(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113774034 A
(43)申请公布日 2021.12.10

(21)申请号 202110711179.1

(22)申请日 2013.08.27

(30)优先权数据
12198352.2 2012.12.20 EP
61/830,273 2013.06.03 US

(62)分案原申请数据
201380066965.4 2013.08.27

(71)申请人 帝斯曼知识产权资产管理有限公司
地址 荷兰海尔伦

(72)发明人 C·法雷尔 M·马约尔加
B·谢弗勒

(74)专利代理机构 北京东方亿思知识产权代理有限责任公司 11258
代理人 肖善强

(51)Int.Cl.
C12N 9/02(2006.01)
C12N 15/53(2006.01)
C12N 1/19(2006.01)
C12P 23/00(2006.01)
C12R 1/645(2006.01)

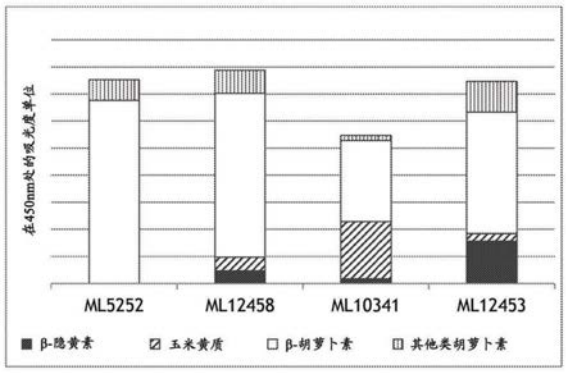
权利要求书1页 说明书14页
序列表4页 附图2页

(54)发明名称

胡萝卜素羟化酶及其用于产生类胡萝卜素的用途

(57)摘要

本发明涉及胡萝卜素羟化酶及其用于产生类胡萝卜素的用途，具体地涉及新的羟化酶、编码它的核酸序列、包含此序列的表达构建体和载体、用其转化的微生物和用于类异戊二烯的微生物羟化的方法，例如，用于将β-胡萝卜素转变成β-隐黄素和玉米黄质、将角黄素转变成虾青素的方法。



CN 113774034 A

1.一种具有类异戊二烯的微生物羟化的羟化酶活性的多肽,其中所述多肽包含SEQ ID NO:2的序列。

2.权利要求1的多肽,其中所述蛋白具有以下酶活性:

(a)用于将β-胡萝卜素转变成β-隐黄素和/或玉米黄质的酶活性,其中基于类胡萝卜素的总量,隐黄素的百分比是至少21%,和/或,

(b)将角黄素转变成金盏花红素和/或虾青素的酶活性,其中基于类胡萝卜素的总量,虾青素的百分比是至少12%。

3.一种编码权利要求1或2的多肽的分离的核酸。

4.权利要求3的分离的核酸,其由SEQ ID NO:1或3描述的序列组成。

5.一种核酸构建体或表达载体,其包含与一个或多个调控序列可操作连接的权利要求3或4的多核苷酸,所述调控序列指导所述多肽在表达宿主细胞中的产生。

6.一种转化的微生物,其中表达权利要求5的核酸构建体。

7.权利要求6的转化的微生物,其类胡萝卜素代谢不同于野生型。

8.权利要求7的转化的微生物,其中所述微生物是含油株(oleaginous strain)。

9.权利要求8的转化的微生物,其中所述含油株是解脂耶氏酵母(Yarrowia lipolytica)的菌株。

10.一种用于产生权利要求6的转化的微生物的方法,其包括引入由SEQ ID NO:1或3中描述的序列组成的核酸或核酸构建体,其与一个或多个调节信号功能性连接。

11.一种用于制备类胡萝卜素或类胡萝卜素衍生物的方法,其包括在权利要求1的多肽存在下将β-紫罗兰酮转变成羟基-β-紫罗兰酮和/或将4-酮-β-紫罗兰酮转变成3-羟基-4-酮-β-紫罗兰酮结构元件。

12.权利要求11的方法,包括如下步骤:

a)在有助于类胡萝卜素产生的条件下培养根据权利要求6至9中任一项的转化的微生物;和

b)回收所述类胡萝卜素或类胡萝卜素衍生物。

13.一种包含根据权利要求1或2的获自雨生红球藻(Haematococcus pluvialis)的多肽的组合物。

胡萝卜素羟化酶及其用于产生类胡萝卜素的用途

[0001] 本申请是申请日为2013年8月27日、名称为“胡萝卜素羟化酶及其用 于产生类胡萝卜素的用途”的中国发明专利申请201380066965.4的分案申请。

[0002] 发明背景

[0003] 本发明涉及新的羟化酶、编码它的核酸序列、包含此序列的表达构建体和 载体、用其转化的微生物、用于类异戊二烯的微生物羟化的方法，例如，用于 将β-胡萝卜素转变成β-隐黄素或玉米黄质以及将角黄素转变成虾青素的方法。

[0004] 类胡萝卜素是颜色在黄色至红色范围内的有机色素，其由一些生物包括 光合生物(例如，植物、藻类、蓝藻)和一些真菌天然产生。

[0005] 类胡萝卜素如叶黄素、玉米黄质或虾青素是在人和牲畜饮食中作为色素 物质和维生素A衍生物的前体的非常重要的添加剂。此外，类胡萝卜素具有 促进健康的作用如增强免疫应答，而且由于其抗氧化剂特性的原因具有预防 癌症的作用，这使其用作感兴趣的保健品(nutraceutical)。因此用于制备类胡 萝卜素的经济方法和具有增加的类胡萝卜素含量的食品至关重要。用于制备 类胡萝卜素的特别经济的方法是生物技术方法，其利用来自产生类胡萝卜素 的生物的类胡萝卜素生物合成的生物合成基因和蛋白。

[0006] 经由β-隐黄素催化β-胡萝卜素酶法转变成玉米黄质的原核β-胡萝卜素羟 化酶和编码这些蛋白的基因已知来自噬夏孢欧文氏菌(Erwinia uredovora) (Misawa等，J.of Bacteriology 1990,6704-6712;EP 393690B1)、草生欧文氏菌 (Erwinia herbicola)(WO 9113078)、橙黄土壤杆菌(Agrobacterium aurantiacum) (Misawa等，J.of Bacteriology 1995,6575-6584;EP 735 137A1)、产碱菌属菌种 (Alcaligenes sp.)PC-1(EP 735 137A1)、黄杆菌属菌种(Flavobacterium sp.)菌株 R1534(Pasamontes等，Gene 1997,185,35-41;EP 747483A2)和来自蓝藻集胞 藻属(Cyanobacterium Synechocystis)PCC6803 (Masamoto等，Plant Cell Physiol. 1998,39(5),560-564)。

[0007] 还已知来自橙黄土壤杆菌、产碱菌属和噬夏孢欧文氏菌的原核β-胡萝卜 素羟化酶还能够将角黄素经由金盏花红素(adonirubin)转变成虾青素(Misawa 等，J.of Bacteriology 1995,6575-6584;Fraser等，J.Biol.Chem.1997,272, 6128-6135)。

[0008] 已知来自真核来源的三种植物β-胡萝卜素羟化酶可催化β-胡萝卜素经由 β-隐黄素酶法转变成玉米黄质。相应的cDNA已从拟南芥(Arabidopsis thaliana) (Cunningham等，J.Biol.Chem.1996,271,24349-24352,WO 9736998)和辣椒 (Capsicum annuum L.)中分离。

[0009] 真核来源的基因相对于原核基因具有优势，其在更高等的转基因生物如植 物中表达更佳。尽管如此，仍然需要改进和增加类胡萝卜素生产力用于通过将 真核核酸并入生物体而制备类胡萝卜素衍生物或具有增加的类胡萝卜素含量 的食品的方法。

[0010] 此外，现有技术中适当的真核β-胡萝卜素羟化酶具有它们仅有窄的底物 范围的缺点使得存在不能由羟化酶转变且可对羟化酶发挥抑制效果的代谢产 物的积累。

[0011] 发明概述

[0012] 本发明的目标之一是改正现有技术的所述缺陷并提供具有改进特性的真 核胡萝

卜素羟化酶。

[0013] 发明人发现通过包含氨基酸序列SEQ ID NO:2的具有羟化类异戊二烯 (例如,β-胡萝卜素或角黄素)酶活性的蛋白可令人惊讶地实现上述目标。

[0014] 因此,本发明涉及包含氨基酸序列SEQ ID NO:2的蛋白或多肽,或源自 此序列通过氨基酸的取代、插入或缺失并与序列SEQ ID NO:2在氨基酸水平 具有至少50%同源性的序列。

[0015] 特别地,根据本发明的羟化酶如果在所选的或能够产生特定类胡萝卜素 (例如玉米黄质或虾青素)的菌株中表达,与未用编码根据本发明的蛋白或多肽 的基因转化的菌株相比,可增加这些类胡萝卜素的水平。

[0016] 本发明还涉及编码本发明的多肽的分离的多核苷酸、包含所述多核苷酸 的核酸构建体、重组表达载体和重组宿主细胞,并涉及产生所述多肽的方法。

[0017] 本发明还提供用于类胡萝卜素生物产生的改进系统。在一个优选的实例中, 本发明提供产生一种或多种类胡萝卜素的含油真菌(包括,例如,酵母)。本发明 还提供构建所述酵母和真菌的方法,使用这种酵母和真菌以产生类胡萝卜素的 方法,和制备包含类胡萝卜素的组合物(如食品或饲料添加剂,或营养补充剂) 的方法,使用所述含油酵母或真菌中产生的类胡萝卜素的方法。特别地,本发 明提供用于生成包含编码本发明多肽的多核苷酸的酵母和真菌的系统和方法。

