

(22) Date de dépôt/Filing Date: 2000/03/06

(41) Mise à la disp. pub./Open to Public Insp.: 2000/09/09

(45) Date de délivrance/Issue Date: 2008/05/13

(30) **Priorités/Priorities:** 1999/03/09 (EP99104668.1);
2000/02/01 (EP00101666.6)

(51) Cl.Int./Int.Cl. *C12N 15/52* (2006.01),
C12N 1/16 (2006.01), *C12N 15/63* (2006.01),
C12N 5/10 (2006.01), *C12N 9/00* (2006.01),
C12P 21/00 (2006.01), *C12P 23/00* (2006.01)

(72) Inventeurs/Inventors:
HOSHINO, TATSUO, JP;
OJIMA, KAZUYUKI, JP;
SETOGUCHI, YUTAKA, JP

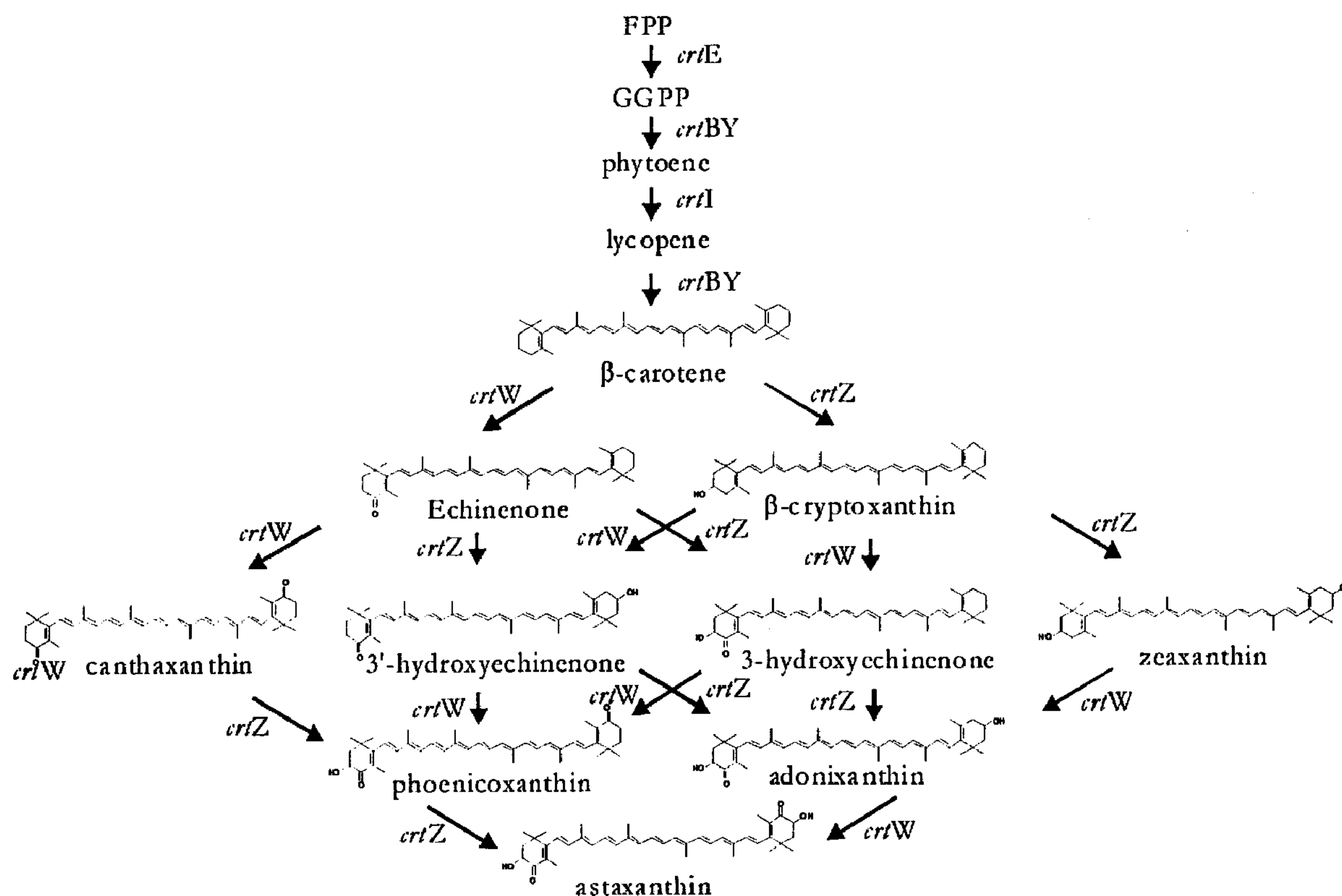
(73) Propriétaire/Owner:
DSM IP ASSETS B.V., NL

(74) **Agent:** GOWLING LAFLEUR HENDERSON LLP

(54) Titre : ASTAXANTHINE-SYNTHETASE

(54) Title: ASTAXANTHIN SYNTHETASE

Biosynthetic pathway from GGPP to astaxanthin
in the bacterial sytem which produces astaxanthin



(57) Abrégé/Abstract:

The present invention is directed to genetic materials useful for the preparation of astaxanthin from beta-carotene, such as polypeptides having astaxanthin synthase activity, DNA fragments coding for astaxanthin synthase, recombinant organisms and

(57) **Abrégé(suite)/Abstract(continued):**

the like. Those novel genetic materials may be originated from *Phaffia rhodozyma*. The present invention also provides a process for the production of astaxanthin.

Abstract

The present invention is directed to genetic materials useful for the preparation of astaxanthin from beta-carotene, such as polypeptides having
5 astaxanthin synthase activity, DNA fragments coding for astaxanthin synthase, recombinant organisms and the like. Those novel genetic materials may be originated from *Phaffia rhodozyma*. The present invention also provides a process for the production of astaxanthin.

The present invention relates to recombinant production of carotenoids and biological materials useful therefor.

5 *Phaffia rhodozyma* (*P. rhodozyma*) is a carotenogenic yeast strain which produces astaxanthin. Astaxanthin is distributed in a wide variety of organisms such as animal (birds such as flamingo and scarlet ibis, and fish such as rainbow trout and salmon), algae and microorganisms. It is also recognized that astaxanthin has a strong antioxidation property against
10 oxygen radical, which is expected to apply for pharmaceutical usage to protect living cells against some diseases such as a cancer. Moreover, industrial need for astaxanthin as a coloring reagent is increasing, especially in the industry of farmed fish like salmon, because astaxanthin imparts distinctive orange-red coloration to the animals and contributes to consumer appeal in the
15 marketplace.

P. rhodozyma is known as a carotenogenic yeast strain which produces astaxanthin. Different from the other carotenogenic yeast, *Rhodotorula* species, *P. rhodozyma* can ferment some sugars such as D-glucose. This is an important feature from a viewpoint of industrial application. In a recent
20 taxonomic study, a sexual cycle of *P. rhodozyma* was revealed and its telemorphic state was designated under the name of *Xanthophyllomyces*

- 2 -

dendrorhous (W. I. Golubev; Yeast 11, 101 – 110, 1995). Some strain improvement studies to obtain hyper producers of astaxanthin from *P. rhodozyma* have been conducted, but such efforts have been restricted to employ the method of conventional mutagenesis and protoplast fusion in this decade. Recently, Wery et al. developed a host vector system using *P. rhodozyma* in which a non-replicable plasmid was used to be integrated in multicopies into the genome of the ribosomal DNA of *P. rhodozyma* (Wery et al., Gene, 184, 89-97, 1997). Verdoes et al. reported more improved vectors to obtain a transformant of *P. rhodozyma* as well as its three carotenogenic genes which code for the enzymes that catalyzes the reactions from geranylgeranyl pyrophosphate to beta-carotene (WO 97/23633).

A specific biosynthetic pathway for carotenogenesis branches from general isoprenoid pathway at the point of important intermediate, farnesyl pyrophosphate (FPP) (FIG.1). FPP and IPP are condensed by geranylgeranyl pyrophosphate (GGPP) synthase which is coded by *crtE* in *P. rhodozyma* to produce GGPP. GGPP is then converted to beta-carotene by the sequential reaction of an enzyme functioning doubly as phytoene synthase and lycopene cyclase which is coded by *crtBY* and phytoene desaturase coded by *crtI*.

In bacteria, enzymes and genes which are involved in xanthophyll formation have been isolated and characterized in detail. beta-Carotene hydroxylase which is coded by *crtZ* is involved in the two steps of hydroxylation for beta-ionone-ring of beta-carotene at both of the ends. The *crtZ* gene was cloned from wide variety of organisms such as *Erwinia uredovora* (Misawa et al., J. Bacteriol., 172, 6704-6712, 1990), *Flavobacter* species (L. Pasamontes et al., 185 (1), 35-41, 1997) and *Agrobacterium aurantiacum* (Misawa et al., J. Bacteriol., 177 (22), 6575-6584, 1995). beta-Carotene ketolase which is coded by *crtW* catalyzes the two steps of introduction of oxo-group into beta-ionone -ring of beta-carotene at both of the ends. Kajiwara et al. cloned and sequenced *bkt* gene corresponding to *crtW* in eubacteria from *Haematococcus bluvialis* (Kajiwara et al., P. Mol. Biol., 29, 343-352, 1995). Harker et al. also cloned and sequenced *crtO* gene

- 3 -

corresponding to crtW in eubacteria from *Synechococcus* PCC7942 (Harker et al., FEBS Letters, 404, 129-134, 1997). Both enzymes the hydroxylase and the ketolase have wide substrate specificity and this ensures the formation of wide variety of xanthophylls in case that both of the enzymes react at the same
5 time depending on the reaction condition. (FIG.1)

As described above, all the genes which were involved in the formation of beta-carotene from FPP were isolated but the enzymes and genes which would be involved in the last step of xanthophyll formation from beta-carotene
10 have not been identified on the protein and DNA level in *P. rhodozyma*. Although Johnson et al. (Crit. Rev. Biotechnol, 11 (4), 297 – 326, 1991) proposed the existence of two independent pathways for astaxanthin formation by assuming that some of the xanthophyll compounds isolated by them would be intermediates of astaxanthin biosynthesis, such two independent pathways
15 could not be proven because enzymes and genes which are involved in such pathways could not be isolated. Furthermore, it can not be excluded that these xanthophyll compounds could have been resulted from an experimental artifact in the isolation step of these compounds. Failure to isolate a mutant from *P. rhodozyma* which accumulates intermediates on the biosynthetic
20 pathway from beta-carotene to astaxanthin made it difficult to clarify the biosynthetic pathway from beta-carotene to astaxanthin.

This invention relates to a gene and enzyme which is involved in the last step of the astaxanthin biosynthesis from beta-carotene to astaxanthin.
25

The present invention, provides an isolated DNA, specifically cDNA comprising a nucleotide sequence coding for astaxanthin synthase which is involved in the reaction from beta-carotene to astaxanthin in *P. rhodozyma*, like the AST gene.

30 In a preferred embodiment, the cloned DNA fragment can be characterized in that

(a) the said nucleotide sequence codes for the said enzyme having an amino

- 4 -

acid sequence described in SEQ ID NO: 1, or

- (b) the said nucleotide sequence codes for a variant of the said enzyme selected from (i) an allelic variant or (ii) an enzyme having one or more amino acid addition, insertion, deletion and/or substitution and having the stated enzyme
5 activity.

In another preferred embodiment, the isolated cDNA fragment can be derived from a gene of *Phaffia rhodozyma* and is selected from:

- (i) a cDNA sequence represented by SEQ ID NO: 2;
10 (ii) an isocoding or an allelic variant for the cDNA sequence represented by SEQ ID NO: 2; and
(iii) a derivative of a cDNA sequence represented by SEQ ID NO: 2 with addition, insertion, deletion and/or substitution of one or more nucleotide(s), and coding for a polypeptide having the said enzyme
15 activity.

In another preferred embodiment the present invention is directed to the isolated cDNA as described above, which is characterized in that the said nucleotide sequence is:

- 20 (i) a nucleotide sequence represented in SEQ ID NO: 2;
(ii) a nucleotide sequence which, because of the degeneracy of the genetic code, encodes an astaxanthin synthase having the same amino acid sequence as that encoded by the nucleotide sequence (i); and
(iii) a nucleotide sequence which hybridizes to the complement of the
25 nucleotide sequence from i) or ii) under standard hybridizing conditions.

In still another preferred embodiment, the isolated genomic DNA fragment can be derived from a gene of *Phaffia rhodozyma* and is selected from:

- 30 (i) a genomic DNA sequence represented by SEQ ID NO: 3;
(ii) an isocoding or an allelic variant for the genomic DNA sequence represented by SEQ ID NO: 3; and

- 5 -

- (iii) a derivative of a genomic DNA sequence represented by SEQ ID NO: 3 with addition, insertion, deletion and/or substitution of one or more nucleotide(s), and coding for a polypeptide having the said enzyme activity.

5

In another preferred embodiment the present invention is directed to the isolated genomic DNA as described above, which is characterized in that the said nucleotide sequence is:

- (i) a nucleotide sequence represented in SEQ ID NO: 3;
- 10 (ii) a nucleotide sequence which, because of the degeneracy of the genetic code, encodes an astaxanthin synthase having the same amino acid sequence as that encoded by the nucleotide sequence (i); and
- (iii) a nucleotide sequence which hybridizes to the complement of the nucleotide sequence from i) or ii) under standard hybridizing
- 15 conditions.

Another aspect of the present invention, provides a recombinant polypeptide having an astaxanthin synthase activity involved in the reaction from beta-carotene to astaxanthin in *P. rhodozyma* which is obtainable by the expression

20 of the cloned DNA fragment as set forth above.

A preferred embodiment of the recombinant polypeptide of the present invention is characterized in that

- (a) the said polypeptide has an amino acid sequence as described in SEQ ID
- 25 NO: 1, or
- (b) the said polypeptide is a variant of the peptide defined in (a) which is selected from (i) an allelic variant or (ii) an enzyme having one or more amino acid addition, insertion, deletion and/or substitution and having the stated enzyme activity.

30

Therefore the present invention is also directed to variants of the polypeptides of the present case. Such variants are defined on the basis of the amino acid

- 6 -

sequence of the present invention by addition, insertion, deletion and/or substitution of one or more amino acid residues of such sequences wherein such derivatives still have the same type of enzymatic activity as the corresponding polypeptides of the present invention or they are the result of
5 the well known phenomenon of allelic variation. Such activities can be measured by any assays known in the art or specifically described herein. Such variants can be made either by chemical peptide synthesis known in the art or by recombinant means on the basis of the DNA sequences as disclosed herein by methods known in the state of the art, such as, e.g. that disclosed by
10 Sambrook et al. (Molecular Cloning, Cold Spring Harbour Laboratory Press, New York, USA, second edition 1989). Amino acid exchanges in proteins and peptides which do not generally alter the activity of such molecules are known in the state of the art and are described, for example, by H. Neurath and R.L. Hill in "The Proteins" (Academix Press, New York, 1979, see especially Figure
15 6, page 14). The most commonly occurring exchanges are: Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/thr, Ser/Asn, Ala/Val, Ser/Gly, Thy/Phe, Ala/Pro, Lys/Art. Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, Asp/Gly as well as the reverse.

