

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁷

[12] 发明专利申请公开说明书

[21] 申请号 98803635.5

C07K 14/405
C12N 9/10 C12N 15/54
C12N 15/31 C12N 1/21
//C12N1/12,(C12N
15/54,
C12R1:89)(C12N9
/10,C12R1:19)
[11]公开号 CN 1251108A

[43]公开日 2000年4月19日

[22]申请日 1998.1.20 [21]申请号 98803635.5

[30]优先权

[32]1997.1.24 [33]JP [31]11430/1997

[86]国际申请 PCT/JP98/00194 1998.1.20

[87]国际公布 WO98/32770 日 1998.7.30

[85]进入国家阶段日期 1999.9.24

[71]申请人 麒麟麦酒株式会社

地址 日本东京都

[72]发明人 S·R·费里 户栗敏博

[74]专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 谭明胜

权利要求书 1 页 说明书 16 页 附图页数 3 页

[54]发明名称 β -酮酰基-ACP合成酶II及其编码基因

[57]摘要

具有能够控制或调节植物细胞饱和和非饱和脂肪酸含量酶活性蛋白质的基因以及该基因表达产物的酶蛋白,更具体地:具有 β -酮酰基-ACP合成酶II(KAS II)活性的蛋白质,并且该蛋白质具有以来源于蓝绿藻(组囊藻, *Anacystis nidulans*)的KAS II氨基酸序列为典型代表的特定氨基酸序列或与其基本上等价的氨基酸序列;编码上述蛋白氨基酸序列的KAS II蛋白基因;含有该基因的重组载体以及转入了该基因的细胞。

MTETGRRW ITLGAITPI GNDPEYWG ILAGRNGIDL IRGFAASRHA CKIAGEVKDF
DPTOMERKD AKFMDFADL AVAASRQAVA DARLDITELN ADAIGLIGS GIGLRVMEI
OOTVLLKGP DRCSPEWPM WIAWMAAGLT AITLGAKGPC NYTYTACAAG SHANGEAFR
IDHGADAMI CGSTESCVTP LAMARFAACK ALSLPNDOPA HACPPFDQGR DGFVMDGEG
ILVLESLEHA DARGAHLYGE IVGYMTCDG YHITSMPGG LGAARATEFG LRDNLOPSS
VSYINAGTS TPANDSTETA AIKCALDEBA YKTVISSTKS MITQLLGGSG GIEVAATLA
IAEDNVPTI NLEDPDPCD LDYVNDARS LPVEVALSNS FGFGRHVTL AFRKPP

ISSN 1008-4274

权 利 要 求 书

1. 一种蛋白质,它具有序列号 2 中所示的氨基酸序列或基本上与序列号 2 中所示氨基酸序列相同的氨基酸序列,且具有 β -酮酰基-ACP 合成酶 II 的活性。
- 5 2. 权利要求 1 的蛋白质,其中基本上相同的氨基酸序列为 1 个或多个氨基酸发生取代,缺失、插入或添加的序列。
 3. 权利要求 1 或 2 的蛋白质,其中氨基酸序列来源于蓝绿藻。
 4. 一种基因,它编码具有序列号 2 中所示氨基酸序列、或具有基本上与序列号 2 所示氨基酸序列相同的氨基酸序列,且具有 β -酮酰基-ACP 合成酶 II 活性的蛋白质。
- 10 5. 权利要求 4 的 β -酮酰基-ACP 合成酶 II 酶活性蛋白质基因,其中基本上相同的氨基酸序列为 1 个或多个氨基酸发生置换、缺失、插入或添加了的序列。
 6. 权利要求 4 或 5 的基因,其中所编码的蛋白质来源于蓝绿藻。
- 15 7. 一种重组载体,它包含权利要求 4 - 6 任一权项中的 β -酮酰基-ACP 合成酶 II 酶活性蛋白质基因。
 8. 一种细胞,它转入了权利要求 4 - 6 任一权项中的 β -酮酰基-ACP 合成酶 II 酶活性蛋白质基因。

说明书

β - 酮酰基 - ACP 合成酶 II 及其编码基因

技术领域

5 本发明涉及与植物中脂肪酸合成相关的合成酶的氨基酸序列和与此相关的 DNA 构造, 即涉及具有特定氨基酸序列的脂肪酸合成酶蛋白质和编码该物质的基因。应用这样的基因和插入了适宜调节序列(调节基因)的嵌合基因转化细胞, 可控制细胞中饱和和不饱和脂肪酸的含量。

10

背景技术

已知脂肪酸合成酶有 2 种类型, 在动物和酵母中的酶是多种酶全体连结成具有一个功能的复合体 - 脂肪酸合成酶复合体 (FAS) (I 型), 与此相对, 在高等植物细胞和原核生物中的酶是各种酶在生物体外独立、散在存在的 (II 型)。用 II 型酶合成脂肪酸的时候, 可溶性蛋白质酰基载体蛋白 (ACP) 是必要的, 合成的脂肪酸以酰基 - ACP 形式存在。所合成的最终产物为棕榈酰基 - ACP。棕榈酰 - ACP 再通过延长链长转变为硬脂酰基 - ACP 后, 经可溶性脂肪酸去饱和酶 (硬脂酰基 - ACP 去饱和酶) 的去饱和作用, 转换成油酰基 - ACP。棕榈酸和油酸合并极性脂质后, 再将后者去饱和 (J. Ohlrogge 和 J. Browse(1995) 脂质的生物合成。植物细胞, 7, p957 - 970)。

