

DOI: 10.3724/SP.J.1005.2013.01403

基于油桐种子 3 个不同发育时期转录组的油脂合成代谢途径分析

陈昊^{1,2}, 蒋桂雄^{1,2}, 龙洪旭^{1,2}, 谭晓风^{1,2}

1. 中南林业科技大学, 经济林培育与保护省部共建教育部重点实验室, 长沙 410004;
2. 中南林业科技大学, 经济林育种与栽培国家林业局重点实验室, 长沙 410004

摘要: 油桐是我国重要的木本油料植物, 过去对油桐的研究主要集中于栽培和常规育种, 与油桐种仁油脂合成相关的分子机理研究还未见报道。文章采用 RNA-seq 技术对油桐种子油脂合成的 3 个不同时期的转录组进行比较, 获得了大量差异表达的 Unigene 序列。在此基础上, 通过 GO 分类和 Pathway 富集性分析将这些差异表达 Unigene 归类于 128 个代谢途径, 其中包含与油脂合成相关的脂肪酸生物合成和甘油磷脂代谢途径。桐酸经脂肪酸生物合成途径合成后通过甘油磷脂代谢途径以桐油的形式贮存。将这两个代谢途径的 Unigene 序列在 KEGG 数据库中进行比对, 获得了一些关键酶的同源蛋白质。文章通过对编码这些同源蛋白质的基因在油桐种子油脂合成期的表达模式进行分析, 以期对油桐油脂合成, 尤其是桐酸合成机理的解析提供理论参考, 并为油桐的遗传改良提供潜在的基因资源, 从而提高油桐的单位面积产量。

关键词: 油桐(*Vernicia fordii*); 油脂合成途径; RNA-seq; 转录组; 差异表达基因

Analysis of oil synthesis metabolism pathways based on transcriptome changes in tung oil tree's seeds during three different development stages

CHEN Hao^{1,2}, JIANG Gui-Xiong^{1,2}, LONG Hong-Xu^{1,2}, TAN Xiao-Feng^{1,2}

1. The Key Lab of Cultivation and Protection for Non-wood Forest Trees of Education Ministry, Central South University of Forestry and Technology, Changsha 410004, China;
2. The Key Lab of Non-wood Forest Products of State Forestry Administration, Central South University of Forestry and Technology, Changsha 410004, China

Abstract: Tung oil tree (*Vernicia fordii*) is one of the important woody oil plants in China. Past researches on tung oil tree mainly focus on the cultivation and conventional breeding while the molecular mechanisms related to tung oil synthesis are still uncovered. We compared transcriptome of tung oil tree's seeds at three different oil synthesis stages using RNA-seq technology and then obtained a lot of differentially expressed Unigenes. Through GO classification and pathway enrichment

收稿日期: 2013-06-27; 修回日期: 2013-09-07

基金项目: 国家林业公益性行业科研专项重大项目(编号: 201204403)资助

作者简介: 陈昊, 博士, 讲师, 研究方向: 林业生物技术。Tel: 0731-85623406; E-mail: chen hao198223@sina.com

通讯作者: 谭晓风, 博士, 教授, 研究方向: 林业生物技术。E-mail: tanxiaofengcn@126.com

网络出版时间: 2013-10-16 19:05:08

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20131016.1905.002.html>

analysis, all of these differentially expressed Unigenes were classified into 128 metabolism pathways including fatty acid biosynthesis and glycerophospholipid metabolism which are involved in oil synthesis. Some homologous proteins of key enzymes were obtained when the sequences of the Unigenes within these two pathways were aligned against KEGG database. Through analysis of expression profiles of these key enzyme genes during seed's oil synthesis stage, this research not only shed light on elucidation of plant oil synthesis but also provides candidate genes for genetic improvement of tung oil tree thereby increasing the yield per unit area of tung oil tree.

Keywords: Tung oil tree; oil synthesis pathway; RNA-seq; transcriptome; differentially expressed genes

