

油棕乙酰 CoA 羧化酶 (ACC) 基因的鉴定与表达分析

肖勇, 雷新涛, 王永, 曹红星, 石鹏, 金龙飞, 夏薇 (中国热带农业科学院椰子研究所, 海南文昌 571339)

摘要 [目的]对油棕乙酰 CoA 羧化酶(ACC)基因进行鉴定与表达分析。[方法]采用生物信息学方法鉴定油棕的 ACC 基因,并应用拟南芥的 ACC 基因蛋白质序列比对油棕的蛋白质库,应用已有油棕转录组信息,研究 ACC 基因在不同组织中的表达情况。[结果]在油棕中鉴定了 2 个 ACC 基因,命名为 *EgACC1* 和 *EgACC2*。*EgACC1* 不含内含子,而 *EgACC2* 含有 32 个内含子;*EgACC1* 在不同组织中都没有表达,而 *EgACC2* 在所有组织中都有表达,其中在组培苗中有较高的表达量,同时,在油棕果发育 21 d 时有高的表达水平,RPKM 值为 70.6,为所有组织中最高。[结论]该研究阐述了油棕 ACC 基因的表达模式,为油棕 ACC 基因的功能分析奠定了基础。

关键词 油棕;ACC 基因;转录组;RPKM

中图分类号 S188 文献标识码 A 文章编号 0517-6611(2017)31-0154-02

Identification and Expression Profile of Acetyl CoA Carboxylase (ACC) in *Elaeis guineensis*

XIAO Yong, LEI Xin-tao, WANG Yong et al (Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang, Hainan 571339)

Abstract [Objective]To study the identification and expression profile of acetyl CoA carboxylase (ACC) in *Elaeis guineensis*. [Method]To identify ACC genes in *E. guineensis* by using bioinformatics, ACC genes protein sequence of *Arabidopsis thaliana* were used to align with protein database of *E. guineensis*. The transcription group information of *E. guineensis* was applied to study the expression of ACC genes in different tissues. [Result] Two ACC genes were identified in the genome of *E. guineensis*, which were named as *EgACC1* and *EgACC2*, respectively. Among them, *EgACC1* do not contain intron, however *EgACC2* contain thirty two introns. The expression of *EgACC1* was not detected in different tissues of *E. guineensis*. However, the expression of *EgACC2* was detected in all tissues. High expression levels were examined in tissue culture seedling. Meanwhile, highest expression level was detected at 21 day of mesocarp development, RPKM value was 70.6. [Conclusion] These study demonstrated that the expression pattern of ACC genes and provided basic work for functional analysis of ACC genes in *E. guineensis*.

Key words *Elaeis guineensis*; ACC gene; Transcriptome; RPKM

油棕是重要的热带木本油料作物,是一种重要的热带木本油料作物,油棕果的含油量在 50% 以上,1 株油棕产油 30~40 kg,平均产油 4 500 kg/hm²,因单位面积产油量高,而被称为“世界油王”^[1]。油棕的中果皮和果仁都能产油,其中果皮压榨的油为棕榈油,棕榈油的主要成分为棕榈酸,十六碳的饱和脂肪酸(C16:0),而油棕果仁压榨的油为棕榈仁油,棕榈仁油的脂肪酸成分类似于椰子油,其主要成分为月桂酸,十二碳的饱和脂肪酸(C12:0)。棕榈油脂肪酸的饱和度非常高,不易被氧化,被认为是奶油和反式脂肪酸的很好替代品,我国 65% 的棕榈油都用于做人造奶油、起酥油、煎炸油、家庭烹调用的调和油及其他食品专用油^[2]。棕榈酸是目前世界上生产量、消费量和国际贸易量最大的植物油种类,年产油量超过 5 000 万 t,2003—2004 年度,我国棕榈油消费总量约为 574.9 万 t,进口量约占世界进口总量的 13.38%,我国棕榈油消费主要依赖于进口,因而培育我国棕榈油产业非常重要。

脂肪酸合成的前体为乙酰 CoA,它首先在乙酰 CoA 羧化酶的作用下合成丙二酰 CoA,然后脂肪酸合成酶以丙二酰 CoA 为底物进行连续的聚合反应,以每次循环增加 2 个碳的频率合成酰基碳链^[3]。ACC 是脂肪酸生物合成的关键酶之一,自然界中有 2 种存在形式:一种为多功能酶,而另一种为多酶复合体。前者含有 3 个功能结构域,被称为真核形式的 ACC,后者可被分解多个功能蛋白,如生物素羧化酶,生物素羧基载体蛋白,羧基转移酶和另外一个功能未知蛋白,被称为原核形式的 ACC^[4-5]。笔者应用生物信息

