



(10) 授权公告号 CN 114174489 B

(45) 授权公告日 2024. 11. 08

(21) 申请号 202080055048.6

(22) 申请日 2020.07.31

(65) 同一申请的已公布的文献号
申请公布号 CN 114174489 A

(43) 申请公布日 2022.03.11

(30) 优先权数据
62/881874 2019.08.01 US

(85) PCT国际申请进入国家阶段日
2022.01.28

(86) PCT国际申请的申请数据
PCT/US2020/044613 2020.07.31

(87) PCT国际申请的公布数据
W02021/022216 EN 2021.02.04

(73) 专利权人 阿迈瑞斯公司
地址 美国加利福尼亚州

(72) 发明人 L·雷茨 C·汉森

(74) 专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 罗文锋 彭昶

(51) Int.Cl.
C12N 1/19 (2006.01)
C12N 15/81 (2006.01)
C12N 15/53 (2006.01)
C12P 7/24 (2006.01)
C12P 19/44 (2006.01)
C12R 1/865 (2006.01)

(56) 对比文件
EP 2957629 A1, 2015.12.23
US 2014245496 A1, 2014.08.28

审查员 郭青

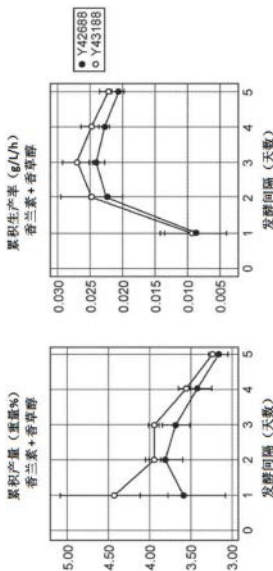
权利要求书1页 说明书28页
序列表5页 附图3页

(54) 发明名称

用于高效生成香兰素的经修饰宿主细胞

(57) 摘要

本发明提供了用于改进香兰素和/或葡萄糖香兰素生成的经遗传修饰的宿主细胞、组合物和方法。本发明所述的宿主细胞、组合物和方法提供了异源生成香兰素和/或葡萄糖香兰素以及可由其中之一或两者合成或生物合成的任何化合物的有效途径。



1.能够生成香兰素(vanillin)或葡糖香兰素(glucovanillin)的经遗传修饰的宿主细胞,其包含脂肪醛脱氢酶HFD1的缺失,其中所述宿主细胞是酿酒酵母(*Saccharomyces cerevisiae*)。

2.根据权利要求1所述的经遗传修饰的宿主细胞,其不表达Hfd1。

3.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含ADH6的缺失。

4.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含GRE2的缺失。

5.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达AroB、AroD或AroZ的一种或多种核酸。

6.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达AroB、AroD、AroF或AroZ的一种或多种核酸。

7.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达大肠杆菌(*E.coli*) AroB、大肠杆菌(*E.coli*) AroD、大肠杆菌(*E.coli*) AroF或疏毛柄孢壳菌(*Podosporapauciseta*) AroZ的一种或多种核酸。

8.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达磷酸泛酰巯基乙胺基转移酶或芳香族羧酸还原酶的一种或多种核酸。

9.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达谷氨酸棒杆菌(*Corynebacterium glutamicum*) 磷酸泛酰巯基乙胺基转移酶或艾阿华诺卡氏菌(*Nocardia iowensis*) 芳香族羧酸还原酶的一种或多种核酸。

10.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达UDP-糖基转移酶(UGT)的核酸。

11.根据权利要求1或2所述的经遗传修饰的宿主细胞,其进一步包含表达拟南芥(*Arabidopsis thaliana*) UGT的核酸。

12.根据权利要求1或2所述的经遗传修饰的宿主细胞,其中每个异源基因均由诱导型启动子表达。

13.根据权利要求1或2所述的经遗传修饰的宿主细胞,其中每个异源基因均由GAL启动子表达。

14.生成香兰素或葡糖香兰素的方法,其包含以下步骤:

(a) 在适于制备香兰素或葡糖香兰素的条件下,在含有碳源的培养基中培养权利要求1-13任一项所述的经遗传修饰的宿主细胞的种群以生成培养液(culturebroth),其中所述宿主细胞是酿酒酵母(*Saccharomyces cerevisiae*);和

(b) 从所述培养液回收所述香兰素或葡糖香兰素。

用于高效生成香兰素的经修饰宿主细胞

发明领域

[0001] 本发明内容涉及特定的遗传修饰、包含其的宿主细胞,以及使用所述宿主细胞生成香兰素和/或葡萄糖香兰素以及可由其中之一或两者合成或生物合成的任何化合物的方法。

[0002] 发明背景

[0003] 香兰素是世界上用量最大的香料成分。仅有约1%的香草风味成分供应来自香草兰花的香草提取物。“天然”香兰素需求旺盛,但供应不足,且价格高昂。可替代的、低成本、大批量的“天然”香兰素来源将是调味品市场一个有利可图的补充。通过酵母对糖进行发酵从头生成的香兰素,与目前市场上的替代品相比,具有以更低的成本生成“天然”香兰素的潜力。

[0004] 有几种方法可用于通过天然前体(包括除葡萄糖以外的前体)的生物转化来生成“天然”香兰素。一种途径是在某些植物的某些部分大量存在的阿魏酸的生物转化。已鉴定出微生物,其可通过生成香兰素作为中间体的途径分解代谢阿魏酸。可对这些微生物进行工程改造,以减少香兰素进一步分解代谢为不需要的副产物,从而优化香兰素的生成。Gallage et al., *Molecular Plant*, 8:40-57 (2015)。在类似的方法中,更具成本效益的底物丁香酚可被微生物分解代谢为阿魏酸,并进一步分解代谢为香兰素(Gallage等人)。

[0005] 未见已知的微生物可天然地将葡萄糖转化为香兰素(Gallage等人)。1998年,开发了一种从葡萄糖到香兰素的酶促路线,所述路线通过3个附加酶促步骤将天然产生的代谢物3-脱氢莽草酸(3-dehydroshikimate)转化为香兰素:1)脱水以生成原儿茶酸(3,4-二羟基苯甲酸),2)3-羟基基团的O-甲基化,和3)使羧酸还原成醛。Li and Frost, *J. Am. Chem. Soc.*, 120:10545-10546 (1998)。通过表达催化3-DHS脱水酶(AroZ)和儿茶酚-O-甲基转移酶(COMT)的异源酶,在大肠杆菌(*E. coli*)中生成香草酸(步骤1和步骤2)证明了这一过程。采用由真菌纯化的芳香族羧酸还原酶(ACAR)进行的酶促转化用于在体外将香草酸转化为香兰素。

[0006] Hansen等人证明了通过表达上述酶与异源PPTase的组合,在单一重组生物体酿酒酵母(*Saccharomyces cerevisiae*)中由葡萄糖从头生物合成香兰素,所述异源PPTase被鉴定为激活此生物体中的ACAR酶所必需的。Hansen et al., *Appl. Environ. Microbiol.* 75: 2765-2774 (2009)。此外,它们表达UDP-葡萄糖基转移酶,将有毒的香兰素产品转化为毒性低得多的葡萄糖香兰素。

[0007] 已报道许多其他修饰以提高酵母中香兰素生物合成的效率。为了提高葡萄糖香兰素的效价(Hansen等人),证明重要的是通过删除天然还原酶(即,ADH6)来减少内源性还原酶活性,以减少香兰素向香草醇的转化。为了减少碳对不需要的异构体异香兰素(通过4-OH而非3-OH的甲基化生成)的损失,人变体Hs.COMT被用作酶进化的起点。US 2014/0245496; WO 2015/121379。获得对正确香兰素异构体具有高度特异性的突变体。为了增加通向PCA的通量并减少通向莽草酸途径代谢物的通量,生成了Aro1的突变体形式,注释为突变体AROM,其中包含E结构域中的突变,并降低了此反应的活性,所述反应使用3-DHS作为底物来制备莽

草酸。

[0008] 可提供低成本、大量来源的“天然”香兰素的进一步遗传修饰将是调味品市场的重要补充。

发明摘要

[0009] 本发明提供了用于改进香兰素和/或葡萄糖香兰素生成的经遗传修饰的宿主细胞、组合物和方法。所述这些组合物和方法部分基于某些基因产物的缺失,所述基因产物包括在已经遗传修饰以生成香兰素和/或葡萄糖香兰素的宿主细胞中的脂肪醛脱氢酶(即,HFD1)的同源物及其同源物类。虽然不打算受任何特定操作理论的束缚,但本发明实施例证明HFD1编码能够将香兰素转化为不太需要的香草酸的酶。HFD1的缺失可减少或消除此反应。

[0010] 一方面,本发明提供了经遗传修饰的宿主细胞及其用于生成香兰素或葡萄糖香兰素的方法。在某些实施方案,本发明提供了能够生成香兰素或葡萄糖香兰素的经遗传修饰的宿主细胞,其中所述宿主细胞具有脂肪醛脱氢酶(HFD1)同源物的缺失。在特定实施方案,所述经遗传修饰的宿主细胞表达足以生成香兰素或葡萄糖香兰素的另外的酶。有用的酶描述在本发明中。

[0011] 另一方面,本发明提供了生成香兰素或葡萄糖香兰素的方法,其包括:在适于制备香兰素或葡萄糖香兰素的条件下,在含有碳源的培养基中培养本发明的宿主细胞的种群以生成培养液(culture broth);从所述培养液回收所述香兰素或葡萄糖香兰素。

[0012] 另一方面,本发明提供了通过本发明提供的方法生成的香兰素。在某些实施方案,本发明提供的香兰素具有:与PDB(Pee Dee Belemnite)标准品的块体(bulk) $\delta^{13}\text{C}$ 偏差为约-14.7‰至约-12.8‰,与SMOW(标准平均大洋水,Standard Mean Ocean Water)标准品的块体 $\delta^2\text{H}$ 偏差为约-150‰至约-124‰,和/或约12.9dpm/g至约14.1dpm/g的 ^{14}C 活性。

[0013] 附图简要说明

[0014] 图1提供了在液体培养基中香兰素及降解产物香草醇和香草酸的效价(g/L),起始浓度为1g/L香兰素,针对所示菌株修饰温育24小时。

[0015] 图2提供了生成香兰素的菌株Y42688及改进的衍生物Y43188的5天发酵的累积产量(重量%;香兰素+香草醇)和累积生产率(g/L/h;香兰素+香草醇)。累积表示从时间0到指定时间的间隔值。

[0016] 图3提供了从蔗糖、葡萄糖或果糖生成香草酸、香兰素、阿魏酸、丁香酚和姜黄素的途径。

[0017] 图4提供了衍生自多种原料的香兰素的块体 $\delta^{13}\text{C}$ 数据比较,表示为与PDB的‰偏差,以及表示为与SMOW的‰偏差的块体 $\delta^2\text{H}$ 数据比较。在图4中,使用以下符号:孜然(■)、丁香酚(●)、阿魏酸(提取自稻米,◆)、阿魏酸(提取自玉米,▲)、愈创木酚(□)、木质素(○)、香草豆(◇)、葡萄糖(提取自玉米,Δ)、蔗糖(提取自甘蔗,x)。

具体实施方式

[0018] 术语

[0019] 本发明使用的术语“约”或“大约”是指关于由本领域技术人员确定的值的合理范围。在某些实施方案,术语约或大约是指±一个、两个或三个标准偏差。在某些实施方案,术

语约或大约是指 $\pm 5\%$ 、 10% 、 20% 或 25% 。在某些实施方案,术语约或大约是指 ± 0.1 、 0.2 或 0.3 对数单位,例如pH单位。

[0020] 本发明使用的术语“异源的/异源性/异源”是指通常在自然界中不存在的物质。术语“异源核苷酸序列”是指自然界中在给定细胞中通常不存在的核苷酸序列。因此,异源核苷酸序列可以是:(a)相对于其宿主细胞是外源的(即,对所述细胞而言是“外源的”);(b)天然存在于所述宿主细胞中(即“内源性/内源的/内源”),但在所述细胞中以非天然量存在(即,比所述宿主细胞中天然存在的量更多或更少);或(c)天然存在于所述宿主细胞中,但位于其天然基因座之外。

[0021] 另一方面,本发明使用的术语“原生的/天然的(native)”或“内源的/内源性/内源”涉及分子,特别是酶和核酸,表示在它们起源或在自然界中发现的生物体中表达的分子。应理解,原生酶或原生多核苷酸的表达可在重组微生物中进行修饰。在特定实施方案,经密码子优化的基因表达天然酶。

[0022] 本发明使用的术语“异源核酸表达盒(cassette)”是指包含编码序列的核酸序列,所述编码序列可操作地连接至足以在宿主细胞中表达所述编码序列的一个或多个调节元件。调节元件的非限制性实例包括启动子、增强子、沉默子、终止子和多聚腺苷酸加尾(poly-A)信号。

[0023] 本发明使用的基因名称通常字母大写并斜体,例如,HFD1。蛋白质名称通常首字母大写但不斜体,例如,Hfd1或Hfd1p。然而,在指明术语蛋白质的地方,则旨在是蛋白质。譬如,技术人员将认识到“HFD1蛋白”旨在指代Hfd1p。

[0024] 本发明使用的术语“脂肪醛脱氢酶的同源物”和“HFD1”或“Hfd1”是指参与泛醌和鞘脂类代谢的编码核酸和脱氢酶,其能够将4-羟基苯甲醛转化为4-羟基苯甲酸以进行泛醌合成代谢和/或在1-磷酸鞘氨醇分解代谢中将十六烯醛转化为十六碳烯酸。在某些实施方案,其EC编号为1.2.1.3。在某些实施方案,其序列是根据NCBI参考序列NP_013828或酿酒酵母(*S.cerevisiae*)YMR110C。

[0025] 本发明使用的术语“NADPH依赖性中链醇脱氢酶”和“ADH6”或“Adh6”是指编码核酸和醇脱氢酶。在某些实施方案,其EC编号为1.1.1.2。在某些实施方案,其序列是根据GenBank(基因库)基因座CAA90836或酿酒酵母(*S.cerevisiae*)YMR318C。

[0026] 本发明使用的术语“3-甲基丁醛还原酶”和“NADPH依赖性甲基乙二醛还原酶”和“GRE2”或“Gre2”是指编码核酸及3-甲基丁醛还原酶和NADPH依赖性甲基乙二醛还原酶。在某些实施方案,其EC编号为1.1.1.265或1.1.1.283。在某些实施方案,其序列是根据NCBI参考序列NP_014490或酿酒酵母(*S.cerevisiae*)YOL151W。

[0027] 本发明使用的术语“YGL039W”是指编码核酸和醛还原酶。其系统名称为YGL039W。在某些实施方案,其序列是根据GenBank(基因库)参考Z72561。

[0028] 本发明使用的术语“二氢叶酸还原酶”和“DHFR”是指编码核酸和二氢叶酸还原酶。在某些实施方案,其EC编号为1.5.1.3。在某些实施方案,DHFR来自小家鼠(*Mus musculus*)。在某些实施方案,DHFR序列是根据NCBI参考序列NP_034179。

[0029] 本发明使用的术语“3-脱氢奎尼酸合成酶”和“AroB”是指编码核酸和3-脱氢奎尼酸合成酶。在某些实施方案,其EC编号为4.2.3.4。在某些实施方案,AroB来自大肠杆菌(*E.coli*)。在某些实施方案,AroB序列是根据UniProtKB P07639。

[0030] 本发明使用的术语“3-脱氢奎尼酸脱水酶”和“AroD”是指编码核酸和3-脱氢奎尼酸脱水酶。在某些实施方案,其EC编号为4.2.1.10。在某些实施方案,AroD来自大肠杆菌(*E.coli*)。在某些实施方案,AroD序列是根据UniProtKB P05194。

[0031] 本发明使用的术语“磷酸-2-脱氢-3-脱氧庚酸醛缩酶,Tyr敏感性”和“AroF”是指编码核酸和磷酸-2-脱氢-3-脱氧庚酸醛缩酶。在某些实施方案,其EC编号为2.5.1.54。在某些实施方案,AroF来自大肠杆菌(*E.coli*)。在某些实施方案,AroF序列是根据UniProtKB P00888。在某些实施方案,AroF是抗反馈的(*J.Bacteriol.* November 1990 172:6581-6584)。

[0032] 本发明使用的术语“3-脱氢莽草酸脱水酶”和“AroZ”是指编码核酸和3-脱氢莽草酸脱水酶。在某些实施方案,其EC编号为4.2.1.118。在某些实施方案,AroZ来自疏毛柄孢壳菌(*Podospira pauciseta*)。在某些实施方案,AroZ序列是根据以下文献:Hansen et al., *Appl Environ Microbiol.* 2009 (May) 75 (9):2765-74。