油桐(*Vernicia fordii*)原产于中国,是世界著名的工业油料树种,也是我国四大木本油料植物(油桐、油茶、核桃、乌桕)之一。油桐作为一种工业油料树种具有栽培历史悠久、分布范围广、生长迅速等优势^[1]。油桐种子制取的桐油是一种重要的工业用油,是制造优质油漆和油墨的基本原料,还是制造生物柴油和优质复合功能新材料的优质原料,具有很高的经济利用价值^[1]。利用色谱分析发现桐油中至少具有 63 种脂肪酸成份,其中油桐特有的桐酸含量高达 72%,桐酸含量的高低决定了桐油的品质及其在工业上的用途^[2]。油脂是 3-磷酸甘油和 3 个脂肪酰-CoA 脱水缩合而成的化合物的总称。植物种子中油脂的合成主要可分为脂肪酸的合成以及脂肪酸与 3-磷酸甘油脱水缩合生成甘油磷脂两个阶段。但在生物体细胞内,游离的脂肪酸含量并不多,绝大多数脂肪酸以甘油磷脂的酰基形式存在。因此,解析脂肪酸合成与甘油磷脂代谢途径将是阐述油桐油脂合成机理的先决条件之一。

植物脂肪酸合成途径属于已知两类脂肪酸合成途径中的第二类(FAS)^[3]。植物脂肪酸合成主要在质体中进行^[4,5]。在种子的发育过程中,合成脂肪酸的主要碳源蔗糖从光合器官转运到种子细胞中,通过糖酵解途径转变为己糖,随后氧化成脂肪酸合成的前体物质乙酰辅酶A(CoA)^[6]。乙酰辅酶A在乙酰辅酶A羧化酶(Acetyl-CoA carboxylase, ACCase)的催化下,生成丙二酰单酰辅酶A^[7]。丙二酰单酰辅酶A在脂肪酸合酶复合体(Fatty acid synthase complex, FAS)的催化下进行连续聚合反应,每次循环使碳链延长 2 个碳原子^[8-12],最后脂肪酰-ACP硫酯酶(Fatty acyl-ACP thioesterase)催化酰基碳链从 ACP 上解离并从脂肪酸合酶复合体中释放,生成游离脂肪酸^[13,14]。不同碳链长度的脂肪酸在酰基辅酶A 合成酶(Acyl-

CoA synthetase)的作用下合成酰基辅酶A^[15]。同时,糖酵解的中间产物磷酸二羟丙酮在细胞质中被 3-磷酸甘油脱氢酶还原,生成 3-磷酸甘油^[15]。随后,酰基辅酶A在内质网上经 3-磷酸甘油酰基转移酶(Glycerol-3-phosphateacyltransferase, GPAT)、溶血性磷脂酰基转移酶(Lyso-phosphatidic acid acyltransferase, LPAAT)、二酰甘油转酰酶(Diacylglycerolacyltransferase, DGAT)以及磷脂酰基磷酸水解酶(Phosphatidatephosphohydrolase, PAPase)的作用,将酰基转移到 3-磷酸甘油骨架上,最终生成甘油磷脂和三脂酰甘油^[16]。由此可见,植物油脂的合成是一个多种酶类共同参与的复杂过程,任何一种酶类的表达调控异常都将影响油脂的合成,这也是油脂含量和组分表现出复杂多样性的分子遗传基础。因此,从整体上研究油脂代谢相关基因的表达变化,有望更全面地阐明植物油脂的代谢网络。

油桐种子发育时大量合成桐油,因此以油桐为材料研究其油脂的代谢不仅能阐述木本植物的油脂代谢机理,同时有望在此基础上通过生物技术的手段对油桐进行遗传改良,提高桐油的产量,改善桐油的品质。然而,相对于杨树等基础研究开展较早的林木而言,目前对油桐的研究主要集中于栽培和常规育种方面,相关的分子生物学研究不够深入。已有的分子生物学研究仅局限于一些同源基因的克隆,还没有涉及到基因的功能鉴定和参与的信号网络分析^[17-21]。造成这一现象的主要原因是缺少有用的参考信息,而新一代测序技术的发展则为解决这一难题提供了必要的技术手段^[22-25]。基于新一代测序技术的转录组测序(RNA-seq)与传统的基因芯片相比具有高通量、低成本、高灵敏度、重复性好且无需已知参考序列等优点,已逐步取代基因芯片成为转录组研究的主要手段^[26-30]。综上所述,本研究

采用RNA-seq技术对处于油脂大量合成期的油桐种子进行转录组测序,并在此基础上获得了与油脂代谢相关基因的表达变化规律,这一研究结果将为油桐乃至植物油脂代谢机理的阐述提供理论参考。此外,将本研究结果应用于油桐的遗传改良,有望提高桐油的单位面积产量,改善桐油的品质。

1 材料和方法

1.1 研究材料

因桐油存在于油桐的种子中,本研究以油桐代表性地方品种——泸溪葡萄桐处于 3 个不同发育阶段(6 月、8 月、10 月)的种子为研究材料,分别提取 3 个时期种子的总 RNA 后对这 3 个时期的种子转录组进行 RNA-seq 分析。

