered cells provided herein can be further engineered to express or overexpress hemA or its variants, and/or hemL to increase the heme precursor ALA production. The nonlimiting examples of hemA gene that can be overexpressed include a mutated hemA (inserting two lysine residuals between Thr-2 and Leu-3 at N terminus of hemA gene from *Salmonella typhimurium* (EC:1.1.1.70). Alternatively, or in addition, a heterologous ALAS gene can be introduced to produce ALA via the C4 pathway (ALS is synthesized by the condensation of glycine and succinyl-CoA). Nonlimiting examples of heterologous ALAS that can be expressed in *E. coli* include ALAS of *Bradyrhizobium japonicum* (EC: 2.3.1.37), ALAS of *Rhodobacter capsulatus*, or an ALAS having at least 50% sequence identity to a naturally occurring ALAS. Further, one or more of the downstream genes (e.g., in *E. coli* hemB, hemC, hemD, hemE, hemF, hemG, hemI, or hemH) that catalyze the synthesis of heme from ALA can be overexpressed to drive the flux from ALA to heme production. Cultures of cells engineered for the production of flavonoids or anthocyanins can in some embodiments include a medium that includes succinate and/or glycine, precursors of heme biosynthesis via the C4 pathway.

In another aspect, provided herein are cell cultures that include engineered cells as provided herein in a culture medium, where the culture medium includes a carbon source that is also an energy source for the cells, where the carbon source can be, for example, glycerol, a sugar, or an organic acid, as nonlimiting examples. The culture medium can further include a feed molecule that is used to produce flavonoids or anthocyanins. A feed molecule can be, for example, acetate, malonate, tyrosine, pantothenate, coumarate, biotin, alpha-ketoglutarate, ascorbate, 5-aminolevulinic acid, succinate, or glycine. In some embodiments, the culture comprises a culture medium that includes a carbon source and at least one supplement that is a cofactor of an enzyme or is a precursor of an enzyme cofactor.

In yet another aspect, methods for producing flavonoids and anthocyanins that include incubating a culture of engineered host cell as provided herein to produce flavonoids or

anthocyanins. The methods can further include recovering at least one of the flavonoids from the cells, the culture medium, or the whole culture.

In yet another aspect, the invention provides an engineered host cell that comprises one or more genetic modifications resulting in production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical intermediates, by the engineered host cell. In certain embodiments, the production of flavonoid or anthocyanin from glycerol occurs through enzymatic transformation. In certain embodiments, the production of flavonoid or anthocyanin from a carbon source that can also be an energy source occurs through enzymatic transformation. In certain embodiments, the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, one or more genetic modifications lead to increase metabolic flux to flavonoid precursors or cofactors. In certain embodiments, one or more genetic modifications cause reduction of formation of byproducts. In certain embodiments, one or more genetic modifications are selected from: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof. In certain embodiments, the engineered host cell is cultured in a medium comprising molecules selected from: tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid, wherein one or more of the selected molecules are the chemical intermediates, including molecules in the biosynthesis pathway or cofactors. In certain embodiments, the engineered host cell comprises at least one or more nucleic acid sequences selected from: (i) nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof. In certain embodiments, the engineered host cell comprises at least one or more peptides selected from: (i)

chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof. In certain embodiments, the engineered cell is *E. coli*. In certain embodiments, one or more genetic modifications decreases fatty acid biosynthesis. In certain embodiments, the engineered host cell comprises an exogenous

5 nucleic acid sequence selected from: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii)

10 nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequences encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase

15 forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-

20 coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms

25 dihydrokaempferol from naringenin; and (viii) any combinations thereof. In certain embodiments, the flavonoid is catechin.

In yet another aspect, the invention provides a method of increasing the production of flavonoids or anthocyanins, the method comprising: providing an engineered host cell that comprises one or more genetic modifications resulting in production of flavonoid or anthocyanin

30 from a carbon source that can also be an energy source, through multiple chemical intermediates, by the engineered host cell. In certain embodiments, the production of flavonoid or anthocyanin

from a carbon source that can also be an energy source occurs through enzymatic transformation. In certain embodiments, the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, the engineered host cell is cultured in a medium

5 comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, one or more genetic modifications lead to increase metabolic flux to flavonoid precursors or cofactors. In certain embodiments, one or more genetic modifications cause increased metabolic flux to flavonoid precursors. In certain embodiments,

10 one or more genetic modifications cause reduction in the formation of byproducts. In certain embodiments, one or more genetic modifications are selected from: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing one or more non-

15 native genes in the engineered host cells; and (iv) a combination thereof. In certain embodiments, the engineered host cell is cultured in a medium comprising molecules selected from: tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid, wherein one or more of the selected molecules are the chemical intermediates, including molecules in the biosynthesis pathway or cofactors. In certain embodiments, the engineered host cell comprises at least one or more

20 nucleic acid sequences selected from: (i) a nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) a nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) cinnamate 4-hydroxylase; and (iv) any combination thereof. In certain embodiments, the engineered host cell comprises at least one or more peptides selected from: (i) chalcone

25 isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof. In certain embodiments, the engineered cell is *E. Coli*. In certain embodiments, one or more genetic modifications decreases fatty acid biosynthesis. In certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from: (i) nucleic acid sequences encoding tyrosine ammonia lyase,

30 wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the

encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof. In certain embodiments, the flavonoid is catechin.

In yet another aspect, the invention provides a plurality of engineered host cells, wherein each of the plurality of the engineered host cells comprises one or more genetic modifications resulting in production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical intermediates. In certain embodiments, the production of flavonoid or anthocyanin from a carbon source that can also be an energy source occurs through enzymatic transformation. In certain embodiments, the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, one or more genetic modifications lead to increase metabolic flux to flavonoid precursors or cofactors. In

certain embodiments, one or more genetic modifications lead to increase metabolic flux to flavonoid precursors or cofactors. In certain embodiments, one or more genetic modifications cause reduction of formation of byproducts. In certain embodiments, one or more genetic modifications are selected from: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof. In certain embodiments, at least one of the engineered cells from the plurality of the engineered host cells is cultured in a medium comprising molecules selected from: tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid, wherein one or more of the selected molecules are the chemical intermediates, including molecules in biosynthesis pathway or cofactors. In certain embodiments, at least one of the engineered cells from the plurality of the engineered host cells comprise at least one or more nucleic acid sequences selected from: (i) nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof. In certain embodiments, at least one of the engineered host cell from the plurality of engineered host cells comprise at least one or more peptides selected from: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof. In certain embodiments, at least one the engineered host cell is *E. coli*. In certain embodiments, one or more genetic modifications decreases fatty acid biosynthesis. In certain embodiments, at least one of the engineered host cell from the plurality of the engineered host cells comprises an exogenous nucleic acid sequence selected from: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol

from naringenin; and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof. In certain embodiments, the flavonoid is catechin.

In yet another aspect, the invention provides a method of increasing the production of flavonoids or anthocyanins, the method comprising: providing a plurality of engineered host cells, wherein each of the plurality of the engineered host cell comprises one or more genetic modifications resulting production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical intermediates, by the engineered host cell. In certain embodiments, the production of flavonoid or anthocyanin from a carbon source that can also be an energy source occurs through enzymatic transformation. In certain embodiments, the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof. In certain embodiments, one or more genetic modifications lead to increase metabolic flux to flavonoid precursors or cofactors. In certain embodiments, one or more genetic modifications lead to increase metabolic flux to flavonoid precursors or cofactors. In certain embodiments, one or more genetic modifications cause reduction of formation of byproducts. In certain embodiments,

one or more genetic modifications are selected from: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof. In certain embodiments, at least one of the engineered cells from the plurality of the engineered host cells is cultured in a medium comprising molecules selected from: tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid, wherein one or more of the selected molecules are the chemical intermediates, including molecules in biosynthesis pathway or cofactors. In certain embodiments, at least one of the engineered cells from the plurality of the engineered host cells comprise at least one or more nucleic acid sequences selected from: (i) nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof. In certain embodiments, at least one of the engineered host cell from the plurality of engineered host cells comprise at least one or more peptides selected from: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof. In certain embodiments, at least one the engineered host cell is *E. coli*. In certain embodiments, one or more genetic modifications decreases fatty acid biosynthesis. In certain embodiments, at least one of the engineered host cell from the plurality of the engineered host cells comprises an exogenous nucleic acid sequence selected from: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine

ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof. In certain embodiments, the flavonoid is catechin.

In yet another aspect, the engineered host cell comprises one or more genetic modifications to increase the production and/or availability of malonyl-CoA. In certain embodiments, the production and/or availability of malonyl-CoA is increased by transformation of acetyl-CoA to malonyl-CoA. In certain embodiments, the engineered host cell comprises one or more genetic modifications selected from: (i) expression of acetyl-CoA carboxylase (ACC); and (ii) overexpression of acetyl-CoA carboxylase. In another embodiment, the engineered host cell is an *E. coli*. In certain embodiments, the *E. coli* cell further comprises genes from fungi. In certain embodiments, the acetyl-CoA carboxylase is from: *Mucor circinelloides*, *Rhodotorula toruloides*, *Lipomyces starkeyi*, and *Ustilago maydis*, and orthologs of acetyl-CoA carboxylase having at least 50% amino acid identity to the acetyl-CoA carboxylase of these aforementioned species. In certain embodiments, one or more genetic modification is deletion or attenuation of one or more fatty biosynthetic genes resulting in decrease in fatty acid biosynthesis. In certain embodiments, one or more genetic modification is overexpression of acetyl-CoA synthase (ACS). In certain embodiments, the acetyl-CoA synthase is selected from: acetyl-CoA synthase gene of *E. coli*, acetyl-CoA synthase gene of *Salmonella typhimurium*, and orthologs of acetyl-CoA synthase gene in any other species having at least 50% amino acid identity to the acetyl-CoA synthase gene of *E. coli* and *Salmonella typhimurium*. In certain embodiments, one or more genetic modification is selected from a group consisting of: (i) overexpression a gene encoding

pyruvate dehydrogenase (PDH), wherein the PDH may include E354K mutation; (ii) exogenous nucleic acid sequence encoding a malonyl-CoA synthetase; (iii) upregulation of endogenous pantothenate kinase (PanK), wherein PanK is not feedback inhibited by coenzyme A; (iv) exogenous nucleic acid sequence encoding a malonate transporter; and (v) any combinations thereof. In certain embodiments, the malonyl-CoA synthetase is selected from of malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a malonyl-CoA synthetase having at least 50% identity to any of these or other naturally occurring malonyl-CoA synthetases. In certain embodiments, one or more genetic modifications to decrease fatty acid biosynthesis is selected from: (i) mutation or downregulation of a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD); (ii) modifications to the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF); (iii) downregulation of beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB); (iv) downregulation of acyl carrier protein (*E. coli* acpP); and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises peptides selected from: (i) acetyl-CoA carboxylase (ACC) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 15 or SEQ ID NO: 16; (ii) malonate CoA-transferase having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 19; (iii) acetyl-CoA synthase (ACS) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 16; (iv) malonyl-CoA synthase having an amino acid sequence at least 80% identical SEQ ID NO: 77, SEQ ID NO: 78, or SEQ ID NO: 79; (v) malonate transporter having an amino acid sequence at least 80% identical to SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 83, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 86, or SEQ ID NO: 87; (vi) pantothenate kinase having an amino acid sequence at least 80% identical to SEQ ID NO: 88, SEQ ID NO: 89, or SEQ ID NO: 90; and (vii) any combinations thereof.

In another aspect, the invention provides a method of increasing the production of flavonoids comprising an engineered host cell, wherein the one or more engineered host cells comprise one or more genetic modifications to increase the production and/or availability of malonyl-CoA. In certain embodiments, the production and/or availability of malonyl-CoA is increased by transformation of acetyl-CoA to malonyl-CoA. In certain embodiments, the engineered host cell comprises one or more genetic modifications selected from: (i) expression of acetyl-CoA carboxylase (ACC); and (ii) overexpression of acetyl-CoA carboxylase. In another embodiment, the engineered host cell is an *E. coli*. In certain embodiments, the *E. coli* cell

further comprises genes from fungi. In certain embodiments, the acetyl-CoA carboxylase is from: *Mucor circinelloides*, *Rhodotorula toruloides*, *Lipomyces starkeyi*, and *Ustilago maydis*, and orthologs of acetyl-CoA carboxylase having at least 50% amino acid identity to the acetyl-CoA carboxylase of these aforementioned species. In certain embodiments, one or more genetic modification is deletion or attenuation of one or more fatty biosynthetic genes resulting in decrease in fatty acid biosynthesis. In certain embodiments, one or more genetic modification is overexpression of acetyl-CoA synthase (ACS). In certain embodiments, the acetyl-CoA synthase is selected from: acetyl-CoA synthase gene of *E. coli*, acetyl-CoA synthase gene of *Salmonella typhimurium*, and orthologs of acetyl-CoA synthase gene in any other species having at least 50% amino acid identity to the acetyl-CoA synthase gene of *E. coli* and *Salmonella typhimurium*. In certain embodiments, one or more genetic modification is selected from a group consisting of: (i) overexpression a gene encoding pyruvate dehydrogenase (PDH), wherein the PDH may include E354K mutation; (ii) exogenous nucleic acid sequence encoding a malonyl-CoA synthetase; (iii) upregulation of endogenous pantothenate kinase (PanK), wherein PanK is not feedback inhibited by coenzyme A; (iv) exogenous nucleic acid sequence encoding a malonate transporter; and (v) any combinations thereof. In certain embodiments, the malonyl-CoA synthetase is selected from of malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a malonyl-CoA synthetase having at least 50% identity to any of these or other naturally occurring malonyl-CoA synthetases. In certain embodiments, one or more genetic modifications to decrease fatty acid biosynthesis is selected from: (i) mutation or downregulation of a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD); (ii) modifications to the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF); (iii) downregulation of beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB); (iv) downregulation of acyl carrier protein (*E. coli* acpP); and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises peptides selected from: (i) acetyl-CoA carboxylase (ACC) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 15 or SEQ ID NO: 16; (ii) malonate CoA-transferase having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 19; (iii) acetyl-CoA synthase (ACS) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 16; (iv) malonyl-CoA synthase having an amino acid sequence at least 80% identical SEQ ID NO: 77, SEQ ID NO: 78, or SEQ ID NO: 79; (v) malonate transporter having an amino acid sequence at least 80%

identical to SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 83, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 86, or SEQ ID NO: 87; (vi) pantothenate kinase having an amino acid sequence at least 80% identical to SEQ ID NO: 88, SEQ ID NO: 89, or SEQ ID NO: 90; and (vii) any combinations thereof.

5 In another aspect, the invention provides an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase endogenous biosynthesis of tyrosine. In certain embodiments, one or more genetic modifications comprises upregulation of 3-deoxy-D-arabino-heptulosonate synthase. In certain embodiments, one or more genetic modifications are selected from: (i) upregulation of chorismate mutase; (ii) upregulation of
10 prephenate dehydrogenase; (iii) overexpression of shikimate kinase; (iv) overexpression of shikimate dehydrogenase; and (v) any combinations thereof. In certain embodiments, one or more genetic modifications comprises downregulation of L-phenylalanine biosynthetic pathway. In certain embodiments, one or more genetic modifications comprises expression of exogenous phosphoenolpyruvate synthase (ppsA). In certain embodiments, one or more genetic
15 modifications comprises expression of exogenous transketolase (tktA). In certain embodiments, wherein the one or more genetic modifications comprises disruption of tyrR gene. In certain embodiments, one or more genetic modifications are selected from a group consisting of: (i) expression or overexpression of (D146N) variant of phospho-2-dehydro-3-deoxyheptonate aldolase; (ii) expression or overexpression of variant of 3-dehydroquinate synthase (aroB); (iii)
20 overexpression of transketolase tktA; (iv) deletion of shikimate kinase (aroK); (v) deletion of tyrR; (vi) expression or overexpression of A354V variant of chorismate mutase (tyrA); (vi) and any combination thereof.

 In another aspect, the invention provides a method of increasing endogenous biosynthesis of tyrosine comprising an engineered cell, wherein the engineered host cell comprises one or
25 more genetic modifications to increase endogenous biosynthesis of tyrosine. In certain embodiments, one or more genetic modifications comprises upregulation of 3-deoxy-D-arabino-heptulosonate synthase. In certain embodiments, one or more genetic modifications are selected from: (i) upregulation of chorismate mutase; (ii) upregulation of prephenate dehydrogenase; (iii) overexpression of shikimate kinase; (iv) overexpression of shikimate dehydrogenase; and (v) any
30 combinations thereof. In certain embodiments, one or more genetic modifications comprises

downregulation of L-phenylalanine biosynthetic pathway. In certain embodiments, one or more genetic modifications comprises expression of exogenous phosphoenolpyruvate synthase (ppsA). In certain embodiments, one or more genetic modifications comprises expression of exogenous transketolase (tktA). In certain embodiments, wherein the one or more genetic modifications

5 comprises disruption of tyrR gene. In certain embodiments, one or more genetic modifications are selected from a group consisting of: (i) expression or overexpression of (D146N) variant of phospho-2-dehydro-3-deoxyheptonate aldolase; (ii) expression or overexpression of variant of 3-dehydroquinate synthase (aroB); (iii) overexpression of transketolase tktA; (iv) deletion of shikimate kinase (aroK); (v) deletion of tyrR; (vi) expression or overexpression of A354V

10 variant of chorismate mutase (tyrA); (vi) and any combination thereof.

In another aspect, the invention provides an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G). In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase. In certain

15 embodiments, the anthocyanin synthase is selected from: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof. In certain embodiments, one or more engineered host cells

20 comprises flavonoid-3-glucosyl transferase (3GT). In certain embodiments, flavonoid-3-glucosyl transferase is selected from: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof. In certain embodiments, one or more genetic modifications comprises

25 overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications are selected from a group consisting of: (i) anthocyanin synthase, (ii) flavonoid-3-glucosyl transferase (3GT), and (iii) a combination thereof.

In another aspect, the invention provides a method for increasing the production of flavonoids comprising an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G). In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase. In certain embodiments, the anthocyanin synthase is selected from: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof. In certain embodiments, one or more engineered host cells comprises flavonoid-3-glucosyl transferase (3GT). In certain embodiments, flavonoid-3-glucosyl transferase is selected from: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof. In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications are selected from a group consisting of: (i) anthocyanin synthase, (ii) flavonoid-3-glucosyl transferase (3GT), and (iii) a combination thereof.

In another aspect, the invention provides a method of increasing the transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G), delphinidin or gallocatechin to delphinidin-3-glucoside (De3G), or afzelechin or pelargonidin to pelargonidin-3-glucoside (Pe3G) comprising anthocyanin synthase, wherein the anthocyanin synthase is selected from: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof. In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications are

selected from a group consisting of: (i) anthocyanin synthase, (ii) flavonoid-3-glucosyl transferase (3GT), and (iii) a combination thereof.

In another aspect, the invention provides a method of increasing the transformation of cyanidin to cyanidin-3-glucoside (Cy3G), delphinidin to delphinidin-3-glucoside (De3G), or pelargonidin to pelargonidin-3-glucoside (Pe3G), comprising flavonoid-3-glucosyl transferase (3GT), wherein the flavonoid-3-glucosyl transferase is selected from: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO: 14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof.

In another aspect, the invention provides an engineered host cell comprises one or more genetic modifications to increase the production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictyol (EDL), and/or pentahydroxyflavone (PHF), wherein the engineered host cell comprises cytochrome P450 reductase (CPR) and at least one of flavanone-3-hydroxylase (F3H), flavanone-3'-hydroxylase (F3'H), or flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the precursor for increase in production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictyol (EDL), and/or pentahydroxyflavanone (PHF) is naringenin and/or dihydrokaempferol (DHK). In certain embodiments, the engineered host cell further comprises peptides selected from a group consisting of: (i) flavonoid 3'-hydroxylase (F3'H); (ii) cytochrome P450 reductase (CPR); and (iii) any combination thereof. In certain embodiments, the engineered host cell produces eriodictyol or taxifolin. In certain embodiments, the engineered host cell further comprises flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the engineered host cell produces pentahydroxyflavone or dihydromyricetin. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is truncated to remove the N-terminal leader sequence. In certain embodiments, cytochrome P450 reductase (CPR) is truncated to remove the N-terminal leader sequence. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 7. In certain embodiments, flavanone-3'-hydroxylase (F3'H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 8. In certain

embodiments, cytochrome P450 reductase (CPR) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 9. In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 10, (ii) SEQ ID NO. 56, and (iii) SEQ ID NO. 57. In certain embodiments, the engineered host cell further comprises cytochrome b₅. In certain embodiments, cytochrome b₅ has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 98. In certain embodiments, wherein the flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 7, (ii) SEQ ID NO. 45, (iii) SEQ ID NO. 46, (iv) SEQ ID NO. 47, and (v) SEQ ID NO. 48.

In another aspect, the invention provides method of increasing the production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictoyl (EDL), and/or pentahydroxyflavone (PHF) comprising an engineered host cell, wherein the engineered host cell comprises cytochrome P450 reductase (CPR) and at least one of flavanone-3-hydroxylase (F3H), flavanone-3'-hydroxylase (F3'H), or flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the precursor for increase in production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictoyl (EDL), and/or pentahydroxyflavone (PHF) is naringenin and/or dihydrokaempferol (DHK). In certain embodiments, the engineered host cell further comprises peptides selected from a group consisting of: (i) flavonoid 3'-hydroxylase (F3'H); (ii) cytochrome P450 reductase (CPR); and (iii) any combination thereof. In certain embodiments, the engineered host cell produces eriodictoyl or taxifolin. In certain embodiments, the engineered host cell further comprises flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the engineered host cell produces pentahydroxyflavone or dihydromyricetin. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is truncated to remove the N-terminal leader sequence. In certain embodiments, cytochrome P450 reductase (CPR) is truncated to remove the N-terminal leader sequence. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 7. In certain embodiments, flavanone-3'-hydroxylase (F3'H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 8. In certain

embodiments, cytochrome P450 reductase (CPR) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 9. In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 10, (ii) SEQ ID NO. 56, and (iii) SEQ ID NO. 57. In certain embodiments, the engineered host cell further comprises cytochrome b₅. In certain embodiments, cytochrome b₅ has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 98. In certain embodiments, wherein the flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 7, (ii) SEQ ID NO. 45, (iii) SEQ ID NO. 46, (iv) SEQ ID NO. 47, and (v) SEQ ID NO. 48.

VI. BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S)

FIG. 1 shows the metabolic pathway of flavonoid and anthocyanin bioproduction in engineered cells and methods of preparing anthocyanins described herein.

FIG. 2 shows structures of the flavonoid and anthocyanin molecules that may be produced using engineered cells and methods of preparing anthocyanins described herein.

FIG. 3 shows HPLC spectra showing peaks corresponding to the molecules prepared using engineered cells and methods of preparing anthocyanins described herein.

FIG. 4 shows the pathway of flavonoid and anthocyanin bioproduction in engineered cells and methods of preparing anthocyanins described herein.

VII. DETAILED DESCRIPTION OF THE INVENTION

The present application provides engineered cells for producing one or more flavonoids, cultures that include the engineered cells, and methods of producing one or more flavonoids, or at least one anthocyanin. The terms “flavonoid”, “flavonoid product”, or “flavonoid compound” are used herein to refer to a member of a diverse group of phytonutrients found in almost all fruits and vegetables. As used herein, the terms “flavonoid”, “flavonoid product”, or “flavonoid compound” are used interchangeably to refer a molecule containing the general structure of a 15-carbon skeleton, which consists of two phenyl rings (A and B) and a heterocyclic ring. Flavonoids may include, but are not limited to, isoflavone type (e.g., genistein), flavone type (e.g., apigenin), flavonol type (e.g., kaempferol), flavanone type (e.g., naringenin), chalcone type

(e.g., phloretin), anthocyanidin type (e.g., cyanidin), catechins, flavanones, and flavanonols. Flavonoid compounds of interest include, without limitation, naringenin, naringenin chalcone, eriodictyol, taxifolin, dihydrokaempferol, dihydroquercetin, dihydromyricetin, leucocyanidin, leucopelargonidin, leucodelphinidin, pentahydroxyflavone, cyanidin, catechin, delphinidin, pelargonidin, and kaempferol. Anthocyanins are in the forms of anthocyanidin glycosides and acylated anthocyanins. Anthocyanin compounds of interest include, without limitation, cyanidin glycoside, delphinidin glycoside, pelargonidin glycoside, peonidin glycoside, and petunidin glycoside.

The terms ‘precursor’ or ‘flavonoid precursor’ as used herein may refer to any intermediate present in the biosynthetic pathway that leads to the production of catechins or anthocyanins. flavonoid precursors may include, but are not limited to tyrosine, phenylalanine, coumaric acid, p-coumaroyl-CoA, malonyl-CoA, pyruvate, acetyl-CoA, and naringenin.

Cells engineered for the production of a flavonoid or an anthocyanin can have one or multiple modifications, including, without limitation, the downregulation, disruption, or deletion of endogenous genes, the upregulation of an endogenous gene, and the introduction of exogenous genes.

The term “non-naturally occurring”, when used in reference to an enzyme is intended to mean that nucleic acids or polypeptides include at least one genetic alteration not normally found in a naturally occurring polypeptide or nucleic acid sequence. Naturally occurring nucleic acids, and polypeptides can be referred to as “wild-type” or “original”. A host cell, organism, or microorganism that includes at least one genetic modification generated by human intervention can also be referred to as “non-naturally occurring”, “engineered”, “genetically engineered,” or “recombinant”.

A host cell, organism, or microorganism engineered to express or overexpress a gene or nucleic acid sequence, or to overexpress an enzyme or polypeptide has been genetically engineered through recombinant DNA technology to include a gene or nucleic acid sequence that does not naturally encode the enzyme or polypeptide or to express an endogenous gene at a level that exceeds its level of expression in a non-altered cell. As nonlimiting examples, a host cell, organism, or microorganism engineered to express or overexpress a gene or a nucleic acid sequence, or to overexpress an enzyme or polypeptide can have any modifications that affect a

coding sequence of a gene, the position of a gene on a chromosome or regulatory elements associated with a gene. Overexpression of a gene can also be by increasing the copy number of a gene in the cell or organism. Similarly, a host cell, organism, or microorganism engineered to under-express or to have reduced expression of a gene, nucleic acid sequence, or to under-express an enzyme or polypeptide can have any modifications that affect a coding sequence of a gene, the position of a gene on a chromosome or regulatory elements associated with a gene. Specifically included are gene disruptions, which include any insertions, deletions, or sequence mutations into or of the gene or a portion of the gene that affect its expression or the activity of the encoded polypeptide. Gene disruptions include “knockout” mutations that eliminate expression of the gene. Modifications to under-express a gene also include modifications to regulatory regions of the gene that can reduce its expression.

The term “exogenous” or “heterologous” is intended to mean that the referenced molecule or the referenced activity is introduced into the host microbial organism. The molecule can be introduced, for example, by introduction of an encoding nucleic acid into the host genetic material such as by integration into a host chromosome or as non-chromosomal genetic material that may be introduced on a vehicle such as a plasmid. Therefore, the term “endogenous” refers to a referenced molecule or activity that is naturally present in the host.

Genes or nucleic acid sequences can be introduced stably or transiently into a host cell using techniques well known in the art including, but not limited to, conjugation, electroporation, chemical transformation, transduction, and transfection. Optionally, for exogenous expression in *E. coli* or other prokaryotic cells, some nucleic acid sequences in the genes or cDNAs of eukaryotic nucleic acids can encode targeting signals such as an N-terminal mitochondrial or other targeting signal, which can be removed before transformation into prokaryotic host cells, if desired. Furthermore, genes can be subjected to codon optimization with techniques well known in the art to achieve optimized expression of the proteins.

The percent identity (% identity) between two sequences is determined when sequences are aligned for maximum homology. Algorithms well known to those skilled in the art, such as Align, BLAST, Clustal Omega, and others compare and determine a raw sequence similarity or identity, and also determine the presence or significance of gaps in the sequence which can be assigned a weight or score. Such algorithms also are known in the art and are similarly

applicable for determining nucleotide or amino acid sequence similarity or identity and can be useful in identifying orthologs of genes of interest. Additional sequences added to a polypeptide sequence, such as but not limited to immunodetection tags, purification tags, localization sequences (presence or absence), etc., do not affect the % identity.

5 A homolog is a gene or genes that have the same or identical functions in different organisms. Genes that are orthologous can encode proteins with sequence similarity of about 45% to 100% amino acid sequence identity, and more preferably about 60% to 100% amino acid sequence identity. Genes can also be considered orthologs if they share three-dimensional structure but not necessarily sequence similarity, of a sufficient amount to indicate that they have evolved from a common ancestor to the extent that the primary sequence similarity is not identifiable. Paralogous are genes related by duplication within a genome, and can evolve new functions, even if these are related to the original one.

 An engineered cell for producing flavonoids include an exogenous nucleic acid sequence encoding tyrosine ammonia lyase (TAL) activity (alternatively or in addition, an exogenous nucleic acid encoding phenylalanine ammonia-lyase (PAL) activity and an exogenous nucleic acid encoding cinnamate-4-hydroxylase (C4H) activity), an exogenous nucleic acid sequence encoding 4-coumarate-CoA ligase (4CL) activity, an exogenous nucleic acid sequence encoding chalcone synthase (CHS) activity, and an exogenous nucleic acid sequence encoding chalcone isomerase (CHI) activity. Optionally, the engineered cell can further include an exogenous nucleic acid sequence encoding an exogenous nucleic acid sequence encoding flavanone-3-hydroxylase (F3H) activity, an exogenous nucleic acid sequence encoding flavonoid 3'-hydroxylase (F3'H) activity or flavonoid 3',5'-hydroxylase (F3'5'H), an exogenous nucleic acid sequence encoding cytochrome P450 reductase (CPR) activity, an exogenous nucleic acid sequence encoding dihydroflavonol-4-reductase (DFR) activity, and/or an exogenous nucleic acid sequence encoding leucoanthocyanidin reductase (LAR) activity.

 Tyrosine ammonia-lyase (TAL) can be, for example, a member of the aromatic amino acid deaminase family that catalyzes the elimination of ammonia from L-tyrosine to yield p-coumaric acid. An exemplary tyrosine ammonia lyase is the *Saccharothrix espanaensis* tyrosine ammonia lyase (TAL; SEQ ID NO: 1). Also considered for use in the engineered cells provided herein are TALs with SEQ ID NOS: 23-26, TALs listed in Table 1, TAL homologs and variants

having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to SEQ ID:1 that have the activity of a tyrosine ammonia lyase that produces p-coumaric acid from tyrosine.

5 Table 1. Tyrosine ammonia-lyase

Organism	GenBank Accession Number
<i>Rhodotorula glutini</i>	AGZ04575.1
<i>Flavobacterium johnsoniae</i>	WP_012023194.1
<i>Herpetosiphon aurantiacus</i>	ABX02653.1
<i>Rhodobacter capsulatus</i>	ADE83766.1
<i>Saccharothrix espanaensis</i>	AKE50820.1
<i>Trichosporon cutaneum</i>	AKE50834.1

Similar to tyrosine ammonia-lyase, phenylalanine ammonia-lyase (PAL) can be a member of the aromatic amino acid deaminase family that catalyzes the non-oxidative deamination of L-phenylalanine to form trans-cinnamic acid. An exemplary phenylalanine ammonia-lyase is the *Brevibacillus laterosporus* phenylalanine ammonia-lyase (PAL; SEQ ID NO :2). Also considered for use in the engineered cells provided herein are PALs with SEQ ID NOS: 27-29, PAL homologs and variants having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to SEQ ID NO: 2 that have the activity of a phenylalanine ammonia lyase that produces trans-cinnamic acid from phenylalanine.

Cinnamate-4-hydroxylase (C4H) belongs to the cytochrome P450-dependent monooxygenase family and catalyzes the formation of p-coumaric acid from trans-cinnamic acid. Considered for use in the engineered cells provided herein are C4H of *Helianthus annuus L.* (C4H; SEQ ID NO: 3), C4Hs with SEQ ID NOS: 30-32, and C4H homologs of other species, as well as variants of naturally occurring C4Hs having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at

least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to the SEQ ID NO: 3 (C4H, *Helianthus annuus* L.) that have the activity of a C4H.

4-coumarate-CoA ligase (4CL) catalyzes the activation of 4-coumarate to its CoA ester. Considered for use in the engineered cells provided herein are 4CLs of *Petroselinum crispum* (SEQ ID NO: 4), 4CLs in Table 2, 4CLs with SEQ ID NOS: 33-36, and 4CL homologs of other species, as well as variants of naturally occurring 4CLs having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID No: 4 (4CL, *Petroselinum crispum*) that have the activity of a 4CL.