催化由棕榈酰基 - ACP 向硬脂酰基 - ACP 进行链延长的酶, 能从棕榈酰基 - ACP、丙二酰基 - ACP 和 NADPH 制备硬脂酰基 - ACP (J. Ohlrogge 和 J. Browse(1995)Lipid Biosynthesis. The Plant Cell 7, p957-970)。这些反应说明了棕榈酰基 - 硫酯和 C2 单位的缩合物经还原、脱水、再还原生成硬脂酰基 - 硫酯的一系列酶反应的全部流程。

25 已对植物中脂质的生物合成进行了很好地研究 (Browse 等, Annu. Rev. Plant Physiol. Mol. Biol.(1991)42:467-506)。通过这些研究得知植物细胞中, 从棕榈酸生成硬脂酸是由 β - 酮酰基 - ACP 合成酶 II (KASII) 进行催化的反应。但是在上述文献中对 KASII 酶及其基因均未进行分离, 它们的序列也是未知的。

另外, 已知 KAS 的同功酶在植物的叶绿体中有 3 种, 除 KASII 以外的 2 种同功酶中, KASIII 催化酰基链合成的启动, KASI 催化酰基

链到碳原子数为 16 的棕榈酰 - ACP 的延长反应。从拟南芥属中分离出多种与植物的脂质合成相关的酶的变异体, 特别对进行去饱和作用的酶进行了详细分析。也有关于 KASII 的报道, 将其变异体命名为 fabI。该变异体中, KASII 酶活性降低到 65%, 结果棕榈酸的含量在叶中增加 7%、根中增加 3% (Wu 等, Plant Physiol. (1994)106:143-150)。关于 KASII 酶的完全纯化, 从蓖麻 (Ricinus communis; 特表平 6 - 500234 号公报) 和大豆 (Glycine max; 特表平 7 - 501446 号公报) 克隆以该纯化酶限制性分解肽的氨基酸序列为基础的基因。上述公报中, 对于用这些基因进行转化所致的脂肪酸变化是, 在大肠杆菌中表达蓖麻的基因时, 碳原子数为 16 的脂肪酸含量约减少 20%, 总的脂肪酸含量增加 30%, 而将大豆的基因导入カノーラ (canola) 的时候, 其种子的棕榈酸减少 0.8%, 导入烟草的时候, 其叶中棕榈酸约减少 2%。而不可思议的是未见到蓖麻和大豆二者的基因间存在明确同源性的序列。

可是, 已知构成生物膜的膜脂, 相转移温度主要是根据与脂质结合的脂肪酸的不饱和程度有所变化, 结果生物的耐低温性也有所变化。为了提高膜脂的不饱和程度, 例如脂肪酸的转移酶 (PCT/JP 92/00024 (PCT/WO 92/13082))。脂肪酸的去饱和酶 (PCT/JP 94/02288 (PCT/WO 95/18222)) 等有效。

发明公开

基于上述情况, 本发明的目的是提供蛋白质的基因及其表达产物酶蛋白质, 该蛋白质具有这样的酶活性: 它能调节或控制植物细胞或微生物细胞中饱和脂肪酸和不饱和脂肪酸的含量。

如果存在具有这样酶活性的蛋白质, 即与脂肪酸合成相关的酶活性降低, 细胞的脂质中棕榈酸含量增加, 而酶活性增加时, 碳原子数为 18 以上的脂肪酸的含量增加, 也许脂质中不饱和脂肪酸的含量增加, 例如作为由于酶活性的增加, 碳原子数为 18 以上的脂肪酸含量增加的结果。

为了解决上述问题, 本发明人进行了认真研究, 结果从蓝绿藻 (Anacystis nidulans) 中成功分离了编码 β - 酮酰基 - ACP 合成酶 II (KASII) 的基因, 并发现把该基因转入大肠杆菌可付与 KASII 的生产能力, 由此链长延长的脂肪酸增加, 基于以上发现, 完成了本发明。

也就是说，本发明涉及一种蛋白质，它具有序列号 2 中所示的氨基酸序列，或所具有的氨基酸序列基本上与序列号 2 中的氨基酸序列是同一氨基酸序列，且具有 KASII 酶活性。

另外本发明还涉及编码蛋白质的 KASII 酶基因，所编码的蛋白质具有序列号 2 中所示的氨基酸序列，或具有基本上与序列号 2 中氨基酸序列相同的氨基酸序列，且具有 KASII 酶活性。

再者，本发明涉及含有上述基因的重组载体和转入了上述基因的细胞。

附图简述

10 图 1 表示由单字符表示的、与序列表 2（3 字符表示的氨基酸）相对应的氨基酸序列。

图 2 是大麦（SWISS - PROT 数据库中登记号为 P23902）KASI 酶的氨基酸与由蓖麻子（GENBANK 登记号 L13241）KASII 的 DNA 序列推测的氨基酸序列的比较图。

15 图 3 是本发明中来源于组囊藻（*Anacystis nidulans*）的 KASII 酶（图中以 KASII 表示）与集胞藻类 PCC6803 株（*Synechocystis* sp. strain PCC6803）的 fabF 或 J（编号 sll 1069）的氨基酸序列（记作 Fab）的比较图。

