

PRODUCTION OF CAROTENOIDS AND APOCAROTENOIDS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of Singapore application No. 10201701500U, filed 24 February 2017, the contents of it being hereby incorporated by reference in its entirety for all purposes.

FIELD OF THE INVENTION

[0002] The invention is in the field of carotenoids and apocarotenoids, in particular, the invention relates to improved methods for producing apocarotenoids and carotenoids.

BACKGROUND OF THE INVENTION

[0003] Carotenoids are one group of natural products synthesized in many plants, algae and certain bacteria and fungi. Carotenoids have many unsaturated carbon bounds and these bounds are conjugated, which contribute to distinct features, the bright colors (ranging from pale yellow to orange to red) and potent anti-UV and antioxidant effects. Because of these features, carotenoids have been widely used as natural colorings and nutraceuticals. Particularly, phytoene has effective UVB- absorbing capability and reduces the melanin synthesis in human skin, thus with a growing market in cosmetics. Lycopene, β -carotene and α -carotene are well-known for their antioxidant effects and have been widely used in food, cosmetics, nutraceutical and animal feeds products.

[0004] Apocarotenoids are a class of compounds derived from carotenoids by carotenoid cleavage oxygenases. Widely distributed in bacteria, fungi, plants and animals, apocarotenoids act as aroma and scent compounds (α - and β -ionone), photosensory pigment (bixin, crocin), hormones (abscisic acid) and signalling compounds (strigolactones). Among various apocarotenoids, α - and β -ionone are two important aromatic compounds. α -ionone has a sweet and violet-like aroma with odour threshold of ~0.4 ppb. Its isomer, β -ionone, has a warm, woody, and violet aroma and an even lower odour threshold of ~0.007 ppb in air and 1 ppb in water. Due to their significantly low odour threshold and pleasant smell, they are widely exploited in cosmetics and perfume industry. Besides ionones, retinol (or vitamin A) is another commercially important apocarotenoid. Retinol plays an essential function in vision, bone development and skin health as antioxidants. As active cosmetic ingredients and

effective medicines for skin diseases, the retinol market size is estimated to be about 1.6 billion dollars.

[0005] Despite their high commercial values, supply of natural ionones and retinoids is severely limited by their extremely low abundance in nature. Ionones are present in sub-ppm
5 level in many flowers and fruits, such as rose, sweet osmanthus, orris root and raspberry. For instance, it requires 100 tons of raspberries, or 20 hectares of agricultural area, to produce merely 1 gram of α -ionone. As for retinoids, there is no natural source from plants. Exceptionally low amount of retinoids exists in some animal-derived food, such as eggs and butter. Hence, the current supply through extraction from natural sources cannot meet the
10 increasing demands for natural ionones and retinoids. Although these compounds could be chemically synthesized, apocarotenoids such as α -ionone have chiral centres, and synthetic ones are usually a mixture of different enantiomers. The different isomers of many fragrance compounds are known to have different odors, thus it is important to synthesize single isomer instead of a mixture of isomers. More importantly, consumers tend to prefer natural to
15 synthetic flavours and thus natural ingredients have significantly higher prices. The production of α -ionone was previously demonstrated in engineered *Escherichia coli*, but with a very low yield.

[0006] Thus, there is a need to provide naturally produced compounds with improved yields.

20

SUMMARY

[0007] In one aspect, there is provided a method for producing a carotenoid or apocarotenoid comprising the step of expressing in a host cell an expression module comprising an expression vector having a coding region encoding at least one optimised carotenoid or apocarotenoid generating enzyme, the coding region being operably linked to a
25 promoter.

[0008] In one aspect, there is provided a host cell comprising an expression vector having a coding region encoding at least one optimised carotenoid or apocarotenoid generating enzyme, the coding region being operably linked to a promoter.

[0009] In one aspect, there is provided a kit when used in the method as described herein,
30 for the production of a carotenoid comprising one or more of: a first vector encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR*; a second vector encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi*

operably linked to a promoter; a third vector encoding one or more optimised third gene products selected from *ispA*, *crtE* or *crtB*, optionally *crtI*, operably linked to a promoter for the production of phytoene or lycopene.

[0010] In one aspect, there is provided a system for producing a carotenoid or apocarotenoid comprising an expression vector having a coding region encoding at least one optimised carotenoid or apocarotenoids generating enzyme, the coding region being operably linked to a promoter, wherein said at least one optimised carotenoid or apocarotenoids generating enzyme is selected from: a. *ΔN50-LsLcyE* and TrxA-CCD1 (preferably TrxA-*Osmanthus fragans* CCD1) for the production of α -ionone; or b. *crtY* and CCD1 (preferably *Petunia hybrid* CCD1) for the production of β -ionone; or c. *ΔN50-LsLcyE* for the production of ϵ -carotene; or d. *crtY* and *blh* for the production of retinal or retinol.

[0011] In one aspect, there is provided a kit when used in the method as described herein, for the production of a carotenoid or apocarotenoid comprising one or more of: a first vector encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY* operably linked to a promoter; a second vector encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi* operably linked to a promoter; a third vector encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI* operably linked to a promoter; and a fourth vector encoding one or more optimised gene products selected from: a. *ΔN50-LsLcyE* and TrxA-*Osmanthus fragans* CCD1 operably linked to a promoter for the production of α -ionone; or b. *crtY* and *phCCD1* operably linked to a promoter for the production of β -ionone; or c. *ΔN50-LsLcyE* operably linked to a promoter for the production of ϵ -carotene; or d. *crtY* and *blh* operably linked to a promoter for the production of retinal or retinol.

[0012] In one aspect, a method for producing a carotenoid or apocarotenoid comprising the steps of: a. contacting a host cell as described herein in a chemically defined media with a substrate for apocarotenoid or carotenoid production; b. incubating the host cell in said chemically defined media to produce one or more preselected carotenoid or apocarotenoid, and c. extracting the one or more preselected carotenoid or apocarotenoids from the chemically defined media using an organic layer.

DEFINITIONS

[0013] As used herein the term “coding region”, also known as the coding sequence or CDS (from coding DNA sequence), is that portion of DNA or RNA, composed of exons, that codes for protein.

- 5 [0014] As used herein an “operon” refers to group of genes or a segment of DNA that functions as a single transcription unit. It may be comprised of an operator, a promoter, and one or more structural genes that are transcribed into one polycistronic mRNA.

BRIEF DESCRIPTION OF THE DRAWINGS

- 10 [0015] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying drawings, in which:

- [0016] **Fig. 1** is a diagram of the ‘plug-n-play’ platform for the biosynthesis of carotenes and apocarotenoids. Feedstock is inexpensive sugars (such as glucose) or glycerol. The platform could produce various carotenes (phytoene, lycopene, α -carotene, β -carotene, δ -carotene and ϵ -carotene) and apocarotenoids (α -ionone, β -ionone, psi-ionone, retinol, dialdehydes (C17 and C19), geranylacetone, β -cyclocitral and 6-methyl-5-hepten-2-one (MHO)).

- [0017] **Fig. 2** is a schematic representation of expression modules for carotene synthesis.
- 20 The upstream MVA pathway (expression module 1), the downstream MVA pathway (module 2), the phytoene or lycopene pathway (expression module 3), the carotene pathway (expression module 4). The genes expressed encode the following enzymes: *atoB*, Acetoacetyl-CoA thiolase; *hmgS*, HMG-CoA synthase; *thmgR*, truncated HMG-CoA reductase; *mevk*, mevalonate kinase; *pmk*, phosphomevalonate kinase; *pmd*, mevalonate pyrophosphate decarboxylase; *idi*, IPP isomerase; *ispA*, FPP synthase; *crtE*, GGPP synthase;
- 25 *crtB*, phytoene synthase; *crtI*, phytoene desaturase; *crtY*, lycopene-beta-cyclase; *LCYe*, lycopene-epsilon-cyclase. Abbreviation for the compounds: G3P, D-glyceraldehyde-3-phosphate; DHAP, Dihydroxyacetone phosphate; HMG-CoA, 3-hydroxy-3-methyl-glutaryl-coenzyme A; MVA, mevalonate; MVAP, phosphomevalonate; MVAPP, diphosphomevalonate;
- 30 IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; HMBPP, (E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate; GPP, geranyl

pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate. Dashed arrow indicates multiple enzymatic steps.

[0018] **Fig. 3** shows the UV absorbance spectrum of various carotenoids. The UV-B and UV-A wavelength is indicated.

5 [0019] **Fig. 4** shows use of an experimental design-aided systematic pathway optimization (EDASPO) method to optimize phytoene production by controlling SAR, MPPI and EBIA modules with T7 promoter variants. The number represents the different strength of T7 promoter variants, with 1, 2 and 3 representing approximately 92%, 37% and 16% of native T7 promoter strength, respectively. Expression module 1, or SAR module, *hmgS-atoB-hmgR*;
10 Expression module 2, or MPPI module, *mevK*, *pmk*, *pmd* and *idi*; Expression module 3, or EBA module, phytoene biosynthetic pathway (*crtEB* and *ispA*).

[0020] **Fig. 5** shows the use of EDASPO method to optimize lycopene production by controlling SAR, MPPI and EBIA expression modules with T7 promoter variants. The number represents the different strength of T7 promoter variants, with 1, 2 and 3 representing
15 approximately 92%, 37% and 16% of native T7 promoter strength, respectively. Expression module 1, or SAR module, *hmgS-atoB-hmgR*; Expression module 2, or MPPI module, *mevK*, *pmk*, *pmd* and *idi*; Expression module 3, or EBIA module, lycopene biosynthetic pathway (*crtEBI* and *ispA*).

[0021] **Fig. 6** shows the production of ϵ -carotene in the lycopene accumulating *E. coli*
20 strain. Carotenoid production of strains expressing different forms of LCYe enzymes. An illustration of promoters and regulated genes for different strains. Different N-terminal truncated LCYe from *Lactuca sativa* was tested. The abbreviations are as follows. LL, wild-type LCYe enzyme from *Lactuca sativa*; L25, LCYe enzyme with the first 25 amino acids removed; L34, LCYe enzyme with the first 34 amino acids removed; L50, LCYe enzyme
25 with the first 50 amino acids removed; L75, LCYe enzyme with the first 75 amino acids removed; L92, LCYe enzyme with the first 92 amino acids removed; L100, LCYe enzyme with the first 100 amino acids removed. See strain description in Table 3. All the measurements were average of triplicates with standard error bar shown in the figure. The percentage of ϵ -carotene in the final total carotenoids (lycopene + δ -carotene + ϵ -carotene)
30 were also indicated.

