

**METHODS FOR PRODUCING BIOCHEMICALS USING ENZYME GENES DERIVED FROM A STRAIN
OF *BREVUNDIMONAS*, AND COMPOSITIONS MADE THEREBY**

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. non-provisional application 17/395,421 filed on August 5, 2021, Japanese patent application No. 2021-033930 filed on March 3, 2021, and also claims priority to U.S. provisional application 63/130,569 filed on December 24, 2020, all three applications are incorporated herein by reference.

STATEMENT REGARDING SEQUENCE LISTING

[0002] The Sequence Listing associated with this application is provided in text format in lieu of a paper copy, and is hereby incorporated by reference into the specification. The name of the text file containing the Sequence Listing is 5627_ST25.txt. The text file is 277.5 KB, was created on September 29, 2021, and is being submitted electronically via EFS-Web.

BACKGROUND

[0003] The present disclosure is generally related to the field of molecular biology and more particularly to genetically-engineering the metabolic pathways of microorganisms to utilize various feedstocks, including gaseous feedstocks, for the biological production of biochemicals.

SUMMARY

[0004] In certain embodiments, a nucleic acid sequence is provided for expressing carotenoid products comprising any one or more of SEQ ID NOS: 1, 5, 6, 7 or 8. In certain frequent embodiments, a vector is provided comprising the nucleic acid of SEQ ID NO: 1 and a heterologous nucleic acid sequence.

[0005] In certain frequent embodiments, a nucleic acid sequence is provided that encodes an enzyme comprising an amino acid sequence that is at least 96% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 97% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 98% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -

carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 99% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin.

[0006] In frequently included embodiments, a vector is provided comprising one or more nucleic acid sequence(s) that encode(s) an enzyme comprising an amino acid sequence that is at least 96% identical to SEQ ID NO: 2, wherein when expressed the enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 97% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 98% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 99% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin.

[0007] In frequently included embodiments, a synthetic nucleic acid construct is provided comprising a promoter, a ribosome binding site, and one or more nucleic acid sequence that encode(s) an enzyme comprising an amino acid sequence that is at least 96% identical to SEQ ID NO: 2, wherein when expressed the enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 97% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 98% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. In certain related embodiments, the amino acid sequence is at least 99% identical or homologous to SEQ ID NO: 2, and the expressed enzyme is capable of converting β -carotene to canthaxanthin. Often the synthetic nucleic acid construct is a vector comprising a plasmid.

[0008] In frequent embodiments a transformed expression host organism is provided comprising the synthetic nucleic acid construct noted above and herein, and the transformed host organism is capable of heterologous expression of the synthetic nucleic acid construct. Often the expression host organism is a transformed bacteria adapted to grow in a

chemoautotrophic metabolic mode. In certain embodiments the expression host organism is *Cupriavidus necator*.

[0009] In certain embodiments a nucleic acid sequence is provided corresponding to a crtW carotenoid ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2, wherein the nucleic acid sequence is comprised in an expression construct adapted to produce carotenoids in a biological host cell. In certain frequent embodiments, the biological host cell is capable of using CO₂ and H₂ to satisfy as least part of the carbon and energy requirements of the host cell.

[0010] In certain embodiments, a nucleic acid sequence is provided corresponding to a crtZ-crtW carotenoid hydroxylase-ketolase gene fusion, wherein the crtW portion of the fusion is a ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid SEQ ID NO: 2.

[0011] In certain embodiments, a nucleic acid sequence is provided encoding a crtZ-crtW carotenoid hydroxylase-ketolase fusion protein of SEQ ID NO: 4, wherein (a) the crtW portion of the fusion is a ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2, and (b) the nucleic acid sequence is part of an expression construct adapted to produce carotenoids when functionally integrated in a biological host cell.

[0012] In certain embodiments, a nucleic acid sequence is provided encoding a crtZ-crtW carotenoid hydroxylase-ketolase fusion protein of SEQ ID NO: 4, wherein (a) the crtW portion of the fusion is a ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2, and (b) the nucleic acid sequence is part of an expression construct adapted to produce carotenoids when functionally integrated in a biological host cell, and (c) the biological host cell is capable of using CO₂ and H₂ to satisfy as least part of its carbon and energy requirements.

[0013] In certain embodiments, a suicide vector construct is provided adapted for inserting a DNA sequence into a genome of a bacterium using a transposon, the suicide vector construct comprising (a) the DNA sequence; (b) an insert-flanking DNA comprising the nucleic acid sequence of SEQ ID NO: 9 that contains the transposon; and (c) a suicide plasmid backbone. In some embodiments the suicide vector construct is adapted for inserting a DNA sequence into a

microbial genome of a bacterium using a transposon. The microbial genome can include organisms such as archaea, bacteria, and yeast.

[0014] In certain embodiments, a transformed host cell is provided comprising a nucleic acid sequence that encodes the amino acid SEQ ID NO: 2, wherein the nucleic acid sequence is part of an expression construct adapted to produce carotenoids in the host cell.

[0015] In certain embodiments, a method of forming a transformed host cell contemplated herein is provided, comprising inserting the expression construct into the genome of the host cell using a transposon. Often such insertion utilizing a transposon is a random insertion.

[0016] In certain embodiments, a nucleic acid sequence is provided corresponding to a crtW carotenoid ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2, wherein the nucleic acid sequence is part of an expression construct adapted to produce carotenoids in a cell-free expression system.

[0017] In certain embodiments, a method of producing ketocarotenoids in a biological host cell is provided by heterologous expression of OB307-crtW in the host cell. Often the biological host cell comprises a hydrogen-oxidizing bacterium. Also often the hydrogen-oxidizing bacterium comprises a strain selected from *Cupriavidus*, *Rhodobacter*, *Rhodococcus*, *Rhodopseudomonas*, *Rhodospirillum*, *Paracoccus* or *Hydrogenophaga*. In certain embodiments, the strain of hydrogen-oxidizing bacterium is *Cupriavidus necator*. In certain often included embodiments the biological host cell is cultivated as part of a consortium of different species of host cells.

[0018] In certain embodiments, a method of producing ketocarotenoids in a biological host cell is provided including transforming the biological host cell with a vector comprising a crtZ-OB307-crtW fusion, and heterologously expressing the crtZ-OB307-crtW fusion in the biological host cell to synthesize the ketocarotenoids; or heterologously expressing a crtZ-OB307-crtW fusion in the biological host cell to synthesize the ketocarotenoids. Often the biological host cell comprises a hydrogen-oxidizing bacterium. Also often the hydrogen-oxidizing bacterium comprises a strain selected from *Cupriavidus*, *Rhodobacter*, *Rhodococcus*, *Rhodopseudomonas*, *Rhodospirillum*, *Paracoccus* or *Hydrogenophaga*. In certain embodiments, the strain of

hydrogen-oxidizing bacterium is *Cupriavidus necator*. In certain often included embodiments the biological host cell is cultivated as part of a consortium of different species of host cells.

[0019] In certain embodiments, a method of producing canthaxanthin from β -carotene in vitro is provided, comprising contacting a protein expression product of a nucleic acid sequence at least 96% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8 in a solution that comprises β -carotene, wherein the protein expression product catalyzes a conversion of at least some of the β -carotene to canthaxanthin. Often the nucleic acid sequence is at least 90% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 91% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 92% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 93% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 94% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 95% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 97% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 98% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. Often the nucleic acid sequence is at least 99% identical to the nucleic acid sequence of any of SEQ ID NOS: 1, 5, 6, 7 or 8. In certain frequent embodiments the host organism is one that naturally produces β -carotene.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Figure 1 depicts a schematic of the individual enzymes and products in the biosynthetic pathway between farnesyl diphosphate (FPP) and astaxanthin. A typical carotenoid metabolic pathway includes genes *crtE*, *B*, *I*, *Y*, *Z* and *W*.

[0021] Figure 2 depicts the components of the System 1 astaxanthin operon with *crtZ* and OB-307 *crtW*.

[0022] Figure 3 depicts the components of the System 2 canthaxanthin operon, which has no *crtZ*.

[0023] **Figure 4** depicts the components of the System 3 astaxanthin operon with the *crtZW* fusion gene.

[0024] **Figure 5** depicts a detailed map of the synthetic carotenoid operon for making astaxanthin (containing the OB307-*crtW* gene) along with a Tn5 transposase gene inserted into a suicide vector, with tetracycline as the antibiotic resistance marker. The transposon is added in order to randomly insert the operon into the genome of the host cell.

[0025] **Figure 6** depicts the process of transformation and chromosomal insertion of the operon into the host cell using transposon mutagenesis. System 1 is used here as an example. (ME=Mosaic Ends (inverted repeat sequences); Pro=Promoter; ori=Origin of replication or transfer; term=Transcriptional terminator).

[0026] **Figure 7** is the HPLC chromatogram showing the canthaxanthin produced by *C. necator* cells that heterologously express the canthaxanthin biosynthesis pathway. Solid line: Cell extract. Dashed line: canthaxanthin standard.

[0027] **Figure 8** is the corresponding UV-Vis spectrum of the canthaxanthin peak shown in Figure 7.

[0028] **Figure 9** is the HPLC chromatogram showing the carotenoid products from *C. necator* cells that heterologously express the astaxanthin biosynthesis pathway. Solid line: Cell extract. Dashed line: Astaxanthin standard.

[0029] **Figure 10** is the corresponding UV-Vis spectrum of the astaxanthin peak shown in Figure 9.

DETAILED DESCRIPTION

[0030] Carotenoids are long-chain isoprenoid molecules that have nutritional advantages as colorants and additives in fish feed, animal feed and nutraceuticals because they provide protection against cellular oxidative damage, in particular against free radicals and reactive oxygen species. Carotenoids can be expressed in plants, algae, archaea, fungi and bacteria, both naturally and through the expression of one or more carotenoid genes that encode the biosynthetic enzymes. Traditional production of forty-carbon (C40) tetraterpene carotenoids, including carotenes and xanthophylls, has involved extraction of native molecules from various

microbes or plants. However, some naturally-occurring producers of astaxanthin, such as the yeast *Xanthophyllomyces* produce a less valuable enantiomer of astaxanthin and the process of growing highly productive, naturally producing microalgae, such as *Haematococcus pluvialis*, is difficult, time-consuming, resource-intensive and expensive.

[0031] Non-biological production of molecules such as astaxanthin and canthaxanthin, via chemical synthesis from petroleum feedstocks, has been achieved (Ernst, 2002). However, these latter methods produce a mixture of astaxanthin enantiomers that are also less valuable because they are less efficient radical quenchers and therapeutics and these synthetic products have faced significant regulatory issues with regard to human and animal consumption in the EU. More recently, genetically-engineered organisms have been used for the production of high-value canthaxanthin, astaxanthin and other C40 carotenoids and xanthophylls. Figure 1 shows the carotenoid biosynthesis pathway from farnesyl diphosphate (FPP) to astaxanthin.

[0032] In addition to astaxanthin, canthaxanthin is a valuable carotenoid product that can be synthesized by ketolase enzymes, such as the bacterial *crtW* ketolase gene acting on beta-carotene as its substrate. Carotenoids such as canthaxanthin and astaxanthin can be produced by ketolases encoded by *crtW* genes from various *Brevundimonas* species, which are considered to be the most active and effective carotenoid ketolases.

[0033] There is also a need for an expression system that can cheaply and efficiently produce carotenoids using this CrtW enzyme, since the yield of carotenoid per gram dry weight of biomass and rate of production is not high in natural or genetically modified organisms.

[0034] Hydrogen-oxidizing bacteria are attractive hosts for carotenoid expression because some species naturally produce larger amounts of internal membranes than many other bacteria, and these membranes are required for accumulating the highly lipophilic C40 carotenoids.

[0035] Extensive membrane capacity is also advantageous because both the CrtZ hydroxylase and the CrtW ketolase enzymes are likely integral membrane proteins that contain transmembrane (TM) helices capable of spanning cell membranes.

[0036] Furthermore, because certain hydrogen-oxidizing bacteria such as *Cupriavidus necator* do not naturally make carotenoids, there is less of a chance of regulatory interference (e.g.,

feedback inhibition) or undesirable enzymatic modification of the product (as in, for example, *Brevundimonas vesicularis* strain DC263, which naturally hydroxylates the astaxanthin product to dihydroxy-astaxanthin because it contains the *crtG* gene).

[0037] The carotenoids so produced are provided as part of the bacterial biomass or extracted from it to create a substantially pure carotenoid product, or through other extraction methods such as super critical CO₂ or solvent based extraction to form a concentrate. Further, carotenoids such as canthaxanthin can be mixed with other ingredients, such as sugars, corn starch, lignosulphonate, binders, oils or others to produce a product (*e.g.*, DSM Carophyll Red 10%).

[0038] The bacterial CrtW enzymes employ 6-8 of the following amino acid residues to bind the di-iron cofactors that catalyze the oxygenation reactions: His69, His73, His107, His 110, His111, His225, His228 and His229, as determined by the presence of the His-rich motifs HX(3 or 4)H, HX(2 or 3)HH, and HX(2 or 3)HH. Asp118 may also be required, based on mutagenesis studies. Thus, although not intending to be bound by any particular theory of operation, it is believed that natural or engineered versions of this enzyme should or must include these ligands in order to have catalytic activity. Likewise, such enzymes may require functional transmembrane sequences since there are putative TM helices that appear to organize the iron binding sites on the inside of the membrane.

[0039] Expressing such codon-optimized gene pathways in bacteria that have high G + C content has previously proved to be challenging, for example, because the GC content makes it difficult to *de novo* synthesize genes and operons for synthetic biology.

[0040] The present disclosure describes a newly discovered *crtW* gene from a new strain of *Brevundimonas*, designated herein as OB307, which encodes a novel ketolase for carotenoid synthesis. The intensely red-colored bacterial strain was isolated from a contaminating colony on a laboratory agar Petri plate. In a microscope, the cells are short rods, approximately 0.5 µm x 2.0 µm in length. This non-pathogenic, aerobic bacterium can be cultivated in a variety of different heterotrophic culture media, including Luria-Bertani broth and agar medium and MR2 minimal medium with 2% (w/v) fructose. Suitable growth conditions include shaking in a glass culture flask with a Bellco cap at 150 rpm at 30°C. This specific *crtW* gene can be utilized for the

production of useful and valuable carotenoids such as asthaxanthin and canthaxanthin from beta-carotene.

[0041] The present disclosure also provides an exemplary synthetic operon containing additional relevant carotenoid gene sequences, the expression of which is used to produce ketocarotenoids. Suitable DNA expression constructs derived from these sequences are inserted into an expression host for expression. The expression product being a ketolase enzyme that is operable for transforming beta-carotene into canthaxanthin and astaxanthin. The carotenoid products of this synthetic operon have been expressed in *Escherichia coli*, *Bacillus subtilis* B-14200, *Bacillus* B-356, *Rhodopseudomonas palustris*, *Rhodobacter sphaeroides* and *Cuprividus necator*. *R. palustris* and *R. sphaeroides* are commonly known as purple non-sulfur (PNS) bacteria. *Rhodobacter capsulatus* is another PNS bacterium that can be used as a host for these DNA expression constructs.

