PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification
C12P 23/00, 7/26, C12N 9/02, 1/20, 15/00, C07H 21/04, C07K 14/00

(21) International Application Number: PCT/US99/10455
(22) International Filing Date: 21 May 1999 (21.05.99)
(30) Priority Data:
60/086,460 22 May 1998 (22.05.98) US

(71) Applicant (for all designated States except US): UNIVERSITY OF MARYLAND [US/US]; Office of Technology Liaison, 4312 Knox Road, College Park, MD 20742 (US).

(72) Inventor; and
(75) Inventor/Applicant (for US only): CUNNINGHAM, Francis, X. [US/US]; 2727 Washington Avenue, Chevy Chase, MD 20815 (US).


(54) Title: CAROTENOID KETOLASE GENES AND GENE PRODUCTS, PRODUCTION OF KETOACAROTENOIDS AND METHODS OF MODIFYING CAROTENOIDS USING THE GENES

(57) Abstract

A purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has the nucleic acid sequence of SEQ ID NO: 1 or 3, or has a sequence which encodes the amino acid sequence of SEQ ID NO: 2 or 4, as well as vectors and host cells containing them. Methods of use of the nucleic acid sequences to produce ketocarotenoid in host cells and methods of use of the nucleic acid sequences to modify the production of carotenoids in a host cell are included.
FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL  | Albania    | ES | Spain       | LS | Lesotho        |
| AM  | Armenia    | FI | Finland     | LT | Lithuania      |
| AT  | Austria     | FR | France      | LU | Luxembourg     |
| AU  | Australia   | GA | Gabon       | LV | Latvia         |
| AZ  | Azerbaijan  | GB | United Kingdom | MC | Monaco         |
| BA  | Bosnia and Herzegovina | GE | Georgia     | MD | Republic of Moldova |
| BB  | Barbados    | GH | Ghana       | MG | Madagascar     |
| BE  | Belgium     | GN | Guinea      | MK | The former Yugoslav Republic of Macedonia |
| BF  | Burkina Faso | GR | Greece      | ML | Mali           |
| BG  | Bulgaria    | HU | Hungary     | MN | Mongolia       |
| BJ  | Benin       | IE | Ireland     | MR | Mauritania     |
| BR  | Brazil      | IL | Israel      | MW | Malawi         |
| BY  | Belarus     | IS | Iceland     | MX | Mexico         |
| CA  | Canada      | IT | Italy       | NE | Niger          |
| CF  | Central African Republic | JP | Japan       | NL | Netherlands    |
| CG  | Congo       | KE | Kenya       | NO | Norway         |
| CH  | Switzerland | KG | Kyrgyzstan  | NZ | New Zealand    |
| CI  | Côte d’Ivoire | KP | Democratic People’s Republic of Korea | PL | Poland         |
| CM  | Cameroon    | KR | Republic of Korea | PT | Portugal       |
| CN  | China       | KE | Kazakhstan  | RO | Romania        |
| CU  | Cuba        | LC | Saint Lucia | RU | Russian Federation |
| CZ  | Czech Republic | LI | Liechtenstein | SD | Sudan          |
| DE  | Germany     | LK | Sri Lanka   | SE | Sweden         |
| DK  | Denmark     | LR | Liberia     | SG | Singapore      |
| EE  | Estonia     | SI | Slovenia    |
| SK  | Slovakia    |
| SN  | Senegal     |
| SE  | Swaziland   |
| TD  | Chad        |
| TG  | Togo        |
| TJ  | Tajikistan  |
| TM  | Turkmenistan |
| TR  | Turkey      |
| TT  | Trinidad and Tobago |
| UA  | Ukraine     |
| UA  | Uganda      |
| US  | United States of America |
| UZ  | Uzbekistan  |
| VN  | Viet Nam    |
| VU  | Yugoslavia  |
| ZW  | Zimbabwe    |
CAROTENOID KETOLASE GENES AND GENE PRODUCTS,  
PRODUCTION OF KETOCAROTENOIDS AND METHODS OF  
MODIFYING CAROTENOID USING THE GENES

BACKGROUND OF THE INVENTION

Carotenoids are widely distributed natural pigments that are responsible for many of the yellow, orange and red colors seen in living organisms. They have important commercial uses as coloring agents in the food industry, as feed and food additives, in cosmetics and as provitamin A precursors.

The plant species Adonis aestivalis produces flowers with petals that are deep red in color and nearly black at the base of the petals due to the accumulation of ketocarotenoid and other carotenoid pigments (Neamtu et al., Rev. Roum. Biochim. 6:157, 1969). This pattern of carotenoid accumulation accounts for the common name of some varieties of this species: summer pheasant's eye.

Among the carotenoids identified in the petals of the red petal varieties of these various species is the ketocarotenoid astaxanthin (3,3'-dihydroxy-4,4'-diketo-b,b-carotene; see Figure 1). Various other ketocarotenoids (see Figure 1) including 3-hydroxyechinenone (3-hydroxy-4-keto-b,b-carotene), adonirubin (3-hydroxy-4,4'-diketo-b,b-carotene) adonixanthin (3,3'-dihydroxy-4-keto-b,b-carotene) and isoxeaxanthin (4,4'-dihydroxy-b,b-carotene; see T.W. Goodwin, The Biochemistry of the Carotenoids, vol I. Plants, 2nd edition, 1980, page 147) have also been reported. The latter compound is consistent with speculation that the 4-hydroxy may be an intermediate in the formation of the 4-keto group.

SUMMARY OF THE INVENTION

There is appreciable interest in the biological production of carotenoids, in particular the orange-colored ketocarotenoids such as astaxanthin and canthaxanthin (Figure 1), and in the modification of carotenoid composition. For this reason, an A. aestivalis flower cDNA library was constructed and screened for cDNAs encoding enzymes (hereinafter referred to as "ketolases" although the specific biochemical activity has not yet been established) involved in the conversion of b-carotene into orange compounds with absorption properties similar to those exhibited by common ketocarotenoids such as canthaxanthin (Figure 1). Two distinctly different Adonis aestivalis cDNAs were obtained from among a number of cDNAs that were selected on this basis.
Thus, a first aspect of the present invention is a purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has the nucleic acid sequence of SEQ ID NO: 1 or 3.

The invention also includes a purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and having the amino acid sequence of SEQ ID NO: 2 or 4.

The invention also includes vectors which comprise any portion of the nucleic acid sequences listed above, and host cells transformed with such vectors.

Another aspect of the present invention is a method of producing a ketocarotenoid in a host cell, the method comprising

- inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having ketolase enzyme activity and comprises (1) SEQ ID NO: 1 or 3 or (2) a sequence which encodes the amino acid sequence of SEQ ID NO: 2 or 4, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and
- expressing the heterologous nucleic acid sequence, thereby producing the ketolase enzyme.

Another subject of the present invention is a method of modifying the production of carotenoids in a host cell, relative to an untransformed host cell, the method comprising

- inserting into a host cell which already produces carotenoids a vector comprising a heterologous nucleic acid sequence which encodes for a protein having ketolase enzyme activity and comprises (1) SEQ ID NO: 1 or 3 or (2) a sequence which encodes the amino acid sequence of SEQ ID NO: 2 or 4, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and
- expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to an untransformed host cell.

**BRIEF DESCRIPTION OF THE DRAWINGS**

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by
reference to the following detailed description when considered in connection with the accompanying drawings.

Figure 1 illustrates structures and biochemical routes leading from b-carotene to various of the ketocarotenoids referred to in the text. Conversion of β-carotene to astaxanthin by a hydroxylase enzyme (Hy) and a ketolase enzyme (keto) could proceed via any one or all of several possible routes depending on the order of the reactions.

Figure 2 illustrates the beta ring structure of b-carotene and various modifications of this parent ring that might be produced through the action of the products of the A. aestivialis ketolase cDNAs. Also shown is the structure of the epsilon ring, not found to be a substrate for the A. aestivialis ketolases and present in carotenoids such as d-carotene, e-carotene, a-carotene and lutein.

Figure 3 illustrate results obtained with TLC (thin layer chromatography) separation of carotenoid pigments extracted from E. coli cultures, previously engineered to produce b-carotene, but that now also contain the A. aestivialis ketolase cDNAs and/or other introduced genes and cDNAs. The Figure indicates the empty plasmid vector pBluescript SK- (SK-), the Adonis aestivialis ketolase 1 cDNA in this plasmid vector (Ad keto1), the Haematococcus pluvialis ketolase cDNA in this plasmid vector Hp keto), or the Arabidopsis β-carotene hydroxylase cDNA (At Ohase). Bands that were orange in color are shown here with a darker fill than those with a yellow color. Identities of various bands are indicated to the right of the band.

Figure 4 illustrates the absorption spectrum of one of the orange carotenoids produced from b-carotene via the action of the Adonis ketolases and makes clear the similarity of the spectrum to that of canthaxanthin. Absorption spectra (in acetone) of β-carotene, canthaxanthin and an unknown orange product (orange band #1; the lower orange band in the first lane of Figure 3) extracted from cultures after introduction of the Adonis aestivalis keto1 cDNA (SEQ ID NO: 1) in cells of E. coli that otherwise produce and accumulate β-carotene. The absorption spectrum of the unknown resembles that of canthaxanthin but the compound migrates to a position below echinenone on RP18
TLC plates developed with a mobile phase of methanol:acetone (1:1 by volume). The absorption spectrum of orange band #2 also is similar to that of canthaxanthin but it migrates more rapidly than canthaxanthin indicating that it is probably a more polar compound.

Figure 5 shows SEQ ID NO: 5 (the sequence shown in this Figure includes SEQ ID NO: 1 and also includes some of the flanking DNA from the adaption DNA and the multiple cloning site (MCS) of the library cloning vector, which sequences are shown in bold).

Figure 6 shows SEQ ID NO: 6 (the sequence shown in this Figure includes SEQ ID NO: 2 and also includes a translation of amino acids resulting from the adaption DNA and the multiple cloning site (MCS) of the library cloning vector and the start codon from the plasmid vector pTrChis, which sequences are shown in bold and capitalized).

Figure 7 shows SEQ ID NO: 7 (the sequence shown in this Figure includes SEQ ID NO: 3 and also includes some of the flanking DNA from the adaption DNA and the multiple cloning site (MCS) of the library cloning vector, which sequences are shown in bold).

Figure 8 shows SEQ ID NO: 8 (the sequence shown in this Figure includes SEQ ID NO: 4 and also includes a translation of amino acids resulting from the adaption DNA and the multiple cloning site (MCS) of the library cloning vector and the start codon from the plasmid vector, which sequences are shown in bold and capitalized).

Figure 9 shows a "Gap" alignment of the two Adonis ketolase sequences of the invention. A truncated version of SEQ ID NO: 1 is shown in this Figure for comparative purposes, and is designated SEQ ID NO: 9. The percentage identity was calculated to be 91.107.

Figure 10 shows a "Gap" alignment of SEQ ID NO: 2 and 4. The following results were found:

| Gap weight: | 12 | average match: 2.912 |
| Length weight: | 4 | average mismatch: -2.003 |
Figure 11 shows a comparison between SEQ ID NO: 2 and the *Arabidopsis thaliana* β-carotene hydroxylase enzyme (GenBank U58919) (SEQ ID NO: 10).