1.2 总 RNA 的提取与测序文库的构建

总 RNA 的提取使用 Qiagen 公司的 RNeasy Plant Mini Kit 试剂盒进行。提取的总 RNA 经 Agilent 2100 Bioanalyzer 检测,达到了进行 RNA-seq 实验的标准($OD_{260/280}$ 为 1.8~2.2, 28S:18S >1.0, 浓度 400 ng/ μ L, 总量 20 μ g, RIN(RNA integrity number)值 8.0)。第一链 cDNA 的合成使用 Invitrogen 公司的 SuperScriptTM RT 试剂盒进行,所用引物为 hexamer 随机引物。以置换合成法合成第二链 cDNA,即首先利用 RNase H 在杂合双链的 mRNA 链上造成缺刻或缺口,再以 RNA 片段为引物使用 Invitrogen 公司的 *E. coli* DNA polymerase 合成第二链 cDNA 取代模板上的 mRNA,这些 DNA 片段经 T4 DNA ligase 连接后形成完整的第二链 cDNA。cDNA 双链合成后按照 Illumina 公司的操作手册构建双末端测序文库,测序平台为 HiSeq 2000。

1.3 测序数据分析

1.3.1 原始数据的处理

测序得到的图像数据经碱基读取(Base calling)后转换成由众多称为 reads 的短核酸序列组成的原始序列数据(Raw reads)。Raw reads 经过去除含接头的 reads,去除 N 的比例大于 5%的 reads 及去除低质量 reads(质量值 Q < 10 的碱基数占整个 read 的 20%以上)的处理后获得高质量序列数据(Clean reads)。后续分析都基于 clean reads 进行。

1.3.2 序列的组装

因为油桐没有可参照的基因组,因此采用 reads 组装软件 Trinity^[31]进行转录组测序数据的从头组装。首先将具有重叠区域的 reads 组装成长片段(Contig)后将这些 contig 连在一起组装成两端不能再延长的序列(Unigene)。将 Unigene 序列与 NCBI 非冗余(Nun-redundant, nr)蛋白质数据库(Nr)、Swiss-Prot (Swiss-Protein)蛋白质数据库、KEGG (Kyoto Encyclopedia of Genes and Genomes)数据库和 COG (Cluster of Orthologous Groups of Proteins)数据库进行比对以确定 Unigene 序列的方向。

1.3.3 Unigene 的功能注释、GO 分类和 Pathway 富集性分析

将 Unigene 序列在蛋白质数据库(Nr、Swiss-Prot、KEGG 和 COG)中进行序列比对,获得该 Unigene 序列对应的蛋白质注释信息。利用 Blast2GO^[32]软件进行 Unigene 的 GO (Gene Ontology)功能注释和分类。利用 KEGG 数据库进一步获得 Unigene 的 Pathway 注释信息。

1.3.4 差异表达 Unigene 的筛选、GO 分类和 Pathway 分析

使用 FPKM 法(Fragments per Kb per Million Fragments)^[33]计算 Unigene 的表达量,找出在不同样本间差异表达的 Unigene^[34]。不同样本比较后 FDR(False discovery rate)值 0.001 且表达量的 log2 转化值 $|\log_2 \text{Ratio}| \geq 1$ 的基因定义为差异表达基因。差异表达 Unigene 的 GO 分类和 Pathway 分析同 1.3.3 所述。

2 结果与分析

2.1 测序数据统计

6 月、8 月和 10 月(以下为叙述方便,这 3 个时期分别称为 期、 期和 期)的油桐种子 RNA 在本次 RNA 测序中共得到 58439 条非冗余的 Unigene 序列。从 期、 期和 期的种子 RNA 中分别获得了 61001、54679 和 44495 条 Unigene 序列。非冗余 Unigene 序列主要长度分布范围为 200~3000 nt,大于 3000 nt 的序列有 1749 条,占总序列数的 0.03%(图 1)。随着序列长度的增加,Unigene 数呈逐级递减的

趋势, 没有明显的间断, 这表明 RNA 测序的质量较高(图 1)。非冗余 Unigene 通过 GO 富集性分析可分为分子功能(Molecular function)、细胞组分(Cellular component)和生物过程(Biological process)三大类, 其中催化活性(Catalytic activity)和代谢过程(Metabolic process)具有很高的富集显著性(图 2), 这可能与种子发育时代谢旺盛有关。非冗余 Unigene 的 Pathway 富集性分析将所有 Unigene 分别归于 128 个