[0042] As disclosed herein, the presently disclosed CrtW ketolase enzyme is often utilized for production of ketocarotenoids such as astaxanthin and canthaxanthin via cloning of the disclosed DNA sequences (including similar sequences having attributes noted herein), arranging the DNA into a construct that includes a ribosome binding site, a promoter, and a terminator, as well as other structural gene elements. Other enzyme genes according to the present embodiments, such as *crtZ*, *crtY*, *crtI*, *crtB*, *crtE*, as well as additional structural and control elements are also optionally incorporated into the construct to form an operon for carotenoid production. This construct is then introduced into a host organism such as a host cell, using methods known to the art, either as one or more small, circularized DNA vectors, such as a plasmid, or via incorporation into the genome of the organism. For organisms that already produce beta-carotene, the gene encoding this single enzyme is introduced to cause the production of this CrtW ketolase enzyme and the transformation of some of the beta-carotene into canthaxanthin. If a *crtZ* gene is also introduced, the gene product (i.e., a hydroxylase) may also be expressed, and it will transform at least some of the canthaxanthin to astaxanthin.

[0043] The product of this *crtW* gene is used, for example, in a cell free expression system in which beta-carotene is enzymatically converted into canthaxanthin. If the *crtZ* and *crtW* genes

are expressed in combination, either simultaneously or sequentially, at least a portion of the beta-carotene substrate will be transformed into canthaxanthin and a portion is transformed into astaxanthin by the action of the enzyme products of the two genes. The novel *crtW* and *crtZ* genes may be provided on two different segments of DNA, or as a single piece of DNA comprising a gene for a fusion protein, which encodes both the CrtW ketolase and CrtZ hydroxylase functions.

[0044] Many different organisms are potential heterologous expression hosts for this novel *crtW* gene. Hosts that are able to utilize H₂ and CO₂ as energy and carbon sources and those that are unable to utilize H₂ and CO₂ as energy and carbon sources are contemplated as suitable heterologous expression hosts. For example, these include bacteria, plants, algae, archaea, and fungi. Bacteria such as *Escherichia coli* and *Bacillus subtilis*, fungi such *Saccharomyces cerevisiae* and *Aspergillus oryzae*, plants such as *Oryza glaberrima*, algae such as *Chlorella vulgaris*, or archaea such as *Sulfolobus solfataricus*, or others species of organism can serve as heterologous expression hosts for this novel *crtW* gene, for the production the enzyme which it encodes and for the production of the carotenoid products through the action of this enzyme.

[0045] The heterologous expression of this enzyme and the synthetic operon disclosed herein have been shown in *Escherichia coli*, *Bacillus subtilis* B-14200, *Bacillus* B-356, *Rhodopseudomonas palustris*, *Rhodobacter sphaeroides* and *Cuprividus necator* initially using a broad host range expression plasmid. In all cases, the heterologous expression of the novel OB307-*crtW* gene was observed via production of canthaxanthin in the transformed bacteria (versus no production of canthaxanthin in the wild type organism). This transformation was achieved using the same plasmid as was used in *C. necator*. The promoter disclosed herein is active in all of these strains. The *E. coli* cells were transformed using electroporation of the plasmid, as described above. The other strains were transformed using conjugation with *E. coli* strain S17-1 according to standard methods (see, e.g, Phornphisutthimas *et al.*, 2007; Gruber *et al.*, 2015). The conjugated cells were first plated on LB agar, then resuspended in sterile liquid medium with serial dilutions and plated on the following agar plates: (1) for *E. coli*, LB plus 50 µg/ml kanamycin or 10 µg/ml tetracycline; (2) for *Bacillus*, MR2 medium plus 2% fructose and 50 µg/ml kanamycin; and (3) for *C. necator* and the PNS bacteria, MR2 medium plus 2%

fructose and 500 µg/ml kanamycin. Surviving transconjugant colonies were then picked and restreaked on fresh plates until pure single colonies were obtained. Growth in liquid cultures was performed by inoculating cells of a given variant into LB plus antibiotic (for all of the strains) or MR2 plus antibiotic (for the H₂-oxidizing PNS bacteria and *C. necator*).

[0046] A fusion gene which comprised of *crtZ* and *crtW* was created by constructing a piece of synthetic DNA in which *crtZ* and *crtW* were joined by a linker sequence, and incorporating this fusion sequence into the synthetic operon in place of the original *crtW* gene in the expression plasmid. This heterologous expression vector was then transformed into *Escherichia coli* and *Cupriavidus necator*. Production of astaxanthin and canthaxanthin was observed in both cases. An allelic exchange system (using NaCl-free agar medium with 6% sucrose (w/v) for the *sacB* levansucrase counterselection) and suicide vector were also used to insert this synthetic operon into the *C. necator* genome and the production of carotenoids was again observed.

[0047] *C. necator* strain H16 has been used as an expression host, as have other *C. necator* strains, and strains of other hydrogen-oxidizing bacteria. The carotenoid products can thus be produced by gas fermentation of the transformed bacterium, using inexpensive feedstocks (e.g., waste CO₂, H₂, O₂ and mineral salts) to improve the economic efficiency of the process.

[0048] Additional genera and species of hydrogen-oxidizing bacteria that can be transformed with the vectors and DNA constructs described herein for heterologous expression in the carotenoid pathway while growing on H₂-CO₂-O₂ include, for example, *Rhodobacter capsulatus* and other *Rhodobacter* species, *Paracoccus*, *Rhodococcus*, *Hydrogenophaga*, *Rhodospirillum*, *Rhodopseudomonas*, and the like.

[0049] The novel strain of *Brevundimonas* OB307 was isolated as a red-orange contaminant colony from an agar plate in the laboratory. Its 16S rRNA genes were sequenced (forward and reverse), and compared using Clustal W to the 16S sequences of other *Brevundimonas* species. This analysis revealed that OB307 has a 99.7-99.8% identity with the 16S sequences from *B. vesicularis* and *B. nasdae*. Genomic DNA was extracted from approximately 100 mg of wet cell paste, the entire genome was sequenced using 60x Illumina paired end sequencing (150 base pair reads), and the sequence contigs were assembled and annotated by SNPsaurus, Inc.

(Eugene, OR). From this sequence, a BLAST search identified multiple genes with high similarity to other published carotenoid biosynthetic genes.

[0050] One of the complete open reading frame sequences was initially identified by the annotating software as a “fatty acid desaturase.” Fatty acid desaturases are known to have a similar structure to carotenoid ketolases, and further analysis revealed that this sequence has high similarity to *CrtW*-type carotenoid ketolases, and our subsequent expression cloning confirmed its activity. The gene sequence is therefore designated herein as OB307-*crtW* (SEQ ID NO: 1). As can be seen from the translated amino acid sequence of OB307-*CrtW*, it contains the eight-histidine motif (highlighted in yellow) and the Asp-118 (highlighted in blue) that define the di-iron binding site for this type of ketolase (SEQ ID NO: 2). SEQ ID NO: 3 shows a Clustal W 2.1 amino acid sequence alignment between OB307-*CrtW* and the *CrtW* from *Brevundimonas* strain DC263 (GenBank accession number ABC50116.1). Both proteins contain 241 amino acids, and there are 11 amino acid differences between them (about 95.5% identity). More recently, a putative *crtW* gene from *Brevundimonas* strain SgAir0440 was published as part of the genome sequence of an air-contaminating bacterium (GenBank accession number QCR00114). The gene has 99.6% similarity to the amino acid sequence of OB307-*crtW*, however, it was not reported to have been cloned and expressed, nor was the function of the enzyme analyzed to confirm that it was indeed a beta-carotene ketolase.

[0051] The native OB307-*crtW* sequence was converted into a new sequence that is codon optimized for expression in *C. necator*. This new sequence was included as part of a codon-optimized synthetic operon comprising *crtE*, *crtY*, *crtI*, *crtB*, and *crtW*, which makes canthaxanthin (Figure 3). Constructs designed to make astaxanthin also included the complete *crtZ* sequence (Figure 2). The other gene sequences in the pathway were sourced from various other bacteria, with the GenBank accession numbers as follows: the genes *crtE*, *crtY*, *crtI*, and *crtB* were synthesized from the sequence of the *Pantoea agglomerans/Erwinia herbicola* pAC-BETA plasmid, M8720/M99707; *crtZ* was synthesized from the sequence of *Pantoea ananatis* Strain AJ13355, NC_017533; and *crtW* was synthesized from the sequence of OB307-*crtW* described herein.

[0052] Synthesis of the operon benefits from a specialized procedure (e.g., as available from Aster Bioscience, Inc.; Livermore, CA) due to the very high G + C content (ca. 61%-70%). A constitutive promoter that is highly active in *C. necator* was placed upstream of the carotenoid genes to direct mRNA synthesis in the cell. Other suitable promoters are well known in the art and contemplated herein. Inducible promoters, which can be used to control the timing of the onset of gene transcription by applying an external inducer molecule (e.g., IPTG for the *lac* or *tac* promoters) or an environmental stimulus (e.g., nitrogen deprivation for the *phaC1* promoter) can also be used, if they are compatible with the metabolism and transport system of the host. Ribosome binding sites (RBSs) optimized for *C. necator* were placed upstream of each gene sequence. Spacer sequences were added between the promoter and the RBS of the *crtE* gene, as well as between the RBS and the start codon of each individual gene, in order to optimize the overall expression. A termination sequence (*E. coli rrnB*) was placed at the end of the operon to prevent unwanted translation of any downstream elements.

[0053] The synthetic operons (SEQ ID NO: 6, 7 and 8) were first tested for activity by cloning them into the broad host range plasmid pBBR1MCS-2 (e.g., kanamycin as a selection), using *NdeI* and *Asel* as the flanking restriction sites. The ligated DNA products were transformed into *E. coli* by electroporation using a Bio-Rad GenePulser II with a Capacitance Extender Plus Pulse Controller II unit (Bio-Rad Inc., Hercules, CA). *E. coli* cells were made electrocompetent using three washes with cold 10% glycerol according to the methods described in the online protocol of Belcher and Knight (https://openwetware.org/wiki/Belcher/Knight:_Electrocompetent_Cells). 50 μ l of electrocompetent cells were added to a chilled 1 mm gap sterile cuvette and mixed with 1 μ l of DNA (approximately 1-50 ng). The electroporator settings were as follows: 1.2 kV, 25 μ F, 200 Ω . The time constant was typically 3-5 msec. After pulsing, the cells were then transferred to pre-warmed SOB medium in a small sterile tube and allowed to recover at 37°C for 1 hour with shaking. Aliquots were then plated on LB agar with 50 μ g/ml kanamycin for antibiotic selection. After incubation at 30°C, colonies were picked and individually grown up in LB broth. Plasmid DNA was isolated from the various clones by standard methods. The DNA was cut with the appropriate restriction enzymes and analyzed by agarose gel electrophoresis to identify the positive clones. Plasmid DNA from one correct clone was transformed into *E. coli* conjugation

strain S17-1. The process described above was then repeated to find correct S17-1 clones. An S17-1 clone containing the synthetic canthaxanthin or astaxanthin operon in the plasmid pBBR1MCS-2 was then conjugated into the *C. necator* host strain or other host strains by standard methods as described above. After plating on solid MR2-fructose medium (Table 1) containing 500 µg /ml kanamycin, *C. necator* colonies appeared. Colonies that displayed a deep orange or red color were picked and re-streaked on kanamycin plates to confirm their colored phenotype and antibiotic resistance. Selected clones were picked and grown up in liquid medium with antibiotic.

TABLE 1. Composition of MR2 medium

Name of Chemical	Formula	Mass (g)	Moles (mol)	Molar Mass (g/mol)
Sodium Phosphate Dibasic	Na ₂ HPO ₄	4.00E+00	2.82E-02	141.96
Potassium Phosphate Monobasic	KH ₂ PO ₄	6.67E+00	4.90E-02	136.09
Ammonium Chloride	NH ₄ Cl	1.80E+00	3.37E-02	53.49
Sodium Bicarbonate	NaHCO ₃	2.00E-01	2.38E-03	84.01
Ammonium Nickel (II) Sulfate Hexahydrate	(NH ₄) ₂ Ni(SO ₄) ₂ *6H ₂ O	1.07E-02	3.05E-05	349.99
<i>Add after autoclaving:</i>				
Iron (II) Sulfate Heptahydrate	FeSO ₄ *7H ₂ O	1.20E-02	4.32E-05	278.01
Calcium Chloride Dihydrate	CaCl ₂ *2H ₂ O	2.00E-02	1.36E-04	147.01
Magnesium Sulfate Heptahydrate	MgSO ₄ *7H ₂ O	1.13E+00	4.59E-03	246.47
<i>Trace Elements</i>				
Disodium EDTA	C ₁₀ H ₁₄ N ₂ Na ₂ O*2H ₂ O	1.00E-03	2.69E-06	372.24
Iron (II) Sulfate Heptahydrate	FeSO ₄ *7H ₂ O	4.00E-04	1.44E-06	278.01
Zinc Sulfate Heptahydrate	ZnSO ₄ *7H ₂ O	2.00E-05	6.96E-08	287.56
Manganese (II) Chloride Tetrahydrate	MnCl ₂ *4H ₂ O	6.00E-06	3.03E-08	197.91
Boric Acid	H ₃ BO ₃	6.00E-05	9.70E-07	61.83

Cobalt (II) Chloride Hexahydrate	CoCl ₂ *6H ₂ O	4.00E-05	1.68E-07	237.93
Copper (II) Chloride Dihydrate	CuCl ₂ *2H ₂ O	2.00E-06	1.17E-08	170.48
Nickel (II) Chloride Hexahydrate	NiCl ₂ *6H ₂ O	4.00E-06	1.68E-08	237.69
Sodium Molybdate Dihydrate	Na ₂ MoO ₄ *2H ₂ O	6.00E-06	2.48E-08	241.95

[0054] As described above, the processivity of the enzymes at the end of the pathway for the production of astaxanthin can be improved by genetically fusing the genes for *crtZ* and *crtW* to encode for a chimeric protein. The fusion protein sequence was created by inserting the DNA sequence for a short linker peptide (encoding amino acid sequence GGGGSGGPGS) between the 3' end of the complete *crtZ* gene from *Pantoea ananatis* and the 5' end of OB307-*crtW* gene (without the N-terminal methionine), as shown in the map of Figure 4, as well as SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 8. The *crtZ-crtW* fusion sequence was codon optimized, synthesized, and used to replace the *crtW* gene in the original operon construct to create the insert known as System 3 (SEQ ID NO: 8). When the expression plasmid encoding this sequence was transformed into a suitable host as described above, the cells expressed astaxanthin (Figure 9 and Figure 10).

[0055] In certain embodiments the pathways contemplated herein are improved by genetic modification, in particular by methods of directed evolution, for example via random mutagenesis and library screening to identify improved variants. Strain engineering of the host genome can also be used to improve expression of the recombinant pathway genes.