Figure 12A shows gDNA (SEQ ID NO: 11) immediately upstream of the cDNA of SEQ ID NO: 3. The sequence was obtained from a PCR product generated using the GenomeWalker kit of Clontech Laboratories, Inc. (1020 East Meadow Circle, Palo Alto, CA 94303-4230) and nested primers specific to the ketolases of *Adonis aestivalis* (cagaatcggctgtcttattgtcctcc (SEQ ID NO: 17) and caatggagaatatcaaggtctttgttct (SEQ ID NO: 18)). The termination codon upstream of and in-frame with initiation codon (TAA at positions 204-206) is shown in bold. Initiation codon (ATG) is also shown in bold.

Figure 12B (SEQ ID NO: 12) indicates that the full length polypeptide of SEQ ID NO: 4 begins with the amino acids MAA (shown in bold) immediately preceding the ketolase sequence shown in Figure 8. A similar MAA amino acid sequence immediately preceding SEQ ID NO: 1 is also expected.

Figure 13 shows an alignment of SEQ ID NO: 2, SEQ ID NO: 12, an *Arabidopsis* β-carotene hydroxylase enzyme (predicted product of GenBank U58919) (SEQ ID NO: 13), a putative second *Arabidopsis* hydroxylase predicted by genomic DNA sequence (GenBank AB025606; the exon/intron junctions were chosen with reference to the product of the *Arabidopsis* β-carotene hydroxylase cDNA u58919) (SEQ ID NO: 14), and two *Capsicum annuum* β-carotene hydroxylases (predicted products of GenBank Y09722 and Y09225) (SEQ ID NO: 15 and 16).

**DESCRIPTION OF THE PREFERRED EMBODIMENTS**

The present invention is directed to a purified nucleic acid sequence which
encodes for a protein having ketolase enzyme activity and has the nucleic acid
sequence of SEQ ID NO: 1 or 3.

The invention also includes a purified nucleic acid sequence which encodes for
a protein having ketolase enzyme activity and having the amino acid sequence of SEQ
ID NO: 2 or 4.

Two different but closely-related nucleic acids have been isolated. The
sequences of the longest example of each are presented herein. Sequencing which
has subsequently been conducted of upstream genomic DNA indicates that SEQ ID
NO: 3 lacks bases encoding the first three amino acids (MAA; see Figure 12). Likely,
this is also the case for SEQ ID NO: 1, but the upstream genomic sequences have not
yet been obtained for this nucleic acid.

The two different Adonis ketolases denoted in SEQ ID NO: 1 and 3 are similar
in sequence, sharing about 91% identity, as determined by the Gap program discussed
below (see Figure 9). The predicted amino acid sequences of the enzymes denoted in
SEQ ID NO: 2 and 4 share about 92% similarity and about 90% identity, also as
determined by the Gap program (see Figure 10).

Therefore, it is clear that certain modifications of SEQ ID NO: 1 or 3 or SEQ ID
NO: 2 or 4 can take place without destroying the activity of the enzyme. Note also that
certain truncated versions of the cDNAs of SEQ ID NO: 1 or 3 were found to be
functional (i.e., these cDNAs retained the property of causing the conversion of b-
carotene to orange compounds). Also, the Arabidopsis β-carotene hydroxylase
(GenBank U58919), aligned with the ketolase SEQ ID NO: 2 in Figure 11, retains
catalytic function when truncated to yield a polypeptide that lacks the first 129 amino
acids (Sun et al., 1996). From the alignment in Figure 11, therefore, this would suggest
that the two ketolases of the invention retain catalytic activity after truncation to remove
bases encoding the first 132 amino acids.

Thus, the present invention is intended to include those ketolase nucleic acid
and amino acid sequences in which substitutions, deletions, additions or other
modifications have taken place, as compared to SEQ ID NO: 1 or 3 or SEQ ID NO: 2
or 4, without destroying the activity of the ketolase enzyme. Preferably, the
substitutions, deletions, additions or other modifications take place at those positions
which already show dissimilarity between the present sequences. For SEQ ID NO: 1,


For SEQ ID NO: 2 and 4, as shown in Figure 10, the following amino acids can be substituted or deleted, or additions or other modifications can be made, without destroying the activity of the ketolase enzyme: positions 7, 8, 12, 18, 21, 22, 25, 26, 36, 37, 45, 47-49, 56, 73, 83, 85, 97, 99, 130, 144, 150, 157, 166, 218, 244, 279, 299 and 304. Therefore, the present invention also intends to cover amino acid sequences where such changes have been made.

In each case, nucleic acid and amino acid sequence similarity and identity are measured using sequence analysis software, for example, the Sequence Analysis, Gap, or BestFit software packages of the Genetics Computer Group (University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wisconsin 53705), MEGAlign (DNASTar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), or MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software uses algorithms to match similar sequences by assigning degrees of identity to various substitutions, deletions, and other modifications, and includes detailed instructions as to useful parameters, etc., such that those of routine skill in the art can easily compare sequence similarities and identities. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the Gap program discussed above. This program finds the alignment of two complete
sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the BestFit program discussed above. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

Conservative (i.e. similar) substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (see Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132 (1982)), or on the basis of the ability to assume similar polypeptide secondary structure (see Chou and Fasman, *Adv. Enzymol.* 47: 45-148 (1978)).

If comparison is made between nucleotide sequences, preferably the length of comparison sequences is at least 50 nucleotides, more preferably at least 60 nucleotides, at least 75 nucleotides or at least 100 nucleotides. It is most preferred if comparison is made between the nucleic acid sequences encoding the enzyme coding regions necessary for enzyme activity. If comparison is made between amino acid sequences, preferably the length of comparison is at least 20 amino acids, more preferably at least 30 amino acids, at least 40 amino acids or at least 50 amino acids. It is most preferred if comparison is made between the amino acid sequences in the enzyme coding regions necessary for enzyme activity.

While the two different Adonis ketolase enzymes of the present invention are similar in sequence, previously-described bacterial (Misawa et al., 1995), cyanobacterial (Fernandez-Gonzalez et al., 1997), and green algal (*Haematococcus pluvialis*; Lotan et al., 1995; Kajiwara et al., 1995) β-carotene ketolase enzymes bear little resemblance to the Adonis ketolases, although certain histidine motifs and features of the predicted secondary structure are common to the polypeptides predicted by both groups (Cunningham and Gantt, 1998).

The present invention also includes vectors containing the nucleic acids of the invention. Suitable vectors according to the present invention comprise a gene encoding a ketolase enzyme as described above, wherein the gene is operably linked
to a suitable promoter. Suitable promoters for the vector can be constructed using techniques well known in the art (see, for example, Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 1991). Suitable vectors for eukaryotic expression in plants are described in Fray et al., (1995; Plant J. 8:693-701) and Misawa et al, (1994; Plant J. 6:481-489). Suitable vectors for prokaryotic expression include pACYC184, pUC119, and pBR322 (available from New England BioLabs, Beverly, MA) and pTrcHis (Invitrogen) and pET28 (Novagen) and derivatives thereof. The vectors of the present invention can additionally contain regulatory elements such as promoters, repressors, selectable markers such as antibiotic resistance genes, etc., the construction of which is very well known in the art.

The genes encoding the ketolase enzymes as described above, when cloned into a suitable expression vector, can be used to overexpress these enzymes in a host cell expression system or to inhibit the expression of these enzymes. For example, a vector containing a gene of the invention may be used to increase the amount of ketocarotenoids in an organism and thereby alter the nutritional or commercial value or pharmacology of the organism. A vector containing a gene of the invention may also be used to modify the carotenoid production in an organism.

Therefore, the present invention includes a method of producing a ketocarotenoid in a host cell, the method comprising

inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having ketolase enzyme activity and comprises (1) SEQ ID NO: 1 or 3 or (2) a sequence which encodes the amino acid sequence of

SEQ ID NO: 2 or 4, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and

expressing the heterologous nucleic acid sequence, thereby producing the ketocarotenoid.

The present invention also includes a method of modifying the production of carotenoids in a host cell, relative to an untransformed host cell, the method comprising

inserting into a host cell which already produces carotenoids a vector comprising a heterologous nucleic acid sequence which encodes for a protein having
ketolase enzyme activity and comprises (1) SEQ ID NO: 1 or 3 or (2) a sequence which encodes the amino acid sequence of SEQ ID NO: 2 or 4, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to an untransformed host cell.

The term "modifying the production" means that the amount of carotenoids produced can be enhanced, reduced, or left the same, as compared to an untransformed host cell. In accordance with one embodiment of the present invention, the make-up of the carotenoids (i.e., the type of carotenoids produced) is changed vis a vis each other, and this change in make-up may result in either a net gain, net loss, or no net change in the amount of carotenoids produced in the cell. In accordance with another embodiment of the present invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) is enhanced by the insertion of the ketolase enzyme-encoding nucleic acid. In yet another embodiment of the invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) may be reduced or inhibited by a number of different approaches available to those skilled in the art, including but not limited to such methodologies or approaches as anti-sense (e.g., Gray et al. (1992), Plant Mol. Biol. 19:69-87), ribozymes (e.g., Wegener et al. (1994) Mol. Gen. Genet. 1994 Nov 15;245(4):465-470), co-suppression (e.g. Fray et al. (1993) Plant Mol. Biol. 22:589-602), targeted disruption of the gene (e.g., Schaefer et al. Plant J. 11:1195-1206, 1997), intracellular antibodies (e.g., see Rondon et al. (1997) Annu. Rev. Microbiol. 51:257-283) or whatever other approaches rely on the knowledge or availability of the nucleic acid sequences of the invention, or the enzymes encoded thereby.

Host systems according to the present invention preferably comprise any organism which is capable of producing carotenoids, or which already produces carotenoids. Such organisms include plants, algae, certain bacteria, cyanobacteria and other photosynthetic bacteria. Transformation of these hosts with vectors according to the present invention can be done using standard techniques. See, for example, Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor

Alternatively, transgenic organisms can be constructed which include the nucleic acid sequences of the present invention. The incorporation of these sequences can allow the controlling of carotenoid biosynthesis, content, or composition in the host cell. These transgenic systems can be constructed to incorporate sequences which allow for the overexpression of the various nucleic acid sequences of the present invention. Transgenic systems can also be constructed which allow for the underexpression of the various nucleic acid sequences of the present invention. Such systems may contain anti-sense expression of the nucleic acid sequences of the present invention. Such anti-sense expression would result in the accumulation of the substrates of the enzyme encoded by the sense strand.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

**EXAMPLE 1**

Isolation of plant cDNAs that convert b-carotene into compounds with ketocarotenoid-like spectra

A flower cDNA library from the plant Adonis aestivalis was introduced into a strain of Escherichia coli engineered to accumulate the yellow carotenoid pigment b-carotene (see Cunningham et al., Plant Cell 8:1613-26, 1996). This strain of E. coli normally forms yellow colonies when cultures are spread on a solid agar growth medium. Ketocarotenoids that are derived from b-carotene, such as echinenone and canthaxanthin (Figure 1), are, in contrast, orange to orange-red in color. Colonies that were orange rather than yellow in color were visually selected, and the DNA sequences of the Adonis aestivalis cDNAs within the plasmid vectors contained in these colonies were ascertained. Two distinct cDNAs were obtained from analysis of cDNA inserts in plasmids obtained from approximately 10 selected colonies. The DNA sequences of these two ketolase cDNAs are presented herein.