[0033] 本发明使用的术语“磷酸泛酰巯基乙胺基转移酶”和“PPTASE”是指编码核酸和磷酸泛酰巯基乙胺基转移酶。在某些实施方案,其EC编号为2.7.8.7。在某些实施方案,PPTASE来自谷氨酸棒杆菌(*Corynebacterium glutamicum*)。在某些实施方案,PPTASE序列是根据UniProtKB Q8NP45。

[0034] 本发明使用的术语“芳香族羧酸还原酶”和“ACAR”是指编码核酸和芳香族羧酸还原酶。在某些实施方案,其EC编号为1.2.1.30。在某些实施方案,ACAR来自艾阿华诺卡氏菌(*Nocardia iowensis*)。在某些实施方案,ACAR序列是根据UniProtKB Q6RKB1。

[0035] 本发明使用的术语“丁香酚醇氧化酶”和“EA0”是指编码核酸和丁香酚醇氧化酶。在某些实施方案,EA0来自约氏红球菌(*Rhodococcus jostii*)。在某些实施方案,EA0序列是根据UniProtKB Q0SBK1。

[0036] 本发明使用的术语“UDP-糖基转移酶”和“UGT”是指编码核酸和UDP-糖基转移酶。在某些实施方案,其EC编号为2.4.1.126。在某些实施方案,UGT来自拟南芥(*Arabidopsis thaliana*)。在某些实施方案,UGT是拟南芥UGT72E2。在某些实施方案,UGT序列是根据UniProtKB Q9LVR1。

[0037] 本发明使用的术语“亲本细胞”是指与本发明公开的经遗传修饰的宿主细胞具有相同遗传背景的细胞,除了其不包含工程化到所述经修饰的宿主细胞中的一种或多种特定遗传修饰,譬如,选自由以下组成的组的一种或多种修饰:香兰素途径酶的异源表达,葡萄糖香兰素途径酶的异源表达;或AroB、AroD、AroF、AroZ、PPTASE或ACAR的异源表达;或HFD1、ADH6、GRE2或YGL039W的缺失。

[0038] 本发明使用的术语“天然存在的”是指天然存在的那些物质。譬如,存在于生物体中的基因产物可从自然界中的来源分离得到并且在实验室中未被人为有意修饰,其就是天然存在的基因产物。相反,本发明使用的术语“非天然存在的”是指在自然界中未发现但通过人为干预生成的那些物质。在某些实施方案,天然存在的基因组序列经修饰,例如,经密码子优化,用于本发明提供的生物体。

[0039] 术语“培养基”是指培养基和/或发酵培养基。

[0040] 术语“发酵组合物”是指组合物,所述组合物包含经遗传修饰的宿主细胞和由所述经遗传修饰的宿主细胞生成的产物或代谢物。发酵组合物的实例是全细胞培养液,其可以

是容器(例如,烧瓶、平板或发酵罐)的全部内容物,包括细胞、水相和由所述经遗传修饰的宿主细胞生成的化合物。

[0041] 本发明使用的术语“生成量”通常是指由本发明提供的经遗传修饰的宿主细胞生成的香兰素或其衍生物的量。衍生物可包括葡萄糖香兰素、香草醇和/或香草酸。在一些实施方案,生成量表示为由所述宿主细胞生成的香兰素或葡萄糖香兰素的产量。在其他实施方案,生成量表示为生成所述香兰素或葡萄糖香兰素时所述宿主细胞的生产率。

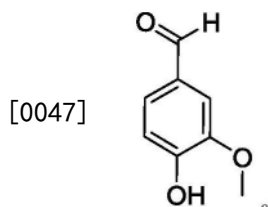
[0042] 本发明使用的术语“生产率/生产力”是指由宿主细胞生成香兰素或其衍生物的量,表示为每单位量的发酵液中生成的香兰素或葡萄糖香兰素的量(按重量计),其中所述宿主细胞根据时间(每小时)进行培养(按体积计)。衍生物可包括葡萄糖香兰素、香草醇和/或香草酸。

[0043] 本发明使用的术语“产量/产率”是指由宿主细胞生成的香兰素或其衍生物的量,表示为宿主细胞消耗的每单位量的碳源生成的香兰素或葡萄糖香兰素的量,按重量计。衍生物可包括葡萄糖香兰素、香草醇和/或香草酸。

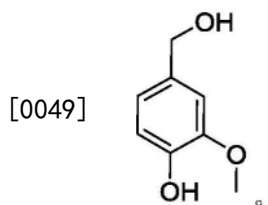
[0044] 本发明使用的术语“效价”是指由宿主细胞生成的香兰素或其衍生物,表示为每体积培养基生成的香兰素或葡萄糖香兰素或其他衍生物的量。衍生物可包括葡萄糖香兰素、香草醇和/或香草酸。

[0045] 本发明使用的术语化合物(例如,香草酸或其他化合物)的“不可检测水平”是指化合物的水平/含量太低而不能通过标准技术进行测定和/或分析以检测所述化合物。譬如,所述术语包括由本领域已知的典型分析方法无法检测到的化合物的水平。

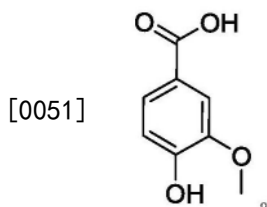
[0046] 术语“香兰素”是指化合物香草醛,包括香兰素的任何立体异构体。香兰素的化学名称是4-羟基-3-甲氧基苯甲醛。在特定实施方案,所述术语是指具有以下结构的化合物:



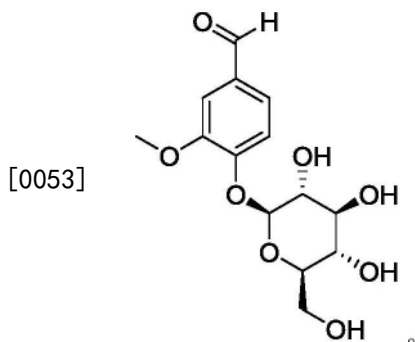
[0048] 术语“香草醇”是指化合物香兰醇,包括香草醇的任何立体异构体。香草醇的化学名称是4-(羟甲基)-2-甲氧基苯酚。在特定实施方案,所述术语是指具有以下结构的化合物:



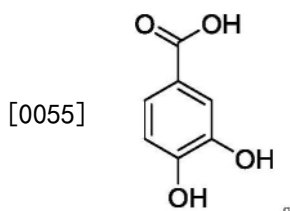
[0050] 术语“香草酸”是指化合物香兰素酸,包括香草酸的任何立体异构体。香草酸的化学名称是4-羟基-3-甲氧基苯甲酸。在特定实施方案,所述术语是指具有以下结构的化合物:



[0052] 术语“葡萄糖香兰素”是指化合物葡萄糖香兰素,包括葡萄糖香兰素的任何立体异构体。葡萄糖香兰素的化学名称是3-甲氧基-4-[(2S,3R,4S,5S,6R)-3,4,5-三羟基-6-(羟甲基)四氢-2H-吡喃-2-基]氧基苯甲醛。在特定实施方案,所述术语是指具有以下结构的化合物:



[0054] 术语“原儿茶酸(protecatechuic acid)”是指化合物原儿茶酸,包括原儿茶酸的任何立体异构体。原儿茶酸的化学名称是3,4-二羟基苯甲酸。在特定实施方案,所述术语是指具有以下结构的化合物:



[0056] 本发明使用的术语“变体”是指通过氨基酸插入、缺失、突变和/或置换而不同于具体列举的“参考”多肽(例如,野生型序列)的多肽,但保留与所述参考多肽基本相似的活性。在一些实施方案,所述变体通过重组DNA技术或通过诱变而产生。在一些实施方案,变体多肽与其参考多肽的不同之处在于一个碱基残基置换另一个碱基残基(即,Arg置换Lys),一个疏水残基置换另一个疏水残基(即,Leu置换Ile),或一个芳香族残基置换另一个芳香族残基(即,Phe置换Tyr)等。在一些实施方案,变体包括类似物,其中实现保守置换导致所述参考序列的基本结构类似。此类保守置换的实例包括但不限于,谷氨酸置换天冬氨酸,反之亦然;谷氨酰胺置换天冬酰胺,反之亦然;丝氨酸置换苏氨酸,反之亦然;赖氨酸置换精氨酸,反之亦然;或者任何异亮氨酸、缬氨酸或亮氨酸之间的彼此置换。

[0057] 在上下文中或两个或更多个核酸或蛋白质序列中,本发明使用的术语“序列同一性”或“同一性百分比”是指两个或更多个序列或子序列是相同,或所述序列或子序列具有特定百分比的相同氨基酸残基或核苷酸。譬如,当进行比较和比对以在比较窗口上进行最大对应,或者使用序列比较算法或通过手动比对和目视检查进行测定的指定区域时,所述序列可在所述指定区域与参考序列具有至少60%,至少65%,至少70%,至少75%,至少80%,至少85%,至少90%,至少91%,至少92%,至少93%,至少94%,至少95%,至少96%,至少97%,至少98%,至少99%,或更高的同一性。譬如,通过计算所述序列中相同核苷酸

(或氨基酸残基)的数量除以总核苷酸(或氨基酸残基)的长度减去任何空位(gap)的长度的比值来确定同一性百分比。

[0058] 为方便起见,可使用本领域已知的计算机程序和数学算法来确定两个序列之间的同一性程度。计算序列同一性百分比的此类算法通常考虑所述比较区域上的序列空位和错配。用于比较和比对序列的程序,如Clustal W(序列比对W)(Thompson et al., (1994) Nucleic Acids Res., 22:4673-4680), ALIGN(Myers et al., (1988) CABIOS, 4:11-17), FASTA(Pearson et al., (1988) PNAS, 85:2444-2448; Pearson(1990), Methods Enzymol., 183:63-98), 以及空位的BLAST(Altschul et al., (1997) Nucleic Acids Res., 25:3389-3402), 均可用于此目的。所述BLAST或BLAST 2.0(Altschul et al., J. Mol. Biol. 215:403-10, 1990) 可从若干来源得到, 包括国家生物信息中心(NCBI)和因特网, 用于与序列分析程序BLASTP、BLASTN、BLASTX、TBLASTN和TBLASTX结合使用。更多信息可从NCBI网站上获悉。

[0059] 在某些实施方案, 序列比对和同一性百分比计算可使用BLAST程序采用其标准默认参数来确定。对于核苷酸序列比对和序列同一性计算, BLASTN程序可以其默认参数(空位开放罚分(Gap opening penalty)=5, 空位延伸罚分(Gap extension penalty)=2, 核匹配(Nucleic match)=2, 核不匹配(Nucleic mismatch)=-3, 期望值(Expectation value)=10.0, 字长(Word size)=11, 查询范围中的最大匹配数=0) 进行使用。对于多肽序列比对和序列同一性计算, BLASTP程序可以其默认参数(比对矩阵(Alignment matrix)=BLOSUM62; 空位罚分(Gap costs): 存在(Existence)=11, 延伸(Extension)=1; 组成调整(Compositional adjustments)=条件组成得分(Conditional compositional score), 矩阵调整; 期望值=10.0; 字长(Word size)=6; 查询范围中的最大匹配数=0) 进行使用。或者, 使用以下程序和参数: 克隆管理组件(Clone Manager Suite)的比对加强版(Align Plus)软件, 版本5(Sci-Ed软件); DNA比较: 总体比较/整体联配(Global comparison), 标准线性打分矩阵(Standard Linear Scoring matrix), 不匹配罚分=2, 开放空位罚分=4, 延伸空位罚分=1。氨基酸比较: 总体比较, BLOSUM 62打分矩阵。在本发明所述的实施方案中, 使用BLASTN或BLASTP程序使用其默认参数来计算序列同一性。在本发明所述的实施方案中, 使用Clustal W使用建议的默认参数进行两个或更多个序列的序列比对(Dealign输入序列: 否; Mbed样聚类引导树(Mbed-like clustering guide-tree): 是; Mbed样聚类迭代(Mbed-like clustering iteration): 是; 组合迭代次数: 默认(0); 最大引导树迭代次数(Max guide tree iterations): 默认; 最大HMM迭代次数: 默认; 指令(Order): 输入(input))。

[0060] 核酸、表达盒和宿主细胞

[0061] 一方面, 本发明提供了表达可用于生成香兰素和/或葡萄糖香兰素的一种或多种酶的核酸、表达载体和宿主细胞。另一方面, 本发明提供了在基因中包含一个或多个缺失的宿主细胞, 其中所述一个或多个缺失可用于生成香兰素和/或葡萄糖香兰素。另一方面, 本发明提供了包含一个或多个缺失并进一步包含一种或多种所述酶的宿主细胞。所述酶和缺失在本发明中详细描述。在某些实施方案, 所述宿主细胞可由培养基中的碳源生成香兰素和/或葡萄糖香兰素。在某些实施方案, 与亲本菌株相比, 所述宿主细胞提供改进的产量和/或生产率。在某些实施方案, 与亲本菌株相比, 所述宿主细胞提供副产物、中间体和/或副产品, 例如, 香草酸。示例性副产物、中间体和/或副产品包括香草酸、香草醇、葡萄糖香草酸、葡萄糖香草

醇和原儿茶醛。

[0062] 在某些实施方案,与亲本菌株相比,根据本发明实施方案的宿主细胞生成多于至少5%、至少10%、至少15%、至少20%、或至少25%的总香兰素或葡萄糖香兰素。在某些实施方案,与亲本菌株相比,根据本发明实施方案的宿主细胞生成多于至少5%、至少10%、至少15%、至少20%、至少25%的总香兰素。在某些实施方案,与亲本菌株相比,根据本发明实施方案的宿主细胞生成多于至少5%、至少10%、至少15%、至少20%、至少25%的总葡萄糖香兰素。在某些实施方案,与亲本菌株相比,根据本发明实施方案的宿主细胞生成的香草酸少2倍、3倍、4倍、5倍或10倍。在某些实施方案,百分比增加是关于香兰素或葡萄糖香兰素效价(g/L)。在某些实施方案,百分比增加是关于香兰素或葡萄糖香兰素产量(重量%)。在某些实施方案,百分比增加是关于香兰素或葡萄糖香兰素生产率(g/L/h)。在某些实施方案,百分比增加是关于所生成的香兰素或葡萄糖香兰素总质量(g)。技术人员将认识到,所生成的总香兰素和/或葡萄糖香兰素可作为生成的实际化合物以及由香兰素和/或葡萄糖香兰素生成的任何下游化合物的总和来测定,如本发明实施例和附图中所示。在某些实施方案,与亲本菌株相比,根据本发明实施方案的宿主细胞生成增加量的香兰素和/或葡萄糖香兰素,并生成较少的香草酸。

[0063] 在有利的实施方案中,宿主细胞包含能够制备香兰素和/或葡萄糖香兰素的一种或多种酶促途径,所述途径可单独进行或一同进行。

[0064] 一方面,本发明提供了包含HFD1缺失的宿主细胞。如以下实施例所述,HFD1编码能够将香兰素转化为香草酸的酶Hfd1。由于香草酸对宿主细胞具有潜在毒性,并且是最终产品中不希望有的杂质,因此其是不需要的发酵副产物。此外,香草酸的积累会使纯化更加困难。此外,香兰素到香草酸的逆反应会导致在香草酸和香兰素之间引入无用的循环。香草酸到香兰素的每一次正向反应均会消耗宝贵的细胞ATP和NADPH,然后,随后又由香兰素转化回香草酸所浪费。在某些实施方案,所述宿主细胞是酿酒酵母(*S.cerevisiae*)。如以下实施例中所述,Hfd1是负责在酿酒酵母(*S.cerevisiae*)中将香兰素转化为香草酸的主要已知酶。在酿酒酵母(*S.cerevisiae*)以外的宿主细胞中,HFD1的同源物被删除。优选地,删除HFD1的所有拷贝。例如,在具有一个HFD1拷贝的单倍体细胞中,所述拷贝被删除。在具有两个HFD1拷贝的二倍体细胞中,所述两个拷贝均被删除。在具有多个HFD1拷贝的任何细胞中,优选地删除每一个拷贝。可通过本领域技术人员显而易见的任何技术删除HFD1基因。有用的技术包括基于同源重组和聚合酶链反应(PCR)的技术。

[0065] 在进一步的实施方案中,上述宿主细胞还包含一个或多个缺失和/或一个或多个可用于生成香兰素和/或葡萄糖香兰素的表达基因。

[0066] 在某些实施方案,宿主细胞进一步包含ADH6的缺失。在酿酒酵母(*S.cerevisiae*)以外的宿主细胞中,ADH6的同源物被删除。优选地,删除ADH6的所有拷贝。例如,在具有一个ADH6拷贝的单倍体细胞中,所述拷贝被删除。在具有两个ADH6拷贝的二倍体细胞中,所述两个拷贝均被删除。在具有多个ADH6拷贝的任何细胞中,优选地删除每一个拷贝。可通过本领域技术人员显而易见的任何技术删除ADH6基因。有用的技术包括基于同源重组和聚合酶链反应(PCR)的技术。