[0022] **Fig. 7** shows a protein and mRNA expression analysis of *LsLcyE* and *OfCCDI* with N-terminal modifications. (A) SDS-PAGE gel image of the whole-cell protein

expression of *LsLcyE* and *OfCCD1* with various N-terminal modifications. The proteins-of-interest were indicated with blue arrow. Lane 1: *E. coli* cells with overexpressed LsLCYe. Lane 2: *E. coli* cells with overexpressed $\Delta 50$ -LsLCYe (L50). Lane 3: *E. coli* cells with overexpressed $\Delta 100$ -LsLCYe (L100). Lane 4: *E. coli* cells with overexpressed OfCCD1.

5 Lane 5: *E. coli* cells with overexpressed SUMO-OfCCD1 (SUMO-O). Lane 6: *E. coli* cells with overexpressed MBP-OfCCD1 (MBP-O). Lane 7: *E. coli* cells with overexpressed TrxA-OfCCD1 (TrxA-O). Lane 8: *E. coli* cells without any overexpression (control). Quantitative analysis of the overexpressed protein yield was based on the band intensity on SDS PAGE gel. (B) LsLCYe expression with various N-terminal truncations. (C) mRNA expression level

10 of LsLCYe and its truncated version. (D) OfCCD1 expression with different fusion partners. (E) mRNA expression level of OfCCD1 with different fusion proteins. The relative fold change of protein expression was normalized by the expression of LsLCYe (B and C) and SUMO-OfCCD1 (D and E). The results were the average of triplicates with standard error bar shown.

15 [0023] **Fig. 8** is a schematic diagram of biosynthetic pathway of apocarotenoids (e.g. α , β -ionone and retinoids). The biosynthetic pathway was grouped into four major modules: the upstream MVA pathway (Expression module 1), the downstream MVA pathway (Expression module 2), the lycopene pathway (Expression module 3) and the apocarotenoid pathway (Expression module 4). The genes expressed encode the following enzymes: *atoB*,

20 Acetoacetyl-CoA thiolase; *hmgS*, HMG-CoA synthase; *thmgR*, truncated HMG-CoA reductase; *mevk*, mevalonate kinase; *pmk*, phosphomevalonate kinase; *pmd*, mevalonate pyrophosphate decarboxylase; *idi*, IPP isomerase; *ispA*, FPP synthase; *crtE*, GGPP synthase; *crtB*, phytoene synthase; *crtI*, phytoene desaturase; *LCYe*, lycopene epsilon-cyclase; *crtY*, lycopene beta-cyclase; *CCD1*, carotenoid cleavage dioxygenase; *BCDO* (or *blh*), β -carotene

25 dioxygenase; *ybbO*, NADP⁺-dependent aldehyde reductase. Abbreviation for the compounds: G3P, D-glyceraldehyde-3-phosphate; DHAP, Dihydroxyacetone phosphate; HMG-CoA, 3-hydroxy-3-methyl-glutaryl-coenzyme A; MVA, mevalonate; MVAP, phosphomevalonate; MVAPP, diphosphomevalonate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; HMBPP, (E)-4-Hydroxy-3-methyl-but-2-enyl

30 pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate. Dashed arrow indicates multiple enzymatic steps.

[0024] **Fig. 9** shows the production of α -ionone in the ϵ -carotene accumulating *E. coli* strain. (A) Optimization of CCD1 with N-terminal fusion partners (FP), a figure insert with a different y-axis scale was supplemented to compare the ionone titers of strains 121-LO and 121-L50-O. (B) Intracellular carotenoids that remained uncleaved. (C) An illustration of promoters and regulated genes for different strains. The abbreviations are as follows. LO, LsLCYe and OfCCD1 were expressed in a polycistronic manner on the same plasmid; L50-O, LsLCYe enzyme with the first 50 amino acids removed (L50) and OfCCD1 were expressed in a polycistronic manner. L50-SUMO-O, L50 and SUMO-fused OfCCD1 were co-expressed. L50-MBP-O, L50 and MBP-fused OfCCD1 were co-expressed. L50-TrxA-O, L50 and TrxA-fused OfCCD1 were co-expressed.

[0025] **Fig. 10** shows the protein solubility of OfCCD1 with N-terminal modifications. The proteins of interest were indicated with arrow. Lane 1: insoluble OfCCD1. Lane 2: soluble OfCCD1. Lane 3: insoluble SUMO-OfCCD1. Lane 4: soluble SUMO-OfCCD1. Lane 5: insoluble MBP-OfCCD1. Lane 6: soluble MBP-OfCCD1. Lane 7: insoluble TrxA-OfCCD1. Lane 6: soluble TrxA-OfCCD1.

[0026] **Fig. 11** shows the homology modelling of OfCCD1 based on the crystal structure of viviparous14 from *Zea mays* (PDB: 3NPE). The two membrane interacting helices are labelled. The membrane is represented by the shaded area. The co-factor, Fe^{2+} and oxygen are labeled.

[0027] **Fig. 12** shows the engineering of the second membrane interacting helix of OfCCD1. (A) A comparison of the titer of α -ionone and psi-ionone between the wide type ofCCD1 and engineered ofCCD1 with mutations (L151F, M152T and N154Y). (B) Biomass of the engineered strain. (C) The solubility and fold change in total expression of TrxA-fused wide-type OfCCD1 and trxA-fused mutant OfCCD1. (D) The SDS PAGE gel image of TrxA-fused wide-type OfCCD1 and trxA-fused mutant OfCCD1. Lane M, molecular marker. Lane 1, the total proteins of *E. coli* cell expressing trxA-fused wide-type OfCCD1. Lane 2, the soluble proteins of *E. coli* cell expressing trxA-fused wide-type OfCCD1. Lane 3, the total proteins of *E. coli* cell expressing trxA-fused mutant OfCCD1. Lane 4, the soluble proteins of *E. coli* cell expressing trxA-fused mutant OfCCD1.

[0028] **Fig. 13** shows the fed-batch fermentation of apocarotenoids. (A) Time-course profiles for ionones and biomass. (B) An illustration of promoters and regulated genes for the strain used in the fed-batch fermentation. Separated by induction, the process consisted of

two major phases, biomass phase accumulating only biomass and production phase that simultaneously produced ionone and biomass. Isopropyl myristate was used as organic layer to extract the α -, β - and psi-ionone. (C) An illustration of the *in situ* separation of apocarotenoids in bioreactor. Food-grade coconut oil (or soybean oil etc) as organic layer to extract the α -, β - and psi-ionone. Chemically defined media without any amino acid and vitamin was used to reduce the cost of medium.

[0029] **Fig. 14** shows the chirality analysis of α -ionone. The product is 100% R-enantiomer of α -ionone and identical to natural α -ionone. In contrast, the chemically synthesized ionone is a mixture of 50% R-enantiomer and 50% S-enantiomer.

10 [0030] **Fig. 15** demonstrates the metabolic engineering strategies employed for the production of β -ionone. Comparison of wild-type OfCCD1, N-terminal modified OfCCD1 and supplementation of *crtY*. In Of, TOF and *crtY* strain, the operon/gene *crtY-OfCCD1*, *crtY-TrxA-OfCCD1* and *crtY* were expressed, respectively. In TOF+Y strain, additional copy of *crtY* gene was supplemented in the p15A-*spec* vector. (A) Ionone production in engineered strain. Strain *crtY* was the control strain that did not express *CCD1* gene thus only produced β -carotene. (B) Accumulated carotenoids in the engineered strain. (C) Biomass of the engineered strains. Although thioredoxin fusion failed to improve β -ionone production, it resulted in a 37-fold increase in the production of the by-product psi-ionone which was converted from lycopene by TOFCCD1. The supplementation of additional *crtY* effectively converted more lycopene towards β -carotene, thus β -ionone production was significantly improved and psi-ionone was significantly reduced.

[0031] **Fig. 16** shows the GC-chromatograms of products in Of, TOF and TOF+Y strain.

[0032] **Fig. 17** shows the mass spectra of α -, β and psi-ionone.

25 [0033] **Fig. 18** shows the alignment of different CCD1 variants in this study. Alignment was done by Clustal Omega multiple sequence alignment tool. The information about CCD1 was summarized in Table 4.

[0034] **Fig. 19** demonstrates the screening of CCDs with higher selectivity. (A) Ionone production using different CCDs. (B) Carotenoids accumulated in *E. coli* cells with different CCDs. In TOF, At, Vv, Ph, TPh strain, the operon/gene *crtY-TrxA-OfCCD1*, *crtY-AtCCD1*, *crtY-VvCCD1*, *crtY-PhCCD1* and *crtY-TrxA-PhCCD1* were expressed, respectively. In Ph+Y and TPh+Y strain, additional copy of *crtY* gene was supplemented in the p15A-*spec* vector. It was found that *PhCCD1* had higher selectivity for β -carotene versus lycopene than *VvCCD1*

and *OfCCD1*, thus produced higher titer of ionone and lower amount of psi-ionone. (C) Fed-batch fermentation of β -ionone with strain TPh+Y. (D) An illustration of promoters and regulated genes for different strains.

[0035] **Fig. 20** shows the mass spectra of dialdehydes and dialcohols produced in the fermentator. C17 and C19 of dialdehydes and dialcohols were detected, but not C14.

[0036] **Fig. 21** shows the screening of CCD2 and CCD4 with better activity and selectivity. The top graph shows the use of β -carotene as substrate for β -ionone production. The bottom graph shows the use of carotene as substrate for α -ionone production.

[0037] **Fig. 22** shows the production of retinoids in the modularized system. (A) Retinoid production using different *blh* genes and increased gene copy of *crtY*. (B) Carotenoids accumulated in different *E. coli* cells. To produce other apocarotenoids, we could merely replace the CCD1 enzyme with other corresponding carotenoid cleavage oxygenases. Here, we used BCDO (*blh*) to produce retinoids. (C) A summary illustration of the modular system for the production of different apocarotenoids.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0038] In a first aspect the present invention refers to a method for producing a carotenoid or apocarotenoid. The method comprises the step of expressing in a host cell an expression module comprising an expression vector having a coding region encoding at least one optimised carotenoid or apocarotenoid generating enzyme, the coding region being operably linked to a promoter.

[0039] The carotenoid may be selected from phytoene, lycopene, α -carotene, γ -carotene, δ -carotene, ϵ -carotene or β -carotene and wherein the apocarotenoid may be selected from α -ionone, β -ionone, pseudo-ionone (or psi-ionone), hydroxy-ionone, β -cyclocitral, trans-geranylacetone, 6-methyl-5-hepten-2-one (MHO), retinal, retinol, 8',10-diapocarotene-8',10-dial (C17) or 10',6-diapocarotene-10',6-dial (C19).

[0040] The lycopene cyclase may be selected from *lycE* or *crtY* or their truncated forms.

[0041] The at least one optimised apocarotenoid generating enzyme, as described herein, may be selected from *crtY*, *CCD1*, *CCD2*, *CCD4*, *BCDO*, *LcyE*, *blh* or *ybbO*.