Furthermore the present invention is not only directed to the DNA sequences
20 as disclosed e.g., in the sequence listing as well as their complementary strands, but also to those which include these sequences, DNA sequences which hybridize under Standard Conditions with such sequences or fragments thereof and DNA sequences, which because of the degeneration of the genetic code, do not hybridize under Standard Conditions with such sequences but
25 which code for polypeptides having exactly the same amino acid sequence.

"Standard Conditions" for hybridization mean in this context the conditions which are generally used by a man skilled in the art to detect specific hybridization signals and which are described, e.g. by Sambrook et al., (s.a.) or
30 preferably so-called stringent hybridization and non-stringent washing conditions, or more preferably so-called stringent hybridization and stringent washing conditions a man skilled in the art is familiar with and which are

- 7 -

described, e.g., in Sambrook et al. (s.a) or more preferably so-called medium stringent conditions, e.g. using the DIG (digoxigenin) labeling kit and luminescent detection kit of Boehringer Mannheim (Mannheim, Germany) following the protocol given by the manufacturer and using as the
5 hybridization solution:

formamide (WAKO, Osaka, Japan) 50% (V/V)
5 x SSC
blocking reagent (Boehringer) 2% (W/V)
10 N-lauroylsarcosine 0.1% (W/V)
SDS 0.3% (W/V)

at a temperature of 42°C over night and subsequently washing and detection as indicated by the manufacturer.

15

DNA sequences which are derived from the DNA sequences of the present invention either because they hybridize with such DNA sequences (see above) or can be constructed by the polymerase chain reaction by using primers designed on the basis of such DNA sequences can be prepared either as
20 indicated namely by the PCR reaction, or by site directed mutagenesis [see e.g., Smith, Ann. Rev. Genet. 19, 423 (1985)] or synthetically as described, e.g., in EP 747 483 or by the usual methods of Molecular Cloning as described, e.g., in Sambrook et al. (s.a.).

25 The present invention is also directed to a vector or plasmid comprising a DNA as described above and a host cell transformed or transfected by a DNA as described above or a vector or plasmid as indicated above.

The present invention also provides a recombinant organism which is
30 obtainable by the transformation of a host using a recombinant DNA carrying the DNA as mentioned above.

- 8 -

The present invention also relates to a method for producing an enzymatic polypeptide capable of catalyzing the reaction from beta-carotene to astaxanthin, which comprises culturing a recombinant organism described above under the condition conducive to the production of said enzymatic
5 polypeptide.

In a further aspect the present invention provides a method for the production of astaxanthin which comprises introducing one or more of the DNA described above into an appropriate host organism and cultivating this transformed
10 organism under the conditions conducive to the production of astaxanthin.

The enzymatic polypeptide of the present invention will be also useful for a method for the production of astaxanthin, which method comprises contacting beta-carotene with the recombinant polypeptide having an astaxanthin
15 synthase activity involved in the reaction from beta-carotene to astaxanthin as set forth above in the presence of an appropriate electron donor in an appropriate reaction mixture containing an appropriate reconstituted membrane. In this method, said recombinant polypeptide may be present in the form of a reconstituted membrane which is prepared from biological
20 membranes like microsomes or mitochondrial membranes. The recombinant polypeptide may be also present in the form of a reconstituted artificial membrane, like liposomes. The electron donor like cytochrome P450 reductase is an appropriate electron donor which can reduce a reaction center of the enzyme of the present invention.

25

The following drawings are included to further illustrate the present invention together with the detailed description given below.

FIG. 1 shows the biosynthetic pathway from acetyl-CoA to astaxanthin in *P. rhodozyma*.

30 FIG. 2 shows a restriction map of the plasmid pR16 which harbors a partial genomic AST gene .

- 9 -

FIG. 3 shows the expression study for AST gene to which 6 x His were added at its amino terminal end on removal of its transmembrane domain. The cells from 0.1 ml of broth were subjected to 10 % sodium dodesyl sulfide - polyacrylamide gel electrophoresis (SDS-PAGE). Lane 1, Molecular weight
5 marker (105 kDa, 82.0 kDa, 49.0 kDa and 33.3 kDa, up to down, Bio-RAD, Richmond, U.S.A.); Lane 2, E. coli (BL21 (DE3) (pLysS) (pAST315) without IPTG); Lane 3, E. coli (BL21 (DE3) (pLysS) (pAST315) with 1.5 mM IPTG); Lane 4, Molecular weight marker).

In general, there are a number of methods to clone a gene coding for
10 biosynthetic enzymes. For example, degenerate PCR can be used. Degenerate PCR is a method to clone a gene of interest which has high homology of its encoded amino acid sequence to the one of a known enzyme from other species which has the same or similar function. A degenerate primer, which is used as a set of primers in degenerate PCR, was designed by a reverse translation
15 of the amino acid sequence to corresponding nucleotides ("degenerated"). In such a degenerate primer, a mixed primer which consists any of A, C, G or T, or a primer containing inosine at an ambiguity code is generally used. After cloning of a partial fragment of the gene of interest the genetic fragment containing the entire gene can be screened by using the cloned and labeled
20 partial DNA fragment as a probe.

In the case of cloning a gene encoding an enzyme whose activity can be measured by an enzymatic assay, purification of such an enzyme by monitoring enzyme activity and determination of its amino acid sequence for the enzyme is a good method. An amino acid sequence thus obtained is easily
25 translated in reverse into the corresponding nucleotide sequence(s). A DNA fragment which has the corresponding nucleotide sequence can be synthesized in vitro with a DNA synthesizer and labeled for direct usage as a hybridization probe. An alternative way to obtain a hybridization probe is the degenerate PCR method using the amino acid sequence information.

30 To clone a gene whose function can not be characterized enzymatically, a method called shot-gun screening has been employed as a conventional

- 10 -

cloning method. This method comprises the isolation of a mutant strain which lacks the specific gene coding for any of the biosynthetic enzymes of interest, [and the transformation of the mutant strain by the DNA prepared from the organism that has an intact gene which corresponds to the gene as such
5 mutated. For an isolation of such a mutant, conventional mutagenesis is often used. Confirmation of the acquired phenotype which is the same as that of the parent strain can be performed by examination of its auxotrophy and the like. In the case that the donor DNA contains the gene which corresponds to the mutated gene in the mutant strain, the transformant by such a gene acquired
10 the same phenotype as the parent strain as a result of genetic complementation.

As a vector, any form of vectors whether they can replicate or not in the cloning host, can be used for shot gun cloning. For the usage of a replicative vector, the capability of complementation is a requisite and it is not necessary
15 that such a vector contains a homologous sequence to the genome of the recipient. In the case of using a non-replicative vector, it is necessary that such a vector contains a homologous sequence to the genome of the host to make a recombination between donor and recipient DNAs.

For the cloning of the DNAs of the present invention the method called
20 "color complementation" can be employed. Non carotenogenic organisms, such as *Escherichia coli* can acquire the carotenogenic ability as a result of transformation by carotenogenic genes which could be cloned from carotenogenic organisms such as *Erwinia uredovora*, *Erwinia herbicola* and the like. *E. coli* harboring *crtE*, *crtB*, *crtI* and *crtY* can produce beta-carotene
25 and color the cells in yellow. Exploiting such a characteristics, numbers of carotenogenic genes have been cloned from various carotenogenic organisms such as bacteria and plants. For example, to clone the *crtY* gene coding for lycopene cyclase, *E. coli* harboring *crtE*, *crtB* and *crtI* on the compatible vector against pUC vector is prepared as a transformation host. Such a host turns
30 red which shows accumulation of lycopene. Next, a cDNA or genomic library from carotenogenic organisms can be constructed using the pUC vector. If the

- 11 -

gene corresponding to crtY in the donor carotenogenic organism is present in the transformed plasmid, genetic complementation would occur and *E. coli* would turn yellow which would show the acquisition of the ability to produce beta-carotene. In fact, crtE, crtBY and crtI gene were cloned from *P. rhodozyma* by this method (Verdoes et al., WO97/23633).

Regarding the cloning of the gene which is involved in the reaction from beta-carotene to astaxanthin, Kajiwarra et al. constructed a cDNA expression library from *P. rhodozyma* in the host *E. coli* harboring crtE, crtB, crtI and crtY genes from *E. uredovora* (Kajiwarra et al., WO 96/28545, 1996). In such a cloning system, a gene which is involved in the reaction from beta-carotene to astaxanthin could be theoretically cloned from *P. rhodozyma* by judging red pigmentation which shows the accumulation of canthaxanthin or astaxanthin. However, such a gene has not been reported so far. Many researchers speculate about the possibility that membrane-bound carotenogenic enzymes would form an enzyme complex. In such a model, the affinity among carotenogenic enzymes would be necessary for efficient carotenogenesis. Based on such an assumption, this color complementation method is not suitable to clone the enzyme involved in the last step of astaxanthin biosynthesis, namely the one from beta-carotene to astaxanthin, because exogeneous enzymes might not have affinity to the *Phaffia*'s carotenogenic enzyme in the sequential reaction of the carotenogenesis.

In this invention, *P. rhodozyma* ATCC96815 which has been redeposited as a Budapest Treaty deposit at the American Type Culture Collection (ATCC) under accession number 74486 on February 18, 1999 and which is blocked for the reaction from beta-carotene to astaxanthin was used as a transformation host (Schroeder, W.A. and Johnson, E.A., J. Ind. Microbiol. 14, 502-507, 1995). Transformation of this mutant by the genomic library prepared from the chromosome of a wild type strain of *P. rhodozyma* ATCC96594 which has also been redeposited as a Budapest Treaty deposit at the American Type Culture Collection (ATCC) under accession number 74438 on April 8, 1998 was used to isolate a clone which produces astaxanthin. In the present invention, such

- 12 -

genetic fragment complementing the reaction from beta-carotene to astaxanthin in *P. rhodozyma* was isolated and its nucleotide sequence was determined.

Such a gene / DNA of the present invention can be used for
5 overproduction of astaxanthin through a gene dosage effect using gene amplification or promoter modification other than complementation of blocked mutation.

In general, a gene consists of several parts which have different functions. In eukaryotes, genes which encode corresponding protein are
10 transcribed to premature messenger RNA (pre-mRNA), differing from the genes for ribosomal RNA (rRNA), small nuclear RNA (snRNA) and transfer RNA (tRNA). Although RNA polymerase II (PolII) plays a central role in this transcription event, PolII cannot solely start a transcription without a cis element covering an upstream region containing a promoter and an upstream
15 activation sequence (UAS), and a trans-acting protein factor. At first, a transcription initiation complex which consists of several basic protein components recognize the promoter sequence in the 5'-adjacent region of the gene to be expressed. In this event, some additional participants are required in the case of the gene which is expressed under some specific regulation, such
20 as a heat shock response, or adaptation to a nutrition starvation, and so on. In such a case, a UAS is required to exist in the 5'-untranslated upstream region around the promoter sequence, and some positive or negative regulator proteins recognize and bind to the UAS. The strength of the binding of transcription initiation complex to the promoter sequence is affected by such a
25 binding of the trans-acting factor around the promoter, and this enables the regulation of the transcription activity.

After the activation of a transcription initiation complex by the phosphorylation, a transcription initiation complex initiates a transcription from the transcription start site. Some parts of the transcription initiation
30 complex are detached as an elongation complex from the promoter region to the 3' direction of the gene (this step is called as a promoter clearance event)

- 13 -

and an elongation complex continues the transcription until it reaches to a termination sequence that is located in the 3'-adjacent downstream region of the gene. Pre-mRNA thus generated is modified in nucleus by the addition of cap structure at the cap site which almost corresponds to the transcription
5 start site, and by the addition of polyA stretches at the polyA signal which locates at the 3'-adjacent downstream region. Next, intron structures are removed from coding region and exon parts are combined to yield an open reading frame whose sequence corresponds to the primary amino acid sequence of a corresponding protein. This modification in which a mature
10 mRNA is generated is necessary for a stable gene expression. cDNA in general terms corresponds to the DNA sequence which is reverse-transcribed from this mature mRNA sequence. It can be synthesized by the reverse transcriptase derived from viral species by using a mature mRNA as a template, experimentally.

15 In this invention, mutation point of the *P. rhodozyma* ATCC96815 strain which rendered beta-carotene production to *P. rhodozyma* wild type strain was determined. From the sequencing result, it was suggested that the base change at the splicing sequence of the eighth intron of AST gene cause such a phenotype as specific beta-carotene accumulation through the improper
20 splicing of mRNA. RT-PCR analysis detected the improper spliced product for AST gene and strongly supported the identification of the mutation point.

This invention also provides the recombinant AST gene which can be expressed in different host organisms such as *E. coli*. In this invention, such recombinant AST gene was expressed in *E. coli* and it was confirmed that AST
25 gene encoded protein product whose size was corresponded to the deduced molecular weight. Biological production of astaxanthin can be realized by using novel AST gene and such recombinant DNA techniques.

According to the present invention, the gene coding for the enzyme which is involved in the last step of astaxanthin biosynthesis was cloned from
30 a cDNA library of *P. rhodozyma*, and their nucleotide sequence was determined. Furthermore, a part of genomic DNA including promoter and

- 14 -

terminator was cloned and can be used for a cloning of its entire gene including the promoter and terminator region.

An entire gene with its coding region, its intron as well as its regulation regions such as a promoter and terminator can be cloned by screening a genomic library, which has been constructed in a phage or plasmid vector in an appropriate host by using a labeled cDNA fragment as a screening probe. Generally, one of the most common host strains for the construction of a genomic library is *E. coli*. As a vector a phage vector, such as a lambda phage vector, or a plasmid vector such as a pUC vector can be used. A genomic library constructed in this way, e.g. from *P. rhodozyma* DNA can be screened by using a labeled DNA fragment with a portion of the gene of interest as a probe. Hybridized plaques or colonies can then be picked and used for subcloning and/or determination of the nucleotide sequence.

There are several strategies to enhance the desired enzymatic activity of the protein of interest by using its DNA sequence.