发明的最佳实施状态

20 以下对本发明进行详细说明。

KASII 酶活性蛋白质及其基因

如上所述，本发明具有 KASII 酶活性的蛋白质具有序列号 2（与图 1 的氨基酸序列（单字符表示）相对应）所示的氨基酸序列，或者具有与序列号 2 所示的氨基酸序列基本上相同的氨基酸序列，另外，本发明的 KASII 酶活性蛋白质基因编码具有序列号 2 中所示氨基酸序列或具有基本上与序列号 2 中所示氨基酸序列相同的氨基酸序列的上述酶蛋白质。本发明中“具有 KASII 酶活性的蛋白质”或“KASII 酶活性蛋白质”是指具有通过延长脂肪酸（特别是棕榈酸）链而生成高级脂肪酸（特别是硬脂酸）这样酶活性的蛋白质。

30 象这样的 KASII 酶活性蛋白质，除可应用自然界中存在的合适基因产物外，只要是保持上述 KASII 酶活性，这些蛋白质的氨基酸序列的一部分变化了的变异基因产物亦可用于本发明。作为 KASII 酶活性蛋白质

基因产物的例子，来源于微生物蓝绿藻（序列号 2）的 KASII 酶可为代表例。

5 本发明中的“基本上相同氨基酸序列”也包括上述变异体的序列，典型的是序列号 2 中所示的来源于蓝绿藻的酶蛋白质的氨基酸序列（序列号 2，图 1），或者是该序列中 1 个或多个，优选 1 个或几个氨基酸发生取代、缺失、插入或添加。

因此，本发明中“编码基本上相同氨基酸序列的基因”中所讲的“基本上”一词表示，不仅包括编码自然界中存在的具有上面定义的 KASII 酶活性的蛋白质的 DNA 序列的基因，而且包括编码如上所述变异型
10 KASII 酶活性蛋白质的 DNA 序列的基因，典型的是编码序列号 2 中所示氨基酸序列或者该序列中 1 个或多个，优选 1 个或几个氨基酸发生取代、缺失、插入或附加了的氨基酸的 DNA 序列的基因。另外，通常情况下 DNA 链编码具有某种氨基酸序列的多肽的时候，与一个氨基酸对应的基因密码（密码子）有多个存在（简并的变异体），不用说在编码
15 本发明的 KASII 酶活性蛋白质的 DNA 链中亦可应用任意的基因密码。

本发明的基因编码的 KASII 酶活性蛋白质具有前面所述原本存在于植物和微生物中的脂肪酸合成的链延长酶的功能，更具体地讲，是具有延长脂肪酸（特别是棕榈酸）的链，生成更高级脂肪酸（特别是硬脂酸）的酶活性。本发明蛋白质的典型例是来源于蓝绿藻的蛋白质，但这种
20 种来源于蓝绿藻的 KASII 酶的化学结构与大肠杆菌和大麦的 KASI 基因编码的蛋白质局部相似，另外与前面所讲的 KASII 相关的申请公开中来源于蓖麻子的酶也局部相似。本发明的 KASII 酶活性蛋白质具有如前所述序列号 2 中所示的氨基酸序列，或者其序列基本上为同一氨基酸序列，并且具有延长脂肪酸（特别是棕榈酸）的链长，生成更高级脂肪酸
25 （特别是硬脂酸）的显著地高酶活性。

获得编码本发明蛋白质的基因的一个手段是根据核酸合成的方法化学合成该链长的至少一部分，但若考虑到结合氨基酸数目的话，优选这样的方法，从天然物质，尤其是从细菌蓝绿藻分离 mRNA，由它合成 cDNA，从该基因的文库中，用基因工程学领域常用的方法来获得，
30 而非化学合成法。

KASII 酶基因可如下获得。

首先，应用众知的方法从高等植物或微生物，特别是从蓝绿藻纯化酶，用肽酶将之片段化，测定该片段的氨基酸序列。接着，合成与测定了氨基酸序列的多肽片段相对应的寡核苷酸。此外，从植物或微生物中提取总 RNA，合成与该 RNA 互补的 DNA (cDNA)。将该 cDNA 连接到象噬菌体 λ gt11 这样的合适的载体中制作 cDNA 文库。这里用于该基因的筛选方法，可应用常用的众知的方法，例如应用抗体的噬斑杂交法、或集落杂交法等免疫学方法，或应用核苷酸探针进行杂交的方法等。

另外也可这样得到目的序列：在已知的 KASII 酶或者与之相关的其它同功酶的共同序列的基础上，设计与位于目的序列两端的短 DNA 序列相对应的引物，以测定全长序列时使用的材料而来的 DNA 作模板进行 PCR。这种情况下为了区别 KAS 基因的同功酶，例如可通过在大肠杆菌中表达该基因来鉴定活性。

可用通常众知的方法测定和确认如此选拔的克隆中本发明基因的碱基序列。例如可用 M13 噬菌体通过双脱氧核苷酸链终止法 (Sambrook 等，分子克隆，第 2 版 (1989)) 进行。

可用通常众知的手段，如用亚磷酸酯法，利用市售的 DNA 合成酶合成如上测定了碱基序列的本基因。

另外，为了表达 DNA 链或其片段，生产其编码的蛋白质或多肽，除与该氨基酸序列相对应的 DNA 序列 (编码区域) 外，表达调节序列也是必要的。因此，本发明的 DNA 链也包括含有这样表达调节序列的 DNA 序列。该表达调节区域中，尤其是为了在高等植物中表达，重要的是编码区域上游的启动子序列 (例如来源于花椰菜花叶病毒的 35S 启动子) 和下游 polyA 加尾信号 (例如来源于胭脂碱合成酶的终止子)。如所得 DNA 为高等植物的基因组基因的时候，若 DNA 序列中含有表达调节区域则可直接应用。