[0067] 在某些实施方案,宿主细胞进一步包含GRE2的缺失。在酿酒酵母(*S.cerevisiae*)以外的宿主细胞中,GRE2的同源物被删除。优选地,删除GRE2的所有拷贝。例如,在具有一个

GRE2拷贝的单倍体细胞中,所述拷贝被删除。在具有两个GRE2拷贝的二倍体细胞中,所述两个拷贝均被删除。在具有多个GRE2拷贝的任何细胞中,优选地删除每一个拷贝。可通过本领域技术人员显而易见的任何技术删除GRE2基因。有用的技术包括基于同源重组和聚合酶链反应(PCR)的技术。

[0068] 在某些实施方案,宿主细胞进一步包含YGL039W的缺失。在酿酒酵母(*S.cerevisiae*)以外的宿主细胞中,YGL039W的同源物被删除。优选地,删除YGL039W的所有拷贝。例如,在具有一个YGL039W拷贝的单倍体细胞中,所述拷贝被删除。在具有两个YGL039W拷贝的二倍体细胞中,所述两个拷贝均被删除。在具有多个YGL039W拷贝的任何细胞中,优选地删除每一个拷贝。可通过本领域技术人员显而易见的任何技术删除YGL039W基因。有用的技术包括基于同源重组和聚合酶链反应(PCR)的技术。

[0069] 在特定实施方案,所述宿主细胞进一步包含可用于生成香兰素或葡萄糖香兰素的途径的酶。此类途径酶先前已进行描述,包括以下文献中记载的那些:Hansen et al., Appl.Environ.Microbiol. (2009) 75 (9):2765-2774;U.S.6,372,461B1;U.S.10,066,252B1;U.S.10,208,293B2;其各自通过引用其全文并入本发明。

[0070] 在某些实施方案,所述宿主细胞进一步包含3-脱氢奎尼酸合成酶或AroB。有用的AroB基因和酶均是已知的。有用的AroB多肽亦是已知的。有用的AroB基因和酶包括大肠杆菌(*E.coli*)的那些基因和酶。示例可于UniProtKB P07639中获悉。在优选实施方案,所述宿主细胞进一步表达或过表达大肠杆菌(*E.coli*)AroB。

[0071] 在某些实施方案,所述宿主细胞进一步包含3-脱氢奎尼酸脱水酶或AroD。有用的AroD基因和酶均是已知的。有用的AroD多肽亦是已知的。有用的AroD基因和酶包括大肠杆菌(*E.coli*)的那些基因和酶。示例可在UniProtKB P05194中获悉。在优选实施方案,所述宿主细胞进一步表达或过表达大肠杆菌(*E.coli*)AroD。

[0072] 在某些实施方案,宿主细胞进一步包含磷酸-2-脱氢-3-脱氧庚酸醛缩酶、Tyr-敏感性或AroF。有用的AroF基因和酶均是已知的。有用的AroB多肽亦是已知的。有用的AroF基因和酶包括大肠杆菌(*E.coli*)的那些基因和酶。示例可在UniProtKB P00888中获悉。在优选实施方案,所述宿主细胞进一步表达或过表达大肠杆菌(*E.coli*)AroF。在某些实施方案,AroF是抗反馈的(J.Bacteriol.November 1990172:6581-6584,incorporated by reference in its entirety)。

[0073] 在某些实施方案,所述宿主细胞进一步包含3-脱氢莽草酸脱水酶或AroZ。有用的AroZ基因和酶均是已知的。有用的3DSD多肽亦是已知的。有用的AroZ基因和酶包括疏毛柄孢壳菌(*Podospira pauciseta*)、玉蜀黍黑粉菌(*Ustilago maydis*)、约氏红球菌(*Rhodococcus jostii*)、不动杆菌属(*Acinetobacter* sp.)、黑曲霉(*Aspergillus niger*)和粗糙脉孢菌(*Neurospora crassa*)的那些基因和酶。示例可在GenBank登录号CAD60599、XP_001905369.1、XP_761560.1、ABG93191.1、AAC37159.1和XM_001392464中获悉。在优选实施方案,所述宿主细胞进一步表达或过表达疏毛柄孢壳菌(*Podospira pauciseta*)AroZ。

[0074] 在某些实施方案,所述宿主细胞进一步包含ACAR。有用的ACAR基因和酶均是已知的。有用的ACAR多肽亦是已知的。有用的ACAR基因和酶包括诺卡氏菌属(*Nocardia* sp.)的那些基因和酶。示例可在GenBank登录号AY495697中获悉。在优选实施方案,所述宿主细胞进一步表达或过表达艾阿华诺卡氏菌(*Nocardia iowensis*)ACAR。

[0075] 在某些实施方案,所述宿主细胞进一步包含PPTASE。有用的PPTASE基因和酶均是已知的。有用的PPTASE多肽亦是已知的。有用的PPTASE基因和酶包括大肠杆菌(*E.coli*)、谷氨酸棒杆菌(*Corynebacterium glutamicum*)和皮疽诺卡氏菌(*Nocardia farcinica*)的那些基因和酶。示例可在GenBank登录号NP_601186、BAA35224和YP_120266中获悉。在优选实施方案,所述宿主细胞进一步表达或过表达谷氨酸棒杆菌(*Cornybacterium glutamicum*) PPTASE。

[0076] 过表达可根据本领域技术人员显而易见的任何技术进行。在某些实施方案,基因由可用于宿主细胞的启动子过表达。在某些实施方案,基因由酿酒酵母(*S.cerevisiae*)启动子过表达。在某些实施方案,所述启动子选自由以下组成的组:pPGK1、pTDH3、pEN02、pADH1、pTPI1、pTEF1、pTEF2、pTEF3、pGAL1、pGAL2、pGAL7、pGAL10、GAL1、pRPL3、pRPL15A、pRPL4、pRPL8B、pSSA1、pSSB1、pCUP1、pTPS1、pHXT7、pADH2、pCYC1和pPDA1。在某些实施方案,基因由GAL启动子过表达。在某些实施方案,基因由选自由以下组成的组的启动子过表达:pGAL1、pGAL2、pGAL7、pGAL10、及其变体。

[0077] 在某些实施方案,宿主细胞中的一个、一些或所有异源启动子均是诱导型。诱导型启动子系统可以是本领域技术人员公认的任何一种。在特定实施方案,启动子可由麦芽糖诱导。在有利的实施方案中,宿主细胞包含可由麦芽糖诱导的GAL调节子。由麦芽糖进一步抑制或诱导的Gal调节子的实例描述于PCT申请公开文本W02015/020649、W02016/210343和W02016210350中,其各自通过引用其全文并入本发明。在某些实施方案,麦芽糖可转换菌株通过在麦芽糖响应性启动子(例如,pMAL32)的控制下染色体整合GAL80的拷贝而构建在不可转换菌株之上。在某些实施方案,GAL80基因产物因温度敏感性而突变,例如,以便于进一步控制。在某些实施方案,GAL80基因产物与温度敏感性多肽融合。在某些实施方案,GAL80基因产物与温度敏感性DHFR多肽或片段融合。生成可转换法呢烯的可转换菌株的其他描述记载于美国专利申请公开号US 2016/0177341和PCT申请公开号W0 2016/210350中,其各自通过引用其全文并入本发明。

[0078] 对于上述每种多肽和核酸,宿主细胞可包含其变体。在某些实施方案,相对于相关多肽,变体可包含多达15、10、9、8、7、6、5、4、3、2或1个氨基酸置换。在某些实施方案,相对于参考多肽,变体可包含多达15、10、9、8、7、6、5、4、3、2或1个保守氨基酸置换。在某些实施方案,可针对宿主细胞优化本发明所述的任何核酸,例如密码子优化。下面详细描述了变体及优化。

[0079] 在某些实施方案,除非上文另有说明,否则另外的酶是原生或天然的。天然酶可由经密码子优化的核酸表达。在有利的实施方案中,所述另外的酶是异源的。在某些实施方案,两种或更多种酶可在一种多肽中进行组合。

[0080] 细胞菌株

[0081] 可用于本发明提供的组合物和方法的宿主细胞包括古细菌细胞、原核细胞或真核细胞。

[0082] 合适的原核宿主包括但不限于,多种革兰氏阳性、革兰氏阴性或革兰氏变种细菌中的任一种。实例包括但不限于,属于以下属的细胞:土壤杆菌属(*Agrobacterium*),脂环酸芽孢杆菌属(*Alicyclobacillus*),鱼腥藻属(*Anabaena*),蓝细菌属(*Anacystis*),节细菌属(*Arthrobacter*),固氮菌属(*Azobacter*),芽孢杆菌属(*Bacillus*),短杆菌属

(*Brevibacterium*), 着色菌属 (*Chromatium*), 梭菌属 (*Clostridium*), 棒状杆菌属 (*Corynebacterium*), 肠杆菌属 (*Enterobacter*), 欧文氏菌属 (*Erwinia*), 埃希氏杆菌属 (*Escherichia*), 乳酸杆菌属 (*Lactobacillus*), 乳球菌属 (*Lactococcus*), 中慢生根瘤菌属 (*Mesorhizobium*), 甲基杆菌属 (*Methylobacterium*), 细杆菌属 (*Microbacterium*), 席藻属 (*Phormidium*), 假单胞菌属 (*Pseudomonas*), 红细菌属 (*Rhodobacter*), 红假单胞菌属 (*Rhodopseudomonas*), 红螺菌属 (*Rhodospirillum*), 红球菌属 (*Rhodococcus*), 沙门氏菌属 (*Salmonella*), 栅藻属 (*Scenedesmun*), 沙雷氏菌属 (*Serratia*), 志贺氏菌属 (*Shigella*), 葡萄球菌属 (*Staphylococcus*), 链霉菌属 (*Streptomyces*), 聚球蓝细菌属 (*Synnecoccus*) 和发酵单胞菌属 (*Zymomonas*)。原核菌株的实例包括但不限于: 枯草芽孢杆菌 (*Bacillus subtilis*), 解淀粉芽孢杆菌 (*Bacillus amyloliquefacines*), 产氨短杆菌 (*Brevibacterium ammoniagenes*), 嗜氨短杆菌 (*Brevibacterium immariophilum*), 拜氏梭菌 (*Clostridium beigerinckii*), 阪崎肠杆菌 (*Enterobacter sakazakii*), 大肠杆菌 (*Escherichia coli*), 乳酸乳球菌 (*Lactococcus lactis*), 百脉根中生根瘤菌 (*Mesorhizobium loti*), 绿脓假单胞菌 (*Pseudomonas aeruginosa*), 迈氏假单胞菌 (*Pseudomonas mevalonii*), 普迪卡假单胞菌 (*Pseudomonas pudica*), 荚膜红细菌 (*Rhodobacter capsulatus*), 类球红细菌 (*Rhodobacter sphaeroides*), 深红红螺菌 (*Rhodospirillum rubrum*), 肠道沙门氏菌 (*Salmonella enterica*), 伤寒沙门氏菌 (*Salmonella typhi*), 鼠伤寒沙门氏菌 (*Salmonella typhimurium*), 痢疾志贺氏菌 (*Shigella dysenteriae*), 福氏志贺菌 (*Shigella flexneri*), 宋内志贺菌 (*Shigella sonne*) 和金黄色葡萄球菌 (*Staphylococcus aureus*)。在特定实施方案, 所述宿主细胞是大肠杆菌 (*Escherichia coli*) 细胞。

[0083] 合适的古细菌宿主包括但不限于属于以下属的细胞: 气火菌属 (*Aeropyrum*), 古球菌属 (*Archaeoglobus*), 盐杆菌属 (*Halobacterium*), 产甲烷球菌属 (*Methanococcus*), 甲烷杆菌属 (*Methanobacterium*), 火球菌属 (*Pyrococcus*), 硫化叶菌属 (*Sulfolobus*), 和热原体属 (*Thermoplasma*)。古细菌菌株的实例包括但不限于: 闪烁古生球菌 (*Archaeoglobus fulgidus*), 盐杆菌属 (*Halobacterium sp.*), 詹氏甲烷球菌 (*Methanococcus jannaschii*), 嗜热自养甲烷杆菌 (*Methanobacterium thermoautotrophicum*), 嗜酸热原体 (*Thermoplasma acidophilum*), 火山热原体 (*Thermoplasma volcanium*), 掘越氏热球菌 (*Pyrococcus horikoshii*), 深海火球菌 (*Pyrococcus abyssi*), 和敏捷气热菌 (*Aeropyrum pernix*)。

[0084] 合适的真核宿主包括但不限于真菌细胞、藻类细胞、昆虫细胞和植物细胞。在一些实施方案, 可用于本发明方法的酵母包括已被微生物保藏中心 (例如, IF0、ATCC等) 保藏并属于以下属的酵母: 芽孢酵母属 (*Aciculoconidium*), 神食酵母属 (*Ambrosiozyma*), 节束酵母属 (*Arthroascus*), *Arxiozyma*, 阿舒囊霉属 (*Ashbya*), *Babjevia*, 本森顿酵母属 (*Bensingtonia*), *Botryoascus*, *Botryozyma*, 酒香酵母属 (*Brettanomyces*), 布勒掷孢酵母属 (*Bullera*), 布勒担孢酵母属 (*Bulleromyces*), 假丝酵母属 (*Candida*), 固囊酵母属 (*Citeromyces*), 棒孢酵母属 (*Clavispora*), 隐球菌属 (*Cryptococcus*), 产黑色素酵母属 (*Cystofilobasidium*), 德巴利氏酵母属 (*Debaryomyces*), 德克酵母属 (*Dekkara*), 拟双足囊菌属 (*Dipodascopsis*), 双足囊菌属 (*Dipodascus*), *Eeniella*, *Endomycopsella*, *Eremascus*,

假囊酵母属(*Eremothecium*),担孢酵母属(*Erythrobasidium*),*Fellomyces*,线黑粉酵母属(*Filobasidium*),耐碱酵母属(*Galactomyces*),地丝菌属(*Geotrichum*),季氏酵母属(*Guilliermondella*),有孢汉逊酵母属(*Hanseniaspora*),汉逊酵母属(*Hansenula*),*Hasegawaea*,胶珊瑚属(*Holtermannia*),*Hormoascus*,生丝毕赤酵母属(*Hyphopichia*),伊萨酵母属(*Issatchenkia*),克勒克酵母属(*Kloeckera*),孢克勒克酵母属(*Kloeckeraspora*),克鲁维酵母属(*Kluyveromyces*),*Kondoa*,*Kuraishia*,克氏担孢酵母属(*Kurtzmanomyces*),白冬孢酵母属(*Leucosporidium*),油脂酵母属(*Lipomyces*),路德酵母属(*Lodderomyces*),马拉色氏霉菌属(*Malassezia*),梅奇酵母属(*Metschnikowia*),木拉克酵母属(*Mrakia*),油脂酵母属无性属(*Myxozyma*),拿逊酵母属(*Nadsonia*),*Nakazawaea*,针孢酵母属(*Nematospora*),甲醇诱导型酵母属(*Ogataea*),卵孢酵母属(*Oosporidium*),管囊酵母属(*Pachysolen*),厚壁孢酵母(*Phachytichospora*),法夫酵母属(*Phaffia*),毕赤酵母属(*Pichia*),红冬孢酵母属(*Rhodosporidium*),红酵母属(*Rhodotorula*),酵母属(*Saccharomyces*),类酵母属(*Saccharomycodes*),覆膜孢酵母属(*Saccharomycopsis*),齐藤酵母属(*Saitoella*),*Sakaguchia*,*Saturnospora*,裂芽酵母孢子菌属(*Schizoblastosporion*),裂殖酵母属(*Schizosaccharomyces*),许旺酵母属(*Schwanniomyces*),锁掷酵母属(*Sporidiobolus*),掷孢酵母属(*Sporobolomyces*),原孢酵母属(*Sporopachydermia*),冠孢酵母属(*Stephanoascus*),梗孢酵母属(*Sterigmatomyces*),拟梗孢酵母属(*Sterigmatosporidium*),*Symbiotaphrina*,合轴酵母属(*Sympodiomyces*),*Sympodiomycopsis*,有孢圆酵母属(*Torulaspora*),*Trichosporiella*,毛孢子菌属(*Trichosporon*),三角酵母属(*Trigonopsis*),*Tsuchiyaea*,*Udeniomyces*,*Waltomyces*,威克酵母属(*Wickerhamia*),拟威克酵母属(*Wickerhamiella*),拟威尔酵母属(*Williopsis*),*Yamadazyma*,耶氏酵母属(*Yarrowia*),接合囊酵母属(*Zygoascus*),接合酵母属(*Zygosaccharomyces*),接合拟威尔酵母属(*Zygowilliopsis*),和*Zygozyma*等等。