One strategy is to use the gene itself in its native form. The simplest approach is to amplify the genomic sequence including its regulatory sequences such as the promoter and the terminator. This can be done by cloning of the genomic fragment coding for the enzyme of interest into an appropriate vector with a selectable marker which functions in *P. rhodozyma*. A drug resistance gene coding for an enzyme that enables the host to survive in the presence of a toxic antibiotic is often used as a selectable marker. The G418 resistance gene harbored in pGB-Ph9 (Wery et al., *Gene*, 184, 89-97, 1997) is an example of such a vector construction. As a vector, two types of vectors can be commonly used. One of these types is an integration vector which does not have an autonomous replicating sequence. The plasmid pGB-Ph9 is an example of this type of vectors. Because such a vector does not have an autonomous replicating sequence, the above vector cannot replicate by itself and can be present only in an integrated form on the chromosome of the host as a result of a single-crossing recombination using the homologous sequence between a vector and the chromosome. By increasing the

- 15 -

concentration of the corresponding drug in the selection medium, the strain in which the integrated gene is amplified on the chromosome can only survive. Another type of vector is a replicable vector which has an autonomous replicating sequence. Such a vector can exist in a multicopy state. In this type of vector, a nutrition complementation maker can be also used in the host which has an appropriate auxotrophy marker. The *P. rhodozyma* ATCC24221 strain which requires cytidine for its growth is one example of such an auxotroph. By using a CTP synthetase as a donor DNA for ATCC24221, a host vector system using a nutrition complementation can be established.

Another strategy to overexpress an enzyme of interest is the placement of the gene of interest under a strong promoter. In such a strategy, the gene of interest must not necessarily be in a multicopy state. Furthermore, a promoter whose promoter activity is induced in an appropriate growth phase and an appropriate timing of cultivation can be also used. Production of astaxanthin accelerates in the late phase of the growth, such as the production phase of a secondary metabolite. For example, by placing carotenogenic genes under the control of a vegetative promoter, the gene expression of these genes could be induced in the exponential growth phase and the production of astaxanthin can become associated to the growth of the production strain.

In this invention, promoter and terminator fragment for triose phosphate isomerase (TPI) gene was cloned from *P. rhodozyma* as one example of such constitutive promoter and terminator. Moreover, a restoration of astaxanthin production was confirmed in the transformants in which AST gene was expressed on the different locus (AMY locus which lies amylase gene) on the chromosome of beta-carotene producing *P. rhodozyma* ATCC96815 driven by constitutive promoter and terminator derived from TPI gene.

Still another strategy to overexpress enzymes of interest is mutation in its regulatory elements. For this purpose, a kind of reporter gene, such as the beta-galactosidase gene, luciferase gene, a gene coding for a green fluorescent protein, and the like is inserted between the promoter and the terminator sequence of the gene of interest so that all the parts including

- 16 -

promoter, terminator and the reporter gene are fused and function together. A transformed *P. rhodozyma* in which said reporter gene is introduced on the chromosome or on the vector can be mutagenized in vivo to induce a mutation within the promoter region of the gene of interest. The mutation can be
5 monitored by detecting the change of the activity coded for by the reporter gene. If the mutation occurs in a cis element of the gene, the mutation point would be determined by the rescue of the mutagenized gene and sequencing. This mutation can then be introduced to the promoter region on the chromosome by the recombination between the native and the mutated
10 promotor sequence. In the same way a mutation in the gene which codes for a trans-acting factor can be made.

A mutation can be also induced by an in vitro mutagenesis of a cis element in the promoter region. In this approach, a gene cassette, containing a reporter gene which is fused to a promoter region derived from a gene of
15 interest at its 5'-end and a terminator region from a gene of interest at its 3'-end, is mutagenized and then introduced into *P. rhodozyma*. By detecting the difference of the activity of the reporter gene, an effective mutation would be screened. Such a mutation can be introduced in the sequence of the native promoter region on the chromosome by the same method as the case of an in
20 vivo mutation approach.

As donor DNA, a gene coding for an enzyme which catalyzes the reaction from beta-carotene to astaxanthin could be introduced. A coding sequence which is identical to its native sequence, as well as its allelic variant, a sequence which has one or more amino acid additions, deletions and/or
25 substitutions as far as its corresponding enzyme has the same type of enzyme activity, can be used. Such a vector can then be introduced into *P. rhodozyma* by transformation and the transformants can be selected by spreading the transformed cells on an appropriate selection medium such as YPD agar medium containing geneticin in the case of pGB-Ph9 or a minimal agar
30 medium omitting cytidine in the case of using auxotroph ATCC24221 as a recipient.

- 17 -

Such a genetically engineered *P. rhodozyma* can be cultivated in an appropriate medium and evaluated for its productivity of astaxanthin. A hyper producer of astaxanthin thus selected would be confirmed in view of the relationship between its productivity and the level of gene or protein expression which is introduced by such a genetic engineering method.

Examples

The following materials and methods were employed in the specific Examples described below:

Strains:

P. rhodozyma ATCC96594 (This strain has been redeposited on April 8, 1998 as a Budapest Treaty deposit under accession No. 74438)

P. rhodozyma ATCC96815 (This strain has been redeposited on February 18, 1999 as a Budapest Treaty deposit under accession No. 74486)

E. coli DH5alpha F⁻, phi80d, lacZdeltaM15, delta(lacZYA-argF)U169, hsd (r_K⁻, m_K⁺), recA1, endA1, deoR, thi-1, supE44, gyrA96, relA1 (Toyobo, Osaka, Japan)

E. coli XL1-Blue MRF[']: delta(mcrA)183, delta(mcrCB-hsdSMR-mrr)173, endA1, supE44, thi-1, recA1, gyrA96, relA1, lac[F['] proAB, lacI[']ZdeltaM15, Tn10 (tet^r)] (Stratagene, La Jolla, USA)

E. coli SOLR: e14['](mcrA), delta(mcrCB-hsdSMR-mrr)171, sbcC, recB, recJ, umuC :: Tn5(kan^r), uvrC, lac, gyrA96, relA1, thi-1, endA1, lambda^R, [F['] proAB, lacI[']ZdeltaM15] Su⁻(nonsuppressing) (Stratagene)

E. coli TOP10: F⁻, mcrA, delta(mrr-hsdRMS-mcrBC), phi80, delta(lacZ M15), delta(lacX74), recA1, deoR, araD139, (ara-leu)7697, galU, galK, rpsL(Str^r), endA1, nupG (Invitrogen, NV Leek, Netherlands)

- 18 -

E. coli BL21 (DE3) (pLysS): *dcm*⁻, *ompTr_Bm_B*⁻, *lon*⁻ lambda(DE3), pLysS (Stratagene)

Vectors:

- 5 pUC19 (Takara Shuzo, Otsu, Japan)
 lambdaZAPII (Stratagene)
 pCR2.1-TOPO (Invitrogen)
 pET11c (Stratagene)

10 Media

P. rhodozyma strain is maintained routinely in YPD medium (DIFCO, Detroit, USA). *E. coli* strain is maintained in LB medium (10 g Bacto-trypton, 5 g yeast extract (DIFCO) and 5 g NaCl per liter). When an agar medium was prepared, 1.5 % of agar (WAKO, Osaka, Japan) was supplemented.

15 Methods

General molecular biology methods were done according to those described in Molecular cloning: a Laboratory Manual, 2nd Edition (Cold Spring Harbor Laboratory Press, 1989). Restriction enzymes and T4 DNA ligase were purchased from Takara Shuzo.

- 20 Isolation of chromosomal DNA from *P. rhodozyma* was performed by using QIAGEN Genomic Kit (QIAGEN, Hilden, Germany) following the protocol supplied by the manufacturer. Mini-prep of plasmid DNA from transformed *E. coli* was performed with the Automatic DNA isolation system (PI-50, Kurabo, Co. Ltd., Osaka, Japan). Midi-prep of plasmid DNA from an
- 25 *E. coli* transformant was performed by using QIAGEN column (QIAGEN). A DNA fragment was isolated and purified from agarose by using QIAquick or QIAEX II (QIAGEN).

- Fluorescent DNA primers for DNA sequencing were purchased from Pharmacia. DNA sequencing was performed with the automated fluorescent
- 30 DNA sequencer (ALFred, Pharmacia, Uppsala, Sweden).

Competent cells of DH5alpha were purchased from Toyobo.

The apparatus and reagent for biolistic transformation of *P. rhodozyma* were purchased from Nippon Bio-Rad Laboratories (Tokyo, Japan).

5

Example 1

Isolation of genomic DNA from *P. rhodozyma*

10 To isolate a genomic DNA from *P. rhodozyma*, ATCC96594 the QIAGEN genomic kit was used according to the method specified by the manufacturer.

At first, cells of *P. rhodozyma* ATCC96594 from 100 ml of overnight culture in YPD medium were harvested by centrifugation (1500 x g for 10 min.) and washed once with TE buffer (10 mM Tris / HCl (pH 8.0) containing 1 mM EDTA). After suspending in 8 ml of Y1 buffer of the QIAGEN genomic
15 kit, lyticase (SIGMA, St. Louis, USA) was added at the concentration of 2 mg/ml to disrupt cells by enzymatic degradation and the reaction mixture was incubated for 90 minutes at 30 °C and then proceeded to the next extraction step. Finally, 20 µg of genomic DNA was obtained.

20

Example 2

Construction of a genomic library from *P. rhodozyma* ATCC96594

As described in the section "detailed description of the invention", a plasmid harboring a drug resistant marker cassette was constructed by
25 inserting a G418 resistant structure gene between the promoter and terminator region of the gene of the glyceraldehyde-3-phosphate dehydrogenase (GAP) and ligating this cassette into the KpnI- and HindIII-digested pUC19. This plasmid was named pUC-G418 and used further on.

- 20 -

Then, a ClaI linker was ligated into the unique EcoRI site of the pUC-G418 vector and the resultant plasmid pUC-G418Cl512 was obtained and used as a vector backbone in the construction of the *Phaffia*'s genomic library.

Then, 10 µg of chromosomal DNA prepared from *P. rhodozyma*
5 ATCC96594 as described above in Example 1 was digested partially with 1.6 units of HpaII for 45 minutes at 37 °C and was subjected to agarose gel electrophoresis. After staining by ethidium bromide, partially digested DNA species from 4 to 10 kb were recovered by electroelution using a dialysis membrane. After ethanol precipitation of recovered HpaII fragments, 1.215
10 µg of DNA was obtained.

Next, 3 µg of pG418Cl512 was digested by 10 units of ClaI for an hour at 37 °C and precipitated with ethanol. ClaI-digested pG418Cl512 was then dephosphorylated using calf intestine alkaline phosphatase. ClaI-digested and dephosphorylated pG418Cl512 vector was then subjected to agarose gel
15 electrophoresis and the DNA fragment was recovered using the QIAquick protocol according to the instructions of the manufacturer. Finally, 2.62 µg of ClaI-digested and dephosphorylated pG418Cl512 was obtained.

2.62 µg of ClaI-digested and dephosphorylated pG418Cl512 was ligated to 1.22 µg of HpaII-partially digested *Phaffia*'s genomic DNA over
20 night at 16 °C and the resultant ligation solution was used as donor DNA for the transformation of an *E. coli* DH5alpha strain. The total ligation mixture (270 µl) was transferred to 1 ml of competent cells of DH5a (Toyobo). After a heat shock treatment at 42 °C for 45 seconds succeeding by maintenance on ice for 30 minutes, the transformed cells were placed on ice for 2 minutes and
25 then incubated at 37 °C for an hour with the addition of 5 ml of SOC medium:

0.5% yeast extract (DIFCO)

2% trypton (DIFCO)

10mM NaCl

2.5 mM KCl

- 21 -

10 mM MgCl₂20 mM MgSO₄

20 mM glucose.

Thus incubated cells were transferred into 100 ml of LB medium
5 containing 100 gamma/ml of ampicillin. Cultivation was continued overnight
at 37 °C and then the cells were harvested for plasmid midi-preparation.

A plasmid library was prepared from harvested cells using QIAGEN
midi-prep columns according to the method supplied by the manufacturer.
Finally, 0.3 mg/ml of Phaffia's genomic library was obtained in the total
10 volume of 5 ml and used as a genomic library in the further study.

Example 3

Transformation of *P. rhodozyma* ATCC96815 with a biolistic method
15

Transformation was done according to the method described in Methods
in Molecular Biology (Johnston et al., 53; 147-153, 1996). As a host strain, *P.*
rhodozyma ATCC96815 was cultured in YPD medium to the stationary phase.
After centrifugation of the broth, cells were concentrated 10-fold with
20 sterilized water and 200 µl of the cell suspension was spread on YPD medium
containing 100 gamma/ml of geneticin, and 0.75 M of mannitol and sorbitol.
Five micrograms of a genomic library, prepared as described in Example 2,
was coated on 1.5 mg of 0.9 µm gold particle, and used as donor DNA for the
biolistic transformation. Approximately twenty thousands of geneticin-
25 resistant clones were yielded (300 to 500 colonies per plate) after one week of
incubation at 20 °C. Although most of the transformants showed a yellow
color (as the host strain, ATCC96815 did), three colonies pigmented red and
were used further on.

Example 4

Analysis of carotenoid obtained from red-pigmented transformants

5

Red-pigmented transformants obtained from *P. rhodozyma* ATCC96815 were cultivated in 10 ml of YPD medium at 20 °C in test tubes. Then, cells were harvested from 0.5 ml of broth and used for the extraction of carotenoids from cells. The carotenoid content of *P. rhodozyma* was measured
10 by HPLC after extraction of the carotenoids from cells of *P. rhodozyma* by disruption with glass beads as described. After extraction, disrupted cells were then collected by centrifugation and the resultant supernatant was analyzed for carotenoid content with HPLC.

HPLC column; Chrompack Lichrosorb si-60 (4.6 mm, 250 mm)

15 Temperature; room temperature

Eluent; acetone / hexane (18 / 82) add 1 ml / L of water to elluent

Injection volume; 10 µl

Flow Rate; 2.0 ml/minute

Detection; UV at 450 nm

20 A sample of beta-carotene was purchased from SIGMA and astaxanthin was obtained from Hoffman La-Roche (Basel, Switzerland).

As a result of HPLC analysis, it was confirmed that all three red transformants produced astaxanthin specifically though the host strain, ATCC96815 produced only beta-carotene.

- 23 -

Example 5

- 5 Plasmid rescue from the chromosome of red transformants which produced astaxanthin

Chromosomal DNA was prepared from all the astaxanthin-producing transformants. For this purpose, the QIAGEN genomic kit was
10 used according to the method specified by the manufacturer, as described in Example 1. 5 µg of chromosomal DNA, thus prepared, was digested by HindIII and then purified according the QIAquick protocol. E. coli DH5alpha competent cells were transformed by the ligated DNA solutions and then spread on LB agar medium containing 100 µg/ml of ampicillin. All of the
15 transformants had the same insert fragments in their plasmids, judging from sequence analysis of the plasmids. This indicated that three independent red transformants derived from P. rhodozyma ATCC96815 were yielded by the same type of recombination event between the donor DNA of the genomic library and chromosomal DNA. One of the plasmids thus rescued was named
20 pR2-4 and used further on.