The products produced by the ketolases of the invention which have been
expressed in a β-carotene-accumulating strain of *Eschericia coli* have not yet been identified. As many as 5 or 6 different colored bands, in addition to the substrate β-carotene, may readily be discerned by C18 TLC separation (see Figure 3). To provide appropriate standards to assist in identification, an *H. pluvialis* ketolase and an Arabidopsis β-carotene hydroxylase were separately introduced into the β-carotene-accumulating *E. coli* to produce echinenone (3-keto-β,β-carotene) and canthaxanthin (3,3'-diketo-β,β-carotene) or β-cryptoxanthin (4-hydroxy-β,β-carotene) and zeaxanthin (4,4'-dihydroxy-β,β-carotene). None of the compounds formed in the presence of the ketolases of the invention (no difference was observed in products formed in the presence of the two different nucleic acid sequences of the invention) both migrate in the TLC system and have the absorption spectrum expected for echinenone, canthaxanthin, β-cryptoxanthin, or zeaxanthin. Two of the colored TLC bands produced in the presence of the Adonis ketolase cDNAs are orange in color. Orange band #1 has an absorption spectrum similar to that of canthaxanthin (see Figure 4) but migrates in a position that indicates a polarity intermediate to echinenone and β-carotene. Orange band #2 also has an absorption spectrum like that of canthaxanthin but migrates in a position that indicates a polarity intermediate to canthaxanthin and zeaxanthin (see Figure 3). The absorption spectra and TLC results suggest that the two orange products could be desaturated at the 3-4 positions of both rings (3,4,-didehydro; see Figure 2). Orange band #1 (see Figure 3) might then be 3,4,3',4'-tetrahydro-β,β-carotene. To substantially affect the absorption spectrum of the substrate β-carotene, any modifications very likely involve a carbon that lies in conjugation with the conjugated chain of carbon-carbon double bonds that constitute the chromophore (Goodwin, 1980; *The Biochemistry of the Carotenoids*, volume I; 2nd edition, Chapman and Hall). For the spectra obtained, only the carbons at the number 4 position of the two rings appear to be plausible locations for modification. The multitude and TLC migrations of the yellow and orange products produced from the symmetrical β-carotene, however, also indicates that the enzymes of the invention carry out more than a single type of reaction. The apparent homology of the ketolases of the invention to the Arabidopsis β-carotene hydroxylase would suggest that compounds with a hydroxyl at the 3 and/or 4 positions of one or both rings are another possible outcome (see Figure 2). In fact, such compounds have been identified in Adonis (see
above), and it has long been conjectured that a hydroxyl at position 4 is an intermediate in the formation of the 4-keto (e.g. crustaxanthin, a 3,3',4,4' tetrahydroxy carotenoid that might be a precursor for astaxanthin in the exoskeleton of the lobster). The histidine motifs and secondary structure in common to the hydroxylase and ketolase enzymes are characteristics of a large group of di-iron oxygenases whose members also include examples of desaturases (J. Shanklin, 1998, *Ann. Rev. Plant Physiol. Plant Mol. Biol.*), therefore a 3-4 desaturation (and/or perhaps a 2-3 desaturation in one or more of the yellow compounds) would also seem a plausible outcome.

To summarize the results of this example for the Adonis ketolases of the invention, a number of different carotenoids, including two with ketocarotenoid-like spectra, are produced from β-carotene via the action of the products of either of the two different nucleic acids of the invention. These orange compounds appear to be the major products. Truncation and fusion of the cDNAs to a stronger promoter in the vector pTrcHis (Invitrogen) was detrimental to growth of *E. coli* but did result in improved yield of the most polar orange product (orange band #2 in Figure 3). Introduction of a cyanobacterial ferredoxin did not change the yield or relative amounts of the various products. Without being bound by theory, it may be that the ketocarotenoids produced in flower petals of Adonis actually include the as yet unidentified orange compounds that are produced in *E. coli* using the nucleic acids of the invention.

**EXAMPLE 2**

**Substrate specificity of the Adonis ketolases**

Carotenoids with ε rings are common in plants. The ε ring differs from the b ring only in the position of the double bond within the ring (Figure 2). The ε ring is reported to be a poor substrate for the Arabidopsis b-carotene hydroxylase (Sun et al., 1996). The Adonis ketolase cDNAs were introduced into strains of *E. coli* engineered (Cunningham et al., 1996) to accumulate carotenoids with one or two ε rings (d-carotene and ε-carotene), or the acyclic carotenoid lycopene. TLC analysis of acetone extracts revealed that these carotenoids were not modified by the Adonis ketolases, as indicated by a lack of any new products formed. Products produced in *E. coli* engineered to accumulate zeaxanthin (Sun et al., 1996) appeared to be the same as
for β-carotene accumulating cultures indicating that a 3-OH is likely to be one of the functional groups introduced to the b ring by the Adonis ketolases. The more polar orange band produced from b-carotene through the action of the Adonis ketolases (e.g., orange band 2 in Figure 3), therefore, could very well be 3,3'-dihydroxy-3,4,3',4'-tetrahydro-b,b-carotene.

The references cited in the application, along with the following references, are incorporated by reference:


Harker M, et al. (1997) Biosynthesis of ketocarotenoids in transgenic cyanobacteria expressing the algal gene for beta-C-4-oxygenase, crtO. FEBS Lett. 404:129-34


I claim:

1. A method of producing a ketocarotenoid in a host cell, the method comprising
   inserting into the host cell a vector comprising a heterologous nucleic acid
   sequence which encodes for a protein having ketolase enzyme activity and has the
   nucleic acid sequence of SEQ ID NO: 1 or 3, wherein the heterologous nucleic acid
   sequence is operably linked to a promoter; and
   expressing the heterologous nucleic acid sequence, thereby producing
   the ketocarotenoid.

2. The method of claim 1, wherein the host cell is selected from the group
   consisting of a bacterial cell, an algal cell and a plant cell.

3. A method of producing a ketocarotenoid in a host cell, the method comprising
   inserting into the host cell a vector comprising a heterologous nucleic acid
   sequence which encodes for a protein having ketolase enzyme activity and has a
   sequence which encodes the amino acid sequence of SEQ ID NO: 2 or 4, wherein the
   heterologous nucleic acid sequence is operably linked to a promoter; and
   expressing the heterologous nucleic acid sequence, thereby producing
   the ketocarotenoid.

4. The method of claim 3, wherein the host cell is selected from the group
   consisting of a bacterial cell, an algal cell and a plant cell.

5. A method of modifying the production of carotenoids in a host cell, relative to an
   untransformed host cell, the method comprising
   inserting into a host cell which already produces carotenoids a vector
   comprising a heterologous nucleic acid sequence which encodes for a protein having
   ketolase enzyme activity and has the nucleic acid sequence of SEQ ID NO: 1 or 3,
   wherein the heterologous nucleic acid sequence is operably linked to a promoter; and
   expressing the heterologous nucleic acid sequence in the host cell to
   modify the production of the carotenoids in the host cell, relative to an untransformed
host cell.

6. The method of claim 5, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell and a plant cell.

7. A method of modifying the production of carotenoids in a host cell, relative to an untransformed host cell, the method comprising inserting into a host cell which already produces carotenoids a vector comprising a heterologous nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has a sequence which encodes the amino acid sequence of SEQ ID NO: 2 or 4, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to an untransformed host cell.

8. The method of claim 7, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell and a plant cell.

9. A purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has the nucleic acid sequence of SEQ ID NO: 1.

10. A purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has the nucleic acid sequence of SEQ ID NO: 3.

11. A purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has a sequence which encodes the amino acid sequence of SEQ ID NO: 2.

12. A purified nucleic acid sequence which encodes for a protein having ketolase enzyme activity and has a sequence which encodes the amino acid sequence of SEQ ID NO: 4.
13. A vector which comprises the nucleic acid sequence of any one of claims 9-12, wherein the nucleic acid sequence is operably linked to a promoter.

14. A host cell which is transformed with the vector of claim 13.

15. The host cell of claim 14, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell and a plant cell.

16. The host cell of claim 14, wherein the host cell is a photosynthetic cell.

17. The host cell of claim 14, wherein the host cell contains a ketocarotenoid.

18. The host cell of claim 14, wherein the host cell contains modified levels of carotenoids, relative to an untransformed host cell.

19. A purified ketolase enzyme which is encoded by the amino acid sequence of SEQ ID NO: 2.

20. A purified ketolase enzyme which is encoded by the amino acid sequence of SEQ ID NO: 4.
FIGURE 3

Solvent front ———

Ad keto 1  SK-keto  Hp keto  At OHase

orange band 2  canthaxanthin

zeaxanthin  β-cryptoxanthin

echinenone

orange band 1  β-carotene

Origin ———
Absorbance

β-carotene

Canthaxanthin

Orange product formed from β-carotene in the presence of an Adonis "ketolase"

Wavelength (nm)

400  440  480  520  560  600
Figure 5 [SEQ ID NO: 5]

```
-23  ggg  ctcgaggaaat  tcggcagcag
   1  agcaatctca  ggtttcagtta  caagtttacct  tttccacaga  aatctcttgt
  51  tgcactcaca  acaagacatt  ctcacccgccc  catgtttgtct  ctttcttcca
101  gttgtgggtgg  agtgcctat  gagaagaaaa  aagacacatc  gtgctgcagt
151  tatctgctct  gttgcagaga  gaacaagagaa  cctgtgatatt  cctcaaatgg
201  aagaagagga  agagaacgag  gaagaactaa  tagaacagag  ggattctggc
251  ataattccata  taaagaaaaac  gctagggggg  aaacaattca  gacggtccac
301  tggctccatt  gtcgcaccccg  tatccttgtct  tgggatctctt  tcaatgtatcg
351  gactgtcgtt  ttatctcacaag  tttttcacgcgc  taatggaggtg  tggagatatt
401  cctgtgcagatg  aagtggggt  tacgtttgctg  gctttttgttg  ctagctgcagat
451  tggcagcggaa  tttttgtcag  gatggtttcag  caaaagaacctt  tggcagcggatt
501  cttttgtgtca  cattcacaag  ttctcaccata  ggtcagcgaaga  agggcgcctc
551  gatgttccatg  atgtgtttgc  tattataaac  ggcctctctctg  ctattgtctt
601  tatcaatattg  gattctcaca  atgaaaggctt  cctttccttgg  gcttcttgttg
651  gtaccgcctct  tggaaacgaca  gtctgtgcca  tggctttatact  tttttttcagc
701  aatggcctct  cacaccggaag  gtttccagta  gggctttattg  caacagtcgcc
751  ttattttcaca  aagcctgggctg  cagctcaccat  aatcccatcaca  tcaggaatat
801  ttcaggggtgt  accatatggcct  cttgtctcttg  gaccccagga  atttggaaaga
851  gtaagagagag  gcactgaaga  attggagagg  gtgtcagttgc  gtacagcat
901  acaagcgcaca  tcatctcatct  gaatcaactct  tttttcatctt  atgggttttt
951  agttttacgag  ttgttacagtt  cacacatgtt  tgcgtcgttctg  taaaaatcaca
1001  gttaccatagcttttttag  aatattttattgtgatag  gtgcgggagt
1051  tacgttacaa  aagggcaaat  ctattttgttga  ggaatccaat  tatataaat
1101  aaaaaatgagagtcgtgtgattcttttcttaatcatcaatatatatatta
1151  ataaagcgcataaaaaaaa  aaaaaa  ctcgag
```

SUBSTITUTE SHEET (RULE 26)
Figure 6 [SEQ ID NO: 6]