学,研究 ACC 基因的结构和保守结构域,并研究不同物种 ACC 基因的进化关系,分析 ACC 基因在油棕组织中的表达情况。

1 材料与方法

1.1 ACC 基因的鉴定 油棕的全基因组序列从 NCBI (the National Center for Biotechnology Information) 网站上下载,拟南芥的 ACC 基因从 TAIR (the Arabidopsis Information Resource) 网站上下载,为了鉴定油棕的 ACC 基因,将拟南芥 ACC 蛋白质序列与油棕的蛋白质数据库进行 Blastp 比对,比对的 *E* 值的阈值为 $1e^{-10}$,将预测的油棕 ACC 蛋白质序列与 CDD (<http://www.ncbi.nlm.gov/cdd>) 和 PFAM 数据库 (<http://pfam.sanger.ac.uk>) 进行比对。

1.2 ACC 基因结构分析 油棕基因组的注释结果从 NCBI (the National Center for Biotechnology Information) 网站上下载,ACC 基因结构采用 GSDS (the Gene Structure Display Server Programme) 软件进行展示。

1.3 ACC 基因的表达分析 油棕的 32 个转录组的原始 reads 数据从 NCBI 网站的 SRA 数据库 (Sequence Read Archive) 中下载,这些转录组分别来自于油棕的中果皮、叶子、果、花、根和茎。为了检测 ACC 基因在油棕不同组织中的表达,对 ACC 基因 RPKM 值进行计算,计算公式如下:

$$RPKM = \frac{10^6 C}{NL/10^3}$$

在该公式中, *C* 代表比对在某一个表达基因序列上的 reads 的条数, *N* 代表比对上所有表达序列的总 reads 数, *L* 代表对应的基因序列的长度^[6]。

2 结果与分析

2.1 油棕 ACC 基因的鉴定 从 TAIR 网站挖掘 ACC 基因,

作者简介 肖勇(1980—),男,湖北监利人,副研究员,博士,从事热带油料作物分子生物学研究。

收稿日期 2017-08-11

有 2 个基因为丙二酰 CoA (ACC), 分别为 *At1g36160* 和 *At1g36180*。用拟南芥的这 2 个基因的蛋白质序列与油棕的蛋白质库进行 Blastp, 设定阈值为 $1e^{-10}$ 。在油棕的基因组上, 发现了 2 个同源基因, 其中 1 个 *EgACC* 基因位于油棕第 3 染色体上, 而另外 1 个 *EgACC* 基因未能定在油棕的染色体上。这 2 个 *EgACC* 基因的长度分别为 1 164 和 18 256 bp。

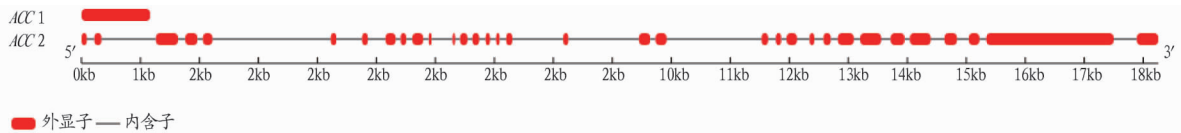


图1 *EgACC1* 基因的结构

Fig.1 Structure of *EgACC1* gene

2.3 不同物种 ACC 基因的聚类分析 下载拟南芥、油菜、油茶、花生、玉米、大豆以及油棕 ACC 基因的蛋白质序列, 并采用 Joining Neighbour 的方法, 对这些蛋白质序列进行聚类, 聚类结构显示: 拟南芥、油茶、玉米、油茶的 ACC 基因聚在一块, 而油棕的 ACC2 和大豆的 ACC2 基因聚在一块, 花生的 ACC 与大豆 ACC1 基因聚在一块, 而油棕的 ACC1 被单独地聚在一类。ACC2 基因在所有组织中都检测到表达量。