10 Table 2. 4-coumarate-CoA ligases

Organism	GenBank Accession Number
<i>Petroselinum crispum</i>	CAA31697.1
<i>Camellia sinensis</i>	ASU87409.1
<i>Capsicum annuum</i>	KAF3620173.1
<i>Castanea mollissima</i>	KAF3954751.1
<i>Daucus carota</i>	AIT52344.1
<i>Gymura bicolor</i>	BAJ17664.1
<i>Ipomoea purpurea</i>	AHJ60263.1
<i>Lonicera japonica</i>	AGE10594.1
<i>Lycium chinense</i>	QDL52638.1
<i>Nelumbo mucifera</i>	XP_010265453.1
<i>Nyssa sinensis</i>	KAA8540582.1
<i>Solanum lycopersicum</i>	NP_001333770.1
<i>Striga asiatica</i>	GER48539.1

The chalcone synthase (CHS) can be, for example, a type III polyketide synthase that sequentially condenses three molecules of malonyl-CoA with one molecule of p-coumaryl-CoA to produce the naringenin precursor naringenin chalcone or naringenin. An exemplary chalcone synthase is the chalcone synthase of *Petunia x hybrida* (CHS, SEQ ID NO: 5). Also considered for use in the engineered cells provided herein are the genes listed in Table 3, CHSs with SEQ ID: 37-40, and CHS homologs and variants having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO: 5 (CHS, *Petunia x hybrida*) that have the activity of a chalcone synthase.

Table 3. Chalcone synthases

Organism	GenBank Accession Number
<i>Petunia hybrida</i>	AAF60297.1
<i>Acer palmatum</i>	AWN08245.1
<i>Callistephus chinensis</i>	CAA91930.1
<i>Camellia japonica</i>	BAI66465.1
<i>Capsicum annuum</i>	XP_016566084.1
<i>Coffea arabica</i>	XP_027118978.1
<i>Curcuma alismatifolia</i>	ADP08987.1
<i>Dendrobium catenatum</i>	ALE71934.1
<i>Garcinia mangostana</i>	ACM62742.1
<i>Iochroma calycinum</i>	AIY22758.1
<i>Iris germanica</i>	BAE53636.1
<i>Lilium speciosum</i>	BAE79201.1

<i>Lonicera caerulea</i>	ALU09326.1
<i>Lycium ruthenicum</i>	ATB56297.1
<i>Magnolia liliiflora</i>	AHJ60259.1
<i>Matthiola incana</i>	BBM96372.1
<i>Morus alba</i> var. <i>multicaulis</i>	AHL83549.1
<i>Nelumbo nucifera</i>	NP_001305084.1
<i>Nyssa sinensis</i>	KAA8548459.1
<i>Paeonia lactiflora</i>	AEK70334.1
<i>Panax notoginseng</i>	QKV26463.1
<i>Ramunculus asiaticus</i>	AYV99476.1
<i>Rosa chinensis</i>	AEC13058.1
<i>Theobroma cacao</i>	XP_007032052.2

Chalcone isomerase (CHI, also referred to as chalcone flavonone isomerase) catalyzes the stereospecific and intramolecular isomerization of naringenin chalcone into its corresponding (2S)-flavanones. Considered for use in the engineered cells provided herein are CHI of *Medicago sativa* (SEQ ID NO: 6), CHI of Table 4, CHIs with SEQ ID NOS: 41-44, and CHI homologs of other species, as well as variants of naturally occurring CHI having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO: 6 (CHI, *Medicago sativa*) that have the activity of a chalcone isomerase.

Table 4. Chalcone Isomerases

Organism	GenBank Accession Number
<i>Medicago sativa</i>	AGZ04578.1

<i>Dendrobium hybrid cultivar</i>	AGY46120.1
<i>Abrus precatorius</i>	XP_027366189.1
<i>Antirrhinum majus</i>	BAO32070.1
<i>Arachis duranensis</i>	XP_015942246.1
<i>Astragalus membranaceus</i>	ATY39974.1
<i>Camellia sinensis</i>	XP_028119616.1
<i>Castanea mollissima</i>	KAF3958409.1
<i>Cephalotus follicularis</i>	GAV77263.1
<i>Clarkia gracilis subsp. sonomensis</i>	QPF47150.1
<i>Dianthus caryophyllus</i>	CAA91931.1
<i>Glycyrrhiza uralensis</i>	AXO59749.1
<i>Handroanthus impetiginosus</i>	PIN05040.1
<i>Lotus japonicus</i>	CAD69022.1
<i>Morus alba</i>	AFM29131.1
<i>Phaseolus vulgaris</i>	XP_007142690.1
<i>Punica granatum</i>	ANB66204.1
<i>Rhodamnia argentea</i>	XP_030524476.1
<i>Spatholobus suberectus</i>	TKY50621.1
<i>Trifolium subterraneum</i>	GAU12132.1

A nucleic acid sequence encoding a CHI can in some embodiments be fused to a nucleic acid sequence encoding a CHS in an engineered cell as provided herein, such that the CHI

activity is fused to the chalcone synthase activity, *i.e.*, a fusion protein is produced in the engineered cell that has both condensing and cyclization activities.

Flavanone 3-hydroxylase (F3H) catalyzes the stereospecific hydroxylation of (2S)-naringenin to form (2R,3R)-dihydrokaempferol. Other substrates include (2S)-eriodictyol, (2S)-dihydrotricetin and (2S)-pinocembrin. Some F3H enzymes are bifunctional and also catalyzes as flavonol synthase (EC: 1.14.20.6). Considered for use in the engineered cells provided herein are F3H of *Rubus occidentalis* (SEQ ID NO: 7), F3Hs with SEQ ID NOS: 45-48, F3Hs listed in Table 5, and other F3H homologs and variants having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO:7 (F3H, *Rubus occidentalis*) that have the activity of a F3H.

Table 5. Flavanone 3-hydroxylases

Organism	GenBank Accession Number
<i>Rubus occidentalis</i>	ACM17897.1
<i>Abrus precatorius</i>	XP_027347564.1
<i>Nyssa sinensis</i>	KAA8547483.1
<i>Camellia sinensis</i>	AAT68774.1
<i>Morella rubra</i>	KAB1219056.1
<i>Rosa chinensis</i>	PRQ47414.1
<i>Malus domestica</i>	AAD26206.1
<i>Vitis amurensis</i>	ALB75302.1
<i>Iochroma ellipticum</i>	AMQ48669.1
<i>Hibiscus sabdariffa</i>	ALB35017
<i>Cephalotus follicularis</i>	GAV71832

Flavonoid 3'-hydroxylases (F3'H) belongs to the cytochrome P450 family with systematic name of flavonoid, NADPH:oxygen oxidoreductase (3'-hydroxylating). In the flavonoid biosynthetic pathway, F3'H converts dihydrokaempferol to dihydroquercetin (taxifolin) or naringenin to eriodictyol. Considered for use in the engineered cells provided

5 herein are F3'H of *Brassica napus* (F3'H; SEQ ID NO: 8), F3'H with SEQ ID NOS: 49-52, those listed in Table 6, and homologs and variants having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to these F3'H. F3'H is a cytochrome P450 enzyme that requires a cytochrome P450 reductase (CPR) to

10 function. Cytochrome P450 reductases are diflavin oxidoreductases that supply electrons to F3'Hs. The P450 reductase can be from the same species as F3'H or different species from F3'H. Considered for use in the engineered cells provided herein are CPR of *Catharanthus roseus* (SEQ ID NO: 9), additional CPRs listed in Table 7, CPRs with SEQ ID NOS: 53-55, CPR homologs of other species, and variants of naturally occurring CPRs having at least 50%, at least

15 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to these CPRs that have the activity of a CPR. In various embodiments, the N-terminal nucleic acid sequences in the genes of F3'H and/or CPR originated from eukaryotic cells can encode targeting leader peptides, which can be removed before introduction into prokaryotic host cells,

20 if desired. In some embodiments, the hydroxylase complex HpaBC from *E. coli* was used to hydroxylate naringenin to eriodictyol or dihydrokaempferol to dihydroquercetin (taxifolin).

Table 6. Flavonoid 3'-hydroxylases

Organism	GenBank Accession Number
<i>Brassica napus</i>	ABC58722.1
<i>Gerbera hybrid cultivar D1</i>	ABA64468.1
<i>Cephalotus follicularis</i>	GAV84063.1
<i>Theobroma cacao</i>	XP_007037548.1
<i>Phoenix dactylifera</i>	XP_008791304.2

Table 7. Cytochrome P450 reductases

Organism	GenBank Accession Number
<i>Catharanthus roseus</i>	CAA49446.1
<i>Brassica napus</i>	XP_013706600.1
<i>Cephalotus follicularis</i>	GAV59576.1
<i>Camellia sinensis</i>	XP_028084858.1

A nucleic acid sequence encoding a F3'H can in some embodiments be fused to a nucleic acid sequence encoding a CPR in an engineered cell as provided herein, such that the F3'H activity is fused to the CPR activity.

In the cells engineered to produce dihydromyricetin, flavonoid 3', 5'-hydroxylase (F3'5'H) can be used to convert dihydrokaempferol to dihydromyricetin or naringenin to pentahydroxyflavone, which is further converted to dihydromyricetin by a F3H. F3'5'H has the systematic name flavanone,NADPH:oxygen oxidoreductase and catalyzes the formation of 3',5'-dihydroxyflavanone from flavanone. An exemplary F3'5'H is the *Delphinium grandiflorum* F3'5'H (SEQ ID NO: 10). Also considered for use in the engineered cells provided herein include F3'5'H with SEQ ID NOS:56-57, F3'5'H homologs of other species, and variants of naturally occurring F3'5'H having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NOS:10 that have the activity of a F3'5'H.

Dihydroflavonol 4-reductase (DFR) acts on (+) – dihydrokaempferol (DHK), (+)-dihydroquercetin (Taxifolin, DHQ), or dihydromyricetin (DHM) to reduce those compounds to the corresponding cis-flavan-3,4-diol (DHK to leucopelargonidin; Taxifolin to leucocyanidin; DHM to leucodelphinidin). An exemplary DFR is the *Anthurium andraeanum* DFR (SEQ ID NO: 11). Also considered for use in the engineered cells provided herein include DFRs in Table

8, DFRs with SEQ ID NOS: 58-61, and DFR homologs of other species, as well as variants of naturally occurring DFR having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO: 11.

5 Table 8. Dihydroflavonol 4-reductases

Organism	GenBank Accession Number
<i>Eustoma grandiflorum</i>	BAD34461.1
<i>Anthurium andraeanum</i>	AAP20866.1
<i>Camellia sinensis</i>	AAT66505.1
<i>Morella rubra</i>	KAB1203810.1
<i>Dendrobium moniliforme</i>	AEB96144.1
<i>Fragaria x ananassa</i>	AHL46451.1
<i>Rosa chinensis</i>	XP_024167119.1
<i>Acer palmatum</i>	AWN08247.1
<i>Nyssa sinensis</i>	KAA8531902.1
<i>Vitis amurens</i>	I82380.1
<i>Abrus precatorius</i>	XP_027329642.1
<i>Angelonia angustifolia</i>	AHM27144.1
<i>Pyrus pyrifolia</i>	Q84KP0.1
<i>Theobroma cacao</i>	XP_017985307
<i>Theobroma cacao</i>	XP_007051597.2
<i>Brassica oleracea</i> var. <i>capitata</i>	QKO29328.1
<i>Rubus idaeus</i>	AXK92786.1

<i>Citrus sinensis</i>	AAY87035.1
<i>Gerbera hybrida</i>	P51105.1
<i>Cephalotus follicularis</i>	GAV76940.1
<i>Ginkgo biloba</i>	AGR34043.1
<i>Dryopteris erythrosora</i>	QFQ61498.1
<i>Dryopteris erythrosora</i>	QFQ61499.1
<i>Cephalotus follicularis</i>	GAV76942.1

Leucoanthocyanidin reductase (LAR) catalyzes the synthesis of catechin from 3,4-cis-leucocyanidin. LAR also synthesizes afzelechin and gallocatechin. Considered for use in the engineered cells provided herein are LAR of *Desmodium uncinatum* (SEQ ID NO: 12), LARs with SEQ ID NOS: 62-65, and LAR homologs of other species, as well as variants of naturally occurring LAR having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO: 12 (LAR, *Desmodium uncinatum*) that have the activity of a LAR.

Optionally, the cells are further engineered to include an anthocyanin synthase (ANS) which catalyzes the conversion of leucoanthocyanidin or catechin to anthocyanidin, leucopelargonidin to pelargonidin, or leucodelphinidin to delphinidin. Considered for use in the engineered cells provided herein are ANS of *Carica papaya* (SEQ ID NO: 13), ANS with SEQ ID NOS: 66-69, and ANS homologs of other species, as well as variants of naturally occurring ANS having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO:13 (ANS, *Carica papaya*) that have the activity of a ANS.

Optionally, the cells are further engineered to include a flavonoid-3-glucosyl transferase (3GT) to generate anthocyanins by transfer of a sugar moiety such as, without limitation, UDP-

α -D-glucose to anthocyanidins to form glycosylated anthocyanins. Considered for use in the engineered cells provided herein are 3GT of *Vitis labrusca* (SEQ ID NO:14), 3GT with SEQ ID NOS: 70-73, and 3GT homologs of other species, as well as variants of naturally occurring 3GT having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to SEQ ID NO: 14 (3GT, *Vitis labrusca*) that have the activity of a 3GT.

In various aspects, host cells may be engineered for enhanced production of flavonoids or anthocyanins by introducing additional exogenous pathways and/or modifying endogenous metabolic pathways to remove or downregulate competitive pathways to reduce carbon loss, increase precursor supply, improve cofactor availability, reduce byproduct formation, or improve cell fitness. Enhancing or improving production of flavonoids or anthocyanins can be increasing yield, titer, or rate of production.

Thus, a host cell engineered for the production of a flavonoid or anthocyanin can be engineered to include any or any combination of: overexpression of an acetyl-CoA carboxylase (ACC) or an ACC variant; expression or overexpression of at least one enzyme for increasing cell's malonyl-CoA supply that does not rely on the ACC step; expression or overexpression of at least one enzyme to increase tyrosine supply; expression or overexpression of at least one enzyme to increase CoA availability for synthesizing precursors malonyl-CoA or p-coumaryl-CoA; expression or overexpression at least one enzyme to increase heme biosynthesis; deletion or downregulation of at least one fatty acid synthesis enzyme; at least one alcohol dehydrogenase, lactate dehydrogenase, pyruvate oxidase, phosphate acetyl transferase, or acetate kinase; at least one enzyme of a fatty acid degradation pathway, at least one thioesterase, or at least one TCA gene. The foregoing list of modifications is nonlimiting.

Malonyl-CoA is the direct precursor for chalcone synthase to perform sequential condensations with p-coumaryl-CoA. Malonyl-CoA supply can be increased by one or more modifications. Malonyl-CoA is synthesized by acetyl-CoA carboxylase (ACC) via the ATP-dependent carboxylation of acetyl-CoA in a multistep reaction. First, the biotin carboxylase domain catalyzes the ATP-dependent carboxylation of biotin using bicarbonate as a CO₂ donor. In the second reaction, the carboxyl-group is transferred from biotin to acetyl-CoA to form

malonyl-CoA. In most eukaryotes, including fungi, both reactions are catalyzed by a large single chain protein, but in *E. coli* and other bacteria, the activity is catalyzed by a multi-subunit enzyme. Host cells can be engineered for example to express an exogenous acetyl-CoA carboxylase or a variant ACC to increase malonyl-CoA synthesis from acetyl-CoA. For example,

5 *Mucor circinelloides* (SEQ ID NO: 15) acetyl-CoA carboxylase can be introduced into the host cells. Additional examples of ACC genes that may be used in the engineered cells provided herein include, without limitation, the genes listed in Table 9, genes with SEQ ID NOS: 74-76, naturally occurring orthologs of these ACCs, or variants having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at

10 least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to referenced genes. Further, naturally occurring acetyl-CoA carboxylase genes can be further engineered to introduce single or multiple amino acid mutations to increase catalytic activity and/or remove feedback inhibition.

Table 9. Acetyl-CoA carboxylases

Organism	GenBank Accession Number
<i>Lipomyces starkeyi</i>	AJT60321.1
<i>Rhodotorula toruloides</i>	GEM08739.1
<i>Ustilago maydis</i>	XP_011390921.1
<i>Mucor circinelloides</i>	EPB82652.1
<i>Kalahariturber pfeilii</i>	KAF8466702.1
<i>Aspergillus fumigatus</i>	KEY77072.1
<i>Rhodotorula diobovata</i>	TNY18634.1
<i>Leucosporidium creatinivorum</i>	ORY74050.1
<i>Microbotryum intermedium</i>	SCV70467.1
<i>Mixia osmundae</i>	GAA98306.1

<i>Puccinia graminis</i>	KAA1079218.1
<i>Suillus occidentalis</i>	KAG1764021.1
<i>Gymnopilus junonius</i>	KAF8909366.1

Additional strategies for increasing malonyl-CoA include increasing acetyl-CoA, which is converted to malonyl-CoA by acetyl-CoA carboxylase (ACC). Acetyl-CoA can be synthesized from acetate by an acyl-CoA ligase in an ATP-dependent reaction. Acetyl-CoA synthetase (ACS) or acetate-CoA ligase (EC 6.2.1.1.) catalyzes the formation of a new chemical bond between acetate and CoA coenzyme A (CoA). ACSs with native activity on acetate will provide the function of increasing acetyl-CoA supply when cells are either supplied with acetate as a co-feed, or where acetate is produced as a by-product. Other acyl-CoA ligases, having their main activity on other acid substrates, may also have substantial activity on acetate, and are viable candidates for providing acetate-CoA ligase activity in the engineered cells provided herein. The ACSs expressed in the host cells can be prokaryotic or eukaryotic. Cultures of engineered host cells that overexpress a nucleic acid sequence encoding ACS can optionally include acetate in the culture medium. Examples of acetyl-CoA synthase that can be expressed in a host cell engineered to produce a flavonoid or anthocyanin include, without limitation, the ACS gene of *E. coli*, the ACS of *Salmonella typhimurium* (SEQ ID NO:16), and orthologs of these ACSs in other species having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% amino acid identity to these ACSs.

Alternatively, or in addition, an engineered host cell can overexpress a gene encoding pyruvate dehydrogenase (PDH), which converts pyruvate to acetyl-CoA, to increase acetyl-CoA supply. PDH catalyzes an irreversible metabolic step, and the control of its activity is complex and involves control by its substrates and products. Nicotinamide adenine dinucleotide hydrogen (NADH), a product of the PDH reaction, is a competitive inhibitor of the PDH complex. The NADH sensitivity of the PDH complex has been demonstrated to reside in LPD, the enzyme that interacts with NAD⁺ as a substrate. Thus, a variant of the Lpd subunit of PDH can be expressed that includes one or more mutations that reduces inhibition of PDH by NADH. Such an example

is a LPD variant in *E. coli* that contains E354K mutation, and the mutated enzyme was less sensitive to NADH inhibition than the native LPD.

Alternatively, or in addition to strategies for increasing ACC activity and strategies for increasing acetyl-CoA, strategies for increasing malonyl-CoA by mechanisms that do not rely on the activity of an ACC can be employed. For example, a cell engineered to produce a flavonoid or an anthocyanin as provided herein can include an exogenous nucleic acid sequence encoding a malonyl-CoA synthetase (EC 6.2.1.14) that generates malonyl-CoA from malonate. Acyl-CoA synthetase catalyzes the conversion of a carboxylic acid to its acyl-CoA thioester through an ATP-dependent two-step reaction. In the first step, the free fatty acid is converted to an acyl-AMP intermediate with the release of pyrophosphate. In the second step, the activated acyl group is coupled to the thiol group of CoA, releasing AMP and the acyl-CoA product. Nonlimiting examples of malonyl-CoA synthetases include the malonyl-CoA synthetases of *Streptomyces coelicolor* (SEQ ID NO:17), matB of *Rhodopseudomonas palustris* (SEQ ID NO: 77), matB of *Rhizobium sp.* BUS003 (SEQ ID NO: 78), matB of *Ochrobacrum sp.* (SEQ ID NO: 79), or other homologs having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the referenced sequences. Malonate can optionally be added to the culture medium of a culture that includes a cell engineered to express a malonyl-CoA synthetase. In *Rhizobium trifolii*, the matB gene is part of the matABC operon, with matA encoding a malonyl-CoA decarboxylase and matC encoding a putative dicarboxylate carrier protein or malonate transporter. An engineered cell that includes an exogenous gene encoding a malonyl-CoA synthetase can also include an exogenous nucleic acid sequence encoding a malonate transporter, such as a malonate transporter encoded by a matC gene, for example of *Streptomyces coelicolor* (SEQ ID NO:18), of *Rhizobiales bacterium* (SEQ ID NO:80), of *Rhizobium leguminosarum* (SEQ ID NO:81), of *Agrobacterium vitis* (SEQ ID NO: 82), of *Neorhizobium sp.* (SEQ ID NO: 83), or a malonate transporter encoded by DctPQM of *Sinorhizobium medicae*, or encoding a malonyl-CoA transporter having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to a naturally-occurring malonate transporter. Cell cultures of a host cell engineered to express a

malonyl-CoA synthetase and a malonate transporter can include a culture medium that includes malonate.

In additional embodiments, a cell engineered to produce a flavonoid or an anthocyanin is further engineered to include an exogenous nucleic acid sequence encoding malonate CoA-transferase (EC:2.8.3.3; also referred to as the alpha subunit of malonate decarboxylase) that makes malonyl-CoA by direct transfer of the CoA from acetyl-CoA. For example, the alpha subunit of malonate decarboxylase from the mdcACDE gene cluster in *Acinetobacter calcoaceticus* has the malonate CoA-transferase activity. The mdcA gene product, the α subunit, is malonate CoA-transferase, and mdcD gene product, the β subunit, is a malonyl-CoA decarboxylase. The mdcE gene product, the γ subunit, may play a role in subunit interaction to form a stable complex or as a codecarboxylase. The mdcC gene product, the δ subunit, was an acyl-carrier protein, which has a unique CoA-like prosthetic group. When the α subunit is removed from the complex and incubated with malonate and acetyl-CoA, the acetyl-CoA moiety of the prosthetic group binds on an α subunit to exchange the acetyl group for a malonyl group. As the thioester transfer should be thermodynamically favorable, the engineered cells can include a nucleic acid encoding a malonate CoA-transferase to increase malonyl-CoA supply. Examples of mdcAs that can be expressed in an engineered cell as provided herein include, without limitation, mdcA of *Acinetobacter calcoaceticus* (SEQ ID NO: 19), mdcAs of Table 10, mdcAs with SEQ ID NOS: 84-87, or a transferase having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to any of these or other naturally occurring malonate CoA-transferases.

Table 10. Malonate CoA-transferases (malonate decarboxylase subunit alpha)

Organism	GenBank Accession Number
<i>Acinetobacter calcoaceticus</i>	AAB97627.1
<i>Geobacillus sp.</i>	QNU36929.1
<i>Acinetobacter johnsonii</i>	WP_087014029.1
<i>Acinetobacter marinus</i>	WP_092618543.1

<i>Acinetobacter rudis</i>	WP_016655668.1
<i>Psychrobacter sp. G</i>	WP_020444454.1
<i>Moraxella catarrhalis</i>	WP_064617969.1
<i>Zoogloea sp.</i>	MBL0283742.1
<i>Dechloromonas sp.</i>	KAB2923906.1
<i>Stenotrophomonas rhizophila</i>	WP_123729366.1
<i>Xanthomonas cucurbitae</i>	WP_159407614.1

In some embodiments, a cell engineered to produce flavonoids or anthocyanins is further engineered to increase the supply of coenzyme A (CoA) to increase its availability for producing acetyl-CoA, malonyl-CoA, and/or p-coumaroyl-CoA. Strategies for increasing CoA supply include expressing or overexpressing at least one enzyme of a CoA biosynthesis pathway. Pantothenate kinase (EC 2.7.1.33, PanK; CoaA) is the first enzyme in the coenzyme CoA biosynthetic pathway. It phosphorylates pantothenate (vitamin B5) to form 4'-phosphopantothenate at the expense of a molecule of adenosine triphosphate (ATP). It is the rate-limiting step in the biosynthesis of CoA. Three distinct types of PanK have been identified - PanK-I (found in bacteria), PanK-II (mainly found in eukaryotes, but also in the *Staphylococci*) and PanK-III, also known as CoaX (found in bacteria). In *E. coli*, pantothenate kinase is competitively inhibited by CoA itself, as well as by some CoA esters. The type III enzymes CoaX are not subject to feedback inhibition by CoA. In some embodiments, a host cell can be engineered to include a nucleic acid sequence encoding type III pantothenate kinase that is not feedback inhibited by coenzyme A, such as, without limitation, CoaX gene of *Pseudomonas aeruginosa* (EC:2.7.1.33, SEQ ID NO: 20), CoaX of *Streptomyces sp. CLI2509* (SEQ ID NO: 88), CoaX of *Streptomyces cinereus* (SEQ ID:89), or CoaX of *Kitasatospora kifunensis* (SEQ ID NO: 90) Cultures of cells engineered for the production of flavonoids or anthocyanins can in some embodiments include a medium that includes pantothenate, a precursor of CoA biosynthesis, and can optionally also include cysteine, used in the CoA biosynthesis.

Additional strategies to increase malonyl-CoA flux to the flavonoid pathway include mutation or downregulation of one or more genes that function in fatty acid biosynthesis. Fatty acid biosynthesis directly competes with flavonoid biosynthesis for the precursor malonyl-CoA and thus limits flavonoid formation. Without limiting the embodiments to any particular
5 mechanism, limiting fatty acid biosynthesis can increase the malonyl-CoA supply available for flavonoid biosynthesis. In some embodiments, the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF) can be disrupted, attenuated or deleted to reduce fatty acid biosynthesis. Another example of a fatty acid biosynthesis gene of a host cell that may be mutated or downregulated is a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD). Other fatty acid biosynthesis genes of
10 the engineered host cell that can be downregulated include a beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB) and/or acyl carrier protein (*E. coli* acpP).

Additional genetic modifications that may be present in a host cell engineered to produce flavonoids or anthocyanins include downregulation, disruption, or deletion of the gene targets that divert carbon flux to form byproducts such as ethanol, acetate, and lactate. They include
15 genes encoding alcohol dehydrogenase, lactate dehydrogenase, pyruvate oxidase, acetyl phosphate transferase and acetate kinase. In an *E. coli* host cell, genes that are downregulated, disrupted, or deleted can include adhE, ldhA, poxB, and ackA-pta.

Further, a cell engineered for the production of flavonoids or anthocyanins can have one or more genes encoding thioesterases downregulated, disrupted, or deleted to prevent hydrolysis
20 of precursors malonyl-CoA, acetyl-CoA, and/or p-coumaroyl-CoA. Acyl-CoA thioesterase enzymes (ACOTs) catalyze the hydrolysis of acyl-CoAs (short-, medium-, long- and very long-chain), bile acid-CoAs, and methyl branched-CoAs, to the free fatty acid and coenzyme A. For example, in an *E. coli* host one or more of the thioesterase genes tesA, tesB, yciA, and/or ybgC can be downregulated, disrupted, or deleted.

In further embodiments, a cell engineered for the production of flavonoids or anthocyanins can have one or more of fatty acid degradation genes downregulated, disrupted, or
25 deleted to improve precursor supply to the flavonoid pathway. In *E. coli*, for example, the acyl-coenzyme A dehydrogenase (fadE) gene encoding acyl-CoA dehydrogenase, adhesion A (fadA) gene encoding 3-ketoacyl-CoA thiolase, and/or gene encoding fatty acid oxidation complex
30 subunit alpha (fadB) can be downregulated, disrupted, or deleted.

Alternatively, or in addition, genes encoding enzymes of the tricarboxylic acid cycle (TCA), such as succinate dehydrogenase, can be disrupted or downregulated to increase alpha-ketoglutarate supply which serves as a cofactor for the flavonoid and anthocyanin pathway enzymes. Other TCA enzymes that can be downregulated include citrate synthase that converts acetyl-CoA to citrate.

Also considered, in further embodiments, is an engineered host cell for the production of flavonoids or anthocyanins to upregulate the endogenous biosynthesis of amino acid tyrosine. Tyrosine is one of the precursors for the flavonoid biosynthesis and its conversion to 4-coumaric acid is the first committed step of the pathway. Efficient biosynthesis of L-tyrosine from feedstock such as glucose or glycerol is necessary to make biological production economically viable. L-tyrosine is one of the three aromatic amino acids derived from the shikimate pathway. The shikimate pathway is the central metabolic route leading to formation of tryptophan (TRP), tyrosine (TYR), and phenylalanine (PHE), this pathway exclusively exists in plants and microorganisms. It starts with the condensation of intermediates of glycolysis and pentosephosphate-pathway, phosphoenolpyruvate (PEP), and erythrose-4-phosphate (E4P), respectively, which enter the pathway through a series of condensation and redox reactions via 3-deoxy-d-arabino-heptulosonate-7-phosphate (DAHP), 3-dehydroquinate (DHQ), 3-dehydroshikimate (DHS) to shikimate. From there the central branch point metabolite chorismate is obtained via shikimate-3-phosphate under ATP hydrolysis and introduction of a second PEP. The initial step of the shikimate pathway is catalyzed by DAHP synthase isozymes and regulated through feedback-inhibition. In *E. coli* three DAHP synthase isozymes exist (aroF, aroG, aroH), which are each feedback inhibited by one of the three aromatic amino acids (TYR, PHE, TRP), in contrast the two DAHP synthases of plants are not subject to feedback-inhibition. In plants and bacteria, the subsequent five steps are catalyzed by single enzymes. From the central intermediate chorismate the pathway branches off to anthranilate and prephenate leading to aromatic amino acid, para-hydroxybenzoic acid (pHBA) and para-aminobenzoic acid (pABA) synthesis, the latter being a precursor for folate metabolism. Strategies to increase L-tyrosine production can include, without limitation, transcriptional deregulation, removing feedback inhibition, overexpression of rate-limiting enzymes, and/or deletion of the L-phenylalanine branch of the aromatic acid biosynthetic pathway. For example, in an *E. coli* host the *tyrR* gene can be disrupted, feedback-inhibition-resistant versions of the DAHP synthase (aroG) and

chorismate mutase (tyrA) can be introduced, and/or rate-limiting enzymes, shikimate kinase (aroK or aroL) and quinate (QUIN)/shikimate dehydrogenase (ydiB) can be overexpressed. Further, the ppsA, aroG, and/or transketolase (tktA) can be overexpressed or exogenously introduced to enhance tyrosine production.

5 Also considered, in further embodiments, is an engineered host cell for the production of flavonoids or anthocyanins further engineered to upregulate the endogenous biosynthesis of cofactor heme. Cytochrome P450 (CYPs), one of the exogenous genes in the engineered cells provided herein, contain heme as a cofactor. Improving heme supply can be an effective strategy to increase flavonoid biosynthesis. 5-aminolevulinic acid (ALA) is the first committed precursor
10 to the heme pathway. There exist two known alternate routes by which this committed intermediate is generated. One route is the C4 pathway (Shemin pathway), which involves the condensation of succinyl-CoA and glycine to D-aminolevulinic acid by ALA synthase (ALAS). The C4 pathway is restricted to mammals, fungi and purple nonsulfur bacteria. The second route is the C5 pathway, which involves three enzymatic reactions resulting in the biosynthesis of
15 ALA from the five-carbon skeleton of glutamate. The C5 pathway is active in most bacteria, all archaea and plants. Seven additional reactions, including assembly of eight ALA molecules into a cyclic tetrapyrrole, modification of the side chains, and incorporation of reduced iron into the molecule, are required to convert ALA to heme. In an *E. coli* host, the three enzymes involved in ALA biosynthesis are glutamyl-tRNA synthetase (GltX), glutamyl-tRNA reductase (hemA), and
20 glutamate-1-semialdehyde aminotransferase (hemL). In an *E. coli* host, the engineered cells provided herein can be further engineered to express or overexpress hemA or its variants, and/or hemL to increase the heme precursor ALA production. The nonlimiting examples of hemA gene that can be overexpressed include, without limitation, a mutated hemA gene from *Salmonella typhimurium* (EC:1.1.1.70, SEQ ID NO: 21) and hemA with SEQ ID NOS: 91-93. Alternatively,
25 or in addition, a heterologous ALAS gene can be introduced to produce ALA via the C4 pathway. Nonlimiting examples of heterologous ALAS that can be expressed in *E. coli* include ALAS of *Rhodobacter capsulatus* (SEQ ID:22), ALAS with SEQ ID NOS: 94-97, or an ALAS having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least
30 99% identity to any of these or other naturally-occurring ALAS. Further, one or more of the downstream genes (*E. coli* hemB, hemC, hemD, hemE, hemF, hemG, hemI, or hemH) that

catalyze the synthesis of heme from ALA can be overexpressed to drive the flux from ALA to heme production. Cultures of cells engineered for the production of flavonoids or anthocyanins can in some embodiments include a medium that includes succinate and/or glycine, precursors of heme biosynthesis via the C4 pathway.