Example 6

- 25 Screening of the original genomic library by using pR2-4 as a hybridization probe

Because the rescued fragment in pR2-4 may have mutations depending on the direction of the recombination event yielding red transformants of P.
30 rhodozyma, screening of the original genomic library was done by using pR2-4 as a hybridization probe.

- 24 -

For this purpose, twenty thousand *E. coli* transformants of the original genomic library, as described in Example 2 were transferred to nylon membrane filters (Hybond-N+, Amersham, Buckinghamshire, UK) and subjected to colony hybridization. Three transformants which harbored the same nucleotide sequence in their insert as that of pR2-4 were isolated. The isolated plasmids from these transformants were named pR3, pR5.1 and pR16.

Next, *P. rhodozyma* ATCC96815 was transformed by pR3, pR5.1 and pR16. All the transformants colored red. This result suggests that the isolated plasmids might contain the gene encoding an enzyme involved in the reaction of beta-carotene to astaxanthin in *P. rhodozyma*. We designated this gene as AST gene. Among these plasmids, pR16 was used further on.

Example 7

Isolation of mRNA from *P. rhodozyma* for cDNA analysis

To analyze the pattern of transcripts from *P. rhodozyma*, total RNA was isolated from *P. rhodozyma* ATCC96594 and ATCC96815 by phenol extraction by combination of the cell disruption with glass beads and purified mRNA using an mRNA separation kit (Clontech, Palo Alto, USA).

At first, cells of ATCC96594 and ATCC96815 strains from 10 ml of a two-day-culture in YPD medium were harvested by centrifugation (1500 x g for 10 min.) and washed once with extraction buffer (100 mM Tris / HCl (pH 7.5) containing 0.1 M LiCl and 0.1 mM EDTA). After filling up to 5.0 ml of cell suspension with the same extraction buffer in 50 ml disposable centrifuge tube (IWAKI Glass, Tokyo, Japan), 1.5 ml of isogen-LS (Nippon gene, Toyama, Japan) and 10 grams of glass beads were added. Centrifuge tubes which contained the cell suspension with isogen-LS and glass beads was shaken with a horizontal table top shaker for an hour. In this step, 300 µg of total RNA was recovered.

- 25 -

Then, mRNA was purified by using an mRNA separation kit (Clontech). On 8.0 µg of mRNA from *P. rhodozyma* ATCC96594 and ATCC96815 strains were obtained.

To synthesize cDNA, we used the SMART cDNA construction kit (Clontech) according to the method specified by the manufacturer. We applied 2 µg of purified mRNA for a first strand synthesis followed by PCR amplification and obtained 1 mg of cDNA.

Example 8

10 Subcloning of pR16 and functional analysis of its insert fragment

The restriction map of pR16 is depicted in FIG. 2. Each EcoRI fragment whose length was 0.7 and 2.7 kb, was subcloned into pUC-G418 and named pRS913 and pRLR913, respectively.

15 Then, the astaxanthin-producing *P. rhodozyma* ATCC96594 strain was transformed with pRS913. As a result of this transformation study, yellow transformants were yielded. This suggested that 0.7 kb EcoRI fragment might contain a truncated AST gene and the transformation via a single-cross recombination between a 0.7 kb EcoRI fragment and its homologous sequence
20 on the chromosome of *P. rhodozyma* would result in a gene disruption of the AST gene on the chromosome of *P. rhodozyma*.

Next, the beta-carotene-producing ATCC96815 strain was transformed with pRLR913 and red transformants were yielded. This suggested that the mutation point of strain ATCC96815 which led the astaxanthin-producing
25 wild type strain to produce beta-carotene would lie in the 2.7 kb EcoRI fragment originally adjacent to the 0.7 kb EcoRI fragment in pR16.

Two hundred µg of cDNA prepared in Example 7 was subjected to agarose gel electrophoresis for virtual Northern analysis. In the case of the

- 26 -

cDNAs prepared from ATCC96594 and 96815, two bands which, namely at 3.2 and 2.0 kb were hybridized in both cases by using the 2.7 kb EcoRI fragment of pRLR913 as a hybridization probe. This suggested that the ast mutation of ATCC96815 would be a point mutation which did not reflect to the change of
5 the length of mRNA such as missense mutation.

In the case of using the 0.7 kb EcoRI fragment of pRS913 as a hybridization probe, a band of 2.0 kb was hybridized. From this study, it seemed that the AST gene might give a 2.0 kb transcript in *P. rhodozyma*.

10 **Example 9**

Cloning of the cDNA of the AST gene

To clone the cDNA for the AST gene from *P. rhodozyma*, we constructed cDNA library from *P. rhodozyma* ATCC96594. Total RNA was isolated by
15 phenol extraction by combination of the cell disruption with glass beads as described in Example 7.

At first, cells of the ATCC96594 strain from 50 ml of a two-day-culture in YPD medium were harvested by centrifugation (1500 x g for 10 min.) and washed once with extraction buffer (100 mM Tris / HCl (pH 7.5) containing 0.1
20 M LiCl and 0.1 mM EDTA). After filling up to 5.0 ml of cell suspension with the same extraction buffer in 50 ml disposable centrifuge tube (IWAKI Glass), 1.5 ml of isogen-LS (Nippon gene) and 10 grams of glass beads were added. Centrifuge tubes which contained cell suspension with isogen-LS and glass beads were shaken with a horizontal table top shaker for an hour. In this
25 step, 1.8 mg of total RNA was recovered.

Then, mRNA was purified by using the PolyATtract mRNA isolation system (Promega corp., Madison, USA) according to the method specified by the manufacturer. Finally, we obtained 8.0 µg of mRNA from the *P. rhodozyma* ATCC96594 strain.

- 27 -

To construct a cDNA library, 8.0 µg of the purified mRNA was used in the COPY kit (Invitrogen, Carlsbad, USA) with the protocol specified by the manufacturer. After ligation of a EcoRI adaptor (Stratagene), synthesized cDNA was subjected to agarose gel electrophoresis. After the excision of the
5 agarose gel piece which covered the length of cDNAs from 1.9 to 2.3 kb, the collected cDNA species were purified by QIAEX II (QIAGEN). This size-fractionated cDNA was ligated to EcoRI-digested and dephosphorylated lambdaZAPII (Stratagene). The over-night ligation mixture was in vitro packaged with Gigapack III gold extract (Stratagene) and used for infection to
10 *E. coli* XL1-Blue MRF' strain.

Conventional plaque screening was performed against 6000 plaques using 2.7 and 0.7 kb EcoRI fragment as described in Example 8 as hybridization probes. One plaque hybridized strongly to these probes, picked up with a sterilized toothpick and the eluted phage particle were used for in
15 vivo excision, according to the method specified by the manufacturer. Finally, infected transformants of *E. coli* SOLR cells which showed resistance against ampicillin were isolated. After sequencing of the isolated plasmids obtained from these transformants, it turned out that these plasmids contained the same fragment as a part of the sequence of pR16 which was described in
20 Example 6.

The entire sequence of the cDNA of the AST gene was determined and is shown as SEQ ID NO: 2 and its deduced amino acid sequence as SEQ ID NO: 1.

Example 10

25 Expression of AST gene in *E. coli*

To confirm that ORF for AST gene actually encode protein, expression study of AST gene was performed in *E. coli* expression system. At first, 6 x histidine (His) tag was added to carboxyl terminal end of AST product in order

- 28 -

to make easy for following purification step. PCR primers whose sequences are listed in TABLE 1 were synthesized.

TABLE 1

PCR primers for cloning 3' portion of AST gene to which 6 x His tag is added

- 5 ast13; GTTCAAAGTTCATTTATGGA (sense primer) (SEQ ID NO: 4)
- ast14;
- GGATCCTCAGTGGTGGTGGTGGTGGTGGTTCGACCGGCTTGACCTGCA
- (antisense primer) (SEQ ID NO: 5)

- Next, 1.5 kb of NdeIEcoRI fragment of pAST1207 and 0.3 kb of
- 10 EcoRIBamHI fragment of pAST114 were ligated into pET11c which was digested by NdeI and BamHI and ligated DNA was transformed into E. coli JM109 strain. Six independent ampicillin resistant clones were examined by restriction analysis and it was found that 5 of 6 clones had correct structure of recombinant expression plasmid containing AST gene. One of which was
- 15 selected for further study (pAST120) and then transformed into E. coli BL21 (DE3) (pLysS) strain. It was revealed that all of the ampicillin resistant clones which were examined by restriction analysis possessed pAST120 properly. Next, expression study was performed by addition of 1mM IPTG to E.coli BL21 (DE3) (pLysS) (pAST120) growing culture when optical density
- 20 (OD) at 600 nm reached to 0.8. After continuation of cultivation at 37 °C for 4 hours, cells were harvested by centrifugation and lysed by boiling in SDS sample buffer (125 mM Tris-HCl, pH 6.8, 20 % glycerol, 4 % SDS, 0.005% bromophenol blue, 5 % mercaptoethanol). The lysate was then subjected to SDS-polyacrylamide gel electrophoresis (PAGE). Expressed protein was not
- 25 observed after staining by coomassie brilliant blue (Rapid stain CBB kit, nacalai tesque, Kyoto, JAPAN) (data not shown).

In general, it is reported that some modifications of amino acid sequence at amino terminal region of P450 protein is required to express P450 protein in E. coli expression system. In fact, AST gene which had an intact

- 29 -

sequence was not expressed in *E. coli* (data not shown) and it was found that some modifications of sequence at amino terminal end was necessary in the case of AST gene as well as other P450 enzymes. As a next strategy for expression of recombinant AST gene, the construction, in which 6 x His tag
 5 sequence was added at amino terminal end of the AST protein on the deletion of the hydrophobic anchor sequences which were located at amino terminal end of AST gene, was made.

In order to add 6 x His tag sequence at amino terminal end of the AST protein on anchor sequences deleted at amino terminal end, following PCR
 10 primers were synthesized and used for PCR cloning.

TABLE 2

PCR primers for cloning AST gene lacking anchor sequence at its 5' portion to which 6 x His tag is added

ast32; CATATGCACCACCACCACCACCTGTATAACCTTCAGGGGCCC
 15 (sense primer for cloning of 5' end of AST gene) (SEQ ID NO: 6)

ast2; GTAACAACACCATCTCCGGT (antisense primer for cloning of 5' end of AST gene) (SEQ ID NO: 7)

ast13; GTTCAAAGTTCATTTATGGA (sense primer for cloning of 3' end of AST gene) (SEQ ID NO: 4)

20 ast33; GGATCCTCAACTCATTCGACCGGCTT (antisense primer for cloning of 3' end of AST gene) (SEQ ID NO: 8)

PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 30 seconds at 72 °C. The plasmid, pAST1207 was used as PCR template. PCR fragments which had desired length were cloned into
 25 pCR2.1-TOPO (Invitrogen) and 6 independent clones which had expected inserts were examined for their insert sequence. As a result, two of the clones had exact insert sequence, respectively and one clone was selected and used for further study (pAST228 for 3' end of AST gene and pAST302#3202 for 5'

- 30 -

end of AST gene, respectively). 0.2kb NdeISphI fragment from pAST302#3202, 1.5kb SphIEcoRI fragment from pAST1207 and 0.05kb KpnIBamHI fragment from pAST228 were ligated into pET11c digested by NdeI and BamHI and ligated mixture was transformed into E. coli DH5alpha.

5 As a result of restriction analysis for 6 independent clones, it was found that all the clones had correct structure harboring AST gene for its expression. One clone was selected and used for further study (pAST315). Next, pAST315 was transferred into expression host E. coli BL21 (DE3) (pLysS). It was confirmed that all the 6 transformants had pAST315 correctly as a result of

10 restriction analysis.

Next, expression study was performed by addition of 1.5 mM IPTG to E.coli BL21 (DE3) (pLysS) (pAST315) growing culture when optical density (OD) at 600 nm reached to 0.93. After continuation of cultivation at 37 °C for 4 hours, cells were harvested by centrifugation and lysed by boiling in SDS

15 sample buffer. The lysate was then subjected to PAGE. Expressed protein whose molecular weight corresponded well with its deduced amino acid sequence (approximate 60 kDa) was observed after staining by coomassie brilliant blue (FIG. 3). From this result, it was confirmed that AST gene encode protein expected from its deduced open reading frame.

20

Example 11

In vitro characterization of the AST gene product

For the enzymatic characterization of the AST gene product, a

25 standard assay which is used for the characterization of P450 enzymes can be applied. For this purpose, it is necessary that the reaction mixture contains a reconstituted membrane. As a reconstituted membrane, natural isolates such as mitochondrial membranes or microsomes and artificial membranes are often used. In any cases, it is necessary that an electron transfer between an

30 electron acceptor and a receptor can occur. As an electron donor, cytochrome

- 31 -

P450 reductase is often added to the reaction mixture. As an electron acceptor, oxygen molecules are involved. Under the presence of an electron source, such as reduced NADPH+, beta-carotene, which is a substrate of astaxanthin synthase can be converted to astaxanthin. Produced astaxanthin
5 can be assayed qualitatively and quantitatively with HPLC analysis.

Example 12

Cloning of genomic fragment containing AST gene

10 To determine the genomic sequence containing AST gene, sequencing experiment performed by using primer-walking procedure. Sequencing analysis of pRS913 showed that pRS913 did not contain the 3' end of AST gene. To obtain the 3'-adjacent genomic fragment to AST gene, genome-walking experiment was performed. To do this, universal genome walker kit
15 (Clontech) was exploited according to the method specified by the manufacturer. As a template of PCR, chromosomal DNA prepared in Example 1 was used. Gene specific primer, ast15 whose sequence was as listed in TABLE 3 was synthesized and used for PCR primer;

TABLE 3

20 Sequence of primer used for genome walking of AST gene

ast15; TAGAGAGAAGGAGGGGTACCAGATGC (SEQ ID NO. 9)

PCR fragments which had appropriate length (smaller than 1 kb) were obtained from EcoRV and StuI library, purified and cloned into pCR2.1-TOPO (Invitrogen). As a result of sequencing, it was found that both fragments
25 contained the genomic fragment contained AST gene. Based on the sequence which was located at 200 bp from polyA site for AST gene, PCR primer was designed as listed in TABLE 4.