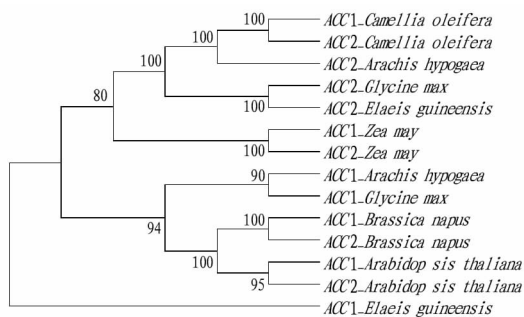


图2 不同物种 ACC 基因的聚类图

Fig.2 Clustering figure of different species of ACC gene

2.4 油棕 ACC 基因在不同组织中的表达 下载油棕的 32 个转录组的数据, 这些转录组分别来自于油棕发育不同时期的中果皮、根、茎、叶、花、果。依据这些转录组的原始数据计算 RPKM 值, 油棕的 ACC1 基因在所有组织中都没有表达, 而 ACC2 在不同的组织都有表达量, ACC2 基因在油棕的组培苗中有高的表达量, 同时在油棕果发育 21 d 时有高的表达水平, RPKM 值为 70.6, 为所有组织中最高。

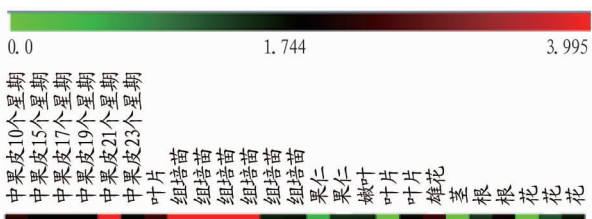


图3 *EgAAC* 基因在油棕不同组织表达的热图

Fig.3 Heat map expressed by *EgAAC* gene in different tissues of *Elaeis guineensis*

3 讨论

油棕是一种重要热带油料作物, 因产油量高, 而被称为“世界油王”。油棕的基因组已被测序, 测序结果显示: 油棕

序列长度较短的 ACC 基因命名为 *EgACC1*, 而序列较长的 ACC 基因命名为 *EgACC2*。

2.2 *EgACC* 基因的结构 分析这 2 个基因的结构, 结果显示 *EgACC1* 并不含有内含子, 总基因的长度为 1 164 bp, 而 *EgACC2* 含有 32 个内含子, 外显子最短的只有 40 bp, 而最长的外显子有 2 157 bp, 平均每个外显子 222.75 bp。

基因组为 1.53 Gb, 预测了 34 802 个基因, 在该研究中, 应用拟南芥 ACC 蛋白质序列比对油棕的蛋白质库, 并鉴定 2 个油棕的 ACC 基因, 分别命名为 *EgACC1* 和 *EgACC2*。对这 2 个 ACC 的基因结构进行分析, 分析结果显示: *EgACC1* 不含内含子, 而 *EgACC2* 含有 32 个内含子, 同时, 研究了这 2 个 *EgACC* 基因在油棕组培苗以及不同组织中的表达, 同时也研究这 2 个 ACC 基因在油棕中果皮发育不同时期的表达变化, 结果显示: *EgACC1* 在所有组织中都没有表达, 而 *EgACC2* 在所有组织中都有表达, 该基因在组培苗中有较高的表达量, 同时, 在油棕果发育 21 d 有高的表达水平, RPKM 值为 70.6, 为所有组织中最高。

在脂类代谢的研究中, 拟南芥是重要的模式植物, 目前已研究了该物种中包括脂肪酸从头合成, 三酰基甘油合成和脂肪酸三酰基甘油降解等 23 条代谢途径, 脂肪酸合成的前体为乙酰 CoA, 它首先在乙酰 CoA 羧化酶的作用下合成丙二酰 CoA, 然后脂肪酸合成酶复合物 (fatty acid synthetase complex, FAS) 以丙二酸单酰 CoA 为底物进行聚合反应, 以每次循环 2 个碳的频率合成酰基碳链^[7]。而不断地增加酰基碳链与酰基载体蛋白结合, 以保护其不受代谢途径中多种酶的侵蚀。一些研究已表明: FAS 常为酮酰基 ACP 合成酶, 多次反应加碳后, 脂肪酸可以在酰基 ACP 硫酯酶 (acyl-ACP thioesterase) 或酰基转移酶 (acyltransferase) 的作用下终止。