5 Engineered cells that produce a flavonoid can be engineered to include multiple pathways to enhance flavonoid production. Those skilled in the art will recognize that the embodiments described herein can be combined in multiple ways. Examples of engineered cells having multiple genetic modifications are exemplary only and do not limit the scope of the invention.

10 Enzymes to be expressed or overexpressed in engineered cells according to the invention are set forth in Table 11.

HOST CELLS

 A host cell as provided herein can be a prokaryotic cell or a eukaryotic cell. Eukaryotic cells may be microbial eukaryotic cells, such as, for example, fungal cells or yeast cells. Prokaryotic cells that can be engineered as provided herein include bacterial cells and
15 cyanobacterial cells.

 Host can be selected based on their ability to take up and utilize particular carbon sources, nitrogen sources, or precursor molecules or may be engineered to take up and utilize molecules that may be added to the culture medium.

 Nonlimiting examples of suitable microbial hosts for the bio-production of a flavonoid
20 include, but are not limited to, any gram-negative organisms, more particularly a member of the family Enterobacteriaceae, such as *E. coli*, any gram-positive microorganism, for example *Bacillus subtilis*, *Lactobacillus* sp. or *Lactococcus* sp.; a yeast, for example *Saccharomyces cerevisiae*, *Pichia pastoris* or *Pichia stipitis*; and other groups or microbial species. More particularly, suitable microbial hosts for the bio-production of a flavonoid generally include, but
25 are not limited to, members of the genera *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula*, and *Saccharomyces*.

CULTURE MEDIUM

In yet another aspect, methods for producing a flavonoid or an anthocyanin that include incubating a culture of an engineered host cell as provided herein to produce a flavonoid or an anthocyanin. The methods can further include recovering the flavonoid or anthocyanin from the culture medium, whole culture, or cells.

The culture comprises cells engineered for the production of flavonoids or anthocyanins in a culture medium. In various embodiments the engineered cells can be prokaryotic or eukaryotic cells. The culture medium includes at least one carbon source that is also an energy source. Exemplary carbon sources include glucose, glycerol, sucrose, fructose, and xylose. Such carbon sources may be purified or crude, including a biomass comprising glycerol, for example, crude glycerol produced as a byproduct of biodiesel production from corn waste. In addition, the culture medium can include one or more other carbon sources or compounds to increase precursor generation or cofactor supply such as, without limitation, tyrosine, phenylalanine, coumaric acid, acetate, malonate, succinate, glycine, bicarbonate, biotin, naringenin, 5-aminolevulinic acid, thiamine, pantothenate, alpha-ketoglutarate, and ascorbate. In some embodiments, tyrosine and coumaric acid are provided in the culture medium. In some embodiments, tyrosine, alpha-ketoglutarate, 5-aminolevulinic acid, and ascorbate are provided in the culture medium.

Culture conditions can include aerobic, microaerobic or any combination alternating aerobic/microaerobic growth conditions. Further, culture conditions can include shake flasks, fermentation, and other large scale culture procedures. An exemplary growth condition for achieving a flavonoid product include aerobic or microaerobic fermentation conditions. The culture conditions can be scaled up and grown continuously for manufacturing flavonoid product. Exemplary growth procedures include, for example, fed-batch fermentation and batch separation. In an exemplary batch fermentation protocol, the cells are grown in a bioreactor that is well controlled for growth temperature, oxygen, pH, carbon sources, and other compounds. The desired temperature can be from, for example, 20-37 °C, depending on the growth characteristics of the production cells and desired conditions for the fermented products. The pH of the bioreactor can be controlled to range from 5-8 or left uncontrolled in some cases. The batch fermentation period can last in the range of several hours to several days, for examples, 8

to 96 hours. Upon completion of the cultivation period, the fermenter contents can be passed through a cell separation unit to remove cells and cell debris. The cells can be lysed or disrupted enzymatically or chemically prior to or after separation of cells from the fermentation broth, as desired, in order to release additional product. To purify the flavonoids and/or anthocyanins to
5 homogeneity the solution containing the flavonoids and/or anthocyanins was concentrated and the product purified via ion exchange or silica-based chromatography. The resulting solution was either lyophilized to yield the products in a solid form or was concentrated into a liquid solution.

In some embodiments, a method of producing a flavonoid or an anthocyanin comprises culturing an engineered cell disclosed herein in a culture medium to produce a flavonoid or an
10 anthocyanin. In some embodiments, glycerol is used as a carbon feedstock. In some embodiments, the glycerol is crude glycerol. In some embodiments, the method comprises isolating naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside. In some embodiments the naringenin,
15 dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of greater than 50%, greater than 55%, greater than 60%, greater than 65%, greater than 70%, greater than 75%, greater than 80%, greater than 85%, greater than 90%, or greater than 95%. In some embodiments, the naringenin,
20 dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 50% to about 99%, e.g., from about 50% to about 95% (for example from: about 50%, 55%, 60%, 65%, 70%, 75%, 80% to about: 85%, 90%, 95%, 97.5%, 99% or 99.9%). In some embodiments, the naringenin,
25 dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 50% to: about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol,
30 leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be

isolated at a purity of from about 55% to: about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 60% to: about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 65% to: about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 70% to: about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 75% to: about 80%, about 85%, about 90%, about 95%, or about 99%, from about 80% to about 85%, about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 85% to: about 90%, about 95%, or about 99%. In some embodiments, the naringenin, dihydrokaempferol, taxifolin, eriodictyol, leucocyanidin, leucodelphinidin, leucopelargonidin, (+)-catechin, cyanidin, delphinidin, pelargonidin, cyanidin glucoside, delphinidin glucoside or pelargonidin glucoside may be isolated at a purity of from about 90% to about 95%, or about 99%, or from about 95% to about 99% or greater.

VIII. EXAMPLES

USING THE MODIFIED CELL TO CREATE PRODUCTS

30 **Example 1 - Production of naringenin in *E. coli***

An *E. coli* cell derived from MG1655 was engineered to overexpress ACC (SEQ ID NO: 15), TAL (SEQ ID NO: 1), 4CL (SEQ ID NO: 4), CHS (SEQ ID NO: 5), and CHI (SEQ ID NO: 6) to produce naringenin when substrates tyrosine and coumaric acid were supplied in culture medium. ACC was expressed on a medium-copy plasmid (15-20 copies) while TAL, 4CL, CHS, and CHI were expressed on the chromosome. Cells of an OD 2.5 were cultured in a 48-well plate at 30 degree for 24 hours with a shaking speed of 600 RPM in minimal medium supplied with trace element, vitamins, 1 mM tyrosine, 1 mM coumaric acid, and 2% glycerol. Cell cultures were extracted with DMSO at 1:1 ratio and centrifuged for 15 mins. The supernatant was analyzed for naringenin with HPLC. The cells produced 232 μ M naringenin.

Variants of the foregoing host cell may be prepared using one or more of ACC (SEQ ID NO: 15), TAL (SEQ ID NO: 1), 4CL (SEQ ID NO: 4), CHS (SEQ ID NO: 5), and CHI (SEQ ID NO: 6) with one or more homologs of ACC (SEQ ID NO: 15), TAL (SEQ ID NO: 1), 4CL (SEQ ID NO: 4), CHS (SEQ ID NO: 5), or CHI (SEQ ID NO: 6), or combinations of two or more thereof, wherein the homologous enzymes have at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the referenced enzymes.

Example 2 - Production of dihydrokaempferol in *E. coli*

An *E. coli* cell derived from MG1655 was engineered to overexpress F3H (SEQ ID NO: 7) on the chromosome to produce dihydrokaempferol when substrate naringenin was supplied in culture medium. Cells of an OD 0.5-0.7 were cultured in a 24-well plate at 30 degree for 18 hours with a shaking speed of 200 RPM in minimal medium supplied with 2% glycerol, trace elements, 0.8 mM naringenin, 65 mg/L 5-aminolevulinic acid, 0.1 mM ferrous sulfate, 0.1 mM 2-oxoglutarate, and 2.5 mM ascorbic acid. Cell cultures were extracted with DMSO and centrifuged for 15 minutes. The supernatant was analyzed for dihydrokaempferol with HPLC.

The cells produced 315 μ M dihydrokaempferol.

Variants of the foregoing host cell may be prepared using a homolog of F3H (SEQ ID NO: 7), wherein the homologous enzyme has at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the referenced enzyme.

Example 3 - Production of taxifolin in *E. coli*

An *E. coli* strain derived from MG1655 was engineered to overexpress F3H (SEQ ID NO: 7), F3'H (SEQ ID NO: 8), and CPR (SEQ ID NO: 9) to produce taxifolin when the substrate naringenin was supplied in culture medium. F3H was overexpressed on the chromosome while
5 F3'H and CPR were overexpressed on a medium-copy plasmid. Cells of an OD 0.5-0.7 were cultured in a 24-well plate at 30 degree for 18 hours with a shaking speed of 200 RPM in minimal medium supplied with 2% glucose, 0.8 mM naringenin, 65 mg/L 5-aminoleuvulinic acid, 0.1 mM ferrous sulfate, 0.1 mM 2-oxoglutarate, and 2.5 mM ascorbic acid. Cell cultures were extracted with 50% DMSO and centrifuged for 15 minutes. The supernatant was analyzed for
10 taxifolin with HPLC. The cells produced 500 µM taxifolin.

Variants of the foregoing host cell may be prepared using one or more of F3H (SEQ ID NO: 7), F3'H (SEQ ID NO: 8), and CPR (SEQ ID NO: 9) along with one or more homologs of F3H (SEQ ID NO: 7), F3'H (SEQ ID NO: 8), and CPR (SEQ ID NO: 9), or combinations of two or more thereof, wherein the homologous enzymes have at least 50%, at least 55%, at least 60%, at
15 least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the referenced enzymes.

Example 4 - Production of anthocyanidins and anthocyanins

An *E. coli* strain derived from MG1655 was engineered to overexpress ANS (SEQ ID NO: 13) and 3GT (SEQ ID NO: 14) to produce cyanidin-3-O-glucoside when the substrate (+)-
20 catechin was supplied in culture medium. ANS and 3GT were overexpressed on the chromosome. Cells of an OD 0.5-0.7 were cultured in a 24-well plate at 30 degree for 18 hours with a shaking speed of 200 RPM in minimal medium supplied with 1.0% glucose, 2.0 mM (+)-catechin, 0.1 mM 2-oxoglutarate, and 2.5 mM ascorbic acid. Cell cultures were acidified with 2M HCL and extracted with 100% Ethanol. The supernatant was analyzed for cyanidin-3-O-
25 glucoside by HPLC. The cells produced 50 mg/L cyanidin-3-O-glucoside.

Variants of the foregoing host cell may be prepared using one or both of ANS (SEQ ID NO: 13) and 3GT (SEQ ID NO: 14) along with a homolog of ANS (SEQ ID NO: 13), 3GT (SEQ ID NO: 14), or both, wherein the homologous enzymes have at least 50%, at least 55%, at least

60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the referenced enzymes.

ANALYTICAL METHODS

Example 5 - Flavonoid precursors and flavonoids

5 For sampling naringenin, eriodictyol, dihydrokaempferol and taxifolin, extraction of total flavonoids from *E. coli* were performed on whole cell broth. 500 μ L of whole cell broth was vortexed for 30 seconds with 500 μ L of DMSO (dimethyl sulfoxide) and centrifuged for 15 minutes. For HPLC analysis, 50 μ L of supernatant was transferred to an HPLC vial.

10 The HPLC method was as follows: An Agilent 1200 HPLC was fitted with an Ascentis C18 Column 150 mm X 4.6 mm, 3 μ m, equipped with a R-18 (3 μ m) guard column. The column was heated to 30 °C with the sample block being maintained at 25 °C. For each sample, 5 μ L was injected and the product was eluted at a flow rate of 1.5 mL/min using 0.1% phosphoric acid in water (solvent A), acetonitrile (solvent B), and methanol (solvent C) with the following gradient:

Time	C		
e	A (%)	B (%)	(%)
0	85	10	5
2.5	85	10	5
7.5	70	25	5
12.5	50	45	5
15	85	10	5

15

The run time was a total of 15 minutes with naringenin, eriodictyol, dihydrokaempferol and taxifolin eluting at 12.50, 11.56, 10.20, and 8.85 minutes respectively. A diode array detector (DAD) was used for the detection of the molecule of interest at 288 nm.

Example 6 - Anthocyanidins and anthocyanins

For sampling (+)-catechin, cyanidin, and cyanidin-3-glucoside the reaction fluid was acidified with 13 M HCl (1:40 v/v), and extracted with 100% ethanol followed by mixing, centrifugation and filtration through a 0.45 µm filter. The HPLC method was as follows: An Agilent 1200 HPLC was fitted with a LiChrospher RP-8 Column 250 mm X 4.6 mm, 5 µm, 5 equipped with a LiChrospher 100 RP-8 (5 µm) LiChroCART 4-4 guard column. The column was heated to 25 °C with the sample block being maintained at 25 °C. For each sample, 10 µL was injected and the product was eluted at a flow rate of 1.0 ml/min using 0.1% phosphoric acid in water (solvent A) and acetonitrile (solvent B) with the following gradient: 90% A to 10% A for 12 min, 90% A for 0.5 min, and 90% A for 3.5 min for column equilibration. The run time 10 was a total of 16 minutes with cyanidin-3-glycoside eluting at 6.95 mins and cyanidin eluting at 8.9 minutes. A diode array detector (DAD) was used for the detection of the molecule of interest at either 280 nm or 530 nm.

Example 7 – Flavonoid Production

The example provides a combination of modifications to the *E.coli* host genome 15 including deletions and overexpression of enzymes from other organisms to recapitulate the bioproduction pathway described in Figure 4. Accordingly, the invention provides an engineered host cell that comprises one or more genetic modifications (as shown in FIG. 4 and described in this Example 7 and herein above in this application) that result in production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical 20 intermediates, by the engineered host cell. In certain embodiments, the production of flavonoid or anthocyanin from a carbon source that can also be an energy source occurs through enzymatic transformation. In certain embodiments, the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, and (v) any combination thereof. In certain embodiments, the engineered host cell is cultured in a medium comprising 25 molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, and (v) any combination thereof. As shown in FIG. 4, in certain embodiments, one or more genetic modifications lead to increase in metabolic flux to flavonoid precursors or cofactors. As shown in FIG. 4, in certain embodiments, one or more of the genetic modifications cause reduction of formation of byproducts. As shown in FIG. 4, in certain embodiments, one or 30 more genetic modifications are selected from: (i) one or more modifications for over-expressing

one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof.

5 As shown in FIG. 4, in certain embodiments, the engineered host cell is cultured in a medium comprising molecules selected from: tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid.

10 As shown in FIG. 4, in certain embodiments, the engineered host cell comprises at least one or more nucleic acid sequences selected from: (i) a nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) a nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) cinnamate 4-hydroxylase; and (iv) any combination thereof. As shown in FIG. 4, in certain embodiments, the engineered host cell comprises at least one or more peptides selected from: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof.

15 As shown in FIG. 4, in certain embodiments, one or more genetic modifications decreases fatty acid biosynthesis. As shown in FIG. 4, in certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms

20 dihydrokaempferol from naringenin; and (v) any combinations thereof.

25 As shown in FIG. 4, in certain embodiments, the engineered host cell comprises at least one or more nucleic acid sequences selected from: (i) nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof.

30

As shown in FIG. 4, in certain embodiments, the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof.

The compositions as described above, can be used in methods described herein for increasing the production of flavonoids or anthocyanins. Such methods involve providing any of the compositions described above to result in enzymatic transformation by the engineered host cell of glycerol through multiple chemical intermediates into a flavonoid or anthocyanin (such as shown in part or in whole in FIG. 4).

In yet another aspect, it is envisioned that the pathway illustrated in FIG. 4 can be carried out using a plurality of engineered host cells, as opposed to a single host cell as described above. In such embodiments, the plurality of the engineered host cells have one or more genetic modifications that result in enzymatic transformation by the engineered host cell of glycerol through multiple chemical intermediates into a flavonoid or anthocyanin (as shown in FIG. 4).

Aspects of the invention are now described with reference herein to FIG. 4.

Step 1: conversion of pyruvate to acetate. *poxB* is deleted to reduce carbon loss and eliminate the byproducts.

Step 2: conversion of pyruvate to lactate. *ldhA* is deleted to reduce carbon loss and eliminate the byproducts.

Step 3: conversion of Acetyl-CoA to acetate. *ackA-pta* is deleted to reduce carbon loss and eliminate the byproducts.

Step 4: conversion of Acetyl-CoA to ethanol (EtOH). *adhE* is deleted to reduce carbon loss and eliminate the byproducts.

5 Step 5: conversion of acetyl-CoA to a substrate for the tricarboxylic acid cycle (TCA).

Step 6: conversion of acetyl-CoA to mal-CoA. Heterologous ACC is expressed to increase the concentration of available mal-CoA. Heterologous ACC may be obtained from fungal species. Accordingly, embodiments of the invention provide an engineered host cell that comprises one or more genetic modifications to increase the production and/or availability of
10 malonyl-CoA. In certain embodiments, the engineered host cell comprises one or more genetic modifications selected from: (i) expression of acetyl-CoA carboxylase (ACC); and (ii) overexpression of acetyl-CoA carboxylase. In another embodiment, the engineered host cell is an *E. coli*. In certain embodiments, the acetyl-CoA carboxylase is from: *Mucor circinelloides*, *Rhodotorula toruloides*, *Lipomyces starkeyi*, and *Ustilago maydis*, and orthologs of acetyl-CoA
15 carboxylase having at least 50% amino acid identity to the acetyl-CoA carboxylase of these aforementioned species. In certain embodiments, one or more genetic modification is deletion or attenuation of one or more fatty biosynthetic genes resulting in decrease in fatty acid biosynthesis. In certain embodiments, one or more genetic modification is overexpression of acetyl-CoA synthase (ACS). In certain embodiments, the acetyl-CoA synthase is selected from:
20 acetyl-CoA synthase gene of *E. coli*, acetyl-CoA synthase gene of *Salmonella typhimurium*, and orthologs of acetyl-CoA synthase gene in any other species having at least 50% amino acid identity to the acetyl-CoA synthase gene of *E. coli* and *Salmonella typhimurium*. In certain embodiments, one or more genetic modification is selected from a group consisting of: (i) overexpression a gene encoding pyruvate dehydrogenase (PDH), wherein the PDH may include
25 E354K mutation; (ii) exogenous nucleic acid sequence encoding a malonyl-CoA synthetase; (iii) upregulation of endogenous pantothenate kinase (PanK), wherein PanK is not feedback inhibited by coenzyme A; (iv) exogenous nucleic acid sequence encoding a malonate transporter; and (v) any combinations thereof. In certain embodiments, the malonyl-CoA synthetase is selected from of malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a
30 malonyl-CoA synthetase having at least 50% identity to any of these or other naturally occurring

malonyl-CoA synthetases. In certain embodiments, one or more genetic modifications to decrease fatty acid biosynthesis is selected from: (i) mutation or downregulation of a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD); (ii) modifications to the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF); (iii) downregulation of beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB); (iv) downregulation of acyl carrier protein (*E. coli* acpP); and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises peptides selected from: (i) acetyl-CoA carboxylase (ACC) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 15 or SEQ ID NO: 16; (ii) malonate CoA-transferase having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 19; (iii) acetyl-CoA synthase (ACS) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 16; (iv) malonyl-CoA synthase having an amino acid sequence at least 80% identical SEQ ID NO: 77, SEQ ID NO: 78, or SEQ ID NO: 79; (v) malonate transporter having an amino acid sequence at least 80% identical to SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 83, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 86, or SEQ ID NO: 87; (vi) pantothenate kinase having an amino acid sequence at least 80% identical to SEQ ID NO: 88, SEQ ID NO: 89, or SEQ ID NO: 90; and (vii) any combinations thereof.

In another aspect, the invention provides a method of increasing the production of flavonoids comprising an engineered host cell, wherein the one or more engineered host cells comprise one or more genetic modifications to increase production and/or availability of malonyl-CoA. In certain embodiments, the engineered host cell comprises one or more genetic modifications selected from: (i) expression of acetyl-CoA carboxylase (ACC); and (ii) overexpression of acetyl-CoA carboxylase. In another embodiment, the engineered host cell is an *E. coli*. In certain embodiments, the acetyl-CoA carboxylase is from: *Mucor circinelloides*, *Rhodotorula toruloides*, *Lipomyces starkeyi*, and *Ustilago maydis*, and orthologs of acetyl-CoA carboxylase having at least 50% amino acid identity to the acetyl-CoA carboxylase of these aforementioned species. In certain embodiments, one or more genetic modification is deletion or attenuation of one or more fatty biosynthetic genes resulting in decrease in fatty acid biosynthesis. In certain embodiments, one or more genetic modification is overexpression of acetyl-CoA synthase (ACS). In certain embodiments, the acetyl-CoA synthase is selected from: acetyl-CoA synthase gene of *E. coli*, acetyl-CoA synthase gene of *Salmonella typhimurium*, and

orthologs of acetyl-CoA synthase gene in any other species having at least 50% amino acid identity to the acetyl-CoA synthase gene of *E. coli* and *Salmonella typhimurium*. In certain embodiments, one or more genetic modification is selected from a group consisting of: (i) overexpression a gene encoding pyruvate dehydrogenase (PDH), wherein the PDH may include E354K mutation; (ii) exogenous nucleic acid sequence encoding a malonyl-CoA synthetase; (iii) upregulation of endogenous pantothenate kinase (PanK), wherein PanK is not feedback inhibited by coenzyme A; (iv) exogenous nucleic acid sequence encoding a malonate transporter; and (v) any combinations thereof. In certain embodiments, the malonyl-CoA synthetase is selected from of malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a malonyl-CoA synthetase having at least 50% identity to any of these or other naturally occurring malonyl-CoA synthetases. In certain embodiments, one or more genetic modifications to decrease fatty acid biosynthesis is selected from: (i) mutation or downregulation of a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD); (ii) modifications to the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF); (iii) downregulation of beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB); (iv) downregulation of acyl carrier protein (*E. coli* acpP); and (v) any combinations thereof. In certain embodiments, the engineered host cell comprises peptides selected from: (i) acetyl-CoA carboxylase (ACC) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 15 or SEQ ID NO: 16; (ii) malonate CoA-transferase having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 19; (iii) acetyl-CoA synthase (ACS) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 16; (iv) malonyl-CoA synthase having an amino acid sequence at least 80% identical SEQ ID NO: 77, SEQ ID NO: 78, or SEQ ID NO: 79; (v) malonate transporter having an amino acid sequence at least 80% identical to SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 83, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 86, or SEQ ID NO: 87; (vi) pantothenate kinase having an amino acid sequence at least 80% identical to SEQ ID NO: 88, SEQ ID NO: 89, or SEQ ID NO: 90; and (vii) any combinations thereof.

Step 7: conversion of mal-CoA to malonyl-ACP (acyl carrier protein). malonyl-coA-ACP transacylase (fabD) is downregulated to increase carbon flux.

Step 8: conversion of malonyl-ACP to 3-ketoacyl-ACP. beta-ketoacyl-ACP synthase II (fabF) is downregulated to increase carbon flux.

Step 9: conversion of mal-CoA to naringenin chalcone; conversion of coumaryl-CoA to naringenin chalcone. A heterologous CHS is overexpressed.

5 Step 10: conversion of naringenin chalcone to naringenin. A heterologous CHI is overexpressed.

Steps 11, 12, and 13: conversion of naringenin to dihydrokaempferol (DHK); conversion of naringenin to eriodictyol (EDL); conversion of eriodictyol (EDL) to dihydroquercetin (DHQ); conversion of (DHK) to dihydroquercetin (DHQ); conversion of dihydrokaempferol (DHK) to
10 dihydromyricetin (DHM); conversion of pentahydroxyflavone (PHF) to dihydromyricetin (DHM). Heterologous F3'5'H, F3H, F3H, and/or CPR are overexpressed. Accordingly, as shown in FIG. 4, in another aspect, the invention provides method of increasing the production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictyol (EDL), and/or pentahydroxyflavone (PHF) comprising an engineered host cell, wherein the engineered host
15 cell comprises cytochrome P450 reductase (CPR) and at least one of flavanone-3'-hydroxylase (F3'H) or flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the precursor for increase in production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictyol (EDL), and/or pentahydroxyflavone (PHF) is naringenin and/or dihydrokaempferol (DHK). In certain embodiments, the engineered host cell further comprises peptides selected from a group
20 consisting of: (i) flavonoid 3'-hydroxylase (F3'H); (ii) cytochrome P450 reductase (CPR); and (iii) any combination thereof. In certain embodiments, the engineered host cell produces eriodictyol or taxifolin. In certain embodiments, the engineered host cell further comprises flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the engineered host cell produces pentahydroxyflavone or dihydromyricetin. In certain embodiments, flavonoid 3'-hydroxylase
25 (F3'H) is truncated to remove the N-terminal leader sequence. In certain embodiments, cytochrome P450 reductase (CPR) is truncated to remove the N-terminal leader sequence. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavanone-3-hydroxylase (F3H) has an amino
30 acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 7. In certain

embodiments, flavanone-3'-hydroxylase (F3'H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 8. In certain embodiments, cytochrome P450 reductase (CPR) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 9. In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 10, (ii) SEQ ID NO. 56, and (iii) SEQ ID NO. 57. In certain embodiments, the engineered host cell further comprises cytochrome b₅. In certain embodiments, cytochrome b₅ has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 98.

As shown in FIG. 4, in another aspect, the invention provides method of increasing the production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictyol (EDL), and/or pentahydroxyflavone (PHF) comprising an engineered host cell, wherein the engineered host cell comprises cytochrome P450 reductase (CPR) and at least one of flavanone-3'-hydroxylase (F3'H) or flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the precursor for increase in production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictyol (EDL), and/or pentahydroxyflavone (PHF) is naringenin and/or dihydrokaempferol (DHK). In certain embodiments, the engineered host cell further comprises peptides selected from a group consisting of: (i) flavonoid 3'-hydroxylase (F3'H); (ii) cytochrome P450 reductase (CPR); and (iii) any combination thereof. In certain embodiments, the engineered host cell produces eriodictyol or taxifolin. In certain embodiments, the engineered host cell further comprises flavonoid 3',5'-hydroxylase (F3'5'H). In certain embodiments, the engineered host cell produces pentahydroxyflavone or dihydromyricetin. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is truncated to remove the N-terminal leader sequence. In certain embodiments, cytochrome P450 reductase (CPR) is truncated to remove the N-terminal leader sequence. In certain embodiments, flavonoid 3'-hydroxylase (F3'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) is fused with cytochrome P450 reductase (CPR). In certain embodiments, flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 7. In certain embodiments, flavanone-3'-hydroxylase (F3'H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 8. In certain embodiments, cytochrome P450 reductase (CPR) has an amino acid sequence at least 80% identical to the polypeptide set

forth in SEQ ID NO. 9. In certain embodiments, flavonoid 3',5'-hydroxylase (F3'5'H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 10, (ii) SEQ ID NO. 56, and (iii) SEQ ID NO. 57. In certain embodiments, the engineered host cell further comprises cytochrome b₅. In certain embodiments, cytochrome b₅ has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 98.

Step 14: conversion of dihydroquercetin (DHQ) to leucocyanidin (LC); conversion of dihydrokaempferol (DHK) to leucopelargonidin (LP); and conversion of dihydromyricetin (DHM) to leucodelphinidin (LD). Heterologous DFR is overexpressed.

Step 15: conversion of leucocyanidin (LC) to catechin; conversion of leucodelphinidin (LD) to galocatechin; and conversion of leucopelargonidin (LP) to afzelechin. Heterologous LAR is overexpressed.

Step 16: conversion of catechin to cyanidin; conversion of leucocyanidin (LC) to catechin; conversion of leucodelphinidin (LD) to delphinidin; conversion of galocatechin to delphinidin; conversion of leucopelargonidin (LP) to pelargonidin; or conversion of afzelechin to pelargonidin. Heterologous ANS is overexpressed. Step 16 could be carried in vivo or in a cell-free medium. Accordingly, as shown in FIG. 4, in another aspect, the invention provides an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G). In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase. In certain embodiments, the anthocyanin synthase is selected from: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO: 13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof. In certain embodiments, one or more engineered host cells comprises flavonoid-3-glucosyl transferase (3GT). In certain embodiments, flavonoid-3-glucosyl transferase is selected from: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO: 14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof. In certain embodiments, one or more

genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications comprises overexpression of
5 anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications are selected from a group consisting of: (i) anthocyanin synthase, (ii) flavonoid-3-glucosyl transferase (3GT), and (iii) a combination thereof.

In another aspect, the invention provides a method for increasing the production of flavonoids comprising an engineered host cell, wherein the engineered host cell comprises one or
10 more genetic modifications to increase transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G). In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase. In certain embodiments, the anthocyanin synthase is selected from: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67,
15 SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof. In certain embodiments, one or more engineered host cells comprises flavonoid-3-glucosyl transferase (3GT). In certain embodiments, flavonoid-3-glucosyl transferase is selected from: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl
20 transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof. In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, one or more genetic modifications comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications
25 comprises overexpression of anthocyanin synthase and flavonoid-3-glucosyl transferase (3GT). In certain embodiments, the one or more genetic modifications are selected from a group consisting of: (i) anthocyanin synthase, (ii) flavonoid-3-glucosyl transferase (3GT), and (iii) a combination thereof.

In another aspect, the invention provides a method of increasing the transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G) comprising anthocyanin synthase, wherein the anthocyanin synthase is selected from: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof.

In another aspect, the invention provides a method of increasing the transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G) comprising flavonoid-3-glucosyl transferase (3GT), wherein the flavonoid-3-glucosyl transferase is selected from: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof.

Step 17: conversion of pelargonidin to callistephin; conversion of delphinidin to myrtillin (De3G); conversion of cyanidin to Cy3G. Heterologous 3GT was overexpressed in *E. coli*. Step 17 could be carried in vivo or as a cell-free reaction.

Step 18: conversion of pyruvate to phosphoenolpyruvate (PEP). ppsA is overexpressed to upregulate tyrosine.

Step 19: conversion of fructose-6-phosphate (F6P) to erythrose-4-phosphate (E4P). tktA is overexpressed to upregulate tyrosine.

Step 20: conversion of phosphoenolpyruvate (PEP) to deoxy-d-arabino-heptulosonate-7-phosphate (DAHP). aroG variant is overexpressed to upregulate tyrosine.

Step 21: conversion of deoxy-d-arabino-heptulosonate-7-phosphate (DAHP) to dehydroquinate (DHQ); conversion of erythrose-4-phosphate (E4P) to dehydroquinate (DHQ).

Step 22: conversion of dehydroquinate (DHQ) to 3-dehydroshikimate (DHS).

Step 23: conversion of 3-dehydroshikimate (DHS) to shikimic acid (SHK). aroE is overexpressed to upregulate tyrosine.

Step 24: conversion of shikimic acid (SHK) to shikimate-3-phosphate (S3P).

Step 25: conversion of shikimate-3-phosphate (S3P) to 5-enolpyruvylshikimate-3-phosphate (EPSP).

Step 26: conversion of 5-enolpyruvylshikimate-3-phosphate (EPSP) to chorismic acid (CHA).

5 Step 27: conversion of chorismic acid (CHA) to prephenate (PPA); conversion of prephenate (PPA) to 4-hydroxy-phenylpyruvate (HPP). tryA variant is overexpressed.