[0056] In certain embodiments the operon is inserted into the genome semi-randomly and then screened for production levels. In the case of carotenoid production, this screening can be done by looking for intense color production in colonies from plated libraries of transformants. Accordingly, a custom suicide vector was constructed (based on the non-replicating, allelic exchange plasmid of Hmelo *et al.* (2015)) so that the operon could be inserted between the mosaic ends (inverted 19-bp inside and outside end sequences) of the phage Tn5 transposon by restriction cloning with *NdeI* and *NsiI*. A Tn5 transposase sequence was also inserted into the plasmid (using Gateway cloning), along with a tetracycline resistance cassette to act as an antibiotic marker (see, e.g., Figure 5, Figure 6 and SEQ ID NO: 9). The transposon suicide vector was assembled, transformed into *E. coli* strain S17-1, and then conjugated into *C. necator* strain

H16 as described above. Transconjugants were plated on MR2 agar plus 2% fructose and 10 µg/ml tetracycline as described above, followed by a second plating on LB agar plus 50 µg/ml kanamycin or MR2 agar plus 2% fructose and 50 µg/ml kanamycin to remove the *E. coli* donor. Orange and red colored colonies were picked for further characterization of their carotenoids as described above. A variety of pale and intensely colored colonies are observed, indicating that the operon has been inserted into a different genome location in each of the clones that expresses carotenoid.

[0057] To rapidly confirm initial expression of the pathway and production of the carotenoid products, *C. necator* clones with the pBBRMCS-2 expression plasmid were inoculated into 50 ml of sterile liquid minimal medium (MR2 at pH 6.8) at 30°C in shake flasks with 20 g/L fructose added as a carbon source. After approximately 48 hours of growth, the cultures achieved an A620 (optical density measured at 620 nm) of approximately 1.4, and they exhibited a deep orange or red color due to production of carotenoids. Other expression hosts transformed with the expression plasmid, such as *Bacillus subtilis* strain NRRL B-14200, *Bacillus subtilis* strain NRRL B-354, *Rhodopseudomonas palustris* strain NRRL B4276, and *Rhodobacter sphaeroides* strain NRRL B1727, have also been cultivated in this way. NRRL strains were obtained from the USDA-ARS Culture Collection (Peoria, IL).

[0058] To evaluate production of carotenoid on gas, cells containing the genomically integrated operon were inoculated into 200-500 ml of sterilized MR2 minimal medium at pH 6.8 (with no carbon source) in a capped, stirred flask (magnetic stir bars) equipped with submerged gas inlets and an exit port. The sterilized external gas inlets, outlets and rubber tubing were capped with sterile disk filters (0.2 µm pore size; cellulose acetate syringe filter, VWR) to prevent contamination from the outside atmosphere. A mixture of H₂:CO₂:O₂ with an approximate ratio of 80:10:10 was supplied by commercial gas cylinders (Praxair, Inc.), or by electrolytic hydrogen from a generator (Parker Dominick Hunter Model 40H; Charlotte, NC). In some embodiments, the CO₂ (often containing other gases, such as H₂, CO, SO_x, NO_x) was collected as waste CO₂ from cement manufacturing, fossil fuel combustion, petrochemical hydrocracking operations and the like, and was supplied in pressurized cylinders. The gas mixing and gas flow rates were controlled by small network of gas flow meters and mass flow controllers (Alicat Scientific, Inc.,

Tucson, AZ). The stir plates and flasks were housed in incubators maintained at 30°C. The exit gas was collected and vented to the outside air. Cultures were grown for 72 hours until the cells reached an A620 of approximately 0.4 and turned noticeably red or orange in color. At commercial scale, this type of cultivation is performed in loop bioreactors specially designed for high-volume cultures grown entirely on gas. An example of a loop bioreactor for gas fermentation of methanotrophs (using methane and oxygen as feedstocks) is provided in Petersen *et al.* (2017, 2020). In another embodiment, the fermentation and cultivation of the host cells expressing the carotenoid genes employs a consortium (*i.e.*, a mixture of different species) so as to improve the growth rate of the carotenoid-containing biomass or improve the overall characteristics of the biomass.

[0059] *Production using cell-free systems.* It is contemplated that the enzymes and constructs provided in the present disclosure are used to express the pathway enzymes and generate the carotenoid products using cell-free expression systems (Schneider *et al.*, 2010; Gregorio *et al.*, 2019; Khambati *et al.*, 2019). Such a system can, for example, be fed with the simple precursors of the carotenoid pathway, such as IPP and DMAP and FPP, and convert these compounds into the more valuable ketocarotenoid products. Cell free expression refers to an agent that, when combined with a polynucleotide, permits *in vitro* translation of the polypeptide or protein encoded by the polynucleotide. These systems are known in the art and exist for both eukaryotic and prokaryotic applications. Exemplary cell free expression systems that can be used in connection with the present disclosure include, for example, commercial kits for various species such as extracts available from Invitrogen Ambion, Qiagen and Roche Molecular Diagnostics, cellular extracts made from hydrogen oxidizing bacteria, including a strain selected from *Cupriavidus*, *Rhodobacter*, *Rhodococcus*, *Rhodopseudomonas*, *Rhodospirillum*, *Paracoccus* or *Hydrogenophaga*, in addition to *E. coli* and other strains.

[0060] Cells were harvested by centrifugation at 6,000 x *g* for 10 minutes. After resuspending in phosphate buffered saline, the cells were centrifuged again. An aliquot of the washed cell pellet was extracted with *n*-hexane/methanol (1:1 v/v) in a 1.5 ml microcentrifuge tube. The solvent extract was separated from the cell debris by centrifugation at 14,000 x *g* for 5 minutes.

Carotenoids can also be efficiently isolated and purified from biomass using supercritical CO₂ extraction (Valderrama *et al.*, 2003; Di Sanzo *et al.*, 2018).

[0061] *Carotenoid analysis.* For identifying and assaying the production of carotenoids, 50 µl of solvent-extracted sample was loaded via syringe onto a Symmetry C18 5 µm (4.6 x 250 mm) HPLC (high-performance liquid chromatography) column, which was pre-equilibrated with a solution containing methanol/water 90:10 (v/v). The running solution was composed of a gradient of water, ethyl acetate, and water. The HPLC instrument was a Beckman System Gold equipped with a 168NM diode array detector. The running conditions were as follows: Flow rate: 1 mL/min; Temperature: 30°C. Peaks were identified by comparing their retention times with solutions of known carotenoid standards dissolved in *n*-hexane. Canthaxanthin standard was obtained from Honeywell Research, Inc.; astaxanthin was from Abcam (Cambridge, MA). Eluted components can also be identified, where possible, by their characteristic absorbance spectra. Sample chromatograms of canthaxanthin (Figure 7) and astaxanthin (Figure 9), as well as their corresponding UV-Vis absorption spectra (Figures 8 and 10), produced using the expression system of the present disclosure are shown. These experiments confirm that the OB307-crtW gene does encode a beta-carotene ketolase, and that the constructs expressing the new OB307-crtW gene do indeed produce canthaxanthin and astaxanthin.

[0062] This *crtW* sequence sometimes requires codon optimization when the gene is heterologously expressed in various expression hosts, in order to produce sufficient amounts of active enzyme to catalyze the transformation of beta-carotene to canthaxanthin. This is also true for the synthetic operon and for constructs where the gene sequences are arranged to produce fusion proteins, such as *crtZ-crtW* fusion proteins. In some embodiments of the present disclosure, the expression host is a plant. In some embodiments the expression host is a fungus, such as *Saccharomyces cerevisiae*. In some embodiments, the expression host is an alga, such as *Chlorella vulgaris*. In some embodiments, the expression host is a bacterium, such as a methylotroph (*e.g.*, *Methylobacterium extorquens*), a methanotroph, (*e.g.*, *Methylococcus capsulatus*), an acetogen (*e.g.*, *Clostridium autoethanogenum*), a hydrogen-oxidizing bacterium (*e.g.*, *Cupriavidus necator*), or a purple non-sulfur bacterium, such as *Rhodospirillum rubrum*, *Rhodobacter sphaeroides*, *Rhodobacter capsulatus*, or *Rhodopseudomonas palustris*. Other

potentially suitable bacterial hosts include *Rhodococcus opacus*, a *Paracoccus* species, such as *Paracoccus zeaxanthinifaciens*, or *Escherichia coli*.

[0063] In the foregoing specification, the invention is described with reference to specific embodiments thereof, but those skilled in the art will recognize that the invention is not limited thereto. Various features and aspects of the above-described invention may be used individually or jointly. Further, the invention can be utilized in any number of environments and applications beyond those described herein without departing from the broader spirit and scope of the specification. The specification and drawings are, accordingly, to be regarded as illustrative rather than restrictive. It will be recognized that the terms “comprising,” “including,” and “having,” as used herein, are specifically intended to be read as open-ended terms of art.

REFERENCES

- [0064] Di Sanzo, G *et al.* (2018) Supercritical Carbon Dioxide Extraction of Astaxanthin, Lutein, and Fatty Acids from *Haematococcus pluvialis* Microalgae. *Mar Drugs* 16:334.
- [0065] Ernst, H (2002) Recent Advances in Industrial Carotenoid Synthesis. *ChemInform* 74:2213-2226.
- [0066] Gregorio, NE *et al.* (2019) A User's Guide to Cell-Free Protein Synthesis. *Methods Protoc.* 2:24.
- [0067] Gruber, S *et al.* (2015) Versatile plasmid-based expression systems for Gram-negative bacteria--General essentials exemplified with the bacterium *Ralstonia eutropha* H16. *New Biotechnol* 32: 552-8.
- [0068] Hmelo, LR, Borlee, BR, Almblad, H, *et al.* (2015) Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nat Protoc* 10:1820-1841.
- [0069] Khambhati, K *et al.* (2019) Exploring the Potential of Cell-Free Protein Synthesis for Extending the Abilities of Biological Systems. *Front Bioen Biotechnol* 7:248.
- [0070] Petersen, LAH *et al.* (2017) Mixing and mass transfer in a pilot scale U-loop bioreactor. *Biotechnol Bioeng.* 114:344-354.
- [0071] Petersen, LAH *et al.* (2020) Modeling and system identification of an unconventional bioreactor used for single cell protein production. *Chem Eng J* 390:124438.

[0072] Phornphisutthimas, S *et al.* (2007) Conjugation in *Escherichia coli*—A laboratory exercise. *Biochem Mol Biol Educ* 35:440-5.

[0073] Schneider, B *et al.* (2010) Membrane Protein Expression in Cell-Free Systems. In: *Heterologous Expression of Membrane Proteins, Methods in Molecular Biology*, vol. 601 (I. Mus-Veteau, *ed.*), Humana Press, Springer Nature, Switzerland.

[0074] Valderrama, J O *et al.* (2003) Extraction of Astaxantine and Phycocyanine from Microalgae with Supercritical Carbon Dioxide *J Chem Eng Data* 48:827-830.

SEQUENCE LISTING

SEQ ID NO: 1 [OB307-crtW beta-carotene ketolase]

213: Unknown

220:

221: Gene (crtW)

222: Derived from Brevundimonas strain OB307

223: Bacterium of the genus Brevundimonas

ATGTCCGCCGTACGCCAATGTCACGGGTCGTCCCGAACCAGGCCCTGATCGGTCTGACG
CTGGCTGGCCTGATCGCGACGGCCTGGCTGAGCCTGCATATCTACGGCGTCTATTTTCAT
CGCTGGACGATGTGGAGCATCCTGACCGTCCGCTAATCGTCGCTTCCAGACCTGGCTG
TCCGTGGCCTGTTTCATCGTCGCCCACGACGCCATGCACGGCTCTCTGGCTCCGGGACGC
CCTCGGCTGAACACGGCGATCGGCAGCCTGGCGCTGGGCCTCTACGCCGTTTTTCGTTTT
GCGCCGTTGAAGACGGCGCACACGCTCATCATGCCGCGCCCGGCACGGCGGACGACCCC
GACTTTCACGCCGACGCCCCGCGCGCCTTCTGCCCTGGTTCTACGGCTTTTTCCGTACC
TATTTTCGGTTGGCGCGAGTTGGCGTTCTGACGGTGCTCGTGGCCGTCGCAGTGCTGATC
CTTGGCGCCCGCATGCCAATCTTCTGGTCTTCTGGGCCGCGCCCGCCCTGCTCTCGGCG
CTACAGCTTTTCACATTCGGCACCTGGCTGCCTCACAGGCATACCGACGACGCCTTCCCC
GACCACCACAACGCGCCGACCGACCCCTTGGGCCGATCCTGTGTTGCTGACCTGCTTC
CACTTCGGCCGCCACCACGAACACCACCTGACCCCTGGAAGCCCTGGTGGCGTCTTTTC
AGCTAG

SEQ ID NO: 2 [OB307-CrtW amino acid sequence]

213: Unknown

220:

221: Amino acid sequence

222: Derived from Brevundimonas strain OB307 crtW

223: Bacterium of the genus Brevundimonas

MetSerAlaValThrProMetSerArgValValProAsnGlnAlaLeu

IleGlyLeuThrLeuAlaGlyLeulleAlaThrAlaTrpLeuSerLeu
HisIleTyrGlyValTyrPheHisArgTrpThrMetTrpSerIleLeu
ThrValProLeulleValAlaPheGlnThrTrpLeuSerValGlyLeu
PhelleValAlaHisAspAlaMetHisGlySerLeuAlaProGlyArg
ProArgLeuAsnThrAlalleGlySerLeuAlaLeuGlyLeuTyrAla
GlyPheArgPheAlaProLeuLysThrAlaHisHisAlaHisHisAla
AlaProGlyThrAlaAspAspProAspPheHisAlaAspAlaProArg
AlaPheLeuProTrpPheTyrGlyPhePheArgThrTyrPheGlyTrp
ArgGluLeuAlaValLeuThrValLeuValAlaValAlaValLeulle
LeuGlyAlaArgMetProAsnLeuLeuValPheTrpAlaAlaProAla
LeuLeuSerAlaLeuGlnLeuPheThrPheGlyThrTrpLeuProHis
ArgHisThrAspAspAlaPheProAspHisHisAsnAlaArgThrSer
ProPheGlyProIleLeuSerLeuLeuThrCysPheHisPheGlyArg
HisHisGluHisHisLeuThrProTrpLysProTrpTrpArgLeuPhe
Ser

SEQ ID NO: 3

213: Unknown

220:

221: Amino acid sequences

222: Sequence alignment of crtW from Brevundimonas strain OB307 and crtW from Brevundimonas strain DC263

223: Bacterium of the genus *Brevundimonas*

CLUSTAL 2.1 multiple sequence alignment

OB307-crtw MSAVTPMSRVVPNQALIGLTLAGLIATAWLSLHIYGVYFHRWTMWSILTVPLIVAFQTWL

DC263-crtW MSAVTPMSRVVPNQALIGLTLAGLIAAWLTLHIYGVYFHRWTIWSVLTVPLIVAGQTWL

*****.*****.*****

OB307-crtw SVGLFIVAHDAMHGSLAPGRPRLNTAIGSLALGLYAGFRFAPLKTAHHAHHAAPGTADDP

DC263-crtW SVGLFIVAHDAMHGSLAPARPRNLNTAIGSLALALYAGFRFTPLKTAHHAHHAAPGTADDP

*****.*****.*****.*****

OB307-crtw DFHADAPRAFLPWFYGGFFRTYFGWRELAVLTVLVAVAVLILGARMNPNNLVFWAAPALLSA

DC263-crtW DFHADAPRAFLPWFYGGFFRTYFGWRELAVLTVLVAVAVLILGARMNPNNLVFWAAPALLSA

OB307-crtw

LQLFTFGTWLPHRHTDDAFPDHHNARTSPFGPILSLTCHFGRHHEHHLTPWKPWWRFLF

DC263-crtW

LQLFTFGTWLPHRHTDDAFPDNHNARTSPFGPVLSLTCHFGRHHEHHLTPWKPWWSLF

*****.*****.*****

OB307-crtw S

DC263-crtW S

*

SEQ ID NO: 4 [CrtZ--Linker--OB307-CrtW amino acid sequence]:

213: Unknown

220:

221: Amino acid sequence

222: Derived from the Pantoea ananatis crtZ amino acid sequence (1-175), a ten amino acid synthetic linker peptide (176-185), and the Brevundimonas strain OB307 crtW sequence without the N-terminal methionine residue (186-425).