KASII 酶活性蛋白质基因的用途

如前所述，本发明涉及含有上述 DNA 链或其片段的重组载体以及转入了该基因的细胞。

重组载体是将上述 DNA 链或其片段连结到载体而形成的，作为载体可应用质粒 (例如 pET17b)、噬菌体 (例如 λ ZAPII) 等众知的载体。

通过将该重组载体 - DNA 转入到如上所述的适宜植物或微生物细胞中进行表达，可在宿主中生产 KASII 酶。

作为细胞，微生物细胞、植物细胞均可、不考虑生物的种类，但微生物细胞推荐大肠杆菌等，植物细胞为烟草等低温敏感性植物等。把基因转入植物的方法，可用通常的众知的方法，例如可用“Plant Molecular Biology Manual, 第 2 版；S. G. Gelvin 和 R. A. Schilperoort 编，Kluwer Academic Publishers, 1995”记载的方法进行。其例子可举出应用病毒的方法，应用土壤杆菌属的方法等生物学方法，电穿孔法、聚乙二醇法、粒子枪法等物理化学方法。

另外，KASII 酶是存在于植物叶绿体包膜的蛋白质，因此有必要将编码叶绿体转移肽的 DNA 链加到 KASII 酶的上游。其实例为豌豆的核酮糖 - 1,5 - 二磷酸羧化酶的小亚单位基因用作编码转移肽的基因。

本发明的 KASII 酶活性蛋白质的基因，代表性的编码来源于蓝绿藻组囊藻的 KASII 酶活性蛋白质（序列号 2）的基因（来源于蓝绿藻的基因的 DNA 序列为序列号 1），通过转化可用于改良植物、微生物的脂质组成，尤其是可用于控制碳原子数 16 和 18 的脂肪酸的含量比。

本发明的 KASII 酶蛋白质作为外来蛋白质在生物体内表达，可使脂肪酸的链长由 16 碳（棕榈酸）伸长到 18 碳（硬脂酸），硬脂酸在生物体内去饱和，不饱和脂肪酸的含量增加。不饱和脂肪酸增加的植物其耐低温性增加（PCT/WO 92/13082，PCT/WO 95/18222），另外期待不饱和脂肪酸增加的微生物（酵母、大肠杆菌等）对培养应急的耐性增加。

实施例

以下列举实施例更详尽地说明本发明，但本发明并不限于这些实施例。

实施例 1 来源于组囊藻的 DNA 的制备以及 DNA 文库的制作

用依照 Shaw 的 Plant Molecular Biology 279 页（IRL 出版，1988）一书中记载的方法配制的大约 100ml 的 BG-11 培养基培养组囊藻（分类号：IAM M - 6，可从东京大学分子细胞学研究所购入）。25℃、1000lux 的荧光灯下每分钟振荡 120 次，充分培养细菌。室温下 500g 离心 10 分钟回收菌体。

分了分离 DNA，将沉淀的菌体悬浮到 50ml 的 50mM Tris Cl (pH 8.0)，1mM EDTA (A 液) 中，再次离心，洗净。接着，再悬浮到 15ml 的 50mM Tris Cl (pH 8.0)、20mM EDTA、50mM NaCl、0.25M 蔗糖的溶液 (B 液) 中，向 B 液中加入溶解的 40mg 溶菌酶 (Sigma 公司)，37 °C 缓慢振荡。1 小时后加入 15mg 的蛋白酶 K 和终浓度为 1% 的 SDS，37 °C 缓慢振荡 1 晚。次日加入 NaClO₄，使浓度为 1M，20ml 的氯仿/异戊醇 (24:1)，缓慢振荡 10 分钟，离心，收集水层。用氯仿/异戊醇的提取再重复一次后，加入 50ml 的乙醇，通过缠绕玻璃棒回收 DNA。将 DNA 溶到 20ml A 液中，NaCl 达到 0.1M，加入 50mg/ml 的 RNase，37 °C 反应 1 小时。然后，用在 A 液中饱和了的等量酚提取 2 次。加入乙醇回收水层中的 DNA，用 70% 的乙醇洗净后，溶到 1ml A 液中作为 DNA 溶液。

为了从所得的 DNA 制作基因组 DNA 文库，用 Sau 3A I 部分消化 100μg DNA 后，依照 Sambrook 等的方法，通过蔗糖密度梯度离心回收到大约 9 - 23kb 的 DNA。用 Bam HI 和 Hind III 进行酶切，克隆到 λDASH II (Stratagene 公司的试剂盒)。

实施例 2 来自蓝绿藻组囊藻的 KASII 酶相似基因的克隆

将大麦的 KASI 酶与蓖麻子的 KASII 酶进行比较，关注二者间同源性高的区域，合成几个短的 DNA 链 (图 2)。观察在如下所示的组合反应中，其中与预测的大小相一致的明显的带。

1. 5' - CC(ACGT)CC(AG)AA(ACGT)CC(AG)AA(ACGT)GA(AG)TT-3'
(序列号 3)
- 25 2. 5'-GA(AG)GA(AG)GT(ACGT)AA(CT)TA(CT)AT(ACT)AA(CT)GC-3'
(序列号 4)