[0042] The CCD2 may be selected from CaCCD2 or CsCCD2. The CCD4 may be selected from the group consisting of AtCCD4, BoCCD4b, CmCCD4, CsCCD4a, MaCCD4, MdCCD4, OfCCD4, PpCCD4, RdCCD4 and VvCCD4a.

[0043] In one embodiment, the apocarotenoid may be α -ionone and the at least one optimised apocarotenoid generating enzyme may be selected from *LcyE* and *CCD1*, and may form an operon having the structure: *LcyE-CCD1*.

[0044] In some embodiments, the *LcyE* may be derived from *Lactuca sativa* and is N-terminal truncated (ranging from 1 to 100 amino acids of *LsLcyE*, especially $\Delta N50$ -*LsLcyE*).

[0045] The *CCD1* may be expressed as a fusion protein selected from *TrxA-CCD1*, *SUMO-CCD1* or *MBP-CCD1*.

[0046] In some embodiments, the fusion protein is *TrxA-CCD1*. The fusion protein may be selected from *TrxA-Osmanthus fragrans* CCD1 or *TrxA-Petunia hybrid* CCD1.

10 [0047] In some embodiments, the CCD1 may be derived from or *Osmanthus fragrans* (OfCCD1). In some embodiments, the OfCCD1, as described herein, may be optimized between amino acid positions F148 to I167. In particular, the OfCCD1 may comprise one or more of the following mutations: N154Y, M152T, L151F.

[0048] In some embodiments, the method may comprise screening for an expression level
15 of α -ionone in an amount of from 10 to 1000 mg/L; from 200 to 800 mg/L; from 300 to 700 mg/L; from 400 to 600 mg/L in a 24 hour period. In some embodiments, the amount of α -ionone is about 500 mg/L in a 24 hour period.

[0049] In some embodiments, the apocarotenoid is β -ionone. The at least one optimised apocarotenoid generating enzyme may be selected from *crtY* and *CCD1*, and may form an
20 operon having the structure *crtY-CCD1*. In some embodiments, the *crtY* as described herein may be derived from *Pantoea ananatis*.

[0050] In some embodiments, the *CCD1* as described herein may be derived from *Petunia hybrida* (PhCCD1).

[0051] In some embodiments, the method may further comprise screening for an
25 expression level of β -ionone in an amount of from 10 to 1000 mg/L; from 200 to 800 mg/L; from 300 to 700 mg/L; from 400 to 600 mg/L in a 24 hour period. In some embodiments, the amount of α -ionone is about 500 mg/L in a 24 hour period.

[0052] The apocarotenoid may be retinal or retinol. In one embodiment, the apocarotenoid may be retinal. The at least one optimised apocarotenoid generating enzyme may be selected
30 from *crtY* and *blh*, and may form an operon having the structure *crtY-blh*. In some embodiments, the apocarotenoid may be retinol. The at least one optimised apocarotenoid generating enzyme may be selected from *crtY*, *blh* or *ybbO*, and may form an operon having

the structure *crtY-blh-ybbO*. In some embodiments, the *crtY* as described herein may be derived from *Pantoea ananatis*. In some embodiments, the *blh* as described herein may be derived from *Uncultured marine bacterium HF10_19P19*. In some embodiments, the method may further comprise screening for an expression level of retinal in an amount of from 10 to 1000 mg/L; from 200 to 800 mg/L; from 300 to 700 mg/L; from 400 to 600 mg/L in a 24 hour period. In some embodiments, the amount of α -ionone is about 500 mg/L in a 24 hour period.

[0053] The carotenoid may be ϵ -carotene. The at least one optimised carotenoid generating enzyme may be *LcyE*. In some embodiments, the *LcyE* may be derived from *Lactuca sativa* (*LsLcyE*) and may be N-terminal truncated. In some embodiments, the *LsLcyE* may comprise an N-terminal truncation of from 1-100 amino acids; from 30-70 amino acids or from 40-60 amino acids. In some embodiments, the *LsLcyE* comprises an N-terminal truncation of 50 amino acids ($\Delta N50$ -*LsLcyE*).

[0054] In some embodiments, the method may further comprise expressing in said host cell: a first expression module comprising an expression vector having a first coding region encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY*, the first coding region being operably linked to a promoter; a second expression module comprising an expression vector having second coding region encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi*, the second coding region being operably linked to a promoter; a third expression module comprising an expression vector having a third coding region encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI*, the third coding region being operably linked to a promoter.

[0055] In some embodiments, the one or more first gene products may form an operon having the structure: *atoB-hmgS-thmgR*.

[0056] In some embodiments, the one or more first gene products may form an operon having the structure: *crtY-atoB-hmgS-thmgR*.

[0057] In some embodiments, the one or more second gene products may form an operon having the structure: *mevK-pmk-pmd-idi*.

[0058] In some embodiments, the one or more third gene products may form an operon having the structure: *crtE-crtB-ispA*.

[0059] In some embodiments, the one or more third gene products may form an operon having the structure: *crtE-crtB-crtI-ispA*.

[0060] The host cell as described herein may be *Escherichia coli* selected from BL21 DE3 or MG1655 DE3.

5 [0061] The promoter may be selected from one or more of TM1, TM2 or TM3, T7 RNA polymerase promoter, a T5 RNA polymerase promoter, a T3 RNA polymerase promoter, an SP6 RNA polymerase promoter or an inducible promoter.

[0062] In some embodiments, the optimisation of the apocarotenoid generating enzyme and optimised gene products is achieved by codon optimization or site-directed mutagenesis.

10 [0063] In another aspect, there is provided a host cell comprising an expression module comprising expression vector having a coding region encoding at least one optimised apocarotenoid or carotenoid generating enzyme, the coding region being operably linked to a promoter.

[0064] In some embodiments, the at least one optimised apocarotenoid generating enzyme
15 may be selected from *crtY*, *CCD1*, *CCD2*, *CCD4*, *BCDO*, *LcyE*, *blh* or *ybbO*.

[0065] The host cell may further comprise a first expression module comprising an expression vector having a first coding region encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY*, the first coding region being operably linked to a promoter; a second expression module comprising an expression vector
20 having a second coding region encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi*, the second coding region being operably linked to a promoter; a third expression module comprising an expression vector having a third coding region encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI*, the third coding region being operably linked to a promoter.

25 [0066] In another aspect, there is provided a vector encoding one or more optimised gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY* operably linked to a promoter.

[0067] In another aspect, there is provided a vector encoding one or more optimised gene products selected from *mevK*, *pmk*, *pmd* or *idi* operably linked to a promoter.

[0068] In another aspect, there is provided a vector encoding one or more optimised gene
30 products selected from *ispA*, *crtE*, *crtB* or *crtI* operably linked to a promoter.

[0069] In another aspect, there is provided a vector encoding one or more optimised gene products selected from *crtY*, *CCD1*, *BCDO*, *LcyE*, *blh* or *ybbO* operably linked to a promoter.

[0070] In some embodiments, the optimised gene products may be selected from the group consisting of *ΔN50-LsLcyE* and *TrxA-Osmanthus fragrans* CCD1; *crtY* and *phCCD1*; *ΔN50-LsLcyE*; *crtY* and *blh*, and *crtY*, *blh* and *ybbO*.

[0071] In another aspect, there is provided a system for producing an carotenoid or apocarotenoid comprising an expression module comprising an expression vector having a coding region encoding at least one optimised carotenoid or apocarotenoid generating enzyme, the coding region being operably linked to a promoter, wherein said at least one optimised carotenoid or apocarotenoid generating enzyme is selected from: a. *ΔN50-LsLcyE* and *TrxA-Osmanthus fragrans* CCD1 for the production of α -ionone; or b. *crtY* and *phCCD1* for the production of β -ionone; or c. *ΔN50-LsLcyE* for the production of ϵ -carotene; or d. *crtY* and *blh* for the production of retinal or retinol; or e. *crtY*, *blh* and *ybbO* for the production of retinol.

[0072] In some embodiments, the system may further comprise: a first expression module comprising an expression vector having a first coding region encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY*, the first coding region being operably linked to a promoter; a second expression module comprising an expression vector having second coding region encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi*, the second coding region being operably linked to a promoter; a third expression module comprising an expression vector having a third coding region encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI*, the third coding region being operably linked to a promoter.

[0073] In another aspect, there is provided a kit when used in the method as described herein, for the production of a carotenoid or apocarotenoid comprising one or more of: a first vector encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR*; a second vector encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi* operably linked to a promoter; a third vector encoding one or more optimised third gene products selected from *ispA*, *crtE* or *crtB*, optionally *crtI*, operably linked to a promoter for the production of phytoene or lycopene.

[0074] In another aspect, there is provided a kit when used in the method as described herein, for the production of an apocarotenoid or carotenoid comprising one or more of: a first vector encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY* operably linked to a promoter; a second vector encoding one or

more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi* operably linked to a promoter; a third vector encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI* operably linked to a promoter; and a fourth vector encoding one or more optimised gene products selected from: a. *ΔN50-LsLcyE* and TrxA-*Osmanthus fragrans* CCD1 operably linked to a promoter for the production of α -ionone; or b. *crtY* and *phCCD1* operably linked to a promoter for the production of β -ionone; or c. *ΔN50-LsLcyE* operably linked to a promoter for the production of ϵ -carotene; or d. *crtY* and *blh* operably linked to a promoter for the production of retinal; or e. *crtY*, *blh* and *ybbO* for the production of retinol.

[0075] In another aspect, there is provided a method for producing an apocarotenoid or carotenoid comprising the steps of: a. contacting a host cell as described herein in a chemically defined media with a substrate for apocarotenoid or carotenoid production; b. incubating the host cell in said chemically defined media to produce one or more preselected apocarotenoids or carotenoids, and c. extracting the one or more preselected apocarotenoids or carotenoids from the chemically defined media using an organic layer.

[0076] In some embodiments, the organic layer may be coconut oil or soybean oil or other edible oils.

[0077] In some embodiments, the apocarotenoid may be selected from α -ionone, β -ionone, pseudo-ionone (or psi-ionone), hydroxy-ionone, trans-geranylacetone, 6-methyl-5-hepten-2-one (MHO), retinal, retinol, 8',10-diapocarotene-8',10-dial (C17) or 10',6-diapocarotene-10',6-dial (C19) and wherein the carotenoid may be selected from phytoene, lycopene, α -carotene, γ -carotene, δ -carotene, ϵ -carotene or β -carotene.

[0078] The invention illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "comprising", "including", "containing", etc. shall be read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the inventions embodied therein herein disclosed may

be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention.

[0079] The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form
5 part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

[0080] Other embodiments are within the following claims and non-limiting examples. In addition, where features or aspects of the invention are described in terms of Markush groups,
10 those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

EXPERIMENTAL SECTION

[0081] Non-limiting examples of the invention and comparative examples will be further
15 described in greater detail by reference to specific Examples, which should not be construed as in any way limiting the scope of the invention.