223: Bacterium of the genus Brevundimonas

MetLeuTrpIleTrpAsnAlaLeuIleValPheValThrValIleGly

MetGluValValAlaAlaLeuAlaHisLysTyrIleMetHisGlyTrp

GlyTrpGlyTrpHisLeuSerHisHisGluProArgLysGlyAlaPhe

GluValAsnAspLeuTyrAlaValValPheAlaAlaLeuSerIleLeu
LeuIleTyrLeuGlySerThrGlyMetTrpProLeuGlnTrpIleGly
AlaGlyMetThrAlaTyrGlyLeuLeuTyrPheMetValHisAspGly
LeuValHisGlnArgTrpProPheArgTyrIleProArgLysGlyTyr
LeuLysArgLeuTyrMetAlaHisArgMetHisHisAlaValArgGly
LysGluGlyCysValSerPheGlyPheLeuTyrAlaProProLeuSer
LysLeuGlnAlaThrLeuArgGluArgHisGlyAlaArgAlaGlyAla
AlaArgAspAlaGlnGlyGlyGluAspGluProAlaSerGlyLysGly
GlyGlyGlySerGlyGlyProGlySerSerAlaValThrProMetSer
ArgValValProAsnGlnAlaLeuIleGlyLeuThrLeuAlaGlyLeu
IleAlaThrAlaTrpLeuSerLeuHisIleTyrGlyValTyrPheHis
ArgTrpThrMetTrpSerIleLeuThrValProLeuIleValAlaPhe
GlnThrTrpLeuSerValGlyLeuPheIleValAlaHisAspAlaMet
HisGlySerLeuAlaProGlyArgProArgLeuAsnThrAlaIleGly
SerLeuAlaLeuGlyLeuTyrAlaGlyPheArgPheAlaProLeuLys
ThrAlaHisHisAlaHisHisAlaAlaProGlyThrAlaAspAspPro
AspPheHisAlaAspAlaProArgAlaPheLeuProTrpPheTyrGly
PhePheArgThrTyrPheGlyTrpArgGluLeuAlaValLeuThrVal
LeuValAlaValAlaValLeuIleLeuGlyAlaArgMetProAsnLeu
LeuValPheTrpAlaAlaProAlaLeuLeuSerAlaLeuGlnLeuPhe
ThrPheGlyThrTrpLeuProHisArgHisThrAspAspAlaPhePro
AspHisHisAsnAlaArgThrSerProPheGlyProlleLeuSerLeu
LeuThrCysPheHisPheGlyArgHisHisGluHisHisLeuThrPro
TrpLysProTrpTrpArgLeuPheSer

SEQ ID NO: 5 [crtZ--Linker--OB307-crtW DNA sequence]:

213: Unknown

220:

221: Nucleic acid sequence

222: Synthetic nucleotide sequence derived from the *Pantoea ananatis* crtZ amino acid sequence (1-525), a synthetic linker sequence(526-555), and the *Brevundimonas* strain OB307 crtW sequence without the N-terminal methionine residue (556-1275).

223: Bacterium of the genus *Brevundimonas*

ATGCTGTGGATCTGGAACGCCCTGATCGTTTTCTGACCGTGATCGGCATGGAAGTGGTG
GCCGCCCTGGCCATAAGTACATCATGCACGGCTGGGGCTGGGGCTGGCACCTGTCGCAC
CACGAACCACGCAAAGGCGCATTTGAGGTGAATGACCTGTATGCCGTGGTGTTCGCCGCC
CTGTGATTCTGTGATCTATCTGGGCTCGACTGGCATGTGGCCGCTGCAGTGGATTGGC
GCCGGCATGACCGCATACGGCCTGCTGTACTTTATGGTTCATGACGGCCTGGTGACCCAG
CGCTGGCCGTTCCGCTACATCCCGCGCAAAGGCTATCTGAAACGCCTGTACATGGCCCCAC
CGCATGCACCATGCAGTGC GCGGCAAGGAGGGCTGTGTGTCATTGGCTTTCTGTACGCC
CCGCCGCTGTGCAAGCTGCAGGCCACTCTGCGCGAGAGACATGGCGCCCGCGCCGGCGCA
GCCCCGCGATGCCCAAGGCGGCGAGGACGAGCCGGCATCGGGCAAAGGCGGGGGCGGGTCC
GGCGGGCCCGGGTCTGCGGCCGTGACCCCGATGTCGAGAGTGGTGCCAAACCAGGCCCTA
ATCGGCCTGACTTTAGCGGGGCTGATAGCCACGGCGTGGCTGAGTCTGCATATTTACGGG
GTGTACTTCCATCGTTGGACAATGTGGTCGATCCTGACGGTGCCGCTGATCGTGGCCTTC
CAGACGTGGCTGTCGGTAGGCCTGTTTCATCGTTGCCACGACGCAATGCACGGCTCCCTA
GCCCCGGGGAGGCCCCGCCTGAACACCGCCATCGGGTCCCTGGCCCTAGGCCTGTACGCT
GGCTTCAGGTTGCCCCCTCTGAAGACCGCCCATGCCACCATGCCGCACCGGGCACA
GCCGACGACCCGGATTTTCACGCGGACGCCCCCGTGCGTTCCTGCCGTGGTCTACGGC
TTTTTCCGTACCTACTTCGGCTGGAGGGAGCTGGCCGTGCTGACCGTGTGGTGCCGTG
GCTGTTTTAATCCTGGGCGCCCGAATGCCGAACCTACTTGTGTTCTGGGCCGCCCCGGCT
CTATTATCGGCCTTGCACTTTTCACCTTCGGCACATGGCTGCCGCACCGACACACCGAC
GACGCCTTCCCGGACCACCACAACGCTCGCACTTACCCTTTGGCCCCATCCTGTCTCTG
CTGACCTGCTTCACTTCGGCCGGCACCATGAGCACCACTGACTCCGTGGAAACCGTGG
TGGAGGCTGTTCTCGTAG

SEQ ID NO: 6 [System 1, insert only, 6449 bp]:

213: Unknown

220:

221: Nucleic acid sequence

222: Synthetic nucleotide sequence derived from the Pj5[E1A1C1C2] promoter (1-327), codon-optimized crtE from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (328-1,251), spacer sequence (1,252-1,291), RBS (1,292-1,305), codon-optimized crtY from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (1,306-2,466), spacer sequence (2,467-2,509), RBS (2,510-2,523), codon-optimized crtI from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (2,524-4,002), spacer sequence (4,003-4,046), RBS (4,047-4,060), codon-optimized crtB from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (4,061-4,990), spacer sequence (4,991-5,031), RBS (5,032-5,045), codon-optimized crtZ from *Pantoea ananatis* Strain AJ13355 NC_017533 in plasmid pEA-320 (5,046-5,573), spacer sequence (5,574-5,612), RBS (5,613-5,626), codon-optimized crtW from *Brevundimonas* strain OB307 (5,627-6,352), ending spacer sequence (6,353-6,371), *E. coli* rrnB terminator (6,372-6,443), and *Asel* restriction site (6,444-6,449).

223: Synthesized

AGTCCATTGTTGCCTTGCAACGCACGCGCTGTCAATGCGGGAATCCGCCTCGGCACTGCA
CGCTTCCCGACCTACCGGACGGTATGCAGCGCTCGCATCTGCCGAGGCCCCAGAGCATAG
GCGAGAAGGATGAATTTTGTATGTACATCGTGGCCATTGCTGCAGAGCGGATATAAAAAC
CGTTATTGACACAGGTGGAAATTTAAAATATACTGTTAGTAAACCTAATGGATCGACCTT
GAATTCAAAAGATCTGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGAATTCAA
AAGATCTTTTAAGAAGGAGATATACATATGGTGTGCGGCTCGAAGGCCGGCGTGTGCGCG
CACCGCGAGATCGAGGTGATGCGCCAGTCGATCGACGACCACCTGGCCGGCCTGCTGCCG
GAGACCGACTCGCAGGACATCGTGTGCTGGCCATGCGCGAGGGCGTGATGGCCCCGGGC
AAGCGCATCCGCCCGCTGCTGATGCTGCTGGCCGCCCGCGACCTGCGCTACCAGGGCTCG
ATGCCGACCCTGCTGGACCTGGCCTGCGCCGTGGAGCTGACCCACACCGCCTCGCTGATG
CTGGACGACATGCCGTGCATGGACAACGCCGAGCTGCGCCGCGGCCAGCCGACCCAC
AAGAAGTTCGGCGAGTCGGTGGCCATCCTGGCCTCGGTGGGCCTGCTGTGAAGGCCTTC
GGCCTGATCGCCGCCACCGGCGACCTGCCGGGCGAGCGCCGCGCCAGGCCGTGAACGAG
CTGTGACCGCCGTGGGCGTGCAGGGCCTGGTGTGCGGCCAGTTCGCGACCTGAACGAC

GCCGCCCTGGACCGCACCCCGGACGCCATCCTGTCGACCAACCACCTGAAGACCGGCATC
CTGTTCTCGGCCATGCTGCAGATCGTGGCCATCGCCTCGGCCTCGTCGCCGTGACCCGC
GAGACCCTGCACGCCTTCGCCCTGGACTTCGGCCAGGCCTTCCAGCTCCTGGACGACCTG
CGCGACGACCACCCGGAGACCGGCAAGGACCGCAACAAGGACGCCGGCAAGTCGACCCTG
GTGAACCGCTGGGCGCCGACGCCGCCGCCAGAAGCTGCGCGAGCACATCGACTCGGCC
GACAAGCACCTGACCTTCGCCTGCCCGCAGGGCGGGCGCCATCCGCCAGTTCATGCACCTG
TGGTTCGGCCACCACCTGGCCGACTGGTCGCCGGTGATGAAGATCGCCTGAGTCATAGCT
GTTTCCTGCCAGTCACGACGTTGTAAACGCAAAGGAGATATAGGTGCGCGACCTGATC
CTGGTGGGCGGGCGCCTGGCCAACGGCCTGATCGCCTGGCGCCTGCGCCAGCGCTACCCG
CAGCTCAACCTGCTGCTGATCGAGGCCGGCGAGCAGCCGGGCGGCAACCACACCTGGTCG
TTCCACGAGGACGACCTGACCCCGGGCCAGCACGCCTGGCTGGCCCCGCTGGTGGCCAC
GCCTGGCCGGGTACGAGGTGCAGTTCGCCGACCTGCGCCGCCGCTGGCCCGGGCTAC
TACTCGATCACCTCGGAGCGCTTCGCCGAGGCCCTGCACCAGGCCCTGGGCGAGAACATC
TGGCTGAAGTCTCGGTGTCGGAGGTGCTGCCGAAGTGGTGCCTGGCCAACGGCGAG
GCCCTGCTGGCCGGCGCCGTGATCGACGGCCGGCGGTGACCGCCTCGTCGGCCATGCAG
ACCGGCTACCAGCTCTTCTGGGCCAGCAGTGGCGCCTGACCCAGCCGCACGGCCTGACC
GTGCCGATCCTGATGGACGCCACCGTGGCCAGCAGCAGGGCTACCGCTTCGTGTACACC
CTGCCGCTGTCGGCCGACACCCTGCTGATCGAGGACACCCGCTACGCCAACGTGCCGAG
CGCGACGACAACGCCCTGCGCCAGACCGTGACCGACTACGCCCACTCGAAGGGCTGGCAG
CTCGCCAGCTCGAACGCGAGGAGACCGGCTGCCTGCCGATCACCTGGCCGGCGACATC
CAGGCCCTGTGGGCCGACGCCCCGGGCGTGCCGCGCTCGGGCATGCGCGCCGGCCTGTT
CACCCGACCACGGCTACTCGCTGCCGCTGGCCGTGGCCCTGGCCGACGCCATCGCCGAC
TCGCCGCGCCTGGGCTCGGTGCCGCTGTACCAGCTCACCCGCCAGTTCGCCGAGCGCCAC
TGGCGCCGCCAGGGCTTCTCCGCCTGCTGAACCGCATGCTGTTCTGGCCGGCCGCGAG
GAGAACCGCTGGCGCGTGATGCAGCGCTTCTACGGCCTGCCGGAGCCGACCGTGGAGCGC
TTCTACGCCGGCCGCTGTCGCTGTTGACAAGGCCCGCATCCTGACCGGCAAGCCGCCG
GTGCCGCTGGGCGAGGCCTGCCGCGCCGCCCTGAACCACTTCCCGGACCGCCGCGACAAG
GGCTGACCTGTGTGAAATTGTTATCCGCTTACCCATACGACGTCCCAGACAAAGGAGATA
TAGATGAAGAAGACCGTGGTGATCGGCGCCGGCTTCGGCGGCCTGGCCCTGGCCATCCGC