MGLQEFGR
aisvfstsys fhknl1l1hsk qdilenpcol fspvvvespm rkkkthraac
icsvaertrn 1dipgieee eeneelietg dsghiikkt lggkgsrfrst
gsivapvsc1 gilsmigpav yfkfsrlmec gdipvaemgi tfaafvaai
gteflsgwvh kelwhdswly ihkshhrsrk grfehndvfa iinalpaial
inygfsnegl lpgacfgtgl gttvcgmayi flhnglshrr fpvgliianvp
yfhklaahq inhsqkgfqqv pfglflgpqee leevrggteee lervisrtak
rtgsst*
-23
ggg ctgcaggaat tcggcagcag
  1 agcaatttca gtgttcagtt caggttatc tttctcaaa agatctctttg
  51 tggactcaaa accaarattt ctcacgaaag caggtcctgtt atttctctca
 101 gttgtgatca tgtgctctat gagaagaaaa aagaaacatg gtgaatcctg
 151 tatctgctcc gtggcagggga gaaacgggaa cctgtatatt cctcaatttg
 201 aagaagaggag agagaatggtg gaagaactaa tagaacagac cgattctgac
 251 atagtgacata taagaagaaac actagggggg aaacaatca gattccccac
 301 tggctccatt tgtcagcccg tctctctctg tgtgatatcc tcaatgatgg
 351 gacctgtgctg ttatctcaag tttccagggc taatggaggg tcggatgata
 401 cctgtacag aaatggggt tacgttttgcc aaccttttgtg ctgctgctgt
 451 tgtgcagggag tttttgctag catgggttca caaagaactc tgtgcagcgt
 501 cttttggtgta cttacacaaag tttccaccact ggtcgagaaa aggccgcttc
 551 gaggctcaatg atgtgctggc tattattac agcatctcccg ctatttgctt
 601 tatcaattat gatttttcca atgaagggct cctttcctgga gcgtgcttttg
 651 gtgtgctgct tgtgaaacaaca gttgctggta tggtctttcag ttttttcac
 701 aatggccttat cacacccgaag gttcccgagta ttggtattaag cgaagctccc
 751 ttatatccac aagcttggtgt cagcttcacca aataacacac tcagggaaaa
 801 tcaggggtgt accaatgggct ctgctttcct gtgcgatcct gcacggaagaa
 851 gtaagaggag gcattgaga agtggagagg gtaatcagtc gtaacaactaa
 901 acgaacgcua ccatctacct gattcataat tttaatatat ataaggttattt
 951 agtttatcgg tgtataaaaa tcacacatcc gttatcgtttt agtaagctcag
1001 agtttaagata ctctctctttt agaatatttt ttgatgtata gtcgccttgtg
1051 atactgtttac actattcgttt gtcgaattcc attatatataa aataaaaaaa
1101 aaaaaaaaaa aa ctgag
Figure 8 [SEQ ID NO: 8]

MGLQEFGR

aisvffsgys fyknllldsk pnilkkpocht fspvvimspm rkkkhhgdpc
icsvgtrtrn ldipqieee enveeliegt dsdivhikkt lggkgskrpt
gsivapvsccl gilsmpav yfkfsrlmeg gdipvaemgi tfatifvaav
gteflsawvh kelweslwv ihykshhrsry grfenvdfya iinalpaial
inygfsnegl lpgacfgvgl gtvgcmayi flhnglsrrr fpvwnianvp
yfhhklaahq ihhskgfqqv pfllflgpgke leevrggtel lervisrttk
rtqpest*
Figure 9: Gap of SEQ ID NO: 9 and SEQ ID NO: 3

1 agcaatctcagtggttcagtaaagttatctttcacaagaatctctttg 50
  1 agcaatgtctcagtggttcagtaaagttatctttcacaagaatctctttg 50
51 tgcactcaaaaacaagcacattctcaaccgcccgtatgtttgctttttctcttca 100
  51 tgcactcaaaaacaagcacattctcaaccgcccgtatgtttgctttttctcttca 100
101 gttggtgggagtgccatatgagaagaaaaagacacatcgtgctgcatg 150
  101 gttggtgggagtgccatatgagaagaaaaagacacatcgtgctgcatg 150
151 tatctgcttcgttgcagagaaagacatttgcattctctttcaaatgg 200
  151 tatctgcttcgttgcagagaaagacatttgcattctctttcaaatgg 200
201 aagaagggaagaggggaagaggggaagagaactaatagaaacagagggattctggc 250
  201 aagaagggaagaggggaagaggggaagagaactaatagaaacagagggattctggc 250
251 ataattcataaaagaaacgtagggggaaacaatcaagacggtcaca 300
  251 ataattcataaaagaaacgtagggggaaacaatcaagacggtcaca 300
301 tggctccattgtgccacccgtatcttgcttgtgctgacgctcttaatgatcg 350
  301 tggctccattgtgccacccgtatcttgcttgtgctgacgctcttaatgatcg 350

SUBSTITUTE SHEET (RULE 26)
Figure 9 (cont.)

351 gacctgctgtttacttcaagtttttcacggctaagatggagatggtggagatatt 400

351 gacctgctgtttacttcaagtttttcacggctaagatggagatggtggagatata 400

401 cctgtgcagaaatgggattacgttggcccctggttgctgtgcgtgat 450

401 cctgtgcagaaatgggattacgttggcccctggttgctgtgcgtgat 450

451 tggcacggaatattttgtcaggatggttcacaagaactctggcagatt 500

451 tggcacggaatattttgtcaggatggttcacaagaactctggcagatt 500

501 ctttggtgtaattttgcacgcaaatttttttctcaccataggtacgcacagttc 550

501 ctttggtgtaattttgcacgcaaatttttttctcaccataggtacgcacagttc 550

551 gaggctcagatgttggctattattaagctgcttctccttgtatgtgctct 600

551 gaggctcagatgttggctattattaagctgcttctccttgtatgtgctct 600

601 tatcaatatttttggatcctcaaatgagctctccttctctgagcctgctttg 650

601 tatcaatatttttggatcctcaaatgagctctccttctctgagcctgctttg 650

651 gtacggcttggaaacgacagtctgtgcatgtccttcatactatatatcttcac 700

651 gtacggcttggaaacgacagtctgtgcatgtccttcatactatatatcttcac 700

SUBSTITUTE SHEET (RULE 26)
Figure 9 (cont.)

701 aatggcctttcaccacgagttccccagttagggctttattgcaaacgtccc 750

701 aatggcctatcaccacgagttccccagttagggctttattgcaaacgtccc 750

751 ttatttcacaagctggctgcagtcaccacaaatccacactcaggaataat 800

751 ttatttcacaagctggctgcagtcaccacaaatccacactcaggaataat 800

801 ttcaaggggtgaccatttggccccctttgaaaaccaggaatttgggaagaa 850

801 ttcaaggggtgaccatttggccccctttgaaaaccaggaatttgggaagaa 850

851 gtaagaggaggcactgaagaatgggagaggtgtatctcagctacaagctaa 900

851 gtaagaggaggcactgaagaatgggagaggtgtatctcagctacaagctaa 900

901 acgaaacgtacaatttacattctctacTGAatcaactttttacattttataggtttt 950

901 acgaaacgtacaatttacattctctacTGAatcaactttttacattttataggtttt 950

951 agtttatcgcgtgtaa.caagtccacacattgtgtgtgtgatgtatatttcaaa 999

951 agtttatcgcgtgtaa.caagtccacacattgtgtgtgatgtatatttcaaa 999

1000 agttaccatatacttttttagaataaaaattgtataggtcgcggag 1049

1001 agttaccatatacttttttagaataaaaattgtataggtcgcggag 1050
Figure 9 (cont.)

1050 ttacggttacaaaggccaaatctattgttgggaattccattatta 1099
     ||||   ||||    ||||                ||||
1051 atactggttac...........actattcgttgtgggaattccattataaaaa 1091

1100 taaaaattaggtttgtagtttttatctgtgatcaatatataatatatat 1149
     ||||   ||
1092 ataaaaaaaaaaaaaaaaaaaaaaa
Figure 10: Gap of SEQ ID NO: 2 and SEQ ID NO: 4

1 AISVFSTSYSFHKNLLLSQDIILRPCLFSSLISPVVEPMKRKKKTHRAAC 50
       |||||  ||||:||||  || || || ||: ||||:|||: |||
1 AISVFSSSYSPYKNLNSHSPKLFTFSLFSPVIMSPMRRKKKHGDPC 50

51 ICSCVAERTNLDIPQIEEEEENEELIEOETDSGIIHIIKTHLGGQSSRT 100
       |||||  ||||:||||  || || || ||: ||||:|||: |||
51 ICSCVAGRTNLDIPQIEEEEENEREELIEOETDSDIVHIIKTHLGGQSKRPT 100

101 GSIVAPVSLGILSMIGPAVYFKFSLMEDCPVLAEMGITFAAFVAAAI 150
       |||||  ||||:||||  || || || ||: ||||:|||: |||
101 GSIVAPVSLGILSMIGPAVYFKFSLMEDCIPVLAEMGFATFVAAAV 150

151 GTEFLSGWVHELWHELWYIHKHSHHRSKGRFEFNDVFAINALPAIAL 200
       |||||  ||||:||||  || || || ||: ||||:|||: |||
151 GTEFLSAWVHELWHELWYIHKHSHHRSKGRFEFNDVFAINALPAIAL 200

201 INYGFSNGLLLPLACFGTGLTTCMAYFLHNLHSRFPVGLIANVP 250
       |||||  ||||:||||  || || || ||: ||||:|||: |||
201 INYGFSNGLLPLACFGVGLTTCMAYFLHNLHSRFPVWLIANVP 250

251 YFHKLAAAHQIHHSGKFQGVFPDGLGFQGELEEERGGEELERTSRTAK 300
       |||||  ||||:||||  || || || ||: ||||:|||: |||
251 YFHKLAAAHQIHHSGKFQGVFPDGLGFPELEEERGGEELERTSTRTCK 300

301 RTQST* 307
      |||||
301 RTQPST* 307

SUBSTITUTE SHEET (RULE 26)
Figure 11: Gap of SEQ ID NO: 2 and Arabidopsis β-carotene hydroxylase (SEQ ID NO: 10)

1 AISVFSTSYS FHKNLLHSHKQ DI LNRPCLLP VVESPMPRKKTHRAAC 50
  | | | | | | | | : | | |
1 MAA XL STAVTFKP...L HRSFSSSSTDFRLRPKL S GFS PS PSLRFKRSV 47
  .
51 IC SvAER TRNLDIPQIEEEEENNEEIIEQTDSGIIH KTLG KPSRRS 100
  | | | | | | | | : : | | | |
48 CVVVEERRQNSPIENDERPE STSSTN A IDAEYLRLAELKLERKKSERST 97
  .
101 GSIVAPVSC LG ILSMI GPM AYVF KFS RLMECGDIPV AEMGITFAFVAAAI 150
  | | | | | | | | : | : | | | |
98 YLIAAMLSSFG ITSMA VMAVYRF SWQMEG GEISML E MGF TFA ISVGAAV 147
  .
151 GTEFLSGW VHELW LSYIHKSHH RSRKGF EFNDVFAIINALPAIAL 200
  | | | | | | | | : | | | | : | | |
148 GMEFWARWAHRLWA SLWMH ES HHPREGPFELNDVFA I VNAGPAI GL 197
  .
201 INYGFSNEG LLPGACFGTGTLGTVCGMA YIFLHNGLSH RRFPVGLI ANVP 250
  : | | | | | | | | | | | | | |
198 LSYGFFNKGLVPGLC FAGAGLTVPGIA MYMVHDGLVHKRFVPVGPIADVP 247
  .
251 YFHKL AAAH IQI IHS GKF QGVFPVFGLPQ ELEEPV RGGTEELERVISRTAK 300
  | | | | | | | | | | | | | |
248 YLRKVA A AHQI IHS H TDKFNGVP YGLFLGPKLEEVE.GGNEELDKEISRRIK 296
  .
301 RTQ SST*....... 307
  .
297 SYKKAGSGSSSSS 311
Figure 12A (SEQ ID NO: 11)

1 CATACCATAA ATAGTAGAGG ACAACCTACA AACCAACCAC CAGAAACCTC 50
51 CAATGGGCAGC

Figure 12B (SEQ ID NO: 12)

MAAAASVFSSGYSFYKQLLLDSKPNILKPPCLLSFSPVIMSMPMKKKKHGDPCICSVAGR
TRNLQPQEEEEEENVEELBQTSEDIVHXTGKQSKRPTGSIVAPVSCGLGIISMIG
PAVYFKFSRLMECSDIPIVAEMGITTFATFVAAAVGTEFLSARVHKEIMHAVELMVYTHKSHHR
SRKGRFEFNDVFAAIALPAIALNYGFSNEGGLLGPGACFGVGLTTGCMAYIFLNLGALS
HRRFPWVLIANVPHYHKLAIAAHQHHSKGKFGVPGVFLGLGPKELEEEIVGRGTEELERVISR
TTRKQPST*
SEQUENCE LISTING

<110> CUNNINGHAM, Francis X.