Step 28: conversion of 4-hydroxy-phenylpyruvate (HPP) to tyrosine; conversion of phenylpyruvate (POPP) to phenylalanine (Phe). Accordingly, as shown in FIG. 4, embodiments of the invention provide an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase endogenous biosynthesis of tyrosine. In certain
10 embodiments, one or more genetic modifications comprises upregulation of 3-deoxy-D-arabino-heptulosonate synthase. In certain embodiments, one or more genetic modifications are selected from: (i) upregulation of chorismate mutase; (ii) upregulation of prephenate dehydrogenase; (iii) overexpression of shikimate kinase; (iv) overexpression of shikimate dehydrogenase; and (v) any
15 combinations thereof. In certain embodiments, one or more genetic modifications comprises downregulation of L-phenylalanine biosynthetic pathway. In certain embodiments, one or more genetic modifications comprises expression of exogenous phosphoenolpyruvate synthase (ppsA). In certain embodiments, one or more genetic modifications comprises expression of exogenous transketolase (tktA). In certain embodiments, wherein the one or more genetic modifications
20 comprises disruption of tyrR gene.

As shown in FIG. 4, in another aspect, the invention provides a method of increasing endogenous biosynthesis of tyrosine comprising an engineered cell, wherein the engineered host cell comprises one or more genetic modifications to increase endogenous biosynthesis of tyrosine. In certain embodiments, one or more genetic modifications comprises upregulation of
25 3-deoxy-D-arabino-heptulosonate synthase. In certain embodiments, one or more genetic modifications are selected from: (i) upregulation of chorismate mutase; (ii) upregulation of prephenate dehydrogenase; (iii) overexpression of shikimate kinase; (iv) overexpression of shikimate dehydrogenase; and (v) any combinations thereof. In certain embodiments, one or more genetic modifications comprises downregulation of L-phenylalanine biosynthetic pathway.
30 In certain embodiments, one or more genetic modifications comprises expression of exogenous

phosphoenolpyruvate synthase (ppsA). In certain embodiments, one or more genetic modifications comprises expression of exogenous transketolase (tktA). In certain embodiments, wherein the one or more genetic modifications comprises disruption of tyrR gene.

Step 29: conversion of tyrosine to coumaric acid. A heterologous TAL is overexpressed.

5 Step 30: conversion of coumaric acid to coumaryl-CoA. A heterologous 4CL is overexpressed.

Step 31: conversion of glutamate (Glu) to glutamyl-tRNA.

Step 32: conversion of glutamyl-tRNA to glutamate semialdehyde (GSA). hemA is overexpressed to upregulate ALA.

10 Step 33: conversion of glutamate semialdehyde (GSA) to δ amino levulinic acid (ALA). hemL is overexpressed to upregulate ALA.

Step 34: conversion of δ amino levulinic acid (ALA) to porphobilinogen (PBG).

Step 35: conversion of porphobilinogen (PBG) to hydroxymethylbilane (HMB).

Step 36: conversion of hydroxymethylbilane (HMB) to uroporphyrinogen III (UPP III).

15 Step 37: conversion of uroporphyrinogen III (UPP III) to coproporphyrinogen III (CPP III).

Step 38: conversion of coproporphyrinogen III (CPP III) to protoporphyrinogen IX (PP IX).

20 Step 39: conversion of protoporphyrinogen IX (PP IX) to protoporphyrin IX, which is subsequently covered to heme.

Step 40: conversion of prephenate (PPA) to phenylpyruvate (POPP).

Step 41: conversion of phenylalanine (Phe) to cinnamate. Heterologous PAL and/or TAL are overexpressed.

25 Step 42: conversion of cinnamate to coumaric acid. Heterologous C4H/CPR are overexpressed.

Table 11: Enzyme Sequences:

Enzyme:	Sequence:	SEQ ID:
Tyrosine ammonia-lyase (TAL) <i>Saccharothrix espanaensis</i> Accession: ABC88669.1	MTQVVERQADRLSSREYLARVVRSAWDAGLTSCDT EEIVRMGASARTIEEYLKSDKPIYGLTQGFGPLVLFDA DSELEQGGSLISHLGTGQGAPLAPEVSRLILWLRIQNM RKGYSASVPVFWQKLADLWNKGFTPAIPRHGTVSAS GDLQPLAHAALAFTGVGEAWTRDADGRWSTVPAVD ALAALGAEPFDWPVREALAFVNGTGASLAVAVLNHR SALRLVRACAVLSARLATLLGANPEHYDVGHGVARG QVGQLTAAEWIRQGLPRGMVRDGSRPLQEPYSLRCA PQVLGAVLDQLDGAGDVLAREVDGCQDNPIYEGEL LHGGNFHAMPVGFASDQIGLAMHMAAYLAERQLGL LVSPVTNGDLPPMLTPRARGAGLAGVQISATSFVSRI RQLVFPASLTTLPTNGWNQDHVPMALNGANSVFEAL ELGWLTVGSLAVGVAQLAAMTGHAAGVWAELAGI CPPLDADRPLGAEVRAARDLLSAHADQLLVDEADGK DFG	1
Phenylalanine ammonia-lyase (PAL) <i>Brevibacillus laterosporus</i> LMG 15441 Accession: WP_003337219.1	MSQVALFEQELMLHGKHTLLNGNDLTITDVAQMAK GTFEAFTFHISEEANKRIEENELKHEIMNQHNPIYGV TTGFGDSVHRQISGEKAWDLQRNLIRFLSCGVGPVAD EAVARATMLIRTNCLVKGNSAVRLEVIHQLIAYMERG ITPIPERGSVGASGDLVPLSYLASILVGEKVLKYGEE REVAEALGAEGLEPLTLEAKEGLALVNGTSFMSAFAC LAYADAEIIFIADICTAMASEALLGNRGHFYSFIHEQ KPHLGQMASAKNIYTLLEGSQLSKEYSQIVGNNEKLD SKAYLELTQSIQDRYSIRCAPHVTGVLYDTLDWVKK WLEVEINSTNDNPIFDVETRDVYNGGNFYGGHVVA MDSLKVAVANIADLLDRQLQLVVDEKFNKDLTPNLIP RFNNDNYEIGLHHGFKGMQIASSALTAALKMSGPVS	2

	VFSRSTEAHNQDKVSMGTISSRDARTIVELTQHVAIIH LIALCQALDLRDSKKMSPQTTKIYNMIRKQVPFVERD RALDGDIEKV VQLIRSGNLKKEIHDQNVND	
Cinnamate-4- hydroxylase (C4H) <i>Helianthus annuus</i> <i>L.</i> Accession: QJC72299.1	MDLLLIEKTLLALFAAII GAIVISKLRGKRFLPPGPLP VPIFGNWLQVGDDL NHRNLTDLAKKFGEIFLLRMGQ RNLVVVSSPD LAKEVLHTQGVEFGSRTRNVVFDIFTG KGQDMVFTVYGEHWRKMRRIMTVPFFTNKVVQQYR YGWEAEAAA VVEDVKKNPAAATEGVVIRRRRLQLMM YNNMFRIMFDRRFESDDPLFVKLKALNGERSRLAQS FEYNYGDFIPILRP FLKGYLKLCKEVKEKRFQLFKDYFVDERKKLESTKSV DNNQLKCAIDHILDAKEKGEINEDNVLYIVENINVAAI ETTLWSIEWGIAELVNHPEIQAKLRNELDTKLGPVQ VTEPDLHKLPYLQAVIKETLRLRMAIPLL VPHMNLHD AKLGGYDIPAESKILVNAWWLANNPEQWKKPEEFRP ERFFEEESKVEANGNDFRYLPFGVGRRSCPGIILALPIL GITIGRLVQNFELLPPPQSKVDTTEKGGQFSLHILKHS TIVAKPRAL	3
4-coumarate-CoA ligase (4CL) <i>Petroselinum</i> <i>crispum</i> Accession: P14912.1	MGDCVAPKEDLIFRSKL PDIYIPKHLPLHTYCFENISKV GDKSCLINGATGETFTYSQVELLSRKVASGLNKLGIQ QGDTIMLLL P NSPEYFFAFLGASYRGAISTMANPFFTS AEVIKQLKASQAKLIITQACYVDKVKDYAAEKNIQIC IDDAPQDCLHFSKLMEADESEMPEVVINSDDVVALPY SSGTTGLPKGVMLTHKGLVTSVAQQVDGDNPNLYM HSEDVMICILPLFHIYSLNAVLCCLRAGVTILIMQKF DIVPFLELIQKYKVTIGPFVPIVLAIKSPVVDKYDLS SVRTVMSGAAPLGKELEDAVRAKFPNAKLGQGYGM TEAGPVLAMCLAFAKEPYEIKSGACGTVVRNAEMKIV DPETNASLPRNQERGEICIRGDQIMKGYLNDPESTRTTI DEEGWLHTGDIGFIDDDDELFI VDRLEIKYKGFQVA	4

	PAELEALLLTHPTISDAAVVPMIDEKAGEVPVAFVVRT NGFTTTEEEIKQFVSKQVVFYKRIFRVFFVDAIPKSPSG KILRKDLRARIASGDLPK	
Chalcone synthase (CHS) <i>Petunia x hybrida</i> Accession: AAF60297.1	MVTVEEYRKAQRAEGPATVMAIGTATPTNCVDQSTY PDYYFRITNSEHKTDLKEKFKRMCEKSMIKKRYMHLT EEILKENPSMCEYMAPSLDARQDIVVVEVPKLGKEAA QKAIKEWGQPKSKITHLVFCTTSGVDMPGCDYQLTKL LGLRPSVKRLMMYQQGCFAGGTVLRLAKDLAENNK GARVLVVCSEITAVTFRGPNDTHLDSL VGQALFGDGA GAIIGSDPIPGVERPLFELVSAAQTLLPDSHG AIDGHL REVGLTFHLLKDV PGLISK NIEKSLEEAFRPLSISD WNS LFWIAHPGGPAILDQVEIKLGLKPEK LKATRNVLSNY GNMSSACVLFILDEM RKASAKEGLGTTGEGLEWGV L FGFGPGLTVETVVLH SVAT	5
Chalcone isomerase (CHI) <i>Medicago sativa</i> Accession: P28012.1	MAASITAITVENLEYPAVVTSPTVGKSYFLGGAGERG LTIEGNFIKFTAIGVYLEDIAVASLAAKWKGKSSEELL ETLDFYRDIISGPFEKLIRGSKIRELSGPEYSRKVMENC VAHLKSVGTYGDAEAEAMQKFAEAFKPVNFPPGASV FYRQSPDGILGLSFSPDTSIPEKEAALIENKAVSSAVLE TMIGEHA VSPDLKRCLAARLPALLNEGAFKIGN	6
Flavanone 3- hydroxylase (F3H) <i>Rubus occidentalis</i> Accession: ACM17897.1	MAPTP TTLTAIAGEKTLQQSFVRDEDERPKVAYNQFS NEIPIISLSGIDEVEGRRAEICNKIVEACEDWGVFQIVD HGVD AKLISEMTRLARDFFALPPEEKLRFDMSGGKKG GFIVSSHLQGEAVQDWREIVTYFSYPVRHRDYSR WPD KPEGWRAVTQQYSDELMGLACKLLEVLSEAMGLEKE ALTKACVDMDQKV VVNFY PKCPQPD LTLGLKRHTDP GTITLLLQDQVGGLQATRDGGKTWITVQPVEGAFVV NLGDHGHFLSNGRFKNADHQAVVNSNHSRLSIATFQ	7

	NPAQEAIIVYPLKVREGEKPILEEPITYTEMYKKKMSK DLELARLKKLAKKEQQPEDSEKAKLEVQVDDIFA	
Flavonoid 3' hydroxylase (F3'H) <i>Brassica napus</i> Accession: ABC58723.1	MTNLYLTILLPTFIFLIVLVLSRRRNRLPPGPNPWPIIG NLPHMGPKPHQTLAAMVTTYGPILHLRLGFADVVA ASKSVAEQFLKVHDANFASRPPNSGAKHMAYNYQDL VFAPYGQRWRMLRKISSVHLFSAKALEDKFKHVRQEE VGTLMLRELARANTKPVNLGQLVNMCVLNALGREMI GRRLFGADADHKAEFEFRSMVTEMMALAGVFNIGDFV PALDCLDLQGVAGKMKRLHKRFD AFLSSILEEHEAM KNGQDQKHTDMLSTLISLKGTDGEGGTLTDTEIKA LLNMFTAGTDTASTVDWAIAELIRHPEIMRKAQEE LDSVVGRGRPINESDLSQLPYLQAVIKENFRLHPPTPLS LPHIASSECEINGYHIPKGSTLLTNIWAIARPDQWSDP LTFRPERFLPGGEKAGVDVKGNDFELIPFGAGRRICAG LSLGLRTIQLLTATLVHGFELWELAGGVTPPEKLNMEET YGITLQRAVPLVVHPKRLRLDMSAYGLGSA	8
Cytochrome P450 reductase (CPR) <i>Catharanthus roseus</i> Accession: Q05001	MDSSSEKLSPFELMSAILKGAKLDGSNSSDSGVAVSPA VMAMLLENKELVMILTTSAVLIGCVVVLWRRSSGS GKKVVEPPKLIVPKSVVEPEEIDEGKKKFTIFFGTQTGT AEGFAKALAEAKARYEKAVIKVIDIDDYAADDEEYE EKFRKETLAFFILATYGDGEPTDNAARFYKWFVEGND RGDWLKNLQYGVFGLGNRQYEHFNKIAKVVDEKVA EQGGKRIVPLVLGDDDDQCIEDDFAAWRENVWPELDN LLRDEDDTTVSTTYTAAIPEYRVVFPDKSDSLISEANG HANGYANGNTVYDAQHPCRSNVAVRKELHTPASDRS CTHLDFDIAGTGLSYGTGDHVGVCNLSLSEVVEAER LLNLPPETYFSLHADKEDGTPLAGSSLPPFPCTLRTA LTRYADLLNTPKKSALLALAAAYASDPNEADRLKYLAS PAGKDEYAQSLVANQRSLLLEVMAEFPSAKPPLGVFFA AIAPRLQPRFYSSSSPRMAPSRIHVTCALVYEKTPGGR	9

	IHKGVCSTWMKNAIPLEESRDCSWAPIFVRQSNFKLP ADPKVPVIMIGPGTGLAPFRGFLQERLALKEEGAELGT AVFFFGCRNRKMDYIYEDELNHFLEIGALSELLVAFSR EGPTKQYVQHKMAEKASDIWRMISDGAYVYVCGDA KGMARDVHRTLHTIAQEQGSMDSTQAEGFVKNLQM TGRYL RDVW	
Flavonoid 3', 5'- hydroxylase (F3'5'H) <i>Delphinium grandiflorum</i> Accession: BAO66642	MSTSLLAAAAAILFFITHLFLRFLSPRRTRKLPPGPKG WPVVGALPMLGNMPHAALADLSRRYGPIVYLKLGSR GMVVASTPDSARAFLKTQDLNFSNRPTDAGATHIAYN SQDMVFADYGPRWKLLRKLSSLHMLGGKAVEDWAV VRRDEVGVMVKAIYESSCAGEAVHVPDMLVFAMAN MLGQVILSRRVFVTKGVESNEFKEMVIELMTSAGLFN VGDFIPSIWMDLQGIVRGMKRLHKKFDALDKILRE HTATRRERKEKPDLDVLMNDRDNKSEQRILTDTNI KALLNLFSAGTDTSSSTIEWALTEMIKNPSIFGRAHA EMDQVIGRNRRLSESDIPKLPYLQAICKETFRKHPSTP LNLPRVAIEPCEVEGYHIPKGTRLSVNIWAIGRDPNVW ENPLEFNPDRLTGKMAKIDPRGNNFELIPFGAGRRIC AGTRMGIVLVEYILGSLVHAFEWKLRDGETLNMEETF GIALQKAVPLAAVVTPLPPSAYVV	10
Dihydroflavonol 4- reductase (DFR) <i>Anthurium andraeanum</i> Accession: AAP20866.1	MMHKGTVCVTGAAGFVGSWLIMRLLEQGYSVKATV RDPSNMKKVKHLLDLPGAANRLTLWKADLVDEGSFD EPIQGCTGVFHVATPMDFESKDPSEMIKPTIEGMLNV LRSCARASSTVRRVFTSSAGTVSIHEGRRHLYDETS WSDVDFCRAKKMTGWMYFVSKTLAEKAAWDFAEK NNIDFISIPTLVNGPFVMPTMPPSMLSALALITRNEPH YSILNPVQFVHLDDLCNAHIFLFECPPDAKGRYICSSHD VTIAGLAQILRQRYPEFDVPTEFGEMEVFDIISYSSKKL TDLGFEFKYSLEDMFDGAIQSCREKGLLPATKEPSYA TEQLIATGQDNH	11

Leucoanthocyanidin reductase (LAR) <i>Desmodium uncinatum</i> Accession: Q84V83.1	MTVSGAIPSMTKNRTL VVG GTGFIGQFITKASLGFGYP TFLVLRPGPVSPSKAVIIKTFQDKGAKVIYGVINDKEC MEKILKEYEIDVVISLVGGARLLDQLTLLEAIKSVKTIK RFLPSEFGHDVDRTDPVEPGLTMYKEKRLVRRAVEEY GIPFTNICCNISASWPYYDNCHPSQVPPPMQDQFIYGD GNTKAYFIDGNDIGKFTMKTIDDIRTLNKNVHFRPSSN CYSINELASLWEKKIGRTLPRFTVTADKLLAHAAENII PESIVSSFTHDIFINGCQVNFSDIEHSDVEIDTLYPDEKF RSLDDCYEDFVPMVHDKIHAGKSGEIKIKDGKPLVQT GTIEEINKDIKTLVETQPNEEIKKDMKALVEAVPISAM G	12
Anthocyanin dioxygenase (ANS) <i>Carica papaya</i> Accession: XP_021901846.1	MFSSVAVPRVEILASSGIESIPKEYVRPQEELTTIGNIFD EEKKDEGPQVPTIDLRDIDSDDQQVRQRCRDELKAA VDWGMHLVNHGIPDHLIDRVKKAGQAFFELPVEVK EKYANDQASGNIQGYGSKLANNASGQLEWEDYYFHL IFPEEKRD LAIWPNNPADYIEVTSEYARQLRRLVSKIL GVL SLGLGLEGRLEKEVGGLDELLQMKINYYPTCP QPELALGVEAHTDISALTFILHNMVPGQLFYEGKWV TAKCVPSIVMHVGD TIEILSNGKYKSILHRGLVNKEK VRISWAVFCEPPKEKIILKPLPETVSENEPPLFPPRTFAQ HIQHKLFRKNQENLEAK	13
Anthocyanidin-3-O-glycotransferase (3GT) <i>Vitis labrusca</i>	MSQTTTNPHVAVLAFPFSTHAAPLLAVVRRRLAVAAPH AVFSFFSTSESNASIFHDSMHTMQCNIKS YDVS DG VPE GYVFTGRPQEGIDLFMRAAPESFRQGMVMAVAETGR PVSCLVADAFIWFAADMAAEMGVAWLPFWTAGPNS LSTHVYIDEIREKIGVSGIQGREDELLNFIPGMSKVRFR DLQEGIVFGNLNSLFSRLLHRMGQVLPKATAVFINSFE ELDDSLTNDLKS KLKTYLNIGPFNLITPPPVPNTTGCL QWLKERKPTSVVYISFGTVTTPPPAELVALAEALEASR	14

Accession: ABR24135	VPFIWSLRDKARMHLPEGFLEKTRGHGMVVPWAPQA EVLAHEAVGAFVTHCGWNSLWESVAGGVPLICRPF GDQRLNGRMVEDVLEIGVRIEGGVFTKSGLMSCFDQI LSQEKGGKLRNLRALRETADRAVGPKGSSSTENFKTL VDLVSKPKDV	
Acetyl-CoA carboxylase (ACC) <i>Mucor circinelloides 1006PhL</i> Accession: EPB82652.1	MVEHRSLPGHFLGGNSLESAPQGPVKDFVQAHEGHT VISKVLIANNMGMAAMKEIRSVRKWAYETFGNERAIEF TVMATPEDLKANA EYIRMADNFVEVPGGSSNNNNYAN VELIVDVAERTAVHAVWAGWGHASENPRLPEMLAKS KHKCLFIGPPASAMRSLGDKISSTIVAQSAQVPTMGW SGDGITETEFDAAGHVIVPDNAYNEACVKTAEQGLKA AEKIGFPVMIKASEGGGGKGIRMVKDGSNFAQLFAQV QGEIPGSPIFIMKLAGNARHLEVQLLADQYGN AISLFG RDCSVQRRHQKIIIEAPVTIAKPDVFEQMEKAAVRLG KLVGYVSAGTVEYLYSHHDDQFYFLELNPRLQVEHPT TEMVSGVNLPA AQLQIAMGIPLHRIRDIRVLYGVQPNS ASEIDFGFEHPTSLTSHRRPTPKGHVIACRITAENPDAG FKPSSGIMQELNFRSSTNVWGYFSVVSAGGLHEYADS QFGHIFAYGENRQQARKNMVIALKELSIRADFRSTVE YIIRLLETPDFEENTINTGWLDM LISKKLTAERPDTML AVFCGAVTKAHMASLDCFQQYKQSLEKGQVPSKGSL KTVFTVDFIYEEVRYNFTVTQSAPGIYTL YLNGTKTQV GIRDLSDGGLLISIDGKSHTTYSRDEVQATRMMVDGK TCLLEKESDPTQLRSPSPGKLVNLLVENGDHLNAGDA YAEIEVMKMYMPLIATEDGHVQFIKQAGATLEAGDII GILSLDDPSRVKHALPFNGTVPAFGAPHITGDKPVQRF NATKLT LQHILQGYDNQALVQTVVKDFADILNNPDL P YSELNSVL SALSGRIPQRLEASIHKL ADESKAANQEFP AAQFEKLVEDFAREHITLQSEATAYKNSVAPLSSIFAR YRNGLTEHAYS NYVELMEAYYDVEILFNQQREEEVIL	15

	SLRDQHKDDLDKVLAVTLSHAKVNIKNNVILMLLDLI NPVSTGSALDKYFTPIKRLSEIESRATQKVTLKAREL LILCQLPSYEERQAQMYQILKNSVTESVYGGGSEYRTP SYDAFKDLIDTKFNVFDVLPFFYHADPYIALAAIEVY CRRSYHAYKILDVAYNLEHKPYVVAWKFLQTAANG IDSNKRIASYSDLTFLNKTETEEPIRTGAMTACNSLAD LQAEPRILTAEEEEPLPPMLQRNAAPKEERMENILNI AVRADEDMDDTAFRTKICEMITANADVFRQAHLRRL SVVVCARDNQWPDYYTFRERENYQEDETIRHIEPAMA YQLELARLSNFDIKPCFIENRQMHVYYAVAKENPSCD RFFIRALVRPGRVKSSMRTADYLISESDRLLTDILDTLE IVSHEYKNSDCNHLFINFIPTFAIEADDVEHALKDFVD RHGKRLWKLRTGAEIRFNVQSKKPDAPIIPMRFTVD NVSGFILKVEVYQEVKTEKSGWILKSVNKIPGAMHM QPLSTPYPTKEWLQPRRYKAHLMGTTYVYDFPELFRQ SVQNQWTQAIKRNPLLKQPSHLVEAKELVLDEDDVL QEIDRAPGTNTVGMVAWIMTIRTPEYPSGRRIIAIANDI TFKIGSFGVAEDQVFYKASELARALGIPRIYLSANSGA RIGLADELISQFRAAWKDASNPTAGFKYLYLTPAEYD VLAQQGDAKSVLVEEIQDEGETRLRITDVIGHTDGLG VENLKGSGLIAGATSRAYDDIFTITLVTCSRSGIGAYL VRLGQRTIQNEGQPIILTGAPALNKVLGREVYTSNLQL GGTQIMYKNGVSHLTAENDLEGIKIVQWLSFVPDVR NAPVSMRLGADPIDRDIEYTPPKGPSDPRFFLAGKSEN GKWLSGFFDQDSFVETLSGWARTVVVGRARLGGIPM GVVSVETRRTVENIVPADPANSDESTQVFMEAGGVWFP NSAYKTAQAINDFNKGEQLPLMIFANWRGFGSGGQRD MYNEVLKYGAQIVDALSNYKQPVFVYIIPNGELRGGA WVVVDPTINKDMMEMYADNNARGGVLEPEGIVEIKY RKPALLATMERLDATYASLKKQLAEEGKTDEEKAAL KVQVEAREQELLPVYQQISIQFADLHDRAGRMKAKG	
--	---	--

	VIRKALDWRRARHYFYWRVRRRLCEEYTFRKIVTATS AAPMPREQMLDLVKQWFTNDNETVNFEDADELVSE WFEKRASVIDQRISKLKSDATKEQIVSLGNADQEAVIE GFSQLIENLSEDARAEILRKLNSRF	
Acetyl-CoA synthase (ACS) <i>Salmonella typhimurium</i> Accession: NP_463140.1	MSQTHKHAIPANIADRCLINPEQYETKYKQSINDPDTF WGEQGKILDWITPYQKVKNTSFAPGNVSIKWYEDGT LNLAANCLDRHLQENGDRTAIWEGDDTSQSKHISYR ELHRDVCRFANTLLDLGIKKGDVVAIYMMPMVPEAAV AMLACARIGAVHSVIFGGFSPEAVAGRIIDSSSRLVITA DEGVRAGRSIPLKKNVDDALKNPNTSVEHVIVLKRT GSDIDWQEGRDLLWRDLIEKASPEHQPEAMNAEDPL FILYTSGSTGKPKGVLHTTGGYLVYAATTFKYVFDYH PGDIYWCTADVGVWTGHSYLLYGPLACGATTLMFEG VPNWPTPARMCQVVDKHQVNILYTAPTALRALMAEG DKAIEGTDRSSRLILGSVGEPINPEAWEWYWKIGKE KCPVVDTWWTETGGFMITPLPGAIELKAGSATRPFF GVQPALVDNEGHPQEGATEGNLVITDSWPGQARTLF GDHERFEQTYFSTFKNMYFSGDGARRDEDGYYWITG RVDDVLNVSGHRLGTAEIESALVAHPKIAEAAVVGIP HAIKGQAIYAYVTNLHGEEPSPELYAEVRNWVRKEIG PLATPDVLHWTDSLPRKTRSGKIMRRILRKIAAGDTSNL GDTSTLADPGVVEKLLLEEKQAIAMPS	16
Malonyl-CoA synthase (matB) <i>Streptomyces coelicolor</i>	MSSLFPALSPAPTGAPADRPALRFGERSLTYAELAAA AGATAGRIGGAGRVAVWATPAMETGVAVVAALLAG VAAVPLNPKSGDKELAHILSDSAPSLVLAPPDAELPPA LGALERVDVDVRARGAVPEDGADDGDPALVVYTS GT TGPPKGAVIPRRALATTLDALADAWQWTGEDVLVQG LPLFHVHGLVLGILGLRRGGSVRHLGRFSTEGAAREL NDGATMLFGVPTMYHRIAETLPADPELAKALAGARL LVSGSAALPVHDHERIAAATGRRVIERYGMTETLMNT	17

Accession: WP_011028356	SVRADGEPRAAGTVGVPLPGVELRLVEEDGTPIAALDG ESVGEIQVRGPNLFTEYLNRPDATAAFAFTEDGFFRTG DMAVRDPDGYVRIVGRKATDLIKSGGYKIGAGEIENA LLEHPEVREAAVTGEPDPLGERIVAWIVPADPAAPP ALGTLADHVAARLAPHKRPRVVRYLDAVPRNDMGKI MKRALNRD	
Malonate transporter (matC) <i>Streptomyces coelicolor</i> Accession: NP_626686.1	MSPELISILVLVVVFVIATTRSVNMGALAFAAAFGVGT LVADLDADGIFAGFPGDLFVVLVGVTYLFAIARANGT TDWL VHAAVRLVRGRVALIPWVMFALTGALTAIGAV SPAAVAIVAPVALSFATRYSSISPLLMGTMVVHGAQAG GFSPISIIYGSIVNGIVEREKLPGEIGLFLASLVANLLIA AVLFAVLGGRKLWARGAVTPEGDGAPGKAGTGTGTS GSDTGTGTGTGTGTSAGTGGTAPTAVAVRSDRETGG AEGTGVRLTPARVATLVALVALVVAVLGFDLDAGLT AVTLAVVLSTAWPDDSRRAVGEIAWSTVLLICGVLT YVGVLEEMGTITWAGEGVGGIGVPLLA VLLCYIGAIV SAFASSVGIMGALIPLAVPFLAQGEIGAVGMVAALAV SATVVDVSPFSTNGALVLAAAPDVDRDRFFRQLMVY GGIVVAAPALAWLVLVVPFGF	18
Malonate CoA- transferase (MdcA) <i>Acinetobacter calcoaceticus</i> Accession: AAB97627.1	MVKKRLWDKQRRRQEKLNLAQQKGFAKQVEHARA IELLETVIASGDRVCLEGNNQKQADFLSKCLSQCNP AVNDLHIVQSVLALPSHIDVFEEKGIASKVDFSFAGPQS LRLAQLVQQKISIGSIHTYLELYGRYFIDLTPNICLITA HAADREGNLYTGPNTEDTPAIVEATAFKSGIVIAQVNE IVDKLPRVDVPADWVDFYIESPKHNYIEPLFTRDPAQI TEVQILMAMMVIKGIYAPYQVQRLNHGIGFDTAAIEL LLPTYAASLGLKGQICTN WALNPHPTLIPAIESGFVDS VHSFGSEVGMEDYIKERPDVFFTGSDGSMRSNRAFSQ TAGLYACDSFIGSTLQIELQGNSSTATVDRISGFGGAP NMGSDPHGRRHASAYTKAGREATDGKLIKGRKLVV	19

	QTVETYREHMHPVFVEELDAWQLQDKMDSELPPIMI YGEDVTHIVTEEGIANLLLCRTDEEREQAIRGVAGYTP VGLKRDAAKVEELRQRGIIQRPEDLGIDPTQVSRDLLA AKSVKDLVKWSGGLYSPPSRFRNW	
Pantothenate kinase (CoaX) <i>Pseudomonas aeruginosa</i> Accession: Q9HWC1.1	MILELDCGNSLIKWRVIEGAARSVAGGLAESDDALVE QLTSQQALPVRACRLVSVRSEQETSQLVARLEQLFPV SALVASSGKQLAGVRNGYLDYQRLGLDRWLALVAA HHLAKKACLVIDLGTAVTSDLVAADGVHLGGYICPG MTLMRSQRLRTHTRRIRYDDAEARRALASLQPGQATA EAVERGCLLMLRGFVREQYAMACELLGPDCEIFLTGG DAELVRDELAGARIMPDLVFVGLALACPIE	20
glutamyl-tRNA reductase (hemA ^m) <i>Salmonella typhimurium</i> Accession: AAA88610.1	MTKKLLALGINHKTAPVSLRERVTFSPDTLDQALDSL LAQPMVQGGVVLSTCNRTELYLSVEEQDNLQEALIR WLCDYHNLNEDDLRNSLYWHQDNDAVSHLMRVASG LDSLVLGEPQILGQVKKAFADSQKGHLNASALRRMF QKSFSVAKRVRTETDIGASAVSVAFAACTLARQIFESL STVTVLLVGAGETIELVARHLREHKVQKMIIANRTRE RAQALADEVGAEVISLSDIDARLQDADIIISSTASPLPII GKGMVERALKSRRNQPMMLVDIAVPRDVEPEVGKLA NAYLYSVDDLQSIISHNLAQRQAAAVEAETIVEQEASE FMAWLRAQGASETIREYRSQSEQIRDELTTKALSALQ QGGDAQAILQDLAWKLTNRLIHAPTKSLQQAARDGD DERLNILRDSLGLE	21
5-aminolevulinic acid synthase (ALAS)	MDYNLALDKAIQKLHDEGRYRTFIDIEREKGAFPKAQ WNRPDGGKQDITVWCGNDYLGMGQHPVLAAMHE ALEAVGAGSGGTRNISGTTAYHRRLEAEIADLHGKEA ALVFSSAYIANDATLSTLRLFPGLIYSDSLNHASMIE GIKRNAGPKRIFRHNDVAHLRELIAADDPAPKLIAFE SVYSMDGDFGPIKEICDIADEFGALTYIDEVHAVGMY	22