CTGCAGGCCGCCGGCATCCCCACCGTGCTGCTGGAGCAGCGCGACAAGCCGGGGCGGCCGC
GCCTACGTGTGGCAGCAGGAGGCTTCACCTTCGACGCCGGCCCGACCGTGATCACCGAC
CCGACCGCCCTGGAGGCCCTGTTACCCCTGGCCGGCCCGCATGGAGGACTACGTGCGC
CTGCTGCCGGTGAAGCCGTTCTACCGCCTGTGCTGGGAGTCGGGCAAGACCCTGGACTAC
GCCAACGACTCGGCCGAGCTGGAGGCCAGATCACCCAGTTCAACCCGCGCGACGTGGAG
GGCTACCGCCGCTTCTTGGCCTACTCGCAGGCCGTGTTCCAGGAGGGCTACCTGCGCCTG
GGCTCGGTGCCGTTCTGTGTTCCGCGACATGCTGCGCGCCGGCCCGCAGCTCCTGAAG
CTGCAGGCCTGGCAGTCGGTGTACCACTCGGTGTGCGGCTTCATCGAGGACGAGCACCTG
CGCCAGGCCTTCTCGTTCCACTCGCTGCTGGTGGGCGGCAACCCGTTACCCACCTCGTCG
ATCTACACCCTGATCCACGCCCTGGAGCGCGAGTGGGGCGTGTGGTTCCCGGAGGGCGGC
ACCGGCGCCCTGGTGAACGGCATGGTGAAGCTGTTACCGACCTGGGCGGCGAGATCGAG
CTGAACGCCCCGCTGGAGGAGCTGGTGGTGGCCGACAACCGCGTGTGCGAGGTGCGCCTG
GCCGACGGCCGCATCTTCGACACCGACGCCGTGGCCTCGAACGCCGACGTGGTGAACACC
TACAAGAAGCTGCTGGGCCACCACCCGGTGGGCCAGAAGCGCGCCGCCGCCCTGGAGCGC
AAGTCGATGTGGAACCTCGCTGTTCTGTGCTGTACTTCGGCCTGAACCAGCCGCACTCGCAG
CTCGCCACCACACCATCTGCTTCGGCCCGCGCTACCGCGAGCTGATCGACGAGATCTTC
ACCGGCTCGGCCCTGGCCGACGACTTCTCGCTGTACCTGCACTCGCCGTGCGTGACCGAC
CCGTGCTGGCCCCGCCGGGCTGCGCCTCGTTCTACGTGCTGGCCCCGGTGCCGCACCTG
GGCAACGCCCCGCTGGACTGGGCCAGGAGGGCCGAAGCTGCGCGACCGCATCTTCGAC
TACCTGGAGGAGCGCTACATGCCGGGCTGCGCTCGCAGCTCGTGACCCAGCGCATCTTC
ACCCCGGCCGACTTCCACGACACCCTGGACGCCACCTGGGCTCGGCCTTCTCGATCGAG
CCGTGCTGACCCAGTCGGCCTGGTTCGCCCGCACAACCGCGACTCGGACATCGCCAAC
CTGTACCTGGTGGGCGCCGGCACCCACCCGGGCGCCGGCATCCCGGGCGTGGTGGCCTCG
GCCAAGGCCACCGCCTCGCTGATGATCGAGGACCTGCAGTGATCTGGGACGTGCTATGGG
TAAGCTGGACATCACCTCCACAACGCAAAGGAGATATAGATGTCGAGCCGCCGCTGCT
GGACCACGCCACCCAGACCATGGCCAACGGCTCGAAGTCGTTCCGCCACCGCCGCCAAGCT
GTTGACCCGGCCACCCGCCGCTCGGTGCTGATGCTGTACACCTGGTGCCGCCACTGCGA
CGACGTGATCGACGACCGACCCACGGCTTCGCCTCGGAGGCCGCCGCCGAGGAGGAGC
CACCCAGCGCCTGGCCCGCTGCGCACCTGACCCTGGCCGCCTTCGAGGGCGCCGAGAT

GCAGGACCCGGCCTTCGCCGCTTCCAGGAGGTGGCCCTGACCCACGGCATCACCCCGCG
CATGGCCCTGGACCACCTGGACGGCTTCGCCATGGACGTGGCCAGACCCGCTACGTGAC
CTTCGAGGACACCCTGCGCTACTGCTACCACGTGGCCGGCGTGGTGGGCCTGATGATGGC
CCGCGTGATGGGCGTGCGGACGAGCGCGTGCTGGACCGCGCCTGCGACCTGGGCCTGGC
CTTCAGCTACCAACATCGCCCGGACATCATCGACGACGCCGCCATCGACCGCTGCTA
CCTGCCGGCCGAGTGGCTGCAGGACGCCGGCCTGACCCGGAGAACTACGCCGCCCCGGA
GAACCGCGCCGCCCTGGCCGCGTGCCGAGCGCCTGATCGACGCCGCCGAGCCGTA
CATCTCGTCGAGGCCGGCCTGCACGACCTGCCGCCGCGCTGCGCCTGGGCCATCGCCAC
CGCCCGCTCGGTGTACCGCGAGATCGGCATCAAGGTGAAGGCCGCCGGCGGCTCGGCCTG
GGACCGCCGCCAGCACACCTCGAAGGGCGAGAAGATCGCCATGCTGATGGCCGCCCCGGG
CCAGGTGATCCGCGCAAGACCACCCGCGTGACCCCGCGCCCGGCCGCTGTGGCAGCG
CCCGGTGTGACTGTCCCCCAGTTCCAGTACCTGGTCATCATCTGCCTTTCAAAGGAGA
TATAGATGCTGTGGATCTGGAACGCCCTGATCGTGTTCTGACCGTGATCGGCATGGAGG
TGGTGGCCGCCCTGGCCACAAGTACATCATGCACGGCTGGGGCTGGGGCTGGCACCTGT
CGCACCACGAGCCGCGCAAGGGCGCCTTCGAGGTGAACGACCTGTACGCCGTGGTGTTCG
CCGCCCTGTGATCCTGCTGATCTACCTGGGCTCGACCGGCATGTGGCCGCTGCAGTGGA
TCGGCGCCGGCATGACCGCTACGGCCTGCTGTACTTCATGGTGCACGACGGCCTGGTGC
ACCAGCGCTGGCCGTTCCGCTACATCCCGCGCAAGGGCTACCTGAAGCGCCTGTACATGG
CCCACCGCATGCACCACGCCGTGCGCGGCAAGGAGGGCTGCGTGTCTCGGCTTCCTGT
ACGCCCCGCCGCTGTGAAGCTGCAGGCCACCCTGCGCGAGCGCCACGGCGCCCCGCGCCG
GCGCCGCCCGCGACGCCAGGGCGGCGAGGACGAGCCGCCCTCGGGCAAGTGAGTTATAT
GGAGGGGGCAAACGCTCTAGAACTAGTGGATCCAAAGGAGATATAGATGTCGGCCGTGAC
CCCGATGTCGAGAGTGGTGCCAAACCAGGCCCTAATCGGCCTGACTTTAGCGGGGCTGAT
AGCCACGGCGTGGCTGAGTCTGCATATTTACGGGGTGTACTTCATCGTTGGACAATGTG
GTCGATCCTGACGGTGCCGCTGATCGTGGCCTTCAGACGTGGCTGTCGGTAGGCCTGTT
CATCGTTGCCACGACGCAATGCACGGCTCCCTAGCCCCGGGAGGCCCCGCCTGAACAC
CGCCATCGGGTCCCTGGCCCTAGGCCTGTACGCTGGCTTCAGGTTCCGCCCTCTGAAGAC
CGCCACCATGCCACCATGCCGACCGGGCACAGCCGACGACCCGGATTTTCACGCGGA
CGCCCCCGTGCGTTCCTGCCGTGGTTCTACGGCTTTTCCGTACCTACTTCGGCTGGAG

GGAGCTGGCCGTGCTGACCGTGTTGGTGGCCGTGGCTGTTTAAATCCTGGGCGCCCGAAT
GCCGAACCTACTTGTGTTCTGGGCCGCCCGGCTCTATTATCGGCCTTGCAGCTTTTCAC
CTTCGGCACATGGCTGCCGCACCGACACACCGACGACGCCTTCCCGGACCACCACAACGC
TCGCACTTCACCCTTTGGCCCCATCCTGTCTCTGCTGACCTGCTTCCACTTCGGCCGGCA
CCATGAGCACCACCTGACTCCGTGGAAACCGTGGTGGAGGCTGTTCTCGTAGCGATACCG
TCGACTTCGAGCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCT
GTTGTTTGTGCGTGAACGCTCTCATTAAT

SEQ ID NO: 7 [System 2, insert only, 5868 bp]:

213: Unknown

220:

221: Nucleic acid sequence

222: Synthetic nucleotide sequence derived from the Pj5[E1A1C1C2] promoter (1-327), codon-optimized crtE from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (328-1,251), spacer sequence (1,252-1,291), RBS (1,292-1,305), codon-optimized crtY from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (1,306-2,466), spacer sequence (2,467-2,509), RBS (2,510-2,523), codon-optimized crtI from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (2,524-4,002), spacer sequence (4,003-4,046), RBS (4,047-4,060), codon-optimized crtB from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (4,061-4,990), spacer sequence (4,991-5,037), RBS (5,038-5,051), codon-optimized crtW from *Brevundimonas* strain OB307 (5,052-5,777), ending spacer sequence (5,778-5,796), and *E. coli* rrnB terminator (5,797-5,868).

223: Synthesized

AGTCCATTGTTGCCTTGCAACGCACGCGCTGTCAATGCGGGAATCCGCCTCGGCACTGCA
CGCTTCCCACCTACCGGACGGTATGCAGCGCTCGCATCTGCCGAGGCCCCAGAGCATAG
GCGAGAAGGATGAATTTTGTATGTACATCGTGGCCATTGCTGCAGAGCGGATATAAAAAC
CGTTATTGACACAGGTGGAAATTTAAATATACTGTTAGTAAACCTAATGGATCGACCTT
GAATTCAAAGATCTGGGAGACCACAACGGTTCCCTCTAGAAATAATTTTGAATTCAA
AAGATCTTTTAAGAAGGAGATATACATATGGTGTGCGGGCTCGAAGGCCGGCGTGTGCGCG

CACCGCGAGATCGAGGTGATGCGCCAGTCGATCGACGACCACCTGGCCGGCCTGCTGCCG
GAGACCGACTCGCAGGACATCGTGTCTGCTGGCCATGCGCGAGGGCGTGATGGCCCCGGG
AAGCGCATCCGCCCCTGCTGATGCTGCTGGCCGCCGCGACCTGCGCTACCAGGGCTCG
ATGCCGACCCTGCTGGACCTGGCCTGCGCCGTGGAGCTGACCCACACCGCCTCGTGATG
CTGGACGACATGCCGTGCATGGACAACGCCGAGCTGCGCCGCGGCCAGCCGACCAACCCAC
AAGAAGTTCGGCGAGTCGGTGGCCATCCTGGCCTCGGTGGGCCTGCTGTGAAGGCCTTC
GGCCTGATCGCCGCCACCGGCGACCTGCCGGGCGAGCGCCGCGCCAGGCCGTGAACGAG
CTGTGACCGCCGTGGGCGTGACGGGCTGGTGTGGGCCAGTTCGCGACCTGAACGAC
GCCGCCCTGGACCGCACCCCGGACGCCATCCTGTGACCAACCACCTGAAGACCGGCATC
CTGTTCTCGGCCATGCTGCAGATCGTGGCCATCGCCTCGGCCTCGTCGCCGTGACCCGC
GAGACCCTGCACGCCTTCGCCCTGGACTTCGGCCAGGCCTTCAGCTCCTGGACGACCTG
CGCGACGACCAACCGGAGACCGGAAGGACCGCAACAAGGACGCCGGCAAGTCGACCCTG
GTGAACCGCCTGGGCGCCGACGCCGCCGCCAGAAGCTGCGCGAGCACATCGACTCGGCC
GACAAGCACCTGACCTTCGCCTGCCCGAGGGCGGCCATCCGCCAGTTCATGCACCTG
TGGTTCGGCCACCACCTGGCCGACTGGTCGCCGGTGATGAAGATCGCCTGAGTCATAGCT
GTTTCCTGCCAGTCACGACGTTGTAAACGCAAAGGAGATATAGGTGCGCGACCTGATC
CTGGTGGGCGGCGGCCTGGCCAACGGCCTGATCGCCTGGCGCCTGCGCCAGCGCTACCCG
CAGCTCAACCTGCTGCTGATCGAGGCCGGCGAGCAGCCGGGCGGCAACCACACCTGGTCG
TTCCACGAGGACGACCTGACCCCGGGCCAGCACGCCTGGCTGGCCCCGCTGGTGGCCAC
GCCTGGCCGGGCTACGAGGTGCAGTTCGCCGACCTGCGCCGCCGCTGGCCCGCGGCTAC
TACTCGATCACCTCGGAGCGCTTCGCCGAGGCCCTGCACCAGGCCCTGGGCGAGAACATC
TGGCTGAAGTCTCGGTGTGGAGGTGCTGCCGAAGTTCGGTGCCTGGCCAACGGCGAG
GCCCTGCTGGCCGGCGCCGTGATCGACGGCCGCGGCGTGACCGCCTCGTCGGCCATGCAG
ACCGGCTACCAGCTCTCCTGGGCCAGCAGTGGCGCCTGACCCAGCCGCACGGCCTGACC
GTGCCGATCCTGATGGACGCCACCGTGGCCAGCAGCAGGGCTACCGCTTCGTGTACACC
CTGCCGCTGTGCGCCGACACCTGCTGATCGAGGACACCCGCTACGCCAACGTGCCGAG
CGCGACGACAACGCCCTGCGCCAGACCGTGACCGACTACGCCCACTCGAAGGGCTGGCAG
CTCGCCAGCTCGAACGCGAGGAGACCGGCTGCCTGCCGATCACCTGGCCGGCGACATC
CAGGCCCTGTGGGCCGACGCCCCGGGCGTGCCGCGCTCGGGCATGCGCGCCGGCCTGTTC

CACCCGACCACCGGCTACTCGCTGCCGCTGGCCGTGGCCCTGGCCGACGCCATCGCCGAC
TCGCCGCGCCTGGGCTCGGTGCCGCTGTACCAGCTCACCCGCCAGTTCGCCGAGCGCCAC
TGGCGCCGCCAGGGCTTCTCCGCCTGCTGAACCGCATGCTGTTCTTGCCGGCCGCGAG
GAGAACCGCTGGCGCGTGATGCAGCGCTTCTACGGCCTGCCGGAGCCGACCGTGGAGCGC
TTCTACGCCGGCCGCTGTCGCTGTTGACAAGGCCCGCATCCTGACCGGCAAGCCGCCG
GTGCCGCTGGGCGAGGCCTGCCGCGCCGCCCTGAACCACTTCCCGGACCGCCGCGACAAG
GGCTGACCTGTGTGAAATTGTTATCCGCTTACCCATACGACGTCCCAGACAAAGGAGATA
TAGATGAAGAAGACCGTGGTGATCGGCGCCGGCTTCGGCGGCCTGGCCCTGGCCATCCGC
CTGCAGGCCGCCGGCATCCCGACCGTGTCTGCTGGAGCAGCGCGACAAGCCGGGCGGCCGC
GCCTACGTGTGGCAGCAGCAGGGCTTCACCTTCGACGCCGGCCCGACCGTGTACCGAC
CCGACCGCCCTGGAGGCCCTGTTACCCCTGGCCGGCCGCCGCATGGAGGACTACGTGCGC
CTGCTGCCGTGAAGCCGTTCTACCGCCTGTGCTGGGAGTCGGGCAAGACCCTGGACTAC
GCCAACGACTCGGCCGAGCTGGAGGCCAGATCACCCAGTTCAACCCGCGCGACGTGGAG
GGCTACCGCCGCTTCTTGCCCTACTCGCAGGCCGTGTTCCAGGAGGGCTACCTGCGCCTG
GGCTCGGTGCCGTTCTGTCGTTCCGCGACATGCTGCGCGCCGGCCCGCAGCTCCTGAAG
CTGCAGGCCTGGCAGTCGGTGTACCACTCGGTGTGCGGCTTCATCGAGGACGAGCACCTG
CGCCAGGCCTTCTCGTTCCACTCGCTGCTGGTGGGCGGCAACCCGTTACCCACCTCGTCG
ATCTACACCCTGATCCACGCCCTGGAGCGCGAGTGGGGCGTGTGGTTCCCGGAGGGCGGC
ACCGGCGCCCTGGTGAACGGCATGGTGAAGCTGTTACCGACCTGGGCGGCGAGATCGAG
CTGAACGCCCCGCTGGAGGAGCTGGTGGTGGCCGACAACCGCGTGTGCGAGGTGCGCCTG
GCCGACGGCCGCATCTTCGACACCGACGCCGTGGCCTCGAACGCCGACGTGGTGAACACC
TACAAGAAGCTGCTGGGCCACCACCCGGTGGGCCAGAAGCGCGCCGCCGCCCTGGAGCGC
AAGTCGATGTGAACTCGCTGTTCTGTGCTGTACTTCGGCCTGAACCAGCCGCACTCGCAG
CTCGCCACCACACCATCTGCTTCGGCCCGCGCTACCGCGAGCTGATCGACGAGATCTTC
ACCGGCTCGGCCCTGGCCGACGACTTCTCGCTGTACCTGCACTGCCCGTGCCTGACCGAC
CCGTCGCTGGCCCCGCCGGGCTGCGCCTCGTTCTACGTGCTGGCCCCGGTGCCGCACCTG
GGCAACGCCCCGCTGGACTGGGCCAGGAGGGCCCGAAGCTGCGCGACCGCATCTTCGAC
TACCTGGAGGAGCGCTACATGCCGGGCCTGCGCTCGCAGCTCGTGACCCAGCGCATCTTC
ACCCCGGCCGACTTCACGACACCCTGGACGCCACCTGGGCTCGGCCTTCTCGATCGAG