[0085] 在一些实施方案,所述宿主微生物是酿酒酵母(*Saccharomyces cerevisiae*),巴斯德毕赤酵母(*Pichia pastoris*),粟酒裂殖酵母(*Schizosaccharomyces pombe*),布鲁赛尔德克酵母(*Dekkera bruxellensis*),乳酸克鲁维酵母(*Kluyveromyces lactis*,此前称为乳酸酵母(*Saccharomyces lactis*)),马克斯克鲁维酵母(*Kluyveromyces marxianus*),食腺嘌呤芽生葡萄孢酵母(*Arxula adeninivorans*),或多形汉逊酵母(*Hansenula polymorpha*) (现称为安格斯毕赤酵母(*Pichia angusta*))。在一些实施方案,所述宿主微生物是假丝酵母属(*Candida*)的菌株,例如解脂假丝酵母(*Candida lipolytica*),吉利蒙假丝酵母(*Candida guilliermondii*),克鲁斯假丝酵母(*Candida krusei*),假热带假丝酵母(*Candida pseudotropicalis*)或产朊假丝酵母(*Candida utilis*)的菌株。

[0086] 在特定实施方案,所述宿主微生物是酿酒酵母(*Saccharomyces cerevisiae*)。在一些实施方案,所述宿主是酿酒酵母(*Saccharomyces cerevisiae*)的菌株,所述酿酒酵母的菌株是选自由贝克氏(*Baker's*)酵母、CEN.PK、CBS 7959、CBS 7960、CBS 7961、CBS 7962、CBS 7963、CBS 7964、IZ-1904、TA、BG-1、CR-1、SA-1、M-26、Y-904、PE-2、PE-5、VR-1、BR-1、BR-2、ME-2、VR-2、MA-3、MA-4、CAT-1、CB-1、NR-1、BT-1、和AL-1组成的组。在一些实施方案,所述宿主微生物是酿酒酵母(*Saccharomyces cerevisiae*)的菌株,所述酿酒酵母的菌株是选自由PE-2、CAT-1、VR-1、BG-1、CR-1、和SA-1组成的组。在特定实施方案,酿酒酵母

(*Saccharomyces cerevisiae*) 的菌株是 PE-2。在另一特定实施方案, 酿酒酵母 (*Saccharomyces cerevisiae*) 的菌株是 CAT-1。在另一特定实施方案, 酿酒酵母 (*Saccharomyces cerevisiae*) 的菌株是 BG-1。

[0087] 在一些实施方案, 所述宿主微生物是适于工业发酵的微生物。在特定实施方案, 所述微生物适应于在高溶剂浓度、高温、高压、扩大的底物利用率、营养限制、由糖和盐引起的渗透应力、酸度、亚硫酸盐和细菌污染、或其组合下存活, 所述这些是公认的工业发酵环境的应激条件。

[0088] 生成香兰素或葡萄糖香兰素的方法

[0089] 另一方面, 本发明提供了生成香兰素或葡萄糖香兰素的方法, 所述方法包含以下步骤: (a) 在适于制备香兰素或葡萄糖香兰素化合物的条件下, 在含有碳源的培养基中培养能够生成所述香兰素或葡萄糖香兰素的任何经遗传修饰的宿主细胞的种群; 和 (b) 从所述培养基回收所述香兰素或葡萄糖香兰素化合物。本领域技术人员将认识到, 所生成的化合物的量可通过测定所述化合物本身的量, 或更优选地, 测定所述化合物及所述化合物的衍生物的量来评估。例如, 可由所生成的香兰素、香草醇、葡萄糖香兰素和葡萄糖香草醇的总量来评估所生成香兰素的量。

[0090] 在一些实施方案, 与不包含一种或多种修饰的亲本细胞或仅包含所述经遗传修饰的宿主细胞的一种或多种修饰的子集, 但在其他方面在遗传上是相同的亲本细胞相比, 所述经遗传修饰的宿主细胞生成增加量的香兰素或葡萄糖香兰素、或其衍生物 (例如香草醇或葡萄糖香草醇)。在一些实施方案, 所述增加量为至少 1%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%、或大于 100%, 例如, 以产量、生成量、和/或生产率计, 或以克/升细胞培养物, 毫克/克干细胞重量计, 或者基于每单位体积的细胞培养物, 基于每单位干细胞重量, 基于每单位时间的每单位体积的细胞培养物, 或基于每单位时间的每单位干细胞重量计, 进行测定。

[0091] 在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 0.25 克/升发酵培养基。在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 0.5 克/升发酵培养基。在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 0.75 克/升发酵培养基。在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 1 克/升发酵培养基。在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 5 克/升发酵培养基。在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 10 克/升发酵培养基。在一些实施方案, 所述香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇) 以从约 10 至约 50 克/升细胞培养物, 从约 10 至约 15 克/升细胞培养物, 超过约 15 克/升细胞培养物, 超过约 20 克/升细胞培养物, 超过约 25 克/升细胞培养物, 或超过约 30 克/升细胞培养物的量生成。

[0092] 在一些实施方案, 所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物 (例如香草醇或葡萄糖香草醇), 其大于约 50 毫克每克干细胞重量。在一些此类实施方案中, 所述香兰素或葡萄糖香兰素以从约 50 至约 1500 毫克, 超过约 100 毫克, 超过约 150 毫克, 超过约

200毫克,超过约250毫克,超过约500毫克,超过约750毫克,或超过约1000毫克每克干细胞重量的量生成。

[0093] 在一些实施方案,所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇),其比由亲本细胞生成的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇)水平高至少约10%,至少约15%,至少约20%,至少约25%,至少约30%,至少约35%,至少约40%,至少约45%,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约2倍,至少约2.5倍,至少约5倍,至少约10倍,至少约20倍,至少约30倍,至少约40倍,至少约50倍,至少约75倍,至少约100倍,至少约200倍,至少约300倍,至少约400倍,至少约500倍,或至少约1000倍或更多,基于每单位体积细胞培养物。

[0094] 在一些实施方案,所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇),其比由亲本细胞生成的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇)水平高至少约10%,至少约15%,至少约20%,至少约25%,至少约30%,至少约35%,至少约40%,至少约45%,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约2倍,至少约2.5倍,至少约5倍,至少约10倍,至少约20倍,至少约30倍,至少约40倍,至少约50倍,至少约75倍,至少约100倍,至少约200倍,至少约300倍,至少约400倍,至少约500倍,或至少约1000倍或更多,基于每单位干细胞重量。

[0095] 在一些实施方案,所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇),其比由亲本细胞生成的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇)水平高至少约10%,至少约15%,至少约20%,至少约25%,至少约30%,至少约35%,至少约40%,至少约45%,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约2倍,至少约2.5倍,至少约5倍,至少约10倍,至少约20倍,至少约30倍,至少约40倍,至少约50倍,至少约75倍,至少约100倍,至少约200倍,至少约300倍,至少约400倍,至少约500倍,或至少约1000倍或更多,基于每单位时间的每单位体积细胞培养物。

[0096] 在一些实施方案,所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇),其比由亲本细胞生成的香兰素或葡萄糖香兰素或其衍生物(例如香草醇或葡萄糖香草醇)水平高至少约10%,至少约15%,至少约20%,至少约25%,至少约30%,至少约35%,至少约40%,至少约45%,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约2倍,至少约2.5倍,至少约5倍,至少约10倍,至少约20倍,至少约30倍,至少约40倍,至少约50倍,至少约75倍,至少约100倍,至少约200倍,至少约300倍,至少约400倍,至少约500倍,或至少约1000倍或更多,基于每单位时间的每单位干细胞重量。

[0097] 在大多数实施方案中,所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素可由含诱导化合物或不含抑制化合物来诱导。在不含诱导化合物或含抑制化合物的情况下,可容易地操纵此类宿主细胞。然后加入诱导化合物,或减少抑制化合物,以诱导宿主细胞生成升高水平的香兰素或葡萄糖香兰素。在其他实施方案,所述宿主细胞生成升高水平的香兰素或葡萄糖香兰素可通过改变培养条件,例如改变生长温度、培养基成分等来诱导。在某些实施方案,生成香兰素的酶在细胞的生长阶段期间受到麦芽糖阻遏,并且生成香兰素的酶在发酵的表达阶段期间被表达。有用的启动子和技术均在US 2018/0171341 A1中有所描述,此文献通过引用其全文并入。

[0098] 在某些实施方案,本发明提供了通过本发明方法生成的香兰素或葡萄糖香兰素或两者。在某些实施方案,本发明提供了与标准品相比具有独特同位素分布的香兰素。在某些实施方案,本发明提供了与标准品相比具有独特碳同位素分布的香兰素。碳同位素分布根据标准技术测定,例如本发明实施例中描述的那些进行测定。所述标准品可以是技术人员认为合适的任何标准品。在某些实施方案,所述标准品是用于测定 ^{14}C 活性的草酸。在某些实施方案,所述 ^{14}C 活性报告为每克碳每分钟的崩解量(dpm/g C),其可用于区分石油衍生的香兰素(dpm/g C接近0)与衍生自植物来源的香兰素,后者通常提供15-16dpm/g C。在某些实施方案,本发明提供了具有约12.9dpm/g至约14.1dpm/g的 ^{14}C 活性的香兰素。在某些实施方案,本发明提供了具有约12.9dpm/g的 ^{14}C 活性的香兰素。在某些实施方案,本发明提供了具有约14.1dpm/g的 ^{14}C 活性的香兰素。在某些实施方案,碳同位素比表示为 $\text{‰} = [(R_{\text{样品}}/R_{\text{标准品}}) - 1] \times 10^3$,其中 $R = ^{13}\text{C}/^{12}\text{C}$ 表示相对于Pee Dee Belemnite (PDB) 标准品。在某些实施方案,本发明提供了香兰素,其具有与PDB标准品的块体 $\delta^{13}\text{C}$ 偏差为约-14.8‰至约-12.8‰(千分之)。在某些实施方案,本发明提供了香兰素,其具有与PDB标准品的块体 $\delta^{13}\text{C}$ 偏差为约-12.8‰(千分之)。在某些实施方案,本发明提供了香兰素,其具有与PDB标准品的块体 $\delta^{13}\text{C}$ 偏差为约-14.8‰(千分之)。根据标准技术,例如本发明实施例中描述的那些,测定氢同位素分布。氢同位素标准品可以是技术人员认为合适的任何标准品。在某些实施方案,本发明提供的标准品是标准平均大洋水(SMOW)。在某些实施方案,氢同位素比表示为 $\text{‰} = [(R_{\text{样品}}/R_{\text{标准品}}) - 1] \times 10^3$,其中 $R = ^2\text{H}/^1\text{H}$ 表示相对于标准平均大洋水(SMOW)标准品。在某些实施方案,本发明提供了香兰素,其具有与SMOW标准品的块体 $\delta^2\text{H}$ 偏差为约-150‰至约-124‰(千分之)。在某些实施方案,本发明提供了香兰素,其具有与SMOW标准品的块体 $\delta^2\text{H}$ 偏差为约-150‰(千分之)。在某些实施方案,本发明提供了香兰素,其具有与SMOW标准品的块体 $\delta^2\text{H}$ 偏差为约-124‰(千分之)。

[0099] 培养基和培养条件

[0100] 用于微生物培养物的维持和生长的物料和方法是微生物学或发酵科学领域的技术人员所熟知的(参见,例如Bailey et al., Biochemical Engineering Fundamentals, second edition, McGraw Hill, New York, 1986)。根据宿主细胞、发酵和过程/方法的特定要求,必须考虑适当的培养基,pH值,温度,以及需氧、微需氧或厌氧条件的要求。

[0101] 本发明提供的生成香兰素或葡萄糖香兰素的方法可在合适的容器(包括但不限于细胞培养板、微量滴定板、烧瓶或发酵罐)中于合适的培养基中进行。此外,所述方法可以本领域已知的任何发酵规模进行,以支持微生物产物的工业生产。可使用任何合适的发酵罐,包括搅拌槽式发酵罐,气升式发酵罐,气泡式发酵罐,或其任何组合。在利用酿酒酵母(*Saccharomyces cerevisiae*)作为宿主细胞的特定实施方案中,菌株可在发酵罐中生长,详细记载如Kosaric, et al, Ullmann's Encyclopedia of Industrial Chemistry, Sixth Edition, Volume 12, pages 398-473, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany中所述。

[0102] 在一些实施方案,所述培养基是其中能够生成香兰素或葡萄糖香兰素的经遗传修饰的微生物可以存活,即保持生长和活力的任何培养基。在一些实施方案,所述培养基是包含可同化的碳源、氮源和磷源(phosphate source)的水性培养基。此种培养基还可包括适当的盐类、矿物质类、金属类和其他营养物类。在一些实施方案,将所述碳源以及一些或所有

必需细胞营养物增量地或连续地添加到发酵培养基中。在某些实施方案,将必需营养物的一个子集保持过量,而少数,例如一种或两种,所需营养物保持在大约通过使细胞生长有效同化所需的最低水平,例如,根据基于将碳源转化为生物质的细胞的代谢或呼吸功能的预定细胞生长曲线。

[0103] 用于培养微生物的合适条件和合适的培养基是本领域熟知的。在一些实施方案,所述合适的培养基补充有一种或多种另外的试剂,例如诱导物(例如,当编码基因产物的一个或多个核苷酸序列受诱导型启动子的控制时),阻抑物(例如,当编码基因产物的一个或多个核苷酸序列受阻抑型启动子的控制时),或选择剂(例如,选择包含所述遗传修饰的微生物的抗生素)。

[0104] 在一些实施方案,所述碳源是单糖(简单糖)、二糖、多糖、不可发酵的碳源、或其一种或多种组合。合适的单糖的非限制性实例包括葡萄糖、半乳糖、甘露糖、果糖、木糖、核糖、和其组合。合适的二糖的非限制性实例包括蔗糖、乳糖、麦芽糖、海藻糖、纤维二糖、和其组合。合适的多糖的非限制性实例包括淀粉、糖原、纤维素、几丁质、和其组合。合适的不可发酵碳源的非限制性实例包括乙酸盐、乙醇和甘油。

[0105] 所述培养基中碳源(例如葡萄糖)的浓度足以促进细胞生长,但不能高到抑制所用微生物的生长。通常,培养物采用碳源(例如葡萄糖)进行,所述碳源以达到所需生长水平和所需生物量的水平进行添加。在其他实施方案,所述培养基中碳源(例如葡萄糖)的浓度大于约1g/L,优选大于约2g/L,更优选大于约5g/L。此外,所述培养基中碳源(例如葡萄糖)的浓度通常小于约100g/L,优选小于约50g/L,更优选小于约20g/L。应当注意,对培养组分浓度的提及可以指初始和/或正在进行的组分浓度。在一些案例中,可能需要在培养期间使所述培养基耗尽碳源。

[0106] 可用于合适培养基的可同化氮的来源包括但不限于简单氮源、有机氮源和复合氮源。此类氮源包括无水氨,铵盐类,以及动物、植物和/或微生物来源的物质。合适的氮源包括但不限于,蛋白质水解产物类,微生物生物质水解产物类,蛋白胨,酵母提取物,硫酸铵,尿素和氨基酸类。通常,所述培养基中所述氮源的浓度大于约0.1g/L,优选大于约0.25g/L,更优选大于约1.0g/L。然而,超过一定浓度,向所述培养基中添加氮源对于微生物的生长是不利的。因此,所述培养基中所述氮源的浓度小于约20g/L,优选小于约10g/L,更优选小于约5g/L。此外,在一些实施例中,可能需要在培养期间使所述培养基耗尽所述氮源。

[0107] 有效的培养基可含有其他化合物,例如无机盐类、维生素类、痕量金属类、或生长促进剂类。此类其他化合物也可存在于有效培养基中的碳源、氮源或矿物源中,或者可特异性地添加至所述培养基中。

[0108] 所述培养基还可含有合适的磷源。此类磷源包括无机磷源和有机磷源两者。优选的磷源包括但不限于磷酸盐类,例如单或二元磷酸钠及单或二元磷酸钾、磷酸铵、和其混合物。通常,所述培养基中磷酸盐的浓度大于约1.0g/L,优选大于约2.0g/L,更优选大于约5.0g/L。然而,超过一定浓度,向所述培养基中添加磷酸盐对于微生物的生长是不利的。因此,所述培养基中所述磷酸盐的浓度通常小于约20g/L,优选小于约15g/L,更优选小于约10g/L。