<120> CAROTENOID KETOLASE GENES AND GENE PRODUCTS, PRODUCTION OF KETO CAROTENOIDs AND METHODS OF MODIFYING CAROTENOIDs USING THE GENES

<130> 8172-9022

<140> Unknown
<141> 1999-05-21

<150> 60/086,460
<151> 1998-05-22

<160> 18

<170> PatentIn Ver. 2.0

<210> 1
<211> DNA
<213> Adonis aestivalis

<400> 1
agcaatctca gtgtctcagta caagtatttc tttcccacaaag aatctcttgtg tcaccccaaa 60
acaagacatt ctcacccgccc cattgtggtgt ctttctccttta gttactgcttga agtgcctat 120
gagaaagaaa aagacacactc tgcctgcatag tattgcggcttgc gcgcaagagaa gaccaagga 180
cctgtatatt ctttacaattag agagaaagaa gaagaacagag gaagaacttacactaa tagaagcac 240
ggattcgcc ataattcata taaaagaaaaa gctagggggtt aacaatcaca gaggtccttcaag 300
tgctctccatt gtgcacccagc tatcttgctcg ctgtgctact tcacgtatacg cacctggtct 360
ttacttcagtg ttttacacgc ttaatacctgtagc ggttagatatt cctgtgccgac aatggaatgt 420
tacgtttggcc gctctttggtcg tgcctgcatag tggacacgggaa ttttgtctcg gatggttttca 480
caaagaacctc tggcagcttc ttttgcgttcg ttcacaccaaatg ctctttccctaatg gttcagaa 540
tagccgcttctg aagttgttgcc tattataacg gcgttccgctct ctattgccttt 600
tatctaatatt gggatctctaa atgaaagccta ccttctctgg aagcggatcct gcctgtctct 660
tggaacgca tgcctgtgcaa tgtgctactat tttttttctcaagttaaagcctttacctg cacacccgga 720
gtttggcttggta ggcctttgatt taaaagagcctttcttaattctcctaac aagctctggtcg cacgtcactaa 780
aatctcacttc tcagaaattg ttcatgcgggtga accatgtgctga cgtgcttttgg gacccggagga 840
atgtgagggct ctaactgagg cattgcagaatgttgaagagg ctgctgcctgc gcgtactgtaa 900
tgcaacgcaaatctcctctgtt cattggcttcg tgtcaatcagtttggctctttt caaatcttgcc 960
tggtaatattg cacacattctgtgcttgtagt gtaatcttcagcttatttttattag 1020
aatatatatatat tcatgttatagtgcctgctgcatag tgggttacccactaggcaaatctatgttgg 1080
gtattccatcattaaaaaacatgattagatttctttctgtaattcata gttttaattacttattgctttgctt 1140
aatatatatatatatattta aaaaagaaaaa aaaaaa aaaaaa

<210> 2
<211> 306
<212> PRT
<213> Adonis aestivalis

<400> 2
Ala Ile Ser Val Phe Ser Thr Ser Tyr Ser Phe His Lys Asn Leu Leu
1      5        10        15
Leu His Ser Lys Gln Asp Ile Leu Asn Arg Pro Cys Leu Leu Phe Ser
20     25       30
Pro Val Val Val Glu Ser Pro Met Arg Lys Lys Thr His Arg Ala
35     40       45
Ala Cys Ile Cys Ser Val Ala Glu Arg Thr Arg Asn Leu Asp Ile Pro
50     55       60
Gln Ile Glu Glu Glu Glu Asn Glu Glu Leu Ile Glu Glu Gln Thr
65     70       75       80
Asp Ser Gly Ile Ile His Ile Lys Thr Leu Gly Gly Lys Gln Ser
85     90       95
Arg Arg Ser Thr Gly Ser Ile Val Ala Pro Val Ser Cys Leu Gly Ile
100    105      110
Leu Ser Met Ile Gly Pro Ala Val Tyr Phe Lys Phe Ser Arg Leu Met
115    120      125
Glu Cys Gly Asp Ile Pro Val Ala Glu Met Gly Ile Thr Phe Ala Ala
130    135      140
Phe Val Ala Ala Ile Gly Thr Glu Phe Leu Ser Gly Trp Val His
145    150      155      160
Lys Glu Leu Trp His Asp Ser Leu Trp Tyr Ile His Lys Ser His His
165    170      175
Arg Ser Arg Lys Gly Arg Phe Glu Phe Asn Asp Val Phe Ala Ile Ile
180    185      190
Asn Ala Leu Pro Ala Ile Ala Leu Ile Asn Tyr Gly Phe Ser Asn Glu
195    200      205
Gly Leu Leu Pro Gly Ala Cys Phe Gly Thr Gly Leu Gly Thr Thr Val
210    215      220
Cys Gly Met Ala Tyr Ile Phe Leu His Asn Gly Leu Ser His Arg Arg
2
Phe  Pro  Val  Gly  Leu  Ile  Ala  Asn  Val  Pro  Tyr  Phe  His  Lys  Leu  Ala  245
Ala  Ala  His  Gln  Ile  His  His  Ser  Gly  Lys  Phe  Gln  Gly  Val  Pro  Phe  260
Gly  Leu  Phe  Leu  Gly  Pro  Gln  Glu  Leu  Glu  Glu  Val  Arg  Gly  Gly  Thr  275
Glu  Glu  Leu  Glu  Arg  Val  Ile  Ser  Arg  Thr  Ala  Lys  Arg  Thr  Gin  Ser  290
Ser  Thr  305

<210>  3
<211>  1112
<212>  DNA
<213>  Adonis aestivalis

<400>  3
agcaatttca  gtgttcagt  caggttatcc  ttcttacaag  aatctcttctg  tggactcaaa  60
accaaatatt  ctcaaacccc  catgctgtct  attctttcca  gtgtgtatca  tggctgcatat  120
gagaaagaa  aagaaacagt  gtgatccatg  tatctgctcc  gttgcaaggg  gaacaagqaa  180
ccttgatatt  ccctcaattg  aagaaagaga  agaaatgtgg  gaagaactaa  tagaaacagc  240
cgattctgac  atatgcata  taaaaaaca  actagggggg  aaacaatcaaa  aacggccccac  300
tgcttcctatt  gtgcaacccc  tatactgctt  tggagatctt  ccaatgtatt  gacatgtgct  360
tttactcaag  ttttcaacgc  taatggaggg  tggagataa  cctgtagcag  aatatggggat  420
tactgtttgcc  accctttgtg  ctgctgctgt  tggcagccag  ttttggtcag  caggtgttca  480
caaagaacac  tggcaacagct  ctctttgctga  catccaaagc  tttcaccactc  ggttcagaaa  540
agggcgccttc  gagaattgtg  atgtgtttgc  tattattacg  gcgtcctccg  ctattgtctt  600
tatcaatatt  ggtttcctca  atgaagcctc  cctctcttgg  gcgtctttgg  tttgctgctt  660
tggaaaccaaa  gcgtgttgta  tggcattttc  ttttctctacc  aatgggctat  cacaaccgaag  720
gttccagcta  tgggcttttg  cgaagcctccc  ttatttcacc  aagcttggtgg  cagctcaccg  780
aatccaccc  tcaagaaaaat  ttgcggtggtg  attccattttg  gccaccagaga  840
attggaagaa  gtatgaggg  gcactgaaga  gttggagagg  gtaatcagttc  gtacaacttac  900
acgaacgcga  ccattctcttt  gatctaaat  ttctctataat  ataaggttttt  agtttacgct  960
tgtatatataa  ttctccatcc  gatctgtttt  agtaatcgaag  acgttcagttt  cttctcttct  1020
agaatatattt  tttgatgtata  ggtgcccggt  atacagttaacc  actatctgttg  tgggaatccc  1080
attatatataaa  aataaaaaaa  aaaaaaa  aa

<210>  4
<211>  306
<212>  PRT
<213>  Adonis aestivalis

3
<400> 4
Ala Ile Ser Val Phe Ser Ser Gly Tyr Ser Phe Tyr Lys Asn Leu Leu
  1  5  10  15
Leu Asp Ser Lys Pro Asn Ile Leu Lys Pro Pro Cys Leu Leu Phe Ser
 20  25  30
Pro Val Val Ile Met Ser Pro Met Arg Lys Lys Lys His Gly Asp
 35  40  45
Pro Cys Ile Cys Ser Val Ala Gly Arg Thr Arg Asn Leu Asp Ile Pro
 50  55  60
Gln Ile Glu Glu Glu Glu Glu Asn Val Glu Glu Leu Ile Glu Gin Thr
 65  70  75  80
Asp Ser Asp Ile Val His Ile Lys Thr Leu Gly Gly Lys Gin Ser
 85  90  95
Lys Arg Pro Thr Gly Ser Ile Val Ala Pro Val Ser Cys Leu Gly Ile
100 105 110
Leu Ser Met Ile Gly Pro Ala Val Tyr Phe Lys Phe Ser Arg Leu Met
115 120 125
Glu Gly Gly Asp Ile Pro Val Ala Glu Met Gly Ile Thr Phe Ala Thr
130 135 140
Phe Val Ala Ala Ala Val Gly Thr Glu Phe Leu Ser Ala Trp Val His
145 150 155 160
Lys Glu Leu Trp His Glu Ser Leu Trp Tyr Ile His Lys Ser His His
165 170 175
Arg Ser Arg Lys Gly Arg Phe Glu Phe Asn Asp Val Phe Ala Ile Ile
180 185 190
Asn Ala Leu Pro Ala Ile Ala Leu Ile Tyr Gly Phe Ser Asn Glu
195 200 205
Gly Leu Leu Pro Gly Ala Cys Phe Gly Val Gly Leu Gly Thr Thr Val
210 215 220
Cys Gly Met Ala Tyr Ile Phe Leu His Asn Gly Leu Ser His Arg Arg
225 230 235 240
Phe Pro Val Trp Leu Ile Ala Asn Val Pro Tyr Phe His Lys Leu Ala