[0018] 序列表概述

[0019] SEQ ID NO:1是编码来自雨生红球藻(Haematococcus pluvialis)的低频胡 萝卜素羟化酶的未优化DNA序列。

[0020] SEQ ID NO:2是从SEQ ID NO:1推导的氨基酸序列。

[0021] SEQ ID NO:3是编码来自雨生红球藻的低频胡萝卜素羟化酶的DNA序列, 如对于在解脂耶氏酵母/解脂西洋蓍霉(Yarrowia lipolytica)中表达而优化的。

[0022] SEQ ID NO:4是编码来自粉尘克罗诺杆菌(Cronobacter pulveris)(之前为 粉尘肠杆菌(Enterobacter pulveris)(Ep)的胡萝卜素羟化酶的DNA序列,如对 于在解脂耶氏酵母中表达而优化的。

[0023] SEQ ID NO:5是编码来自肠杆菌科(Enterobacteriaceae)细菌DC404(Dc) 的胡萝卜素羟化酶的DNA序列,如对于在解脂耶氏酵母中表达而优化的。

[0024] 定义

[0025] 分离的多肽:术语“分离的多肽”意指相对于天然发现的多肽由人工进 行修饰的多肽。在一个方面中,所述多肽为至少1%纯,例如,至少5%纯、 至少10%纯、至少20%纯、至少40%纯、至少60%纯、至少80%纯,和至少 90%纯,如通过SDS-PAGE确定的。

[0026] 基本上纯的多肽:术语“基本上纯的多肽”意指制备物,其包含按重量 计至多10%、至多8%、至多6%、至多5%、至多4%、至多3%、至多2%、 至多1%和至多0.5%的与所述制备物天然或重组结合的其他多肽材料。优选 地,所述多肽按制备物中存在的全部多肽材料的重量计为至少92%纯,例如, 至少94%纯、至少95%纯、至少96%纯、至少97%纯、至少98%纯、至少99% 纯、至少99.5%纯,和100%纯。本发明的多肽优选为基本上纯的形式。这可 例如通过由已知的重组方法或由经典的纯化方法制备多肽来实现。

[0027] 序列同一性:两个氨基酸序列之间或两个核苷酸序列之间的相关性通过 参数“序

列同一性”描述。

[0028] 就本发明而言,两个氨基酸序列之间的序列同一性的程度使用如在 EMBOSS包(EMBOSS:The European Molecular Biology Open Software Suite, Rice等,Trends Genet.2000,16:276-277)的Needle程序,优选3.0.0或以后的版 本中执行的Needleman-Wunsch算法(Needleman和Wunsch,1970,J.Mol.Biol. 48:443-453)来确定。使用的任项参数为缺口开放罚分10、缺口延伸罚分0.5 和EBLOSUM62(BLOSUM62的EMBOSS版本)取代矩阵。使用标记“最长同 一性”的Needle输出(使用-nobrief选项获得)作为百分比同一性并计算如下:

[0029] (相同的残基x100)/(比对长度-比对中缺口的总数)

[0030] 就本发明而言,两个脱氧核糖核酸序列之间的序列同一性的程度使用如 在EMBOSS包(EMBOSS:The European Molecular Biology Open Software Suite, Rice等,2000,见上文)的Needle程序,优选3.0.0或以后的版本中执行的 Needleman-Wunsch算法(Needleman和Wunsch,1970,见上文)来确定。使用 的任选参数为缺口开放罚分10、缺口延伸罚分0.5和EDNAFULL(NCBI NUC4.4的EMBOSS版本)取代矩阵。使用标记“最长同一性”的Needle输 出(使用-nobrief选项获得)作为百分比同一性并计算如下:

[0031] (相同的脱氧核糖核酸x100)/(比对长度-比对中缺口的总数)

[0032] 片段:术语“片段”意指在成熟多肽的氨基和/或羧基末端缺失一个或多 个(几个)氨基酸的多肽,其中所述片段具有羟化酶活性。

[0033] 等位变体:术语“等位变体”意指占据相同染色体基因座的基因的两种 或多种可替换形式中的任何。等位变异通过突变自然出现,且可导致种群内 的多态性。基因突变可为沉默的(在编码的多肽中无变化)或可编码具有改变的 氨基酸序列的多肽。多肽的等位变体是由基因的等位变体编码的多肽。

[0034] 分离的多核苷酸:术语“分离的多核苷酸”意指相对于天然发现的多核 苷酸由人工进行修饰的多核苷酸。在一个方面中,所述分离的多核苷酸为至 少1%纯,例如,至少5%纯、至少10%纯、至少20%纯、至少40%纯、至少 60%纯、至少80%纯、至少90%纯和至少95%纯,如通过琼脂糖凝胶电泳确 定的。所述多核苷酸可以是基因组、cDNA、RNA、半合成、合成来源的, 或其任何组合。

[0035] 基本上纯的多核苷酸:术语“基本上纯的多核苷酸”意指无其他外来或 不想要的核苷酸且其处于适用于基因工程多肽产生系统中的形式的多核苷酸 制备物。因此,基本上纯的多核苷酸包含按重量计至多10%、至多8%、至多 6%、至多5%、至多4%、至多3%、至多2%、至多1%和至多0.5%的与其 天然或重组结合的其他多核苷酸材料。然而,基本上纯的多核苷酸可包括天 然存在的5’和3’非翻译区,如启动子和终止子。优选地,所述多核苷酸为按 重量计至少90%纯,例如,至少92%纯、至少94%纯、至少95%纯、至少96% 纯、至少97%纯、至少98%纯、至少99%纯,和至少99.5%纯。本发明的多 核苷酸优选为基本上纯的形式。

[0036] 编码序列:术语“编码序列”意指直接指定多肽的氨基酸序列的多核苷 酸。编码序列的边界通常由开放阅读框确定,其通常以ATG起始密码子或可 替换的起始密码子如CTG和TTG开始并以终止密码子如TAA、TAG和TGA 结束。编码序列可为DNA、cDNA、合成的或重组的多核苷酸。

[0037] cDNA:术语“cDNA”意指可由自真核细胞获得的成熟、剪接的mRNA 分子通过反转录制备的DNA分子。cDNA缺乏可存在于相应基因组DNA中 的内含子序列。原始的初级RNA转录物是mRNA前体,其通过一系列步骤, 包括剪接进行加工,然后作为成熟的剪接的mRNA出现。

[0038] 核酸构建体:术语“核酸构建体”意指单链或双链核酸分子,其分离自 天然存在的基因,或经修饰以自然界中本不存在的方式包含核酸的区段,或 其可为合成的。当核酸构建体包含表达本发明的编码序列所需的调控序列时, 术语核酸构建体与术语“表达盒”同义。

[0039] 调控序列:术语“调控序列”意指用于表达编码本发明多肽的多核苷酸 所需的全部组件。每个调控序列对于编码所述多肽的多核苷酸可为天然或外 来的,或调控序列彼此间可为天然或外来的。所述调控序列包括但不限于前 导序列、多聚腺苷酸化序列、前肽序列、启动子、信号肽序列和转录终止子。 最少的情况,所述调控序列包括启动子,和转录和翻译终止信号。出于引入 促进调控序列与编码多肽的多核苷酸编码区连接的特定限制位点的目的,所 述调控序列可与接头一起提供。

[0040] 可操作连接:术语“可操作连接”意指构型,其中将调控序列置于相对 于多核苷酸的编码序列适当位置使得调控序列指导编码序列的表达。

[0041] 表达:术语“表达”包括涉及多肽产生的任何步骤,其包括但不限于转 录、转录后修饰、翻译、翻译后修饰和分泌。

[0042] 表达载体:术语“表达载体”意指线性或环状DNA分子,其包含编码多 肽的多核苷酸并与提供其表达的额外核苷酸可操作连接。

[0043] 宿主细胞:术语“宿主细胞”意指对用包含本发明多核苷酸的核酸构建体 或表达载体的转化、转染、转导等易感的任何细胞类型。术语“宿主细胞”涵 盖亲代细胞的任何后代,其由于在复制过程中发生突变而与亲代细胞不同。

[0044] 变体:术语“变体”意指具有羟化酶活性的多肽,其在一个或多个(几个) 位置包含改变(即一个或多个(几个)氨基酸残基的取代、插入和/或缺失)的。取 代意指用不同氨基酸替代占据位置的氨基酸,缺失意指去除占据位置的氨基 酸的,而插入意指在占据位置的氨基酸邻近添加1-3个氨基酸。

附图说明

[0045] 图1示出了根据本发明向解脂耶氏酵母角黄素产生株ML6804引入三种 胡萝卜素羟化酶基因后的类胡萝卜素生产情况;

[0046] 图2示出了根据本发明向解脂耶氏酵母虾青素产生株引入雨生红球藻LF 胡萝卜素羟化酶后的类胡萝卜素生产情况;