ACC 酶有 2 种形式, 即多功能酶和多酶复合体, 前者被称为真核形式的 ACC, 而后者被称为原核形式的 ACC, 原核形式的 ACC 可以被分为几个多功能的亚基, 即生物羧化酶、生物素羧基载体蛋白、羧基转移酶和另一个未知功能的蛋白。在单子叶作物 (如小麦、玉米和水稻) 中的 ACC 酶为真核形式, 这些 ACC 酶对 aryloxy phenoxy propionates 和 cyclohexane diones 这 2 种除草剂非常敏感。但是在双子叶植物豌豆和烟草中, ACC 为原核形式。Gengenbach 等^[8] 检测了 ACC 基因在高油黄豆和低油黄豆中的表达量, 结果显示, 在高油黄豆中 ACC 基因的表达量是低油黄豆的 2 倍, 特别是在发育的早期和中期, 这些结果暗示 ACC 基因的表达与黄豆脂肪酸的积累具有较高的相关性。Ohlrogge 等^[9] 将油棕种子储藏蛋白的启动子与拟南芥的 ACC 基因连接, 并转化油菜, 在 T₁

(下转第 159 页)

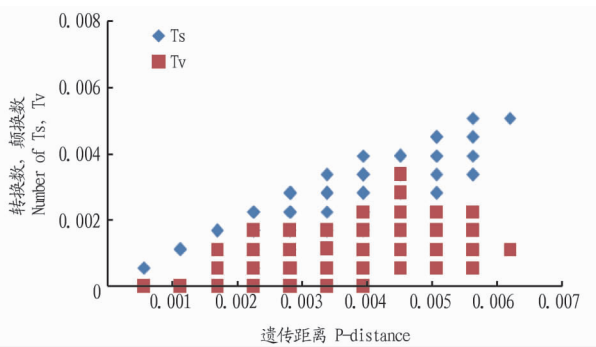


图2 青凤蝶碱基替换饱和和分析

Fig. 2 The substitution saturation analysis of 3 subspecies of *G. sarpedon*

央贯穿 1 列略呈方形的蓝绿色斑(后翅前面的 1 个为白色),后翅外缘有 1 列绿蓝色的新月斑。后翅反面近翅基有 1 条红色短线,翅中部至后缘处有数条红色斑纹。

3 个亚种之间形态特征存在一些差异,青凤蝶蓝斑亚种的绿色斑偏蓝,前翅顶角有 1 个绿斑特别小;斑带亚种的后翅色带不全。有学者提出青凤蝶的斑带亚种是一种变型,而不是亚种。为此,该研究联合 2 个基因序列的整合数据对青凤蝶 3 亚种间遗传变异进行了分析。之所以选择 *CO I* 和 *EF-1 α* 基因联合分析,一方面是增加系统发生分析的置信度^[7-8],另一方面也与 2 个基因各自的特点有关。线粒体 *CO I* 是编码线粒体细胞色素氧化酶亚基 I 的基因,是常用的分子标记之一,它既相对保守又存在高变区,高变区内遗传进化速率更快,种间的遗传差异也更明显,适合于种间和种内分类鉴定及系统发生研究^[9]。*EF-1 α* 基因进化速度比较快,它也用于分析低阶元的系统发育,已有学者做了相关的研究,如 Reed 等^[10]用线粒体 *CO I*、*CO II* 基因的全长和核基因的 *EF-1 α* 基因联合分析了凤蝶科凤蝶属 23 个种和亚种的个体。

在鳞翅目同种个体之间,*CO I* 基因的差异为 0.25%^[11],从试验结果来看,我国青凤蝶 3 个亚种间 *CO I* 基因相似度为 99.15%,核苷酸差异很小,表明 *CO I* 基因在个体间变异率较低。3 个亚种间 *EF-1 α* 基因相似度 96.43%,表明存在一定的变异率。3 亚种青凤蝶在形态上的差异可能

是由于地理隔离形成,其自身的遗传变异性非常低^[12]。

从系统发育树的拓扑结构看,其未显示 3 个亚种间有很好的单系性,而出现了并系性,且各支自举检验值较低(< 50%),因而没有列出发育树,置信度低下可能是因为分子数据所提供的信息位点较少^[13],或者是 *CO I* 基因和 *EF-1 α* 基因序列突变达到饱和。结合对 3 个亚种基因序列两两距离进行 P-distance 计算,并按照亚种分组后计算遗传距离,可以看出指名亚种和斑带亚种的亲缘关系较近,但不足以认定斑带亚种是属于指名亚种的一种变型。