信号途径, 其中代谢途径(Metabolic pathways)的 Unigene 数最多, 有 4736 个, 占有非冗余 Unigene 数的 21.4%, 甜菜红碱生物合成(Betalain biosynthesis)途径的 Unigene 数最少, 只有 1 个。对油桐种子 3 个发育时期 Unigene 的表达量进行两两比较后发现, 随着种子的发育, 越来越多的 Unigene 基因表达量呈上调趋势, 而下调基因数远低于上调基因数且其数量变化无明显规律(图 3)。

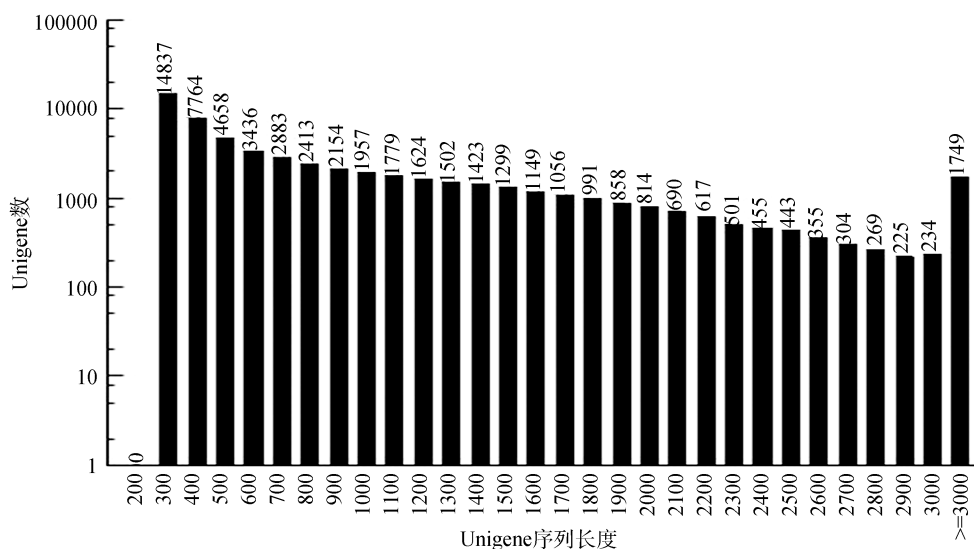


图 1 Unigene 序列的长度分布

横坐标为序列的长度, 纵坐标为具有相同序列长度的 Unigene 数。

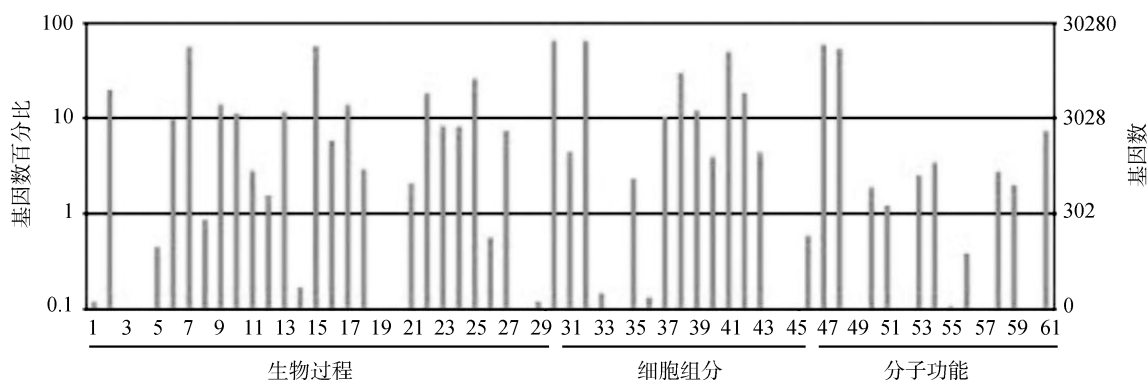


图 2 非冗余 Unigene 的 GO 分类

1: 生物黏附; 2: 生物调控; 3: 碳利用; 4: 细胞杀伤; 5: 细胞增殖; 6: 细胞组分组织或生物发生; 7: 细胞过程; 8: 死亡; 9: 发育过程; 10: 定位的建立; 11: 生长; 12: 免疫系统过程; 13: 定位; 14: 运动; 15: 代谢过程; 16: 多有机体过程; 17: 多细胞生物过程; 18: 生物过程负调控; 19: 氮利用; 20: 色素沉着; 21: 生物过程正调控; 22: 生物过程调控; 23: 繁殖; 24: 繁殖过程; 25: 刺激响应; 26: 节律过程; 27: 信号途径; 28: 硫利用; 29: 病毒繁殖; 30: 细胞; 31: 细胞连接; 32: 细胞部分; 33: 细胞外基质; 34: 细胞外基质部分; 35: 细胞外基质区域; 36: 细胞外区域部分; 37: 大分子复合物; 38: 膜; 39: 膜部分; 40: 膜围绕的腔; 41: 细胞器; 42: 细胞器部分; 43: 共质体; 44: 病毒颗粒; 45: 病毒颗粒部分; 46: 抗氧化活性; 47: 结合; 48: 催化活性; 49: 通道调控因子活性; 50: 电子载体活性; 51: 酶调控因子活性; 52: 金属分子伴侣活性; 53: 分子传感器活性; 54: 核酸结合转录因子活性; 55: 营养库活性; 56: 蛋白结合转录因子活性; 57: 蛋白标签; 58: 受体活性; 59: 结构分子活性; 60: 翻译调控因子活性; 61: 转运蛋白活性。

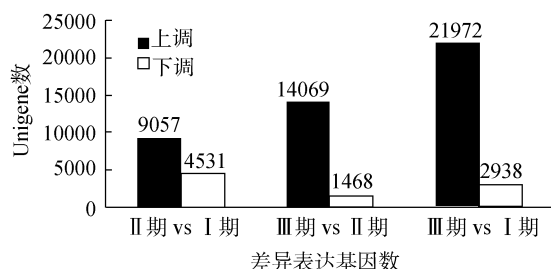