[0109] 培养基还可含有合适的硫源。优选的硫源包括但不限于硫酸盐类,例如硫酸铵((NH₄)₂SO₄)、硫酸镁(MgSO₄)、硫酸钾(K₂SO₄)和硫酸钠(Na₂SO₄)、及其混合物。通常,所述培养

基中硫酸盐的浓度大于约1.0g/L,优选大于约3.0g/L,更优选大于约10.0g/L。然而,超过一定浓度时,向所述培养基中添加硫酸盐不利于微生物的生长。因此,所述培养基中所述硫酸盐的浓度通常小于约50g/L,优选小于约30g/L,更优选小于约20g/L。

[0110] 合适的培养基还可包括镁源,优选地以生理学上可接受的盐的形式,例如七水合硫酸镁,尽管可使用浓度为贡献相似量的镁的其他镁源。通常,所述培养基中镁的浓度大于约0.5g/L,优选大于约1.0g/L,更优选大于约2.0g/L。然而,超过一定浓度,向所述培养基中添加镁对于微生物的生长是不利的。因此,所述培养基中镁的浓度通常小于约10g/L,优选小于约5g/L,更优选小于约3g/L。此外,在一些实施例中,可能需要在培养期间使所述培养基耗尽镁源。

[0111] 在一些实施方案,所述培养基还可包含生物学上可接受的螯合剂,例如二水合柠檬酸三钠。在此类实施例中,所述培养基中螯合剂的浓度大于约0.2g/L,优选大于约0.5g/L,更优选大于约1g/L。然而,超过一定浓度,向所述培养基中添加螯合剂对于微生物的生长是不利的。因此,所述培养基中螯合剂的浓度通常小于约10g/L,优选小于约5g/L,更优选小于约2g/L。

[0112] 所述培养基最初还可包括生物学上可接受的酸或碱以维持所述培养基的所需pH值。生物学上可接受的酸包括但不限于,盐酸、硫酸、硝酸、磷酸、和其混合物。生物学上可接受的碱包括但不限于,氢氧化铵、氢氧化钠、氢氧化钾、和其混合物。在一些实施方案,所使用的碱是氢氧化铵。

[0113] 所述培养基还可包括生物学上可接受的钙源,包括但不限于氯化钙。通常,所述培养基中所述钙源(例如氯化钙二水合物)的浓度在约5mg/L至约2000mg/L的范围内,优选在约20mg/L至约1000mg/L的范围内,更优选在约50mg/L至约500mg/L的范围内。

[0114] 所述培养基还可包括氯化钠。通常,所述培养基中氯化钠的浓度在约0.1g/L至约5g/L的范围内,优选在约1g/L至约4g/L的范围内,更优选在约2g/L至约4g/L的范围内。

[0115] 在一些实施方案,所述培养基还可包含痕量金属。此类痕量金属可作为储备溶液添加至所述培养基中,为方便起见,可与培养基的其余部分分开制备。通常,添加至所述培养基中的此痕量金属溶液的量大于约1mL/L,优选大于约5mL/L,更优选大于约10mL/L。然而,超过一定浓度,向所述培养基中添加痕量金属对于微生物的生长是不利的。因此,添加至所述培养基中的此痕量金属溶液的量通常小于约100mL/L,优选小于约50mL/L,更优选小于约30mL/L。应注意的是,除了在储备溶液中添加痕量金属之外,各个组分可单独进行添加,各自在与上述痕量金属溶液范围所规定的组分的量相对应的范围内。

[0116] 在一些实施方案,所述培养基可包括其他维生素类,例如泛酸、生物素、钙、泛酸盐、肌醇、吡哆醇-HCl和硫胺素-HCl。此类维生素可作为储备溶液添加至所述培养基中,为方便起见,可与培养基的其余部分分开制备。然而,超过一定浓度,向所述培养基中添加维生素类不利于微生物的生长。

[0117] 本发明所述的发酵方法可以常规培养模式进行,所述培养模式包括但不限于分批、补料分批、细胞再循环、连续和半连续。在一些实施方案,所述发酵以补料分批模式进行。在此类案例中,所述培养基中的一些组分在培养期间在发酵的生成阶段过程中被耗尽。在一些实施方案,所述培养物可在开始时(例如,生成阶段)补充相对高浓度的此类组分,使得在需要添加之前支持生长和/或香兰素或葡萄糖香兰素生成一段时间。所述这些组分的优

选范围在整个培养过程中通过添加来维持,所述添加以培养物耗尽的水平进行添加。可通过例如定期对培养基取样并测定浓度来监测所述培养基中组分的水平。或者,一旦开发出标准培养程序,所述添加可在整个培养期间的特定时间对应于已知水平以一定时间间隔进行。如本领域技术人员将认识到的,随着所述培养基的细胞密度增加,培养期间营养物的消耗速率亦将增加。此外,为了避免将外来微生物引入培养基中,可使用本领域已知的无菌添加方法进行添加。此外,在培养期间可加入少量消泡剂。

[0118] 所述培养基的温度可以是适于经遗传修饰的细胞生长和/或香兰素或葡萄糖香兰素生成的任何温度。譬如,在用接种物接种培养基之前,所述培养基可置于并保持在约20°C至约45°C的温度范围,优选保持在约25°C至约40°C的温度范围。在某些实施方案,所述细胞是真核的,例如酵母,温度是在约28°C至约34°C的范围。在某些实施方案,所述细胞是原核的,例如细菌,温度是在约35°C至约40°C的范围,例如37°C。

[0119] 可通过向所述培养基中添加酸或碱来控制所述培养基的pH值。在此类案例中,当氨用于控制pH时,其也方便地用作所述培养基中的氮源。优选地,所述pH值保持在约3.0至约8.0,更优选保持在约3.5至约7.0。在某些实施方案,所述细胞是真核的,例如酵母,pH值优选为约4.0至约6.5。在某些实施方案,所述细胞是原核的,例如细菌,pH值为约6.5至约7.5,例如约7.0。

[0120] 在一些实施方案,在培养期间监测所述培养基的碳源浓度,例如葡萄糖、果糖、或蔗糖浓度。可使用已知技术监测所述培养基的葡萄糖浓度,例如,采用葡萄糖氧化酶试验或高压液相色谱,其可用于监测上清液(例如,所述培养基的无细胞组分)中的葡萄糖浓度。所述碳源浓度通常保持低于发生细胞生长抑制的水平。虽然此浓度可能因生物体而异,但对于葡萄糖作为碳源,细胞生长抑制发生在葡萄糖浓度大于约60g/L时,并可通过试验容易地确定。因此,当葡萄糖、果糖、或蔗糖用作碳源时,优选将葡萄糖、果糖、或蔗糖加入发酵罐中并保持在检测限以下。或者,所述培养基中的葡萄糖浓度维持在约1g/L至约100g/L的范围内,更优选地维持在约2g/L至约50g/L的范围内,更优选地维持在约5g/L至约20g/L的范围内。尽管通过添加例如碳源溶液,可将所述碳源浓度维持在所需水平,但通过添加初始培养基的等分试样来维持所述培养基的所述碳源浓度是可接受的,并且可能是优选的。使用初始培养基的等分试样是可取的,因为可同时维持所述培养基中其他营养物(例如,氮源和磷源)的浓度。同样,通过添加痕量金属溶液的等份试样,亦可在所述培养基中维持所述痕量金属浓度。

[0121] 其他合适的发酵培养基和方法记载在例如WO 2016/196321中。

[0122] 发酵组合物

[0123] 另一方面,本发明提供了发酵组合物,其包含本发明所述的经遗传修饰的宿主细胞和由所述经遗传修饰的宿主细胞生成的香兰素和/或葡萄糖香兰素。发酵组合物还可包含培养基。在某些实施方案,发酵组合物包含经遗传修饰的宿主细胞,并进一步包含香兰素或葡萄糖香兰素。在某些实施方案,本发明提供的发酵组合物包含香兰素,作为由所述经遗传修饰的宿主细胞生成的香兰素和/或葡萄糖香兰素的主要成分。在某些实施方案,本发明提供的发酵组合物包含葡萄糖香兰素,作为由所述经遗传修饰的宿主细胞生成的香兰素和/或葡萄糖香兰素的主要成分。

[0124] 香兰素和/或葡萄糖香兰素的回收

[0125] 一旦香兰素或葡萄糖香兰素由所述宿主细胞生成,便可使用本领域已知的任何合适的分离和纯化方法将其回收或分离用于后续应用。在一些实施方案,通过离心或过滤将包含香兰素或葡萄糖香兰素的澄清水相从发酵物分离得到。在其他实施方案,将絮凝剂和凝结剂添加至澄清水相,例如,添加至澄清水相中。

[0126] 在所述这些细胞中生成的香兰素或葡萄糖香兰素可存在于培养物上清液中和/或与所述宿主细胞结合。在其中一些香兰素或葡萄糖香兰素与宿主细胞结合的实施方案中,香兰素或葡萄糖香兰素的回收可包括改善香兰素和/或葡萄糖香兰素从细胞中释放的方法。在一些实施方案,这可采取用热水或缓冲液处理,含或不含表面活性剂,以及含或不含添加的缓冲液或盐类的形式来清洗所述细胞。在一些实施方案,所述温度是认为适合释放香兰素和/或葡萄糖香兰素的任何温度。在一些实施方案,所述温度是从40°C至95°C;或从60°C至90°C;或从75°C至85°C的范围。在一些实施方案,所述温度是40°C、45°C、50°C、55°C、65°C、70°C、75°C、80°C、85°C、90°C、或95°C。在一些实施方案,物理或化学细胞破碎可用于增强香兰素和/或葡萄糖香兰素从宿主细胞释放。或者和/或随后,可使用分离单元操作回收所述培养基中的香兰素或葡萄糖香兰素,所述分离单元操作包括但不限于溶剂萃取法、膜澄清法、膜浓缩法、吸附法、色谱法、蒸发法、化学衍生化法、结晶法和干燥法。

[0127] 制备经遗传修饰的细胞的方法

[0128] 本发明还提供了用于生成宿主细胞的方法,所述宿主细胞经遗传工程改造以包含一种或多种上述修饰,例如,一种或多种异源核酸和/或生物合成途径酶(例如,用于香兰素或葡萄糖香兰素化合物的一种或多种异源核酸和/或生物合成途径酶)。异源酶在宿主细胞中的表达可通过在所述宿主细胞中引入包含编码所述酶的核苷酸序列的核酸来实现,所述酶受允许在所述宿主细胞中表达的调节元件的控制。在一些实施方案,所述核酸是染色体外质粒。在其他实施方案,所述核酸是染色体整合载体,其可将所述核苷酸序列整合到所述宿主细胞的染色体中。在其他实施方案,所述核酸是线性双链DNA片段,其可通过同源性将核苷酸序列整合到宿主细胞的染色体中。

[0129] 可通过本领域技术人员已知的任何方法,且不限于这些方法,将编码这些蛋白质的核酸引入所述宿主细胞中(参见,例如Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1292-3; Cregg et al. (1985) Mol. Cell. Biol. 5:3376-3385; Goeddel et al. eds, 1990, Methods in Enzymology, vol. 185, Academic Press, Inc., CA; Krieger, 1990, Gene Transfer and Expression--A Laboratory Manual, Stockton Press, NY; Sambrook et al., 1989, Molecular Cloning--A Laboratory Manual, Cold Spring Harbor Laboratory, NY; 和 Ausubel et al., eds., Current Edition, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, NY)。示例性技术包括但不限于原生质球法、电穿孔法、PEG1000介导的转化、和乙酸锂或氯化锂介导的转化。

[0130] 可通过修饰编码所述酶的基因的转录来改变宿主细胞中酶的量。其可通过例如以下方式来实现:通过修饰编码所述酶的所述核苷酸序列的拷贝数(例如,通过使用包含所述核苷酸序列的更高或更低拷贝数的表达载体,或通过所述核苷酸序列另外的拷贝引入所述宿主细胞的基因组中,或通过删除或破坏所述宿主细胞基因组中的所述核苷酸序列),通过改变操纵子的多顺反子mRNA上的编码序列的顺序或将操纵子分解成各自具有其自身控

制元件的单个基因,或通过增加核苷酸序列可操作连接的启动子或操纵子的强度。或者,此外,可通过修饰编码所述酶的mRNA的翻译水平来改变宿主细胞中酶的拷贝数。其可通过譬如以下方式来实现:通过修饰mRNA的稳定性,修饰核糖体结合位点的序列,修饰核糖体结合位点与酶编码序列的起始密码子之间的距离或序列,修饰位于酶编码区起始密码子“上游”或邻接5'侧的整个顺反子间区,使用发夹和特化序列稳定mRNA转录物的3'端,修饰酶的密码子用法,改变用于酶的生物合成的稀有密码子tRNA的表达,和/或通过例如突变其编码序列来提高酶的稳定性。

[0131] 宿主细胞中酶的活性可以多种方式改变,包括但不限于,表达所述酶的经修饰形式(其在宿主细胞中表现出增加或降低的溶解度);表达所述酶的经改变形式(其缺乏通过其抑制所述酶的活性的结构域);表达所述酶的经修饰形式(其具有较高或较低K_{cat}或较低或较高K_m的底物);或表达所述酶的经改变形式(其或多或少受到所述途径中另一分子的反馈或前馈调节的影响)。

[0132] 在一些实施方案,用于遗传修饰宿主细胞的核酸包含一种或多种选择标记,所述选择标记可用于选择转化的宿主细胞和对所述宿主细胞施加选择性压力以维持外源DNA。

[0133] 在一些实施方案,所述选择标记是抗生素抗性标记。抗生素抗性标记的示例性实例包括但不限于,BLA、NAT1、PAT、AUR1-C、PDR4、SMR1、CAT、小鼠dhfr、HPH、DSDA、KAN^R、和SH BLE基因产物。来自大肠杆菌(*E. coli*)的BLA基因产物赋予对β-内酰胺抗生素(例如,窄谱头孢菌素类、头孢霉素类、和碳青霉烯类(厄他培南)、头孢孟多和头孢哌酮)的抗性,以及赋予对除了替莫西林(*temocillin*)之外的所有抗革兰氏阴性细菌青霉素类的抗性;来自诺尔斯链霉菌(*S. noursei*)的NAT1基因产物赋予对诺尔丝菌素的抗性;来自产绿色链霉菌(*S. viridochromogenes*) Tu94的PAT基因产物赋予对双丙氨膦(*bialophos*)的抗性;来自酿酒酵母的AUR1-C基因产物赋予对Auerobasidin A (AbA)的抗性;PDR4基因产物赋予对浅蓝菌素的抗性;SMR1基因产物赋予对甲噻磺隆的抗性;来自Tn9转座子的CAT基因产物赋予对氯霉素的抗性;小鼠dhfr基因产物赋予对甲氨蝶呤的抗性;肺炎克雷伯氏菌(*Klebsiella pneumonia*)的HPH基因产物赋予对潮霉素B的抗性;大肠杆菌的DSDA基因产物使细胞在D-丝氨酸作为唯一氮源的平板上生长;Tn903转座子的KAN^R基因赋予对G418的抗性;和来自印度斯坦异壁链霉菌(*Streptoalloteichus hindustanus*)的SH BLE基因产物赋予对Zeocin(博来霉素)的抗性。在一些实施方案,在分离本发明公开的经遗传修饰的宿主细胞后,删除所述抗生素抗性标记。

[0134] 在一些实施方案,所述选择标记拯救所述经遗传修饰的微生物中的营养缺陷型(例如,营养性营养缺陷型)。在此类实施方案中,亲本微生物包含一种或多种基因产物中的功能性破坏,所述一种或多种基因产物在氨基酸或核苷酸生物合成途径中起作用,并且当无功能性时使得亲本细胞无法在不补充一种或多种营养物的培养基中生长。此类基因产物包括但不限于酵母中的HIS3、LEU2、LYS1、LYS2、MET15、TRP1、ADE2、和URA3基因产物。然后可通过用编码所述经破坏的基因产物的功能性拷贝的表达载体或染色体整合构建体来转化亲本细胞,从而拯救营养缺陷型表型,并可基于所述亲本细胞的所述营养缺陷型表型的缺失来选择生成的经遗传修饰的宿主细胞。利用URA3、TRP1和LYS2基因作为选择标记具有显著的优势,因为正选择和负选择均是可能的。通过URA3、TRP1和LYS2突变的营养缺陷型互补进行正选择,而负选择则是基于特异性抑制剂,即5-氟-乳清酸(FOA)、5-氟邻氨基苯甲酸、

和氨基己二酸(aAA),其分别阻止原养型菌株生长但分别使URA3、TRP1和LYS2突变体生长。在其他实施方案,所述选择标记拯救可通过已知选择方法鉴定的其他非致死性缺陷或表型。