<i>Rhodobacter capsulatus</i> Accession: CAA37857	GPRGAGVAERDGLMHRIDIFNGTLAKAYGVFGGYIA ASAKMVDAVRSYAPGFIFSTSLPPAIAAGAQAASIAFLK TAEGQKLRDAQQMHAQVLKMLKALGMPIIDHGSHI VPVIGDPVHTKAVSDMLLSYGVYVQPINFPTVPRG TERLRFTSPVHDLKQIDGLVHAMDLLWARCA	
Tyrosine ammonia- lyase (TAL) <i>Rhodobacter capsulatus</i> SB 1003 Accession: ADE84832.1	MTLQSQTAKDCLALDGLTLVQCEAIATHRSRISVTP ALRERCARAHARLEHAIAEQRHIYGITTGFGPLANRLI GADQGAELQQNLIYHLATGVGPKLSWAEARALMLAR LNSILQGASGASPETIDRIVAVLNAGFAPEVPAQGTVG ASGDLTPLAHMVLALQGRGRMIDPSGRVQEAGAVM DRLCGGPLTLAARDGLALVNGTSAMTAIAALTGVEA ARIDAALRHSAVLMEVLSGHAEAWHPAFELRPHP GQLRATERLAQALDGAGRVCRTLTAARRLTAADLRP EDHPAQDAYSLRVVPQLVGAVWDTLDWHDRVVTCE LNSVTDNPIFPEGCAVPALHGGNFMGVHVALASDAL NAALVTLAGLVERQIARLTDEKLNKGLPAFLHGGQA GLQSGFMGAQVTATALLAEMRANATPVSQSLSTNG ANQDVVSMGTIAARRARAQLPLSQIQAILALALAQA MDLLDDPEGQAGWSLTARDLRDRIRAVSPGLRADRP LAGHIEAVAQGLRHPSAAADPPA	23
Tyrosine ammonia- lyase (TAL) <i>Trichosporon cutaneum</i> Accession:	MITETNVAKPASTKVMNGDAAKAAPVEPFATYAHSQ ATKTVVIDGHNMKVGDVVAVARHGAKVELAASVAG PVQASVDFKESKKHTSIYGVTTGFGGSADTRTSDTEA LQISLLEHQLCGYLPDPTYEGMLLAAMPIPIVRGAM AVRVNSCVRGHSGVRLEVLQSFADFINIGLVPCVPLR GTISASGDLSPLSYIAGAICGHPDVKVFDTAASPPTVLT APEAIAKYKLKTVRLASKEGLGLVNGTAVSAAAGAL ALYDAECLAMMSQTNTALTVEALDGHVGSFAPFIQEI RPHVGQIEAAKNIRHMLSNSKLAVHEEPELLADQDAG	24

XP_018276715	ILRQDRYALRTSAQWIGPQLEMLGLARQQIETELNSTT DNPLIDVEGGMFHHGGNFQAMAVTSAMDSTRIVLQN LGKLSFAQVTELINCEMNHGLPSNLAGSEPSTNYHCK GLDIHCGAYCAELGFLANPMSNHVQSTEMHNQSVNS MAFASARKTMEANEVLSLLLGSQMYCATQALDLRV MEVKFKMAIVKLLNDTLTKHFSTFLTPEQLAKLNTTA AITLYKRLNQTPSWDSAPRFEDAAKHLVGCIMDALM VNDDITDLTNLPKWKEFAKDAGDLYRSILTATTADG RNDLEPAEYLGQTRAVYEAIRSDLGVKVRRGDVAEG KSGKSIGSNVARIVEAMRDGRLMGAVSKMFF	
Tyrosine ammonia- lyase (TAL) <i>Flavobacterium johnsoniae</i> Accession: WP_012023194	MNTINEYLSLEEFEAIIIFGNQKVTISDVVVNRVNESFNF LKEFSGNKVIYGVNTGFGPMAQYRIKESDQIQLQYNLI RSHSSGTGKPLSPVCAKAAAILARLNTLSLGNSGVHPSV INLMSELINKDITPLIFEHGGVGASGDLVQLSHLALVLI GEGEVFYKGERRPTPEVFEIEGLKPIQVEIREGLALING TSVMTGIGVVNVYHAKKLLDWSLKSSCAINELVQAY DDHFSaelnqtkrhkgqQEIAlKMRQNLSDSTLIRKR EDHLYSGENTEEIFKEKVQEYYSLRCVPQILGPVLETI NNVASILEDEFNSANDNPIIDVKNQHVVYHGGNFHGDY ISLEMDKLKIVITKL TMLAERQLNYLLNSKINELLPPFV NLGTLGFNFGMQGVQFTATSTTAESQMLSNPMYVHSI PNNNDNQDIVSMGTNSAVITSKVIENAFEVLAIEMITIV QAIDYLGQKDKISSVSKKWYDEIRNIPTFKEDQVMYP FVQKVKDHLINN	25
Tyrosine ammonia- lyase (TAL)	MSTTLILTGEGLGIDDVVRVARHQDRVELTTPAILA QIEASCAYINQAVKEHQPVYGVTTGFGGMANVIISPEE AAELQNNAIWYHKTGAGKLLPFTDVRAAMLLRANSH MRGASGIRLEIIQRMVTFNLNANVTPHVREFGSIGASGD LVPLISITGALLGTDQAFMVDNFNGETLDCISALERLGL PRLRLQPKEGLAMMNGTSVMTGIAANCVHDARILLA	26

<i>Herpetosiphon</i> <i>aurantiacus</i> DSM 785 Accession: ABX04526.1	LALEAHALMIQGLQGTNQSFHPFIHRHKPHTGQVWA ADHMLELLQGSQLSRNELDGSHDYRDGDLIQDRYSL RCLPQFLGPIIDGMAFISHHLRVEINSANDNPLIDTASA ASYHGGNFLGQYIGVGMDQLRYYMGLMAKHLDVQI ALLVSPQFNNGLPASLVGNIQRKVNMGKGLQLTANS IMPILTFLGNSLADRFPHTAEQFNQNSQGFGSANLA RQTIQTLQQYIAITLMFGVQAVDLRTHKLAGHYNAAE LLSPLTAKIYHAVRSIVKHPPSPERPYYIWNDDQVLEA HISALAHDIANDGSLVSAVEQTLSGLRSIILFR	
Phenylalanine ammonia-lyase (PAL) <i>Physcomitrella</i> <i>patens</i> Accession: XP_001758374.1	MHDDNTSPYCIGQLGNGAVHGADPLNWAKTAKAME CSHLEEIKRMVDTYQNATQVMIEGATLTPVQVAAIAR RPEVHVVLDAANARSRVDESSNWVLDRIMGGGDIYG VTTGFGATSHRRTQQGVELQRELIRFLNAGVLSKGN LPSETARAAMLVRTNTLMQGYSGIRWEILHAMEKLL NAHVTPKLPLRGTTITASGDLVPLSYIAGLLTGRPN SKA VTEDGREVSALEALRIAGVEKPFELAPKEGLALVNGT AVGSALASTVCYDANIMVLLAEVLSALFCEVMQGP EFADPLTHKLKHHPGQMEAAVMEWVLDGSSFMKA AAKFNETDPLRKPKQDRYALRTSPQWLGPQVEVIRNA THAIEREINSVNDNPIIDAARGIALHGGNFQGTPIGVSM DNMRLSLAAIAKLMFAQFSELVNDYYNGLPSNLSG GPNPSLDYGMKGAEIAMASYLSEINYLANPVTTHVQS AEQHNQDVNSLGLVSARKTEEAMEILKLMSATFLVG LCQAIDLRHVEETMQSAVKQVVTQVAKKTLFMGSDG SLLPSRFCEKELLMVVDROPVFSYIDDSTSDSYPLMEK LRGVLVSRALKSADKETSNAVFRQIPVFEAEKLQLSR VPAVREAYDTKGLSLVPNRIQDCRTYPLYKLVRGDL KTQLLSGQRTVSPGQEIEKVFNAISAGQLVAPLLEC VQ GWTGTPGPFSARASC	27

Phenylalanine ammonia-lyase (PAL) <i>Dictyostelium discoideum AX4</i> Accession: XP_644510.1	MIETNHKDNFLIDGENKNLEINDIISISKGEKNIIFTNEL LEFLQKGRDQLENKLENVAIYGINTGFGGNGDLIIPF DKLDYHQSNLLDFLTCGTGDDFNDQYVRGIQFIIIALS RGWSGVRPMVIQTLAKHLNKGIPQVPMHGSSVGASG DLVPLSYIANVLCGKGMVKYNEKLMNASDALKITSIE PLVLKSKEGLALVNGTRVMSSVSCISINKFETIFKAAIG SIALAVEGLLASKDHYDMRIHNLKNHPGQILIAQILNK YFNTSDNNTKSSNITFNQSENVQKLDKSVQEVYSLRC APQILGIISENISNAKIVIKREILSVNDNPLIDPYYGDVL SGGNFMGNHIARIMDGIKLDISLVANHLHSLVALMMH SEFSKGLPNSLSPNPGIYQGYKGMQISQTSLVVWLRQE AAPACIHSLTTEQFNQDIVSLGLHSANGAASMLIKLCD IVSMTLIIAFQAISLRMKSIENFKLPNKVQKLYSSIIKIPI LENDRRTDIDVREITNAILQDKLDFINLNL	28
Phenylalanine ammonia-lyase (PAL) <i>Brevibacillus laterosporus LMG 15441</i> Accession: WP_003337219.1	MSQVALFEQELMLHGKHTLLNGNDLTITDVAQMAK GTFEAFTFHISEEANKRIEENELKHEIMNQHNPIYGV TTGFGDSVHRQISGEKAWDLQRNLIRFLSCGVGPVAD EAVARATMLIRTNCLVKGNSAVRLEVIHQLIAYMERG ITPIPERGSGVGASGDLVPLSYLASILVGEKVLKYKTEE REVAEALGAEGLEPLTLEAKEGLALVNGTSFMSAFAC LAYADAEIEAFIADICTAMASEALLGNRGHFYSFIHEQ KPHLGQMASAKNIYTLLEGSQLSKEYSQIVGNNEKLD SKAYLELTQSIQDRYSIRCAPHVTGVLYDTLDWVKK WLEVEINSTNDNPIFDVETRDVYNGGNFYGGHVQQA MDSLKVAVANIADLLDRQLQLVVDEKFNKDLTPNLIP RFNNDNYEIGLHHGFKGMQIASSALTAEALKMSGPVS VFSRSTEAHNQDKVSMGTISSRDARTIVELTQHVAIIH LIALCQALDLRDSKKMSPQTTKIYNMIRKQVPFVERD RALDGDIEKVVQLIRSGNLKKEIHDQNVND	29

Cinnamate-4-hydroxylase (C4H) <i>Rubus sp. SSL-2007</i> Accession: ABX74781.1	MDLLLMEKTLGLFVAVVVAITVSKLRGKKFKLPPGP IPVPVFGNWLQVGDDLNRNLTEMAKKFGEVFMLR MGQRNLVVVSSPDLAKEVLHTQGVEFGSRTRNVVFDI FTGKGQDMVFTVYGEHWRKMRRIMTVPFFTNNKVVQ QYRYGWESEAAAVVEDVKKHPEAATNGMVLRRRLQ LMMYNNMYRIMFDRRFESEDDPLFVKLKGNGERSR LAQSFEYNYGDFIPVLRPFLRGYKICKEVKEKRIQLF KDYFVDERKKLSSTQATTNEGLKCAIDHILDAQQKGE INEDNVLYIVENINVAAIETTLWSIEWGIAELVNHPEIQ KKLRDELDTVLGRGVQITEPEIQKLPYLQAVVKETLR LRMAIPLLPHMNLHDAKLGGFDIPAESKILVNAWWL ANNPAHWKKPEEFRPERFLEESKVEANGNDFRYLPF GVGRRSCPGIILALPILGITLGRLVQNFELLPPPGQTQL DTTEKGGQFSLHILKHSPIVMKPRT	30
Cinnamate-4-hydroxylase (C4H) <i>Fragaria vesca</i> Accession: XP_004294725.1	MDLLLLEKTLIGLFIAIVVAIIVSKLRGKKFKLPPGPIPV PVFGNWLQVGDDLNRNLTDMAKKFGDVFMRLMG QRNLVVVSSPDLAKEVLHTQGVEFGSRTRNVVFDIFT GKGQDMVFTVYGEHWRKMRRIMTVPFFTNNKVVQQY RHGWEAEAAAVVEDVKKHPEAATSGMVLRRRLQLM MYNNMYRIMFDRRFESEEDPLFVKLKGNGERSRLA QSFEYNYGDFIPVLRPFLRGYKICKEVKEKRIQLFKD YFVDERKKLASTQVTTNEGLKCAIDHILDAQQKGEIN EDNVLYIVENINVAAIETTLWSIEWGIAELVNHPEIQK KLRDELDTVLGHGVQVTEPELHKLPYLQAVVKETLR LRMAIPLLPHMNLHDAKLGGFDIPAESKILVNAWWL ANNPAHWKKPEEFRPERFLEESKVEANGNDFRYLPF GVGRRSCPGIILALPILGVTLGRLVQNFEMLPPPGQTQ LDTTEKGGQFSLHILKHSTIVMKPRA	31
Cinnamate-4-hydroxylase (C4H)	MDLLLLEKTLIGLFFAILIAIIVSKLRSKRFLKPPGPIPV VFGNWLQVGDDLNRNLTEYAKKFGDVFLLRMGQR	32

<i>Solanum tuberosum</i> Accession: ABC69046.1	NLVVVSSPELAKEVLHTQGVEFGSRTRNVVFDIFTGK GQDMVFTVYGEHWRKMRRIMTVPFFTNTKVVQQYRG GWESEAAASVVEDVKKNPESATNGIVLRKRLQLMMYN NMFRI MFDRRFESEDDPLFVKLRALNGERSRLAQSFE YNYGDFIPILRPFLRGYLKICKEVKEKRLKLFKDYFVD ERKKLANTKSMDSNALKCAIDHILEAQQKGEINEDNV LYIVENINVAAIETTLWSIEWGIAELVNHPHIQKKLRD EIDTVLGPGMQVTEPDMPKLPYLQAVIKETLRLMAI PLLVPHMNLHDAKLAGYDIPAESKILVNAWWLANNP AHWKKEEFRRPERFFEEKHVEANGNDFRFLPFGVGR RSCPGIILALPILGITLGRVLQNFEMLPPPGQSKLDTSE KGGQFSLHILKHSTIVMKPRSF	
4-coumarate-CoA ligase (4CL) <i>Daucus carota</i> Accession: AIT52344.1	MGDCAAPKQEIIFRSKLPDIYIPKHLPLHSYCFENISKV SDRACLINGATGETFSYAQVELISRRVASGLNKLGIHQ GDTMMILLPNTPEYFFAFLGASYRGAVSTMANPFFTS PEVIKQLKASQAKLIITQACYVEKVKEYAAENNITVVC IDEAPRDCLHFTTLMEADEAEMPEVAIDSDDVVALPY SSGTTGLPKGVMLTHKGLVTSVAQRVDGENPNLYIHS EDVMICILPLFHIYSLNAVLCCLAGATILIMQKFDIV PFLELIQKYKVTIGPFVPPIVLAIKSPVVDNYDLSSVR TVMSGAAPLGKELEDAVRAKFPNAKLGQGYGMTEA GPVLAMCLAFAPKEPYEIKSGACGTVVRNAEMKIVDPE THASLPRNQSGEICIRGDQIMKGYLNDPESTKTTIDEE GWLHTGDIGFIDEDDELFIVDRLKEIKYKGFQVAPAEI EALLLTHPTISDAAVVPMIDEKAGEVPVAFVVRLNGS TTTEEEIKQFVSKQVVFYKRVFRVFFVDAIPKSPSGKIL RKELRARIASGDLPK	33
4-coumarate-CoA ligase (4CL)	MEPTTKSKDIIFRSKLPDIYIPKHLPLHTYCFENISRFGS RPCLINGSTGEILTYDQVELASRRVGSGLHRLGIRQGD TIMLLLPNSPEFVLAFLGASHIGAVSTMANPFFTPAEV	34

<i>Striga asiatica</i> Accession: GER48539.1	VKQAAASRAKLIVTQACHVDKVRDYAAEHGVKVVC VDGAPPEECLPFSEVASGDEAELPAVKISPDDVVALPY SSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYIHS DDVIMCVLPLFHIYSLNSIMLCGLRVGAAILIMQKFEIV PFLELIQRYRVITIGPFVPPIVLAIEKSPVVEKYDLSSVRT VMSGAAPLGRELEDAVRLKFPNAKLGGYGMTEAGP VLAMCLAFAKEPFEIKSGACGTVVRNAEMKIVDTETG ASLGRNQPGEICIRGDQIMKGYLNDPESTERTIDKEGW LHTGDIGFIDDDDELFIVDRLKEIKYKGFQVAPAELEA LLNHPNISDAAVVSMKDEQAGEVPVAYVVKSNGSTI TEDEIKQFVSKQVIFYKRINRVFFIDAIPKSPSGKILRKD LRARLAAGVPN	
4-coumarate-CoA ligase (4CL) <i>Capsicum annuum</i> Accession: KAF3620179.1	MPMENEAKQGDIFRSKLPDIYIPNHLSLHSYCFENISE FSSRPCLINGANNQIYTYADVELNSRKVAAGLHKQFGI QQKDTIMILLPNSPEFVFAFLGASYLGAISTMANPLFTP AEVVKQVKASNAEIIVTQACHVNKVKDYALENDVKI VCIDSAPEGCVHFSELIQADEHDIPEVQIKPDDVVALP YSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYI HSEDVMLCVLPLFHIYSLNSVLLCGLRVGAAILIMQKF DIVPFLELIQNYKVTIGPFVPPIVLAIKSPMVDNYDLS SVRTVMSGAAPLGKELEDTVRAKFPNAKLGGYGMT EAGPVLAMCLAFAKEPFEIKSGACGTVVRNAEMKIVD PDTGNSLHRNQSGEICIRGDQIMKGYLNDPEATAGTID KEGWLHTGDIGYIDNDDELFIVDRLKELIKYKGFQVA PAELEALLNHPNISDAAVVPMKDEQAGEVPVAFVVR SNGSTITEDEVKEFISKQVIFYKRIKRVFFVDAVPKSPS GKILRKDLRAKLAAGFPN	35
4-coumarate-CoA ligase (4CL)	MDTKTTQQEIIFRSKLPDIYIPKQLPLHSYCFENISQFSS KPCLINGSTGKVYTYSDVELTSRKVAAGFHNLIQQR DTIMLLLPCPEFVFAFLGASYLGAITMANPFFTPAET	36

<i>Camellia sinensis</i> Accession: ASU87409.1	IKQAKASNSKLIITQSSYTSKVLDYSSENN VKIICIDSP DGCLHFSELIQSNETQLPEVEIDSNEVVALPYSSGTTGL PKGVM LTHKGLVTSVAQQVDGENPNLYIHSEDMMM CVLP LFHIYSLNSVLLCGLRVGAAILIMQKFEIGSFLKL IQR YKVTIGPFVPPIVLAIKSEVVDDYDLSTIRTMMS GAAPLGKELEDAVR AKFPHAKLGQGYGMTEAGPVLA MCLAF AKKPFEIKSGACGTVVRNAEMKIVDPDAGFSL PRNQPGEICIRGDQIMKGYLNDPEATERTIDKQGWLH TGDIGYIDDDDEL FIVDRLKELIKYKGFQVAPAELEAL LLNHPTISDAAV VPMKDESAGEVPVAFVVRTNGFEVT ENEIKKYISEQVVFYKKINRVYFVDAIPKAPSGKILRK DLRARLAAGIPS	
Chalcone synthase (CHS) <i>Capsicum annuum</i> Accession: XP_016566084.1	MVTVEEYRKAQRAEGPATVMAIGTATPSNCVDQSTY PDYYFRITNSEHKTELKEKFKRMCEKSMIKTRYMHLT EEILKENPNMCAYMAPSLDARQDIVVVEVPKLGKEA AQKAIKEWGQPKSKITHLVFCTTSGVDMPGCDYQLA KLLGLRPSVKRLMMYQQGCFAGGTVLR LAKDLAEN NKGARVLVVCSEITAVTFRGPSESHLDSL VGQALFGD GAAAIIMGSDPIPGVERPLFQLVSAAQTLLPDSEGAID GHLREVGLTFHLLKDVPGLISKNIEKSLVEAFQPLGISD WNSLFWIAHPGGPAILDQVELKLGLKPEKLKATREVL SNYGNMSSACVLFILDEMRASTKEGLGTSGEGLEW GVLF GFGPGLTVETVVLHSVAI	37
Chalcone synthase (CHS) <i>Rosa chinensis</i>	MVTVEEVRKAQRAEGPATVLAIGTATPPNCIDQSTYP DYYFRITKSEHKAELEKEKFKRMCDKSMIKKRYMYLT EEILKENPSMCEYMAPSLDARQDMVVVEIPKLGKEAA TKAIKEWGQPKSKITHLVFCTTSGVDMPGADYQLTKL LGLRPSVKRLMMYQQGCFAGGTVLR LAKDLAENNK GARVLVVCSEITAVTFRGPSDTHLDSL VGQALFGDGA AAIIVGSDPLPEVEKPLFELVSAAQTILPDSDGAIDGHL	38

Accession: AEC13058.1	REVGLTFHLLKDVPGLISKNIEKSLNEAFKPLNITDWN SLFWIAHPGGPAILDQVEAKLGLKPEKLEATRHILSEY GNMSSACVLFILDEVRRKSAANGHKTGEGLEWGVLF FGFGPGLTVETVVLHVSAA	
Chalcone synthase (CHS) <i>Morus alba</i> var. <i>multicaulis</i> Accession: AHL83549.1	MSMTPSVHEIRKAQRSEGPATVLSIGTATPTNFVPQAD YPDYYFRITNSDHMTDLKDKFKRMCEKSMITKRHMY LTEEILKENPKMCEYMAPSLDARQDIVVVEVPKLGKE AAAKAIKEWGQPKSKITHLIFCTTSGVDMPGADYQLT KLLGLRPSVKRFMMYQQGCFAGGTVLRLAKDLAENN KGARVLVVCSEITAVTFRGPSHTHLDVLGQALFGDG AAAVILGADPDTSVERPifelVSAAQTILPDSEGAIDGH LREVGLTFHLLKDVPGLISKNIEKSLVEAFTPIGISDWN SIFWIAHPGGPAILDQVEAKLGLKQEKLSATRHVLSEY GNMSSACVLFILDEVRRKKSVEEGKATTGEGLEWGVLF FGFGPGLTVETIVLHSLPAV	39
Chalcone synthase (CHS) <i>Dendrobium</i> <i>catenatum</i> Accession: ALE71934.1	MAPPAMEEIRRAQRAEGPATVLAIGASTPPNALYQAD YPDYYFRITKSEHLTELKEKFKQMCDKSMIRKRYMYL TEEILKENPNICAFMAPSLDARQDIVVTEVPKLAKES ARAIKEWGQPKSRITHLIFCTTSGVDMPGADYQLTRL LGLRPSVNRIMLYQQGCFAGGTVLRLAKDLAENNAG ARVLVVCSEITAVTFRGPSHLDVLGQALFGDGAA AIIVGSDPDLTTERPLFQLVSASQTILPESEGAIDGHLRE MGLTFHLLKDVPGLISKNIQKSLVETFKPLGIHDWNSI FWIAHPGGPAILDQVEIKLGLKEEKLASSRNVLAEYG NMSSACVLFILDEMRRRSAGQATTGEGLEWGVLF FGFGPGLTVETVVLRSVPIAGAV	40

Chalcone isomerase (CHI) <i>Trifolium pratense</i> Accession: PNX83855.1	MSAITAIHVENIEFPAVITSPVTGKSYFLGGAGERGLTI EGNFIKFTAIGVYLEDVAVASLATKWKGKSSEELLET LDFYRDIISGPFEKLIRGSKIRELSGPEYSRKVTENCVA HLKSVGTYGDAEVEAMEKFVEAFKPINFPPGASVFYR QSPDGILGVSIHFFP	41
Chalcone isomerase (CHI) <i>Abrus precatorius</i> Accession: XP_027366189.1	MAAASLTAVQVENLEFPAVVTSPATGKTYFLGGAGV RGLTIEGNFIKFTGIGVYLEDQAVASLATKWKGKSSEE LVESLDFFRDIISGPFEKLIRGSKIRQLSGPEYSKKVME NCVAHMKSVGTYGDAEAAGIEEFAQAFKPVNFPPGA SVFYRQSPDGVLGLSFSQDATIPEEEAAVIKNKPVSA VLETMIGEHA VSPDLKRS LAARLPAVL SHGVFKIGN	42
Chalcone isomerase (CHI) <i>Arachis duranensis</i> Accession: XP_015942246.1	MAAEPSITAIQFENLVFPAVVTTPPGSSKSYFLAGAGER GLTIDGKFIKFTGIGVYLEDKAVPSLAGKWKDKSSQQ LLQTLHFYRDIISGPFEKLIRGSKILALSGVEYSRKVME NCVAHMKSVGTYGDAEAEAIQQFAEAFKNVNFKPGA SVFYRQSPLGHLGLSFSQDGNIPKEEAAVIENKPLSSA VLETMIGEHA VSPDLKCS LAARLPAVLQQGHI VTPPQH N	43
Chalcone isomerase (CHI) <i>Cephalotus follicularis</i> Accession: GAV77263.1	MGPSPSVTELQVENVTFPPSVKPPGSTKTLFLGGAGER GLEIQGKFIKFTAIGVYLEGDAVASLAVKWKGKSKEE LTDSVEFFRDI VTPFEKFTQVTTILPLTGQQYSEKVSE NCVAFWKSVGIYTDAAEAKAIEKFIEVFKEETFPPGSSIL FTQSPNGALTIAFSKDGVIPEVGKAVIENKLLAEGLE SIIGKHGVSPVAKQCLATRLSELL	44

Flavanone 3-hydroxylase (F3H) <i>Abrus precatorius</i> Accession: XP_027329642.1	MGSASETVCVTGAAGFIGSWLVMRLIQNGYKVRATV RDPANMKKVKHLLLELPNAKTNLSLWKADLAEEGSFD EAIKGCTGVFHVATPMDFESKDPENEVIKPTINGLIDI MKACMKAKTVRRLVFTSSAGTVDVTEHPKPLFDESC WSDVQFCRRVRMTGWMYFVSKTLAEQEA WKFAKEN NIDFISVIPPLVVGPFLLVPTMPPSLITALSLITGNESHYAI IKQGQFVHLDDLCLAHIFLFQHPKAQGRYICCSHEATI HDIASLLNQKYPEFNVPTKFKNIPDQLEIIRFSSKKITDL GFKFKYSLEDMFTGAVETCCEKRLSETAEISGTTQK	45
Flavanone 3-hydroxylase (F3H) <i>Camellia sinensis</i> Accession: AAT66505.1	MKDSVASATASAPGTVCVTGAAGFIGSWLVMRLLER GYIVRATVRDPANLKKVKHLLDLPKADTNLTLWKAD LNEEGSFDEAIEGCSGVFHVATPMDFESKDPENEVIKP TINGVLSIIRSCTKAKTVKRLVFTSSAGTVNVQEHQQP VFDENNWSDLHFINKKKMTGWMYFVSKTLAEKAAW EAAKENNIDFISIPTLVGGPFIMPTFPPSLITALSPITRN EGHYSIIKQGQFVHLDDLCESHIFLYERPQAEGRYICSS HDATIHDLAKLMREKWPEYNVPTEFKGIDKDLVVVSF SSKKLIGMGFEFKYSLEDMFRGAIDTCREKGLLPHSFA ENPVNGNKV	46
Flavanone 3-hydroxylase (F3H) <i>Nyssa sinensis</i> Accession: KAA8531902.1	MVDMKDDDDSPATVCVTGAAGFIGSWLIMRLLQGGYI VRATVRDPANMKKVKHLQELEKADKNLTLWKADLT EEGSFDEAIKGCSGVFHVATPMDFESKDPENEVIKPTI NGVLSIVRSCVKAKTVKRLVFTSSAGTVNLQEHQQLV YDENNWSDLDLIYAKKMTGWMYFVSKILAEKAAWE ATKENNIDFISIPTLVGGPFITPTFPPSLITALSLITGNEA HYSIIKQGQFVHLDDLCEAHIFLYEQPKAEGRYICSSH DATIIDLAKMIREKWPEYNVPTELKGIEKDLQTVSFSS KKLIGMGFEFKYSLEDMYKGAIIDTCREKGLLPYSTHE TPANANANANANVKKNQNTTEI	47

Flavanone 3-hydroxylase (F3H) <i>Rosa chinensis</i> Accession: XP_024167119.1	MASESESVCTGASGFVGSWLVMRLLDRGYTVRATV RDPANKKKVKHLLDLPKAATHLTLWKADLAEEGSFD EAIKGCTGVFHVATPMDFESKDPENEVIKPTINGVLDI MKACLKAKTVRRLVFTASAGSVNVEETQKPVYDESN WSDVEFCRRVKMTGWMYFASKTLAEQEA WKFAKEN NIDFITIPTLVIGPFLMPAMPPSLITGLSPLTGNESHYSII KQGQFIHLDDLCCQSHIYLYEHPKAEGRYICSSHDATIH EIAKLLREKYPEYNVPTTFKGIEENLPKVHFSSKKLLE TGFEFKYSLED MFVGAVDACKAKGLLPPTERVEKQE VDESSVVGKVTG	48
Flavonoid 3'-hydroxylase (F3'H) <i>Cephalotus follicularis</i> Accession: GAV84063.1	MSPLILYSIALAIFLYCLRTLLKRHPHRLPPGPRPWPIIG NLPHMGQMPHSLAAMARTYGPLMHLRLGFVDVIV AASASVASQLLKTHDANFSSRPHNSGAKYIAYNYQDL VFAPYGPRWRMLRKISSVHLFSGKALDDYRHVRQEE VAVLIRALARAESKQAVNLGQLLNVCTANALGRVML GRRVFGDGSVSDPMAEEFKSMVVEVMALAGVFNI DFIPALDWLDLQGVAAKMKNLHKRFDFTLTGLLEEH KKMLVGDGGSEKHKDLLSTLISLKDSADDEGLKLTDT EIKALLNMFTAGTDTSSSTVEWAIAELIRHPKILAQV LKELDTVVGRDRLVTDLDLPQLTYLQAVIKETFRLHP STPLSLPRVAAESCEIMGYHIPKGSTLLVNVWAIARDP KEWAEPLEFRPERFLPGGEKPNVDIKGNDFEVIPFGAG RRICAGMSLGLRMVQLLTATLVHAFDWDLTSGLMPE DLSMEEAYGLTLQRAEPLMVHPRPRLSPNVY	49
Flavonoid 3'-hydroxylase (F3'H) <i>Theobroma cacao</i>	MASFLLYSILSAVFLYFIFATLRKRHRLPLPPGPKPWPII GNLPHMGVPVPHSLAALAKVYGPLMHLRLGFVDVV VAASASVAAQFLKVHDANFSSRPPNSGAKYVAYNYQ DLVFAPYGPRWRMLRKISSVHLFSGKALDDFRHVRQ DEVGVLVRALADAKTKVNLGQLLNVCTVNALGRVM LGKRVFGDGSKADPEADEFKSMVVELMVLAGVVNI	50

Accession: EOY22049.1	GDFIPALEWLDLQGVQAKMKKLHKRFDRFLSAILEEH KIKARDGSGQHKDLLSTFISLEDADGEGGKLTDTTEIKA LLNMFTAGTDTSSSTVEWAIAELIRHPKILAQVRKEL DSVVGRDRLVSDLDLPNLTYFQAVIKETFRLHPSTPLS LPRMASESCEINGYHIPKGATLLVNVWAIARDPDEWK DPLEFRPERFLPGGERPNADV RGND FEVIPFGAGRRIC AGMSLGLRMVQLLAATLVHAFDWELADGLMPEKLN MEEAFGLTLQRAAPLMVHPRPRLSPRAY	
Flavonoid 3' hydroxylase (F3'H) <i>Gerbera hybrida</i> Accession: ABA64468.1	MTPLTLLIGTCVTGLFLYVLLNRCTRNP NRLPPGPTPW PVVGNLPHLGTIPHHSLAAMAKKYGPLMHLRLGFVD VVVAASASVAAQFLKTHDANFADRPPNSGAKHIAYN YQDLVFAPYGPRWRMLRKICSVHLFSTKALDDFRHV RQEEVAILARALVGAGKSPVKLGQLLN VCTTNALAR VMLGRRVFDSGDAQADEFKDMVVELMVLAGEFNIG DFIPVLDWLDLQGVTKKMKKLHAKFDSFLNTILEEHK TGAGDGVASGKVDLLSTLISLKDDADGEGGKLS DIEI KALLNLFTAGTDTSSSTIEWAIAELIRNPQLLNQARK EMDTIVGQDRLVTESDLGQLTFLQAIHKETFRLHPSTPL SLPRMALESCEVGGYYIPKGSTLLVNVW AISRDPKIW ADPLEFQPTRFLPGGEKPNTDIKGND FEVIPFGAGRRIC VGMSLGLRMVQLLTATLIHAFDWELADGLNP KKLNM EEAYGLTLQRAAPLVVHPRPRLAPHVYETTKV	51
Flavonoid 3' hydroxylase (F3'H) <i>Phoenix dactylifera</i>	MAPLLLLFFTL LLSYLLYYYFFSKERTKGSRAPLPPGP RGWPVLGNLPQLGPKPHHTLHALSRAHGPLFRLRLGS VDVVVAASA AVAAQFLRAHDANFSNRPPNSGAEHIA YNYQDLVFAPYGPGRARRKLLNVHLFSGKALEDLR PVREGELALLVRALRDRAGANELVDLGRAANKCATN ALARAMVGRRVFQEEDEKAAEFENMVVELMRLAG VFNVGDFVPGIGWLDLQGVVRRMKELHRRYDGF LDG LIAAHRRAAEGGGGGGKDLLSVLLGLKDEDLDFDGE	52