其中序列号 4 是与氨基酸序列 EEVNYINA 相对应的有义引物，序列号 3 是编码与氨基酸序列 NSFQFGG 相对应的反义链的引物。应用有义和反义引物进行 PCR 反应。反应条件是 100μl 反应液中加入各 20mM 的引物、来自组囊藻的 DNA 1μg，应用 Gene Amp™ 反应试剂盒 (宝酒造) 进行。反应程序为 95 °C (1 分)、50 °C (1 分)、72 °C (2

分)为1个循环,进行35个循环。但仅在第1个循环,95℃反应延长到3分,50℃的反应变为37℃。反应结束后,用100μl的氯仿提取,回收水层,然后用100μl的乙醚除去氯仿得到水层,取10μl用2%琼脂糖凝胶电泳进行分析。

5 结果检测出与预想的大小相同(约330bp)的DNA。将该DNA片段克隆到pCRII载体(Invitrogen公司)中。通过双脱氧法(Applied Biosystem公司)应用荧光测序仪来检测DNA的碱基序列。用Multiprime DNA labelling Kit(Amersham公司)将该DNA制作成³²P-dCTP标记的探针,用于以下的杂交实验。

10 用DNA文库的噬菌体感染大肠杆菌P2392,在放入了NZYM培养基直径约15cm的平皿上形成大约1万个噬斑后,转移到尼龙膜上。依照Sambrook等书中的方法(Molecular Cloning: Second edition, Cold Spring Harbor Laboratory Press, 1989)进行杂交。其条件是在由5×SSC(1×SSC: 0.15M NaCl, 15mM柠檬酸钠)、10mM EDTA、10×Denhardt液(50×Denhardt液: Ficoll(400型, Pharmacia)、聚乙烯吡咯烷酮、牛血清的蛋白(fraction V, Sigma)各10g/L)和250μg/ml鲑鱼精子DNA构成的溶液中,60℃、15个小时。其后,在5×SSC、0.1%SDS溶液中,45℃、15分钟将膜洗2次,进行放射自显影。将10个阳性克隆纯化后得到噬菌体DNA。用几个限制性酶进行酶切,用琼脂糖凝胶电泳后,按照常规方法进行Southern印迹。按与上述噬斑杂交相同的条件将膜进行杂交,比较杂交的强度和DNA片段的长度。结果,其中2个克隆(λB、F)无论从强度或片段长度哪一些点来考虑均很满意,因此再用几种限制性酶进行酶切,进行Southern杂交。结果,用Sal I进行酶切的时候,检测出约5kbp的杂交片段,因此把它亚克隆到pUC19(宝酒造)的Sal I位点(分别称作pB、pF)。当用限制性酶对各克隆进行更详细的酶谱分析时,发现pB和pF为同一DNA片段,因此按照常规方法用限制性酶对后者制作缺失质粒,对杂交的DNA片段周围边2kbp的DNA链的碱基序列用荧光测序仪进行检测(序列号1)。其中有1251bp组成的开放阅读框架(ORF),推断是417个残基的氨基酸序列(序列号2)。将该氨基酸序列的同源性数据库记录的蛋白质进行比较,结果显示与脂肪酸合成酶有明显的同源性。尤其是与上总 DNA研究所制备的集胞藻类



PCC6803 株的总基因组序列中 fabF 或 J (数据库 DDBJ 登记号 D90905、PID, g1652389) 形成的蛋白质显示最大的 74 % 的同源性 (图 3)。其它的蛋白质中, 与来源于蓖麻籽的 KASII 基因、来源于大麦的 KASI 基因有 43 - 46 % 的同源性, 另外与来源于大肠杆菌的 KASI 有 35 % 的同源性。但是由基因的同源性不能判别本 ORF 的功能对应 KASI、II 或 III 哪一个。

实施例 3 来源于组囊藻的 KASII 相似基因在大肠杆菌中的活性测定

为了特定上述基因的功能试着在大肠杆菌中表达。首先, 为了去除 ORF 前后多余的 DNA 序列, 制作这样的 DNA, N 端将 Nde I 位点转入由 DNA 合成 (Applied Biosystem 公司) 的起始密码子 (ATG) 的 5' 端, 而在 ORF 的在下方转入 Hind III 位点。

1. 5' - CGCACATATGACTGAAACCGGACGCC (序列号 5)
- 15 2. 5' - CCGCAAGCTTGCAGCAGCGCGTACTGC (序列号 6)

以两个合成的 DNA 作引物, 以 pF 为模板 DNA 进行 PCR 反应。反应条件按照 Perkin-Elmer 公司的手册进行, 以 94 °C (1 分)、60 °C (1 分)、72 °C (2 分) 为一个循环, 反应进行 30 个循环。用 Nde I 和 Hind III 酶切反应产物后, 将在相同限制性酶切位点预先酶切了的 pET17b 克隆到大肠杆菌 DH5, 然后将之克隆到 BL21 (DE3) plys S (Novagen)。