Materials and Methods

[0082] *E. coli* Bl21-Gold DE3 strain (Stratagene) was used in this study. The genes *hmgS*,
hmgR, *mevK*, *pmk* and *pmd* (or *MVD1*) were amplified by PCR using the chromosomal DNA
20 of *Saccharomyces cerevisiae*. The genes *atoB* and *idi* were amplified from *E. coli* genomic DNA. All the genes were cloned into two plasmids, p15A-*spec-hmgS-atoB-hmgR* (L2-8) and p15A-*cam-mevK-pmk-pmd-idi* (L2-5). The genes in the lycopene biosynthetic pathway (*crtEBI*) amplified from the pAC-LYC plasmid was introduced into p15A-*kan-crtEBI-ispA* plasmid. The *LCYe* gene from *Lactuca sativa* (LsLCYe enzyme), the *crtY* gene from *Pantoea*
25 *ananatis*, the *CCD1* gene from *Arabidopsis thaliana*, *Osmanthus fragrans*, *Vitis vinifera* and *Petunia hybrid* and the *blh* genes from *Uncultured marine bacterium* HF10_19P19 (*blh1*) and *Uncultured marine bacterium* 66A03 (*blh2*) were codon optimized and synthesized by Integrated DNA Technologies. The *LsLCYe* gene and *crtY* gene was first cloned into the plasmid p15A-*amp-LsLCYe* (L2-9), and p15A-*amp-crtY* (L2-9), respectively. *OfCCD1* gene
30 was inserted into p15A-*amp-LsLCYe* (L2-9) plasmid. Different *CCD1* or *blh* genes were later inserted into p15A-*amp-crtY* (L2-9) strain. Site-directed mutagenesis was introduced from primers synthesized by Integrated DNA Technologies. Additional copy of *crtY* was inserted

into p15A-*spec-hmgS-atoB-hmgR* (L2-8) plasmid. All the p15A plasmids were derived from pAC-LYC plasmid. The stem-loop structure of RNA I in the p15A plasmid origin was mutated to make the plasmids compatible to each other. The T7 promoter variants are TM1, TM2 and TM3 was summarized in Table 1. The information about plasmids and strains was summarized in Table 2 and 3.

[0083] Table 1 summarizes the promoter sequences used.

Promoter Name	Sequence
T7	TAATACGACTCACTATAGGGGAATTGTGAGCG
TM1	TAATACGACTCACTAATGGGGAATTGTGAGCG
TM2	TAATACGACTCACTCGAGGGGAATTGTGAGCG
TM3	TAATACGACTCACTATAAAGGAATTGTGAGCG

[0084] Table 2 summarizes the strains and plasmids used.

Strains or plasmids	Descriptions
BL21 DE3	<i>E. coli</i> BL21-Gold DE3 strain (Stratagene).
p15A- <i>spec-hmgS-atoB-hmgR</i> (L2-8)	P15A vector carrying <i>atoB-hmgS-hmgR</i> operon controlled by TM1 promoter, with p15A origin mutated
p15A- <i>crtY-spec-hmgS-atoB-hmgR</i> (L2-8)	P15A vector carrying <i>crtY-atoB-hmgS-hmgR</i> operon controlled by TM1 promoter, with p15A origin mutated
p15A- <i>cam-mevK-pmk-pmd-idi</i> (L2-5)	P15A vector carrying <i>mevK-pmk-pmd-idi</i> operon controlled by TM2 promoter, with p15A origin mutated
p15A- <i>kan-crtEBI-ispA</i>	P15A vector carrying <i>crtEBI-ispA</i> operon controlled by TM1 promoter
p15A- <i>amp-LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-ΔN25LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene, with the first 25 amino acid removed, controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-ΔN34LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene, with the first 34 amino acid removed, controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-ΔN50LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene, with the first 50 amino acid removed, controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-ΔN75LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene, with the first 75 amino acid removed, controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-ΔN92LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene, with the first 92 amino acid removed, controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-ΔN100LsLCYe</i> (L2-9)	P15A vector carrying <i>LsLCYe</i> gene, with the first 100 amino acid removed, controlled by TM1 promoter, with p15A origin mutated

9)	amino acid removed, controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-LsLCYe-OfCCD1</i> (L2-9)	p15A- <i>amp-LsLCYe</i> (L2-9) containing <i>OfCCD1</i> gene
p15A- <i>amp-ΔN50LsLCYe-OfCCD1</i> (L2-9)	p15A- <i>amp-ΔN50LsLCYe</i> (L2-9) containing <i>OfCCD1</i> gene
p15A- <i>amp-ΔN50LsLCYe-SUMO-OfCCD1</i> (L2-9)	p15A- <i>amp-ΔN50LsLCYe</i> (L2-9) containing <i>OfCCD1</i> gene with <i>SUMO</i> fusion
p15A- <i>amp-ΔN50LsLCYe-MBP-OfCCD1</i> (L2-9)	p15A- <i>amp-ΔN50LsLCYe</i> (L2-9) containing <i>OfCCD1</i> gene with <i>MBP</i> fusion
p15A- <i>amp-ΔN50LsLCYe-TrxA-OfCCD1</i> (L2-9)	p15A- <i>amp-ΔN50LsLCYe</i> (L2-9) containing <i>OfCCD1</i> gene with <i>trxA</i> fusion
p15A- <i>amp-ΔN50LsLCYe-TrxA-OfCCD1_I151F_M152T_N154Y</i> (L2-9)	p15A- <i>amp-ΔN50LsLCYe</i> (L2-9) containing <i>OfCCD1</i> gene with <i>trxA</i> fusion, and three mutations, namely I151F, M152T, N154Y
p15A- <i>amp-crtY</i> (L2-9)	P15A vector carrying <i>crtY</i> gene controlled by TM1 promoter, with p15A origin mutated
p15A- <i>amp-crtY-AtCCD1</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>AtCCD1</i> gene
p15A- <i>amp-crtY-OfCCD1</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>OfCCD1</i> gene
p15A- <i>amp-crtY-TrxA-OfCCD1</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>OfCCD1</i> gene with the <i>trxA</i> fusion
p15A- <i>amp-crtY-PhCCD1</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>PhCCD1</i> gene
p15A- <i>amp-crtY-TrxA-PhCCD1</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>PhCCD1</i> gene with the <i>trxA</i> fusion
p15A- <i>amp-crtY-blh1</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>blh</i> gene from <i>Uncultured marine bacterium HF10_19P19</i>
p15A- <i>amp-crtY-blh2</i> (L2-9)	p15A- <i>amp-TM1-crtY</i> (L2-9) containing <i>blh</i> gene from <i>Uncultured marine bacterium 66A03</i>
pET11a- <i>trxA-CaCCD2</i>	pET11a vector containing <i>trxA</i> -fused <i>CaCCD2</i> gene from <i>Crocus angustifolius</i>
pET11a- <i>trxA-CsCCD2</i>	pET11a vector containing <i>trxA</i> -fused <i>CsCCD2</i> gene from <i>Crocus sativus</i>
pET11a- <i>trxA-AtCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>AtCCD4</i> gene from <i>Arabidopsis thaliana</i>
pET11a- <i>trxA-BoCCD4b</i>	pET11a vector containing <i>trxA</i> -fused <i>BoCCD4b</i> gene from <i>Bixa orellana</i>
pET11a- <i>trxA-CmCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>CmCCD4</i> gene from <i>Chrysanthemum morifolium</i>
pET11a- <i>trxA-CsCCD4a</i>	pET11a vector containing <i>trxA</i> -fused <i>CsCCD4a</i> gene from <i>Crocus sativus</i>
pET11a- <i>trxA-MaCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>MaCCD4</i> gene from <i>Musa acuminata</i> AAA Group
pET11a- <i>trxA-MdCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>MdCCD4</i> gene from <i>Malus domestica</i>
pET11a- <i>trxA-OfCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>OfCCD4</i> gene from <i>Osmanthus fragran</i>

pET11a- <i>trxA-PpCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>PpCCD4</i> gene from <i>Prunus persica</i>
pET11a- <i>trxA-RdCCD4</i>	pET11a vector containing <i>trxA</i> -fused <i>RdCCD4</i> gene from <i>Rosa damascena</i>
pET11a- <i>trxA-VvCCD4a</i>	pET11a vector containing <i>trxA</i> -fused <i>VvCCD4a</i> gene from <i>Vitis vinifera</i>

[0085] Table 3 describes the strains used in the study.

No.	Strains name	Descriptions
1	121	Strain producing lycopene , with three plasmids p15A- <i>spec-hmgS-atoB-hmgR</i> (L2-8), p15A- <i>cam-mevK-pmk-pmd-idi</i> (L2-5) and p15A- <i>kan-crtEBI-ispA</i> .
2	121Y	Strain producing β-carotene , with three plasmids p15A- <i>crtY-spec-hmgS-atoB-hmgR</i> (L2-8), p15A- <i>cam-mevK-pmk-pmd-idi</i> (L2-5) and p15A- <i>kan-crtEBI-ispA</i> .
3	121+LL	Strain producing ϵ-carotene , Strain 121 with p15A- <i>amp-LsLCYe</i> (L2-9), expressing wildtype <i>LCYe</i> .
4	121+L50	Strain producing ϵ-carotene , Strain 121 with p15A- <i>amp-AN50LsLCYe</i> (L2-9), expressing <i>LCYe</i> with first 50 amino acid truncated.
5	121+L100	Strain producing ϵ-carotene , Strain 121 with p15A- <i>amp-AN100LsLCYe</i> (L2-9), expressing <i>LCYe</i> with first 100 amino acid truncated.
6	121-LO	Strain producing α-ionone , Strain 121 with p15A- <i>amp-LsLCYe-OfCCD1</i> (L2-9), expressing wildtype <i>LCYe</i> and wildtype <i>OfCCD1</i> .
7	121-L50-O	Strain producing α-ionone , Strain 121 with p15A- <i>amp-LsLCYe-OfCCD1</i> (L2-9), expressing <i>LCYe</i> with first 50 amino acid truncated and wildtype <i>OfCCD1</i> .
8	121-L50-SUMO-O	Strain producing α-ionone , Strain 121 with p15A- <i>amp-AN50LsLCYe-SUMO-OfCCD1</i> (L2-9), expressing <i>LCYe</i> with first 50 amino acid truncated and SUMO-fused <i>OfCCD1</i> .
9	121-L50-MBP-O	Strain producing α-ionone , Strain 121 with p15A- <i>amp-AN50LsLCYe-MBP-OfCCD1</i> (L2-9), expressing <i>LCYe</i> with first 50 amino acid truncated and MBP-fused <i>OfCCD1</i> .
10	121-L50-TrxA-O	Strain producing α-ionone , Strain 121 with p15A- <i>amp-AN50LsLCYe-TrxA-OfCCD1</i> (L2-9), expressing <i>LCYe</i> with first 50 amino acid truncated and <i>TrxA</i> -fused <i>OfCCD1</i> .
11	121-L50-TrxA-O-L151F_M152T_N154Y	Strain producing α-ionone , Strain 121 with p15A- <i>amp-AN50LsLCYe-TrxA-OfCCD1_L151F_M152T_N154Y</i> (L2-9), expressing <i>LCYe</i> with first 50 amino acid truncated and <i>TrxA</i> -fused <i>OfCCD1</i> with three mutations.
12	Of	Strain producing β-ionone , Strain 121 with p15A- <i>amp-crtY-OfCCD1</i> (L2-9), expressing <i>crtY</i> and wildtype <i>OfCCD1</i> .