- 32 -

TABLE 4

Sequence of primer for cloning of 3'-adjacent fragment to AST gene

ast18; CCCCGGATTGTGGAGAACT (SEQ ID NO: 10)

5 By using ast15 and ast18 primers as PCR primers and chromosomal DNA prepared in Example 1 as PCR template, PCR was conducted. Proof-reading polymerase (HF polymerase, Clontech) ensured the amplification of PCR fragment which had exact sequence. PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 30 seconds at 72 °C. Six
10 independent clones which had 400 bp inserts showed the identical sequence.

By combined with sequences for pRS913 and pRL913, 3.9 kb sequence containing AST gene containing 474 bp of promoter region and 269 bp of terminator region was determined (SEQ ID NO. 3). As a result, AST gene showed intron-rich structure (17 introns).

15

Example 13

Determination of mutation point in beta-carotene producing strain, *P. rhodozyma* ATCC96815

To confirm the fact that beta-carotene production by *P. rhodozyma*
20 ATCC96815 strain was caused from the mutation within the AST gene, genomic sequence containing AST gene obtained from ATCC96815 and its parent strain *P. rhodozyma* ATCC24230 were determined. To do this, PCR primers whose sequences were as listed in TABLE 5 were synthesized and used for PCR cloning;

25

TABLE 5

PCR primers for cloning entire genomic AST gene;

ast21; ATGTTTCATCTTGGTCTTGCT (sense primer) (SEQ ID NO: 11)

ast4; ACGTAGAAGTCATAGCGCCT (antisense primer) (SEQ ID NO: 12)

- 5 By using HF polymerase (Clontech) as PCR polymerase and chromosomal DNA prepared from strains, ATCC96815 and ATCC24230 by the same protocol as Example 1 as PCR template, PCR was performed under the condition as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 4 minutes at 72 °C. PCR fragments obtained whose length were approximate
- 10 3.5 kb were cloned into pCR2.1-TOPO and sequenced for their entire sequences by primer walking procedure. Between the sequence for *P. rhodozyma* ATCC96594 strain and ATCC24230 strain, 7 base changes were found. Four base changes were found in its exon sequence but those did not give any amino acid changes. Three base changes were found in its intron
- 15 structure. In comparison between beta-carotene producing strain ATCC96815 and its parent strain, ATCC24230, one base change which was located at 5'-splicing sequence (GTAAGT > GTAAAT) within the eighth intron was found. This might indicate that mutation which conferred the phenotype of beta-carotene accumulation on astaxanthin-producing *P. rhodozyma* was caused
- 20 from improper splicing of mRNA for AST gene.

To confirm this assumption, RT-PCR was performed by using cDNA prepared from *P. rhodozyma* ATCC96815 as PCR template. mRNA was isolated from *P. rhodozyma* ATCC96815 by the same protocol as Example 9 and used for the synthesis of cDNA. To obtain cDNA from this mRNA prepared from

25 ATCC96815 by PCR method, SMART PCR cDNA library construction kit (Clontech) was exploited according to the method specified by supplier. The following primers whose sequence were as listed in TABLE 6 and which covered eighth intron were synthesized and used for PCR primers.

TABLE 6

PCR primers for RT-PCR to detect the improper splicing product for AST gene

ast7; TTTGACTCAAGGATTAGCAG (sense primer) (SEQ ID NO: 13)

ast26; TGTCTTCTGAGAGTCGGTGA (antisense primer) (SEQ ID NO: 14)

- 5 RT-PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 30 seconds at 72 °C. As a result of PCR, 300 bp of PCR products were amplified and cloned into pCR2.1-TOPO. Two independent clones which had 300 bp insert were sequenced. As a result, it was confirmed that improper splicing products for AST gene was synthesized in the *P. rhodozyma*
- 10 ATCC96815 strain. Improper splicing for the eighth intron of AST gene might cause the production of shorter truncated AST protein than AST protein spliced properly because stop codon lay in the eighth intron. This result indicated that mutation point lay in AST gene which failed in the proper splicing.

15

Example 14

Expression of AST gene in beta-carotene-producing *Phaffia rhodozyma*

- To confirm the fact that AST gene encoded the enzyme which was involved in the conversion from beta-carotene to astaxanthin, AST gene was cloned into beta-carotene-producing strain. To exclude the possibility of recombination at native locus of AST gene on the chromosome, expression plasmid for AST gene on AMY locus of *Phaffia rhodozyma*'s chromosome was constructed. To do this, cloning of some genetic elements from *Phaffia rhodozyma* was required.

- 25 1) Cloning of constitutive promoter and terminator from *Phaffia rhodozyma*

To clone a constitutive promoter and terminator from *Phaffia rhodozyma*, degenerate PCR method was exploited. Among the genes which are often used

- 35 -

as constitutive promoter and terminator in yeast's genetics, TPI gene which encode triose phosphate isomerase was tried to be cloned. Among the conserved amino acid sequence registered in Blocks database

(blocks.fhcrc.org), two motif sequences (Arg-Thr-Phe-Phe-Val-Gly-Gly-Asn and Asp-Val-Asp-Gly-Phe-Leu-Val-Gly-Gly-Ala) were selected and their degenerate primers were synthesized as follows.

TABLE 7

Degenerate PCR primers for cloning of TPI gene from *P. rhodozyma*

tp1; MGNACNTTYTTYGTNGGNGGNAAY (sense primer) (SEQ ID NO: 15)
 10 tp6; GCNCCNCCNACNARRAANCCRTCACRTC (antisense primer) (SEQ ID NO: 16)

(M=A or C; N=A, C, G or T; Y=C or T; R=A or G)

PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 46 °C and 15 seconds at 72 °C. ExTaq polymerase (Takara Shuzo) was used as
 15 PCR polymerase. As a PCR template, cDNA pool was prepared from mRNA isolated from *P. rhodozyma* ATCC96594 by using SMART PCR cDNA library construction kit (Clontech). 0.7 kb PCR fragment was purified and cloned into pCR2.1-TOPO. Six independent clones had inserts having desired length, judging from restriction analysis. Two of which were sequenced and it was
 20 confirmed that both of them had insert sequence which had striking homology to known TPI genes from various organisms. One of which was selected for further study (pTPI923).

Next, based on the insert sequence of pTPI923, several PCR primers whose sequences are listed in TABLE 8 were synthesized for genome walking
 25 to clone promoter and terminator of TPI gene. For this experiment, universal genome walker kit (Clontech) was exploited according to the method specified by the manufacturer.

- 36 -

TABLE 8

PCR primers for genome walking to clone TPI promoter and terminator

tp9; GCTTACCTCGCTTCCAACGTTTCCCAG (terminator cloning, primary)
(SEQ ID NO: 17)

5 tp10; GGATCTGTCTCTGCCTCCAAGTCAAG (terminator cloning, nested)
(SEQ ID NO: 18)

tp11; GGGTCAATGTCGGCAGCGAGAAGCCCA (promoter cloning, primary)
(SEQ ID NO: 19)

tp12; ATGTACTCGGTAGCACTGATCAAGTAG (promoter cloning, nested)
10 (SEQ ID NO: 20)

PCR condition was as follows; 7 cycles of 4 seconds at 94 °C and 3 minutes at 74 °C, followed by 32 cycles of 4 seconds at 94 °C and 3 minutes at 69 °C and succeeded to extension at 69 °C for 4 minutes. KOD polymerase (TOYOBO) was used as PCR polymerase. Chromosomal DNA prepared from P.

15 rhodozyma ATCC96594 was used as a PCR template. As a result, candidate for terminator region was obtained from EcoRV and StuI library. Sequencing analysis for these candidates revealed that both clones had the downstream sequence for TPI gene containing deduced 3' end of the TPI structure gene and terminator region. In case of cloning for promoter region, candidates which
20 were obtained from EcoRV library contained deduced 5' end of the TPI structure gene and promoter region.

Then, PCR primers whose sequences are listed in TABLE 9 were synthesized for the construction of promoter cassette and terminator cassette derived from TPI gene.

- 37 -

TABLE 9

PCR primers to construct TPI promoter and TPI terminator cassette

tp13; GCGGCCGCATCCGTCTCGCCATCAGTCT (sense primer for promoter cassette) (SEQ ID NO: 21)

5 tp14; CCTGCAGGTCTAGAGATGAATAAATATAAAGAGT (antisense primer for promoter cassette) (SEQ ID NO: 22)

tp15; CCTGCAGGTAAATATATCCAGGGATTAACCCCTA (sense primer for terminator cassette) (SEQ ID NO: 23)

10 tp16; GGTACCCGTGCGCAGTCGACCGAGACAT (antisense primer for terminator cassette) (SEQ ID NO: 24)

PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 30 seconds at 72 °C. HF polymerase (Clontech) was used as PCR polymerase and yielded PCR fragments were cloned into pCR2.1-TOPO. As a result of restriction and sequencing analysis, it was found that clones which
15 had identical sequences were obtained. Each one clone was selected for further study (pTPIP1104 for promoter cassette and pTPIT1104 for terminator cassette, respectively).

2) Cloning of partial amylase gene from *Phaffia rhodozyma*

To locate and express a foreign gene on the chromosome of *P. rhodozyma*,
20 amylase gene was cloned from *P. rhodozyma*. In case that expression vector on which foreign gene would be cloned could contain homologous genetic fragment to the chromosomal sequence of *P. rhodozyma* such as amylase gene, expression vector can be integrated on the homologous region on chromosome of *P. rhodozyma* after the single cross recombination.

25 Eleven amino acid sequences for amylase from various organisms were selected from Entrez database (ncbi.nlm.nih.gov/Entrez) and used for amino acid alignment by clustal W (Thompson, J.D., Higgins, D.G. and

- 38 -

Gibson, T.J., Nucleic Acids Research, 22: 4673-4680, 1994). The eleven organisms whose amylase sequences were registered on the database are as listed in TABLE 10;

TABLE 10

5 Various amylase genes which were registered on database for clustal W analysis

Aspergillus niger var. awamori amyA gene (accession number X52755)

Aspergillus niger var. awamori amyB gene (accession number X52756)

Aspergillus kawachii acid-stable alpha-amylase gene (accession number
10 AB008370)

Aspergillus oryzae amy1 gene (accession number X12725)

Aspergillus shirousamii alpha-amylase gene (accession number P30292)

Cryptococcus species alpha-amylase gene (accession number D83541)

Lipomyces kononenkoae subsp. spencermartinsiae alpha-amylase gene
15 (accession number U30376)

Debaryomyces occidentalis amy1 gene (accession number X16040)

Saccharomycopsis fibuligera ALP1 gene (accession number X05791)

Schizosaccharomyces pombe alpha-amylase gene (accession number Z64354)

Two conserved amino acid sequences (Asp-Tyr-Ile-Gln-Gly-Met-Gly-Phe-
20 Asp/Thr-Ala-Ile-Trp and Asp-Gly-Ile-Pro-Ile-Ile-Tyr-Tyr-Gly-Thr-Glu-Gln) for
amylase were selected to clone amylase gene from *P. rhodozyma* by degenerate
PCR method. Then, PCR primers whose sequences are listed in TABLE 11
were synthesized for the cloning of AMY gene from *P. rhodozyma*.

- 39 -

TABLE 11

Degenerate PCR primers for cloning of amylase (AMY) gene from *P. rhodozyma*

amy1; GAYTAYATHCARGGNATGGGNTTYRMNGCNATHTG (sense primer)
5 (SEQ ID NO: 25)

amy2; TGYTCNGTNC CRTARTADATDATNGG DATNCCRTC (antisense
primer) (SEQ ID NO: 26)

(Y=C or T; H=A, C or T; R=A or G; N=A, C, G or T; M=A or C; D=A, G or T)

PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at
10 50 °C and 2 minutes at 72 °C. ExTaq polymerase (Takara Shuzo) was used as
PCR polymerase. As a PCR template, chromosomal DNA prepared in
Example 1 and cDNA pool prepared from mRNA isolated from *P. rhodozyma*
ATCC96594 by using SMART PCR cDNA library construction kit (Clontech)
were used. 1.7 kb and 0.9 kb PCR fragments were yielded when chromosome
15 and cDNA were used as PCR template, respectively. Both fragments were
purified and cloned into pCR2.1-TOPO. Six independent clones had inserts
having desired length, judging from restriction analysis. Two of which were
sequenced and it was confirmed that both of them had insert sequence which
had striking homology to known amylase genes from various organisms. One
20 of which contained chromosomal AMY fragment was selected for further study
(pAMY216). To construct a partial amylase cassette, two PCR primers whose
sequences are listed in TABLE 12 were synthesized based on the internal
sequence of insert fragment of pAMY216.

- 40 -

TABLE 12

PCR primers to construct a partial AMY cassette

amy101; CCGCGGCATTGATACCTCTACCCCGT (sense primer for AMY cassette) (SEQ ID NO:27)

- 5 amy102; GCGGCCGCCTGCAATCCTGGATCCACCG (antisense primer for AMY cassette) (SEQ ID NO: 28)

PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 2 minutes at 72 °C. HF polymerase (Clontech) and chromosomal DNA were used as PCR polymerase and PCR template, respectively. The
10 yielded PCR fragment was cloned into pCR2.1-TOPO. As a result of restriction and sequencing analysis, it was found that the clone which had correct sequence was obtained. One clone was selected for further study (pAMY1113).

- 3) Construction of expression vector for AST gene which functioned in Phaffia
15 rhodozyma

Expression plasmid for AST gene was constructed by restriction digestion and ligation of each genetic component. At first, 0.3 kb KpnIPstI fragment from pTPIT1104 and 1.7 kb SacIKpnI fragment from pG418Sa512 were ligated into pGEM-T plasmid which was digested by SacI and PstI. It was found that 9
20 clones among 12 transformants had correct structure as a result of restriction digestion and one of those was selected for further study (pTPITG1120).

Next, PCR cloning for AST gene was performed to add the appropriate restriction site to both ends. PCR primers whose sequences are listed in TABLE 13 were synthesized.

TABLE 13

PCR primers to clone entire AST gene cassette

ast11; TCTAGAATGTTTCATCTTGGTCTTGCTCA (sense primer) (SEQ ID NO: 29)

5 ast12; CCTGCAGGTCATTCGACCGGCTTGACCT (antisense primer) (SEQ ID NO: 30)

PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C and 2 minutes at 72 °C. HF polymerase (Clontech) and pAST1207 were used as PCR polymerase and PCR template, respectively. The yielded PCR
10 fragment was cloned into pCR2.1-TOPO. As a result of restriction and sequencing analysis, it was found that one clone which had correct sequence was obtained. This clone was selected for further study (pAST113).