在加入了 100 µg/ml 氨苄青霉素和 30 µg/ml 氯霉素的 75ml LB 培养基中培养 (32 °C) 后者的大肠杆菌的重组体, 直到培养液的浓度在波长 25 600nm 处达到 0.5OD。其后, 加入终浓度达 0.4mM 的 IPTG, 再继续培养 2 小时。10000 × g 离心 10 分钟, 从培养液中回收大肠杆菌, 在 50mM Tris HCl (pH 7.4) 中洗涤细菌, -20 °C 冻结。在由 20mM Tris HCl (pH 8.0)、20mM 二硫苏糖醇、10mM MgCl₂ 和 1 µg/ml DNaseI 组成的溶液中, 于冰上溶解菌体。其后, 10000 × g、4 °C 离心 1 小时, 30 将上清中的蛋白溶液在聚丙烯酰胺构成 10 ~ 20 % 浓度梯度的板电泳胶上进行 SDS 电泳, 用考马斯亮兰染色。结果检测出来源于组囊藻的蛋白质为分子量约 50kDa 的蛋白质。

至于该大肠杆菌脂肪酸的组成，从上述培养的菌体中将脂肪酸作为甲基酯进行回收、分析。甲基化是将大约 5mg 的脂质与 5 % 溶于无水乙醇的盐酸 1ml 一起在封闭的管中在沸水中加热 4 小时，冷却后，用己烷抽提脂肪酸甲基酯。用氢火焰离子检测器，在毛细柱（聚酯液相； 10 % EGSS - X， 175 °C）上分析甲基化脂肪酸酯。通过与标准甲基化脂肪酸的相对保留时间来测定脂肪酸。结果如下表所示。

大肠杆菌的脂肪酸组成

样品	14:0	16:0	16:1	18:0	18:1	16:0 + 16:1/18:0 + 18:1
对照	2	37	21	1	38	1.49
重组体 1	0	24	13	6	57	0.59
重组体 2	0	21	11	7	62	0.46

10 结果表明，碳原子数 16 的脂肪酸（16:0 和 16:1）减少，相反碳原子数 18 的脂肪酸（18:0 和 18:1）的比例大幅度增加。

产业上利用的可能性

15 本发明提供编码以来源于组囊藻的 KASII 酶所代表的具有 β - 酮酰基 - ACP 合成酶 II 活性的蛋白质的 DNA 链。如上所述，编码本发明酶蛋白质的基因是 KASII 酶的基因，它将脂肪酸（尤其是 C16 的棕榈酸）转变为更高级脂肪酸（尤其是 C18 硬脂酸）的活性明显的高，通过转化，它有益于植物脂质的改良、微生物脂质的改良、特别是碳原子数 16 和 18 的脂肪酸含量比例的调整或不饱和脂肪酸含量的增加。

20

25



序列表

申请人：麒麟麦酒株式会社

发明名称： β -酮酰基-ACP合成酶II酶及编码该酶的基因

档案号：113702-432

5 申请日期：平成10年1月20日

序列数：6

序列号：1

序列长度：1251

10 序列类型：核酸(DNA)

链数：双链

拓扑学：直链

序列种类：基因组DNA

起源：

15 生物名：组囊藻

株名：IAM M-6

序列

ATGACTGAAA CCGGACGCCA GCGTGTGTT ATTACTGGTT TGGGAGCCAT TACTCCCATC 60

GGTAATGATC CAACGGAATA TTGGCAGGGA ATCCTTGCCG GTCGCAACGG CATCGATCTG 120

20 ATTCTGGGGCT TTGATGCGTC TCGTCACGCC TGCAAAATTG CCGGGGAGGT CAAGGACTTT 180

GACCCACCC AGTACATGGA CCGCAAGGAT GCTAAGCGGA TGGATCGGTT TGCACAACCTG 240

GCGGTTGCTG CCAGTCGCCA AGCAGTCGCC GATGCCAAGC TGGACATCAC TGAAGTGAAT 300

GCGGATGCCA TCGGGGTGCT GATCGGCTCA GGCATTGGTG GTTTGAGGGT GATGGAGGAC 360

25 CAGCAGACGG TTTTGCTGGA AAAAGGCCCC GATCGCTGCA GCCCCTTCAT GGTGCCGATG 420

ATGATCGCCA ACATGGCGGC AGGACTGACG GCCATCCAGT TGGGTGCCAA AGGCCCTTGC 480

AATGTCACGG TGAAGTCTG CGCTGCGGGT TCTAATGCGG TGGGTGAAGC CTTCGGGCTG 540

ATTCAGCAGG GCTATGCCCA AGCCATGATC TGTGGCGGAA CTGAATCCTG TGTGACCCCA 600

30 CTGGCTATGG CCGGTTTTGC GGCCTGTAAG GCACTGTCCG TGCCCAACGA TGACCCGGCC 660

CATGCTTGCC GTCCCTTTGA CCAAGGCCGT GATGGTTTTG TGATGGGCGA AGGCCGAGGG 720

ATTTTGGTCT TGAATCCTT GGAGCATGCC CAAGCGAGGG GCGCTCACAT CTATGGCGAA 780



ATCGTCGGCT ATGGCATGAC CTGTGATGCC TATCACATCA CCTCGCCGGT CCCAGGTGGT 840
TTGGGTGCCG CCCGGGCGAT CGAGTTCGGG CTCCGCGATG CCAATCTGCA GCCCAGCCAA 900
5 GTCAGCTACA TCAATGCTCA CGGCACCAGC ACACCGGCCA ACGACAGCAC CGAAACGGCA 960
GCTATTAAGA AAGCCCTAGG TGAGCAGGCC TACAAAACCG TGATCAGCTC GACTAAGTCG 1020
ATGACCGGTC ACCTGTTAGG GGGCTCCGGC GGAATTGAGG CGGTAGCGGC AACCTCGCG 1080
ATCGCTGAGG ACATGGTGCC GCCGACGATT AACCTGGAAG ATCCCGATCC CGATTGCCAC 1140
10 TTGGACTATG TCCCAATCA GGCGCGATCG CTACCGGTGG AAGTGGCTTT GTCCAATTCC 1200
TTCGGCTTTG GTGGGCACAA CGTCACGCTG GCCTTCCGGA AATTCATCC C 1251