13	TOF	Strain producing β-ionone , Strain 121 with p15A- <i>amp-crtY-TrxA-OfCCD1</i> (L2-9), expressing <i>crtY</i> and <i>TrxA</i> -fused <i>OfCCD1</i> .
14	TOF+Y	Strain producing β-ionone , Strain 121Y with p15A- <i>amp-crtY-TrxA-OfCCD1</i> (L2-9), expressing <i>crtY</i> and <i>TrxA</i> -fused <i>OfCCD1</i> .
15	<i>crtY</i>	Strain producing β-carotene , strain 121 with p15A- <i>amp-crtY</i> (L2-9).
16	Ph	Strain producing β-ionone , Strain 121 with p15A- <i>amp-crtY-PhCCD1</i> (L2-9), expressing <i>crtY</i> and <i>PhCCD1</i> .
17	At	Strain producing β-ionone , Strain 121 with p15A- <i>amp-crtY-AtCCD1</i> (L2-9), expressing <i>crtY</i> and <i>PhCCD1</i> .
18	Vv	Strain producing β-ionone , Strain 121 with p15A- <i>amp-crtY-VvCCD1</i> (L2-9), expressing <i>crtY</i> and <i>PhCCD1</i> .
19	Ph+Y	Strain producing β-ionone , Strain 121Y with p15A- <i>amp-crtY-PhCCD1</i> (L2-9), expressing <i>crtY</i> and <i>PhCCD1</i> .
20	TPh	Strain producing β-ionone , Strain 121 with p15A- <i>amp-crtY-TrxA-PhCCD1</i> (L2-9), expressing <i>crtY</i> and <i>TrxA</i> -fused <i>PhCCD1</i> .
21	TPh+Y	Strain producing β-ionone , Strain 121Y with p15A- <i>amp-crtY-TrxA-PhCCD1</i> (L2-9), expressing <i>crtY</i> and <i>TrxA</i> -fused <i>PhCCD1</i> .
22	Blh1	Strain producing retinoids , Strain 121 with p15A- <i>amp-crtY-blh1</i> (L2-9), expressing <i>crtY</i> and <i>blh1</i> .
23	Blh2	Strain producing retinoids , Strain 121 with p15A- <i>amp-crtY-blh1</i> (L2-9), expressing <i>crtY</i> and <i>blh1</i> .
24	Blh1+Y	Strain producing retinoids , Strain 121Y with p15A- <i>amp-crtY-blh2</i> (L2-9), expressing <i>crtY</i> and <i>blh2</i> .
25	Blh2+Y	Strain producing retinoids , Strain 121Y with p15A- <i>amp-crtY-blh2</i> (L2-9), expressing <i>crtY</i> and <i>blh2</i> .

Media and culture conditions

[0086] All the cells were grown in 2XPY media (20 g/L Peptone, 10 g/L Yeast extract and 10 g/L NaCl), supplemented with 10 g/L glycerol, 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES). For prolonged incubation (48 h), 5 mg/L Tween80 was also added into the media to prevent cell aggregation. Briefly, 10 μ L fresh cell culture was inoculated into 1 mL fresh media in 14 mL BD Falcon™ tube. Cells were initially grown at 37°C with the shaking speed of 300 rpm and were induced by a range of IPTG concentrations (as indicated in text) when OD₆₀₀ reached around 0.6. After induction, 200 μ L of dodecane was supplemented onto the culture to extract ionone or retinoids, and the cells were incubated at 28°C for another 20 h or 48 h before harvest. The media were

supplemented with appropriate antibiotics (100 mg/L ampicillin, 34 mg/L chloramphenicol, 50 mg/L kanamycin and 50 mg/L spectinomycin) to maintain corresponding plasmids.

Quantification of carotenoids and retinoids

[0087] Intracellular carotenoids were extracted from cellular pellets according to the acetone extraction method. Briefly, 10–50 μ L bacterial culture (depending on the content of carotenoids in the cells) was collected and centrifuged. Cell pellets were washed with PBS and were resuspended in 20 μ L of water, followed by addition of 180 μ L of acetone. The HPLC method employed an Agilent 1260 Infinity LC System equipped with a ZORBAX, Eclipse Plus C18, 4.6 mm X 250 mm, 5 μ m column and diode array detector (DAD). Isocratic condition (50% methanol, 48% ethyl acetate and 2% water) was maintained at 1.5 mL/min for 5 min. The carotenoids were detected at wavelength of 450 nm. Standard curves were generated using commercial standards including ϵ -carotene, β -carotene, lycopene, retinol and retinal. The extracellular retinoid samples were prepared by diluting 20–50 μ L of organic layer into 1000 μ L hexane and analysed by the same HPLC system. Isocratic condition was used as follows. Mobile phase was 95% methanol and 5% acetonitrile and the flow rate was 1.5 ml/min. Retinoids were detected at wavelength of 340 nm.

Quantification of α , β -Ionone and ψ -ionone

[0088] The α , or β -ionone samples were prepared by diluting 20–50 μ L of organic layer into 1000 μ L hexane. The samples were analysed on an Agilent 7980B gas chromatography equipped with Agilent VF-WAXms column and an Agilent 7200 Accurate-Mass Quadrupole Time-of-Flight (GC/MS). Injection of samples was performed in splitless mode at 240°C. The oven program started at 100°C for 2 min, then the temperature was raised up at 30°C/min until 240°C and maintained at 240°C for another 2 min. The ionone concentrations were calculated by interpolating with a standard curve prepared by commercial standards. Mass spectrometer was operated in EI mode with full scan analysis (m/z 30–300, 1 spectra/s).

Chirality analysis

[0089] The chirality of α -ionone from the samples was analyzed with Agilent 7980B gas chromatography equipped with Agilent Cyclosil-B GC Column and an Agilent 7200 Accurate-Mass Quadrupole Time-of-Flight (GC/MS). The oven program started at 80°C for 2 min, then the temperature was raised up at 5°C/min until 210°C and to 250°C at 20°C /min and finally maintained at 250°C for another 2 min. The ionone concentrations were

calculated by interpolating with a standard curve prepared by commercial standards. Mass spectrometer was operated in EI mode with full scan analysis (m/z 30–300, 2 spectra/s).

Fed-batch fermentation

[0090] Starting medium was a chemically defined medium modified which contained 15 g/L glucose, 2 g/L $(\text{NH}_4)_2\text{SO}_4$, 4.2 g/L KH_2PO_4 and 11.24 g/L K_2HPO_4 , 1.7 g/L citric acid, 0.5 g/L MgSO_4 and 10 mL/L trace element solution. The trace element solution (100X) contained 0.25 g/L $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 1.5 g/L $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$, 0.15 g/L $\text{CuSO}_4 \cdot 2\text{H}_2\text{O}$, 0.3 g/L H_3BO_3 , 0.25 g/L $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.8 g/L $\text{Zn}(\text{CH}_3\text{COO})_2$, 5 g/L Fe(III) citrate, and 0.84 g/L EDTA, pH 8.0. Feed medium (500 g/L glucose and 5 g/L MgSO_4) was pumped into the 250 mL Mini Bioreactor (Applikon Biotechnology) at an initial rate of 0.6 mL/h and approximately exponentially increased to at 1.8 mL/h within 12 h and maintained at 1.5 mL/h for another 24 h. Cells were induced by 0.03 mM IPTG when OD reached about 40. After induction, 30 mL of isopropyl myristate was supplemented into the bioreactor to extract ionones.

EXAMPLES

Production of carotenes by “plug-n-play”

Example 1. Optimization of phytoene production

[0091] The “plug-n-play” as described herein platform produces carotenes and apocarotenoids from inexpensive feedstock (Fig. 1). It is very challenging to maintain high metabolic flux along a recombinant pathway that simultaneously overexpresses 13 enzymes in *E. coli* cells to produce carotenes (Fig. 2). Ultraviolet-B (UVB, 280-310 nm) is the most acute sunlight-induced damage to the skin. The colorless carotene, phytoene, has the maximum UV absorbance at 280 nm (Fig. 3), rendering it an ideal ingredient for topical UV protection agent. Phytoene is the first key intermediate that can be further converted to various C40 carotenoids. To produce phytoene, the biosynthetic pathway from acetyl-CoA was divided into three expression modules (Fig. 2, Module 1, upstream mevalonate pathway including *hmgS-atoB-hmgR*; Module 2, downstream mevalonate pathway including *mevK*, *pmk*, *pmd* and *idi*; Module 3, phytoene biosynthetic pathway including *crtEB* and *ispA*), and systematically optimized by EDASPO approach. Briefly, the 3 expression modules were controlled by three different T7 promoter variants with different strength, and the phytoene yield was then correlated with the promoter strength of each module to locate the optimal transcriptional range for each expression module (Ternary diagram, Fig. 4). As a

result, the optimal range for each expression module was identified and an optimized strain (221 strain) was obtained, that produced ~45 mg/L of phytoene in a shake flask, or 38,000 ppm (Fig. 4).

Example 2. Optimization of lycopene production

5 [0092] Lycopene is another valuable carotenoid that has potent anti-oxidant effect. The biosynthetic pathway is extremely long and challenging to optimize in heterologous microbial production strain. Here, the pathway was divided into three different modules (Fig. 2, Module 1, upstream mevalonate pathway including *hmgS-atoB-hmgR*; Module 2, downstream mevalonate pathway including *mevK*, *pmk*, *pmd* and *idi*; Module 3, lycopene
10 biosynthetic pathway including *crtEBI* and *ispA*), and overproduce lycopene by EDASPO approach. Briefly, the 3 expression modules were controlled by three different T7 promoter variants with different strength, and the lycopene yield was then correlated with the promoter strength of each expression module to locate the optimal transcriptional range for each expression module (Ternary diagram, Fig. 5). As a result, the optimal range for each
15 expression module was identified and an optimized strain (121 strain) was obtained, that produced ~107 mg/L of lycopene in a shake flask, or 26,000 ppm (Fig. 5). This lycopene-producing strain served as parental platform strain for further production of downstream carotenoids and apocarotenoids.