WO 99/61652

PCT/US99/10455

245

Ala Ala His Gln Ile His His Ser Gly Lys Phe Gln Gly Val Pro Phe
260

Gly Leu Phe Leu Gly Pro Lys Glu Leu Glu Glu Val Arg Gly Gly Thr
275

Glu Glu Leu Glu Arg Val Ile Ser Arg Thr Thr Lys Arg Thr Gln Pro
290

Ser Thr

305

<210> 5
<211> 1205
<212> DNA
<213> Adonis aestivalis

<400> 5
gggctgcagg aatccggcac gagagcaact tcagtttcca gtacaaagtta ttctttccac 60
aagaactctct tgtgctaccc aaaaacagac attctcaacc gcgccagtgtt gctttctctct 120
ccaagtgtgg tggagttgcgc tatgagaaag aaaaaagcac atctgtgctgc atgtatctgc 180
tctgtgctcag agagacacag gaacctgtgat attctccaaa ttgaaagaga ggaagagaa 240
gagagataac taataagaca cacggtattct ggcataattc atataaaaag aaccgctagg 300
ggggagcaat cagagaggct gcctgtgcctc atctgcgcaac cggatagtct ttttggtgtg 360
cctctacctg agttctacctc tgtttacttc aagttctttcag ggttagtgg ggtagtggat 420
attctgtgtg cagaaatggg gattagtttt gcggctttttg tttgctggctgc gattggcagct 480
gaatatttggt caggatgggc tcacaaagaa acacaggcagc attcttttttg gttggtcgac 540
aagctcacc atagctacag aaaaaagcgcgc ctctaggtttca atgtatggtt gttgtttattt 600
aaccggtttct cttgtatttt gttttatattt cattacaatt tattgattttt ctaaatgaggg cctcttttctct 660
gagcgtgttttc cttgtagtcgg cagatgtgtg gcagttgtgta cattttttcattttttttc 720
cacatgtgcc ttctccacatat aaatgctttta tttgccacagtt ccttttattttc 780
cacacagctgg ctgagctcct cacaaatctac cattcagaaa aatcagggg taggtgat 840
gggctgtttgc tgtgacccca gaaatgggaa gaagtaagag gaggacactga atggtttgag 900
agggtgatcc tgtgtaacat taaaggaagc caatcactca catgaatccaa ctttttttaca 960
atatggtgcc ttccatatct ccggttgctac atacctagcgt tggatgtatcc 1020
aaatctccac tattcttttt tgtttagatga taggtgccag atgtaccagttt 1080
acaaagccta aacttatggt gttgagatcc cattataaaa aataaataatgc agaatattgta 1140
gttcttctcg gttgatcaata tcaatataa ttatataag caaaaaaa aaaaaa 1200
tcag

<210> 6
<211> 315
<212> PRT
<213> Adonis aestivalis
Met Gly Leu Gln Glu Phe Gly Thr Arg Ala Ile Ser Val Phe Ser Thr
  1   5    10   15

Ser Tyr Ser Phe His Lys Asn Leu Leu Leu His Ser Lys Gln Asp Ile
  20  25    30

Leu Asn Arg Pro Cys Leu Leu Phe Ser Pro Val Val Val Glu Ser Pro
  35  40    45

Met Arg Lys Lys Thr His Arg Ala Ala Cys Ile Cys Ser Val Ala
  50  55    60

Glu Arg Thr Arg Asn Leu Asp Ile Pro Gln Ile Glu Glu Glu Glu Glu
  65  70    75   80

Asn Glu Glu Leu Ile Glu Gln Thr Asp Ser Gly Ile Ile His Ile
  85  90    95

Lys Lys Thr Leu Gly Gly Lys Gln Ser Arg Arg Ser Thr Gly Ser Ile
 100 105   110

Val Ala Pro Val Ser Cys Leu Gly Ile Leu Ser Met Ile Gly Pro Ala
 115 120   125

Val Tyr Phe Lys Phe Ser Arg Leu Met Glu Cys Gly Asp Ile Pro Val
 130 135   140

 Ala Glu Met Gly Ile Thr Phe Ala Ala Phe Val Ala Ala Ala Ile Gly
 145 150   155   160

Thr Glu Phe Leu Ser Gly Trp Val His Lys Glu Leu Trp His Asp Ser
 165 170   175

Leu Trp Tyr Ile His Lys Ser His His Arg Ser Arg Lys Gly Arg Phe
 180 185   190

Glu Phe Asn Asp Val Phe Ala Ile Asn Ala Leu Pro Ala Ile Ala
 195 200   205

Leu Ile Asn Tyr Gly Phe Ser Asn Glu Gly Leu Leu Pro Gly Ala Cys
 210 215   220

Phe Gly Thr Gly Leu Gly Thr Thr Val Cys Gly Met Ala Tyr Ile Phe
 225 230   235   240

Leu His Asn Gly Leu Ser His Arg Arg Phe Pro Val Gly Leu Ile Ala
 245 250   255
Asn Val Pro Tyr Phe His Lys Leu Ala Ala Ala His Gln Ile His His
260 265 270

Ser Gly Lys Phe Gln Gly Val Pro Phe Gly Leu Phe Leu Gly Pro Gln
275 280 285

Glu Leu Glu Glu Val Arg Gly Gly Thr Glu Glu Leu Glu Arg Val Ile
290 295 300

Ser Arg Thr Ala Lys Arg Thr Gln Ser Ser Ser Thr
305 310 315

<210> 7
<211> 1141
<212> DNA
<213> Adonis aestivalis

<400> 7
gggctgcaag aattcggccac gagagcaatt tcaggttctca gttcaggtta ttctttctac 60
aaagatcctc tggtggacct aaacccaaat attctctaaac ccccatgtgct gtatatctctctct 120
ccgaggacac tctgtgctgg tctgagaaag aaaaagaaaa atgggtatc cctgtatgcg 180
tccgttgcag ggagaaacaag gaaaccttgatt tgcctctcaca ttgagaagaga ggaagagaaat 240
gtgggaagatc taataagacaa gacgccattc gacatagtgc atataaaagaa aacactaggg 300
gggaaacact ccaaacgGCC cactggctcc atggtcgCAC cccgtatcctg ttctttgaccc 360
ccttcactgtg ttggacacctg cttttacactc aaggttttcac ggttagattga ggttagagat 420
attactgtg tagaataatggt gattacctttg gcccaccttgg ttgctgtgctg tgtggacacc 480
gagttttttgt cacgcttgggt tcataaagaa ctccccgctgtt agtcttctgtgt gtagattcactac 540
aagtctcacc atccggtcag cccgatgtgctc ccctatatga ctataatgtgcg caagagccag 600
aaccgctttt ccgtatagtc ccctaatcgc ctgagatgtgc ctatggtttcg gccctctctct 660
gggactgcttg ttggtgctgg tgtggacaaa ccacgctgtggt ggtggtgtct cattttttcttt 720
ccacactgcg ttagcactca agggtactcc cttgatgttctc cccctatttc ggcctttttg 780
cacaaagcttg ctgctgctgca cccaatcaca cacatcgagaa ctttcggcag tgtacatttt 840
ggcctgttcc cggagcaccag cggaaattggaa gaaggtaaag cggacactga agagttgagc 900
aggaggtactgc gtcgctacac taacaggaac ccacactcttctgagac cccctttttttc 960	tatataaatgt tgtatttttt cgggtgtaat aaatccacaca cttgtgctgtc tttttttat 1020
catttaaatg atctctctct cttcagataat tttttttatgt ataggtgctgg gttatatcattgt 1080
tacaccttc gtttgtggttt cccattattaa aaaaataaaaa aaaaaaaaaa aaaaactctga 1140

<210> 8
<211> 315
<212> PRT
<213> Adonis aestivalis

<400> 8
Met Gly Leu Gln Glu Phe Gly Thr Arg Ala Ile Ser Val Phe Ser Ser
<table>
<thead>
<tr>
<th>Gly</th>
<th>Tyr</th>
<th>Ser</th>
<th>Phe</th>
<th>Tyr</th>
<th>Lys</th>
<th>Asn</th>
<th>Leu</th>
<th>Leu</th>
<th>Leu</th>
<th>Leu</th>
<th>Asp</th>
<th>Ser</th>
<th>Lys</th>
<th>Pro</th>
<th>Asn</th>
<th>Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Lys</td>
<td>Pro</td>
<td>Pro</td>
<td>Cys</td>
<td>Leu</td>
<td>Leu</td>
<td>Phe</td>
<td>Ser</td>
<td>Pro</td>
<td>Val</td>
<td>Val</td>
<td>Ile</td>
<td>Met</td>
<td>Ser</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Arg</td>
<td>Lys</td>
<td>Lys</td>
<td>Lys</td>
<td>His</td>
<td>Gly</td>
<td>Asp</td>
<td>Pro</td>
<td>Cys</td>
<td>Ile</td>
<td>Cys</td>
<td>Ser</td>
<td>Val</td>
<td>Ala</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Arg</td>
<td>Thr</td>
<td>Arg</td>
<td>Asn</td>
<td>Leu</td>
<td>Asp</td>
<td>Ile</td>
<td>Pro</td>
<td>Gln</td>
<td>Ile</td>
<td>Glu</td>
<td>Glu</td>
<td>Glu</td>
<td>Glu</td>
<td>Glu</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Val</td>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
<td>Ile</td>
<td>Glu</td>
<td>Gln</td>
<td>Thr</td>
<td>Asp</td>
<td>Ser</td>
<td>Asp</td>
<td>Ile</td>
<td>Val</td>
<td>His</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Lys</td>
<td>Thr</td>
<td>Leu</td>
<td>Gly</td>
<td>Gly</td>
<td>Lys</td>
<td>Gly</td>
<td>Lys</td>
<td>Gln</td>
<td>Ser</td>
<td>Lys</td>
<td>Arg</td>
<td>Pro</td>
<td>Thr</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Pro</td>
<td>Val</td>
<td>Ser</td>
<td>Cys</td>
<td>Leu</td>
<td>Gly</td>
<td>Ile</td>
<td>Leu</td>
<td>Ser</td>
<td>Met</td>
<td>Ile</td>
<td>Gly</td>
<td>Pro</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Tyr</td>
<td>Phe</td>
<td>Lys</td>
<td>Phe</td>
<td>Ser</td>
<td>Arg</td>
<td>Leu</td>
<td>Met</td>
<td>Glu</td>
<td>Gly</td>
<td>Gly</td>
<td>Asp</td>
<td>Ile</td>
<td>Pro</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Glu</td>
<td>Met</td>
<td>Gly</td>
<td>Ile</td>
<td>Thr</td>
<td>Phe</td>
<td>Ala</td>
<td>Thr</td>
<td>Phe</td>
<td>Val</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Val</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Glu</td>
<td>Phe</td>
<td>Leu</td>
<td>Ser</td>
<td>Ala</td>
<td>Trp</td>
<td>Val</td>
<td>His</td>
<td>Lys</td>
<td>Glu</td>
<td>Leu</td>
<td>Trp</td>
<td>His</td>
<td>Glu</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Trp</td>
<td>Tyr</td>
<td>Ile</td>
<td>His</td>
<td>Lys</td>
<td>Ser</td>
<td>His</td>
<td>His</td>
<td>Arg</td>
<td>Ser</td>
<td>Arg</td>
<td>Lys</td>
<td>Gly</td>
<td>Arg</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Phe</td>
<td>Asn</td>
<td>Val</td>
<td>Phe</td>
<td>Ala</td>
<td>Ile</td>
<td>Asn</td>
<td>Ala</td>
<td>Leu</td>
<td>Pro</td>
<td>Ala</td>
<td>Ile</td>
<td>Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ile</td>
<td>Asn</td>
<td>Tyr</td>
<td>Gly</td>
<td>Phe</td>
<td>Ser</td>
<td>Asn</td>
<td>Glu</td>
<td>Gly</td>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
<td>Gly</td>
<td>Ala</td>
<td>Cys</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Gly</td>
<td>Val</td>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
<td>Cys</td>
<td>Gly</td>
<td>Met</td>
<td>Ala</td>
<td>Tyr</td>
<td>Ile</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>His</td>
<td>Asn</td>
<td>Gly</td>
<td>Leu</td>
<td>Ser</td>
<td>His</td>
<td>Arg</td>
<td>Arg</td>
<td>Phe</td>
<td>Pro</td>
<td>Val</td>
<td>Trp</td>
<td>Leu</td>
<td>Ile</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Val</td>
<td>Pro</td>
<td>Tyr</td>
<td>Phe</td>
<td>His</td>
<td>Lys</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>His</td>
<td>Gln</td>
<td>Ile</td>
<td>His</td>
<td>8</td>
</tr>
</tbody>
</table>
260  Ser Gly Lys Phe Gln Gly Val Pro Phe Gly Leu Phe Leu Gly Pro Lys
265
270
275
280
285