[0047] 图3示出了根据本发明向解脂耶氏酵母β-胡萝卜素产生株引入雨生红球 藻LF羟化酶后的类胡萝卜素生产情况。

[0048] 发明详述

[0049] 下文的胡萝卜素羟化酶意指根据本发明的蛋白或多肽,即具有例如将β- 胡萝卜素转变成玉米黄质或将角黄素转变成虾青素的酶活性的蛋白,其包含 氨基酸序列SEQ ID NO:2,或源自此序列通过氨基酸的取代、插入或缺失并 与序列SEQ ID NO:2在氨基酸水平具有至少50%同源性的序列。

[0050] SEQ ID NO:2中描述的氨基酸序列源自SEQ ID NO:1中描述的cDNA序 列的翻译。SEQ ID NO:1的优化的序列示于SEQ ID NO:3中。

[0051] 取代意指通过一个或多个氨基酸被一个或多个氨基酸替代。所述替代优 选为称为保守的那些,其中替代的氨基酸具有与原始氨基酸相似的性质,例 如,用Asp替代Glu、用Asn替代Gln、用Ile替代Val、用Ile替代Leu、用 Thr替代Ser。

[0052] 缺失是用直接的键/连接基(linkage)替代氨基酸。优选的用于缺失的位置 是多肽的末端和单独的蛋白结构域之间的连接。

[0053] 插入是向多肽链中引入氨基酸,形式上存在用一个或多个氨基酸替代直 接的键/连接基。

[0054] 两个蛋白间的同源性意指每个蛋白全长上氨基酸的同一性,其用计算机 程序GAP(UWGCG,University of Wisconsin,Genetic Computer Group,program algorithm of Needleman和Wunsch(J.Mol.Biol.1970,48:443-453)的辅助通过 比较进行计算,设置如下参数:

[0055] 缺口权重:12

[0056] 长度权重:4

[0057] 平均匹配:2.912

[0058] 平均错配:-2.003

[0059] 与序列SEQ ID NO:2在氨基酸水平具有至少50%同源性的蛋白意指蛋白 在其序列与序列SEQ ID NO:2在使用上述程序算法以上组参数的比较中具有 至少50%,优选60%,特别优选70%的同一性。

[0060] 根据本发明的蛋白能够催化β-紫罗兰酮结构元件转变成3-羟基-β-紫罗兰 酮结构元件,如将β-胡萝卜素转变成玉米黄质、将β-胡萝卜素转化成β-隐黄 素、将β-隐黄素转变成玉米黄质、将海胆酮(echinenone)转变成3'-羟基海胆 酮、将3-羟基海胆酮转变成侧金盏花黄质(4-酮玉米黄质)、将α-胡萝卜素转变 成α-隐黄素或其他具有多至40C原子的化学化合物,并包括β-紫罗兰酮环转 变成相应的3-羟基-β-紫罗兰酮化合物或4-酮-β-紫罗兰酮结构元件转变成3- 羟基-4-酮-β-紫罗兰酮结构元件,如角黄素转变成虾青素、角黄素转变成芬尼 黄质(phoenicoxanthin)(金盏花红素(adonirubin))、芬尼黄质(金盏花红素)转变 成虾青素、海胆酮转变成3-羟基海胆酮、3'-羟基海胆酮转变成侧金盏花黄质 (4-酮玉米黄质)或其他具有多至40C原子的化学化合物并包含4-酮-β-紫罗兰 酮环转变成相应的3-羟基-4-酮-β-紫罗兰酮化合物。

[0061] 可如下文描述通过编码来自天然或遗传操纵的生物的胡萝卜素羟化酶的 适当核酸的基因表达来制备这些蛋白。

[0062] 本发明进一步涉及用于制备类胡萝卜素和类胡萝卜素衍生物的方法,其 包括在根据本发明的蛋白的存在下将β-紫罗兰酮结构元件转变成3-羟基-β-紫 罗兰酮结构元件和/或将4-酮-β-紫罗兰酮结构元件转变成3-羟基-4-酮-β-紫罗 兰酮结构元件。

[0063] 类胡萝卜素和类胡萝卜素衍生物为例如玉米黄质、β-隐黄素、3'-羟基海 胆酮、3-羟基海胆酮、侧金盏花黄质(4-酮玉米黄质)、虾青素、芬尼黄质(金盏 花红素)、α-隐黄素、β-隐黄素或叶黄素及其具有多至40C原子并在分子中包 含至少一个3-羟基-β-紫罗兰酮或至少一个3-羟基-4-酮-β-紫罗兰酮结构元件 的衍生物,如例如,3-羟基-6-乙烯基-β-紫罗

兰酮、3-羟基-4-酮-6-乙烯基-β- 紫罗兰酮、3-羟基视黄醇、3-羟基-4-酮视黄醇、3-羟基视黄醛、3-羟基-4-酮视 黄醛、3-羟基视黄酸、3-羟基-4-酮视黄酸。

[0064] 在根据本发明的方法中，在根据本发明的蛋白的存在下，存在将β-紫罗 兰酮结构元件转变成3-羟基-β-紫罗兰酮结构元件，如β-胡萝卜素经由β-隐黄 素转变成玉米黄质、β-胡萝卜素转变成β-隐黄素、β-隐黄素转变成玉米黄质、海胆酮转变成3'-羟基海胆酮或3-羟基海胆酮、3'-羟基海胆酮或3-羟基海胆 酮转变成侧金盏花黄质(4-酮玉米黄质)、α-胡萝卜素转变成玉米次黄质 (zeinoxanthin)或α-隐黄素或具有多至40C原子的化学化合物，并含有β-紫罗 兰酮环转变成相应的3-羟基-β-紫罗兰酮化合物或4-酮-β-紫罗兰酮结构元件 转变成3-羟基-4-酮-β-紫罗兰酮结构元件，如角黄素经由金盏花红素转变成虾 青素、角黄素转变成芬尼黄质(金盏花红素)、芬尼黄质(金盏花红素)转变成虾 青素、海胆酮转变成3-羟基海胆酮、3'-羟基海胆酮转变成侧金盏花黄质(4- 酮玉米黄质)或具有多至40C原子的化学化合物，并包含4-酮-β-紫罗兰酮环 转变成相应的3-羟基-4-酮-β-紫罗兰酮化合物。

[0065] 本发明还涉及用于编码根据本发明的羟化酶的核酸序列。优选的核酸具 有SEQ ID NO:1的序列和SEQ ID NO:3的序列。

[0066] 而且本发明涉及根据SEQ ID NO:1和SQ ID NO:3的核酸功能类似物，其通过进一步编码具有合意的特异性的羟化酶的单个或多个核苷酸的添加、取代、插入和/或缺失而获得。

[0067] 本发明还涵盖那些核酸序列，其包含所谓沉默突变或其与依照特定来源 或宿主生物的密码子选择特定提及的序列和所述核酸序列的天然存在的变体 相比而经过修饰。

[0068] 本发明还涵盖由遗传密码的简并而获得的核酸序列的修饰(即，相应氨基 酸序列中无任何变化)或保守的核苷酸取代(即，用相同电荷、大小、极性和/ 或溶解性的另一氨基酸替代相应的氨基酸)，和通过核苷酸添加、插入、倒位 或缺失修饰的序列，该序列编码根据本发明的具有“修饰的底物概貌”的羟 化酶，和相应的互补序列。

[0069] 本发明进一步涉及包含在调节核酸序列的遗传控制下的根据本发明的核 酸序列的表达构建体；和包含至少一个这些表达构建体的载体。

[0070] 优选地，根据本发明的构建体涵盖所述编码序列的5'-上游启动子和3'- 下游终止子，和任选地，其他惯用的调节元件，并在每种情况中与编码序列 可操作连接。可操作连接可理解为意指启动子、编码序列、终止子和(如果合 适的话)其他调节元件以这样的方式顺序安排，使得各调节元件可满足其对于 编码序列表达的意欲的功能。可操作连接的序列的实例为靶向序列，或其他 翻译增强子、增强子、多聚腺苷酸化信号等。进一步的调节元件涵盖选择性 标志、扩增信号、复制起点等。

[0071] 除了人工调节序列，天然调节序列仍可存在于实际的结构基因上游。如 何合意，该天然调节可由遗传修饰关闭，且该基因的表达可增强或降低。然 而，所述基因构建体也可在构建中更加简单，即在结构基因上游未插入额外 的调节信号且并未移除天然启动子及其调节。取而代之，对天然调节序列以 这样的方式进行突变使得调节不再发生且所述基因表达增加或降低。一个或 多个拷贝的核酸序列可存在于所述基因构建体中。

[0072] 合适的启动子的实例为：有利地采用于革兰氏阴性细菌中的cos、tac、trp、 tet、trp-tet、lpp、lac、lpp-lac、laclq、T7、T5、T3、gal、trc、ara、SP6、I-PR 或I-PL启动子；和革

兰氏阳性启动子amy和SPO2、酵母启动子ADC1、MFa、Ac、P-60、CYC1、GAPDH、TEF1或植物启动子CaMV/35S、SSU、OCS、lib4、usp、STLS1、B33、nos或泛素或菜豆蛋白启动子。特别的优选给予使 用诱导型启动子,例如光-特别是温度-诱导型启动子,如PrP1启动子。

[0073] 原则上,可使用具有其调节序列的所有天然启动子。此外,也可以有利 的方式使用合成的启动子。

[0074] 上述调节序列意欲允许蛋白表达和核酸序列的靶向表达。依赖于宿主生 物,其可意指例如,基因仅在诱导发生后表达或过表达,或其立即和/或组成 型地表达和/或过表达。