Accession: XP_008791304.2	GAKLTDTDIKALLLNLF TAGTDTTSSTVEWALSELVK HPDILRKAQLELDSVVG DRLVSESDLPNLPFMQAIIK ETFRLHPSTPLSLPRMAAEECEVAGYCIPKGATLLVNV WAIARDPAVWRDPLEFRPARFLPDGGCEGMDVKGND FGIIPFGAGRRICAGMSLGIRMVQFMTATLAHAFHWD LPEGQMPEKLDMEAYGLTLQRATPLMVHPVPR LAP TAYQS	
Cytochrome P450 reductase (CPR) <i>Camellia sinensis</i> Accession: XP_028084858	MASNSNLIRAIESALGV SFGSELVSDTAIVVTTSVAVI IGLLFFLLKRSSDRSKESKPVVISKPLLVEEEEEDEVE AGSGKTKVTMFYGTQTGTAE GFAKSLAKEIKARYEK AIVKVVDLDDYAADDDQYE QKLKKETLVFFMLATYG DGEPTDDAARFYKWFTEENERGAWLQQLTYGVFSLG NRQYEHFNKIGKVVD EQLSKQGA KRLIPVGLGDDDQ CIEDDFAAWRETLWPELDQ LLRDEDDANTVSTPYAA AIPEYRVVIHDPLSGRGEAP SFIDSHTICEIWSTSREG SNQQISEYFWTSNSLKTMA SNSNLIRSIESALGV SFGSE SVSDTAIVVTTSVAVIIG LLFFLLKRSSDRSKESKPVV ISKPLLVEEEEEDEVEAGSGKTKVTLFYGTQTGTAE GFA KSLAEEIKARYEK AIVKVVDLDDYAADDDQYE QKLK KETLVFFMLATYGDGEPTD NAARFYKWFTEENERGA WLQQLTYGVFSLGNRQYEHFNKIGKVVD EQLSKQGA KRLIPVGLGDDDQCIEDDFAAWRETLWPELDQ LLRDE DDANTVSTPYTAAIPEYRVVIHDPTTTSYEDKNLNMA NGNASYDIHHPCR VNVAVQRELHKPESDRSCIHFEDI SGTGIIYETGDHVG VYADNFDEVVEEAANLLGQPLEL LFSVHADKDDGTS LGGSLPPFP GPCTLRDALAHYAD LLNPPRKAALSALAAH AVEPSEAERLKFLSSPQ GKED YSQWVVASQRSLL EIMAEFPSAKPPLGVFFAAVAPRL QPRYYSISSSPRFV PNRVHVTCALVYGPSPTGRIHKG V CSTWMKNAVPLEKSHDCSSAPIFTRTSNFKLPTDPSIPI	53

	IMVGPGTGLAPFRGFLQERLALKEDGVQLGHAMLFFG CRNRRMDFIYEDELNNFVDQGAVSELVVAFSREGPEK EYVQHKLNAKAAQVWGLISQGGYLYVCGDAKGMAR DVHRMLHTIVEQQENVDSRKAEVIVKKLQMEGRYLR DVW	
Cytochrome P450 reductase (CPR) <i>Cephalotus follicularis</i> Accession: GAV59576.1	MASNSNLIRAIESALGVSGSELVSDTAIVVTTTSVAVI IGLLFFLLKRSSDRSKESKPVVISKPLLVEEEEEDEVE AGSGKTKVTMFYGTQTGTAEFAKSLAKEIKARYEK AIVKVVDLDDYAADDDQYEQKLKKETLVFFMLATYG DGEPTDDAARFYKWFTEENERGAWLQQLTYGVFSLG NRQYEHFNKIGKVVDLQSKQGA KRLIPVGLGDDDQ CIEDDFAAWRETLWPELDQLLRDEDDANTVSTPYAA AIPEYRVVIHDPLSGRGEAPSFIDSHTICEIWSTSREG SNQQISEYFWTSNSLKTMASNSNLIRSIESALGVSGSE SVSDTAIVVTTTSVAVIIGLLFFLLKRSSDRSKESKPVV ISKPLLVEEEEEDEVEAGSGKTKVTLFYGTQTGTAEFA KSLAEEIKARYEKAIVKVVDLDDYAADDDQYEQKLK KETLVFFMLATYGDGEPTDNAARFYKWFTEENERGA WLQQLTYGVFSLGNRQYEHFNKIGKVVDLQSKQGA KRLIPVGLGDDDQCIEDDFAAWRETLWPELDQLLRDE DDANTVSTPYAAIPEYRVVIHDPTTTSYEDKNLNMA NGNASYDIHHPCRNVNAVQRELHKPESDRSCIHFEDI SGTGIIYETGDHVG VYADNFDEVVEEAANLLGQPLEL LFSVHADKDDGTSLGGSLPPFPGPCTLRDALAHYAD LLNPPRKAALSALAAH AVEPSEAERLKLFLSSPQKED YSQWVVASQRSLEIMAEFPSAKPPLGVFFAAVAPRL QPRYYSISSSPRFV PNRVHVTCALVYGPSPTGRIHKG V CSTWMKNAVPLEKSHDCSSAPIFTRTSNFKLPTDPSIPI IMVGPGTGLAPFRGFLQERLALKEDGVQLGHAMLFFG CRNRRMDFIYEDELNNFVDQGAVSELVVAFSREGPEK	54

	EYVQHKLNAAQVWGLISQGGYL YVCGDAKGMAR DVHRMLHTIVEQQENVDSRKA EVIVKKLQMEGRYLR DVW	
Cytochrome P450 reductase (CPR) <i>Brassica napus</i> Accession: XP_013706600.1	MSSSSSPFDLMSAIIKGEPVVVSDPANASAYESVAAE LSSMLIENRQFAMIISTSI AVLIGCIVMLLWRRSGGSGS SKRAETLKPLVLKPPREDEVDDGRKKVTIFFGTQTGT AEGFAKALGEEARARYEKTRFKIVDLDDYAADDDEY EEKLKKEDVAFFFLATYGDGEPTDNAARFYKWFTEG DDRGEWLKNLKYGVFGLGNRQYEHFNK VAKVVDDI LVEQGAQRLVHVGLGDDDQCIEDDFTAWREALWP EL DTILREEGDTAVTPYTA AVLEYRVSIHNSADALNEKN LANGNGHAVFDAQHPYRANVAVRRELHTPESDRSCT HLEFDIAGSGLTYETGDHVGVLSDNLNETVEEALRLL DMSPDTYFSLHSDKEDGTPISSSLPPTFPFPCSLRTALTR YACLLSPPKKSALLALAAHASDPTEAERLKH LASPAG KDEYSKWVVESQRSLLLEVMAEFPSAKPPLGVFFAAV APRLQPRFYSISSSPKIAETRIHVTCALVYEKMP TGRIH KGV CSTWMKSAVPYEKSENCCSAPIFVRQSNFKLPSD SKVPIIMIGPGTGLAPFRGFLQERLALVESGVELGPSVL FFGCRNRRMDFIYEEELQRFLESGALSEL SVAFSREGP TKEYVQHKMMDKASDIWNMISQGAYVYVCGDAKG MARDVHRS LHTIAQEQGSMDSTKAESFVKNLQMSG R YLRDVW	55
Flavonoid 3', 5'- hydroxylase (F3'5'H)	MALDTFLLRELA AAAVLFLISHYLIHSL LKKSTPPLPPG PKGWPFGALPLLGTMPHVALAQMAKKYGPV MYLK MGTCGMVVASTPDAARAFLKTLDLNFSNRPPNAGAT HLAYNAQDMVFADYGPRWKLLRKLSNLHMLGGKAL EDWTQVRTVELGHMIQAMCEASRAKEPVV VPEMLTY AMANMIGKVILGHRVFVTQGSSESNEFKDMVVELMTS	56

<i>Cephalotus follicularis</i> Accession: GAV62131	AGYFNIGDFIPSIWMDLQGIERGMMKKLHKRFDALLT KMFEEHMATAHERKGNPDLLDIVMANRDNSEGERLT TTNIKALLLNLFSAAGTDTSSSIIEWSLAEMLNPSILKR AHEEMDQVIGRNRRLSESDIKKLPYLQAICKESFRKHP STPLNLPRVSSQACQVNGYYIPKDTRLSVNIWAIGRDP EVWENPLDFTPERFLSGKNAKIDPRGNDFELIPFGAGR RICAGTRMGIVLVEYILGTLVHSFDWSLPHGVKLNMD EAFGLALQKAVPLAAIVSPRLAPTAYVV	
Flavonoid 3', 5'- hydroxylase (F3'5'H) <i>Dendrobium moniliforme</i> Accession: AEB96145	MSIFLITSLLLCLSLHLLRRRHISRLPLPPGPPNLPPIGA LPFIGPMPHSGLALLARRYGPIMFLKMGIRRVVVASSS TAARTFLKTFDSHFSRPSGVISKEISYNGQNMVFADY GPKWKLLRKVSSLHLLGSKAMSRWAGVRRDEALSMI QFLKKHSDSEKPVLLPNLLVCAMANVIGRIAMSKRVF HEDGEEAKEFKEMIKELLVGQGASNMEDLVPAIGWL DPMGVRKKMLGLNRRFDRMVSKLLVEHAETAGERQ GNPDLDDLVAASEVKGEDGEGLCEDNIKGFISDLFVA GTDTSIVIEWAMAEMLNPSILRRRAQEETDRVIGRH RLLDESDIPNLPYLQAICKEALRKHPPTPLSIPHYASEP CEVEGYHIPGETWLLVNIWAIGRDPDVWENPLVFDPE RFLQGEMARIDPMGNDFELIPFGAGRRICAGKLAGMV MVQYYLGTLVHAFDWSLPEGVGELDMEEGPGLVLPK AVPLAVMATPRLPAAAYGLL	57
Dihydroflavonol 4- reductase (DFR) <i>Acer palmatum</i>	MGSEAETVCVTGASGFIGSWLIMRLLERGYTVRATVR DPDNEKKVKHLVELPKAKTHLTLWKADLSDEGSFDE AIHGCTGVFHVATPMDFESKDPENEVIKPTINGVLGIM KACKKAKTVKRLVFTSSAGTVDVEEHKKPVYDENS SDLDFVQSVKMTGWMYFVSKTLAEKAAWKFAEENSI DFISVIPPLVVGPFLLMPSPMPPSLITALSPITRNEGHYAI KQGNVYHLDDLGMGHIFLYEHAESKGRYFCSSHSATI LELSKFLRERYPEYDLPTEYKGVDDSLNVVFCSSKIL	58

Accession: AWN08247.1	DLGFQFKYSLEDMFTGAVETCREKGLIPLTNIDKKHV AAKGLIPNNSDEIHVAAAEEKTTATA	
Dihydroflavonol 4- reductase (DFR) <i>Abrus precatorius</i> Accession: XP_027329642.1	MGSASETVCVTGAAGFIGSWLVMRLIQNGYKVRATV RDPANMKKVKHLLLPNAKTNLSLWKADLAEEGSFD EAIKGCTGVFHVATPMDFESKDPENEVIKPTINGLIDI MKACMKAKTVRRLVFTSSAGTVDVTEHPKPLFDESC WSDVQFCRRVRMTGWMYFVSKTLAEQEA WKFAKEN NIDFISVIPPLVVGPFLLVPTMPPSLITALSLITGNESHYAI IKQGQFVHLDDLCLAHIFLFQHPKAQGRYICCSHEATI HDIASLLNQKYPEFNVPTKFKNIPDQLEIIRFSSKKITDL GFKFKYSLEDMFTGAVETCKEKRLSETAEISGTTQK	59
Dihydroflavonol 4- reductase (DFR) <i>Dendrobium moniliforme</i> Accession: AEB96144.1	MENEKKGPPVVVTGASGYVGSWLVMLLQKGYEVRA TVRDPTNLKKVKPLLDLPRSNELLSIWKADLDGIEGSF DEVIRGSIGVFHVATPMNFQSKDPENEVIQPAINGLLGI LRSCKNAGSVQRVIFTSSAGTVNVEEHQAAAYDETC WSDLDFVNRVKMTGWMYFLSKTLAEKAAWEFVKD NHIHLITIPTLVVGSFITSEMPPSMITALSLITGNDAHY SILKQIQFVHLDDLCDAHIFLFEHPKANGRYICSSYDST IYGLAEMLNRYPTYAIPHKFKEIDPDIKCVSFSSKKL MELGFKYKYTMEEMFDDAIKTCREKKLIPLNTEEIVL AAEKFEVKEQIAVK	60
Dihydroflavonol 4- reductase (DFR) <i>Rosa chinensis</i>	MASESESVCVTGASGFVGSWLVMRLLDRGYTVRATV RDPANKKKVKHLLDLPKAATHLTLWKADLAEEGSFD EAIKGCTGVFHVATPMDFESKDPENEVIKPTINGVLDI MKACLKAKTVRRLVFTASAGSVNVEETQKPVYDESN WSDVEFCRRVKMTGWMYFASKTLAEQEA WKFAKEN NIDFITIPTLVIGPFLMPAMPPSLITGLSPLTGNESHYSII	61

Accession: XP_024167119.1	KQGQFIHLDDLDCQSHIYLYEHPKAEGRYICSSHDATIH EIAKLLREKYPEYNVPTTFKGIEENLPKVHFSSKKLLE TGFEEKYSLEDMFVGAVDACKAKGLLPPPTERVEKQE VDESSVVGKVTG	
Leucoanthocyanidin reductase (LAR) <i>Camellia sinensis</i> Accession: XP_028127206.1	MTVSSPCVGEGQGRVLIIGASGFIGEFIAQASLDSGRTT FLLVRSLDKGAIPSKSKTINSLHDKGAILIHGVIEDQEF VEGILKDHKIDIVISAVGGANILNQLTIVKAIAVGTIK RFLPSEFGHDVDRANPVEPGLAMYKEKRMVRRRIEES GVPYTYICCNISASWPYYDNTHPSEVIPPLDRFQIYGD GTVKAYFVDGSDIGKFTMKVVDDIRTLNKSVMHFRPSC NFLNMNELSSLWEKKIGYMLPRLTVTEDDLLAAAE NIIPQSIVASFTHDIFIKGCQVNFSDGPNEVEVSNLYPD ETFRMTDECFDDFVMKMDRWN	62
Leucoanthocyanidin reductase (LAR) <i>Coffea arabica</i> Accession: XP_027097479.1	MTRSPSPNGQAEKGSRLIIGATGFIGHFIAQASLASGK STYILSRAAARCPKARAIKALEDQGAISIHGSVNDQE FMEKTLKEHEIDIVISAVGGGNLLEQVILIRAMKAVGT IKRFLPSEFGHDVDRAEPVEPGLTMYNEKRRVRRRIEE SGVPYTYICCNISASWPYYDNTHPSEVSPPLDQFQIYG DGSVKAYFVAGADIGKFTVKATEDVRTLNKIVHFRPS CNFLNINELATLWEKKIGRTLPRVVSEDDLLAAAE NIIPQSVVASFTHDIFIKGCQVNFVDPGPNEIEVSSLYP DEPFQTMDECFNEFAGKIEEDKKHVVGTKGKNIAHRL VDVLTAPKLCA	63
Leucoanthocyanidin reductase (LAR) <i>Theobroma cacao</i>	MKSTNMNGSSPNVSEETGRTLVVGSGGFMGRFVTEA SLDSGRPTYILARSSNSPSKASTIKFLQDRGATVIYGS TDKEFMEKVLKEHKIEVVISAVGGGSILDQFNLEAIR NVDTVKRFLPSEFGHDTDRADPVEPGLTMYEQKRQIR RQIEKSGIPYTYICCNISAAWPYHDNTHPADVLPLDR FKIYGDGTVKAYFVAGTDIGKFTIMSIEDDRTLNKTVH	64

Accession: ADD51357.1	FQPPSNLLNINEMASLWEEKIGRTLPRVTITEEDLLQM AKEMRIPQSVVAALTHDIFINGCQINFSLDKPTDVEVC SLYPDTPFRTINECFEDFAKKIIDNAKAVSKPAASNNAI FVPTAKPGALPITAICT	
Leucoanthocyanidin reductase (LAR) <i>Fragaria x ananassa</i> Accession: ABH07785.2	MTVSPSIASAAKSGRVLIIGATGFIGKFVAEASLDGLP TYVLVRPGPSRPSKSDTIKSLKDRGAILHGVMSDKPL MEKLLKEHEIEIVISAVGGATILDQITLVEAITSVGTVK RFLPSEFGHDVDRADPVEPGLTMYLEKRKVRRAIEKS GVPYTYICCNISIASWPYYDNKHPSEVVPPLDQFQIYGD GTVKAYFVDGPDIGKFTMKTVDIRTMNKNVHFRPSS NLYDINGLASLWEKKIGRTLPKVTITENDLLTMAAEN RIPESIVASFTHDIFIKGCTNFPIEGPNDVDIGTLYPEE SFRTLDECFNDFLVKVGKLETDKLAACKNKA AVGVE PMAITATCA	65
Anthocyanin dioxygenase (ANS) <i>Chenopodium quinoa</i> Accession: XP_021735950.1	MTQNKEPVNQKSEHDEQRVESLASSGIESIPKEYVRL NEELTSMGNVFEEEEKKEEGSQVPTIDIKDIASEDPEVR GKAIQELKRAAMEWGVMHLVNHGISDELIDRVKVAG QTFFELPVEEKEKYANDQASGNVQGYGSKLANSASG RLEWEDYYFHLSPEDKRDLSIWPETPADYIPAVSEYS KELRYLATKILSALSALGLEEGRLEKEVGGLEELLQ FKINYYPKCPQPELALGVEAHTDVSALTFILHNMVPG LQLFYEGKWVTAKCVPNSIIMHIGDTIEILSNGKYKSIL HRGLVNKEKVRISWAVFCEPPKEKIILKPLPDLVSDEE PARYPRTFAQHVQYKLFKRKTQGPQTITKN	66

Anthocyanin dioxygenase (ANS) <i>Iris sanguinea</i> Accession: QCI56004.1	MASSKVMPAPARVESLASSGLASIPTEYVRPEWERDD SLGDALEEIKKTEEGPQIPIVDLRGFDGDEKERLHCM EEVKEAAVEWGVMMHIVNHGIAPELIERVRAAGKGFFD LPVEAKERYANNQSEGKIQGYGSKLANNASGQLEWE DYFFHLIFPSDKVDLSIWPKEPADYTEVMMEFAKQLR VVVTKMLSILSLGLGFEEKLEKKLGGMEELLMQMKI NYYPKCPQPELALGVEAHTDVSSLSFILHNGVPGLQV FHGGRWVNARLVPGSLVVHVGDLEILSNGRYKSVL HRGLVNKEKVRISWAVFCEPPKEKIVLEPLAELVDKR SPAKYPPRTFAQHIQHKLFKKAQEQLAGGVHIPEAIQN	67
Anthocyanin dioxygenase (ANS) <i>Magnolia sprengeri</i> Accession: AHU88620.1	MATQVASIPRVEMLASAGIQAIPTTEYVRPEAERNSIGD VFEEKKLEGPQIPVVDLMGLEWENEEVFKKVEEDM KKAASEWGVMMHIINHGISMELMDRVRIAGKAFFDLPI EEKEMYANDQASGKIAGYGSKLANNASGQLEWEDYF FHLIFPEDKRDMSIWPKQPSDYVEATEEFAKQLRGLV TKVLVLLSRGLGVEEDRLEKEFGGMEELLQMKINYY PKCPQPDALGVEAHTDVSALTIFILHNMVPGLQVFFD DKWVTAKCIPGALVVHIGDSLEILSNGKYRSILHRGLV NKEKVRISWAI FCEPPKEKVVLQPLPELVSEAEPARFT PRTFSQHVRQKL FKKQQDALENLKSE	68
Anthocyanin dioxygenase (ANS) <i>Prosopis alba</i> Accession:	MVSSAAVVATRVERLATSGIKSIPKEYVRPQEELTNIG NVFEEKKKEGPEVPTIDLTEIESEDEVVRARCHETLKK AAQEWGVMMNLVNHGIPPELLNQLRKAGETFFSLPIEE KEYANDQASGKIAGYGSKLANNASGQLEWEDYFFH LVFPEDKCDLSIWPRTPSDYIEVTSEYARQLRGLATKI LGALSLGLGLEKGRLEEEVGGMEELLQMKINYYPIC PQPELALGVEAHTDVSSLTFLHNMVPGLQLFYNGQ WITAKCVPNSIFMHIGDTVEILSNGRYKSILHRGLVNK	69

XP_028787846.1	EKVRISWAVFCEPPKEKIILKPLPELVTDDPARFPRT FAQHIQHKLFRKCQEGLSK	
Anthocyanidin-3- O-glycotransferase (3GT) <i>Cephalotus follicularis</i> Accession: GAV66155.1	MPQFTTNEPHVAVLAFPFGTHAAPLITIIHRLAVASPN THFSFLNTSQSNNSIFSSDVYNRQPNLKAHNVWDGVP EGYVFGKPKQESIELFVKAAPETFRKGVEAAVAETGR KVSCLVTDAFFWFAAEIAGELGVPWVPFWTAGPCSL THVYTDLIRKTIGVGGIEGREDESLEFIPGMSQVVIRDL QEGIVFGNLESVFSMDVHRMGIVLPQAAAFINSFEEL DLTITNDLKSKFKQFLSIGPLNLASPPRVPDTNGCLP WLDQQKVASVAYISFGTVMAPSPELVALAEALEASK IPFIWSLGEKLVHLPKGFLDKTRTHGIVVPWAPQSDV LENGAVGVFITHCGWNSLLESIAGGVPMICRPFFGDQ RLNGRMVQDVWEIGVTATGGPFTTEGVMGDLDLILS QARGKKMKDNISVLKTLAQTAVGPEGSSAKNYEALL NLVRLSI	70
Anthocyanidin-3- O-glycotransferase (3GT) <i>Prunus cerasifera</i> Accession: AKV89253.1	MAPQPIDDDHVVEHHVAALAFPFSTHASPTLALVRR LAAASPNTLFSFFSTSQSNNSLFSNTITNLPRNIKVF ADGVDPDGYVFAGKPKQEDIELFMKAAPHNFTTSLDTCV AHTGKRLTCLITDAFLWFGAHLAHDLGVPWLPLWLS GLNSLSLHVHTDLLRHTIGTQSIAGRENELITKNVNIPG MSKVRIKDLPEGVIFGNLDSVFSRMLHQMGQLPRAN AVLVNSFEELDITVTNDLKSKFNKLLNVGPFNLAAAA SPPLPEAPTAADDVTGCLSWLDKQKAASSVVYVSFGS VARPPEKELLAMAQALEASGVPLWSLKDSFKTPLL ELLIKASNGMVVPWAPQPRVLAHASVGAFVTHCGWN SLETIAGGVPMICRPFFGDQQRVNARLVEDVLEIGVT	71

	EDGVFTKHGLIKYFDQVLSQQRGKKMRDNINTVKLL AQQPVEPKGSSAQNFKLLLDVISGSTKV	
Anthocyanidin-3- O-glycotransferase (3GT) <i>Scutellaria baicalensis</i> Accession: A0A482AQV3	MVFQSHIGVLAFFPGTHAAPLLTVVQRLATSSPHTLFS FFNSAVSNSTLFNNGVLDSYDNIRVYHVWDGTPQGG AFTGSHFEAVGLFLKASPGNFDKVIDEAEVETGLKISC LITDAFLWFGYDLAEKRGVPWLAFWTSAQCALSAHM YTHEILKAVGSNGVGETAEEELIQLIPGLEMAHLSL PPEIFFDKNPPLAITINKMVLKLPKSTAVILNSFEEIDP IITDLKSKFHHFLNIGPSILSSPTPPPPDDKTGCLAWLD SQTRPKSVVYISFGTVITPPENELAALSEALETCNYPFL WSLNDRAKKS LPTGFLDRTKELGMIVPWAPQPRVLA HRSVGVFVTHCGWNSILESICSGVPLICRPFFGDQKLN SRMVEDSWKIGVRLEGGVLSKTATVEALGRVMMSEE GEIIRENVNEMNEKAKIAVEPKGSSFKNFNKLLEIINAP QSS	72
Anthocyanidin-3- O-glycotransferase (3GT) <i>Vitis vinifera</i> Accession: P51094	MSQTTTNPHVAVLAFFPSTHAAPLLAVVRRLAAAAPH AVFSFFSTSQSNASIFHDSMHTMQCNIKSYDISDGVE GYVFAGRPQEDIELFTRAAPESFRQGMVMAVAETGRP VSCLVADAFIWFAADMAAEMGLAWLPFWTAGPNSLS THVYIDEIREKIGVSGIQGREDELLNFIPGMSKVRFRDL QEGIVFGNLNSLFSRMLHRMGQVLPKATAVFINSFEEL DDSLTNDLKSKLKTYLNIGPFNLITPPPVVPNTTGCLQ WLKERKPTSVVYISFGTVTTPPPAEVVALSEALEASRV PFIWSLRDKARVHLPEGFLEKTRGYGMVVPWAPQAE VLAHEAVGAFVTHCGWNSLWESVAGGVPLICRPFFG DQRLNGRMVEDVLEIGVRIEGGVFTKSGLMSCFDQIL	73

	SQEKGGKKLRENLRALRETADRAVGPKGSSTENFITLV DLVSKPKDV	
Acetyl-CoA carboxylase (ACC) <i>Ustilago maydis</i> 521 Accession: XP_011390921.1	MPPPDHKAVSQFIGGNPLETAPASPVADFIRKQGGHS VITKVLICNNGIAAVKEIRSIRKWAYETFGDERAIEFTV MATPEDLKVNADYIRMADQYVEVPGGSNNNNYANV DLIVDVAERAGVHAVWAGWGHASENPRLPESLAASK HKIIFIGPPGSAMRSLGDKISSTIVAQHADVPCMPWSG TGIKETMMSDQGFLTVSDDVYQQACIHTAEEGLEKAE KIGYPVMIKASEGGGGKGIRKCTNGEEFKQLYNAVLG EVPGSPVFMKLAGQARHLEVQLLADQYGNAISIFGR DCSVQRRHQKIIEEAPVTIAPEDARESMEKAAVRLAK LVGYVSAGTVEWLYSPESGEFAFLELNPRLQVEHPTT EMVSGVNIPAAQLQVAMGIPLYSIRDITLYGMDPRG NEVIDFDFSSPESFKTQRKPQPQGHVVACRITAENPDT GFKPGMGALTELNFRSSTSTWGYFSVGTSGALHEYAD SQFGHIFAYGADRSEARKQMVISLKELSIRGDFRTTVE YLIKLETDAFESNKITTGWLDGLIQDRLTAERPPADL AVICGAAVKAHLLARECEDEYKRILNRGQVPPRDTIK TVFSIDFIYENVKYNFTATRSSVSGWVLYLNGGRTL QLRPLTDGGLLIGLSGKSHPVYWREEVGMTRLMIDSK TCLIEQENDPTQIRSPSPGKLVRFLVDSGDHVKANQAI AEIEVMKMYLPLVAAEDGVVSFVKTAGVALSPGDIIG ILSLDDPSRVQHAKPFAGQLPDFGMPVIVGNKPHQRY TALVEVLNDILDGYDQSFMRQAVIKELIETLRNPELPY GQASQILSSLGGRIPARLEDVVRNTIEMGHSKNIEFPA ARLRKLTENFLRDSVDPAIRGQVQITIAPLYQLFETYA GGLKAHEGNVLASFLQKYYEVESQFTGEADVLELR LQADGDLKVVVALQTSRNGINRKNALLLTLLDKHIKG TSLVSRITSGATMIEALRKLASLQKSTAPIALKAREVS	74

	LDADMPSLADRSAQMMAILRGSVTSSKYGGDDEYHA PSLEVLRELSDSQSVYDVLHSFFGHREHHVAFALC TYVVRAYRAYEIVNFDYAVEDFDVEERAVLTWQFQL PRSASSLKERERQVSISDLMMMDNNRRARPIRELRTGA MTSCADVADIPELLPKVLKFFKSSAGASGAPINVLNV AVVDQTDVFDAEVRSQLALYTNACSKESAAARVRRV TYLLCQPGLYPFFATFRPNEQGIWSEEKAIRNIEPALA YQLELDRVSKNFELTPVPVSSSTIHL YFARGIQNSADT RFFVRSLVRPGRVQGDMAAYLISESDRIVNDILNVIEV ALGQPEYRTADASHIFMSFIYQLDVSLVDVQKAIAAGFL ERHGTRFFRLRITGAEIRMILNGPNGEPRPIRAFTNET GLVVRYETYEETVADDGSVILRGIEPQGKDATLNAQS AHFPYTTKVALQSRRSRAHALQTTFFVYDFIDVLGQAV RASWRKVAASKIPGDVIKSAVELVFDEQENLREVKRA PGMNNIGMVAWLVEVLTPEYPAGRKL VVIGNDVTIQ AGSFGPVEDRFFAAASKLARELGVPRLYISANS GARIG LATEALDLFKVKFVGDDPAKGFEYIYLDDESLQAVQA KAPNSVMTKPVQAADGSVHNIITDIIGKPQGGLGVEC LSGSGLIAGETSRAKDQIFTATIITGRSVGIGAYLARLG ERVIQVEGSPLIL TGYQALNKLLGREVYTSNLQLGGPQ IMYKNGVSHLTAQDDLDVRSFVNWISYVPAQRGGP LPIMPTTDSWDRAVTYQPPRGPDPRWLINGTKAEDG TKLTGLFDEGSFVETLGGWATSVVTGRARLGGIPVG IAVETRTLERVVPADPANPNSTEQRIMEAGQVWYPNS AYKTAQAIWDFDKEGLPLVILANWRGFSGGQQDMYD EILKQGSKIVDGLSSYKQPVFVHIPPMGELRGGSWVV VDSAINDNNGMIEMSADVNSARGGVLEASGLVEIKYRA DKQRATMERLDSVYAKLSKEAAEATDFTAQTTARKA LAEREKQLAPIFTAIATEYADAHDRAGRMLATGVLRS ALPWENARRYFYWRLRRRLTEVAAERTVGEANPTLK	
--	--	--