Example 3. Optimization of carotene production

20 [0093] A fourth biosynthetic expression module can be introduced based on the parental lycopene-overproducing strain (Fig. 2). Depending on the expression levels of crtY and LCYe, α -, β -, δ - γ -and ϵ -carotene can be produced (Fig. 2). However, ϵ -carotene is known to be accumulated at very low levels in nature. To date, only the LCYe from *Lactuca sativa* (LsLCYe) is known to preferentially form the bicyclic ϵ -ring from lycopene. When the wild
25 type LsLCYe enzyme in the lycopene strain (121 strain) was overexpressed, a mixture of δ -carotene (carotene with an ϵ -ring at one end) and ϵ -carotene (carotene with two ϵ -rings at both ends) was produced, with more than 70% lycopene uncyclized (0.1 mM, Fig. 6). The ϵ -carotene production was initially optimized by tuning the transcription levels and obtained highest amount of ϵ -carotene with moderate IPTG induction (0.03 mM). However, the
30 highest ϵ -carotene titer was merely 18 mg/L (2500 ppm), corresponding to less than 30% in the total carotenoids (Fig. 6), suggesting that the activity of LsLCYe enzyme was insufficient.

[0094] Structural prediction by Phobius showed a likely membrane interaction region (104th-128th amino acids) in the LsLCYe enzyme and a weak signal peptide region (1st-30th amino acids). Modifications of N-terminal residues involved in membrane interactions and signal peptide coding regions have previously been reported to enhance solubility and activity of recombinant protein. Therefore, the truncation of the N-terminus sequences involved in anchoring to membrane was examined to determine if it would increase the catalytic efficiency of LsLCYe. A series of N-terminal truncated LsLCYe was constructed by systematically removing 50 residues at a time using the truncate cross-lapping *in vitro* assembly method. The enzymatic expression level was increased by ~40% when the first 50 amino acids (Fig. 7A and B) were removed, with a concomitant greater than two-fold increase in ϵ -carotene (>40 mg/L, 5500 ppm and 80% of final products) production (Fig. 6, 121+L50 strain). Remarkably, removing the first 100 residues (Fig. 6, 121+L100 strain) led to an increase in the protein concentration by ~140% (Fig. 7A and B). However, the cyclase activity was completely lost. In order to better understand how truncation affected LsLCYe expression, the transcription level and mRNA secondary structure of truncated LsLCYe and its native form were compared. The results indicated that increased expressions of truncated LsLCYe did not appear to be due to significant increases in mRNA expressions (Fig. 7C) nor modifications in secondary mRNA structures.

[0095] In order to demonstrate the broader utility of this modular approach, other apocarotenoids were produced by replacing the LCYe enzyme with the expression of lycopene beta cyclase (or *crtY*) from *Pantoea ananatis*. Unlike LsLCYe, the native lycopene beta cyclase was highly active and converted greater than 70% of lycopene into β -carotene (65 mg/L), with residual amount of lycopene (24 mg/L) (*crtY* strain in Fig. 15).

Production of apocarotenoids by “plug-n-play”

25 Example 4. Optimization of α -ionone production

[0096] By modular approach, carotenoid cleavage oxygenase(s) can be added to the fourth expression module to produce various apocarotenoids (Fig. 8). LsLCYe with N-terminus 50 amino acids truncated (Δ N50-LsLCYe) was further explored for α -ionone production. It has been demonstrated previously that CCD1 from *Osmanthus fragrans* (OfCCD1) can cleave carotene at C9-C10 position to produce both the α - and β -ionone. Thus, OfCCD1 enzyme was overexpressed in ϵ -carotene strains (121-LL and 121-L50 in Fig. 6). However, even with enhanced availability of ϵ -carotene in strain 121-L50-O, the amount of α -ionone detected was

less than 500 $\mu\text{g/L}$ (Fig. 9A, insert). As previously noted, lycopene and ϵ -carotene were the main intermediates accumulated in the LsLCYe and $\Delta\text{N50-LsLCYe}$ overexpressed strains, respectively (Fig. 9B). SDS-PAGE analysis showed that hardly any OfCCD1 protein was observed 24 h after induction (Fig. 7A and D), even though mRNA was well detected (Fig. 7E). The results indicated that OfCCD1 expression may be limited by insufficient translation, consistent with previous studies on other CCD1 overexpression.

[0097] Literature suggested that fusion with glutathione-S-transferase (GST) would significantly increase expression and solubility of CCD1 in *E. coli*. Thus, the fusion partners were examined for improvements in the functional expression of OfCCD1. Three commonly used N-terminal fusion partners, small ubiquitin-like modifier (SUMO) protein, maltose binding protein (MBP) and thioredoxin (TrxA) were tested. OfCCD1 when fused with the three fusion partners significantly increased the titers of α -ionone (Fig. 9A). Among them, the strain with TrxA-fused OfCCD1 (or TOfCCD1) was able to produce 29.7 mg/L α -ionone with low induction (0.03 mM IPTG), which is more than 50-fold improvement (Fig. 9A). To study the underlying mechanism of fusion partners, the protein expression was further investigated by denaturing SDS-PAGE gel. Without fusion partner, OfCCD1 was barely detectable (Fig. 7A and D). In contrast, OfCCD1 enzymes fused with SUMO, MBP and TrxA were well expressed. The relative protein band intensities of SUMO-, MBP- and TrxA-fused OfCCD1 were well correlated with their corresponding α -ionone production in Fig. 9A. For example, TOfCCD1 expressing strain had the best expression and produced the highest amount of α -ionone. The effects of fusion partners on solubility of OfCCD1 were further investigated. All the fused CCD1 have clear soluble fractions (Fig. 10), however, there was no clear correlation between α -ionone production and solubility of SUMO-, MBP- and TrxA-fused OfCCD1. As transcription and mRNA secondary structure remained relatively similar, the increased expression of OfCCD1 was likely to be due to an enhancement in translation efficiency or increase in protein stability. Interestingly, even with the highest ionone producer strain (Fig. 9, 121-L50-TrxA-O strain with 0.03 mM IPTG), both ϵ -carotene and lycopene (~30 mg/L total carotenoids) were still accumulated to the similar levels as α -ionone (Fig. 9B), suggesting that α -ionone yield could be further improved. Hence, a progressive improvement in the production of α -ionone from 500 $\mu\text{g/L}$ to 29 mg/L was achieved.

[0098] To further improve the α -ionone yield, homology modelling was used to predict the 3D structure of OfCCD1 (Fig. 11). There are two membrane-interaction helices based on

the OfCCD1 model. The first helix is located at the N-terminal of the protein, and the second helix is found within the sequence (F148-I167). Sequence alignment indicated the main discrepancies occurred at leucine-151, methionine-152, asparagine-154, and isoleucine-167 (Fig. 18). By simultaneously mutating leucine-151 to phenylalanine, methionine-152 to threonine and asparagine-154 to tyrosine, the α -ionone titer was further improved to ~35 mg/L (Fig. 12A). Protein expression analysis showed that these mutations have further improved the trxA-fused OfCCD1 expression (Fig. 12 C and D). The results highlighted that CCD enzyme engineering is critical for improving ionone production.

[0099] Next, the performance in fed-batch fermentation was tested using the best α -ionone producing strain (121-L50-TrxA-O strain). As α -ionone is volatile, isopropyl myristate was used to entrap the compound during fermentation. As shown in Fig. 13, about 480 mg/L of α -ionone was produced within 48 h. In sum, a ~1000-fold increase in the α -ionone production from initial 500 μ g/L to 480 mg/L was achieved. Thus, the α -ionone titer achieved in this study is about 1428 times higher than previously reported.

[00100] Furthermore, the produced α -ionone was subjected to chirality check and found to be 100% R-enantiomer of α -ionone (Fig. 14), identical to natural α -ionone. In contrast, the chemically synthesized ionone is a mixture of 50% R-enantiomer and 50% S-enantiomer. More importantly, different enantiomers of ionone have different odors and potential different activities. The result clearly showed the advantage of biosynthesis through the “plug-n-play” platform over chemical synthesis.

Example 5. Optimization of β -ionone production

[00101] To demonstrate the broader utility of this modular approach, other apocarotenoids were produced by co-expressing carotenoids cleavage dioxygenase with lycopene beta cyclase (or *crtY*) from *Pantoea ananatis*. The β -ionone production with wild-type *OfCCD1* expressing strain (Of strain) was 140 μ g/L, corresponding to merely less than 1% conversion yield from β -carotene (Fig. 15). As the N-terminal TrxA fused OfCCD1 (TOfCCD1) improved α -ionone (Fig. 9), the possibility that it may similarly improve β -ionone production was tested. Disappointingly, TOfCCD1 expression did not result in a significant increase in the production of β -ionone (Fig. 15A). Instead, another major peak at 9.46 min was detected in the GC chromatogram (Fig. 16 and 17). This peak was predicted to be psi-ionone with the National Institute of Standards and Technology (NIST) mass spectrum library and further

validated with authentic standard of psi-ionone. The expression of TOfCCD1 resulted in a 37-fold increase in psi-ionone production from ~0.4 to ~14.6 mg/L. Noteworthy, psi-ionone was also produced by the α -ionone production strain, although it was a minor product (Fig. 9A and 12A). Hence, it appeared that lycopene was also cleaved by TOfCCD1 and produced psi-ionone (Fig. 8), consistent with the observation that certain types of CCD1 could convert lycopene to psi-ionone or 6-methyl-5-hepten-2-one (MHO). In addition, it was observed that no β -carotene but large amount of lycopene accumulated in TOfCCD1 expressing strain (TOf strain, Fig. 15B). One possible reason could be that lycopene was preferentially converted by TOfCCD1 rather than cyclized by the lycopene beta cyclase in exponential phase.

[00102] As large amount of lycopene and lycopene-derived psi-ionone were produced in the TOf strain, it was hypothesized that increasing the gene dosage of *crtY* (TOf+Y strain) may have two potential benefits. Firstly, it may increase the carbon fluxes from lycopene to β -carotene and hence increase β -ionone production. Secondly, the production of psi-ionone may be reduced by the decrease in the availability of intracellular lycopene. To test the hypothesis, *crtY* was firstly inserted into the first expression module of our system, *atoB-hmgS-hmgR*, resulting in a new operon, *crtY-atoB-hmgS-hmgR* (illustrated in Fig. 15). As the plasmids used in the study are about 10 copies per cell, the insertion increased the *crtY* copy number from 10 to 20 copies per cell. Interestingly, gene dosage of *crtY* effectively resulted in more lycopene converted to β -carotene, hence, from undetectable amount of β -carotene (TOf strain, Fig. 15) to a significant yield of 42.4 mg/L (TOf+Y strain, Fig. 15). Consequently, the titer of β -ionone was significantly increased from 0.3 to 20.9 mg/L, and psi-ionone was decreased from 14.6 to 4.4 mg/L (Fig. 15), indicative of a higher carbon flux diverted into β -ionone production through β -carotene.