Finally, 1.6 kb SacII/NotI fragment from pAMY1113, 0.3 kb NotI/XbaI fragment from pTPIP1104 and 1.5 kb XbaI/Sse8387I fragment from pAST113
15 were ligated into pTPITG1120 which was digested by SacII and Sse8387I. It was confirmed that all the five transformants tested had correct structure as a result of restriction analysis and one of which was selected for further study (pAATG123).

4) Restoration of astaxanthin production in beta-carotene-producing *Phaffia*
20 *rhodozyma*

The expression plasmid for AST gene (pAATG123) was transformed into beta-carotene-producing *Phaffia rhodozyma* ATCC96815. Biolistic transformation was performed as described in Example 3. Two geneticin-resistant colonies which colored red were picked up and selected for further study. In order to
25 confirm the integration at AMY locus on the chromosome of *P. rhodozyma*, PCR primer whose sequence is listed in TABLE 14 was synthesized.

- 42 -

TABLE 14

PCR primer to confirm the integration of expression plasmid at AMY locus on
the chromosome of *P. rhodozyma*

amy5; CTCTCCTGTTACAAAAACA (sense primer) (SEQ ID NO: 31)

5

Chromosome was prepared from those transformants and used for PCR
template. PCR condition was as follows; 25 cycles of 15 seconds at 94 °C, 30
seconds at 55 °C and 2 minutes at 72 °C. ExTaq polymerase (Takara Shuzo)
was used as PCR polymerase. Positive 2.0 kb PCR band was yielded in the
10 PCR reaction in which chromosome obtained from red-colored transformants
was used as a template DNA. No PCR band was observed in the PCR reaction
mixture in which chromosome derived from host strain, *P. rhodozyma*
ATCC96815 was used as a PCR template.

5) Flask fermentation by recombinants in which recombinant AST gene was
15 integrated on the chromosome of beta-carotene-producing *P. rhodozyma*

The productivity of astaxanthin was evaluated in the flask fermentation. The
medium formulation for flask fermentation is as follows.

TABLE 15

20 Seed medium formulation for flask fermentation

	Glucose	30.0 g / l
	NH ₄ Cl	4.83 g / l
	KH ₂ PO ₄	1.0 g / l
25	MgSO ₄ -7H ₂ O	0.88 g / l
	NaCl	0.06 g / l

- 43 -

	CaCl ₂ -2H ₂ O	0.2 g / l
	KH phthalate	20.0 g / l
	FeSO ₄ -7H ₂ O	28 mg / l
	Citric acid-1H ₂ O	15.0 mg / l
5	ZnSO ₄ -7H ₂ O	40.0 mg / l
	CuSO ₄ -5H ₂ O	0.75 mg / l
	MnSO ₄ -4,5H ₂ O	0.6 mg / l
	H ₃ BO ₃	0.6 mg / l
	Na ₂ MoO ₄ -2H ₂ O	0.6 mg / l
10	KI	0.15 mg / l
	Myo-inositol	60.0 mg / l
	Nicotinic acid	3.0 mg / l
	Ca D-pantothenate	3.0 mg / l
	Vitamin B1 (thiamin HCl)	3.0 mg / l
15	p-Aminobenzoic acid	1.8 mg / l
	Vitamin B6 (pyridoxine HCl)	0.3 mg / l
	Biotin	0.048 mg / l
	7 ml / Test Tube (21 mm diameter)	

20

TABLE 16

Medium formulation for flask fermentation

	MgSO ₄ -7H ₂ O	2.1 g / l
	CaCl ₂ -2H ₂ O	0.865 g / l
25	(NH ₄) ₂ SO ₄	3.7 g / l
	FeSO ₄ -7H ₂ O	0.28 g / l
	Glucose (sterilized separately)	22 g / l

- 44 -

	KH_2PO_4 (sterilized separately)	14.25 g / l
	Citric acid- $1\text{H}_2\text{O}$	0.21 g / l
	$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	70.14 mg / l
	$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	10.5 mg / l
5	$\text{MnSO}_4 \cdot 4,5\text{H}_2\text{O}$	8.4 mg / l
	H_3BO_3	8.4 mg / l
	$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	8.4 mg / l
	KI	2.1 mg / l
	Myo-inositol	0.374 g / l
10	Nicotinic acid	18.7 mg / l
	Ca D-pantothenate	28.05 mg / l
	Vitamin B1 (thiamin HCl)	18.7 mg / l
	p-Aminobenzoic acid	11.22 mg / l
	Vitamin B6 (pyridoxine HCl)	1.87 mg / l
15	Biotin	1.122 mg /
	CaCO_3	10 g / l

1 drop of Actcol (Takeda Chemical Industries Ltd., Osaka, JAPAN) was added to each flask.

20 50ml (final volume with 5% of seed inoculum) / 500 ml flask with baffles

Cells were harvested from fermented broth after 7-day fermentation and analyzed for their accumulation of astaxanthin and beta-carotene by HPLC as described in Example 4. Results are summarized in TABLE 17.

- 45 -

TABLE 17

Restoration of astaxanthin production by the recombinants in which AST gene was integrated. (Data is indicated as relative titer of astaxanthin and beta-carotene against the titer of beta-carotene accumulated by *P. rhodozyma*

5 ATCC96815)

Relative titer (%)		
Strain	Astaxanthin	Beta-carotene
ATCC96815 :: pR16	34.0 %	18.6 %
ATCC96815 :: pAATG123	16.3 %	56.3 %
10 ATCC96815	0 %	100 %

Partial restoration of astaxanthin production by ATCC96815 :: pAATG123 indicated that promoter strength by TPI promoter is not strong enough for perfect restoration of astaxanthin production.

15

-46-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: F. Hoffmann-La Roche AG.
- (B) STREET: 124 Grenzacherstrasse
- (C) CITY: Basle
- (E) COUNTRY: Switzerland
- (F) POSTAL CODE (ZIP): CH-4070

(ii) TITLE OF INVENTION: Astaxanthine Synthetase

(iii) NUMBER OF SEQUENCES: 31

(iv) CORRESPONDENCE ADDRESS

- (A) NAME: GOWLING, STRATHY & HENDERSON
- (B) STREET: 160 ELGIN STREET, SUITE 2600
- (C) CITY: OTTAWA
- (D) PROVINCE: ONTARIO
- (E) COUNTRY: CANADA
- (F) POSTAL CODE: K1P 1C3

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: 2,298,869
- (B) FILING DATE: 06-MAR-2000

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: EPO 00101666.6
- (B) FILING DATE: 01-FEB-2000

- (A) APPLICATION NUMBER: EPO 99104668.1
- (B) FILING DATE: 09-MAR-1999

(viii) ATTORNEY/AGENT INFORMATION

- (A) NAME: GOWLING, STRATHY & HENDERSON
- (B) REFERENCE NUMBER: 08-886312CA

(ix) TELECOMMUNICATION INFORMATION

- (A) TELEPHONE: 613-233-1781
- (B) TELEFAX: 613-563-9869

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 557 amino acids
- (B) TYPE: amino acid

-47-

(C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
 (B) STRAIN: Phaffia rhodozyma

(ix) FEATURE:
 (A) NAME/KEY: Region
 (B) LOCATION:1..26
 (D) OTHER INFORMATION:/label= Transit

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

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

-48-

Gly	Phe	Asp	Tyr	Lys	Ser	Asp	Ser	Leu	Gln	Asn	Lys	Thr	Asn	Glu	Leu	210	215	220	
Tyr	Val	Ala	Phe	Val	Gly	Leu	Thr	Asp	Gly	Phe	Ala	Pro	Thr	Leu	Asp	225	230	235	240
Ser	Phe	Lys	Ala	Ile	Met	Trp	Asp	Phe	Val	Pro	Tyr	Phe	Arg	Thr	Met	245	250	255	
Lys	Arg	Arg	His	Glu	Ile	Pro	Leu	Thr	Gln	Gly	Leu	Ala	Val	Ser	Arg	260	265	270	
Arg	Val	Gly	Ile	Glu	Leu	Met	Glu	Gln	Lys	Lys	Gln	Ala	Val	Leu	Gly	275	280	285	
Ser	Ala	Ser	Asp	Gln	Ala	Val	Asp	Lys	Lys	Asp	Val	Gln	Gly	Arg	Asp	290	295	300	
Ile	Leu	Ser	Leu	Leu	Val	Arg	Ala	Asn	Ile	Ala	Ala	Asn	Leu	Pro	Glu	305	310	315	320
Ser	Gln	Lys	Leu	Ser	Asp	Glu	Glu	Val	Leu	Ala	Gln	Ile	Ser	Asn	Leu	325	330	335	
Leu	Phe	Ala	Gly	Tyr	Glu	Thr	Ser	Ser	Thr	Val	Leu	Thr	Trp	Met	Phe	340	345	350	
His	Arg	Leu	Ser	Glu	Asp	Lys	Ala	Val	Gln	Asp	Lys	Leu	Arg	Glu	Glu	355	360	365	
Ile	Cys	Gln	Ile	Asp	Thr	Asp	Met	Pro	Thr	Leu	Asp	Glu	Leu	Asn	Ala	370	375	380	
Leu	Pro	Tyr	Leu	Glu	Ala	Phe	Val	Lys	Glu	Ser	Leu	Arg	Leu	Asp	Pro	385	390	395	400
Pro	Ser	Pro	Tyr	Ala	Asn	Arg	Glu	Cys	Leu	Lys	Asp	Glu	Asp	Phe	Ile	405	410	415	
Pro	Leu	Ala	Glu	Pro	Val	Ile	Gly	Arg	Asp	Gly	Ser	Val	Ile	Asn	Glu	420	425	430	
Val	Arg	Ile	Thr	Lys	Gly	Thr	Met	Val	Met	Leu	Pro	Leu	Phe	Asn	Ile	435	440	445	
Asn	Arg	Ser	Lys	Phe	Ile	Tyr	Gly	Glu	Asp	Ala	Glu	Glu	Phe	Arg	Pro	450	455	460	
Glu	Arg	Trp	Leu	Glu	Asp	Val	Thr	Asp	Ser	Leu	Asn	Ser	Ile	Glu	Ala	465	470	475	480
Pro	Tyr	Gly	His	Gln	Ala	Ser	Phe	Ile	Ser	Gly	Pro	Arg	Ala	Cys	Phe	485	490	495	
Gly	Trp	Arg	Phe	Ala	Val	Ala	Glu	Met	Lys	Ala	Phe	Leu	Phe	Val	Thr				

-49-

	500		505		510										
Leu	Arg	Arg	Val	Gln	Phe	Glu	Pro	Ile	Ile	Ser	His	Pro	Glu	Tyr	Glu
		515					520					525			
His	Ile	Thr	Leu	Ile	Ile	Ser	Arg	Pro	Arg	Ile	Val	Gly	Arg	Glu	Lys
	530					535					540				
Glu	Gly	Tyr	Gln	Met	Arg	Leu	Gln	Val	Lys	Pro	Val	Glu			
545					550					555					

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1932 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: *Phaffia rhodozyma*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:33..1706

(ix) FEATURE:

- (A) NAME/KEY: polyA_site
- (B) LOCATION:1871

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION:14..1891

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GAATTCGGCA CGAGGCCACC TACTTTCTCC ATATGTTTCAT CTTGGTCTTG CTCACAGGTG	60
CTTTAGGCCT GGCTGCTTTC TCATGGGCAT CCATAGCGTT CTTCACTCTT TACCTCGCTC	120
CGAGGCGATC TTTACTGTAT AACCTTCAGG GCCCGAATCA TACCAACTAC TTTACAGGCA	180
ATTTTTTTAGA CATCCTCTCA GCTCGTACAG GTGAAGAGCA TGCGAAGTAC AGAGAAAAAT	240
ACGGAAGCAC CCTCCGGTTT GCTGGGATCG CTGGAGCACC CGTCTTGAAC TCGACCGATC	300
CGAAAGTCTT CAACCATGTG ATGAAAGAAG CCTACGACTA TCCGAAACCT GGTATGGCCG	360
CTCGAGTGCT CAGAATTGCT ACCGGAGATG GTGTTGTTAC GGCGGAAGGT GAAGCTCATA	420

- 50 -

AGCGACATCG AAGGATCATG ATCCCCTCTC TGTCCGCTCA GGCCGTTAAG TCGATGGTCC	480
CAATTTTCTT AGAAAAAGGT ATGGAACTTG TCGACAAGAT GATGGAGGAT GCGGCTGAGA	540
AGGATATGGC CGTGGGAGAG TCGGCCGGTG AAAAGAAGGC AACCAGACTC GAGACCGAAG	600
GAGTCGATGT AAAGGATTGG GTCGGTCGAG CTACTCTGGA CGTCATGGCT CTTGCAGGAT	660
TTGACTATAA GAGCGACTCG CTCCAGAACA AGACCAATGA GCTCTATGTC GCTTTTGTCG	720
GACTTACCGA TGGGTTTGCT CCTACCTTGG ACTCGTTCAA GGCTATCATG TGGGATTTTG	780
TACCTTACTT CCGAACTATG AAACGGAGAC ATGAGATACC TTTGACTCAA GGATTAGCAG	840
TTTCCCGACG AGTTGGGATC GAGCTTATGG AGCAAAAGAA GCAGGCCGTG CTTGGCTCAG	900
CTTCCGATCA GGCTGTTGAT AAAAAGGATG TTCAAGGTCG GGATATCCTA AGTCTCCTAG	960
TGAGAGCAAA CATCGCCGCC AACCTGCCTG AATCTCAAAA GCTGTCCGAT GAGGAGGTAC	1020
TCGCTCAGAT CAGTAACCTG TTATTTGCTG GATATGAAAC TTCTTCGACA GTCTTGACAT	1080
GGATGTTTCA CCGACTCTCA GAAGACAAAG CCGTTCAGGA TAAACTTCGA GAAGAAATTT	1140
GTCAGATCGA CACGGATATG CCTACGCTAG ACGAACTTAA TGC GTTGCCT TATCTCGAAG	1200
CGTTTGTTAA GGAGTCTCTT CGTCTAGACC CTCCTAGTCC GTATGCTAAC CGTGAATGCT	1260
TAAAGGATGA AGACTTCATC CCACTTGCCG AGCCTGTCAT TGGTCGAGAT GGGTCGGTCA	1320
TCAACGAGGT CCGGATCACG AAAGGAACGA TGGTCATGCT TCCGTTGTTC AACATCAATC	1380
GTTCAAAGTT CATTTATGGA GAAGATGCAG AAGAATTCAG ACCGGAGAGG TGGCTTGAGG	1440
ACGTAACAGA CTCGCTCAAC AGTATTGAAG CACCCTATGG ACACCAGGCG AGCTTTATCT	1500
CTGGACCCAG AGCTTGCTTT GGTGGCGAT TTGCTGTCGC CGAGATGAAG GCCTTCTTGT	1560
TTGTCACTCT CCGTCGGGTC CAGTTCGAGC CCATCATCTC TCATCCAGAG TACGAGCACA	1620
TCACCTTGAT CATTTCCCGT CCTCGAATCG TTGGTAGAGA GAAGGAGGGG TACCAGATGC	1680
GTTTGCAGGT CAAGCCGGTC GAATGAGTTG ATTCTTCATA TGTTAAGAGA AGTTCTATAT	1740
CTGAGAATGT GTGACTAGGA CAATGCCTTC TTTGTATCGA TTTGTTTCTC ATACCCGGGC	1800
AGGCGCTATG ACTTCTACGT CGTCTATCGT CGCTCTGGAC TCTCTTCTTA CCCTATATAT	1860
TATTCCATCC GAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AGCGGCCGCT CGAGCCGGCT	1920
CGTGCCGAAT TC	1932