[0075] 所述调节序列或因子可优选对表达具有阳性效果且以这种方式增加或降 低后者。因此,调节元件的增强可在转录水平通过使用强转录信号如启动子 和/或“增强子”有利地发生。此外,翻译可通过改善例如mRNA的稳定性得 以增强。

[0076] 通过将合适的启动子与合适的羟化酶核苷酸序列和终止子信号或多聚腺 苷酸化信号融合来生成表达盒。为此,使用惯用的重组和克隆技术,例如, 如它们在T.Maniatis,E.F.Fritsch和J.Sambrook,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Laboratory,Cold Spring Harbor,N.Y. (1989)和in T.J.Silhavy,M.L.Berman和L.W.Enquist,Experiments with Gene Fusions,Cold Spring Harbor Laboratory,Cold Spring Harbor,N.Y.(1984)和in Ausubel,F.M.等,Current Protocols in Molecular Biology,Greene Publishing Assoc.and Wiley Interscience(1987)中所述的。

[0077] 对于在合适的宿主生物中表达,所述重组核酸构建体或基因构建体有利 地插入宿主特异性载体,其允许所述宿主中最优的基因表达。载体为本领域 的技术人员熟知且可以在例如“Cloning Vectors”(Pouwels P.H.等编,Elsevier, Amsterdam-New York-Oxford,1985)中找到。载体应理解为意指不仅是质粒, 还有本领域的技术人员已知的所有其他载体如例如,噬菌体、病毒如SV40、CMV、杆状病毒和腺病毒、转座子、IS元件、质粒、粘粒和线性或环状DNA。这些载体可在宿主生物中自主复制或染色体复制。

[0078] 根据本发明的载体允许转化的重组微生物(例如,用至少一个根据本发明 的载体转化)的生成且其可用于产生突变体。将上述根据本发明的重组构建体 有利地引入合适的宿主生物并进行表达。其优选使用本领域的技术人员已知 的通常的克隆和转染方法以在上述表达系统中带来上述核酸的表达。合适的 系统描述于,例如,current protocols in molecular biology,F.Ausubel等编,Wiley Interscience,New York 1997中。

[0079] 合适的宿主生物原则上为允许表达根据本发明的核酸、其等位变体和其 功能等价物或衍生物的所有生物。优选的初始生物是那些天然能够合成类胡 萝卜素的那些。然而,由于引入类胡萝卜素生物合成基因而能够合成类胡萝 卜素的初始生物也是合适的。初始生物意指原核或真核生物如,例如,微生 物或植物。优选的微生物为细菌、酵母、藻类或真菌。

[0080] 因此,本发明进一步涉及用于制备下文描述的遗传修饰的微生物的方法, 其中将根据本发明的胡萝卜素羟化酶基因引入初始生物的基因组。初始生物 意指在根据本发明的遗传修饰前的生物。

[0081] 根据本发明的胡萝卜素羟化酶基因原则上可通过本领域的技术人员已知 的所有方法引入下文所述的初始生物,由此对其进行遗传修饰。

[0082] 通过转化、转染、电穿孔、使用所谓粒子枪或通过显微注射将它们有利 地引入初始生物或其细胞。

[0083] 技术人员可在Sambrook,J.等(1989)Molecular cloning:A laboratory manual,Cold Spring Harbor Laboratory Press,by F.M.Ausubel等(1994) Current protocols in molecular biology,John Wiley and Sons,by D.M.Glover等, DNA Cloning Vol.1,(1995),IRL Press(ISBN 019-963476-9),by Kaiser等 (1994)Methods in Yeast Genetics,Cold Spring HarborLaboratory Press or Guthrie等,Guide to Yeast Genetics and Molecular Biology,Methods in Enzymology,1994,Academic Press的教科书中找到用于微生物的适当方法。

[0084] 可提及的有利方法的实例是那些如通过同源或异源重组引入DNA的那 些,例如使用URA3基因,特别是来自阿舒囊霉属(Ashbya)的URA3基因,如 德国申请DE 19801120.2中所述,和/或通过下文描述的REMI方法(＝"限制酶 介导的整合)。

[0085] REMI技术是基于用相同的限制内切酶在两末端剪切的线性DNA构建体，连同用于该DNA构建体的此限制的限制内切酶一起共转化入生物。所述限制性 内切酶随后切割生物的基因组DNA,所述生物中已引入DNA构建体连同限制 酶。这导致细胞自身修复机制的激活。这些修复机制可修复由内切酶引起的基 因组DNA中的链断裂,且在此过程中还以一定频率将共转化的DNA构建体并 入基因组。通常,在该过程中将限制裂解位点保留在DNA的两端。

[0086] Bolker等(Mol.Gen.Genet.,1995,248:547-552)描述了用于真菌插入诱 变的此技术。Von Schiestl和Petes(Proc.Natl.Acad.Sci.USA,1991,88: 7585-7589)使用了该方法以查明在酵母属中是否存在异源重组。Brown等(Mol. Gen.Genet.1996,251:75-80)描述了用于诱导型报告基因的稳定转化和调节 表达的该方法。

[0087] 可能的是使用REMI方法以在基因组中的转录活性位点处定位根据本发 明的核酸片段或根据本发明的上述胡萝卜素羟化酶基因。

[0088] 可能且是有利的是将核酸连同至少一个报告基因克隆入DNA构建体,所 述构建体被引入基因组。该报告基因应通过生长、荧光、化学或生物发光试 验或通过光度测量使检测轻而易举。可提及的报告基因的实例为抗生素抗性 基因、水解酶基因、荧光蛋白基因、生物发光基因、葡糖苷酶基因、荧光素 酶基因、β-半乳糖苷酶基因、gfp基因、脂肪酶基因、酯酶基因、过氧化物酶 基因、β-内酰胺酶基因、乙酰-、磷酸-或腺苷酸转移酶基因。这些基因使得可 能容易地测量和定量转录活性和由此的基因表达。这意味着可能在基因组中 识别位点,其具有区别为多至因数2的生产力。

[0089] 如果意欲向生物中引入多个基因,如例如,类胡萝卜素生物合成的更多基 因,其可与报告基因一起全部引入单一载体中,或在每种情况中可将各单独的 基因与报告基因引入一个载体,再引入生物中,能够同时或连续引入多个载体。还能够使用REMI技术插入编码各种活性的基因片段。

[0090] 原则上适于将根据本发明的胡萝卜素羟化酶基因或核酸构建体整合入初 始生物的基因组的限制酶全部为技术人员已知的。由于限制酶在待整合的基 因组或载体中剪切过于频繁,仅识别4个碱基对作为限制裂解位点的限制酶 是较不优选的,而优选的酶识别6、7、8或更多碱基对作为裂解位点,如 BamHI、EcoRI、BglII、SphI、SpeI、XbaI、XhoI、NcoI、

SalI、ClaI、KpnI、 HindIII、SacI、PstI、BpnI、NotI、SrfI或SfiI，仅提及几个可能的酶。有利的 是使用的酶在待引入的DNA中不再具有裂解位点，这增加了整合的效力。通 常，在REMI混合物中使用5至500U，优选10至250，特别优选10至100U 的酶。所述酶有利地用于水溶液，其包含用于稳定渗透压的物质(例如糖如蔗 糖、海藻糖或葡萄糖，多元醇如甘油或聚乙二醇)、在pH5至9、优选6至8、 特别优选7至8的范围具有有利缓冲能力的缓冲剂(如Tris、MOPS、HEPES、 MES或PIPES)和/或稳定核酸的物质(如Mg、Cu、Co、Fe、Mn或Mo的无机 或有机盐)。适于存在的其他物质还可能为如EDTA、EDDA、DTT、β-巯基 乙醇或核酸酶抑制剂。然而，还可能在无这些添加的情况下实施REMI技术。

[0091] 所述方法在5至80℃，优选10至60℃，特别优选20至40℃的温度范 围实施。其他已知的用于使细胞膜脱稳定化的方法适于所述方法，如，例如， 电穿孔、与装载的泡融合或用多种碱金属或碱土金属盐如锂、铯或钙盐脱稳 定化，锂盐为优选的。

[0092] 本发明进一步涉及相应的遗传修饰的生物，当初始生物包含胡萝卜素羟 化酶基因的情况下，其通过遗传修饰而使根据本发明的胡萝卜素羟化酶基因 的表达与野生型生物相比增加；或在初始生物不包含胡萝卜素羟化酶基因的 情况下，其通过遗传修饰引起根据本发明的胡萝卜素羟化酶基因的表达。