综上所述,结合目前已有的 2 个基因部分序列和各亚种之间形态特征的差异,仍然适宜采用传统的基于形态学的青凤蝶亚种分类体系。

参考文献

- [1] 武春生. 中国动物志:昆虫纲 第 25 卷 鳞翅目 凤蝶科[M]. 北京:科学出版社,2001.
- [2] 魏忠民,武春生. 中国青凤蝶变异型初步观察[J]. 昆虫知识,2006,43(3):431-432.
- [3] 寿建新. 国内外蝴蝶分类认识总结[J]. 西安文理学院学报(自然科学版),2014,17(4):21-27.
- [4] 袁锋,袁向群. 蝶类分子系统学研究进展[J]. 西北农业学报,2013,22(12):1-14.
- [5] KIM K J, JANSEN R K. A chloroplast DNA phylogeny of lilacs (*Syringa*, Oleaceae): Plastome groups show a strong correlation with crossing groups [J]. *American J Bot*, 1998, 85(9):1338-1351.
- [6] AUSTIN C C, ZUG G R. Molecular and morphological evolution in the south-central Pacific skink *Emoia tongana* (Reptilia: Squamata): Uniformity and human-mediated dispersal [J]. *Australian J Zool*, 1999, 47(5):425-437.
- [7] MONTEIRO A, PIERCE N E. Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from *COI*, *COII* and *EF-1 α* gene sequences [J]. *Mol Phylogenet Evol*, 2001, 18(2):264-281.
- [8] NORMARK B B. Molecular systematics and evolution of the aphid family Lachnidae [J]. *Mol Phylogenet Evol*, 2000, 14(1):131-140.
- [9] 吴冬霞, 朱国萍, 陈娜, 等. 线蛭蝶亚科蝶类部分群线粒体 *COI* 基因的系统发生分析 [J]. 生命科学研究, 2007, 11(1):64-71.
- [10] REED R D, SPERLING F A. Interaction of process partitions in phylogenetic analysis: An example from the swallowtail butterfly genus *Papilio* [J]. *Mol Biol Evol*, 1999, 16(2):286-297.
- [11] HEBERT P D N, STOECKLE M Y, ZEMBLAK T S, et al. Identification of birds through DNA barcodes [J]. *PLoS Biol*, 2004, 2(10):1657-1663.
- [12] 叶维萍, 叶海燕, 卢慧蕊, 等. 基于线粒体 12S rRNA 和 ND5 基因序列的中国飞蝗属 3 亚种系统发育关系研究 [J]. 昆虫分类学报, 2005, 27(1):5-12.
- [13] 刘晓燕, 吴孝兵, 诸立新. 中国黄粉蝶亚科六属间基于 *COII* 和 *EF-1 α* 基因部分序列的系统发育关系 (鳞翅目: 粉蝶科) [J]. 昆虫学报, 2007, 50(6):604-609.
- [14] ALBAN C, BALDET P, DOUCE J. Localization and characterization of two structurally different forms of acetyl-CoA carboxylase in young pea leaves, of which one is sensitive to aryloxyphenoxy propionate herbicides [J]. *Biochem J*, 1994, 30:557-565.
- [15] MORTAZAVI A, WILLIAMS B A, MCCUE K, et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq [J]. *Nat Methods*, 2008, 5(7):621-628.
- [16] OHLROGGE J, BROWSE J. Lipid biosynthesis [J]. *Plant cell*, 1995, 7(7):957-970.
- [17] GENGENBACH B G, SOMERS D A, WYSE D L, et al. Transgenic plants expressing maize acetyl CoA carboxylase gene and method of altering oil content; 6222099 [P]. 2001-04-24.
- [18] OHLROGGE J B, ROESLER K R, SHORROSH B S. Method of increasing oil content of seeds; 5925805 [P]. 1999-07-20.
- [19] SELLWOOD C, SLABAS A R, RAWSTHORNE S. Effects of manipulating expression of acetyl-CoA carboxylase in *Brassica napus* L. embryos [J]. *Biochemical society*, 2000, 28(6):598-600.

(上接第 155 页)

代,拟南芥的脂肪酸含量增加 5.0%~6.4%。Sellwood 等^[10]通过反义表达技术,抑制 *ACC* 基因的表达,转基因植株的含油量显著降低。

参考文献

- [1] BASRI W M, ABDULLAH S N A, HENSON I E. Oil palm: Achievements and potential [J]. *Plant production science*, 2005, 8(3):288-297.
- [2] KOH L P, WILCOVE D S. Cashing in palm oil for conservation [J]. *Nature*, 2007, 448:993-994.
- [3] 李昌珠, 李正茂. 植物脂肪酸的生物合成及其生理功能的研究进展 [J]. 湖南林业科技, 2009, 36(6):45-49.
- [4] ASHTON A R, JENKINS C L D, WHITFIELD P R. Molecular cloning of two different cDNAs for maize acetyl CoA carboxylase [J]. *Plant Mol Biol*, 1994, 24(1):35-49.