Glu Leu Glu Glu Val Arg Gly Gly Thr Glu Glu Leu Glu Arg Val Ile
290
295
300

Ser Arg Thr Thr Lys Arg Thr Gln Pro Ser Thr
305
310
315

<210> 9
<211> 1149
<212> DNA
<213> Adonis aestivalis

<400> 9
agcaatctca gtttcaagga caagttattc ttacctcaacag aatctctttgt tgcacctcaaa 60
aacaagcatc tctcaacgcc catgtttttct cttctctccaa gtttgggttgg agtgctcctat 120
gagaagaaa aagacacact gttgctgtcat tattctgtct gttgcaagaga gaacagagaa 180
cctgctatttt cctcaaaatttg aagaagagag agagaacagag gaagaacactaa tagaaacagac 240
gatctgtgac ataattcata taagagaaac gctagggggg aacaatcagac gacggtccac 300
tgctctcatt gtgcacccggc ttttctgctct tggatctcct tcaatgattc gacctgtctgt 360
ttatctcaag ttctcagggc taatggagtg ttgagatttatt cttgtgcacag aataatgggat 420
tacgtttgcc gcctttgttg ctgctgctgat tggcaacggaa ttttggctcag gtgggttctta 480
caaagactc tgagcaagttt ttttgtgtgt cattcacaag ccctcaacacaa ggtcagaaa 540
agggcctttc gatgtccagat atgctttttgc tattatatc ggcctctctct cttgtgtct 600
tatcaatttat ggttttctcaaa atgaaggcctt cttctctgga gctgcttttg gtaccgggtct 660
tggaacacga gttgtgtaga ttgtcttcat ttttctccac aatggcctttt caccacgagaaa 720
gttccccactg gggcatttttg caaagctcgc cttatttttccc aagcttggctg cagctcaccac 780
aatccatctc tcaagaaaaat ttcaggggtc accattttggc cttgcttttg gaccccaagag 840
attggagaggt gtaagagggg gcactgaaga atggagagag gttcgcaagtc gcagctgcta 900
acgaaaccacaa tcatcttcatg ctaactactc cttactaattt atgagttcttg agttttctgg 960
tgtttacataat cacacatttttg tttcggtgtaa gtaatactaa gttacccatat ccctttcttg 1020
aatatatatat gtttgtgattgt gcggctgtgt cttgctttctc aagcccaactt ctatgcttgtt 1080
ggaattcccc aatagaaaatg gttgtgtatttt tacatccttt aataagtaa 1140
atatatatatt

<210> 10
<211> 310
<212> PRT
<213> Arabidopsis

<400> 10
Met Ala Ala Xaa Leu Ser Thr Ala Val Thr Phe Lys Pro Leu His Arg
1  5  10  15

9
<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser</td>
<td>Ser Ser Ser Ser Ser Thr Asp Phe Arg Leu Arg Leu Pro Lys Ser</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser Gly Phe Ser Pro Ser Leu Arg Phe Lys Arg Phe Ser Val Cys</td>
</tr>
<tr>
<td></td>
<td>35</td>
</tr>
<tr>
<td>Tyr</td>
<td>Val Val Glu Glu Arg Arg Gln Asn Ser Pro Ile Glu Asn Asp Glu</td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Arg</td>
<td>Pro Glu Ser Thr Ser Ser Thr Asn Ala Ile Asp Ala Glu Tyr Leu</td>
</tr>
<tr>
<td></td>
<td>65</td>
</tr>
<tr>
<td>Ala</td>
<td>Leu Arg Leu Ala Glu Lys Leu Glu Arg Lys Lys Ser Glu Arg Ser</td>
</tr>
<tr>
<td></td>
<td>85</td>
</tr>
<tr>
<td>Thr</td>
<td>Tyr Leu Ile Ala Ala Met Leu Ser Ser Phe Gly Ile Thr Ser Met</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Ala</td>
<td>Val Met Ala Val Tyr Tyr Arg Phe Ser Trp Gln Met Glu Gly Gly</td>
</tr>
<tr>
<td></td>
<td>115</td>
</tr>
<tr>
<td>Glu</td>
<td>Ile Ser Met Leu Glu Met Phe Gly Thr Phe Ala Leu Ser Val Gly</td>
</tr>
<tr>
<td></td>
<td>130</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala Val Gly Met Glu Phe Trp Ala Arg Trp Ala His Arg Ala Leu</td>
</tr>
<tr>
<td></td>
<td>145</td>
</tr>
<tr>
<td>Trp</td>
<td>His Ala Ser Leu Trp Asn Met His Glu Ser His His Lys Pro Arg</td>
</tr>
<tr>
<td></td>
<td>165</td>
</tr>
<tr>
<td>Glu</td>
<td>Gly Pro Phe Glu Leu Asn Asp Val Phe Ala Ile Val Asn Ala Gly</td>
</tr>
<tr>
<td></td>
<td>180</td>
</tr>
<tr>
<td>Pro</td>
<td>Ala Ile Gly Leu Leu Ser Tyr Gly Phe Phe Asn Lys Gly Leu Val</td>
</tr>
<tr>
<td></td>
<td>195</td>
</tr>
<tr>
<td>Pro</td>
<td>Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Ile</td>
</tr>
<tr>
<td></td>
<td>210</td>
</tr>
<tr>
<td>Ala</td>
<td>Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val</td>
</tr>
<tr>
<td></td>
<td>225</td>
</tr>
<tr>
<td>Gly</td>
<td>Pro Ile Ala Asp Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His</td>
</tr>
<tr>
<td></td>
<td>245</td>
</tr>
<tr>
<td>Gln</td>
<td>Leu His His Thr Asp Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe</td>
</tr>
<tr>
<td></td>
<td>260</td>
</tr>
</tbody>
</table>

10
Leu Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Asn Glu Glu Leu Asp
275  280  285

Lys Glu Ile Ser Arg Arg Ile Lys Ser Tyr Lys Lys Ala Ser Gly Ser
290  295  300

Gly Ser Ser Ser Ser Ser
305  310

<210> 11
<211> 60
<212> DNA
<213> Adonis aestivalis

<400> 11
cataccataa atagtagagg acaacctaca aaccaaccac cagaaacccg caatggcagc 60

<210> 12
<211> 309
<212> PRT
<213> Adonis aestivalis

<400> 12
Met Ala Ala Ala Ile Ser Val Phe Ser Ser Ser Gly Tyr Ser Phe Tyr Lys
1  5  10  15

Asn Leu Leu Leu Asp Ser Lys Pro Asn Ile Leu Lys Pro Pro Cys Leu
20  25  30

Leu Phe Ser Ser Val Val Ile Met Ser Pro Met Arg Lys Lys Lys Lys
35  40  45

His Gly Asp Pro Cys Ile Cys Ser Val Ala Gly Arg Thr Arg Asn Leu
50  55  60

Asp Ile Pro Gln Ile Glu Glu Glu Glu Glu Glu Asn Val Glu Glu Leu Ile
65  70  75  80

Glu Gln Thr Asp Ser Asp Ile Val His Ile Lys Thr Leu Gly Gly
85  90  95

Lys Gln Ser Lys Arg Pro Thr Gly Ser Ile Val Ala Pro Val Ser Cys
100  105  110

Leu Gly Ile Leu Ser Met Ile Gly Pro Ala Val Tyr Phe Lys Phe Ser
115  120  125

11
<table>
<thead>
<tr>
<th>130</th>
<th>Arg</th>
<th>Leu</th>
<th>Met</th>
<th>Gly</th>
<th>Gly</th>
<th>Asp</th>
<th>Ile</th>
<th>Pro</th>
<th>Val</th>
<th>Ala</th>
<th>Glu</th>
<th>Met</th>
<th>Gly</th>
<th>Ile</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>135</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>Phe</td>
<td>Ala</td>
<td>Thr</td>
<td>Phe</td>
<td>Val</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Val</td>
<td>Gly</td>
<td>Thr</td>
<td>Glu</td>
<td>Phe</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>155</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>Trp</td>
<td>Val</td>
<td>His</td>
<td>Lys</td>
<td>Glu</td>
<td>Leu</td>
<td>Trp</td>
<td>His</td>
<td>Glu</td>
<td>Ser</td>
<td>Leu</td>
<td>Trp</td>
<td>Tyr</td>
<td>Ile</td>
<td>His</td>
</tr>
<tr>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>Ser</td>
<td>His</td>
<td>His</td>
<td>Arg</td>
<td>Ser</td>
<td>Arg</td>
<td>Lys</td>
<td>Gly</td>
<td>Arg</td>
<td>Phe</td>
<td>Glu</td>
<td>Phe</td>
<td>Asn</td>
<td>Asp</td>
<td>Val</td>
</tr>
<tr>
<td>185</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>195</td>
<td>Ala</td>
<td>Ile</td>
<td>Ile</td>
<td>Asn</td>
<td>Ala</td>
<td>Leu</td>
<td>Pro</td>
<td>Ala</td>
<td>Ile</td>
<td>Ala</td>
<td>Leu</td>
<td>Ile</td>
<td>Asn</td>
<td>Tyr</td>
<td>Gly</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>Ser</td>
<td>Asn</td>
<td>Glu</td>
<td>Gly</td>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
<td>Gly</td>
<td>Ala</td>
<td>Cys</td>
<td>Phe</td>
<td>Gly</td>
<td>Val</td>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td>215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>225</td>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
<td>Cys</td>
<td>Gly</td>
<td>Met</td>
<td>Ala</td>
<td>Tyr</td>
<td>Ile</td>
<td>Phe</td>
<td>Leu</td>
<td>His</td>
<td>Asn</td>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td>230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>245</td>
<td>His</td>
<td>Arg</td>
<td>Arg</td>
<td>Phe</td>
<td>Pro</td>
<td>Val</td>
<td>Trp</td>
<td>Leu</td>
<td>Ile</td>
<td>Ala</td>
<td>Asn</td>
<td>Val</td>
<td>Pro</td>
<td>Tyr</td>
<td>Phe</td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>255</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>Lys</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>His</td>
<td>Gln</td>
<td>Ile</td>
<td>His</td>
<td>His</td>
<td>Ser</td>
<td>Gly</td>
<td>Lys</td>
<td>Phe</td>
<td>Gln</td>
</tr>
<tr>
<td>265</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>275</td>
<td>Val</td>
<td>Pro</td>
<td>Phe</td>
<td>Gly</td>
<td>Leu</td>
<td>Phe</td>
<td>Leu</td>
<td>Gly</td>
<td>Pro</td>
<td>Lys</td>
<td>Glu</td>
<td>Leu</td>
<td>Glu</td>
<td>Val</td>
<td>Arg</td>
</tr>
<tr>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>285</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td>Gly</td>
<td>Gly</td>
<td>Thr</td>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
<td>Glu</td>
<td>Arg</td>
<td>Val</td>
<td>Ile</td>
<td>Ser</td>
<td>Arg</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>295</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>305</td>
<td>Thr</td>
<td>Gln</td>
<td>Pro</td>
<td>Ser</td>
<td>Thr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|<210> | 13 |<211> | 310 |<212> | PRT |<213> | Arabidopsis |

|<400> | 13 | Met | Ala | Ala | Xaa | Leu | Ser | Thr | Ala | Val | Thr | Phe | Lys | Pro | Leu | His | Arg |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 5   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 10  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 15  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 20  | Ser | Phe | Ser | Ser | Ser | Thr | Asp | Phe | Arg | Leu | Arg | Leu | Pro | Lys | Ser |     |
| 25  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 30  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
Leu Ser Gly Phe Ser Pro Ser Leu Arg Phe Lys Arg Phe Ser Val Cys
35 40 45