[0093] 遗传修饰的生物意指生物，其中(优选通过上述方法之一)已插入根据本发 明的胡萝卜素羟化酶基因或核酸构建体。

[0094] 遗传修饰的生物包含至少一种根据本发明的胡萝卜素羟化酶基因或至少 一种根据本发明的核酸构建体。取决于初始生物，所述核酸可存在于染色体 内或染色体外。

[0095] 与野生型相比优选地改变遗传修饰的生物中的类胡萝卜素代谢。

[0096] 优选的生物为重组真菌或酵母。在一个具体的实施方案中，所述重组真菌 是含油的(oleaginous)，因为它可将脂质积累至其干细胞重量的至少约20％；并 产生至少一种选自下组的类胡萝卜素：表氧化玉米黄质(antheraxanthin)、金盏花 红素(adonirubin)、侧金盏花黄质(adonixanthin)、虾青素(astaxanthin)、角黄素 (canthaxanthin)、capsorubrin、β-隐黄素(β-cryptoxanthin)、α-胡萝卜素(α-carotene)、 β、ψ-胡萝卜素、δ-胡萝卜素、ε-胡萝卜素、海胆酮、3-羟基海胆酮、3'-羟基海 胆酮、γ-胡萝卜素、4-酮-γ-胡萝卜素、ζ-胡萝卜素、α-隐黄素、脱氧屈曲黄素 (deoxyflexixanthin)、硅藻黄质(diatoxanthin)、7,8-二脱氢虾青素、二脱氢番茄红 素(didehydrolycopene)、岩藻黄素(fucoxanthin)、岩藻黄醇(fucoxanthinol)、异胡 萝卜素(isorenieratene)、β-异胡萝卜素、山莴苣黄素(lactucaxanthin)、叶黄素 (lutein)、番茄红素(lycopene)、myxobactone、新黄质(neoxanthin)、链孢红素 (neurosporene)、羟基链孢红素、多甲藻素(peridinin)、八氢番茄红素(phytoene)、 视紫红质(rhodopin)、视紫红质葡糖苷、4-酮玉红黄质(4-keto rubixanthin)、管藻 黄质(siphonaxanthin)、球形烯(spheroidene)、球形酮(spheroidenone)、螺菌黄质 (spirilloxanthin)、红酵母烯(torulene)、4-酮-红酵母烯、3-羟基-4-酮-红酵母烯、 尿酸酯(uriolide)、尿酸醋酸酯(uriolide acetate)、紫黄质(violaxanthin)、玉米黄质-β- 二葡糖苷、玉米黄质、C 30类胡萝卜素及其组合，且可将产生的类胡萝卜素积 累至其干细胞重量的至少约1％。甚至更优选地，重组真菌是选自下组的属的成 员：曲霉属(Aspergillus)、布拉霉属(Blakeslea)、葡萄孢属(Botrytis)、假丝酵母属 (Candida)、尾孢属(Cercospora)、隐球菌属(Cryptococcus)、小克银汉霉属

(Cunninghamella)、镰孢属(Fusarium)(赤霉属(Gibberella))、克鲁维酵母属(Kluyveromyces)、油脂酵母属(Lipomyces)、被孢霉属(Mortierella)、毛霉属 (Mucor)、脉孢菌属(Neurospora),青霉属(Penicillium)、须霉属(Phycomyces)、毕 赤酵母属(Pichia)(汉逊酵母属(Hansenula))、柄锈菌属(Puccinia)、腐霉属 (Pythium)、红冬孢酵母属(Rhodosporidium)、红酵母属(Rhodotorula)、酵母属 (Saccharomyces)、小核菌属(Sclerotium)、木霉属(Trichoderma)、丝孢酵母属 (Trichosporon)、Xanthophyllomyces(发夫酵母属(Phaffia))和耶氏酵母/西洋蓍霉属 (Yarrowia)或为选自下组的菌种:土曲霉(Aspergillus terreus)、构巢曲霉 (Aspergillus nidulans)、黑曲霉(Aspergillus niger)、三孢布拉霉(Blakeslea trispora)、 灰葡萄孢(Botrytis cinerea)、日本假丝酵母(Candida japonica)、铁红假丝酵母 (Candida pulcherrima)、拉考夫假丝酵母(Candida revkaufi)、热带假丝酵母 (Candida tropicalis)、产朊假丝酵母(Candida utilis)、烟草尾孢(Cercospora nicotianae)、Cryptococcus curvatus、刺孢小克银汉霉(Cunninghamella echinulata)、 秀丽小克银汉霉(Cunninghamella elegans)、藤仓镰孢(Fusarium fujikuroi)(玉米赤 霉(Gibberella zeae))、乳酸克鲁维酵母(Kluyveromyces lactis)、斯达氏油脂酵母 (Lipomyces starkeyi)、Lipomyces lipoferus、高山被孢霉(Mortierella alpina)、拉曼 被孢霉(Mortierella ramanniana)、深黄被孢霉(Mortierella isabellina)、葡酒色被 孢霉(Mortierella vinacea)、卷枝毛霉(Mucor circinelloides)、粗糙脉孢霉 (Neurospora crassa)、布拉克须霉(Phycomyces blakesleanus)、巴斯德毕赤酵母 (Pichia pastoris)、Puccinia distincta、畸雌腐霉(Pythium irregulare)、对圆红冬孢 酵母(Rhodosporidium toruloides)、胶红酵母(Rhodotorula glutinis)、禾本红酵母 (Rhodotorula graminis)、粘质红酵母(Rhodotorula mucilaginosa)、Rhodotorula pinicola、纤细红酵母(Rhodotorula gracilis)、酿酒酵母(Saccharomyces cerevisiae)、 齐整小核菌(Sclerotium rolfsii)、里氏木霉(Trichoderma reesei)、皮肤丝孢酵母 (Trichosporon cutaneum)、茁芽丝孢酵母(Trichosporon pullulans)、 Xanthophyllomyces dendrorhous(红发夫酵母(Phaffia rhodozyma))和解脂耶氏酵 母/解脂西洋蓍霉(Yarrowia lipolytica)。

[0097] 在这些天然的含油株中,一些还天然地产生类胡萝卜素而一些则不产生, 这些株可通过引入如美国专利7 851 199中公开的类胡萝卜素生物合成基因 被额外地用作宿主细胞。

[0098] 在其他实施方案中,本发明提供产生类胡萝卜素的方法,所述方法包括 在允许类胡萝卜素产生的条件下培养真菌并分离产生的类胡萝卜素的步骤。

[0099] 根据本发明遗传修饰的微生物的培养以本身已知的方式发生,如适当的 野生型的培养,例如在合适的培养基中(如例如,在琼脂平板上或在悬浮培养 物中)的微生物的情况,或在土壤或合适的营养培养基中的植物的情况。收获 意指在微生物的情况中微生物的分离,和在植物的情况中植物(适当时,包含 类胡萝卜素的特定的植物部分)的切断。类胡萝卜素以本身已知的方式分离, 例如,通过生物细胞的破坏,类胡萝卜素的提取,和通过化学或物理分离方 法如提取或层析对类胡萝卜素的随后的纯化。

[0100] 特别地,本申请涉及以下一些具体实施方案:

[0101] 实施方案1.一种用于类异戊二烯的微生物羟化的具有羟化酶活性的多 肽,其选

自下组：

[0102] a)SEQ ID NO:2的多肽；

[0103] b)源自SEQ ID NO:2通过氨基酸的取代、插入或缺失并与SEQ ID NO:2 的序列在氨基酸水平具有至少50％同源性的多肽。

[0104] 实施方案2.实施方案1的多肽，其中所述蛋白具有用于将β-胡萝卜素转 变成β-隐黄素和/或玉米黄质的酶活性，和/或将角黄素转变成金盏花红素和/ 或虾青素的酶活性。

[0105] 实施方案3.一种编码实施方案1的多肽的分离的核酸。

[0106] 实施方案4.实施方案3的分离的核酸，其分别由SEQ ID NO:1和SEQ ID NO:3描述的序列组成。

[0107] 实施方案5.一种核酸构建体或表达载体，其包含与一个或多个(几个)调 控序列可操作连接的实施方案3或4的多核苷酸，所述调控序列指导所述多 肽在表达宿主细胞中的产生。

[0108] 实施方案6.一种转化的微生物(宿主细胞)，其中表达实施方案3或4的 核酸。

[0109] 实施方案7.实施方案6的转化的微生物，其类胡萝卜素代谢不同于野生 型。

[0110] 实施方案8.实施方案7的转化的微生物，其中所述微生物是含油株 (oleaginous strain)。

[0111] 实施方案9.实施方案8的转化的微生物，其中所述含油株是解脂耶氏酵 母(Yarrowia lipolytica)的菌株。

[0112] 实施方案10.一种用于产生实施方案6的转化的微生物的方法，其包括 引入由SEQ ID NO:1中描述的序列组成的核酸或核酸构建体，其与一个或多 个调节信号功能性连接。

[0113] 实施方案11.一种用于制备类胡萝卜素或类胡萝卜素衍生物的方法，其 包括在实施方案1的多肽存在下将β-紫罗兰酮转变成3-羟基-β-紫罗兰酮和/ 或将4-酮-β-紫罗兰酮转变成3-羟基-4-酮-β-紫罗兰酮结构元件。

[0114] 实施方案12.实施方案11的方法，包括如下步骤：

[0115] a)在有助于类胡萝卜素产生的条件下培养根据实施方案6至9中任一项 的转化的微生物；和

[0116] b)回收所述类胡萝卜素和类胡萝卜素衍生物。

[0117] 实施方案13.一种包含实施方案1或2任一项的多肽的组合物。

[0118] 下列实施例对本发明进行了阐述。

实施例

[0119] 下列表1描述某些在下列例示中使用的解脂耶氏酵母菌株：

[0120] 表1:解脂耶氏酵母菌株

ML5252	*MATA erg9-4789::ura3 tef1P-{HMG-tr GGS carB carRP}* 原养型	经典和标准分子遗传技术
ML6804	*MATB erg9-4789::ura3 tef1P-{HMG-tr GGS carB carRP crtW}*原养型	经典和标准分子遗传技术
ML9335	*MATA erg9-4789::URA3 tef1P-{HMG-tr GGS carB carRP crtW Dc-crtZ}*原养型	经典和标准分子遗传技术
ML9863	*MATB erg9-4789::URA3 tef1P-{HMG-tr GGS carB carRP crtW Xa-crtZ Dc-crtZ}* 原养型	经典和标准分子遗传技术
ML11218	*ML9863 crtW-Δ6180*	用 Hyg^R 盒靶向破坏；随后用 cre-lox 系统移除标记
ML11453	*ML9335 tef1P-carRP/*Hyg^R	非靶向转化;还可并入之前存在的活性的额外非标记拷贝
ML11584	*ML9863 tef1P-crtZ-Ep/*NAT^R	非靶向转化;还可并入之前存在的活性的额外非标记拷贝
ML11956	*MATB erg9-4789::ura3 tef1P-{HMG-tr GGS carB carRP crtW Dc-crtZ}*	ML11453 x ML11584
ML12526	*ML9335 with tef1P-HMG-tr tef1P-carB 3X tef1P-carRP*	非靶向转化随后使用 cre-lox 系统移除 Hyg^R 和 NAT^R