[0135] 本发明描述了可用于本发明公开的方法、组合物和生物体的特定基因和蛋白质;然而,人们将认识到这些基因的绝对同一性是不必要的。譬如,可对包含编码多肽或酶的序列的特定基因或多核苷酸进行变化并筛选活性。通常,此类变化包括保守突变和沉默突变。可使用本领域已知的方法筛选此类经修饰或经突变的多核苷酸和多肽以表达功能性酶。

[0136] 由于遗传密码的固有简并性,编码基本上相同或功能等同的多肽的其他多核苷酸亦可用于克隆和表达编码此类酶的多核苷酸。

[0137] 如本领域技术人员将理解的,修饰编码序列以增强其在特定宿主中的表达可能是有利的。所述遗传密码是冗余的,具有64个可能的密码子,但大多数生物体通常使用这些密码子的子集。在物种中最常使用的密码子称为最佳密码子,而那些未经常使用的密码子被分类为稀有密码子或低使用率密码子。在有时被称为“密码子优化”或“控制物种密码子偏性”的过程中,密码子可被置换以反映所述宿主的优选密码子用法。可使用密码子用法表容易地确定其他宿主细胞的密码子优化,或者可使用商业上可获得的软件,例如来自Integrated DNA Technologies的CodonOp(www.idtdna.com/CodonOptfrom)进行密码子优化。

[0138] 可制备含有由特定原核或真核宿主优选的密码子的优化编码序列(Murray et al.,1989,Nucl Acids Res.17:477-508),例如,与由非优化序列生成的转录物相比,以提高翻译速率或以生成具有期望性质(例如更长的半衰期)的重组RNA转录物。还可修饰翻译终止密码子以反映宿主偏好。譬如,酿酒酵母和哺乳动物的典型终止密码子分别是UAA和UGA。单子叶植物类的典型终止密码子是UGA,而昆虫和大肠杆菌通常使用UAA作为终止密码子(Dalphin et al.,1996,Nucl Acids Res.24:216-8)。

[0139] 本领域技术人员将认识到,由于遗传密码的简并性质,可使用与其核苷酸序列不同的多种DNA分子来编码本发明给定的酶。引用编码上述生物合成酶的原生DNA序列在本发明中仅用于说明本发明的实施方案,并且本发明包括任何序列的DNA分子,所述序列编码本发明方法中所用酶的多肽和蛋白质的氨基酸序列。以类似的方式,多肽通常可在其氨基酸序列中耐受一个或多个氨基酸置换、缺失和插入,而不会损失或显著损失所需活性。本发明包括具有与本发明所述特定蛋白质不同的氨基酸序列的此类多肽,只要所述经修饰的多肽或变体多肽具有所述参考多肽的酶促合成代谢活性或分解代谢活性即可。此外,由本发明所示的DNA序列编码的氨基酸序列仅阐明了本发明的实施方案。

[0140] 此外,可用于本发明提供的组合物和方法的酶的同源物均包含在本发明公开内容中。在一些实施方案,当氨基酸序列具有至少约30%、40%、50%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性时,两种蛋白质(或所述蛋白质的区域)基本上是同源的。为了确定两个氨基酸序列或两个核酸序列的同一性百分比,比对所述序列以达到最佳比较目的(例如,为了最佳比对,可在第一和第二氨基酸或核酸序列中的一个或两个中引入空位,并且出于比较目的,可忽略非同源序列)。在一个实施方案,出于比较目的而比对的参考序列的长度为所述参考序列长度的至少30%,通常至少40%,更通常至少50%,甚至更通常至少60%,甚至更通常至少70%、80%、

90%、100%。然后比较相应氨基酸位置或核苷酸位置处的氨基酸残基或核苷酸。当所述第一序列中的位置被与所述第二序列中的相应位置的相同氨基酸残基或核苷酸占据时,那么所述分子在此位置均是相同的(本发明使用的氨基酸或核酸“同一性”等同于氨基酸或核酸“同源性”)。两个序列之间的同一性百分比是所述序列共有的相同位置的数量的函数,考虑到空位的数量和每个空位的长度,需引入这些空位以实现所述两个序列的最佳比对。

[0141] 当“同源/同源性”用于提及蛋白质或肽时,将认识到不相同的残基位置通常因保守氨基酸置换而不同。“保守氨基酸置换”是其中氨基酸残基被具有相似化学性质(例如,电荷或疏水性)的侧链(R基团)的另一氨基酸残基置换的氨基酸置换。通常,保守氨基酸置换不会显著改变蛋白质的功能特性。在通过保守置换使两个或两个以上氨基酸序列彼此不同的情况下,可以向上调节序列同一性百分比或同源性程度以校正所述置换的保守性质。进行此种调节的方法是本领域技术人员所熟知的(参见,例如Pearson W.R.,1994,Methods in Mol Biol 25:365-89)。

[0142] 以下六组各自含有彼此保守置换的氨基酸:1) 丝氨酸(S), 苏氨酸(T);2) 天冬氨酸(D), 谷氨酸(E);3) 天冬酰胺(N), 谷氨酰胺(Q);4) 精氨酸(R), 赖氨酸(K);5) 异亮氨酸(I), 亮氨酸(L), 丙氨酸(A), 缬氨酸(V);和6) 苯丙氨酸(F), 酪氨酸(Y), 色氨酸(W)。

[0143] 通常使用序列分析软件来测定多肽的序列同源性,其亦称为序列同一性百分比。用于将分子序列与含有来自不同生物体的大量序列的数据库进行比较的典型算法是计算机程序BLAST。当搜索含有来自大量不同生物体的序列的数据库时,通常对氨基酸序列进行比较。

[0144] 此外,编码前述酶(或本发明提及的任何其他酶类(或控制或调节其表达的任何调节元件))的任何基因可通过遗传/蛋白质工程技术进行优化,例如本领域普通技术人员已知的定向进化或合理诱变。此种作用使本领域普通技术人员能够优化所述酶在酵母中的表达和活性。

[0145] 此外,编码所述这些酶的基因可从其他真菌和细菌物种中鉴定得到,并且可表达对此途径的调节。多种生物体可作为所述这些酶的来源,包括但不限于,酵母属(*Saccharomyces* spp.),包括酿酒酵母(*S.cerevisiae*)和葡萄汁酵母(*S.uvarum*);克鲁维酵母菌属(*Kluyveromyces* spp.),包括耐热克鲁维酵母(*K.thermotolerans*),乳酸克鲁维酵母(*K.lactis*),和马克斯克鲁维酵母(*K.marxianus*);毕赤酵母属(*Pichia* spp.);汉逊酵母属(*Hansenula* spp.),包括多形汉逊酵母(*H.polymorpha*);假丝酵母属(*Candida* spp.);丝孢酵母属(*Trichosporon* spp.);接合糖酵母属(*Yamadazyma* spp.),包括树干接合糖酵母(*Y.spp.Stipitis*),比勒陀利亚有孢圆酵母(*Torulaspora pretoriensis*),东方伊萨酵母(*Issatchenkia orientalis*);裂殖酵母属(*Schizosaccharomyces* spp.),包括粟酒裂殖酵母(*S.pombe*);隐球酵母属(*Cryptococcus* spp.);曲霉属(*Aspergillus* spp.);脉孢菌属(*Neurospora* spp.);或黑粉菌属(*Ustilago* spp.)。来自厌氧真菌的基因来源包括但不限于,梨囊鞭菌属(*Piromyces* spp.),根囊鞭菌属(*Orpinomyces* spp.),或新美鞭菌属(*Neocallimastix* spp.)。可用的原核酶的来源包括但不限于,大肠杆菌(*Escherichia coli*),运动发酵单胞菌(*Zymomonas mobilis*),金黄色葡萄球菌(*Staphylococcus aureus*),芽孢杆菌属(*Bacillus* spp.),梭菌属(*Clostridium* spp.),棒状杆菌属(*Corynebacterium* spp.),假单胞菌属(*Pseudomonas* spp.),乳球菌属(*Lactococcus*

spp.), 肠杆菌属 (*Enterobacter* spp.) 和沙门氏菌属 (*Salmonella* spp.)。

[0146] 本领域技术人员已知的技术可适于鉴定其他同源基因和同源酶。通常, 类似基因和/或类似酶可通过功能分析进行鉴定, 并具有功能相似性。本领域技术人员已知的技术可适用于鉴定类似基因和类似酶。譬如, 为了鉴定同源或类似的UDP糖基转移酶, 或任何生物合成途径基因、蛋白质、或酶, 技术可包括但不限于使用基于目的基因/酶的公开序列的引物进行PCR, 或通过使用设计用于扩增目的基因中的保守区域的简并引物进行简并PCR, 来克隆基因。此外, 本领域技术人员可使用技术来鉴定具有功能同源性或相似性的同源或类似的基因、蛋白质、或酶。技术包括通过针对所述活性的体外酶测定法来检测细胞或细胞培养物中酶的催化活性 (例如, 如本发明所述或如Kiritani, K., *Branched-Chain Amino Acids Methods Enzymology*, 1970中所述), 然后通过纯化技术来分离具有所述活性的酶, 通过诸如埃德曼 (Edman) 降解等技术确定所述酶的所述蛋白质序列, 设计可能的核酸序列的PCR引物, 通过PCR来扩增所述DNA序列, 以及克隆所述核酸序列。为了鉴定同源或类似基因和/或同源或类似酶、类似基因和/或类似酶或蛋白质, 技术还包括将关于候选基因或酶的数据同诸如BRENDA、KEGG或MetaCYC的数据库进行比较。可根据本发明的教导, 在上述数据库中识别候选基因或酶。

[0147] 实施例

[0148] 实施例1. 酵母转化方法

[0149] 在优化的乙酸锂 (LiAc) 转化中, 使用标准分子生物学技术将各DNA构建体整合到酿酒酵母 (*Saccharomyces cerevisiae*) (CEN.PK2) 中。简言之, 细胞在30°C、振摇 (200rpm) 的酵母提取物蛋白胨右旋糖 (YPD) 培养基中生长过夜, 在100mL YPD中稀释至OD₆₀₀为0.1, 然后生长至OD₆₀₀为0.6-0.8。对于每次转化, 通过离心收获5mL培养物, 在5mL无菌水中清洗, 再次离心, 重悬于1mL 100mM LiAc中, 并转移至微量离心管中。将细胞离心 (13,000x g) 30s, 移除上清液, 并将细胞重悬于由240μL 50% PEG、36μL 1M LiAc、10μL 煮鲑鱼精DNA和74μL 供体DNA组成的转化混合物中。在42°C下热激40分钟后, 将细胞在YPD培养基中回收过夜, 然后铺板至选择性培养基。DNA整合通过菌落PCR确证, 其中使用了对所述整合具有特异性的引物。

[0150] 实施例2: 通过消除原生酵母酶Hfd1将香兰素转化为香草酸来降低发酵液中香草酸的浓度

[0151] Hfd1作为香兰素脱氢酶的鉴定

[0152] 众所周知, 酿酒酵母 (*Saccharomyces cerevisiae*) 具有内源性酶活性, 其能够将香兰素转化为毒性较低的香草醇。在酿酒酵母中还尚未报道能够将香兰素转化为香草酸的酶。为了鉴定负责此种活性的原生酵母酶, 我们采用酵母菌株Y33653, 其中几个负责将香兰素还原为香草醇的基因已被消除, 随后敲除其中每一个被鉴定为醛脱氢酶的酵母基因。所述这些酶的活性降低, 例如, 如图1所示。这些敲除菌株在含有4%蔗糖的液体培养基中培养72小时。然后将培养物离心并弃去上清液。然后将细胞重新悬浮在相同体积的含有0.75g/L香兰素的4%蔗糖液体培养基中。细胞于96孔板中在30度和1000rpm下培养24小时。然后通过UPLC分析上清液以确定培养基中香兰素、香草醇和香草酸的浓度。在这些测试的菌株中, Y33653 Δ hfd1未生成可检测的香草酸。参见图1。此结果清楚地表明, Hfd1是酿酒酵母 (*Saccharomyces cerevisiae*) 中负责将香兰素转化为香草酸的唯一的酶。

[0153] Y33653中醛脱氢酶的敲除

[0154] 通过将潮霉素抗性标记整合到靶基因的开放阅读框中,由亲本菌株Y33651构建醛脱氢酶敲除菌株。使用R21作为模板 (SEQ ID NO:2),使用含有与敲除靶基因的紧邻侧翼区域具有40bp同源性的引物,通过HygA的PCR扩增来生成转化盒。引物见表1。

[0155]	靶基因	正向引物	序列	反向引物	序列
	HFD1	LR726	CAGTACAGGCACACCAGAAC	LR732	GCAAGAGCTGCTGCCAATG
	ALD2	LR727	GAAGCGTTCCAGCTCGAAAG	LR733	GGAACGATTTGAGCAACAAC
	ALD3	LR728	CTTTCTTCCACGCCCTTTAG	LR734	CCCACAACGGAACCATAACC
	ALD4	LR729	AGCACGAGCTTGCAATTACCG	LR735	TGCCCACAAACACCCAAAGG
	ALD5	LR730	GGACTAGCAAGGGCTTAATG	LR736	GCAAACGCCTAATGGTTCC
	ALD6	LR731	GGCGTATCCAAGCCGAAAC	LR737	TTGACCACAGACACCGATTG

[0156] 表1.实施例5中使用的引物

[0157] 培养基和培养条件

[0158] 将每个菌株从单个菌落接种到含有360μL BSM 2.0 (8g/L KH_2PO_4 、15g/L $(\text{NH}_4)_2\text{SO}_4$ 、6.15g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.0575g/L $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.0032g/L CuSO_4 、0.0032g/L $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 、0.0047g/L $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ 、0.0048g/L $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 、0.028g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.029g/L $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 、0.117g/L EDTA、0.0006g/L生物素、0.0024g/L对氨基苯甲酸、0.012g/L烟酸、0.03g/L肌醇(myoinositol)、0.012g/L吡哆醇(pyridoxine)HCl、0.012g/L硫胺素HCl、0.012g/L泛酸钙、6g/L琥珀酸) 4%蔗糖的96孔板中。将平板在30度和1000rpm下培养3天。然后将所述这些平板离心并重新悬浮在含有1g/L香兰素的BSM 2.0 4%蔗糖中,并于30度下培养24小时。

[0159] UPLC检测条件

[0160] 为了量化所生成的香兰素的量,在配备有二极管阵列检测器的Agilent Vanquish™ Flex二元UHPLC系统上使用以下程序,对样品进行分析:

[0161] 流动相(A):1.4%硫酸v/v水溶液

[0162] 流动相(B):100%乙腈

[0163] 梯度如下[梯度时间(min),流动相A(%)]:[(0.00,88),(0.05,88),(1.25,85),(2.25,83),(3.0,82),(3.5,88),(4.0,88)]。流速为1mL/min。

[0164] 通过缺失生成香兰素的酿酒酵母(*Saccharomyces cerevisiae*)菌株中的HFD1来提高香兰素的生成量

[0165] 为了证实HFD1的缺失将会通过减少原生Hfd1酶将香兰素转化回香草酸而改善制备发酵衍生的香兰素的方法,我们测试了具有(Y42688)和不具有(Y43188)原生HFD1基因的香兰素生成菌株在发酵过程中的作用,如下所述。每24小时取样一次并测定香兰素、香草醇的浓度。HFD1的缺失使得0-3天产量增加7%,0-3天生产率增加12%(图2)。

[0166] 生成Y43188的菌株构建

[0167] Y43188

[0168] Y43188是通过一次整合由Y42688产生。MS135802(SEQ ID NO:1)包含与HFD1开放阅读框上游和下游区域的同源性,导致所得菌株中所述基因的缺失。如下图表所示。

[0169] US_HFD1 DS_HFD1

[0170] Y42688/Y43188-的发酵培养基和条件

[0171] 使用在琼脂板上生长的酵母菌落接种含有60mL BSM 2.0的500mL带挡板种子烧瓶,所述BSM 2.0含有4%蔗糖、2%麦芽糖、5g/L赖氨酸,并于28°C、200RPM的振荡器中生长21小时。然后将60mL种子烧瓶培养物接种到含有240mL上述MF培养基的0.5-L制备发酵罐(MFA)中。发酵罐的营养供给原料是100g/L纯蔗糖进料。初始脉冲为2g TRS/L,速率为5g/L/h。然后使用基于培养物对碳的需求的算法来调整发酵罐进料速率,如溶解氧的增加所示。在30°C的恒定温度和5.0的恒定pH值(通过添加氢氧化铵控制)有氧条件下进行发酵,直至溶解氧达到0%。然后控制搅拌以在发酵的剩余时间保持15mmol O₂/L/h的氧利用率。需要时去除培养物以防止溢出。每天还添加盐类、痕量金属类和维生素类。在开始时将0.1mL L-61消泡剂添加至发酵培养基,随后视需要添加。每天监测所生成的香兰素的量和细胞消耗的总糖量,并且每24小时测定此两个值的比率(即糖的产物产率)。发酵罐运行5天。