(2) INFORMATION FOR SEQ ID NO: 3:

-51-

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3969 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: *Phaffia rhodozyma*
- (ix) FEATURE:
 - (A) NAME/KEY: 5'UTR
 - (B) LOCATION:517..518
 - (D) OTHER INFORMATION:/label= Experimental
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:535..652
- (ix) FEATURE:
 - (A) NAME/KEY: intron
 - (B) LOCATION:653..734
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:735..783
- (ix) FEATURE:
 - (A) NAME/KEY: intron
 - (B) LOCATION:784..898
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:899..1015
- (ix) FEATURE:
 - (A) NAME/KEY: intron
 - (B) LOCATION:1016..1087
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:1088..1179
- (ix) FEATURE:
 - (A) NAME/KEY: intron
 - (B) LOCATION:1180..1302
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:1303..1517
- (ix) FEATURE:
 - (A) NAME/KEY: intron
 - (B) LOCATION:1518..1600

-52-

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:1601..1634

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:1635..1723

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:1724..1866

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:1867..1939

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:1940..1999

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:2000..2081

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:2082..2181

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:2182..2257

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:2258..2354

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:2355..2431

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:2432..2542

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:2543..2618

(ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION:2619..2652

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION:2653..2742

-53-

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION:2743..2814

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION:2815..2962

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION:2963..3050

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION:3051..3113

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION:3114..3171

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION:3172..3247

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION:3248..3321

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION:3322..3398

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION:3399..3423

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION:3424..3513

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION:3514..3700

(ix) FEATURE:

(A) NAME/KEY: polyA_site

(B) LOCATION:3865..3866

(D) OTHER INFORMATION:/label= Experimental

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CGGACCGAAG CCTCGCCAGC AGTTGATCAA GCGAACCAAG CCGAACAATC CTGGCGCGCC

60

TGGAGGAGCG GGAGCGGGAG GAGCAGCAGG TGATGCATCG GGTGGACAGA ATCAGTAGTG

120

- 54 -

TGTGTGTATG TGTGTAGTGT AGTTGGGTTG TCCCATGTGC TTCTTCTTAT CATCATCATT	180
TCTTTAAAAT CTCTACATTG AATGTTTACC GGAACGGGCT TTGATGATAC TACGGACCAC	240
GTTGTGTAAAC CAGTTCGATT GAGATTACGA TTAGATAGCC GATCCGTCGA TCAGATCTCG	300
ATCTAGAGCG ACATCTGGCT CGATCGGTCC TTGCCGAAAA TCAGGGCACC GATCAGGGCA	360
GAGGAACGCC GAGGCCGAAC GAGACAGACA CACCATCATC ATCAGCCATG TCTTTTTTGT	420
GATCGTTTTT ACATACTACC CGTCGATTCT AACCTTCTTT CTTCTTCTCT TGCCATCTTT	480
GCATTCTCTA TCTCGTGTAAC CATCGATCCG ATTCTTGCCA CCTACTTTCT CCATATGTTC	540
ATCTTGGTCT TGCTCACAGG TGCTTTAGGC CTGGCTGCTT TCTCATGGGC ATCCATAGCG	600
TTCTTCAGTC TTTACCTCGC TCCGAGGCGA TCTTCACTGT ATAACCTTCA GGGTAAGAAT	660
TGAGCTCTGG AATCATGCTT GTGTAAATCC TATAATCTCA TTCATCCTAT TCCTCTTCTT	720
CATCCTCTCT TCAGGCCCGA ATCATACCAA CTACTTTACA GGCAATTTTT TAGACATCCT	780
CTCGTGAGTT TTCATCATTG GCTCAGTCGT CCAATCTTAA CGATCATCGC TAACGACCTT	840
TCGGACGCGT TCTTCTTTCT ATGTGAAATC TGATCTTTGG TTTGTTACGA GAGCACAGAG	900
CTCGTACAGG TGAAGAGCAT GCGAAGTACA GAGAAAAATA CGGAAGCACC CTCCGGTTTG	960
CTGGGATCGC TGGAGCACCC GTCTTGAACT CGACCGATCC GAAAGTCTTC AACCAGTTTG	1020
TCCATCCGAA CCCTCATCCT CCTCTGCTGA TCAATTCAAC TGTAGTTAAC GCACTTTGAA	1080
TGGACAGTGT GATGAAAGAA GCCTACGACT ATCCGAAACC TGGTATGGCC GCTCGAGTGC	1140
TCAGAATTGC TACCGGAGAT GGTGTTGTTA CGGCGGAAGG TGCTTTTCAA GTTCTCTTAT	1200
ATCACATCTA ATCCACTCGG CGCGATTGAA CTCAACATTT CTGACGAGCC TGTCACCTTG	1260
TTTTCACTTC ATGGTCTCGG TGCATCTTGT CTCATCTCAT AGGTGAAGCT CATAAGCGAC	1320
ATCGAAGGAT CATGATCCCC TCTCTGTCCG CTCAGGCCGT TAAGTCGATG GTCCCAATTT	1380
TCTTAGAAAA AGGTATGGAA CTTGTCGACA AGATGATGGA GGATGCGGCT GAGAAGGATA	1440
TGGCCGTGGG AGAGTCGGCC GGTGAAAAGA AGGCAACCAG ACTCGAGACC GAAGGAGTCG	1500
ATGTAAAGGA TTGGGTCGTG AGTACCCGCC TATTCCTTCA CCTTGATGGA CGAAGCATAT	1560
CAAGGAAAGG TTCATTGACT GACAAACACT ATCTTACCAG GGTCGAGCTA CTCTGGACGT	1620
CATGGCTCTT GCAGGTCAGT CTA CTCTCTC TTATAAATGC TCCACATATG TATGCATGTA	1680
CTGACATGCT CTTCTATAT TCGATACGAC GTCATATGTC CAGGATTTGA CTATAAGAGC	1740
GACTCGCTCC AGAACAAGAC CAATGAGCTC TATGTCGCTT TTGTCGGACT TACCGATGGG	1800

- 55 -

TTTGCTCCTA CCTTGGACTC GTTCAAGGCT ATCATGTGGG ATTTTGTACC TTACTTCCGA	1860
ACTATGGTAT GTCTGCCATT CTTTGATATC CAAAGATTAT GGATAGGTTA CTTGCTAAAA	1920
TTTCACCTAT CGTGAACAGA AACGGAGACA TGAGATACCT TTGACTCAAG GATTAGCAGT	1980
TTCCCGACGA GTTGGGATCG TAAGTGCCAG ATCAAGCCTC TCTGAATATT CTTGGTCATC	2040
ATCTTAACCT CCTAGGCTCA TTCATCCATG GTGCGCAATA GGAGCTTATG GAGCAAAAGA	2100
AGCAGGCCGT GCTTGGCTCA GCTTCCGATC AGGCTGTTGA TAAAAAGGAT GTTCAAGGTC	2160
GGGATATCCT AAGTCTCCTA GGTTAGTAAC GTTTTTTAAAC GTATATACAG AGCGGCGACA	2220
TTCTTTCCCT GACAACTGTC AACATGCTCG TTACTAGTGA GAGCAAACAT CGCCGCCAAC	2280
CTGCCTGAAT CTCAAAAGCT GTCCGATGAG GAGGTACTCG CTCAGATCAG TAACCTGTTA	2340
TTTGCTGGAT ATGAGTGTGT ATCCTTTCCC CTCTCTATCC TTAGCTGATT AAAAGCACTA	2400
ATAGAGGTCT TTATGTTTCC TGTTTGATCA GAACTTCTTC GACAGTCTTG ACATGGATGT	2460
TTCACCGACT CTCAGAAGAC AAAGCCGTTT AGGATAAACT TCGAGAAGAA ATTTGTCAGA	2520
TCGACACGGA TATGCCTACG CTGTGAGGAT GTTTTTGATG CTAAATTACT TCTTCTTGCA	2580
AATGACTAAA ACGGCCTTCC ATTCTTGATC CATTTTAGAG ACGAACTTAA TGCGTTGCCT	2640
TATCTCGAAG CGGTTGGTTC TCGATTCTTG GTCTTGTCTT CCAAATACAA TACGGATTAT	2700
TGCTCATCTG ATTTGCGTCT ACGGGCTGTG GAATTTAACT AGTTTGTTAA GGAGTCTCTT	2760
CGTCTAGACC CTCCTAGTCC GTATGCTAAC CGTGAATGCT TAAAGGATGA AGACGTATGT	2820
TGGCTTCATC ACGCATAATT TTCATTTTCA ATTCCCTTGT ACATACGCAT ACAGGCTGAC	2880
CGAGCTCAAA TTCCGGCTTC CTCTTCTGTG CTTCTTTTTT TGGCCTTTCT TATCTTCATT	2940
CTTCAACCAA AATTTGTCAC AGTTCATCCC ACTTGCCGAG CCTGTCATTG GTCGAGATGG	3000
GTCGGTCATC AACGAGGTCC GGATCACGAA AGGAACGATG GTCATGCTTC GTAAGTTTTT	3060
CTTTATTTCA TCTCGTCCAT GAAATAGTTT CTGATAGACG CGGACCAATT CAGCGTTGTT	3120
CAACATCAAT CGTTCAAAGT TCATTTATGG AGAAGATGCA GAAGAATTCA GGTACAATTC	3180
GTTTTCTTTT AAAAGCCAAT CGGTTTCGTA TCGTAATTGA CCGGGCTCTC TTTTAATTTT	3240
TCGAAAGACC GGAGAGGTGG CTTGAGGACG TAACAGACTC GCTCAACAGT ATTGAAGCAC	3300
CCTATGGACA CCAGGCGAGC TGTATGTTTT ATTGATTTTA TCTTTGTGAA TTTTGCAAAA	3360
CGTTGAACTT CGCGCTTCCC TTGTTGTTGA AATCCCAGTT ATCTCTGGAC CCAGAGCTTG	3420
CTTGTAAGTT TCTTCTCATC TGGCGCCTTA GCAGTATCCG ATCAGCCATC TAGTTCTTTG	3480

-56-

TACGATTGTT TCTGACTCTC TCGACTTTCG CAGTGGTTGG CGATTTGCTG TCGCCGAGAT 3540
 GAAGGCCTTC TTGTTTGTC A CTCTCCGTCG GGTCCAGTTC GAGCCCATCA TCTCTCATCC 3600
 AGAGTACGAG CACATCACCT TGATCATTTC CCGTCCTCGA ATCGTTGGTA GAGAGAAGGA 3660
 GGGGTACCAG ATGCGTTTGC AGGTCAAGCC GGTCGAATGA GTTGATTCTT CATATGTTAA 3720
 GAGAAGTTCT ATATCTGAGA ATGTGTGACT AGGACAATGC CTTCTTTGTA TCGATTTGTT 3780
 TCTCATACCC GGGCAGGCGC TATGACTTCT ACGTCGTCTA TCGTCGCTCT GGACTCTCTT 3840
 CTTACCCTAT ATATTATTCC ATCCGTCTGT ATATTTGTCT ATCACGACGT CTGTGTCGTC 3900
 AACTCAATAT TCAGCCTCTT CATGCTTCTG TGTCTCCATA GATGTGATCT TCATGTTTGT 3960
 CGACTGCAG 3969

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial Sequence: Sense primer for expression of AST gene in E.coli"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTTCAAAGTT CATTTATGGA 20

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: Antisense primer for expression of AST gene in E.coli."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGATCCTCAG TGGTGGTGGT GGTGGTGTTC GACCGGCTTG ACCTGCA 47

-57-

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: 5' sense primer for expression of modified AST gene in E.coli."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CATATGCACC ACCACCACCA CCACCTGTAT AACCTTCAGG GGCCC

45

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: 5' antisense primer for expression of modified AST gene in E. coli."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTAACAACAC CATCTCCGGT

20

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: 3' antisense primer for expression of modified AST gene in E.coli"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

-58-

GGATCCTCAA CTCATTCGAC CGGCTT

26

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence:
genome walking primer for cloning of AST gene."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TAGAGAGAAG GAGGGGTACC AGATGC

26

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence:
Antisense primer for cloning of terminator region of AST gene."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CCCCGGATTG TGGAGAAACT

20

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: sense
primer for cloning genomic AST gene."

-59-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATGTTTCATCT TGGTCTTGCT

20

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
antisense primer for cloning genomic AST gene."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ACGTAGAAGT CATAGCGCCT

20

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence: sense
primer for RT-PCR of AST gene."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTTGACTCAA GGATTAGCAG

20

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
antisense primer for RT-PCR of AST gene."