[0121]

[0122] 以如美国专利7,851,199所述的ML350和ATCC201249开始,通过引入 在内源TEF1启动子控制下的异源基因,加上数代杂交构建了耶氏酵母株 ML5252、ML6804、ML9863和ML9335。GGS基因和截短的HMG基因 ("HMG-tr")分别源自对应天然的香叶基香叶基焦磷酸合酶(geranylgeranyl pyrophosphate synthase)和羟甲基戊二酰辅酶A还原酶(hydroxymethylglutaryl-CoA reductase)基因的耶氏酵母属序列。carRP和carB 基因源自卷枝毛霉,且它们分别编码双功能性八氢番茄红素合酶/番茄红素环 化酶和八氢番茄红素脱氢酶。合成crtW基因以编码Parvularcula bermudensis 的胡萝卜素酮酶(US 2012/0156718)。

[0123] 从自养黄杆菌(Xanthobacter autotrophicus)(Xa)扩增crtZ基因或进行合成 以编码粉尘肠杆菌(Ep)(SEQ ID NO:4)、肠杆菌科细菌DC404(Dc)(SEQ ID NO:5)或黄杆菌属菌种R1534(Fb)(美国专利6,087,152)的胡萝卜素羟化酶。这 些基因有时但不总是与营养缺陷标志(URA3、LEU2、URA2、LYS1、ADE1) 或loxP位点、Hyg^R(潮霉素抗性)或NAT^R(诺尔丝菌素抗性)标志的残留相关。

[0124] 表2:质粒

[0125]

质粒	骨架	插入	寡聚核苷酸或来源
pMB6486	pMB6157(Hyg^R tef1P-xprT)	Hp^{LF}-crtZ	合成的NheI-MluI片段
pMB6487	pMB6157(Hyg^R tef1P-xprT)	Fb-crtZ	合成的NheI-MluI片段
pMB6056	pMB6157(Hyg^R tef1P-xprT)	Ep-crtZ	合成的NheI-MluI片段

[0126] 本文描述的所有基础分子生物学和DNA操纵方法基本上根据Sambrook等 或

Ausubel等(J.Sambrook,E.F.Fritsch,T.Maniatis(编).1989.Molecular Cloning: A Laboratory Manual.Cold Spring Harbor Laboratory Press:New York;F.M. Ausubel, R.Brent,R.E.Kingston,D.D.Moore,J.G.Seidman,J.A.Smith,K.Struhl (编).1998.Current Protocols in Molecular Biology.Wiley:New York)实施。

[0127] 实施例1a:编码雨生红球藻(H.pluvialis)胡萝卜素羟化酶LF的pMB6486(Hyg^Rtef-crtZ-Hp^{LF})的产生。

[0128] 如使用雨生红球藻(UTEX No.2505)的mRNA序列分析的结果,发明人令 人惊奇地发现了新的β-胡萝卜素羟化酶,将其命名为Hp^{LF}-CrtZ(雨生红球藻 LF(低频))羟化酶。

[0129] 使用如SEQ ID NO:3中详明的解脂耶氏酵母密码子偏好,基于雨生红球 藻基因的推导的蛋白序列从头合成密码子优化的胡萝卜素羟化酶(Hp^{LF}-CrtZ) ORF序列。在从头合成的过程中,将包含NheI限制位点和允许有效翻译的典 型Kozak序列的序列5'-TGCTAGCCACAAAA添加至紧接ATG的上游。将包 含MluI限制位点的序列ACGCGT-3'添加至紧接终止密码子的下游。该序列使 用NheI和MluI裂解并连接至用NheI和MluI剪切的pMB6157以产生 pMB6486。所得pMB6486编码的Hp^{LF}-CrtZ蛋白详明于SEQ ID NO:2中。

[0130] 实施例1b:编码粉尘克罗诺杆菌(之前的粉尘肠杆菌)胡萝卜素羟化酶的pMB6056(Hyg^Rtef-crtZ-Ep)的产生。

[0131] 使用如SEQ ID NO:4中详明的解脂耶氏酵母密码子偏好,基于粉尘克罗 诺杆菌(之前称为粉尘肠杆菌(EMBL登录号CAZ90621.1))基因的推导的蛋白 序列从头合成密码子优化的胡萝卜素羟化酶(Ep-CrtZ)ORF序列。在从头合成 的过程中,将包含NheI限制位点和允许有效翻译的典型Kozak序列的序列 5'-TGCTAGCCACAAAA添加至紧接ATG的上游。将包含MluI限制位点的序 列ACGCGT-3'添加至紧接终止密码子的下游。该序列使用NheI和MluI裂解并 连接至用NheI和MluI剪切的pMB6157以产生pMB6056。

[0132] 实施例1c:编码黄杆菌属R1534胡萝卜素羟化酶的pMB6487(Hyg^R tef-crtZ-Fb)的产生。

[0133] 使用解脂耶氏酵母密码子偏好,基于如美国专利6,087,152中详明的黄杆 菌属R1534的蛋白序列从头合成密码子优化的胡萝卜素羟化酶(Fb-CrtZ)。在 从头合成的过程中,将包含NheI限制位点和允许有效翻译的典型Kozak序列 的序列5'-TGCTAGCCACAAAA添加至紧接ATG的上游。将包含MluI限制 位点的序列ACGCGT-3'添加至紧接终止密码子的下游。

[0134] 该序列使用NheI和MluI裂解并连接至用NheI和MluI剪切的pMB6157以 产生pMB6487。

[0135] 实施例2A:向解脂耶氏酵母角黄素产生株ML6804引入三种胡萝卜素羟化酶基因以产生虾青素

[0136] 为测试Hp^{LF}CrtZ的羟化潜力并与Fb CrtZ和Ep CrtZ进行比较,用三种 不同的构建体转化菌株ML6804。

[0137] 1)携带tef1p控制下的黄杆菌属crtZ基因和潮霉素抗性标记Hyg^R的 pMB6487的HindIII-XbaI片段;

[0138] 2)携带tef1p控制下的粉尘克罗诺杆菌(之前为粉尘肠杆菌)crtZ基因和 潮霉素抗性标记Hyg^R的pMB6056的PvuII片段;

[0139] 3)携带tef1p控制下的雨生红球藻“低频”crtZ基因和潮霉素抗性标记 Hyg^R的pMB6486的HindIII-XbaI片段。

[0140] 在30℃生长3-4天后在具有100mg/L潮霉素的YPD上选择转化体。将 来自每种构建体的十个转化体在摇瓶中的YPD内生长4天。全部转化体产生 虾青素。来自每种转化的代表性转化体与其亲代菌株ML6804示于图1。包 含黄杆菌属R1534 crtZ基因的菌株ML12471产生3％的虾青素和9％的金盏 花红素(作为总类胡萝卜素的百分比)。包含粉尘克罗诺杆菌(之前为粉尘肠杆 菌)crtZ基因的菌株ML11622产生7％的虾青素和11％的金盏花红素。且包含 雨生红球藻LF crtZ基因的菌株ML12466产生12％的虾青素和22％的金盏花 红素。

[0141] 实施例2B:向解脂耶氏酵母虾青素产生株引入雨生红球藻LF胡萝卜素羟化酶以增加虾青素的纯度

[0142] 用携带tef1启动子、Hp^{LF} crtZ基因和针对潮霉素抗性Hyg^R的选择性标志 的pMB6486的HindIII-XbaI片段转化菌株ML9863和ML11956。在30℃生 长3-4天后,在转化平板(YPD,具有100mg/L潮霉素)上选择来自与亲代相 比表现较暗的每种株的二十个Hyg^R转化体。转化体在YPD中于摇床中生长 4天。与其亲代菌株相比两个代表性的转化体示于图2。

[0143] ML12562源自ML9863且ML12566源自ML11956。作为总类胡萝卜素 的百分比,ML12562和ML12566相比它们的亲代分别产生3倍(40％对13％) 和2倍(27％对13％)更多的虾青素(图2)。

[0144] 实施例2c:向解脂耶氏酵母β-胡萝卜素产生株引入雨生红球藻LF羟化酶以产生β-隐黄素:

[0145] 用三种不同的构建体转化菌株ML5252:

[0146] 1)携带tef1p控制下的黄杆菌属R1534 crtZ基因和潮霉素抗性标记Hyg^R的pMB6487的HindIII-XbaI片段;

[0147] 2)携带tef1p控制下的粉尘克罗诺杆菌(之前为粉尘肠杆菌)crtZ基因和 潮霉素抗性标记Hyg^R的pMB6056的PvuII片段;