CCGCTGCTGACCCAGTCGGCCTGGTTCCGCCCGCACAAACCGCGACTCGGACATCGCCAAC
CTGTACCTGGTGGGCGCCGGCACCCACCCGGGCGCCGGCATCCCGGGCGTGGTGGCCTCG
GCCAAGGCCACCGCCTCGCTGATGATCGAGGACCTGCAGTGATCTGGGACGTCTGTATGGG
TAAGCTGGACATCACCTCCCACAACGCAAAGGAGATATAGATGTCGCAGCCGCCGCTGCT
GGACCACGCCACCCAGACCATGGCCAACGGCTCGAAGTCGTTGCCACCGCCGCAAGCT
GTTTCGACCCGGCCACCCGCCGCTCGGTGCTGATGCTGTACACCTGGTGCCGCCACTGCGA
CGACGTGATCGACGACCAGACCCACGGCTTCGCCTCGGAGGCCGCCGCCGAGGAGGAGGG
CACCCAGCGCCTGGCCCGCCTGCGCACCCCTGACCCTGGCCGCCTTCGAGGGCGCCGAGAT
GCAGGACCCGGCCTTCGCCGCCTTCAGGAGGTGGCCCTGACCCACGGCATCACCCCGCG
CATGGCCCTGGACCACCTGGACGGCTTCGCCATGGACGTGGCCAGACCCGCTACGTGAC
CTTCGAGGACACCCTGCGCTACTGCTACCACGTGGCCGGCGTGGTGGGCCTGATGATGGC
CCGCGTGATGGGCGTGCGGACGAGCGCGTGCTGGACCGCGCCTGCGACCTGGGCCTGGC
CTTCAGCTACCAACATCGCCCGCGACATCATCGACGACGCCGCCATCGACCGCTGCTA
CCTGCCGGCCGAGTGGCTGCAGGACGCCGGCCTGACCCCGGAGAACTACGCCGCCCGCGA
GAACCGCGCCGCCCTGGCCCGCGTGGCCGAGCGCCTGATCGACGCCGCCGAGCCGTACTA
CATCTCGTCGAGGCCGGCCTGCACGACCTGCCGCCGCGCTGCGCCTGGGCCATCGCCAC
CGCCCGCTCGGTGTACCGCGAGATCGGCATCAAGGTGAAGGCCGCCGGCGGCTCGGCCTG
GGACCGCCGCCAGCACACCTCGAAGGGCGAGAAGATCGCCATGCTGATGGCCGCCCGGG
CCAGGTGATCCGCGCAAGACCACCCGCGTGACCCCGCGCCCGGCCGCTGTGGCAGCG
CCCGGTGTGACTGTCCCCGTTATATGGAGGGGGCAAACGCTCTAGAACTAGTGGATCCAA
AGGAGATATAGATGTCGGCCGTGACCCCGATGTCGAGAGTGGTGCCAAACCAGGCCCTAA
TCGGCCTGACTTTAGCGGGGCTGATAGCCACGGCGTGGCTGAGTCTGCATATTTACGGGG
TGTACTTCATCGTTGGACAATGTGGTCGATCCTGACGGTGCCGCTGATCGTGGCCTTCC
AGACGTGGCTGTGGTAGGCCTGTTTCATCGTTGCCACGACGCAATGCACGGCTCCCTAG
CCCCGGGAGGCCCGCCTGAACACCGCCATCGGGTCCCTGGCCCTAGGCCTGTACGCTG
GCTTCAGGTTCCGCCCTCTGAAGACCGCCACCATGCCACCATGCCGACCGGGCACAG
CCGACGACCCGGATTTTACGCGGACGCCCCCGTGCGTTCCTGCCGTGGTCTACGGCT
TTTTCCGTACCTACTTCGGCTGGAGGGAGCTGGCCGTGCTGACCGTGTGGTGGCCGTGG
CTGTTTTAATCCTGGGCGCCCGAATGCCGAACCTACTTGTGTTCTGGGCGCCCGGCTC

TATTATCGGCCTTGCAGCTTTTCACCTTCGGCACATGGCTGCCGCACCGACACACCGACG
 ACGCCTTCCCGGACCACCACAACGCTCGCACTTCACCCTTTGGCCCCATCCTGTCTCTGC
 TGACCTGCTTCCACTTCGGCCGGCACCATGAGCACCACCTGACTCCGTGGAACCGTGGT
 GGAGGCTGTTCTCGTAGCGATACCGTCGACTTCGAGCAAATAAAACGAAAGGCTCAGTCG
 AAAGACTGGGCCTTTCGTTTTATCTGTTGTTGTCGGTGAACGCTCT

SEQ ID NO: 8 [System 3, insert only, 6462 bp]:

213: Unknown

220:

221: Nucleic acid sequence

222: Synthetic nucleotide sequence derived from the Pj5[E1A1C1C2] promoter (1-327), codon-optimized crtE from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (328-1,251), spacer sequence (1,252-1,291), RBS (1,292-1,305), codon-optimized crtY from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (1,306-2,466), spacer sequence (2,467-2,509), RBS (2,510-2,523), codon-optimized crtI from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (2,524-4,002), spacer sequence (4,003-4,046), RBS (4,047-4,060), codon-optimized crtB from *Pantoea agglomerans* M87280/M99707 pAC-BETA plasmid (4,061-4,990), spacer sequence (4,991-5,080), RBS (5,081-5,093), a codon-optimized crtZW fusion containing the crtZ gene from *Pantoea ananatis* Strain AJ13355 NC_017533, a 30-bp sequence encoding a linker peptide, and the crtW gene from *Brevundimonas* strain OB307 without the N-terminal methionine (5,094-6,371), ending spacer sequence (6,372-6,390), and *E. coli* rrnB terminator (6,391-6,462).

223: Synthesized

AGTCCATTGTTGCCTTGCAACGCACGCGCTGTCAATGCGGGAATCCGCCTCGGCACTGCA
 CGCTTCCCACCTACCGGACGGTATGCAGCGCTCGCATCTGCCGAGGCCCCAGAGCATAG
 GCGAGAAGGATGAATTTTGTATGTACATCGTGGCCATTGCTGCAGAGCGGATATAAAAAC
 CGTTATTGACACAGGTGGAAATTTAAATATACTGTTAGTAAACCTAATGGATCGACCTT
 GAATTCAAAGATCTGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGAATTCAA
 AAGATCTTTTAAGAAGGAGATATACATATGGTGTGCGGGCTCGAAGGCCGGCGTGTGCGCG

CACCGCGAGATCGAGGTGATGCGCCAGTCGATCGACGACCACCTGGCCGGCCTGCTGCCG
GAGACCGACTCGCAGGACATCGTGTCTGCTGGCCATGCGCGAGGGCGTGATGGCCCCGGG
AAGCGCATCCGCCCCTGCTGATGCTGCTGGCCGCCGCGACCTGCGCTACCAGGGCTCG
ATGCCGACCCTGCTGGACCTGGCCTGCGCCGTGGAGCTGACCCACACCGCCTCGTGATG
CTGGACGACATGCCGTGCATGGACAACGCCGAGCTGCGCCGCGGCCAGCCGACCAACCCAC
AAGAAGTTCGGCGAGTCGGTGGCCATCCTGGCCTCGGTGGGCCTGCTGTGAAGGCCTTC
GGCCTGATCGCCGCCACCGGCGACCTGCCGGGCGAGCGCCGCGCCAGGCCGTGAACGAG
CTGTGACCGCCGTGGGCGTGACGGGCTGGTGTGGGCCAGTTCGCGACCTGAACGAC
GCCGCCCTGGACCGCACCCCGGACGCCATCCTGTGACCAACCACCTGAAGACCGGCATC
CTGTTCTCGGCCATGCTGCAGATCGTGGCCATCGCCTCGGCCTCGTCGCCGTGACCCGC
GAGACCCTGCACGCCTTCGCCCTGGACTTCGGCCAGGCCTTCAGCTCCTGGACGACCTG
CGCGACGACCAACCGGAGACCGGAAGGACCGCAACAAGGACGCCGGCAAGTCGACCCTG
GTGAACCGCCTGGGCGCCGACGCCGCCGCCAGAAGCTGCGCGAGCACATCGACTCGGCC
GACAAGCACCTGACCTTCGCCTGCCCGAGGGCGGCCATCCGCCAGTTCATGCACCTG
TGGTTCGGCCACCACCTGGCCGACTGGTCGCCGGTGATGAAGATCGCCTGAGTCATAGCT
GTTTCCTGCCAGTCACGACGTTGTAAACGCAAAGGAGATATAGGTGCGCGACCTGATC
CTGGTGGGCGGCGGCCTGGCCAACGGCCTGATCGCCTGGCGCCTGCGCCAGCGCTACCCG
CAGCTCAACCTGCTGCTGATCGAGGCCGGCGAGCAGCCGGGCGGCAACCACACCTGGTCG
TTCCACGAGGACGACCTGACCCCGGGCCAGCACGCCTGGCTGGCCCCGCTGGTGGCCAC
GCCTGGCCGGGCTACGAGGTGCAGTTCCTGGACCTGCGCCGCCGCTGGCCCGCGGCTAC
TACTCGATCACCTCGGAGCGCTTCGCCGAGGCCCTGCACCAGGCCCTGGGCGAGAACATC
TGGCTGAAGTCTCGGTGTGCGAGGTGCTGCCGAAGTTCGGTGCCTGGCCAACGGCGAG
GCCCTGCTGGCCGGCGCCGTGATCGACGGCCGCGGCGTGACCGCCTCGTCGGCCATGCAG
ACCGGCTACCAGCTCTCCTGGGCCAGCAGTGGCGCCTGACCCAGCCGCACGGCCTGACC
GTGCCGATCCTGATGGACGCCACCGTGGCCAGCAGCAGGGCTACCGCTTCGTGTACACC
CTGCCGCTGTGCGCCGACACCTGCTGATCGAGGACACCCGCTACGCCAACGTGCCGAG
CGCGACGACAACGCCCTGCGCCAGACCGTGACCGACTACGCCCACTCGAAGGGCTGGCAG
CTGCCCAGCTCGAACGCGAGGAGACCGGCTGCCTGCCGATCACCTGGCCGGCGACATC
CAGGCCCTGTGGGCCGACGCCCCGGGCGTGCCGCGCTCGGGCATGCGCGCCGGCCTGTTC

CACCCGACCACCGGCTACTCGCTGCCGCTGGCCGTGGCCCTGGCCGACGCCATCGCCGAC
TCGCCGCGCCTGGGCTCGGTGCCGCTGTACCAGCTCACCCGCCAGTTCGCCGAGCGCCAC
TGGCGCCGCCAGGGCTTCTCCGCCTGCTGAACCGCATGCTGTTCTTGCCGGCCGCGAG
GAGAACCGCTGGCGCGTGATGCAGCGCTTCTACGGCCTGCCGGAGCCGACCGTGAGCGC
TTCTACGCCGGCCGCTGTCGCTGTTGACAAGGCCCGCATCCTGACCGGCAAGCCGCCG
GTGCCGCTGGGCGAGGCCTGCCGCGCCGCCCTGAACCACTTCCCGGACCGCCGCGACAAG
GGCTGACCTGTGTGAAATTGTTATCCGCTTACCCATACGACGTCCCAGACAAAGGAGATA
TAGATGAAGAAGACCGTGGTGATCGGCGCCGGCTTCGGCGGCCTGGCCCTGGCCATCCGC
CTGCAGGCCGCCGGCATCCCAGCGTGCTGCTGGAGCAGCGCGACAAGCCGGGCGGCCGC
GCCTACGTGTGGCAGCAGCAGGGCTTACCTTCGACGCCGGCCCGACCGTGATCACCGAC
CCGACCGCCCTGGAGGCCCTGTTACCCCTGGCCGGCCGCCGCATGGAGGACTACGTGCGC
CTGCTGCCGTGAAGCCGTTCTACCGCCTGTGCTGGGAGTCGGGCAAGACCCTGGACTAC
GCCAACGACTCGGCCGAGCTGGAGGCCAGATCACCCAGTTCAACCCGCGCGACGTGGAG
GGCTACCGCCGCTTCTTGCCCTACTCGCAGGCCGTGTTCCAGGAGGGCTACCTGCGCCTG
GGCTCGGTGCCGTTCTGTCGTTCCGCGACATGCTGCGCGCCGGCCCGCAGCTCCTGAAG
CTGCAGGCCTGGCAGTCGGTGTACCACTCGGTGTGCGGCTTCATCGAGGACGAGCACCTG
CGCCAGGCCTTCTCGTTCCACTCGCTGCTGGTGGGCGGCAACCCGTTACCCACCTCGTCG
ATCTACACCCTGATCCACGCCCTGGAGCGCGAGTGGGGCGTGTGGTTCCCGGAGGGCGGC
ACCGGCGCCCTGGTGAACGGCATGGTGAAGCTGTTACCGACCTGGGCGGCGAGATCGAG
CTGAACGCCCCGCTGGAGGAGCTGGTGGTGGCCGACAACCGCGTGTGCGAGGTGCGCCTG
GCCGACGGCCGCATCTTCGACACCGACGCCGTGGCCTCGAACGCCGACGTGGTGAACACC
TACAAGAAGCTGCTGGGCCACCACCCGGTGGGCCAGAAGCGCGCCGCCGCCCTGGAGCGC
AAGTCGATGTGAACTCGCTGTTCTGTGCTGTACTTCGGCCTGAACCAGCCGCACTCGCAG
CTCGCCACCACACCATCTGCTTCGGCCCGCGCTACCGCGAGCTGATCGACGAGATCTTC
ACCGGCTCGGCCCTGGCCGACGACTTCTCGCTGTACCTGCACTGCCCGTGCGTGACCGAC
CCGTCGCTGGCCCCGCCGGGCTGCGCCTCGTTCTACGTGCTGGCCCCGGTGCCGACCTG
GGCAACGCCCCGCTGGACTGGGCCAGGAGGGCCCCGAAGCTGCGCGACCGCATCTTCGAC
TACCTGGAGGAGCGCTACATGCCGGGCCTGCGCTCGCAGCTCGTGACCCAGCGCATCTTC
ACCCCGGCCGACTTCACGACACCCTGGACGCCACCTGGGCTCGGCCTTCTCGATCGAG