序列号： 2

序列长度： 417

15 序列类型： 蛋白质

拓扑学： 直链

序列类型： 肽

起源：

生物名： 组囊藻

20 株名： IAM M - 6

表示特征的记号： CDS

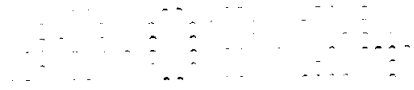
确定特征的方法： P

序列

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



180 185 190
Gly Thr Glu Ser Cys Val Thr Pro Leu Ala Met Ala Gly Phe Ala Ala
195 200 205
Cys Lys Ala Leu Ser Leu Arg Asn Asp Asp Pro Ala His Ala Cys Arg
210 215 220
Pro Phe Asp Gln Gly Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly
225 230 235 240
Ile Leu Val Leu Glu Ser Leu Glu His Ala Gln Ala Arg Gly Ala His
245 250 255
Ile Tyr Gly Glu Ile Val Gly Tyr Gly Met Thr Cys Asp Ala Tyr His
260 265 270
Ile Thr Ser Pro Val Pro Gly Gly Leu Gly Ala Ala Arg Ala Ile Glu
275 280 285
Phe Gly Leu Arg Asp Ala Asn Leu Gln Pro Ser Gln Val Ser Tyr Ile
290 295 300
Asn Ala His Gly Thr Ser Thr Pro Ala Asn Asp Ser Thr Glu Thr Ala
305 310 315 320
Ala Ile Lys Lys Ala Leu Gly Glu His Ala Tyr Lys Thr Val Ile Ser
325 330 335
Ser Thr Lys Ser Met Thr Gly His Leu Leu Gly Gly Ser Gly Gly Ile
340 345 350
Glu Ala Val Ala Ala Thr Leu Ala Ile Ala Glu Asp Met Val Pro Pro
355 360 365
Thr Ile Asn Leu Glu Asp Pro Asp Pro Asp Cys Asp Leu Asp Tyr Val
370 375 380



Pro Asn Gln Ala Arg Ser Leu Pro Val Glu Val Ala Leu Ser Asn Ser

385 390 395 400

5 Phe Gly Phe Gly Gly His Asn Val Thr Leu Ala Phe Arg Lys Phe His

 405 410 415

Pro

417

10 序列号： 3

序列长度： 20

序列类型： 核酸（DNA）

链数： 双链

拓朴学： 直链

15 序列种类： 其它的核酸（合成DNA）

序列

5' -CC (ACGT) CC (AG) AA (ACGT) CC (AG) AA (ACGT) GA (AG) TT-3' 20

序列号： 4

20 序列长度： 23

序列类型： 核酸（DNA）

拓朴学： 直链

序列种类： 其它的核酸（合成DNA）

序列

25 5' -GA (AG) GA (AG) GT (ACGT) AA (CT) TA (CT) AT (ACT) AA (CT) GC-3' 23

序列号： 5

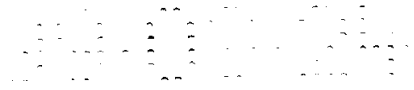
序列长度： 26

序列类型： 核酸（DNA）

30 拓朴学： 直链

序列种类： 其它的核酸（合成DNA）

序列



5' -CGCACATATGACTGAAACCGGACGCC-3'

26

序列号： 6

序列长度： 27

5 序列类型： 核酸（DNA）

拓扑学： 直链

序列种类： 其它的核酸（合成DNA）

序列

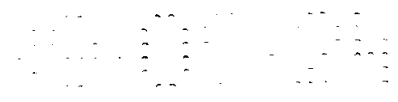
5' -CCGCAAGCTTGCAGCAGCGGTACTGC-3'

27

说明书附图

MTETGRQRV ITGLGAIPI GNDPTEYWOG ILAGRNGIDL IRGFDASRHA CKIAGEVKDF
DPTQYMDRKD AKRMDRFAQL AVAASRQAVA DAKLDITELN ADAIGVLIGS GIGGLRVMED
QQTVLLEKGP DRCSPFMVPM MIANMAAGLT AIQLGAKGPC NVTVTACAAG SNAVGEAFRL
IQHGYAQAMI CGGTESCVTP LAMAGFAACK ALSLRNDDPA HACRPFDOGR DGFVMGEGAG
ILVLESLEHA QARGAHYGE IVGYGMTDA YHITSPVPGG LGAARAIEFG LRDANLOPSQ
VSYINAHGTS TPANDSTETA AIKKALGEHA YKTVISSTKS MTGHLLGGSG GIEAUAATLA
IAEDMVPPTI NLEDPPDCD LDYVPNOARS LPVEVALSNS FGFGGHNVTL AFRKFHP