[0148] 3)携带tef1p控制下的雨生红球藻LF crtZ基因和潮霉素抗性标记Hyg^R的pMB6486的HindIII-XbaI片段。

[0149] 在30℃生长3-4天后在具有100mg/L潮霉素的YPD上选择转化体。将 来自每种构建体的十株转化体在摇瓶中的YPD内生长4天。全部转化体都产 生玉米黄质和β-隐黄素。代表性转化体与亲代菌株ML5252相比示于图3。 包含黄杆菌属crtZ基因的菌株ML12458产生6％的β-隐黄素和6％的玉米黄 质(作为总类胡萝卜素的百分比)。包含粉尘克罗诺杆菌(之前为粉尘肠杆菌) crtZ基因的菌株ML10341产生3％的β-隐黄素和39％的玉米黄质(作为总类胡 萝卜素的百分比)。包含雨生红球藻LF crtZ基因的菌株ML12453产生21％的 β-隐黄素和4％的玉米黄质(作为总类胡萝卜素的百分比)。

[0150] 从解脂耶氏酵母细胞提取类胡萝卜素并通过HPLC定量类胡萝卜素生产:

[0151] 根据之前美国专利号7,851,199B2中描述的方法对生成的菌株实施摇瓶 测试和类胡萝卜素分析。

序列表

<110> 帝斯曼知识产权资产管理有限公司(DSM IP Assets B.V.)

<120> 胡萝卜素羟化酶及其用于产生类胡萝卜素的用途

<130> Case 28352 PCT

<160> 5

<170> SIPOSequenceListing 1.0

<210> 1

<211> 894

<212> DNA

<213> 雨生红球藻(Haematococcus pluvialis)

<400> 1

atggctttgt cgaagatgca gccaatgggg gtgcaaggac ggcgcatcgc attgcctcgt 60
gacatctcaa ggcctcagat gccctctcaa aggcagccat tggcccaggt gctacgcgta 120
gcagccccag acacagaggc ggtggagtcc tcacgggtgc agccgtctgt cggtgcaggc 180
aatcagcaaa ctgcagacga cgccctccaa cagctcgacc ggtccatcgc tctgcgccgc 240
gctcaacgca aaagggagca gctgacctac caggctgcag ccatagcagc gtccgtgggc 300
atatcaagcc ttgccatcct tgccacctac acgcggtttt caatgcacat cacatcagat 360
ggcgagatgc cttggagcga cctcttgggc accctctcac tggtggcagg tggcgcgttt 420
ggcatggaga tgtatgcgcg ttatgcacac aaggccctat ggcatgagtc ccccctgggt 480
tggctgctgc acaagagcca ccacactccc cgcaccggac cctttgaagc caatgactta 540
ttcgccatca tcaatggcct gccagccatg ctgctgtgct attatgggtt ctgggagcct 600
ggcatggtcg gcgcctcatg ctttggtgct gggctgggca tcaccctgta cgggatggcg 660
tacatgttca tccatgacgg cctagtccac cggcgctttc ccaccgggcc catctcagac 720
ctgccctaca tgaagcgcct gacagtggcc caccagatcc accacagcgg caagtatgat 780
ggcgcgccct ggggcatgtt cctggggcct caggagctgg aggcaatatc aggtgccaca 840
gaggagctgg accgcttggt ggcagacctg gactggtcaa agcgcagggt gtga 894

<210> 2

<211> 297

<212> PRT

<213> 雨生红球藻(Haematococcus pluvialis)

<400> 2

Met Ala Leu Ser Lys Met Gln Pro Met Gly Val Gln Gly Arg Arg Ile
1 5 10 15
Ala Leu Pro Arg Asp Ile Ser Arg Pro Gln Met Pro Ser Gln Arg Gln
20 25 30
Pro Leu Ala Gln Val Leu Arg Val Ala Ala Pro Asp Thr Glu Ala Val
35 40 45
Glu Ser Ser Arg Val Gln Pro Ser Val Gly Ala Gly Asn Gln Gln Thr

50 55 60

Ala Asp Asp Ala Leu Gln Gln Leu Asp Arg Ser Ile Ala Leu Arg Arg
65 70 75 80
Ala Gln Arg Lys Arg Glu Gln Leu Thr Tyr Gln Ala Ala Ala Ile Ala
85 90 95
Ala Ser Val Gly Ile Ser Ser Leu Ala Ile Leu Ala Thr Tyr Thr Arg
100 105 110
Phe Ser Met His Ile Thr Ser Asp Gly Glu Met Pro Trp Ser Asp Leu
115 120 125
Leu Gly Thr Leu Ser Leu Val Ala Gly Gly Ala Phe Gly Met Glu Met
130 135 140
Tyr Ala Arg Tyr Ala His Lys Ala Leu Trp His Glu Ser Pro Leu Gly
145 150 155 160
Trp Leu Leu His Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu
165 170 175
Ala Asn Asp Leu Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu
180 185 190
Cys Tyr Tyr Gly Phe Trp Glu Pro Gly Met Val Gly Ala Ser Cys Phe
195 200 205
Gly Ala Gly Leu Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Ile
210 215 220
His Asp Gly Leu Val His Arg Arg Phe Pro Thr Gly Pro Ile Ser Asp
225 230 235 240
Leu Pro Tyr Met Lys Arg Leu Thr Val Ala His Gln Ile His His Ser
245 250 255
Gly Lys Tyr Asp Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu
260 265 270
Leu Glu Ala Ile Ser Gly Ala Thr Glu Glu Leu Asp Arg Leu Val Ala
275 280 285
Asp Leu Asp Trp Ser Lys Arg Arg Val
290 295

<210> 3
<211> 891
<212> DNA
<213> 雨生红球藻(Haematococcus pluvialis)
<400> 3
atggctctgt ccaagatgca gcccatgggt gtccagggcc gacgaattgc tctcccccga 60
gacatctccc gaccccagat gccttcccag cgacagcctc tcgcccaggt tctgcgagtt 120
gctgcccccg acaccgaggc cgtcgagtct tctcgagtcc agccctccgt cggtgccggc 180

aaccagcaga ccgccgacga cgccctccag cagctcgacc gatccattgc tctgcgacga 240
gcccagcgaa agcgagagca gctcacctac caggctgctg ccattgccgc ctccgttggt 300
atctcctctc tcgccatcct cgccacctac acccgattct ccatgcacat cacctccgac 360
ggcgagatgc cctggtccga tctcctcggt actctgtctc tcgtcgccgg tggtgccttt 420
ggcatggaga tgtacgcccg atacgcccac aaggctctgt ggcacgagtc tcccctcggc 480
tggctcctcc acaagtccca ccacactccc cgaaccggcc ccttcgaggc caacgacctc 540
tttgccatca tcaacggtct gcccgccatg ctcctctgct actacggctt ctgggagccc 600
ggtatggtcg gtgcttcttg tttcggtgct ggcctcggta tcaccctcta cggtatggcc 660
tacatgttca tccacgacgg tctggtccac cgacgattcc ccactggccc catctccgat 720
ctgccctaca tgaagcgact caccgtcgcc caccagatcc accactccgg caagtacgac 780
ggtgctccct ggggtatgtt cctcggcccc caggagctgg aggccatttc cggtgccacc 840
gaggagcttg accgactcgt tgccgatctc gactggtcca agcgacgagt c 891

<210> 4
<211> 567
<212> DNA
<213> Enterobacter pulveris
<400> 4
atgcttgtgc tgtggaatac gcttatcgtc atcacgacat tttgcctgat ggagattgtg 60
gcaacgcttg cccataaata catcatgcac ggctggggat ggggctggca ccgttcgcac 120
cacgagccgc ggcacggctg gtttgaggtc aacgatctct acgcagtggt gttcgcactg 180
ctcgccatcg tgctgatcgc actgggaaca gcaggctggt ggccgctgca atggcttggc 240
gcaggcatga cgctctacgg tctgttctat ttcctcgtgc acgacggtct ggtgcaccgg 300
agatggccgt ttaactacat tccgcggaga ggctacctta agcggcttta ccaggcacac 360
cggctgcacc acgcagtgaa tggcaaagag ggttgcgtct ccttcggttt tctctacgca 420
gctaagccgg aagcacttca ggcagagctg cggagacgtc acggtcggaa acctaaagca 480
gacgccgcca gagcccggcg ggacggtaaa cctgtcgcac agagcgagag ccgagcacaa 540
gacctgccgc cgaaatcggc agagtaa 567

<210> 5
<211> 558
<212> DNA
<213> Enterobacteriacea DC404
<400> 5
atgcttgcgt tgtggaatac cgggatcgtg ctactgacta tcatcatcat ggaaggggtg 60
gcaacgttcg cacacaagta catcatgcac ggctggggat ggggctggca tcattcgcac 120
cataccccgc gcaaaggggc gtttgagcgt aacgatctct atgcggtggt gtttgcgcta 180
ctggccattg cgctgattta cgcgggcagc gaagggtact ggccgcttca gtggattggc 240
gcgggaatga ccggctacgg cgtgatctac tttatcgttc acgatggtct tgtccaccag 300
cgctggccgt tccgttacgt gccgcgccgc ggctatctgc gccgcctcta catggcacac 360
cggctgcatc acgcggtgcg ggggcgcgaa gggtgcgtct ccttcgggtt tatctacgcc 420

ccaccggtgg acaagctgca ggcggtgctg cgcgaacgta acggcagacc cgccagcgcg 480
ggcgctgcca gaggtgcgga tcgcgcggcg gccagctcgc cttccgggaa gccatcgcct 540
gcttcgcgcc ggaaataa 558