[0172] 香兰素Y42688/Y43188-的定量

[0173] 为了量化所生成的香兰素的量,在配备有二极管阵列检测器的Agilent Vanquish™ Flex二元UHPLC系统上使用以下程序,对样品进行分析:

[0174] 流动相(A):1.4%硫酸v/v水溶液

[0175] 流动相(B):100%乙腈

[0176] 梯度如下[梯度时间(min),流动相A(%)]:[(0.00,88),(0.05,88),(1.25,85),(2.25,83),(3.0,82),(3.5,88),(4.0,88)]。流速为1mL/min。

[0177] 实施例3:在工程微生物中由蔗糖或葡萄糖生物合成香兰素

[0178] 经由3-脱氢莽草酸的中间体由葡萄糖生物合成香兰素可通过在大肠杆菌(E.coli)(Kunjapur et al.,J.Am.Chem.Soc.2014,136:11644-11654)或酵母(Hansen et al.,Appl.Environ.Microbiol.200975:2765-2774)中进行途径工程以在发酵液中积累香兰素而进行。以甘蔗来源的蔗糖或玉米来源的葡萄糖为原料生成所述经分离的香兰素样品。

[0179] 用于同位素分析的样品制备和仪器用法根据已公开发表的方法进行(Culp&Noakes,J.Agric.Food Chem.1990,38,1249-1255;Culp&Noakes J.Agric.Food Chem.1992,40,1892-1897)。¹⁴C活性报告为每克碳每分钟的崩解量(dpm/g C),可用于区分石油衍生的香兰素(dpm/g C接近0)与衍生自植物来源的香兰素,后者通常提供15-16dpm/g C。¹³C/¹²C的稳定同位素比表示为:‰=[(R_{样品}/R_{标准品})-1]×10³,其中R=¹³C/¹²C表示相对于碳放射性同位素的Pee Dee Belemnite(PDB)标准品。²H/¹H的稳定同位素比表示为:‰=[(R_{样品}/R_{标准品})-1]×10³,其中R=²H/¹H表示相对于氢放射性同位素的标准平均大洋水(Standard Mean Ocean Water,SMOW)标准品。

[0180] 香兰素是通过从芳香族氨基酸生物合成中虹吸3-脱氢莽草酸,分3步由蔗糖或葡萄糖发酵生成(Li and Frost,J.Am.Chem.Soc.1998 120:10545-10546)。来自甘蔗的蔗糖和来自玉米的葡萄糖均可大量提供,并代表了微生物生长和天然香料成分制造成本最低的碳源。蔗糖通过常见的中间体6-磷酸果糖进入糖酵解(图3)。如表2所示,发酵衍生的香兰素可追溯到提取自玉米的葡萄糖或提取自甘蔗的蔗糖,其¹⁴C数据分别为12.9dpm/g C或14.0dpm/g C至14.1dpm/g C。所述这些测定值均与光合作用过程中CO₂摄取的15-16dpm/g C的大气稳定状态一致,表明香兰素中的碳可追溯到最近收获的植物来源。尽管仅凭这些数

据无法证明香兰素样品是天然的,但其至少排除了所述香兰素衍生自石油的可能性。

[0181] 表2. 各种来源的香兰素的同位素分析

[0182]

香兰素来源	$\delta^{13}\text{C}$, ‰	$\delta^2\text{H}$, ‰	参考	
姜黄素	-30.4 至 -27.8	-155 至 -128	Geißler 等人	
丁香酚	-32.4 至 -30.4	-114 至 -62	Geißler 等人	
提取自稻米的阿魏酸	-37.3 至 -35.4	-174 至 -158	Geißler 等人	
提取自玉米的阿魏酸	-19.9 至 -19.0	-99 至 -97	Geißler 等人	
愈创木酚	-30.7 至 -25.9	-25 至 117	Geißler 等人	
木质素	-28.1 至 -27.4	-182 至 -175	Geißler 等人	
香草豆	-21.8 至 -20.2	-83 至 -59	Geißler 等人	
香兰素来源	^{14}C , dpm/g of C	$\delta^{13}\text{C}$, ‰	$\delta^2\text{H}$, ‰	参考
提取自玉米的葡萄糖	12.9	-14.7	-132	本研究
提取自甘蔗的蔗糖	14.1	-12.8	-150	本研究
提取自甘蔗的蔗糖	14.0	-14.1	-125	本研究
提取自甘蔗的蔗糖	14.0	-14.5	-124	本研究
提取自甘蔗的蔗糖	14.1	-14.4	-124	本研究
提取自甘蔗的蔗糖	14.1	-14.8	-124	本研究

[0183] 表2提供了多种天然来源和人工来源的香兰素的块体 $\delta^{13}\text{C}$ 数据和 $\delta^2\text{H}$ 值。相对于PDB,通过发酵由提取自玉米的葡萄糖或提取自甘蔗的蔗糖衍生的香兰素的碳同位素比为-14.8‰至-12.8‰。鉴于玉米和甘蔗在光合作用过程中遵循Hatch-Slack CO_2 固定(C4),两个样品所测定的与PDB的‰偏差均在预期范围内(Culp&Noakes, J. Agric. Food Chem. 1990, 38, 1249-1255)。通过提取自稻米的阿魏酸的生物转化而得到的香兰素传统上被认为是一种天然风味成分,与PDB的偏差为-37.3‰至-35.4‰。水稻在光合作用过程中遵循卡尔文(Calvin) CO_2 固定(C3),其与PDB的‰偏差在组别中处于最极端范围。(Culp&Noakes, Agric. Food Chem. 1990, 38, 1249-1255)。正如Geißler及其同事所披露的,当阿魏酸衍生自玉米时,偏差会移至-19.9‰至-19.0‰,这符合对C4植物的预期(Geißler et al. Flavour Frag. J. 2017, 32, 228-2)。姜黄素、丁香酚、木质素和愈创木酚与PDB的块体 $\delta^{13}\text{C}$ ‰偏差变得不那么明显,需要使用块体 $\delta^2\text{H}$ 值进行聚类。相对于SMOW,测得由提取自玉米的葡萄糖衍生的香兰素的氢同位素比为-132‰。相对于SMOW,测得由提取自甘蔗的蔗糖衍生的香兰素的氢同位素比为-150‰至-124‰的范围。尽管通过发酵由蔗糖或葡萄糖衍生的香兰素产生了先前报道中独特的碳同位素比,但结合氢同位素比数据,产生了来自其他已知途径所得到香兰素的独特聚类(图4)。虽然姜黄素、丁香酚和木质素均可衍生自天然来源,但愈创木酚主要衍生自石油。姜黄素向香兰素的生物转化符合天然风味成分的标准,而由丁香酚衍生的香兰素则取决于用于生成所述成分的过程(Gallage et al. Mol. Plant 2015, 8, 40-57.)。另一方面,转化木质素的过程和愈创木酚的来源均被认为是非自然的。衍生自其任一原料的香兰素使得大部分香兰素作为人造香草调味剂进行生产和销售。从香荚兰(*Vanilla planifolia*)或塔希提香草兰(*Vanilla tahitensis*)的提取物中分离香兰素会使得块体 δ

^{13}C ‰偏差为-21.8‰至-20.2‰。这可归因于香草兰花 (*Vanilla orchid*) 在光合作用期间表现出卡尔文 (Calvin) 与 Hatch-Slack CO_2 固定的组合, 其是 CAM 植物的独特特征 (Geißler et al. *Flavour Frag. J.* 2017, 32, 228-237)。

[0184] 风味成分香兰素可通过合成、生物合成或生物转化方法从多种原材料衍生得到。除蔗糖或葡萄糖经由 3-脱氢莽草酸的中间体生物合成之外, 多个香兰素样品的块体 $\delta^{13}\text{C}$ 和 $\delta^2\text{H}$ 值先前已公开发表 (Culp & Noakes, *J. Agric. Food Chem.* 1992, 40, 1892-1897; Geißler et al. *Flavour Frag. J.* 2017, 32, 228-237)。本实施例表明, 由从两种 C4 植物提取的蔗糖生物合成的香兰素的块体 $\delta^{13}\text{C}$ 偏差位于 -14.7 至 -12.8 的独特位置, 而与其他香兰素样品的数据相比, 其块体 $\delta^2\text{H}$ 值重叠。

[0185] 本说明书中引用的所有出版物、专利和专利申请均通过引用并入本发明, 如同每个单独的出版物或专利申请被具体和单独地指出通过引用并入。尽管出于清楚理解的目的, 已通过举例说明和示例的方式对上述发明进行了一些详细描述, 但根据本发明教导, 对于本领域普通技术人员而言将显而易见的是, 在不脱离所附权利要求书的精神或范围的情况下, 可对其进行某些改变/变化以及修改/修饰。

[0186] SEQ ID NO:1

[0187] 长度:1081

[0188] 类型:

[0189] 生物体:酿酒酵母 (*Saccharomyces cerevisiae*)

[0190] 其他信息:MS135802 序列

[0191] GACGGCACGGCCACGCGTTTAAACCGCCAAGAGAGAGAAGCTAGATTATCATTACAGCAGCCACATAGT
ATACCAAATTCCAGTACAGGCACACCAGAACATGATCAAGACACTTAGAGGAAATGGAACAACGAATTTCCAGCCAA
AAATTCGAGTAGTTCATGATGAAAGATTTTACATGCATTTTATATATAAATATATACCGTCCTATATGGATTTCA
TGCCAACAGGGTATATAATAGACAATTACCGGTGTACTGATATATCAACTATCGACTCCAAGCCTTTTATCTATCAG
TCAATTTTACATCAAGATCCCACTTTTAGATAGGTTGCGAAAATTCATCTAATATTAGTGATTTAATTAGATGGTGG
ATTGCTTACCCTTTTTTTTGTCTGTTTATAGGAGGAGATTCTTCGGATTTTAGGGATAAACGGATACTCCATATATAAA
AAACAAAACCTTCAGGCATATTGATTATCTAAAAGGAATATTCTAAAACCATAGCCATAGTAATTTATCACCAACACG
CTCGTCCAACGCCGGCGGACCTACGTAATGTTTAAGGTTAATTAATTATTTGATGTATAAGTAACCTTTTCGTTTAAA
AATTTTCATATGGGCGATAATATATCAATATTTATTAATTACAATTTACTCTATTTGCTCGTATAGAGTATATACTCG
CTAAATACATTTTGATTACCAATCATCTTCCTCTTCATCTTCGTTAGCCTCCTCGTCATCACTCTCTAATCTCATTT
TATTGGCCCTAGCAGCAAGCGCATTCGCCAAACCATCACCCAAACTTGCTTGGGCTGCAGCTGGAATAACGTTTGCA
GAGGTGTTATCATGTTGTACAGAATTCAGAACGGGTTTAGGAACATCCCTGTTTTGAGGAGGAGTAGGGATATTATT
ATTTCCGGAATGTGGTTTCATGTAAGCTGTTGGCACCCAACCTTCCTTAGAGCCATCCAGAAGTTTCCCGAGAGACC
AACCGCTTGTTCTTCTCTTGTAATGTAAATTACGTCTCCTTTCTTTAAAGGTAACCTCAGCGGTGTTTAAACCCAG
CGCCTGGCGGG

[0192] SEQ ID NO:2

[0193] 长度:1912

[0194] 类型:

[0195] 生物体:人工序列

[0196] 其他信息:R21 序列

[0197] TCGACACTAGTAATACACATCATCGTCCTACAAGTTCATCAAAGTGTTGGACAGACAACCTATACCAGCA
TGGATCTCTTGTATCGGTTCTTTTCTCCCGCTCTCTCGCAATAACAATGAACACTGGGTCAATCATAGCCTACACAG
GTGAACAGAGTAGCGTTTATACAGGGTTTATACGGTGATTCTACGGCAAAAATTTTTCATTTCTAAAAAAG
AAAAATTTTCTTTCCAACGCTAGAAGGAAAAAGAAAAATCTAATTAAATTGATTTGGTGATTTTCTGAGAGTTCCCT
TTTTCATATATCGAATTTTGAATATAAAAGGAGATCGAAAAATTTTCTATTCAATCTGTTTTCTGGTTTTATTG
ATAGTTTTTTTGTGTATTATTATTATGGATTAGTACTGGTTTATATGGGTTTTTCTGTATAACTTCTTTTTATTTTA
GTTTGTTTAATCTTATTTTGTAGTTACATTATAGTTCCTAACTGCAAGAGAAGTAACATTAATAATGAAAAAGCCTG
AACTCACC GCGACGTCTGTGCGAGAAGTTTCTGATCGAAAAGTTGACAGCGTCTCCGACCTGATGCAGCTCTCGGAG
GGCGAAGAATCTCGTGCTTTTACGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGG
TTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGG
AATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCGGTGCACAGGGTGTACGTTGCAAGACCTGCCTGAAACCGAA
CTGCCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTT
CGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTTCATATGCGCGATTGCTGATCCCCATG
TGTATCACTGGCAAATGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGG
GCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCG
CATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGA
GGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGAGCTTGCAGGATCGCCGCGG
CTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTTCGATGATGCAGC
TTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAA
GCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGAAACCGACGCCCCAGCACTCGTCCGAGG
GCAAAGGAATAGGTTTAACTTGATACTACTAGATTTTTTCTCTTCATTTATAAAATTTTGGTTATAATTGAAGCTT
TAGAAGTATGAAAAATCCTTTTTTTTCATTCTTTGCAACCAAAATAAGAAGCTTCTTTTATTGAAATGATGA
ATATAAACCTAACAAAAGAAAAAGACTCGAATATCAAACATTAATAAATAAAGAGGTTATCTGTTTTCCCAT
TAGTTGGAGTTTGCATTTTCTAATAGATAGAACTCTCAATTAATGTGGATTTAGTTTCTCTGTTCTGTTTTTTTGT
TTTGTCTCACTGTATTTACATTTCTATTTAGTATTTAGTTATTCATATAATCTTAACCTTCTCGAGGAGCTC

序列表

<110>	阿迈瑞斯公司 (AMYRIS, INC.)	
<120>	用于高效生成香兰素的经修饰宿主细胞	
<130>	107345.00796	
<140>	PCT/US2020/044613	
<141>	2020-07-31	
<160>	14	
<170>	PatentIn 版本3.5	
<210>	1	
<211>	1081	
<212>	DNA	
<213>	酿酒酵母 (<i>Saccharomyces cerevisiae</i>)	
<220>		
<221>	尚未归类特征区 (misc_feature)	
<223>	MS135802序列	
<400>	1	
	gacggcacgg ccacgcgttt aaaccgcaa gagagagaag ctagattatc attacagcag	60
[0001]	ccacatagta taccaaattc cagtacaggc acaccagaac atgatcaaga cacttagagg	120
	aaatggaaca acgaatttcc agccaaaaat tccgagtagt tcatgatgaa agatttttac	180
	atgcatttta tatataaata tataccgtcc tatatggatt tcatgccaac agggtatata	240
	atagacaatt accggtgtac tgatatatca actatcgact ccaagccttt tatctatcag	300
	tcaattttac atcaagatcc cacttttaga taggttcgaa aattcaatct aatattagtg	360
	atttaattag atggttgatt gcttaccctt tttttgtcg ttttaggagg agattcttcg	420
	gatttttaggg ataaacggat actccatata taaaaaaca aacttcaggc atattgatta	480
	tctaaaagga atattctaaa accatagcca tagtaattta tcaccaacac gtcgtccaa	540
	cgccggcgga cctacgtaat gttaaagggt aattaattat ttgatgtata agtaaccttt	600
	cgttttaaaa tttcatatgg gcgataatat atcaatattt attaattaca atttactcta	660
	tttgctcgta tagagtatat actcgctaaa tacattttga ttaccaatca tcttcctctt	720
	catcttcggt agcctcctcg tcatcactct ctaatctcat tttattggcc ctagcagcaa	780
	gcgcattcgc caaacatca cccaaacttg cttgggctgc agctggaata acgtttgcag	840
	agggtgttatc atgttgtaca gaattcagaa cgggtttagg aacatccctg ttttgaggag	900
	gagtagggat attattattt ccggaatgtg gtttcatgta agctgttgge acccaacctt	960