-60-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TGTCTTCTGA GAGTCGGTGA

20

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
degenerate sense primer for cloning of TPI gene."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

MGNACNTTYT TYGTNGGNGG NAAY

24

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
degenerate antisense primer for cloning of TPI gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GCNCCNCCNA CNARRAANCC RTCNACRTC

29

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
primary walking primer for cloning of TPI terminator"

-61-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GCTTACCTCG CTTCCAACGT TTCCCAG

27

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
nested walking primer for cloning of TPI terminator"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGATCTGTCT CTGCCTCCAA CTGCAAG

27

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
primary walking primer for cloning of TPI promoter"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GGGTCAATGT CGGCAGCGAG AAGCCCA

27

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid

-62-

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Artificial sequence:
nested walking primer for cloning of TPI promoter"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ATGTACTCGG TAGCACTGAT CAAGTAG

27

- (2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Artificial sequence: sense
primer for construction of TPI promoter cassette"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GCGGCCGCAT CCGTCTCGCC ATCAGTCT

28

- (2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Artificial sequence:
antisense primer for construction of TPI promoter cassette"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CCTGCAGGTC TAGAGATGAA TAAATATAAA GAGT

34

-63-

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: sense primer for construction of TPI terminator cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CCTGCAGGTA AATATATCCA GGGATTAACC CCTA

34

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: antisense primer for construction of TPI terminator cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GGTACCCGTG CGCAGTCGAC CGAGACAT

28

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Artificial sequence: degenerate sense primer for cloning of AMY gene"

-64-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GAYTAYATHC ARGGNATGGG NTTYRMNGCN ATHTG

35

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
degenerate antisense primer for cloning of AMY gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TGYTCNGTNC CRTARTADAT DATNGGDATN CCRTC

35

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence: sense
primer for construction of a partial AMY cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CCGCGGCATT GATACCTCTA CCCCCT

26

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs

-65-

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
antisense primer for construction of a partial AMY cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GCGGCCGCCT GCAATCCTGG ATCCACCG

28

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence: sense
primer for construction of AST cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

TCTAGAATGT TCATCTTGGT CTTGCTCA

28

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence:
antisense primer for construction of AST cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CCTGCAGGTC ATTCGACCGG CTTGACCT

28

-66-

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Artificial sequence: sense primer for confirmation of integration at AMY locus by PCR analysis"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CTCTCCTGTT CACAAAAACA

20

WHAT IS CLAIMED IS:

1. An isolated DNA comprising a nucleotide sequence coding for an enzyme having astaxanthin synthase activity catalyzing the reaction from beta-carotene to astaxanthin, characterized in that said nucleotide sequence is:
 - (a) selected from the group consisting of the nucleotide sequences represented in SEQ ID NO:2 and SEQ ID NO:3;
 - (b) a nucleotide sequence which, because of the degeneracy of the genetic code, encodes an astaxanthin synthase having the same amino acid sequence as that encoded by the nucleotide sequences as shown in SEQ ID NO:2 or SEQ ID NO:3; or
 - (c) a nucleotide sequence which hybridizes to the complement of the nucleotide sequences of SEQ ID NO:2 or SEQ ID NO:3 coding for an astaxanthin synthase under the medium stringent hybridizing conditions 50% (V/V) formamide, 5 x SSC, 2% (W/V) blocking reagent, 0.1% (W/V) N-lauroylsarcosine and 0.3% (W/V) SDS at a temperature of 42° over night.
2. A vector or plasmid comprising the isolated DNA as claimed in claim 1.
3. A host cell transformed or transfected by the isolated DNA as claimed in claim 1 or a vector or plasmid as claimed in claim 2.
4. A polypeptide encoded by the isolated DNA as claimed in any one of claims 1 to 3.
5. A process for the production of a polypeptide having an activity of astaxanthin synthase which comprises culturing a transformed host cell as claimed in claim 3 under the conditions conducive to the production of said enzyme.
6. A process for the biological production of astaxanthin which comprises introducing one or more isolated DNAs as claimed in claim 1 into an appropriate host organism, cultivating the obtained organism under the condition conducive to the production of astaxanthin and recovering astaxanthin from the culture.

FIG.1 Biosynthetic pathway from GGPP to astaxanthin in the bacterial sytem which produces astaxanthin

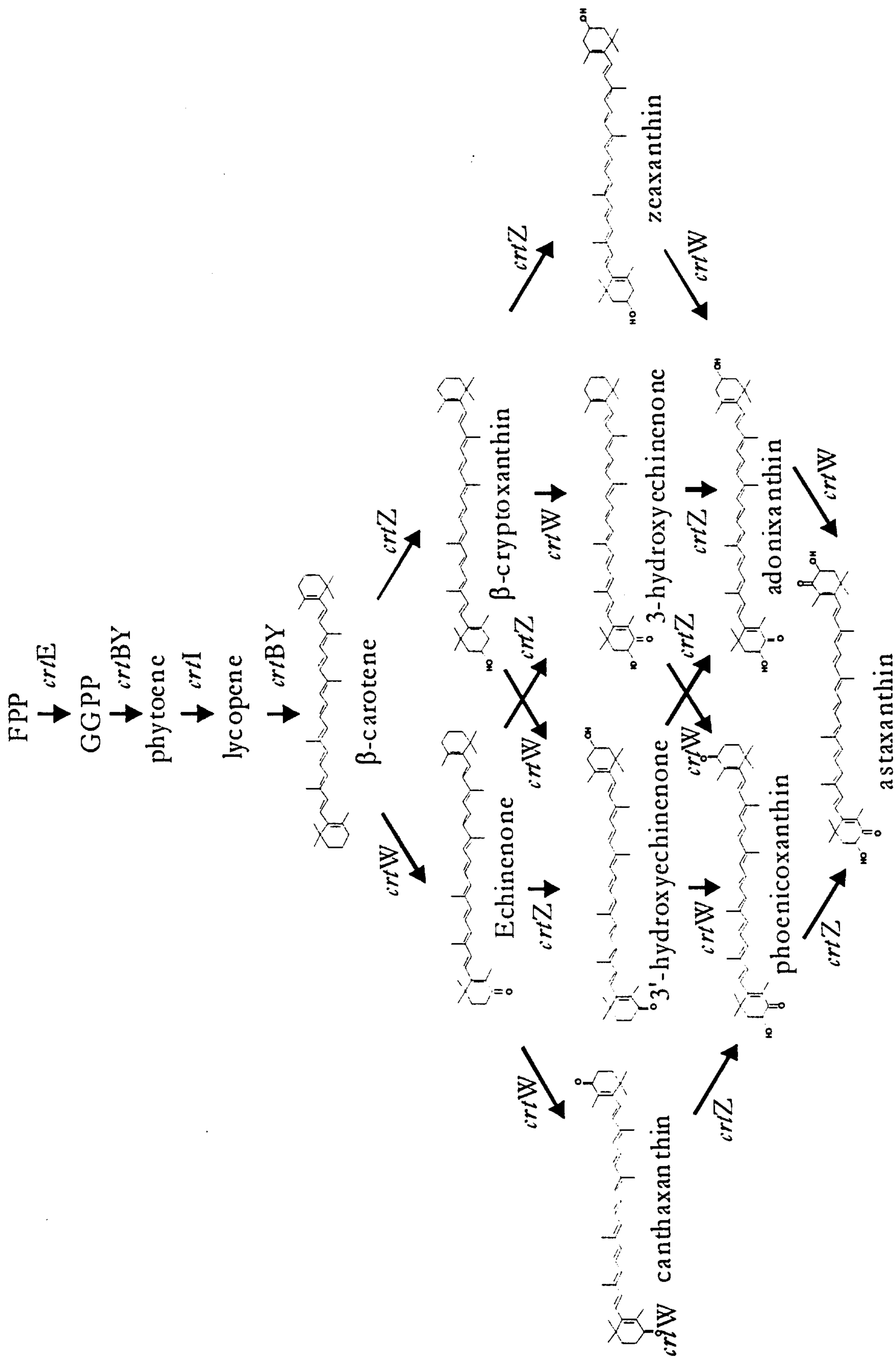
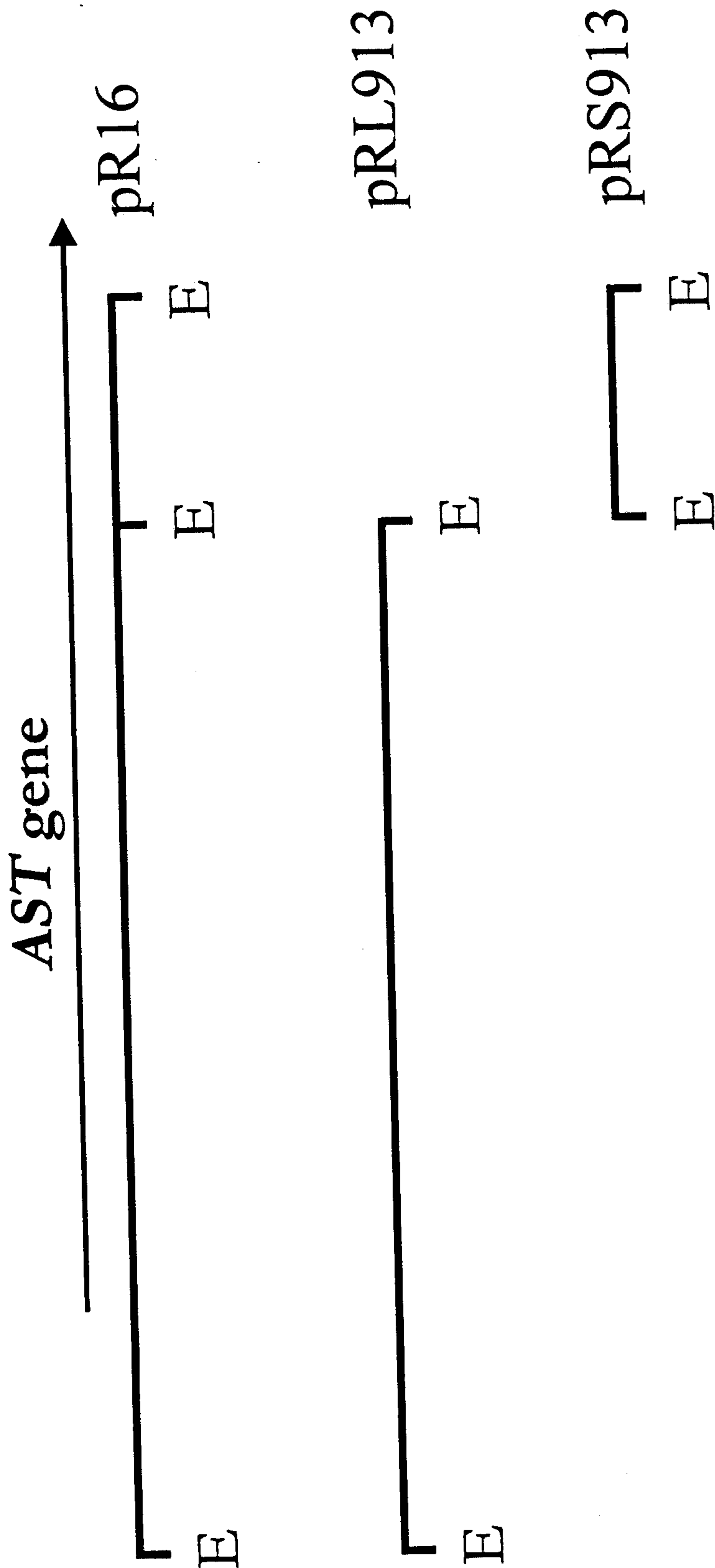


FIG. 2 Restriction map of DNA fragments which were used for functional analysis of *AST* gene from *Phaffia rhodozyma*

Vector backbone of these plasmids was pUC-G418. Restriction enzyme indicated as E is *EcoRI*.



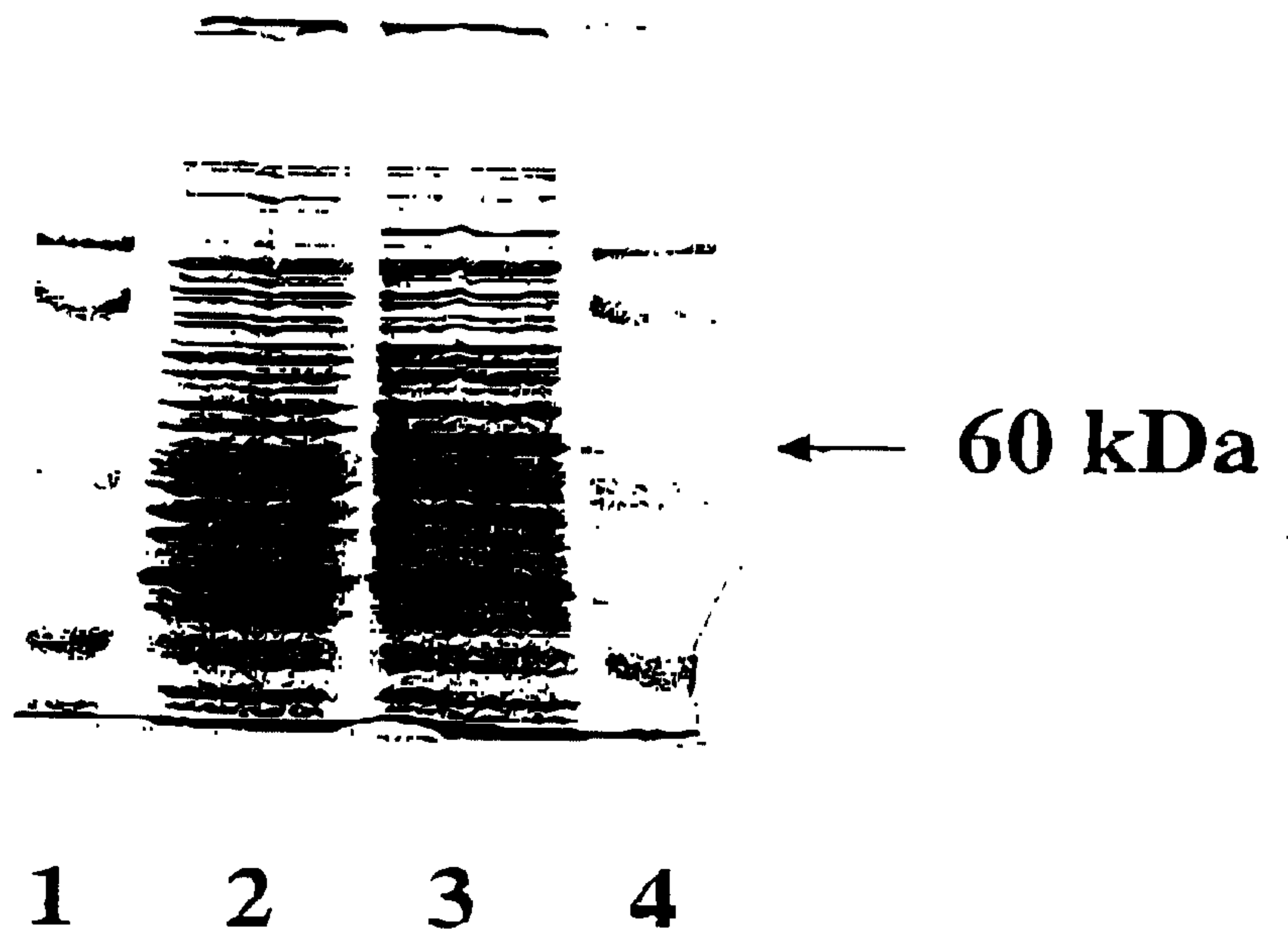


FIG. 3 Expression of *AST* gene in *E. coli*

Biosynthetic pathway from GGPP to astaxanthin in the bacterial system which produces astaxanthin