[00103] Although supplementation of *crtY* markedly improved β -ionone production, there was still 4.4 mg/L of psi-ionone produced. To maximize the β -ionone production and minimize by-product/intermediates formation, CCD with higher selectivity for β -carotene was screened. Another three CCD1 enzymes, from *Arabidopsis thaliana* (AtCCD1), *Vitis vinifera* (VvCCD1) and *Petunia hybrid* (PhCCD1) were tested (Fig. 18). All CCD1 enzymes used in this study are shown in Table 4. With the expression of wild-type AtCCD1, VvCCD1 and PhCCD1, β -ionone was produced at 0.1, 13.0 and 27.0 mg/L, respectively, accompanied by 0.2, 14.8 and 4.8 mg/L psi-ionone, respectively (Fig. 19A). The data suggested that wild-type VvCCD1 and PhCCD1 had higher activities in *E. coli* than AtCCD1 and OfCCD1. In

addition, PhCCD1 also had higher selectivity for β -carotene than VvCCD1 and OfCCD1. Nevertheless, there were still 13.3 mg/L β -carotene and 13.8 mg/L lycopene accumulated in the PhCCD1 expressing strain (Fig. 19B). The fusion of TrxA to PhCCD1 and supplementation of additional copy of *crtY* was next investigated. With additional

5 supplementation of *crtY* (Ph+Y strain), the β -ionone titer increased from 27 to 32.4 mg/L and psi-ionone titer decreased from 4.8 to 1.6 mg/L. Concomitantly, β -carotene production increased from 13.3 (Ph strain) to 24.3 (Ph+Y strain) mg/L and lycopene production decreased from 13.8 to 1.5 mg/L. However, the fusion of TrxA (TPh strain) merely resulted in a moderate increase in β -ionone production from 27 to 30.4 mg/L and a slight decrease in

10 psi-ionone production from 4.8 to 3.2 mg/L. Furthermore, there was no synergistic benefit with the combination of *crtY* supplementation and TrxA fusion (TPh+Y strain, Fig. 19). Collectively, the results suggested that unlike OfCCD1, native PhCCD1 activity was sufficiently efficient in converting β -carotene to β -ionone. In sum, the β -ionone production from initial 140 μ g/L to 32.4 mg/L was improved, a ~230-fold increase in tubes and flasks

15 over a period of 24 h. Further in fed-batch fermentation, ~500 mg/L of β -ionone was produced with Ph+Y strain (Fig. 19C), which was about 3700-fold increase over that of the initial strain.

[00104] Table 4 summarizes the CCD1 enzymes in this study.

Entry	Enzymes	Organisms
O65572	AtCCD1	<i>Arabidopsis thaliana</i> (Mouse-ear cress)
U3M7F6	VvCCD1	<i>Vitis vinifera</i> (Grape)
D4QE74	OfCCD1	<i>Osmanthus fragrans</i>
Q6E4P3	PhCCD1	<i>Petunia hybrida</i> (Petunia)

20 [00105] The organic phase from the bioreactor was yellow in color. To identify the yellow compound(s), ultra performance liquid chromatography coupled with time-of-flight mass spectrometry analysis was carried out. A few dialdehydes (C17 and C19) and alcohols were identified (Fig. 20). These are the by-products due to the non-specific cleavage of PhCCD1. To circumvent the problem, screening for a more specific CCD would was carried out. Since

25 CCD4 has demonstrated remarkable improvement in the substrate and product specificity, we have screened several CCD4 enzymes listed in Table 5, and found that PpCCD4 is most active against β -carotene, and CmCCD4 is most active against ϵ -carotene (Fig. 21).

[00106] Table 5 summarizes the CCD2 and CCD4 enzymes in this study.

CCD enzymes	Organisms
CaCCD2	<i>Crocus angustifolius</i>
CsCCD2	<i>Crocus sativus</i>
AtCCD4	<i>Arabidopsis thaliana</i>
BoCCD4b	<i>Bixa orellana</i>
CmCCD4	<i>Chrysanthemum morifolium</i>
CsCCD4a	<i>Crocus sativus</i>
MaCCD4	<i>Musa acuminata</i> AAA Group
MdCCD4	<i>Malus domestica</i>
OfCCD4	<i>Osmanthus fragran</i>
PpCCD4	<i>Prunus persica</i>
RdCCD4	<i>Rosa damascena</i>
VvCCD4a	<i>Vitis vinifera</i>

Example 6. Optimization of retinoids production

[00107] Retinoids are biosynthesized by cleaving β -carotene at C15-C15' position. To further illustrate the utility of the "plug-n-play" system to produce other apocarotenoids, CCD1 gene was replaced with two *blh* genes (BCDO enzyme, Fig. 8), from *Uncultured marine bacterium HF10_19P19* (*blh1*) and *Uncultured marine bacterium 66A03* (*blh2*). Without any optimization of the upstream pathway, the "plug-n-play" system produced 18.5 mg/L retinol and 5.1 mg/L retinal in the *blh1* strain and 3.0 mg/L retinol and 1.1 mg/L retinal in *blh2* strain (Fig. 22A). In *blh1* strain, lycopene accumulated (28.2 mg/L) but no β -carotene was detected (Fig. 22B). Thus, it was hypothesized that retinoid production was limited by lycopene beta-cyclase instead of BCDO in *blh1* strain. As expected, the co-expression of additional *crtY* gene (*blh1*+Y) increased retinoid production (from 23.6 to 33.2 mg/L). As for *blh2* strain, retinoid production was likely to be limited by the activity of BCDO instead of lycopene beta-cyclase, as the increase in *crtY* gene dosage failed to improve the retinoid production despite that >90% lycopene was converted into β -carotene (Fig. 22). It is worth noting that in the retinoid-producing strains, the *ybbO* gene, predicted as an oxidoreductase in

E. coli and validated experimentally with high retinol dehydrogenase activity (Genbank Accession ID is ECD_00444, Fig. 8), was not overexpressed but greater than 80% of retinal was converted into retinol. While the underlying mechanism was unknown, it could be due to the relatively high intracellular expression of ybbO enzyme or other unknown endogenous proteins with retinol dehydrogenase activity. The retinoid and ionone results suggested that this “plug-n-play” system is versatile and will be useful for the production of many apocarotenoids which share a common upstream pathway (Fig. 22C).

Discussion

[00108] Modular metabolic engineering has been successfully applied to rationally improve small molecule production by microbial cell factories. It is a powerful approach to reduce the parameter optimization space and allow sequential optimization of subsets of expression modules before amalgamating with the full pathway. Generally, as expression module number increases, the accuracy of pathway optimization will improve but the full combination will also increase exponentially. Thus, a balanced expression module number has to be decided. As described herein, the multi-gene pathway containing 13 genes was cast into 4 different expression modules: upstream mevalonate pathway (expression module 1), downstream mevalonate pathway (expression module 2), lycopene biosynthetic pathway (expression module 3) and apocarotenoids biosynthetic pathway (expression module 4). By optimizing the first three metabolic expression modules with EDASPO approach, a stable parental strain that accumulated high content of lycopene was obtained (strain 121 in Fig. 6). Subsequently, the fourth expression module that carried lycopene cyclase and CCD was introduced to produce various apocarotenoids. Having the apocarotenoid expression module independent from the rest of the expression modules offers three major advantages over non-modular approaches. Firstly, it minimizes the potential negative effects on the upstream system (the mevalonate pathway and lycopene biosynthetic pathway). Secondly, it significantly simplifies the strain engineering workload. As the upstream system was optimized and remained efficient, engineering efforts could be focused on the downstream apocarotenoid biosynthesis expression module. Lastly, it makes the system a ‘plug-n-play’ chassis for the production of valuable apocarotenoids. As such, ‘plugging’ lycopene cyclase (*LCYe* or *crtY*) and CCD1 produces α - or β - ionone and ‘plugging’ *crtY* and *blh* produces retinol. More importantly, similar titers of the apocarotenoids were obtained in flasks (Fig. 9, 15, 19 and 22) indicating the robustness of this modular system.

[00109] Collectively, these results demonstrate that modulating key protein expression is essential for the heterologous biosynthesis of apocarotenoids in *E. coli*. Broadly, this highlights the benefits of combining protein engineering and modular pathway design for the overproduction of valuable chemicals.

- 5 [00110] The recent changes in the regulation of natural ingredient labelling have resulted in the increase in demand of natural flavours and fragrances. Consequently, there are growing interests in engineering microbes to produce the ingredients from renewable resources. However, the complex metabolic characteristics of apocarotenoid pathway have hampered the development of highly efficient microbial processes. Advantageously, the present
- 10 disclosure provides modular metabolic engineering and enzyme engineering strategies that were systematically applied to effectively minimize the metabolic burden imposed by overexpression of 13 enzymes to overcome the challenge from critical enzymes of low activities. This strategy has enabled the development of a robust *E. coli* strain capable of producing unprecedented yields of α -ionone and β -ionone, demonstrating the great potential
- 15 of using microbes in production of natural flavours and fragrances.

Claims

1. A method for producing a carotenoid or apocarotenoid comprising the step of expressing in a host cell an expression module comprising an expression vector having a coding region encoding at least one optimised apocarotenoid or carotenoid generating enzyme, the coding region being operably linked to a promoter.
2. The method as claimed in claim 1, wherein the carotenoid is selected from phytoene, lycopene, α -carotene, γ -carotene, δ -carotene, ϵ -carotene or β -carotene and wherein the apocarotenoid is selected from α -ionone, β -ionone, pseudo-ionone (or psi-ionone), hydroxy-ionone, β -cyclocitral, trans-geranylacetone, 6-methyl-5-hepten-2-one (MHO), retinal, retinol, 8',10-diapocarotene-8',10-dial (C17) or 10',6-diapocarotene-10',6-dial (C19).
3. The method as claimed in claim 1 or 2, wherein said at least one optimised carotenoid or apocarotenoid generating enzyme is selected from *crtY*, *CCD1*, *CCD2*, *CCD4*, *BCDO*, *LcyE*, *blh* or *ybbO*.
4. The method as claimed in claim 3, wherein said CCD2 is selected from CaCCD2 or CsCCD2.
5. The method as claimed in claim 3, wherein said CCD4 is selected from the group consisting of AtCCD4, BoCCD4b, CmCCD4, CsCCD4a, MaCCD4, MdCCD4, OfCCD4, PpCCD4, RdCCD4 and VvCCD4a.
6. The method as claimed in claim 2 or claim 3, wherein the apocarotenoid is α -ionone and wherein said at least one optimised apocarotenoid generating enzyme is selected from *LcyE* and *CCD1*, and form an operon having the structure: *LcyE-CCD1*.
7. The method as claimed in claim 6, wherein said *LcyE* is derived from *Lactuca sativa* and is N-terminal truncated (ranging from 1 to 100 amino acids of *LsLcyE*, especially Δ N50-*LsLcyE*).
8. The method as claimed in claim 6, wherein said *CCD1* is expressed as a fusion protein selected from *TrxA-CCD1*, *SUMO-CCD1* or *MBP-CCD1*.
9. The method as claimed in claim 8, wherein the fusion protein is *TrxA-CCD1*.
10. The method as claimed in claim 9, wherein fusion protein is selected from *TrxA-Osmanthus fragrans CCD1* or *TrxA-Petunia hybrid CCD1*.