图 1



	10	20	30	40	50	
大麦	MHAHAHALGLRVPPAFPRRRARPRR—PAAAVLATSAAPORE——TDP——RKRVV					
	::	::	::	::	::	
蓖麻	PCSHYSSNGLFPNTPLLPKRHPRLHRLPRSGEAMAVAVQPEKEVATNKKPLMKORRVV					
	80	90	100	110	120	130
	60	70	80	90	100	110
大麦	ITGMGLASVFGSDVDTFYDRLLAGESGVGPIDRFDASSFPTRFAGQIRGFSSEGYIDGKN					
	::	::	::	::	::	::
蓖麻	VTGMGVVSPLGHDIDVYYNLLDGSSGISOISDFCAQFPTRIAGEIKSFSTDGWVAPKL					
	140	150	160	170	180	190
	120	130	140	150	160	170
大麦	DRRLDDCIRYCILSGKKALESAGLGAGSDAHVKLDVGRAGVLVGTGMGGLSVFSDGVQNL					
	::	::	::	::	::	::
蓖麻	SKRMDKFMLYMLTAGKKALADGGI—TEDMMDELDKARCGVILGSAMGGMKVFNDALIEAL					
	200	210	220	230	240	
	180	190	200	210	220	230
大麦	IEKGYRKISPFPIPYAITNMGSAALLIDVGFMGPNYSISTACATSNYCFYAAANHIRRGE					
	.X::	::	::	::	::	::
蓖麻	-RISYRKMNPFCVFPFATTNMGSAALLAMD LGWMGPNYSISTACATSNFCILNAAANHIRRGE					
	250	260	270	280	290	300
	240	250	260	270	280	290
大麦	ADII VAGGTEAAIIPILGGFVACRALSORNDDPI TACRPWDKERDGFVMGEGAGVLVME					
	::	::	::	::	::	::
蓖麻	ADIMLCGGSDAAIIPILGGFVACRALSORNDDPTKASRPWDMNRDGFVMGEGAGVLLLE					
	310	320	330	340	350	360
	300	310	320	330	340	350
大麦	SLEHAMKRDAPIIAEYLGAVNCDAYHMTDPRADGLGVSSCI TMSLRDAGVAPEEVNYIN					
	::	::	::	::	::	::
蓖麻	ELEHAKKRGANIYAEFLGGSFTCDAYHMTDPRPDGVGVILCIEKALARSGVSKEEVNYIN					
	370	380	390	400	410	420
	360	370	380	390	400	410
大麦	AHATSTLAGDLAEVRAIKOVFKNPSEIKINSTKSMIGHCLGAAGGLEAIATIKSITTGWV					
	::	::	::	::	::	::
蓖麻	AHATSTPAGDLKEYEALMRCFSONPDLRVNSTKSMIGHLLGAAGAVEAIATIQAIRTGWV					
	430	440	450	460	470	480
	420	430	440	450	460	
大麦	HPTINQFNPEPEVDFDTVANEEKOH-EVNVGINSFSGFGGHNSVWFAPFK					
	::	::	::	::	::	::
蓖麻	HPNINLENPEEGVDTKVLVGPCKERLDIKVALSNSFGFGGHNSIIFAPYK					
	490	500	510	520	530	

图 2
2

Fab	MANLEKKRVVVTGLGAI TP GNTLQDYWQGLMEGRNG GP TRFDASDQACRFGGGEVKDFX:::.....	10	20	30	40	50	60
KAS I I	MTETGRQRVV TGLGAI TP GNDPTEYWQG LAGRNG DL IRGFDASRHACK AGEVKDF	10	20	30	40	50	60
Fab	DATQFLDRKEAKRMDRFCHFAVCASQQA NDAKLV NELNADE GVL GTG GGLKVLED :::.....	70	80	90	100	110	120
KAS I I	DPTQYMDRKDAKRMDFEAQLAVAASROAVADAKLD TELNADA GVL GSG GGLRVMED	70	80	90	100	110	120
Fab	QQT LLDKGPSRCSPPFI PMMI ANMASGLTA NLGAKGPNNCTVTACAAGSNA GDAFRL :::.....	130	140	150	160	170	180
KAS I I	QQTVLLEKGPDRCSPPFMVPMI ANMAAGLTA QLGAKGPCNVTVTACAAGSNAVGEAFRL	130	140	150	160	170	180
Fab	VQNGYAKAWI CGGTEAAI TPLSYAGFASARAL SFRNDDPLHASRPFDKDRDGFVMGEGSG :::.....	190	200	210	220	230	240
KAS I I	IQHGYAQAMI CGGTESCVTPLAMAGFAACKAL SLRNDOPAHACRPFDOGRDGFVMGEGAG	190	200	210	220	230	240
Fab	ILILEELESALARGAK YGEMVGYAMTCDAYH IAPVPDGRGATRA AWALKDSGLKPEM :::.....	250	260	270	280	290	300
KAS I I	ILVLESLEHAQARGAHI YGEI VGYGMTCDAYH I TSPVPGGLGAARA EFGLRDANLQPSO	250	260	270	280	290	300
Fab	VSY NAHGTSTPANDVTETRA KDALGNHAYN AVSSTKSMTGHLLGGSGG EAVATVMA :::.....	310	320	330	340	350	360
KAS I I	VSY NAHGTSTPANDSTETAA K KALGEHAYKTV SSTKSMTGHLLGGSGG EAVAATLA	310	320	330	340	350	360
Fab	IAEDKV PPT NLENDPDCDLDYVPGOSRAL VDVALSNSFGFGGHNVTLAFKKYQ :::.....X...	370	380	390	400	410	
KAS I I	IAEDMVPPT NLEDPDPCDLDYV PNOARSLPVEVALSNSFGFGGHNVTLAFRKFH	370	380	390	400	410	

图 3