	HVERLAVLRQFVGAAASDDDKAVAEHLEASADQLLA ASKQLKAQYILAQISTLDPELRAQLAASLK	
Acetyl-CoA carboxylase (ACC) <i>Hesseltinella vesiculosa</i> Accession: ORX57605.1	MVDHKS L P G H F L G G N S V D T A P Q D P V C E F V K S H Q G H T V I S K V L I A N N G M A A M K E I R S V R K W A Y E T F G N E R A I E F T V M A T P E D L K A N A E Y I R M A D N Y I E V P G G T N N N N Y A N V E L I V D V A E R T G V H A V W A G W G H A S E N P R L P E M L A K S K N K C V F I G P P A S A M R S L G D K I S S T I V A Q S A D V P T M G W S G D G V S E T T D H N G H V L V N D D V Y N S A C V K T A E A G L A S A E K I G F P V M I K A S E G G G G K G I R K V E D P S T F K Q A F A Q V Q G E I P G S P I F I M K L A G N A R H L E V Q L L A D Q Y G N A I S L F G R D C S V Q R R H Q K I I E E A P V T I A K P D I F E Q M E K A A V R L G K L V G Y V S A G T V E Y L Y S H H D E K F Y F L E L N P R L Q V E H P T T E M V S G V N L P A A Q L Q I A M G I P M H R I R D I R V L Y G V Q P N S A S E I D F D L E H P T A L Q S Q R R P M P K G H V I A V R I T A E N P D A G F K P S G G V M Q E L N F R S S T N V W G Y F S V V S S G A M H E Y A D S Q F G H I F A Y G E N R Q Q A R K N M V I A L K E L S I R G D F R T T V E Y I I R L L E T P D F T D N T I N T G W L D M L I S K K L T A E R P D T M L A V F C G A V T K A H L A S V E C W Q Q Y K N S L E R G Q I P S K E S L K T V F T V D F I Y E N I R Y N F T V T R S A P G I Y T L Y L N G T K T Q V G V R D L S D G G L L I S L N G R S H T T Y N R E E V Q A T R L M I D G K T C L L E K E S D P T Q L R S P S P G K L V S L L L E N G D H I R T G Q A Y A E I E V M K M Y M P L V A S E D G H V Q F I K Q V G A T L E A G D I I G I L S L D D P S R V K H A L P F T G Q V P K Y G L P H L T G D K P H Q R F T H L K Q T L E Y V L Q G Y D N Q G L I Q T I V K E L S E V L N N P E L P Y S E L S A S M S V L S G R I P G R L E Q Q L H D L I N Q A H A Q N K G F P A V D I Q Q A I D T F A R D H L T T Q A E V N A Y K T A V A P I M T I A A S Y S N G L K Q H E H S V Y V D L M E Q Y Y N V E V L F N S N Q S R D E E V I L A L R D Q H K D D L E K V I N I L S H A K V N I K N N L I L M L L D I I P A T S S E A L D R C F L P I L K H L S E I D S R G T Q K V T L K A R E Y L I L C Q L P S L E E R Q S Q M Y N I L K S S V T E S V Y G G G T E Y R	75

	TPSYDAFKDLIDTKFNVFDVLPNFFYHPDSYVSLAALE VYCRRSYHAYKILDVAYNLEHQPYIVAWKFLQSSA GGGFNNQRIASYSDLTFLLNKTEEEPIRTGAMVALKTL EELEAELPRIMTAFEEEEPLPPMLMKQPPDPKTEERMEN ILNISIQQQDMEDDTLRKNMTTLIAHSDAFRKAALR RITLVVCRDNQTPDYITFRERNGYEEDETIRHIEPALA YQLELARLSNFDIKPCFIENRQMHVYYAVAKENPSDC RFFIRALVRPGRVKSSMRTADYLISESDRLLTDILTLE IVSHDYKNSDCNHLFINFIPTFAIEADEVETALKDFVDR HGKRLWKLRTGAEIRFNIQSKRPDAPVIPLRFTVDNV SGYILKVDVYQEVKTDKNGWILKSVGKIPGAMHMQP LSTPYPTKEWLQPRRYKAHLMGTTYVYDFPELFRQAI HNLWAQACKADA AVKIPSQVIEAKELVLDDDNQLQA IDRAPGTNTVGMVAWLLTLRTPDYPRGRRVIAIANDI TFKIGSFGVQEDLVFYKASEYARELGVP RVYLSANS ARIGLADELISR FHV AWKDEDQPGSGFEYLYLLPEEY DALIQQGD AQSVLVQEVQDKGERRFRITDIIGHTDGL GVENLRGSGLIAGATS RAYDDIFTITLVTCR SVGIGAY LVRLGQRTVQNEGQPIILTGAPALNKVLGREVYTSNL QLGGTQIMYKNGVSHLTAENDLEGINKIMQWLSFVPE CRGAPLPMRAGADPIDREIEYLPPKGPSDPRFFLAGKQ ENGKWLSGFFDHGSFVETLSGWARTVVVGRARLGGI PMGVVAVETRTRVENIVPADPANADSQEQVVM EAGGV WFPNSAYKTAQAINDFNKGEQLPLMIFANWRGFSGG QRDMYNEVLKYGAQIVDALSNYKQPVFVYVVPNGEL RGGAWVVVDSTINEDMMEMYADTQARGGVLEPEGI VEIKYRRPQLLATMERLDPVYSDLKRRLAALDDSQKE QADELIAQVEAREQALLPVYQQVAIQFADLHDRSGR MEAKGVIRKTLEWRTARHYFYWRVRRRLLEEYAIRK MDES RDQAKTLLQQWFQADTNLDDFDKNDQAVVA	
--	---	--

	WFDANKLLLDQRIAKLKSEKLKDHVVQLASVDQDAV VEGFSKLMESLSVDQRKEVLHKLATRF	
Acetyl-CoA carboxylase (ACC) <i>Rhodotorula toruloides</i> NBRC10032 Accession: GEM08739.1	MASTTPHDSRVVSVSSGKKLYIEVDDGAGKDAPAIVF MHGLGSSTSFWEAPFSRNLSSRFRLIRYDFDGHGLSP VSLDAADDGAMIPLVDLVEDLAAVMEWTGVDKVA GIVGHSMGLVASTFAAKYPQKVEKL VLLGAMRSLN PTVQTNMLKRADTVLESGLSAIVAQVVSAAALSDKSKQ DSPLAPAMVRTLVLGTDPLGYAAACRALAGAKDPDY STIKAKTLVVSGESDYLSNKETTEALVNDIPGAKEVQ MDGVGHWHAVEDPAGLAKILDGFFLQGKFSGEAKA VNGSHAVDETPKKPKYDHGRVVKYLGGNSLESAPPS NVADWVRERGGHTVITKILIANNGIAAVKEIRSVRKW AYETFGSERAIEFTVMATPEDLKVNADYIRMADQYVE VPGGTNNNNYANVDVIVDVAERAGVHAVWAGWGH ASENPRLPESLAASKHKIVFIGPPGSAMRSLGDKISSTI VAQHAEVPCMDWSGQGVDQVTQSLEGYVTVADDVY QQACVHDADEGLARASRIGYPVMIKASEGGGGK GIR KVEREQDFKQAFQAVLTEVPGSPVFIMKLAGAARHLE VQVLADQYGN AISLFGRDCSVQRRHQKIIEEAPVTIAK PDTFEQMEKSAVRLAKLVGYVSAGTVEFLYSAADDK FAFLELNPRQLQVEHPTTEMVSGVNLPAACLQVAMGV PLHRIRDITLYGKAPNGSSEIDFEFENPESAKTQRKPS PKGHVVAVRITAENPDAGFKPSMGTQLQELNFRSSTNV WGYFSVGSAGGLHEFADSQFGHIFAYGSDRSES RKN MVVALKELSIRGDFRTTVEYLIKLLTDAFEQNTITTA WLDSLISARLTAERPDTTLAIICGAVTKAHLASEANIA EYKRILEKGQSPPKELLATVVPLEFVLEDVKYRATASR SSPSSWSIYVNGSNVSVGIRPLADGGLLILLDGRSYTC YAKEEVGALRLSIDSRTVLVAQENDPTQLRSPSPGKL VRYFIESGEHISKGEAYAEIEVMKMIMPLIAAEDGIAQ	76

	FIKQPGATLEAGDILGILSLDDPSRVHHAKPFDGQLPA LGLPSIIGTKPHQRFAYLKDVLSNILMGYDNQAIMQSS IKELISVLRNPELPYGEANAVLSTLSGRIPAKLEQTLRQ YIDSAHESGAEFPSAKCRKAIDTTLEQLRPAAEQTVRN FLVAFDDIVYRYRSGLKHHEWSTLAGIFAAAYAETEKP FSGKSDVVLRLDAHRDSLDSVVKIVLSHYKAASKN SLVLALLDVVKSDSVPLIEQVVSPALKDLADLDSKA TTKVALKAREVLIHIQLPSLDERLGQLEQILKASVTPT VYGEPGHDRTPRGEVLKDVIDSRFTVFDVLPSSFQHQ DQWVSLAALDITYVRRAYRSYNLLNIEHIEADAAEDEP ATVAWSFRMRKAASESEPPTPTTGLTSQRTASYSDLT FLLNNAQSEPIRYGAMFSVRSLDGFRQELGTVLRHFP DSNKGKLQQQPAASSSQEQWNVINVALTPASAQVD EDALRADFAAHVNAMSAEIDARGMRRLTLICREGQ YPSYYTVRKQDGTWKELETIRDIEPALAFQLELGRLSN FHLEPCPVENRQVHIYYATAKGNSSDCRFFVRALVRP GRLRGNMKTADYLVSEADRLVTDVLDSEVASSQRR AADGNHISLNFLYSLRLDFDEVQAALAGFIDRHGKRF WRLRVTGAEIRIVLEDAQGNIQPIRAIENVSGFVVKYE AYREVTTDKGQVILKSIGPQGALHLQPVNFPYPTKEW LQPKRYKAHVVGTTYVYDFPDLFRQAIKQWKAVGK TAPAEELLVAKELVLDEFGKPQEVARPPGTNNIGMVG WIYTIFTPEYPSGRRVVVIANDITFKIGSFGPEEDRYFY AVTQLARQLGLPRVYLSANSGARLGIAEELVDLFSVA WADSSRPEKGFKYLYLTAEKLGELKNKGEKSVITKRI EDEGETRYQITDIIGLQEGLGVESLKGSGLIAGETSRA YDDIFTITLVTARSVGIGAYLVRLGQRAVQVEGQPIILTG AGALNKVLGREVYSSNLQLGGTQIMYKNGVSHLTAA NDLEGVLSIVQWLAFVPEHRGAPLPVLPSPVDPWDRSI DYTPIKGAYDPRWFLAGKTDEADGRWLSGFFDKGSF QETLSGWAQTVVVGRARLGGIPMGAIAVETRTIERIIP	
--	---	--

	ADPANPLSNEQKIMEAGQVWYPNSSFKTGQAIFDFNR EGLPLIIFANWRGFSGGQQDMFDEV LKRGSLIVDGLS AYKQPVFVYIVPNGELRGGAWVVLDP SINAEGMMEM YVDETARAGVLEPEGIVEIKLRKDKLLALMDRLDPTY HALRVKSTDASLSPTDAAQAKTELAAREKQLMPIYQQ VALQFADSHDKAGRILSKGCAREALEWSNARRYFYA RLRRRLAEEAAVKRLGEADPTLSRDERLAIVHDAVGQ GVDLNNDLAAAAAFEQGA AITERVKLARATTVAST LAQLAQDDKEAFAASLQQVLGDKLTAADLARILA	
Malonyl-CoA synthase (matB) <i>Rhodopseudomonas palustris</i> Accession: WP_011661926.1	MNANLFSRLFDGLVEADKLAIETLEGERISYGDLVAR SGRMANVLVARGVKPGDRVAAQAEKSVAALVLYLA TVRAGAVYLPLNTAYTLHELDYFIGDAEPKLVVCDPA KREGIAALAQKVGAGVETLDAKGQGS LSEAAAQASV DFATVPREGDDLAAILYTS GTTGRSKGAMLSHDNLS NSLTLVEFWRFTPD DVLIALPIYHTHGLFVASNVTLF ARASMIFLPKFDPDAIIQLMSRASVLMGVPTFYTRLLQ SDGLTKEAARHMR LFISSAPLLADTHREWASRTGHA VLERYGMTETNMNTSNPYD GARVPGAVGPALPGVSL RVVDPETGAELSPGEIGMIEVKGPNVFQGYWRMPEKT KAEFRDDGFFITGDLGKIDADGYVFIVGRGKDLVITGG FNVYPKEVESEIDAISGVVESAVIGVPHADLGEGVTAV VVRDKGASVDEAAVLGALQGQLAKFKMPKRVLFVD DLPRNTMGKVQKNVLREAYAKLYAK	77
Malonyl-CoA synthase (matB) <i>Rhizobium sp. BUS003</i>	MVNHLFDAIRLSITSPESTFIELEDGKVW TYGAMFNCS ARITHVLVKLGVS PGDRVAVQVEKSAQALMLYL GCL RAGAVYLPLNTAYTPAELEYFLGDATPKLVVVS PCAA EQLEPLARRVGTRLLTLGVNGDGSLMDMASLEPVEF ADIERKADDLAAILYTS GTTGRSKGAMLT HDNLLSNA QTLREHWRFTSADRLIHALPIFH THGLFVATNV TLLAG GAIYLLSKFDPDQIFALMTRATVMMGVPTFYTRLLQD	78

Accession: NKF42351.1	ERLNKANTRHMRLFISGSAPLLAETHRLFEEYTGHAIL ERYGMTETNMITSNPCD GARVP GTVGYALPGVSVRIT DPVSGEPLAAGEPGMIEVKGPNVFQGYWNMPDKTKE EFRSDGYFTTGDIGVMETDGRISIVGRGKDLIISGGYNI YPKEIENEIDAIEGVVESAVIGVPHPD LGEGVTAIVVG QPKAHLDTTITNNLQGRLARFKQPKNVIFVDELPRNT MGKVQKNVLRDRYRDLYLK	
Malonyl-CoA synthase (matB) <i>Ochrobactrum</i> sp. 3-3 Accession: WP_114216069.1	MANHLFDLVRANATDLTKTFIETETGLKLTYYDDLMT GTARYANVLVGLGVKPGDRVAVQVEKSAGAIFLYLA CVRAGAVFLPLNTAYTLTEIEYFLGDAEPALVVC DPA RRDGITEVAKKTGVP AVETLGKGQDGS LFDKAAAAP ETFADVARGPGDLAAILYTS GTTGRSKGAMLSHDNLA SNALTLKDYWRF GADDVLLHALPIFHTHGLFVATNTI LVAGASMLFLPKFDADKVFELMPRATTMMGVPTFYV RLVQDARLTREATKHMRLFISGSAPLLAETHKLFREK TGVSILERYGMTETNMNTSNPYDGDRVAGTVGFPLPG VALRVADPETGAAIPQGEIGVIEVKGPNVFSGYWRMP EKTAAEFRQDGGFFITGDLGKIDDQGYVHIVGRGKDLV ISGGYNVYPKEVETEIDGMAGVVESAVIGVPHPDFGE GVTAVVVAEKGASLDEATHIKTLEQRLARYKLPKRVI VDDLPRNTMGKVQKNLLRDAYKGLYGG	79
Malonate transporter (matC) <i>Rhizobiales</i> <i>bacterium</i> Accession:	MSPELISILVLVVVFVIATTRSVNMGALAFAAAFGVGT LVADLDADGIFAGFP GD L FVVLVGVTYLF AIARANGT TDWL VHAAVRLVRGRVALIPWVMFALTGALTAIGAV SPA AVAIVAPVALSFATRY SISPLLMGT MVVHGAQAG GFSPISYGSIVNGIVEREKLP GSEIGLFLASLVANLLIA AVLFAVLGGRKLWARGAVTPEGDGAPGKAGTGTTGS GSDTGTGTGTGTGTSAGTGGTAPTAVAVRSDRETGG AEGTGVRLTPARVATLVALVALVAVLGF DL DAGLT AVTLAVVLSTAWPDDSRRAVGEIAWSTVLLICGVLT Y	80

MBN8942514.1	VGVL EEMGTITWAGEGVGGIGVPLLA AVLLCYIGAIV SAFASSVGIMGALIPLAVPFLAQGEIGAVGMVAALAV SATVVDVSPFSTNGALVLAAAPDVDRDRFFRQLMVY GGIVVAAVPALAWLVLVVPFGF	
Malonate transporter (matC) <i>Rhizobium leguminosarum</i> Accession: AAC83457.1	MGIELLSIGLLIAMFIIATI QPINMGALAFAGAFVLGSMI IGMKTNEIFAGFPSDLFLTLVAVTYLFAIAQINGTIDWL VECAVRLVRGRIGLIPWVMFLVAAIITGFGALGPAAV AILAPVALSFAVQYRIHPVMMGLMVIHGAQAGGFSP SIYGGITNQIVAKAGLPFAPTSFLSSFFFNLAIAVLVFF VFGGARVMKHDPASLGPLPELHPEGVSASIRGHGGTP AKPIREHAYGTAADTATTLRLNNERITTLIGLTALGIG ALVFKFNVGLVAMTVAVVLALLSPKTQKAAIDKVS STVLLIAGIITYVGVMEKAGTVDYVANGISSLGMPLLV ALLLCFTGAIVSAFASSTALLGAIPLAVPFLQGHISAI GVVAAIAISTTIVDTSPFSTNGALVVANAPDDSRQVL RQLLIYSALIAIIGPIVAWL VFVVPGLV	81
Malonate transporter (matC) <i>Agrobacterium vitis</i> Accession: WP_180575084.1	MNIEILSIGLLVAIFIIATI QPINMGVLAFGCTFVLGSLII GMKPADIFAGFPADLFLTLVAVTYLFAIAQINGTIDWL VERSVRMVRGRVGVIPWVMFLVAAIITGFGALGPAA VAILAPVALSFAVQYRIHPVLMGLMVIHGAQAGGFSP SIYGGITNQIVAKAGLPFAPTSFLSSFFFNLAIAVLIFI FGGLSILKQRSSVKGPLPELHPEGISASIKGHGGTPAKP FREHAYGTAADTQSKVRLTTEKVTTLIGLTALGVGAL VFKFNVGLVAITVAVLLALLSPTTQKAAIDKVS STVLLISGIITYVGVMEKAGTIDYVAHGSSLGMPLLVALL LCFTGAIVSAFASSTALLGAIPLAVPFLQGHISAVGV VAAIAISTTIVDTSPFSTNGALVVANAPDDQQRDKVMR QMLIYSALIALIGPIVAWL VFVVPGLI	82

Malonate transporter (matC) <i>Neorhizobium sp.</i> Accession: WP_105370917.1	MSIEILSILLLVAMFVIATIQPINMGALAFACFTFVLGSLI IGMKTSDFAGFPSDLFLTLVAVTYLFAIAQINGTIDWL VECAVRMVRGHVAWIPWVMFVVAIIITGFGALGPAA VAILAPVALSFAVQYRIHPVMMGLMVIHGAQAGGFSP ISVYGGITNQIVAKAGLPFAPTSFLSSFFFNLAIAVLVF FVFGGARIMKQAAGPTGPLPELHPEGVSAAIRGHGGT PAKPIREHAYGTAADTLQTLRLTPEKVFTLIGLTALGI GALVFKFNVGLVAITVAVALALISPKTQKAAVDKVS WSTVLLIAGIITYVGVLEKAGTVNYVANGISSLGMPLL VALLLCFTGAIVSAFASSTALLGAIPLAVPFLQGHIS AVGVVAAIAISTTIVDTSPFSTNGALVVANAPDETREQ VLRQLLIYSALIAIGPVVAWLVFVVPGLV	83
Malonate CoA-transferase (MdcA) <i>Moraxella catarrhalis</i> Accession: WP_064617969.1	MTTWNQKQQRKAQKLAKACDSGFDKYVPHERIIALL ETVIDRGDRVCLEGNNQKQADFLSKSLSSCNPDIVNG LHIVQSVLALPSHIDVFERGIASKVDFSAGPQSLRLAQ LVQAQKITIGAIHTYLELYGRYFIDLTPNVALITAHAA DKRGNLYTGANTEDTPAIVEATTFKSGIVIAQVNEIVD ELPRVDIPSDWVDYYTQSPKHNYIEPLFTRDPAQITEIQ ILMAMMAIKGIYAPYKINRLNHGIGFDTAAIELLLPTY AESLGLKGEICHTHWALNPHPTLIPAIESGFIHVSFSGS EVGMENYVKARSDVFFTGADGSMRSNRAFSQTAGLY ACDLFIGSTLQIDLQGNSSTATADRIAGFGGAPNMGSD PHGRRHASAYMKAGREAVDGSPIKGRKLVVQMVE TYREHMQSVMFVNELDAFKLQKMGADLPPIMIYGDD VTHIVTEEGIANLLLCRTPDEREQAIRGVAGYTPIGLG RDDTMVARLRERKVIQRPEDLGINPMHATRDLLAAKS VKDLVRWSDRLYEPPSRFRNW	84

Malonate CoA-transferase (MdcA) <i>Dechloromonas aromatica</i> Accession: WP_011289741.1	MNAPQPRQWDSLRQNRARRLERAASLGLAGQNGKEI PVDRIIDLLEAVIQPGDRVCLEGNNQKQADFLSESLAD CDPARINHLSMVQSVLALPSHVDLFERGLATRLDFSFS GPQGARLAKLVQEQRIEIGAIHTYLELFGRYFMDLTPN VALIAAQAADAEGNLYLGPNTEDTPAIVEATAFKGGI VIAQVNERLDKLPVDPADWVDFTVLAPKPNYIEPL FTRDPAQITEVQVLMAMMAIKGIYAEYGVTRLNHGIG FDTAAIELLLPTYAADLGLKGKICHTHWALNPHPTLIPA IEAGFVESVHCFGSEVGMDDYISARSDIFFTGADGSMR SNRAFSQTAGLYACDMFIGSTLQMDLAGNSSTATLGR ITGFGGAPNMGSDPHGRRHASPAWLKAGREAYGPQA IRGRKLVVQMVETFREHMAPVFVDDLD AWKLQASM GSDLPPIMYGDDVSHIVTEEGIANLLLCRTPAEREQAI RGVAGFTPVGMDKGTVENLRDRGIIRRPEDLGIDP RQASRDLLAARSIKDLVRCSGGLYAPPSRFRNW	85
Malonate CoA-transferase (MdcA) <i>Pseudomonas cissicola</i> Accession: WP_078590875.1	MSRQWDTQADSRRQRLQRAAALAPQGRVVAADDVV ALLEAVIEPGDRVCLEGNNQKQADFLARCLTEVDPAR VHDLHMOVSVLSLAAHLDVFERGIAKRLDFSFSGPQA ARLAGLVSEGRIEIGAIHTYLELFGRYFIDLTPRIALVT AQAADRHGNLYTGPNTEDTPVIVEATAFKGGIVIAQV NEILDTLPRVDIPADWVDFVTQAPKPNYIEPLFTRDPA QISEIQVLMAMMAIKGIYAEYGVDRNLNHGIGFDTAAIE LLLPTYAQSLGLKGKICRHWALNPHPALIPAIESGFVQ SVHSFGSELGMENYIAARPDIFFTGADGSMRSNRALS QTAGLYACDMFIGSTLQIDLQGNSSTATRDRIAGFGG APNMGSDARGRRHASAAWLKAGREAA TP GEMPRGR KLVVQMVETFREHMAPAFVDRLDAWEL AERANMPL PPVMYIGDDVSHVLTEEGIANLLLCRTPEEREQAIRGV SGYTAVGLGRDKRMVENLRDRGVIKRPDDL G IRPRD ATRDLLAARTVKDLVRWSGGLYDPPKRFRNW	86

Malonate CoA-transferase (MdcA) <i>Geobacillus subterraneus</i> Accession: WP_184319829.1	MNKIYREKRSWRTRRDRKAKRIEHMKQIAKGKIIPTE KIVEALTALIFPGDRVIEGNNQKQASFLSKALSQVNP EKVNGLHIIMSSVSRPEHLDFEKGIAARKIDFSYAGPQS LRMSQMLEDGKL VIGEIH TYLELYGRLFIDLTPSVALV AADKADASGNLYTGPNTEETPTLVEATAFRDGIVIAQ VNELADELPRVDIPGSWIDFVVAADHPYELEPLFTRDP RLITEIQILMAMMVIKGIYERHNIQSLNHGIGFNTAAIE LLLPTYGESLGLKGKICKHWALNPHPTLIPAIETGWVE SIHCFGGEVGMKEYIARPDIFFTGKDGNLRSNRTL SQ VAGQYAVDLFIGSTLQIDRDGNSSTVTNGRLAGFGGA PNMGHDPRGRRHSSPAWLDMITSDHPAAKGRKL VVQ MVETFQKGNRPVFVESLDAIEVGRSARLATTPI MIYGE DVTHIVTEEGIA YLYKASSLEERRQAIAAIAGVTPIGLE RDPRKTEQLRRDGVVAFPEDLGIRRTDAKRSL LAAKSI EELVEWSEGLYEPPARFRSW	87
Pantothenate kinase (CoaX) <i>Streptomyces sp. CL12509</i> Accession: WP_095682415.1	MLLTIDVGNTHTVLGLFDGEEIVEHWRISTDSRRTADE LAVLLQGLMGTHPLLGMELGEGIDGIAICSTVPAVLH ELREVSRRYYGDVPAILVEPGVKTGVPILMDNPKEVG TDRIINAVAAQHLYGGPAIVVDFGTATTFD AVSARGE YTGGVIAPGIEISVEALGLRGAQLRKIELARPRSVIGKS TVEAMQSGILYGFAGQVDGVVQRMACELAPDPADV T VIATGGLAPMVLGEAAVIDHHEPWLT LIGLRLVYERN AGR	88
Pantothenate kinase (CoaX) <i>Streptomyces cinereus</i> Accession: WP_188874884.1	MTKLWLDLGNTRLKYWLTD DSGQVLDHAAEQHLQA PAELLKGLTFRLERLNPDFIGVSSVLGQAVNNHVAESL ERLQKPFEEAQVHAKHALMSSDYNPAQLGVDRWLQ MLGIIEPSKKQCVIGCGTAVTIDLVDQGHHLGGYIFPSI YLQRESLFSGTRQISIIDGT FDSIDSGTNTQDAVHHGIM LSIVGAINETIHRYPQFEITMTGGDAHTFEPHLSASVEI RQDLVLAGLQRFFAAKNNTKNQN	89

Pantothenate kinase (CoaX) <i>Kitasatospora kifunensis</i> Accession: WP_184936930.1	MLLTIDVGNTQTTLGLFDGEEVVDHWRISTDPRRTAD ELAVLMQGLMGRQPGGAGRERVDGLAICSSVPAVLH ELREVTRRYYGDLPAVLVAPGVKTGVHVLMDNPKEV GADRIVNALAANHLYGGPCIVVDFGTATTFDAINERG DYVGGAIAPGIEISVEALGVRGAQLRKIELAKPRNVIG KNTVEGMQSGVLYGFAGQVDGLVTRMAKELSPTDPE DVQVIATGGLAPLVLDEASSIDVHEPWLTIGLRLVYE RNTAS	90
glutamyl-tRNA reductase (hemA) <i>Citrobacter freundii</i> Accession: NTY05430.1	MTLLALGINHKTAPVSLRERVTFSPDTLDQALDSLQA LPMVQGGVVLSTCNRTIYLSVEEQDNLREALIRWLC EYHNLNEEDLRNSLYWHQDNDAVSHLMRVASGLDS LVLGEPQILGQVKKAFADSQKGHNASALERMFQKS FSVAKRVRTETDIGSSAVSVAFAACTLARQIFESLSTV TVLLVGAGETIELVARHLREHKVKKMIANRTRERAQ VLADEVGAEVISLSDIDARLQDADIIISSTASPLPIIGKG MVERALKNRRNQPMMLLDIAVPRDVEPEVGKLSNAY LYSVDDLQSIISHNLAQRKAAAVEAETIVEQEASEFMA WLRAQGASDTIREYRSQSEQIRDEL TAKALAAALQQGG DAQAIMQDLAWKLTNRLIHAPTKSLQQAARDGDSER LNILRDSLGLE	91
glutamyl-tRNA reductase (hemA) <i>Pseudomonas reactans</i> Accession: NWA43040.1	MTLLALGINHKTAPVSLRERVTFSPETIEQALSSLLQQP LVQGGVVLSTCNRTIELYLSVEQQENLQEQLVKWLCD YHLSADEVRKSLYWHQDNAAVSHLMRVASGLDSL VVGEPQILGQVKKAFASQHGQAVSGELERLFQKSFS VAKRVRTETDIGASAVSVAFAACTLARQIFESLSDVSV LLVGAGETIELVARHLREHKVRHMMIANRTRERAQV LASEVGAEVITLQDIDARLADADIIISSTASPLPIIGKGM VERALKARRNQPMMLMVDIAVPRDIEPEVGKLANAYL YSVDDLHSIIQNNMAQRKAAAVQAESIVEQESSNFMA WLRSGAVEIIRDYRSRADLVRAEAEAKALAAIAQGA	92

	DVSAVIHEL AHKLTNRLIHAPTRSLQQAASDGDVERL QILRDSLGLDQQ	
glutamyl-tRNA reductase (hemA) <i>Gammaproteobacte ria</i> Accession: WP_193016510.1	MTLLALGINHKTAPVALREKVSFSPDTMGDALNNLLQ QPAVRGGVVLSTCNRTELYLSMEDKENSHEQLIRWLC QYHQIEPNELQSSIWYHQNQAVSHLMRVASGLDSL VLGEPQILGQVKKAFADSQNYDSLSELERLFQKSFSV AKRVRTETQIGANAVSVAFAACTLARQIFESLSSLTILL VGAGETIELVARHLREHQVKKIIIANRTKERAQRLASE VDAEVITLSEIDECLAQADIVISSTASPLPIIGKGMVER ALKKRRNQPMMLVDIAVPRDIEQDVEKLNNVYLYSV DDLEAIIQHNREQRQAAAVQAEHIVQQESGQFMDWL RAQGAVGAIREYRDSAETLRAEMTEKAITLIQNGADA EKVIQQLSHQLMNRLIHTPTKSLQQAASDGDIERLNL RESLGITHN	93
5-aminolevulinic acid synthase (ALAS) <i>Schizophyllum commune H4-8</i> Accession: XP_003036856.1	MGPALDVRGKQLAAGYASVAGQADVEKIHQDQGITI PPNATVEMCPHAKAARDAARIAEDLAAAASKQQPA KKAGGCPFHAAQAQAQAKPAAAPKETVATADKKGK SPRAAGGFDYEKFYEEELDKKHQDKSYRYFNNINRLA ARFPTAHTAKVTDEVEVWCSNDYLGMMGNPVVLET MHRVLDKYGHGAGGTRNIAGNGALHLSLEQELARLH RKEGALVFTSCYVANDATLSTLGSKMPGCVIFS DRMN HASMIQGIRHSGTKKVIFEHNDLADLEKKLAEYPKETP KIIAFESVYSMCGSIGPIKEICDLAEKYGAITFLDEVHA VGLYGPRGAGVAEHLDDYDLHKAAGDSPDAIPGTVMD RVDIITGTLGKSYGAIGGYIAGSARFVDMIRSYAPGFIF TTSLPATVAGAQASVVYQKEYLGDRQLKQVNVREV KRRFAELDIPVVPGPSHIVPVLVGDAALAKQASDKLL AEHDIYVQAINYPTVARGEERLRITVTQRHTLEQMDH LIGAVDQVFNELNINRVQDWKRLGGRASVGVPGGQD FVEPIWTDEQVGLADGSAPLTLRNGQPNEVSHDAVV	94

	AARSRFDWLLGPIPSHIQAKRLGQSLEGTPIAPLAPKQ SSGLKLPVEEMTMGQTIAVAA	
5-aminolevulinic acid synthase (ALAS) <i>Crassisporium funariophilum</i> Accession: KAF8165006.1	MDKIARFKQTCPFLLGRKTNSTLRNLSTSSSPRFPSTAL TERATKCPVMGPALNVRSEIVAGYASVAANSVALI HKEKGVFPPPGATVEMCPHASAARAAARMADDLAA AAEKKKGHFTSAAPRDEAAQAAAAGCPFHVKAAD AAAARKAAAAPAPVKAKEDGGFNYESFYVNELDKK HQDKSYRYFNNINRLAAKFPVAHTSNVKDEVEVWCA NDYLGMMGNPVVLETMHRTLDDKYGHGAGGTRNIAG NGAMHLSLEQELATLHRKPAALVFSSCYVANDATLST LGAKLPGCIIFFSDTMNHASMIQGMRHSGAKRVLFKH NDLEDLENKLKQYPKDTPKVIAFESVYSMCGSIGPIKE ICDLAEQYGALTFLDEVHAVGLYGPARGAGVAEHLDY DAHVAAGESPHPIKGSVMMDRVDIITGTLGKAYGAVGG YIAGSDDFVDMIRSYAPGFITTSPPATVAGARASVV YQKHYVGDRQLKQVNVREVKRRFAELDVPVVPGPSH IVPVLVGDAALAKAASDKLLAEHNIYVQSINYPTVAR GEERLRITVTPRHTEQMDKLVRAVDKIFAELKINRLA DWKALGGRAGVGLTAGAEEAHVDPMWTEEQGLLD GTSPRTLNRGEAAVVDAMAVGQARAVFDNLLGPISG KLQSERSVLASSTPAAANPARPAARKVVKMKTGGVP MSEDIPLPPPDVSASA	95
5-aminolevulinic acid synthase (ALAS) <i>Dendrothele bispora</i> CBS 962.96	MDKLSSLRFRKASCPFLGRKTKSTLRTLCTSSSPRFPSS ILTERATKCPVMGPALNVRSEITAGYASVAGSSEVD QIHKQQGVTPPVNATVEMCPHASAARAAARMADDL AAAAAQKKVGSGASSAKAAAAGCPFHKSVAAAGASA STASKPSAPIHKASVPGGFDYDNFYNNLEKKHKDKS YRYFNNINRLASKFPVAHTGDVKDEVQVWCSNDYLG MGNNPVVLETMHRTLDDKYGHGAGGTRNIAGNGALH LGLAQELAAALHRKEAALVFSSCYVANDATLSTLGSKL	96