11. The method as claimed in claim 10, wherein the CCD1 is derived from or *Osmanthus fragrans* (OfCCD1) and comprise one or more of the following mutations: N154Y, M152T, L151F.
12. The method as claimed in any one of claims 6 to 11, comprising screening for an expression level of α -ionone in an amount of up 100-1000 mg/L in a 24 hour period.
13. The method as claimed in claim 2 or claim 3, wherein the apocarotenoid is β -ionone and wherein said at least one optimised apocarotenoid generating enzyme is selected from *crtY* and *CCD1*, and form an operon having the structure *crtY-CCD1*.
14. The method as claimed in claim 13, wherein said *CCD1* is derived from *Petunia hybrida* (PhCCD1).
15. The method as claimed in claim 13, wherein said *crtY* is derived from *Pantoea ananatis*.
16. The method as claimed in any of claims 13 to 15, further comprising screening for an expression level of β -ionone in an amount of up 10-1000 mg/L in a 24 hour period.
17. The method as claimed in claim 2 or claim 3, wherein the apocarotenoid is retinal and wherein said at least one optimised apocarotenoid generating enzyme is selected from *crtY* and *blh*, and form an operon having the structure *crtY-blh*.
18. The method as claimed in claim 17, wherein said *crtY* is derived from *Pantoea ananatis*.
19. The method as claimed in claim 17, wherein said *blh* is derived from *Uncultured marine bacterium HF10_19P19*.
20. The method as claimed in claims 17 to 19, further comprising screening for an expression level of retinal in an amount of up 10-1000 mg/L in a 24 hour period.
21. The method as claimed in claim 2 or claim 3, wherein the carotenoid is ϵ -carotene and wherein said at least one optimised carotenoid generating enzyme is *LcyE*.
22. The method as claimed in claim 21, wherein said *LcyE* is derived from *Lactuca sativa* and is N-terminal truncated (*AN50-LsLcyE*).
23. The method as claimed in any of claims 21 to 22, further comprising screening for an expression level of ϵ -carotene in an amount of up 10-1000 mg/L in a 24 hour period.
24. The method as claimed in claim 2, wherein the carotenoid is phytoene.
25. The method as claimed in claim 24, further comprising screening for an expression level of phytoene in an amount of up 10-1000 mg/L in a 24 hour period.

26. The method of any one of claims 1 to 25, further comprising expressing in said host cell: a first expression module comprising an expression vector having a first coding region encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY*, the first coding region being operably linked to a promoter; a second expression module comprising an expression vector having second coding region encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi*, the second coding region being operably linked to a promoter; a third expression module comprising an expression vector having a third coding region encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI*, the third coding region being operably linked to a promoter.
27. The method as claimed in claim 26, wherein said one or more first gene products form an operon having the structure: *atoB-hmgS-thmgR*.
28. The method as claimed in claim 26, wherein said one or more first gene products form an operon having the structure: *crtY-atoB-hmgS-thmgR*.
29. The method as claimed in claim 26, wherein said one or more second gene products form an operon having the structure: *mevK-pmk-pmd-idi*.
30. The method as claimed in claim 26, wherein said one or more third gene products form an operon having the structure: *crtE-crtB-ispA*.
31. The method as claimed in claim 26, wherein said one or more third gene products form an operon having the structure: *crtE-crtB-crtI-ispA*.
32. The method of any of claims 1 to 31, wherein the host cell is *Escherichia coli* selected from BL21 DE3 or MG1655 DE3.
33. The method as claimed in any one of claims 1 to 32, wherein the promoter is selected from one or more of TM1, TM2 or TM3, T7 RNA polymerase promoter, a T5 RNA polymerase promoter, a T3 RNA polymerase promoter, an SP6 RNA polymerase promoter or an inducible promoter.
34. The method as claimed in any one of claims 1 to 33, wherein the optimisation of the apocarotenoid generating enzymes and optimised gene products is achieved by codon optimisation or site-directed mutagenesis.
35. A host cell comprising an expression module comprising expression vector having a coding region encoding at least one optimised apocarotenoid or carotenoid generating enzyme, the coding region being operably linked to a promoter.

36. The host cell as claimed in claim 35, wherein said at least one optimised apocarotenoid generating enzyme is selected from *crtY*, *CCD1*, *CCD2*, *CCD4*, *BCDO*, *LcyE*, *blh* or *ybbO*.
37. The host cell as claimed in claim 36, wherein the optimised gene products are selected from the group consisting of $\Delta N50$ -*LsLcyE* and TrxA-*Osmanthus fragrans* CCD1; *crtY* and *phCCD1*; $\Delta N50$ -*LsLcyE*; *crtY* and *blh*, and *crtY*, *blh* and *ybbO*.
38. The host cell as claimed in claim 37, wherein the *Osmanthus fragrans* CCD1 comprises one or more of the following mutations: N154Y, M152T, L151F.
39. The host cell as claimed in any one of claims 26 to 38, further comprising:
- 10 a first expression module comprising an expression vector having a first coding region encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY*, the first coding region being operably linked to a promoter;
- a second expression module comprising an expression vector having a second coding region encoding one or more optimised second gene products selected from *mevK*,
15 *pmk*, *pmd* or *idi*, the second coding region being operably linked to a promoter;
- a third expression module comprising an expression vector having a third coding region encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI*, the third coding region being operably linked to a promoter.
40. The host cell as claimed in any one of claims 26 to 39, wherein the host cell is
20 *Escherichia coli* selected from BL21 DE3 or MG1655 DE3.
41. A vector encoding one or more optimised gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY* operably linked to a promoter.
42. A vector encoding one or more optimised gene products selected from *mevK*, *pmk*, *pmd* or *idi* operably linked to a promoter.
- 25 43. A vector encoding one or more optimised gene products selected from *ispA*, *crtE*, *crtB* or *crtI* operably linked to a promoter.
44. A vector encoding one or more optimised gene products selected from *crtY*, *CCD1*, *BCDO*, *LcyE*, *blh* or *ybbO* operably linked to a promoter.

45. The vector as claimed in claim 44, wherein the optimised gene products are selected from the group consisting of *ΔN50-LsLcyE* and *TrxA-Osmanthus fragrans* CCD1; *crtY* and *phCCD1*; *ΔN50-LsLcyE*; *crtY* and *blh*, and *crtY*, *blh* and *ybbO*.
46. The vector as claimed in claim 45, wherein the *Osmanthus fragrans* CCD1 comprises one or more of the following mutations: N154Y, M152T, L151F.
47. A system for producing a carotenoid or apocarotenoid comprising an expression module comprising an expression vector having a coding region encoding at least one optimised carotenoid or apocarotenoid generating enzyme, the coding region being operably linked to a promoter, wherein said at least one optimised carotenoid or apocarotenoid generating enzyme is selected from:
- a. *ΔN50-LsLcyE* and *TrxA-Osmanthus fragrans* CCD1 for the production of α -ionone; or
 - b. *crtY* and *phCCD1* for the production of β -ionone; or
 - c. *ΔN50-LsLcyE* for the production of ϵ -carotene; or
 - d. *crtY* and *blh* for the production of retinal; or
 - e. *crtY*, *blh* and *ybbO* for the production of retinol.
48. The system as claimed in claim 47, further comprising:
- a first expression module comprising an expression vector having a first coding region encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY*, the first coding region being operably linked to a promoter;
 - a second expression module comprising an expression vector having second coding region encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi*, the second coding region being operably linked to a promoter;
 - a third expression module comprising an expression vector having a third coding region encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI*, the third coding region being operably linked to a promoter.

49. A kit when used in the method of any of claims 1 to 34, for the production of an apocarotenoid or carotenoid comprising one or more of:
- a first vector encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR* and optionally *crtY* operably linked to a promoter;
 - 5 a second vector encoding one or more optimised second gene products selected from *mevK*, *pmk*, *pmd* or *idi* operably linked to a promoter;
 - a third vector encoding one or more optimised third gene products selected from *ispA*, *crtE*, *crtB* or *crtI* operably linked to a promoter; and
 - a fourth vector encoding one or more optimised gene products selected from:
- 10 a. *AN50-LsLcyE* and *TrxA-Osmanthus fragrans* CCD1 operably linked to a promoter for the production of α -ionone; or
 - b. *crtY* and *phCCD1* operably linked to a promoter for the production of β -ionone; or
 - c. *AN50-LsLcyE* operably linked to a promoter for the production of ϵ -carotene; or
 - 15 d. *crtY* and *blh* operably linked to a promoter for the production of retinal; or
 - e. *crtY*, *blh* and *ybbO* for the production of retinol.
50. A method for producing a carotenoid or apocarotenoid comprising the steps of:
- 20 a. contacting a host cell as claimed in any one of claims 35 to 40 in a chemically defined media with a substrate for carotenoid or apocarotenoid production;
 - b. incubating the host cell in said chemically defined media to produce one or more preselected carotenoids or apocarotenoids, and

c. extracting the one or more preselected carotenoids or apocarotenoids from the chemically defined media using an organic layer.

51. The method as claimed in claim 50, wherein the organic layer is coconut oil or soybean oil.
- 5 52. The method as claimed in claim 50 or 51, wherein the apocarotenoid is selected from α -ionone, β -ionone, pseudo-ionone (or psi-ionone), hydroxy-ionone, β -cyclocitral, trans-geranylacetone, 6-methyl-5-hepten-2-one (MHO), retinal, retinol, 8',10-diapocarotene-8',10-dial (C17) or 10',6-diapocarotene-10',6-dial (C19) and wherein the carotenoid is selected from phytoene, lycopene, α -carotene, γ -carotene, 10 δ -carotene, ϵ -carotene or β -carotene.
53. A kit when used in the method as claimed in any of claims 1 to 34, for the production of a carotenoid comprising one or more of: a first vector encoding one or more optimised first gene products selected from *atoB*, *hmgS*, *thmgR*; a second vector encoding one or more optimised second gene products selected from *mevK*, *pmk*, 15 *pmd* or *idi* operably linked to a promoter; a third vector encoding one or more optimised third gene products selected from *ispA*, *crtE* or *crtB*, optionally *crtI*, operably linked to a promoter for the production of phytoene or lycopene.