CCGCTGCTGACCCAGTCGGCCTGGTTCCGCCCGCACAAACCGCGACTCGGACATCGCCAAC
CTGTACCTGGTGGGCGCCGGCACCCACCCGGGCGCCGGCATCCCGGGCGTGGTGGCCTCG
GCCAAGGCCACCGCCTCGCTGATGATCGAGGACCTGCAGTGATCTGGGACGTCTGTATGGG
TAAGCTGGACATCACCTCCCACAACGCAAAGGAGATATAGATGTCGCAGCCGCCGCTGCT
GGACCACGCCACCCAGACCATGGCCAACGGCTCGAAGTCGTTGCCACCGCCGCCAAGCT
GTTGACCCGGCCACCCGCCGCTCGGTGCTGATGCTGTACACCTGGTGCCGCCACTGCGA
CGACGTGATCGACGACCAGACCCACGGCTTCGCCTCGGAGGCCGCCGCCGAGGAGGAGG
CACCCAGCGCCTGGCCCGCTGCGCACCCCTGACCCTGGCCGCCTTCGAGGGCGCCGAGAT
GCAGGACCCGGCCTTCGCCGCCTTCAGGAGGTGGCCCTGACCCACGGCATCACCCCGCG
CATGGCCCTGGACCACCTGGACGGCTTCGCCATGGACGTGGCCAGACCCGCTACGTGAC
CTTCGAGGACACCCTGCGCTACTGCTACCACGTGGCCGGCGTGGTGGGCCTGATGATGGC
CCGCGTGATGGGCGTGCGCGACGAGCGCTGCTGGACCGCGCCTGCGACCTGGGCCTGGC
CTTCAGCTACCAACATCGCCCGCGACATCATCGACGACGCCGCCATCGACCGCTGCTA
CCTGCCGGCCGAGTGGCTGCAGGACGCCGGCCTGACCCCGGAGAACTACGCCGCCCGCGA
GAACCGCGCCGCCCTGGCCCGCTGGCCGAGCGCCTGATCGACGCCGCCGAGCCGTACTA
CATCTCGTCGAGGCCGGCCTGCACGACCTGCCCGCGCTGCGCCTGGGCCATCGCCAC
CGCCCGCTCGGTGTACCGCGAGATCGGCATCAAGGTGAAGGCCGCCGGCGGCTCGGCCTG
GGACCGCCGCCAGCACACCTCGAAGGGCGAGAAGATCGCCATGCTGATGGCCGCCCGGG
CCAGGTGATCCGCGCAAGACCACCCGCGTGACCCCGCGCCCGGCCGCTGTGGCAGCG
CCCGGTGTGACTGTCCCCGTTATATGGAGGGGGCAAACGCTCTAGAACTAGTGGATCCCT
GTCCCCCAGTTCCAGTACCTGGTCATCATCTGCCTTTCAAAGGAGATATAGATGCTGT
GGATCTGGAACGCCCTGATCGTTTTCTGACCGTGATCGGCATGGAAGTGGTGGCCGCC
TGGCCATAAGTACATCATGCACGGCTGGGGCTGGGGCTGGCACCTGTCGCACCACGAAC
CACGCAAAGGCGCATTTGAGGTGAATGACCTGTATGCCGTGGTGTTCGCCGCCCTGTCGA
TTCTGCTGATCTATCTGGGCTCGACTGGCATGTGGCCGCTGCAGTGGATTGGCGCCGGCA
TGACCGCATACGGCCTGCTGTACTTTATGGTTCATGACGGCCTGGTGCACCAGCGCTGGC
CGTTCCGCTACATCCCGCGCAAAGGCTATCTGAAACGCCTGTACATGGCCACCGCATGC
ACCATGCAGTGCGCGGCAAGGAGGGCTGTGTGTCATTGGCTTTCTGTACGCCCCGCCG
TGTCGAAGCTGCAGGCCACTCTGCGCGAGAGACATGGCGCCCGCGCCGGCGCAGCCCCGCG

ATGCCCAAGGCGGCGAGGACGAGCCGGCATCGGGCAAAGGCGGGGGCGGGTCCGGCGGCC
 CGGGGTCTGCGCCGTGACCCCGATGTCGAGAGTGGTGCCAAACCAGGCCCTAATCGGCC
 TGACTTTAGCGGGGCTGATAGCCACGGCGTGGCTGAGTCTGCATATTTACGGGGTGTACT
 TCCATCGTTGGACAATGTGGTCGATCCTGACGGTGCCGCTGATCGTGGCCTTCCAGACGT
 GGCTGTCGGTAGGCCTGTTTCATCGTTGCCACGACGCAATGCACGGCTCCCTAGCCCCGG
 GGAGGCCCCGCCTGAACACCGCCATCGGGTCCCTGGCCCTAGGCCTGTACGCTGGCTTCA
 GGTTGCCCCCTCTGAAGACCGCCACCATGCCACCATGCCGCACCGGGCACAGCCGACG
 ACCCGGATTTTACGCGGACGCCCCCGTGCGTTCTGCGGTGGTTCTACGGCTTTTCC
 GTACCTACTTCGGCTGGAGGGAGCTGGCCGTGCTGACCGTGTGGTGCCGTGGCTGTTT
 TAATCCTGGGCGCCGAATGCCGAACCTTACTTGTGTTCTGGGCCGCCCCGGCTCTATTAT
 CGGCCTTGCAGCTTTTACCTTCGGCACATGGCTGCCGCACCGACACACCGACGACGCCT
 TCCCGGACCACACAACGCTCGCACTTACCCTTTGGCCCCATCCTGTCTCTGCTGACCT
 GCTTCCACTTCGGCCGGCACCATGAGCACCACTGACTCCGTGGAAACCGTGGTGGAGGC
 TGTTCCTGTCGATACCGTCGACTTCGAGCAAATAAAACGAAAGGCTCAGTCGAAAGAC
 TGGGCCTTTCGTTTTATCTGTTGTTGTCGGTGAACGCTCTC

SEQ ID NO: 9 [pDONRPEX18TC-Tn5 Insert with OB307-crtW (from attL1 to attL2), 8,861 bp]:

213: Unknown

220:

221: Nucleic acid sequence

222: Synthetic nucleotide sequence derived from the attL1 sequence (1-100), a spacer
 sequence (101-112), a Tn5 Mosaic End sequence (113-131), spacer sequence (132-236), the
 Pj5[E1A1C1C2] promoter (237-563), codon-optimized crtE from Pantoea agglomerans
 M87280/M99707 pAC-BETA plasmid (564-1,487), spacer sequence (1,488-1,525), RBS (1,526-
 1,541), codon-optimized crtY from Pantoea agglomerans M87280/M99707 pAC-BETA plasmid
 (1,542-2,702), spacer sequence (2,703-2,745), RBS (2,746-2,759), codon-optimized crtI from
 Pantoea agglomerans M87280/M99707 pAC-BETA plasmid (2,760-4,238), spacer sequence
 (4,239-4,282), RBS (4,283-4,296), codon-optimized crtB from Pantoea agglomerans
 M87280/M99707 pAC-BETA plasmid (4,297-5,226), spacer sequence (5,227-5,267), RBS (5,268-

5,281), codon-optimized crtZ from *Pantoea ananatis* Strain AJ13355 NC_017533 in plasmid pEA-320 (5,282-5,809), spacer sequence (5,810-5,848), RBS (5,849-5,862), codon-optimized crtW from *Brevundimonas* strain OB307 (5,863-6,588), spacer sequence (6,589-6,607), *E. coli* rrnB terminator (6,608-6,679), and *Asel* restriction site (6,680-6,685), spacer sequence (6,686-7,093), a Tn5 Mosaic End sequence (7,094-7,112), a *SpeI* restriction site sequence (7,113-7,118), T0 terminator (7,119-7,221), spacer and promoter sequence (7,222-7,321), Tn5 transposase sequence (7,322-8,752), spacer sequence (8,753-8,761), and an attL2 sequence (8,762-8,861).

223: Synthesized

CAAATAATGATTTTATTTTGAAGTATAGTGACCTGTTGTTGCAACAMATTGATGAGCAA
TGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTTCAGGCCGAGGCCTGTCTCT
TATACACATCTTTGTGTCTCAGGCCGCCTAGGCCGCGCCGCGCAATTCGAGCTCGGTA
CCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTACCGGTTTATTATTAAGTC
CATTGTTGCCTTGCAACGCACGCGCTGTCAATGCGGGAATCCGCCTCGGCACTGCACGCT
TCCCGACCTACCGGACGGTATGCAGCGCTCGCATCTGCCGAGGCCCCAGAGCATAGGCGA
GAAGGATGAATTTTGTATGTACATCGTGGCCATTGCTGCAGAGCGGATATAAAAACCGTT
ATTGACACAGGTGGAAATTTAAATATACTGTTAGTAAACCTAATGGATCGACCTTGAAT
TCAAAAGATCTGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGAATTCAAAAGA
TCTTTTAAGAAGGAGATATACATATGGTGTGCGGCTCGAAGGCCGGCGTGTGCGCCGACC
GCGAGATCGAGGTGATGCGCCAGTCGATCGACGACCACCTGGCCGGCCTGCTGCCGGAGA
CCGACTCGCAGGACATCGTGTGCTGGCCATGCGCGAGGGCGTGTGCCCCGGGCAAGC
GCATCCGCCCCTGCTGATGCTGCTGGCCGCCCGCGACCTGCGCTACCAGGGCTCGATGC
CGACCTGCTGGACCTGGCCTGCGCCGTGGAGCTGACCCACACCGCCTCGCTGATGCTGG
ACGACATGCCGTGCATGGACAACGCCGAGCTGCGCCGCGGCCAGCCGACCACCCACAAGA
AGTTCGGCGAGTCGGTGGCCATCCTGGCCTCGGTGGGCCTGCTGTGAAGGCCTTCGGCC
TGATCGCCGCCACCGGCGACCTGCCGGGCGAGCGCCGCGCCAGGCCGTGAACGAGCTGT
CGACCGCCGTGGGCGTGCAGGGCCTGGTGTGGGCCAGTTCCGCGACCTGAACGACGCCG
CCCTGGACCGCACCCCGGACGCCATCCTGTGACCAACCACCTGAAGACCGGCATCCTGT
TCTCGGCCATGCTGCAGATCGTGGCCATCGCCTCGGCCTCGTCGCCGTGACCCGCGAGA

CCCTGCACGCCTTCGCCCTGGACTTCGGCCAGGCCTTCCAGCTCCTGGACGACCTGCGCG
ACGACCACCCGGAGACCGGCAAGGACCGCAACAAGGACGCCGGCAAGTCGACCCTGGTGA
ACCGCCTGGGCGCCGACGCCGCCGCCAGAAGCTGCGCGAGCACATCGACTCGGCCGACA
AGCACCTGACCTTCGCCTGCCCCGAGGGCGGCGCCATCCGCCAGTTCATGCACCTGTGGT
TCGGCCACCACCTGGCCGACTGGTCGCCGGTGATGAAGATCGCCTGAGTCATAGCTGTTT
CCTGCCCAGTCACGACGTTGTAAAACGCAAAGGAGATATAGGTGCGCGACCTGATCCTGG
TGGGCGGCGGCCTGGCCAACGGCCTGATCGCCTGGCGCCTGCGCCAGCGCTACCCGAGC
TCAACCTGCTGCTGATCGAGGCCGGCGAGCAGCCGGGCGGCAACCACACCTGGTCGTTCC
ACGAGGACGACCTGACCCCGGGCCAGCACGCCTGGCTGGCCCCGCTGGTGGCCACGCCT
GGCCGGGCTACGAGGTGCAGTTCGCCGACCTGCGCCGCCGCTGGCCCGCGGCTACTACT
CGATCACCTCGGAGCGCTTCGCCGAGGCCCTGCACCAGGCCCTGGGCGAGAACATCTGGC
TGAAGTCTCGGTGTCGGAGGTGCTGCCGAAGTGGTGCCTGGCCAACGGCGAGGCCC
TGCTGGCCGGCGCCGTGATCGACGGCCGCGCGTGACCGCCTCGTCGGCCATGCAGACCG
GCTACCAGCTCTTCTGGGCCAGCAGTGGCGCCTGACCCAGCCGCACGGCCTGACCGTGC
CGATCCTGATGGACGCCACCGTGGCCAGCAGCAGGGCTACCGCTTCGTGTACACCCTGC
CGCTGTGCGCCGACACCCTGCTGATCGAGGACACCCGCTACGCCAACGTGCCGAGCGCG
ACGACAACGCCCTGCGCCAGACCGTGACCGACTACGCCCACTCGAAGGGCTGGCAGCTCG
CCCAGCTCGAACGCGAGGAGACCGGCTGCCTGCCGATCACCTGGCCGGCGACATCCAGG
CCCTGTGGGCCGACGCCCCGGGCGTGCCGCGCTCGGGCATGCGCGCCGGCCTGTTCCACC
CGACCACCGGCTACTCGCTGCCGCTGGCCGTGGCCCTGGCCGACGCCATCGCCGACTCGC
CGCGCCTGGGCTCGGTGCCGCTGTACCAGCTACCCGCCAGTTCGCCGAGCGCCACTGGC
GCCGCCAGGGCTTCTCCGCCTGCTGAACCGCATGCTGTTCTGGCCGGCCGCGAGGAGA
ACCGCTGGCGCGTGATGCAGCGCTTCTACGGCCTGCCGGAGCCGACCGTGGAGCGCTTCT
ACGCCGGCCGCTGTGCTGTTTCGACAAGGCCCGCATCCTGACCGGCAAGCCGCCGTGC
CGCTGGGCGAGGCCTGCCGCGCCGCCCTGAACCACTTCCCGGACCGCCGCGACAAGGGCT
GACCTGTGTGAAATTGTTATCCGCTTACCCATACGACGTCCCAGACAAAGGAGATATAGA
TGAAGAAGACCGTGGTGATCGGCGCCGGCTTCGGCGGCCTGGCCCTGGCCATCCGCTGC
AGGCCGCCGCGATCCCGACCGTGCTGCTGGAGCAGCGCGACAAGCCGGGCGGCGCGCCT
ACGTGTGGCACGACCAGGGCTTCACCTTCGACGCCGGCCGACCGTGATCACCGACCCGA