	ccttagagcc atccagaagt ttcccgagag accaaccgct tggttcttct cttgtaatgt	1020
	aaattacgtc tcctttcttt aaaggtaact cagcggtggt taaaccccag cgcctggcgg	1080
	g	1081
	<210> 2	
	<211> 1912	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的: R21序列	
	<400> 2	
	tcgacactag taatacacat catcgtccta caagttcatc aaagtgttgg acagacaact	60
	ataccagcat ggatctcttg tatcggttct tttctccgc tctctcgca taacaatgaa	120
	cactgggtca atcatagcct acacaggtga acagagtagc gtttatacag gttttatacg	180
	gtgattccta cggcaaaaat ttttcatttc taaaaaaaaa aagaaaaatt tttctttcca	240
	acgctagaag gaaaagaaaa atctaattaa attgatttgg tgattttctg agagttccct	300
	ttttcatata tcgaattttg aatataaaag gagatcgaaa aaatttttct attcaatctg	360
[0002]	ttttctggtt ttatttgata gtttttttgt gtattattat tatggattag tactggttta	420
	tatgggtttt tctgtataac ttctttttat tttagtttgt ttaatcttat tttgagttac	480
	attatagttc cctaactgca agagaagtaa cattaaaaat gaaaaagcct gaactcacgc	540
	cgacgtctgt cgagaagttt ctgatcgaaa agttcgacag cgtctccgac ctgatgcagc	600
	tctcggaggg cgaagaatct cgtgctttca gcttcgatgt aggagggcgt ggatatgtcc	660
	tgccggtaaa tagctgcgcc gatggtttct acaaagatcg ttatgtttat cggcactttg	720
	catcgccgc gctcccgatt ccggaagtgc ttgacattgg ggaattcagc gagagcctga	780
	cctattgcat ctcccgcgt gcacagggtg tcacgttgca agacctgcct gaaaccgaac	840
	tgcccgtgt tctgcagccg gtcgcggagg ccatggatgc gatcgtgcg gccgatctta	900
	gccagacgag cgggttcggc ccattcggac cgcaaggaat cggtaatac actacatggc	960
	gtgatttcat atgcgcgatt gctgatcccc atgtgtatca ctggcaaac gtgatggacg	1020
	acaccgtcag tgcgtccgtc gcgcaggctc tcgatgagct gatgcttttg gccgaggact	1080
	gccccgaagt ccggcacctc gtgcacgcgg atttcggctc caacaatgtc ctgacggaca	1140
	atggccgcat aacagcggtc attgaactgga gcgaggcgat gttcggggat tcccaatacg	1200

	aggctcgccaa catctttcttc tggaggccgt ggttggcttg tatggagcag cagacgcgct	1260
	acttcgagcg gaggcattccg gagcttgcag gatcgccgcg gctccgggcg tatatgctcc	1320
	gcattggtct tgaccaactc tatcagagct tggttgacgg caatttcgat gatgcagctt	1380
	gggcgcaggg tcgatgcgac gcaatcgtec gatccggagc cgggactgtc gggcgtagac	1440
	aaatcgcccc cagaagcgcg gccgtctgga ccgatggctg tgtagaagta ctcgccgata	1500
	gtggaaaccg acgccccagc actcgtccga gggcaaagga ataggtttaa cttgatacta	1560
	ctagatTTTT tctcttcatt tataaaattt ttggttataa ttgaagcttt agaagtatga	1620
	aaaaatcctt ttttttcatt ctttgcaacc aaaataagaa gcttctttta ttcattgaaa	1680
	tgatgaatat aaacctaa aaagaaaaag actcgaatat caaacattaa aaaaaataa	1740
	aagaggttat ctgttttccc atttagttgg agtttgcatt ttctaataga tagaactctc	1800
	aattaatgtg gatttagttt ctctgttcgt ttttttttgt tttgttctca ctgtatttac	1860
	atttctatTT agtatttagt tattcatata atcttaactt ctcgaggagc tc	1912
[0003]	<210> 3	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物	
	<400> 3	
	cagtacaggc acaccagaac	20
	<210> 4	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物	
	<400> 4	
	gcaagagctg ctgccaatg	19
	<210> 5	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成引物	

	<400> 5 gaagcgttcc agctcgaaag	20
	<210> 6 <211> 20 <212> DNA <213> 人工序列	
	<220> <223> 合成引物	
	<400> 6 ggaacgattt gagcaacaac	20
	<210> 7 <211> 20 <212> DNA <213> 人工序列	
	<220> <223> 合成引物	
	<400> 7 ctttcttcca cgcccttttag	20
[0004]	<210> 8 <211> 20 <212> DNA <213> 人工序列	
	<220> <223> 合成引物	
	<400> 8 cccacaacgg aaccataacc	20
	<210> 9 <211> 20 <212> DNA <213> 人工序列	
	<220> <223> 合成引物	
	<400> 9 agcacgagct tgcattaccg	20
	<210> 10 <211> 20 <212> DNA <213> 人工序列	

	<220>		
	<223>	合成引物	
	<400>	10	
		tgcccacaaa cacccaaagg	20
	<210>	11	
	<211>	20	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物	
	<400>	11	
		ggactagcaa gggcttaatg	20
	<210>	12	
	<211>	19	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物	
[0005]	<400>	12	
		gcaaacgcct aatggttcc	19
	<210>	13	
	<211>	19	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物	
	<400>	13	
		ggcgtatcca agccgaaac	19
	<210>	14	
	<211>	20	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成引物	
	<400>	14	
		ttgaccacag acaccgattg	20

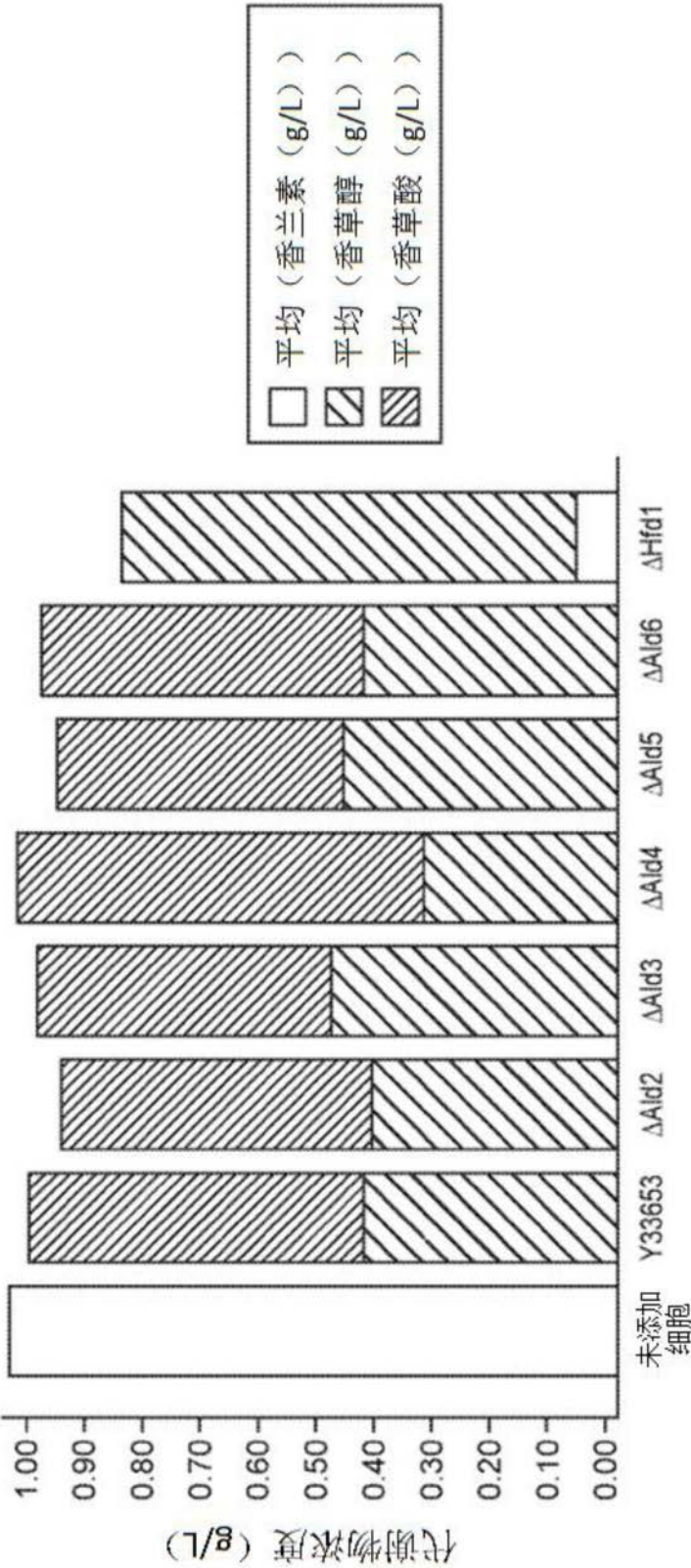


图1

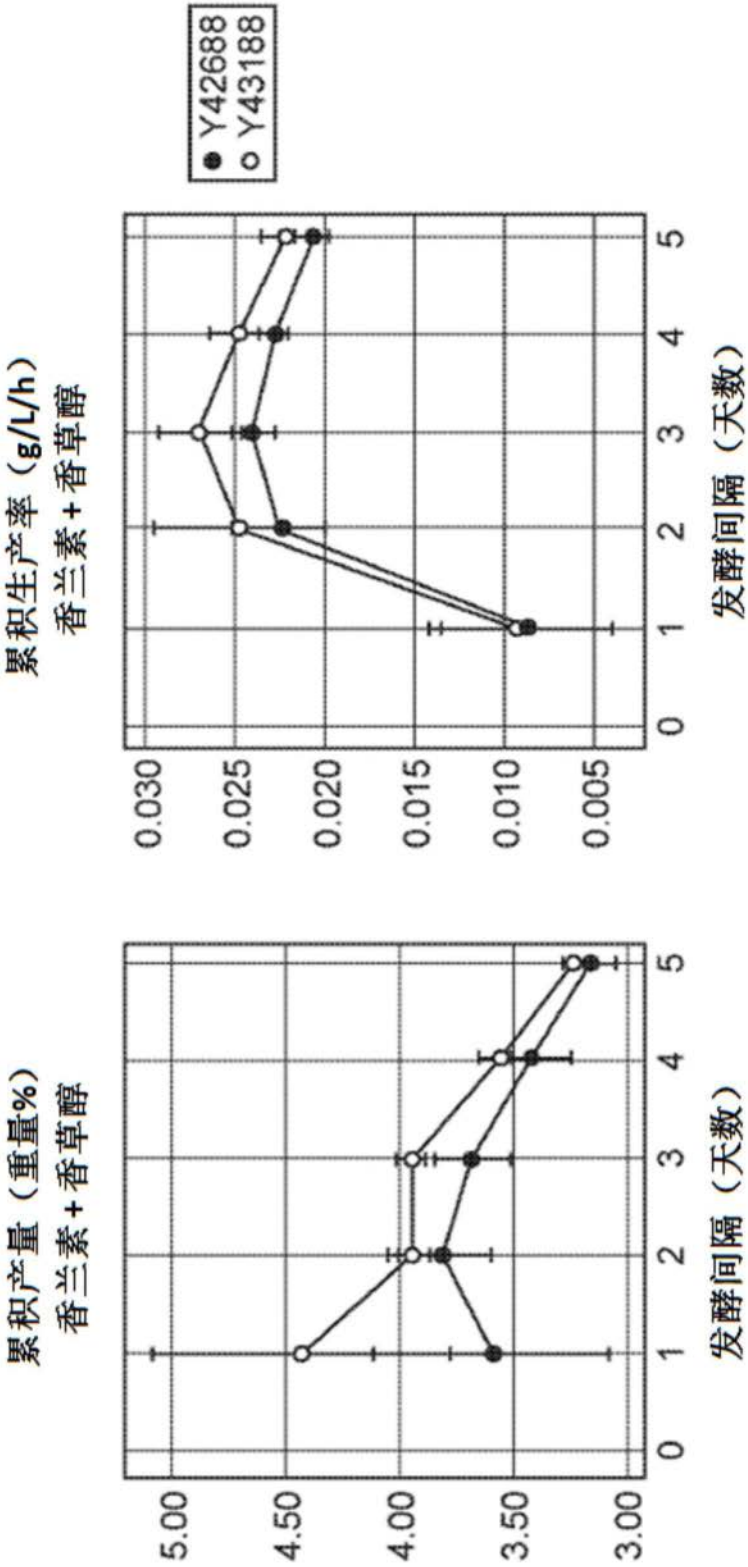


图2

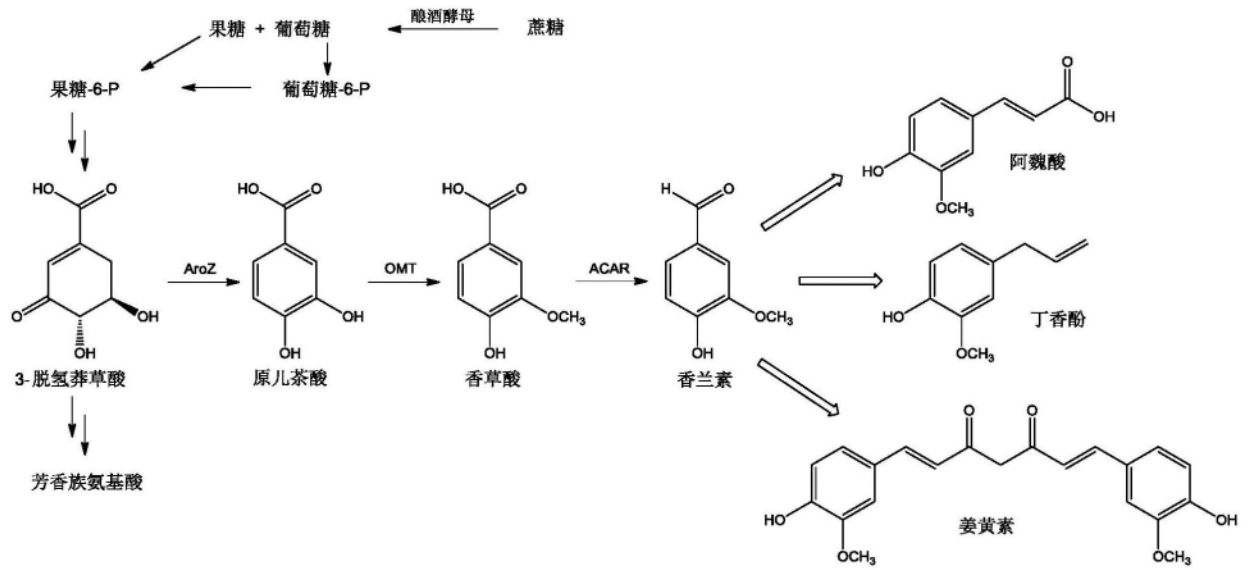


图3

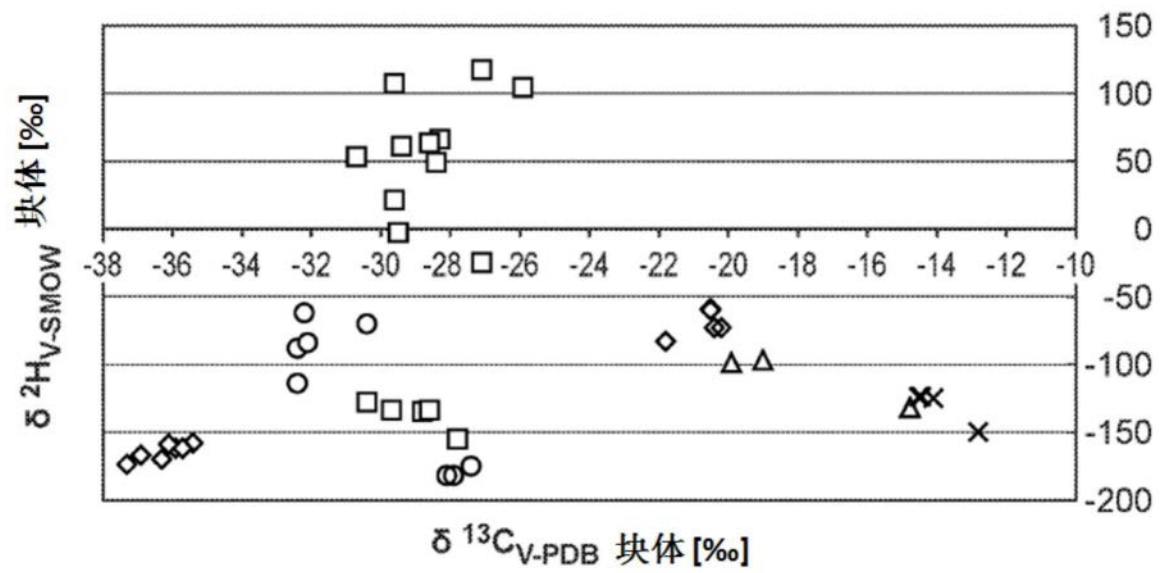


图4