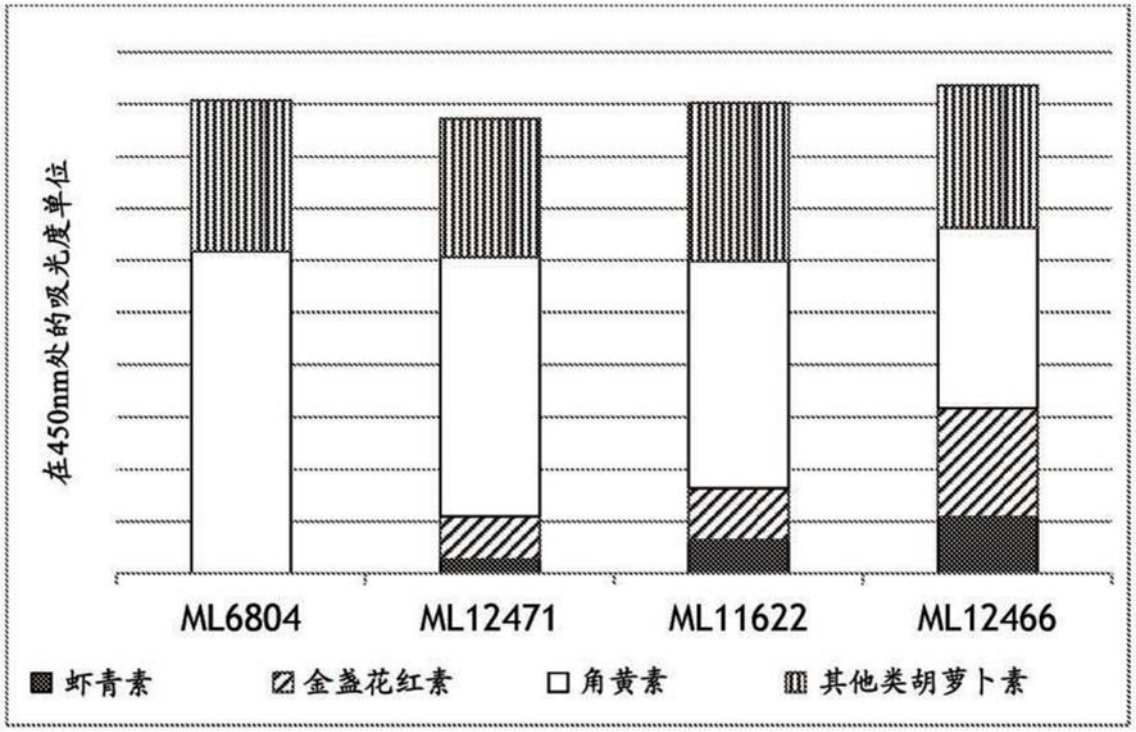

在450nm处的吸光度单位
ML6804
ML12471
ML11622
ML12466
虾青素
金盏花红素
角黄素
其他类胡萝卜素

图1

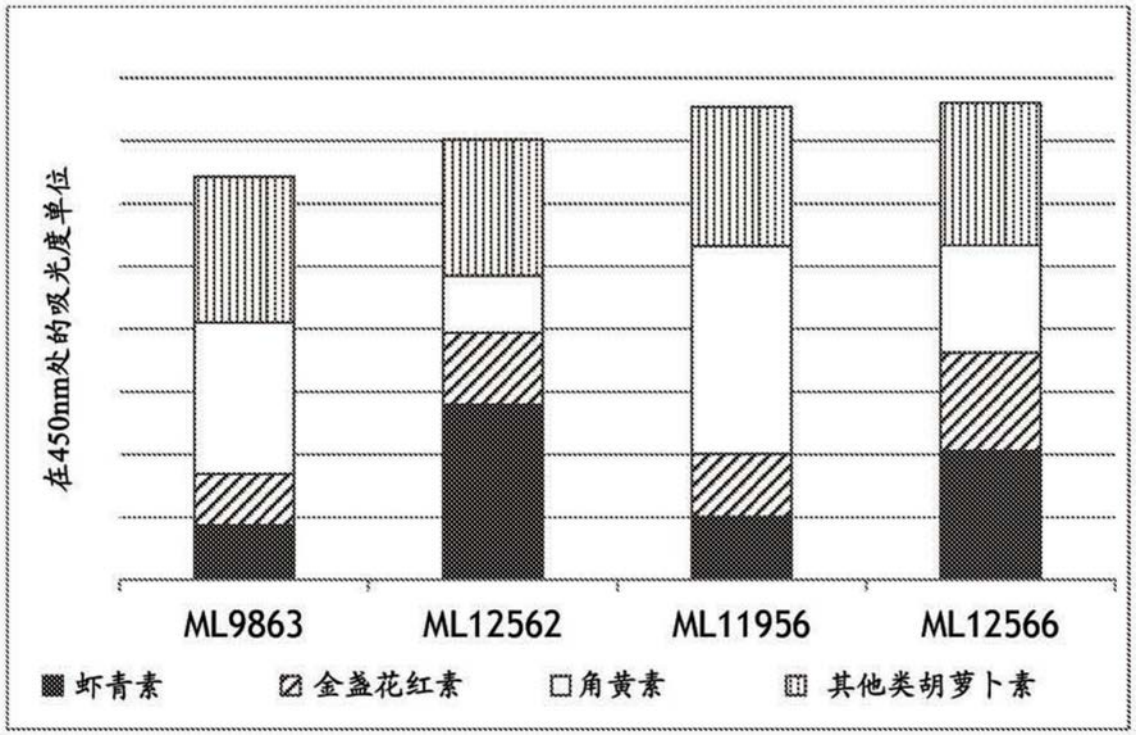

在450nm处的吸光度单位
ML9863
ML12562
ML11956
ML12566
虾青素
金盏花红素
角黄素
其他类胡萝卜素

图2

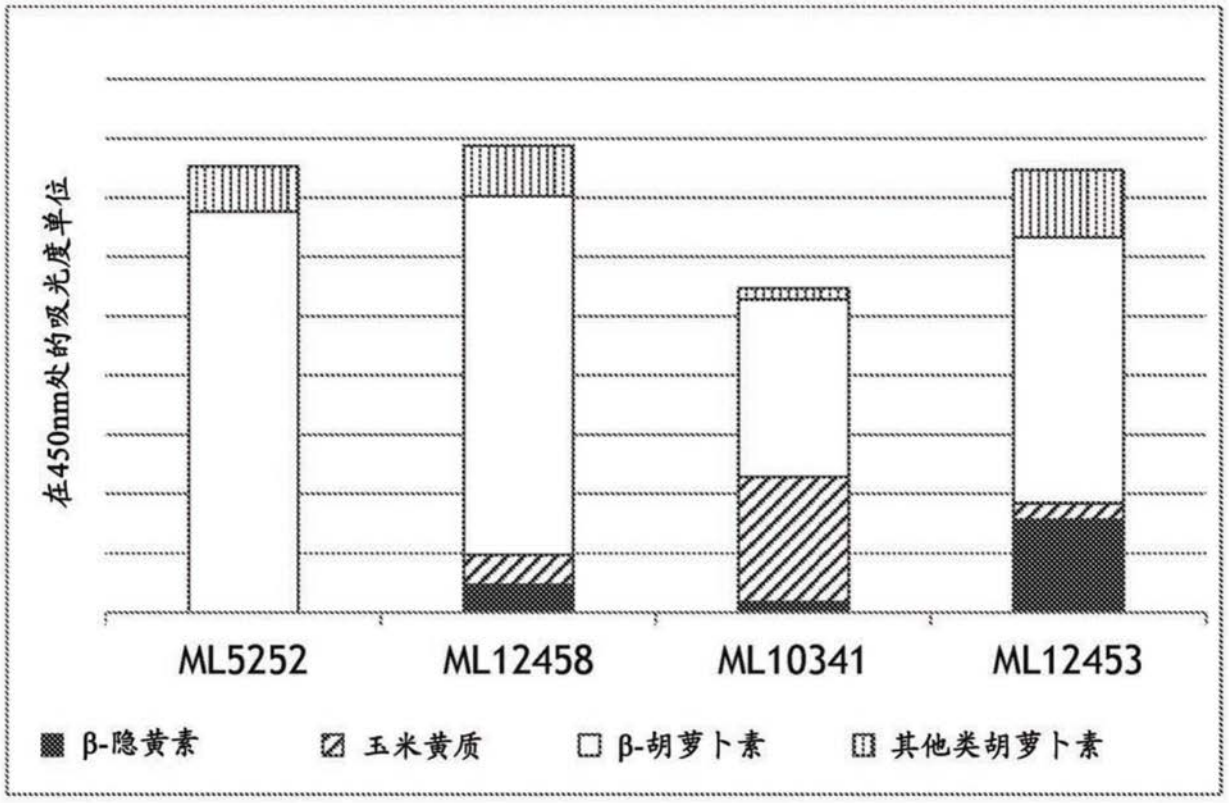

在450nm处的吸光度单位
ML5252
ML12458
ML10341
ML12453
β-隐黄素
玉米黄质
β-胡萝卜素
其他类胡萝卜素

图3