CCGCCCTGGAGGCCCTGTTACCCTGGCCGGCCGCCGCATGGAGGACTACGTGCGCCTGC
TGCCGGTGAAGCCGTTCTACCGCCTGTGCTGGGAGTCGGGCAAGACCCTGGACTACGCCA
ACGACTCGGCCGAGCTGGAGGCCAGATCACCCAGTTCAACCCGCGCGACGTGGAGGGCT
ACCGCCGCTTCTGTCCTACTCGCAGGCCGTGTTCCAGGAGGGCTACCTGCGCCTGGGCT
CGGTGCCGTTCTGTGTTCCGCGACATGCTGCGCGCCGGCCCGCAGCTCCTGAAGCTGC
AGGCCTGGCAGTCGGTGTACCAAGTCGGTGTGCGCTTCATCGAGGACGAGCACCTGCGCC
AGGCCTTCTGTTCCACTCGTGCTGGTGGGCGCAACCCGTTACCACCTCGTCGATCT
ACACCCTGATCCACGCCCTGGAGCGCGAGTGGGGCGTGTGGTTCGCGAGGGCGGCACCG
GCGCCCTGGTGAACGGCATGGTGAAGCTGTTACCGACCTGGGCGGCGAGATCGAGCTGA
ACGCCCCGCGTGGAGGAGCTGGTGGTGGCCGACAACCGCGTGTGCGAGGTGCGCCTGGCCG
ACGGCCGCATCTTCGACACCGACGCCGTGGCCTCGAACGCCGACGTGGTGAACACCTACA
AGAAAGTGTGGGCCACCAACCCGGTGGGCCAGAAGCGCGCCGCCCGCCTGGAGCGCAAGT
CGATGTGCAACTCGCTGTTCTGTGCTGTACTTCGGCCTGAACCAGCCGCACTCGCAGCTCG
CCCACCACACCATCTGCTTCGGCCCCGCGCTACCGCGAGCTGATCGACGAGATCTTACCG
GCTCGGCCCTGGCCGACGACTTCTCGCTGTACCTGCACTCGCCGTGCGTGACCGACCCGT
CGCTGGCCCCCGCGGGCTGCGCCTCGTTCTACGTGCTGGCCCCGGTGCCGCACCTGGGCA
ACGCCCCGCTGGACTGGGCCAGGAGGGCCGAAGCTGCGCGACCGCATCTTCGACTACC
TGGAGGAGCGCTACATGCCGGGCTGCGCTCGCAGCTCGTGACCCAGCGCATCTTACCC
CGGCCGACTTCCACGACACCCTGGACGCCACCTGGGCTCGGCCTTCTCGATCGAGCCGC
TGCTGACCCAGTCGGCCTGGTTCGCCCCGACAACCGCGACTCGGACATCGCCAACCTGT
ACCTGGTGGGCGCCGGCACCCACCCGGGCGCCGGCATCCCGGGCGTGGTGGCCTCGGCCA
AGGCCACCGCCTCGCTGATGATCGAGGACCTGCAGTGATCTGGGACGTGATGGGTAAG
CTGGACATCACTCCACAACGCAAAGGAGATATAGATGTCGAGCCGCCGCTGCTGGAC
CACGCCACCCAGACCATGGCCAACGGCTCGAAGTCGTTGCCACCGCCGCAAGCTGTTT
GACCCGGCCACCCGCCGCTCGGTGCTGATGCTGTACACCTGGTGCCGCCACTGCGACGAC
GTGATCGACGACGACCCACGGCTTCGCCTCGGAGGCCGCCGCCGAGGAGGAGGCCACC
CAGCGCCTGGCCCGCTGCGCACCTGACCCTGGCCGCTTCGAGGGCGCCGAGATGCAG
GACCCGGCCTTCGCCGCTTCCAGGAGGTGGCCCTGACCCACGGCATCACCCGCGCATG
GCCCTGGACCACCTGGACGGCTTCGCCATGGACGTGGCCAGACCCGCTACGTGACCTTC

GAGGACACCCTGCGCTACTGCTACCACGTGGCCGGCGTGGTGGGCCTGATGATGGCCCGC
GTGATGGGCGTGCGCGACGAGCGCGTGCTGGACCGCGCCTGCGACCTGGGCCTGGCCTTC
CAGCTCACCAACATCGCCCGGACATCATCGACGACGCCGCCATCGACCGCTGCTACCTG
CCGGCCGAGTGGCTGCAGGACGCCGGCCTGACCCCGGAGAACTACGCCGCCCGGAGAAC
CGCGCCGCCCTGGCCCGCGTGGCCGAGCGCCTGATCGACGCCGCCGAGCCGTACTACATC
TCGTGCGAGGCCGGCCTGCACGACCTGCCGCCGCGCTGCGCCTGGGCCATCGCCACCGCC
CGCTCGGTGTACCGCGAGATCGGCATCAAGGTGAAGGCCGCCGGCGGCTCGGCCTGGGAC
CGCCGCCAGCACACCTCGAAGGGCGAGAAGATCGCCATGCTGATGGCCGCCCCGGGCCAG
GTGATCCGCGCCAAGACCACCCGCGTGACCCCGCGCCCGGCCGCTGTGGCAGCGCCCG
GTGTGACTGTCCCCCAGTTCCAGTACCTGGTCATCATCCTGCCTTCAAAGGAGATATA
GATGCTGTGGATCTGGAACGCCCTGATCGTGTTCTGTGACCGTGATCGGCATGGAGGTGGT
GGCCGCCCTGGCCACAAGTACATCATGCACGGCTGGGGCTGGGGCTGGCACCTGTGCGA
CCACGAGCCGCGCAAGGGCGCCTTCGAGGTGAACGACCTGTACGCCGTGGTGTTCGCCGC
CCTGTGATCCTGCTGATCTACCTGGGCTCGACCGGCATGTGGCCGCTGCAGTGGATCGG
CGCCGGCATGACCGCCTACGGCCTGCTGTACTTCATGGTGCACGACGGCCTGGTGCACCA
GCGCTGGCCGTTCCGCTACATCCCGCGCAAGGGCTACCTGAAGCGCCTGTACATGGCCCA
CCGCATGCACCACGCCGTGCGCGGCAAGGAGGGCTGCGTGTGTTTGGTTCCTGTACGC
CCCGCCGCTGTGAAGCTGCAGGCCACCCTGCGCGAGCGCCACGGCGCCCGCGCCGGCGC
CGCCCGGACGCCAGGGCGGCGAGGACGAGCCGGCCTCGGGCAAGTGAGTTATATGGAG
GGGGCAAACGCTCTAGAACTAGTGGATCAAAGGAGATATAGATGTCGGCCGTGACCCCG
ATGTCGAGAGTGGTGCCAAACCAGGCCCTAATCGGCCTGACTTTAGCGGGGCTGATAGCC
ACGGCGTGGCTGAGTCTGCATATTTACGGGGTGTACTTCATCGTTGGACAATGTGGTCG
ATCCTGACGGTGCCGCTGATCGTGGCCTTCAGACGTGGCTGTCGGTAGGCCTGTTATC
GTTGCCACGACGCAATGCACGGCTCCCTAGCCCCGGGGAGGCCCCGCTGAACACCGCC
ATCGGGTCCCTGGCCCTAGGCCTGTACGCTGGCTTCAGGTTGCCCCCTCTGAAGACCGCC
CACCATGCCACCATGCCGACCGGGCACAGCCGACGACCCGGATTTTACGCGGACGCC
CCCCGTGCGTTCTGCGTGTTCTACGGCTTTTCCGTACCTACTTCGGCTGGAGGGAG
CTGGCCGTGCTGACCGTGTGGTGGCCGTGGCTGTTTAAATCCTGGGCGCCGAATGCCG
AACTTACTTGTGTTCTGGGCCGCCCGGCTCTATTATCGGCCTTGACGCTTTTCACCTTC

GGCACATGGCTGCCGCACCGACACACCGACGACGCCTTCCCGGACCACCACAACGCTCGC
ACTTCACCCTTTGGCCCCATCCTGTCTCTGCTGACCTGCTTCCACTTCGGCCGGCACCAT
GAGCACCACTGACTCCGTGGAAACCGTGGTGGAGGCTGTTCTCGTAGCGATACCGTCGA
CTTCGAGCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTG
TTTGTCGGTGAACGCTCTCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGT
ATTGGGCGCATGCATAAACTGCTGCCGTTTAGCCCGATAGCGTGGTGACCCACGGCGAT
TTTAGCCTGGATAACCTGATTTTCGATGAAGGCAAACCTGATTGGCTGCATTGATGTGGC
CGTGTGGGCATTGCGGATCGTTATCAGGATCTGGCCATTCTGTGGAACCTGCCTGGGCGAA
TTTAGCCCCGAGCCTGCAAAAACGTCTGTTTCAGAAATATGGCATTGATAATCCGGATATG
AACAACTGCAATTTCTCTGATGCTGGATGAATTTTTCTAAGACCCTGTCTAATCAAT
GCGGACCCTAGAGGTCCCCTTTTTATTTAAAAATTTTTTACAAAACGGTTTACAAGC
ATAAAATCTCTGAAGATGTGTATAAGAGACAGACTAGTCTTGGACTCCTGTTGATAGATC
CAGTAATGACCTCAGAACTCCATCTGGATTTGTTTCAAGACGCTCGGTTGCCGCCGGGCGT
TTTTTATTGGTGAGAATCCAGGGTCCCCTGGTTTAACTACACAAGTAGCGTCCTGAAC
GGAACCTTTCCGTTTTCCAGAATCTGATGTTCCATGTGACCTCCTAACATGGTAACGTT
CATGATTACCACTGCACTGCATCGTGCGGCGGATTGGGCGAAAAGCGTGTTTTCTAGTGC
TGCGCTGGGTGATCCGCGTCTACCGCGCGTCTGGTGAATGTTGCGGCGCAACTGGCCAA
ATATAGCGGCAAAAGCATTACCATTAGCAGCGAAGGCAGCAAAGCCATGCAGGAAGGCGC
GTATCGTTTTATTCGTAATCCGAACGTGAGCGCGGAAGCGATTCTGTAAGCGGGTGCCAT
GCAGACCGTGAACTGGCCCAGGAATTTCCGGAACCTGCTGGCAATTGAAGATACCACCTC
TCTGAGCTATCGTCATCAGGTGGCGGAAGAACTGGGCAAACCTGGGTAGCATTACAGATAA
AAGCCGTGGTTGGTGGGTGCATAGCGTGCTGCTGCTGGAAGCGACCACCTTTCGTACCGT
GGGCCTGCTGCATCAAGAATGGTGGATGCGTCCGGATGATCCGGCGGATGCGGATGAAAA
AGAAAGCGGCAAATGGCTGGCCGCTGCTGCAACTTCGCGTCTGAGAATGGGCAGCATGAT
GAGCAACGTGATTGCGGTGTGCGATCGTGAAGCGGATATTCATGCGTATCTGCAAGATAA
ACTGGCCCATAACGAACGTTTTGTGGTGGTAGCAAACATCCGCGTAAAGATGTGGAAAG
CGGCCTGTATCTGTATGATCACCTGAAAAACCAGCCGGAACCTGGGCGGCTATCAGATTAG
CATTCCGCAGAAAGGCGTGGTGGATAAACGTGGCAAACGTAAAAACCGTCCGGCGCGTAA
AGCGAGCCTGAGCCTGCGTAGCGGCCGTATTACCCTGAAACAGGGCAACATTACCCTGAA

CGCGGTGCTGGCCGAAGAAATCAATCCGCCGAAAGGCGAAACCCCGCTGAAATGGCTGCT
GCTGACCAGCGAGCCGGTGGAAAAGTCTGGCCCAAGCGCTGCGTGTGATTGATATTTATAC
CCATCGTTGGCGCATTGAAGAATTTACAAAAGCGTGGAACCGGGTGCGGGTGCGGAACG
TCAGCGTATGGAAGAACCGGATAACCTGGAACGTATGGTGAGCATTCTGAGCTTTGTGGC
GGTGCGTCTGCTGCAACTGCGTGAATCTTTTACTCCGCCGCAAGCACTGCGTGCGCAGGG
CCTGCTGAAAGAAGCGGAACACGTTGAAAGCCAGAGCGCGGAAACCGTGCTGACCCCGGA
TGAATGCCAACTGCTGGGCTATCTGGATAAAGGCAAACGCAAACGCAAAGAAAAAGCGGG
CAGCCTGCAATGGGCGTATATGGCGATTGCGCGTCTGGGCGGCTTTATGGATAGCAAACG
TACCGGCATTGCGAGCTGGGGTGCGCTGTGGGAAGGTTGGGAAGCGCTGCAAAGCAAAC
GGATGGCTTTCTGGCCGCGAAAGACCTGATGGCGCAGGGCATTAAAAATCTAATGGAATCG
AACCCAGCTTTCTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGC
AACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTG

CLAIMS

What is claimed:

1. An expression construct comprising:
 - a nucleic acid sequence for a *crtW* carotenoid ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2, the expression construct adapted to produce carotenoids in a biological host cell.
2. The expression construct of claim 1, wherein the expression construct is a plasmid.
3. The expression construct of claim 1 or 2, wherein the expression construct is integrated into the genome of the biological host cell.
4. A method of expressing a *crtW* carotenoid ketolase protein of SEQ ID NO: 2 the method comprising:
 - obtaining a *crtW* gene from a species of *Brevundimonas*,
 - creating an expression construct by operably linking a promoter sequence to the *crtW* gene,
 - introducing the expression construct into a biological host cell using a replicating plasmid or by genomic integration and propagating the cells under conditions that result in expression of the *crtW* carotenoid ketolase protein.
5. The method of claim 4, wherein the expression construct is adapted to produce carotenoids when functionally integrated into the biological host cell.
6. The method of claim 4 or 5, wherein the protein is further expressed in a biological host cell capable of using CO₂ and H₂ to satisfy at least part of its carbon and energy requirements.
7. The method of claim 4, 5, or 6, wherein the *crtW* gene is a ketolase gene from *Brevundimonas* strain OB307.
8. A method of producing a nucleic acid sequence encoding a *crtZ-crtW* carotenoid hydroxylase-ketolase fusion protein of SEQ ID NO: 4, the method comprising:
 - obtaining a *crtW* gene from a species of *Brevundimonas*,
 - adding a sequence encoding a ten amino acid linker peptide of SEQ ID NO: 5 to the 3'-end of the *crtW* gene;

- adding a sequence encoding a *crtZ* gene lacking the *N*-terminal methionine codon and containing a 3' stop codon to the 3'-end of the linker peptide to produce a DNA construct; and
- inserting the DNA construct into an expression vector.
9. The method of claim 8, wherein the nucleic acid sequence is part of an expression construct adapted to produce carotenoids when functionally introduced into a biological host cell.
 10. The method of claim 8 or 9, wherein the *crtZ-crtW* carotenoid hydroxylase-ketolase fusion protein is further expressed in a biological host cell capable of using CO₂ and H₂ to satisfy at least part of its carbon and energy requirements.
 11. The method of claim 8, 9, or 10, wherein the *crtW* sequence is a ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2.
 12. A suicide vector construct adapted for inserting a DNA sequence into a microbial genome of a bacterium using a transposon, the suicide vector construct comprising:
 - (a) the DNA sequence;
 - (b) an insert-flanking DNA comprising the nucleic acid sequence of SEQ ID NO: 9 that contains the transposon; and
 - (c) a suicide plasmid backbone.
 13. The suicide vector construct of claim 12, wherein the DNA sequence encodes a gene for producing a carotenoid.
 14. An expression construct encoding a *crtZ-crtW* carotenoid hydroxylase-ketolase fusion protein of SEQ ID NO: 4, wherein
 - (a) the *crtW* portion of the fusion is a ketolase gene from *Brevundimonas* strain OB307 that encodes the amino acid sequence of SEQ ID NO: 2, and
 - (b) the nucleic acid sequence is adapted to produce carotenoids when functionally integrated into a biological host cell.
 15. A food or feed additive product derived from biological cells containing expression constructs, expression vectors, or using the methods referenced in claims 1-14.