Accession: THV05492.1	PGCILFSDKMNHASMIQGMRHSGAKKVIFNHNLEDL ENKLLKQYPKETPKIIAFESVYSMCGSIGPIKEICDLAEK YGALTFLDEVHAVGLYGPHGAGVAEHLDYNAQKAA GKSPEPIPGSVMDRVDIITGTLGKAYGAVGGYIAGSM DFVDTIRSYAPGFIFTTSLPPATVSGAQASVAYQKEYL GDRQLKQVNVREVKKRFAELDIPVIPGSPHILPVLVGD AALAKAASDKLLTDHDIYVQSINYPTVAVGEERLRIT VTPRHTEQMDKLVRAVNQVFTELNINRISDWKVAG GRAGVGMGVESVEPIWTDEQLGITDGTTPKTLRDGQR FLVDAQGVTAARGRFDLTLLGPMMSGSLQANPTLPLVD DELKVPLPTLVAAAA	
5-aminolevulinic acid synthase (ALAS) <i>Bradyrhizobium japonicum</i> Accession: A0A0A3YXD2	MDYAQFFNTALDRLHTERRYRVFADLERIAGRFPHAL WHSPKGKRDVVIWCSNDYLGMGQHPKVVGAMVETA TRVGTGAGGTRNIAGTHHPLVQLEAELADLHGKEASL LFTSGYVSNQTGIATIAKLIPNCLILSDELNHNSMIEGIR QSGCERVVFRHNDLADLEEKLLKAAGPNRPKLIACESL YSMDGDVAPLAKICDLAEKYGAMTYVDEVHAVGMY GPRGGGIAERDGMHRIDILEGTLAKAFGCLGGYIAA NGQIDAVRSYAPGFIFTTALPPAICSAATAAIRHLKTS NWERERHQDRAARVKAILNAAGLPVMSSDTHIVPLFI GDAEKCKQASDLLLEQHGIYIQPINYPTVAKGTERLRI TPSPYHDDGLIDQLAEALLQVWDRLGLPLKQKSLAAE	97
Cytochrome b5 <i>Petunia x hybrida</i> , Accession: AAD10774.1	MDKQRVFTLSQVAEHKSKQDCWIIINGRVVDVTKFLE EHPGGEEVLIESAGKDATKEFQDIGHSAKAKNLLFKY QIGYLQGYKASDDSELELNLVTDSIKEPNKAKEMKAY VIKEDPKPKYLTFVEYLLPFLAAAFYLYRYLTGALQ F	98

Table 12: Glossary of abbreviations

Abbreviation	Full Name
3GT	anthocyanidin-3-O-glycotransferase
4CL	4-coumarate-CoA ligase
ACC	acetyl-CoA carboxylase
ACOT	acyl-CoA thioesterase
acpP	acyl carrier protein
ACS	acetyl-CoA synthase
adhE	aldehyde-alcohol dehydrogenase
ADP	adenosine diphosphate
ALA	5-aminolevulinic acid
ALAS	ALA synthase
ANS	anthocyanin dioxygenase
aroG	DAHP synthase
aroK	shikimate kinase
aroL	shikimate kinase
ATP	adenosine triphosphate
C3G	cyanidin-3-O-glycoside
C4H	cinnamate-4-hydroxylase
CHI	chalcone isomerase
CHS	chalcone synthase
CoA	coenzyme A
CPR	cytochrome P450 Reductase
DAD	diode array detector
DAHP	deoxy-d-arabino-heptulosonate-7-phosphate
DctPQM	a malonate transporter
DFR	dihydroflavonol 4-reductase
DHK	dihydrokaempferol

DHM	dihydromyricetin
DHQ	dihydroquercetin
DMSO	dimethyl sulfoxide
E4P	erythrose-4-phosphate
F3'H	flavonoid 3' hydroxylase
F3H	flavanone 3-hydroxylase
fabB	beta-ketoacyl-ACP synthase I
fabD	malonyl-coA-ACP transacylase
fabF	beta-ketoacyl-ACP synthase II
FadA	3-ketoacyl-CoA thiolase
FadB	fatty acid oxidation complex subunit alpha
FadE	acyl-CoA dehydrogenase
GltX	glutamyl-tRNA synthetase
hemA	glutamyl-tRNA reductase
hemL	glutamate-1-semialdehyde aminotransferase
HPLC	high performance liquid chromatography
ldhA	lactate dehydrogenase
LAR	leucoanthocyanidin reductase
matB	malonyl-CoA synthase
matC	malonate transporter
mdcA	malonate coA-transferase
mdcC	acyl-carrier protein, subunit of mdc
mdcD	malonyl-CoA decarboxylase, subunit of mdc
mdcE	co-decarboxylase, subunit of mdc
pABA	para-aminobenzoic acid
PAL	phenylalanine ammonia-lyase
PanK	pantothenase kinase
Pdh	pyruvate dehydrogenase
PEP	phosphoenolpyruvate

pHBA	para-hydroxybenzoic acid
PHE	phenylalanine
pheA	chorismate mutase / prephenate dehydrogenase
poxB	pyruvate dehydrogenase
ppsA	phosphoenolpyruvate synthase
TAL	tyrosine ammonia-lyase
TCA	tricarboxylic acid cycle
tesA	thioesterase I
tesB	thioesterase II
tktA	transketolase
TRP	tryptophan
TYR	tyrosine
TyrA	chorismate mutase
tyrR	transcriptional regulator
ybgC	a thioesterase
yciA	a thioesterase
ydiB	QUIN/shikamate dehydrogenase
ackA-pta	Acetate kinase-phosphate acetyltransferase

CLAIMS

1. An engineered host cell that comprises one or more genetic modifications resulting in production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical intermediates, by the engineered host cell.
5
2. The engineered host cell of claim 1, wherein the production of flavonoid or anthocyanin from a carbon source that can also be an energy source occurs through enzymatic transformation.
- 10 3. The engineered host cell of claim 1, wherein the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol; and (vi) any combination thereof.
- 15 4. The engineered host cell of claim 1, wherein the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.
- 20 5. The engineered host cell of claim 1, wherein the one or more genetic modifications lead to increase in metabolic flux to flavonoid precursors or cofactors.
6. The engineered host cell of claim 1, wherein the one or more genetic modifications cause reduction of formation of byproducts.
- 25 7. The engineered host cell of claim 1, wherein the one or more genetic modifications are at least one genetic modification selected from the group consisting of: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing
30 one or more non-native genes in the engineered host cells; and (iv) a combination thereof.

8. The engineered host cell of claim 1, wherein the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid.
- 5 9. The engineered host cell of claim 1, wherein the engineered host cell comprises at least one or more nucleic acid sequences selected from the group consisting of: (i) nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof.
- 10 10. The engineered host cell of claim 1, wherein the engineered host cell comprises at least one or more peptides selected from the group consisting of: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof.
- 15 11. The engineered host cell of claim 1, wherein the engineered host cell is *E. coli*.
- 20 12. The engineered host cell of claim 1, wherein the one or more genetic modifications decrease fatty acid biosynthesis.
- 25 13. The engineered host cell of claim 1, wherein the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and
- 30

p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof.

5

14. The engineered host cell of claim 1, wherein the flavonoid is catechin.

10

15. A method of increasing the production of flavonoids or anthocyanins, the method comprising: providing an engineered host cell that comprises one or more genetic modifications resulting in production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical intermediates, by the engineered host cell.

15

16. The method of claim 15, wherein the production of flavonoid or anthocyanin from glycerol occurs through enzymatic transformation.

20

17. The method of claim 15, wherein the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.

25

18. The method of claim 15, wherein the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.

30

19. The method of claim 15, wherein the one or more genetic modifications lead to increase in metabolic flux to flavonoid precursors or cofactors.

20. The method of claim 15, wherein the one or more genetic modifications cause reduction in the production of byproducts.

21. The method of claim 15, wherein the one or more genetic modifications are at least one genetic modification selected from the group consisting of: (i) one or more

- modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification for expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof.
- 5
22. The method of claim 15, wherein the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid.
- 10
23. The method of claim 15, wherein the engineered host cell comprises at least one or more nucleic acid sequences selected from the group consisting of: (i) nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof.
- 15
24. The method of claim 15, wherein the engineered host cell comprises at least one or more peptides selected from the group consisting of: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof.
- 20
25. The method of claim 15, wherein the engineered host cell is *E. coli*.
- 25
26. The method of claim 15, wherein the one or more genetic modifications decrease fatty acid biosynthesis.
27. The method of claim 15, wherein the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase
- 30

- converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof.
- 5
- 10
28. The method of claim 15, wherein the flavonoid is catechin.
29. A plurality of engineered host cells, wherein each of the plurality of the engineered host cells comprises one or more genetic modifications resulting in production of flavonoid or anthocyanin from a carbon source that can also be an energy source, through multiple chemical intermediates, by the engineered host cells.
- 15
30. The plurality of engineered host cells of claim 29, wherein the production of flavonoid or anthocyanin from glycerol occurs through enzymatic transformation.
- 20
31. The plurality of engineered host cells of claim 29, wherein the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.
- 25
32. The plurality of engineered host cells of claim 29, wherein the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.
- 30
33. The plurality of engineered host cells of claim 29, wherein the one or more genetic modifications lead to increase in metabolic flux to flavonoid precursors or cofactors.

34. The plurality of engineered host cells of claim 29, wherein the one or more genetic modifications cause a reduction in the formation of byproducts.
- 5 35. The plurality of engineered host cells of claim 29, wherein the one or more genetic modifications are at least one genetic modification selected from the group consisting of: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic
10 modification is expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof.
36. The plurality of engineered host cells of claim 29, wherein at least one of the engineered host cell is cultured in a medium comprising molecules selected from a
15 group consisting of tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid.
37. The plurality of engineered host cells of claim 29, wherein at least one of the engineered host cell comprises at least one or more nucleic acid sequences selected from the group consisting of: (i) nucleic acid sequences encoding tyrosine ammonia
20 lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any
25 combination thereof.
38. The plurality of engineered host cells of claim 29, wherein at least one of the engineered host cell comprises at least one or more peptides selected from the group consisting of: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein
30 comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof.

39. The plurality of engineered host cells of claim 29, wherein at least one of the engineered host cell is *E. coli*.
40. The plurality of engineered host cells of claim 29, wherein the one or more genetic
5 modifications decrease fatty acid biosynthesis.
41. The plurality of engineered host cells of claim 29, wherein at least one of the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase,
10 wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv)
15 nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms
20 naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof.
42. The plurality of engineered host cells of claim 29, wherein the flavonoid is catechin.
25
43. A method of increasing the production of flavonoids or anthocyanins, the method comprising: providing a plurality of engineered host cells, wherein each of the plurality of the engineered host cell comprises one or more genetic modifications that resulting in production of flavonoid or anthocyanin from a carbon source that can also be an
30 energy source, through multiple chemical intermediates, through multiple chemical intermediates, by the engineered host cell.

44. The method of claim 43, wherein the one or more genetic modifications lead to increase in metabolic flux to flavonoid precursors.
45. The method of claim 43, wherein the carbon source is selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.
46. The method of claim 43, wherein the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of: (i) glycerol, (ii) a sugar, (iii) an organic acid, (iv) an amino acid, (v) a biomass comprising glycerol, and (vi) any combination thereof.
47. The method of claim 43, wherein the one or more genetic modifications lead to increase in metabolic flux to flavonoid precursors or cofactors.
48. The method of claim 43, wherein the one or more genetic modifications cause reduction in the production of byproducts.
49. The method of claim 43, wherein the one or more genetic modifications are at least one genetic modification selected from the group consisting of: (i) one or more modifications for over-expressing one or more endogenous genes in the engineered host cells; (ii) one or more modifications for under-expressing one or more endogenous genes in the engineered host cells; (iii) one or more genetic modification is expressing one or more non-native genes in the engineered host cells; and (iv) a combination thereof.
50. The method of claim 43, wherein at least one of the engineered host cell is cultured in a medium comprising molecules selected from a group consisting of tyrosine, phenylalanine, malonate, p-coumarate, bicarbonate, acetate, pantothenate, biotin, thiamine, alpha-ketoglutarate, ascorbate, and 5-aminolevulinic acid.
51. The method of claim 43, wherein at least one of the engineered host cell comprises at least one or more nucleic acid sequences selected from the group consisting of: (i)

nucleic acid sequences encoding tyrosine ammonia lyase activity; (ii) nucleic acid sequences encoding phenylalanine ammonia lyase activity; (iii) nucleic acid sequences encoding cinnamate 4-hydroxylase activity; (iv) nucleic acid sequences encoding 4-coumarate-CoA ligase (4CL) activity; and (v) any combination thereof.

5

52. The method of claim 43, wherein at least one of the engineered host cell comprises at least one or more peptides selected from the group consisting of: (i) chalcone isomerase; (ii) chalcone synthase; (iii) a fusion protein comprises a chalcone synthase and a chalcone isomerase; and (iv) any combination thereof.

10

53. The method of claim 43, wherein at least one of the engineered host cell is *E. coli*.

54. The method of claim 43, wherein the one or more genetic modifications decreases fatty acid biosynthesis.

15

55. The method of claim 43, wherein at least one of the engineered host cell comprises an exogenous nucleic acid sequence selected from the group consisting of: (i) nucleic acid sequence encoding tyrosine ammonia lyase, wherein the encoded tyrosine ammonia lyase forms 4-coumaric acid using tyrosine as a substrate; (ii) nucleic acid sequence encoding phenylalanine ammonia lyase, wherein the encoded phenylalanine ammonia lyase converts phenylalanine to trans-cinnamic acid; (iii) nucleic acid sequence encoding cinnamate-4-hydroxylase, wherein the cinnamate-4-hydroxylase produces 4-coumaric acid from trans-cinnamic acid; (iv) nucleic acid sequence encoding 4-coumarate-CoA ligase activity, wherein 4-coumarate-CoA ligase forms p-coumaroyl-CoA from coumaric acid (v) nucleic acid sequence encoding chalcone synthase activity, wherein chalcone synthase forms naringenin chalcone from malonyl-CoA and p-coumaroyl-CoA; (vi) nucleic acid sequence encoding chalcone isomerase activity, wherein chalcone isomerase forms naringenin from naringenin chalcone; (vii) nucleic acid sequence encoding flavanone-3-hydroxylase, wherein flavanone-3-hydroxylase forms dihydrokaempferol from naringenin; and (viii) any combinations thereof.

30

56. The method of claim 43, wherein the flavonoid is catechin.

57. An engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase the production and/or availability of malonyl-CoA.
- 5 58. The engineered host cell of claim 57, wherein the production and/or availability of malonyl-CoA is increased by transformation of acetyl-CoA to malonyl-CoA.
59. The engineered host cell of claim 57, wherein the engineered host cell comprises one or more genetic modifications selected from a group consisting of: (i) expression of
10 acetyl-CoA carboxylase (ACC); and (ii) overexpression of acetyl-CoA carboxylase.
60. The engineered host cell of claim 57, wherein the engineered host cell is *E. coli*.
61. The engineered host cell of claim 60, wherein the *E. coli* further comprises genes from
15 fungi.
62. The engineered host cell of claim 59, wherein the acetyl-CoA carboxylase is from a species selected from the group consisting of *Mucor circinelloides*, *Rhodotorula toruloides*, *Lipomyces starkeyi*, and *Ustilago maydis*, and orthologs of acetyl-CoA
20 carboxylase having at least 50% amino acid identity to the acetyl-CoA carboxylase of these aforementioned species.
63. The engineered host cell of claim 57, wherein the one or more genetic modification is deletion or attenuation of one or more fatty biosynthetic genes resulting in decrease in
25 fatty acid biosynthesis.
64. The engineered host cell of claim 57, wherein the one or more genetic modification is overexpression of acetyl-CoA synthase (ACS).
- 30 65. The engineered host cell of claim 64, wherein the acetyl-CoA synthase is selected from the group consisting of, acetyl-CoA synthase gene of *E. coli*, acetyl-CoA synthase gene of *Salmonella typhimurium*, and orthologs of acetyl-CoA synthase gene in any other

species having at least 50% amino acid identity to the acetyl-CoA synthase gene of *E. coli* and *Salmonella typhimurium*.

- 5 66. The engineered host cell of claim 57, wherein the one or more genetic modification is selected from a group consisting of: (i) overexpression a gene encoding pyruvate dehydrogenase (PDH), wherein the PDH may include E354K mutation; (ii) exogenous nucleic acid sequence encoding a malonyl-CoA synthetase; (iii) upregulation of endogenous pantothenate kinase (PanK), wherein PanK is not feedback inhibited by coenzyme A; (iv) exogenous nucleic acid sequence encoding a malonate transporter; 10 and (v) any combinations thereof.
67. The engineered host cell of claim 66, wherein the malonyl-CoA synthetase is selected from the group consisting of malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a malonyl-CoA synthetase having at least 50% 15 identity to any of these or other naturally occurring malonyl-CoA synthetases.
68. The engineered host cell of claim 63, wherein the wherein one or more genetic modifications to decrease fatty acid biosynthesis is selected from the group consisting of: (i) mutation or downregulation of a gene encoding malonyl-CoA-ACP transacylase 20 (*E. coli* fabD); (ii) modifications to the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF); (iii) downregulation of beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB); (iv) downregulation of acyl carrier protein (*E. coli* acpP); and (v) any combinations thereof.
69. The engineered host cell of claim 57, wherein the engineered host cell comprises 25 peptides selected from a group consisting of: (i) acetyl-CoA carboxylase (ACC) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 15 or SEQ ID NO: 16; (ii) malonate CoA-transferase having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 19; (iii) acetyl-CoA synthase (ACS) having an amino acid sequence at least 80% identical to the 30 polypeptide set forth in SEQ ID NO: 16; (iv) malonyl-CoA synthase having an amino acid sequence at least 80% identical SEQ ID NO: 77, SEQ ID NO: 78, or SEQ ID NO: 79; (v) malonate transporter having an amino acid sequence at least 80% identical to

SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 83, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 86, or SEQ ID NO: 87; (vi) pantothenate kinase having an amino acid sequence at least 80% identical to SEQ ID NO: 88, SEQ ID NO: 89, or SEQ ID NO: 90; and (vii) any combinations thereof.

5

70. A method of increasing the production of flavonoids comprising an engineered host cell, wherein the one or more engineered host cells comprise one or more genetic modifications to increase the production and/or availability of malonyl-CoA.

10

71. The method of claim 70, wherein the production and/or availability of malonyl-CoA is increased by transformation of acetyl-CoA to malonyl-CoA.

15

72. The method of claim 70, wherein the engineered host cell comprises one or more genetic modifications selected from a group consisting of: (i) expression of acetyl-CoA carboxylase (ACC); and (ii) overexpression of acetyl-CoA carboxylase.

73. The method of claim 70, wherein the engineered host cell is *E. coli*.

74. The method of claim 73, wherein the *E. coli* further comprises genes from fungi.

20

75. The method of claim 70, wherein the acetyl-CoA carboxylase is from a species selected from the group consisting of *Mucor circinelloides*, *Rhodotorula toruloides*, *Lipomyces starkeyi*, and *Ustilago maydis*, and orthologs of acetyl-CoA carboxylase having at least 50% amino acid identity to the acetyl-CoA carboxylase of these aforementioned species.

25

76. The method of claim 70, wherein the one or more genetic modification is deletion or attenuation of one or more fatty biosynthetic genes resulting in decrease in fatty acid biosynthesis.

30

77. The method of claim 70, wherein the one or more genetic modification is overexpression of acetyl-CoA synthase (ACS).

78. The method of claim 77, wherein the acetyl-CoA synthase (ACS) is selected from the group consisting of, acetyl-CoA synthase gene of *E. coli*, acetyl-CoA synthase gene of *Salmonella typhimurium*, and orthologs of acetyl-CoA synthase gene in any other species having at least 50% amino acid identity to the acetyl-CoA synthase gene of *E. coli* and *Salmonella typhimurium*.
79. The method of claim 14, wherein the one or more genetic modification is selected from a group consisting of: (i) overexpression a gene encoding pyruvate dehydrogenase (PDH), wherein the PDH may include E354K mutation; (ii) exogenous nucleic acid sequence encoding a malonyl-CoA synthetase; (iii) upregulation of endogenous pantothenate kinase (PanK), wherein PanK is not feedback inhibited by coenzyme A; (iv) exogenous nucleic acid sequence encoding a malonate transporter; and (v) any combinations thereof.
80. The method of claim 79, wherein the malonyl-CoA synthetase is selected from the group consisting of malonyl-CoA synthetases of *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, or a malonyl-CoA synthetase having at least 50% identity to any of these or other naturally occurring malonyl-CoA synthetases.
81. The method of claim 76, wherein the wherein one or more genetic modifications to decrease fatty acid biosynthesis is selected from the group consisting of: (i) mutation or downregulation of a gene encoding malonyl-CoA-ACP transacylase (*E. coli* fabD); (ii) modifications to the gene beta-ketoacyl-ACP synthase II (*E. coli* fabF); (iii) downregulation of beta-ketoacyl-ACP synthase I enzyme (*E. coli* fabB); (iv) downregulation of acyl carrier protein (*E. coli* acpP); and (v) any combinations thereof.
82. The method of claim 70, wherein the engineered host cell comprises peptides selected from a group consisting of: (i) acetyl-CoA carboxylase (ACC) having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 15 or SEQ ID NO: 16; (ii) malonate CoA-transferase having an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO: 19; (iii) acetyl-CoA synthase (ACS) having an amino acid sequence at least 80% identical to the polypeptide set forth

- in SEQ ID NO: 16; (iv) malonyl-CoA synthase having an amino acid sequence at least 80% identical SEQ ID NO: 77, SEQ ID NO: 78, or SEQ ID NO: 79; (v) malonate transporter having an amino acid sequence at least 80% identical to SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 82, SEQ ID NO: 83, SEQ ID NO: 84, SEQ ID NO: 85, SEQ ID NO: 86, or SEQ ID NO: 87; (vi) pantothenate kinase having an amino acid sequence at least 80% identical to SEQ ID NO: 88, SEQ ID NO: 89, or SEQ ID NO: 90; and (vii) any combinations thereof.
- 5
83. An engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase endogenous biosynthesis of tyrosine.
- 10
84. The engineered host cell of claim 83, wherein the one or more genetic modifications comprises upregulation of 3-deoxy-D-arabino-heptulosonate synthase.
85. The engineered host cell of claim 83, wherein the one or more genetic modifications are selected from a group consisting of: (i) upregulation of chorismate mutase; (ii) upregulation of prephenate dehydrogenase; (iii) overexpression of shikimate kinase; (iv) overexpression of shikimate dehydrogenase; and (v) any combinations thereof.
- 15
86. The engineered host cell of claim 83, wherein the one or more genetic modifications comprises downregulation of L-phenylalanine biosynthetic pathway.
- 20
87. The engineered host cell of claim 83, wherein the one or more genetic modifications comprises expression of exogenous phosphoenolpyruvate synthase (ppsA).
- 25
88. The engineered host cell of claim 83, wherein the one or more genetic modifications comprises expression of exogenous transketolase (tktA).
89. The engineered host cell of claim 83, wherein the one or more genetic modifications comprises disruption of tyrR gene.
- 30
90. The engineered host of claim 83, wherein the one or more genetic modifications are selected from a group consisting of: (i) expression or overexpression of (D146N) variant of phospho-2-dehydro-3-deoxyheptonate aldolase; (ii) expression or

overexpression of variant of 3-dehydroquinate synthase (aroB); (iii) overexpression of transketolase tktA; (iv) deletion of shikimate kinase (aroK); (v) deletion of tyrR; (vi) expression or overexpression of A354V variant of chorismate mutase (tyrA); (vi) and any combination thereof.

5

91. A method of increasing endogenous biosynthesis of tyrosine comprising an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase endogenous biosynthesis of tyrosine.

10

92. The method of claim 91, wherein the one or more genetic modifications comprises upregulation of 3-deoxy-D-arabino-heptulosonate synthase.

15

93. The method of claim 91, wherein the one or more genetic modifications are selected from a group consisting of: (i) upregulation of chorismate mutase; (ii) upregulation of prephenate dehydrogenase; (iii) overexpression of shikimate kinase; (iv) overexpression of shikimate dehydrogenase; and (v) any combinations thereof.

20

94. The method of claim 91, wherein the one or more genetic modifications comprises downregulation of L-phenylalanine biosynthetic pathway.

95. The method of claim 91, wherein the one or more genetic modifications comprises expression of exogenous phosphoenolpyruvate synthase (ppsA).

25

96. The method of claim 91, wherein the one or more genetic modifications comprises expression of exogenous transketolase (tktA).

97. The method of claim 91, wherein the one or more genetic modifications comprises disruption of tyrR gene.

30

98. The method of claim 91, wherein the one or more genetic modifications are selected from a group consisting of: (i) expression or overexpression of (D146N) variant of phospho-2-dehydro-3-deoxyheptonate aldolase; (ii) expression or overexpression of variant of 3-dehydroquinate synthase (aroB); (iii) overexpression of transketolase tktA;

(iv) deletion of shikimate kinase (aroK); (v) deletion of tyrR; (vi) expression or overexpression of A354V variant of chorismate mutase (tyrA); (vi) and any combination thereof.

- 5 99. An engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G).
- 10 100. The engineered host cell of claim 99, wherein the one or more genetic modifications comprising overexpression of anthocyanin synthase.
- 15 101. The engineered host cell of claim 100, wherein the anthocyanin synthase is selected from a group consisting of: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof.
- 20 102. The engineered host cell of claim 100, wherein the engineered host cell further comprises flavonoid-3-glucosyl transferase (3GT).
- 25 103. The engineered host cell of claim 102, wherein the flavonoid-3-glucosyl transferase is selected from a group consisting of: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof.
- 30 104. A method of increasing the production of flavonoids comprising an engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G).

105. The method of claim 104, wherein the one or more genetic modifications comprising overexpression of anthocyanin synthase.
106. The method of claim 104, wherein the anthocyanin synthase is selected from a group consisting of: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof.
107. The method of claim 105, wherein the engineered host cell further comprises flavonoid-3-glucosyl transferase (3GT).
108. The method of claim 106, wherein the flavonoid-3-glucosyl transferase is selected from a group consisting of: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof.
109. A method of increasing the transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G) comprising anthocyanin synthase, wherein the anthocyanin synthase is selected from a group consisting of: (i) anthocyanin synthase of *Carica papaya* (SEQ. ID NO:13); (ii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 66, SEQ. ID NO: 67, SEQ. ID NO: 68, or SEQ. ID NO: 69; (iii) the anthocyanin synthase has an amino acid sequence at least 80% identical to SEQ. ID NO: 13; and (iv) any combinations thereof.
110. A method of increasing the transformation of leucocyanidin or catechin to cyanidin-3-glucoside (Cy3G) comprising flavonoid-3-glucosyl transferase (3GT), wherein the flavonoid-3-glucosyl transferase is selected from a group consisting of: (i) flavonoid-3-glucosyl transferase in *Vitis labrusca* (SEQ. ID NO:14); (ii) the flavonoid-3-glucosyl

transferase has an amino acid sequence at least 80% identical to SEQ. ID NO: 70, SEQ. ID NO: 71, SEQ. ID NO: 72, or SEQ. ID NO: 73; and (iii) any combinations thereof.

- 5 111. An engineered host cell, wherein the engineered host cell comprises one or more genetic modifications to increase the production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictoyl (EDL), and/or pentahydroxyflavone (PHF), wherein the engineered host cell comprises cytochrome P450 reductase (CPR) and at least one of flavanone-3-hydroxylase (F3H), flavanone-3'-hydroxylase (F3'H), or flavonoid 3',5'-hydroxylase (F3'5'H).
- 10 112. The cell of claim 111, wherein the precursor for increase in production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictoyl (EDL), and/or pentahydroxyflavanone (PHF) is naringenin and/or dihydrokaempferol (DHK).
- 15 113. The engineered host cell of claim 111, wherein the flavonoid 3'-hydroxylase (F3'H) is truncated to remove the N-terminal leader sequence.
114. The engineered host cell of claim 111, wherein the flavonoid 3',5'-hydroxylase (F3'5'H) is truncated to remove the N-terminal leader sequence.
- 20 115. The engineered host cell of claim 111, wherein the cytochrome P450 reductase (CPR) is truncated to remove the N-terminal leader sequence.
116. The engineered host cell of claim 111, wherein the flavonoid 3'-hydroxylase (F3'H) is fused with cytochrome P450 reductase (CPR).
- 25 117. The engineered host cell of claim 111, wherein the flavonoid 3',5'-hydroxylase (F3'5'H) is fused with cytochrome P450 reductase (CPR).
- 30 118. The engineered host cell of claim 111, wherein the flavanone-3'-hydroxylase (F3'H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 8.

119. The engineered host cell of claim 111, wherein the cytochrome P450 reductase (CPR) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 9.
- 5 120. The engineered host cell of claim 111, wherein the flavonoid 3',5'-hydroxylase (F3'5'H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 10, (ii) SEQ ID NO. 56, and (iii) SEQ ID NO. 57.
- 10 121. The engineered host cell of claim 111, wherein the engineered host cell further comprises cytochrome b₅.
122. The engineered host cell of claim 121, wherein the cytochrome b₅ has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 98.
- 15 123. The engineered host cell of claim 111, wherein the flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 7, (ii) SEQ ID NO. 45, (iii) SEQ ID NO. 46, (iv) SEQ ID NO. 47, and (v) SEQ ID NO. 48.
- 20 124. A method of increasing the production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictoyl (EDL), and/or pentahydroxyflavone (PHF) comprising an engineered host cell, wherein the engineered host cell comprises cytochrome P450 reductase (CPR) and at least one of flavanone-3-hydroxylase (F3H), flavanone-3'-hydroxylase (F3'H), or flavonoid 3',5'-hydroxylase (F3'5'H).
- 25 125. The method of claim 124, wherein the precursor for increase in production of dihydroquercetin (DHQ), dihydromyricetin (DHM), eriodictoyl (EDL), and/or pentahydroxyflavone (PHF) are naringenin and/or dihydrokaempferol (DHK).
- 30 126. The method of claim 124, wherein the flavonoid 3'-hydroxylase (F3'H) is truncated to remove the N-terminal leader sequence.

127. The method of claim 124, wherein the flavonoid 3',5'-hydroxylase (F3'5'H) is truncated to remove the N-terminal leader sequence.
128. The method of claim 124, wherein the cytochrome P450 reductase (CPR) is truncated to remove the N-terminal leader sequence.
129. The method of claim 124, wherein the flavonoid 3'-hydroxylase (F3'H) is fused with cytochrome P450 reductase (CPR).
130. The method of claim 124, wherein the flavonoid 3',5'-hydroxylase (F3'5'H) is fused with cytochrome P450 reductase (CPR).
131. The method of claim 124, wherein the flavanone-3'-hydroxylase (F3'H) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 8.
132. The method of claim 124, wherein the cytochrome P450 reductase (CPR) has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 9.
133. The method of claim 124, wherein the flavonoid 3',5'-hydroxylase (F3'5'H) has an amino acid sequence at least 80% identical to the polypeptides selected from the group consisting of: (i) SEQ ID NO. 10, (ii) SEQ ID NO. 56, and (iii) SEQ ID NO. 57.
134. The method of claim 124, wherein the engineered host cell further comprises cytochrome b₅.
135. The method of claim 134, wherein the cytochrome b₅ has an amino acid sequence at least 80% identical to the polypeptide set forth in SEQ ID NO. 98.
136. The method of claim 124, wherein the flavanone-3-hydroxylase (F3H) has an amino acid sequence at least 80% identical to the polypeptides selected from a group consisting of: (i) SEQ ID NO. 7, (ii) SEQ ID NO. 45, (iii) SEQ ID NO. 46, (iv) SEQ ID NO. 47, and (v) SEQ ID NO. 48.