Tyr Val Val Glu Glu Arg Arg Gln Ser Pro Val Glu Asn Asp Glu
50 55 60

Arg Pro Glu Ser Thr Ser Ser Thr Asn Ala Ile Asp Ala Glu Tyr Leu
65 70 75 80

Ala Leu Arg Ala Glu Lys Leu Glu Arg Lys Lys Ser Glu Arg Ser
85 90 95

Thr Tyr Leu Ile Ala Ala Met Leu Ser Ser Phe Gly Ile Thr Ser Met
100 105 110

Ala Val Met Ala Val Tyr Arg Phe Ser Trp Glu Met Glu Gly Gly
115 120 125

Glu Ile Ser Met Leu Glu Met Phe Gly Thr Phe Ala Leu Ser Val Gly
130 135 140

Ala Ala Val Gly Met Glu Phe Trp Ala Arg Trp Ala His Arg Ala Leu
145 150 155 160

Trp His Ala Ser Leu Trp Asn Met His Glu Ser His His Lys Pro Arg
165 170 175

Glu Gly Pro Phe Glu Leu Asn Asp Val Phe Ala Ile Val Asn Ala Gly
180 185 190

Pro Ala Ile Gly Leu Leu Ser Tyr Gly Phe Phe Asn Lys Gly Leu Val
195 200 205

Pro Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Ile
210 215 220

Ala Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val
225 230 235 240

Gly Pro Ile Ala Asp Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His
245 250 255

Gln Leu His His Thr Asp Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe
260 265 270

Leu Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Asn Glu Glu Leu Asp
275 280 285
Lys Glu Ile Ser Arg Arg Ile Lys Ser Tyr Lys Lys Ala Ser Gly Ser  
290 295 300  
Gly Ser Ser Ser Ser  
305 310  

<210> 14  
<211> 305  
<212> PRT  
<213> Arabidopsis  

<400> 14  
Met Ala Ala Gly Leu Ser Thr Ile Ala Val Thr Leu Lys Pro Leu Asn  
1 5 10 15  
Arg Ser Ser Phe Ser Ala Asn His Pro Ile Ser Thr Ala Val Phe Pro  
20 25 30  
Pro Ser Leu Arg Phe Asn Gly Phe Arg Arg Arg Lys Ile Leu Thr Val  
35 40 45  
Cys Phe Val Val Glu Glu Arg Lys Gln Ser Ser Pro Met Asp Asp Asp  
50 55 60  
Asn Lys Pro Glu Ser Thr Thr Ser Ser Glu Ile Leu Met Thr Ser  
65 70 75 80  
Arg Leu Leu Lys Lys Ala Glu Lys Lys Ser Glu Arg Phe Thr Tyr  
85 90 95  
Leu Ile Ala Ala Val Met Ser Ser Phe Gly Ile Thr Ser Met Ala Ile  
100 105 110  
Met Ala Val Tyr Tyr Arg Phe Ser Trp Gln Met Lys Gly Gly Glu Val  
115 120 125  
Ser Val Leu Glu Met Phe Gly Thr Phe Ala Leu Ser Val Gly Ala Ala  
130 135 140  
Val Val Gly Met Glu Phe Trp Ala Arg Trp Ala His Arg Ala Leu Trp  
145 150 155 160  
His Asp Ser Leu Trp Asn Met His Glu Ser His His Lys Pro Arg Glu  
165 170 175  
Gly Ala Phe Glu Leu Asn Asp Val Phe Ala Ile Thr Asn Ala Val Pro  
180 185 190  

14
Ala Ile Gly Leu Leu Tyr Tyr Gly Phe Leu Asn Lys Gly Leu Val Pro
180 185 190

Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Met Phe Gly Met Ala
195 200 205

Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly
210 215 220

Pro Ile Ala Asn Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His Gln
225 230 235 240

Leu His His Thr Asp Lys Phe Lys Gly Val Pro Tyr Gly Leu Phe Leu
245 250 255

Gly Pro Lys Gln Glu Val Glu Glu Val Gly Gly Lys Glu Glu Leu Glu
260 265 270

Lys Glu Ile Ser Arg Arg Ile Lys Leu Tyr Asn Lys Gly Ser Ser Thr
275 280 285

Ser
290 295 300

<210> 15
<211> 315
<212> FRT
<213> Capsicum annuum

<400> 15

Met Ala Ala Glu Ile Ser Ile Ser Ala Ser Ser Arg Ala Ile Cys Leu
1 5 10 15

Gln Arg Asn Pro Phe Pro Ala Pro Lys Tyr Phe Ala Thr Ala Pro Pro
20 25 30

Leu Leu Phe Phe Ser Pro Leu Thr Cys Asn Leu Asp Ala Ile Leu Arg
35 40 45

Ser Arg Arg Lys Pro Arg Leu Ala Ala Cys Phe Val Leu Lys Asp Asp
50 55 60

Lys Leu Tyr Thr Ala Gln Ser Gly Lys Glu Ser Asp Thr Glu Ala Ile
65 70 75 80
Gly Asp Glu Ile Glu Val Glu Thr Asn Glu Glu Lys Ser Leu Ala Val

85 90 95

Arg Leu Ala Glu Lys Phe Ala Arg Lys Lys Ser Glu Arg Phe Thr Tyr

100 105 110

Leu Val Ala Ala Val Met Ser Ser Leu Gly Ile Thr Ser Met Ala Val

115 120 125

Ile Ser Val Tyr Tyr Arg Phe Ser Trp Gln Met Glu Gly Gly Glu Met

130 135 140

Pro Phe Ser Glu Met Phe Cys Thr Phe Ala Leu Ala Phe Gly Ala Ala

145 150 155 160

Ile Gly Met Glu Tyr Trp Ala Arg Trp Ala His Arg Ala Leu Trp His

165 170 175

Ala Ser Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu Gly

180 185 190

Pro Phe Glu Leu Asn Asp Ile Phe Ala Ile Ile Asn Ala Val Pro Ala

195 200 205

Ile Ala Phe Phe Ser Phe Gly Phe Asn His Lys Gly Leu Ile Pro Gly

210 215 220

Ile Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala Tyr

225 230 235 240

Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly Pro

245 250 255

Ile Ala Lys Val Pro Tyr Phe Gln Arg Val Ala Ala Ala His Gln Leu

260 265 270

His His Ser Asp Lys Phe Asp Gly Val Pro Tyr Gly Leu Phe Leu Gly

275 280 285

Pro Lys Glu Leu Glu Glu Val Gly Val Ile Glu Glu Leu Glu Lys Glu

290 295 300

Val Asn Arg Arg Ile Lys Ser Leu Lys Arg Leu

305 310 315

<210> 16

<211> 316
Thr Thr Gly Arg Tyr His Tyr Gln Leu Val Trp Cys Gln Ile Ser Phe
  1  5    10  15
Ser Ser Thr Ser Arg Thr Ser Tyr Tyr Arg His Ser Pro Phe Leu Gly
  20  25  30
Pro Lys Pro Thr Pro Thr Pro Thr Pro Ser Val Tyr Pro Ile Thr Pro Phe
  35  40  45
Ser Pro Asn Leu Gly Ser Ile Leu Arg Cys Arg Arg Arg Arg Pro Ser Phe
  50  55  60
Thr Val Cys Phe Val Leu Glu Asp Asp Lys Phe Lys Thr Gln Phe Glu
  65  70  75  80
Ala Gly Glu Glu Asp Ile Glu Met Lys Ile Glu Glu Gln Ile Ser Ala
  85  90  95
Thr Arg Leu Ala Glu Lys Leu Ala Arg Lys Ser Glu Arg Phe Thr
 100 105 110
Tyr Leu Val Ala Ala Val Met Ser Ser Phe Gly Ile Thr Ser Met Ala
 115 120 125
Val Met Ala Val Tyr Tyr Arg Phe Tyr Trp Gln Met Glu Gly Gly Glu
 130 135 140
Val Pro Phe Ser Glu Met Phe Gly Thr Phe Ala Leu Ser Val Gly Ala
 145 150 155 160
Ala Val Gly Met Glu Phe Trp Ala Arg Trp Ala His Lys Ala Leu Trp
 165 170 175
His Ala Ser Leu Trp His Met His Glu Ser His His Lys Pro Arg Glu
 180 185 190
Gly Pro Phe Glu Leu Asn Asp Val Phe Ala Ile Ile Asn Ala Val Pro
 195 200 205
Ala Ile Ala Leu Leu Asp Tyr Gly Phe Phe His Lys Gly Leu Ile Pro
 210 215 220
Gly Leu Cys Phe Gly Ala Gly Leu Gly Leu Ile Thr Val Phe Gly Met Ala
 225 230 235 240
Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly
245 250 255
Pro Val Ala Asn Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His Ser
260 265 270
Leu His His Ser Glu Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe Leu
275 280 285
Gly Pro Lys Glu Leu Glu Glu Val Gly Glu Leu Glu Leu Glu Lys
290 295 300
Glu Val Asn Arg Arg Thr Arg Tyr Ile Lys Gly Ser
305 310 315

<210> 17
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 17
cagaatcggt ctgttctatt agttcttcc 29

<210> 18
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 18
caatgcgaggaaatatcaagggttccttaccttc 32
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12P 23/00, 7/26; C12N 9/02, 1/20, 15/00; C07H 21/04; C07K 14/00
US CL : 435/67, 148, 189, 252.3, 252.33, 320.1; 536/23.2, 23.6; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/67, 148, 189, 252.3, 252.33, 320.1; 536/23.2, 23.6; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 5,453,565 A (MAWSON) 26 September 1995, see abstract and claims.</td>
<td>1-20</td>
</tr>
<tr>
<td>A,E</td>
<td>US 5,910,433 A (KAJIWARA et al.) 08 June 1999, see the entire patent.</td>
<td>1-20</td>
</tr>
</tbody>
</table>

☐ Further documents are listed in the continuation of Box C.  ☐ See patent family annex.

Date of the actual completion of the international search: 13 AUGUST 1999

Date of mailing of the international search report: 29 OCT 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized Officer
TEKCHAND SAIDHA
Telephone No. (703) 308-0196

* Special categories of cited documents:
  "A" document defining the general state of the art which is not considered to be of particular relevance
  "E" earlier document published on or after the international filing date
  "L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  "O" document referring to an oral disclosure, use, exhibition or other means
  "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"A" document member of the same patent family

Form PCT/ISA/210 (second sheet)(July 1992)
B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN Files: Medline, Caplus, Biosis, Agricola, Embase & Scisearch. Search terms used: beta carotene and ketolase, ketocarotenoid, Adonis aestivalis, carotenoid biosynthesis, gene? or dna or rna or nucleic acid? in various permutations and combinations.