图 3 油桐种子 3 个不同发育时期差异表达基因数
期、期和期分别表示 6 月、8 月和 10 月的油桐果实发育期。

2.2 油桐种子中脂肪酸合成途径关键基因表达模式分析

对 3 个时期的转录组数据进行 Pathway 富集性分析发现共有 54 个 Unigene 可归类于脂肪酸生物合成(Fatty acid biosynthesis)途径, 占有非冗余 Unigene 的 0.24%。脂肪酸合成是油脂合成的基础, 因此对脂肪酸合成途径基因的研究有助于更深入的认识油脂合成的机理。对每组两两比较获得的差异表达 Unigene 分别进行 Pathway 富集性分析后发现, 每组差异表达 Unigene 中都有一些被归类于脂肪酸生物合成途径(图 4)。将 54 个 Unigene 序列在 KEGG 数据库中比对后获得 14 个与其他物种同源的脂肪酸合成相关基因并对这些基因在油桐油脂合成期的表达模式进行了分析(表 1)。

乙酰辅酶A羧化酶(Acetyl-CoA carboxylase)催化乙酰辅酶A转化成丙二酸单酰辅酶A这一脂肪酸生物合成的起始反应^[7]。高等植物体内存在两种类型的乙酰辅酶A羧化酶——同质型和异质型乙酰辅酶A羧化酶。在双子叶植物如油桐中, 乙酰辅酶A的羧化反应由位于质体中的异质型乙酰辅酶A羧化酶催化^[35]。异质型乙酰辅酶A羧化酶由 4 个亚基组成, 分别为生物素羧基载体蛋白(Biotin carboxyl carrier protein, BCCP)、生物素羧化酶(Biotin carboxylase, BC)、羧基转移酶(Carboxyl transferase, CT)的 α 亚基(α -CT)和 β 亚基(β -CT)。BCCP基因(*accB*)从期到期持续上调表达, BC基因(*ACAC*)和 β -CT基因(*accD*)则持续下调, α -CT基因(*accA*)先上调再下调且在期的表达量低于期(表 1)。酰基载体蛋白-丙二酰转移酶(ACP-malonyltransferase)催化丙二酸单酰辅酶A的丙二酸转移到酰基载体蛋白(Acyl-carrier-protein, ACP)上形成丙二酸单酰ACP, 从而为后续的缩合反应提供底物^[8]。酰基载体蛋白-

丙二酰转移酶基因的表达量在期时上调, 在期时的表达量与期相比无明显变化(表 1)。植物体内的脂肪酸代谢途径为生物界到目前为止已知两类脂肪酸代谢途径中的第二类(FAS)^[3]。这一途径中存在 3 种 3-氧酰基-酰基载体蛋白合酶(Ketoacyl-ACP synthase), 分别为KAS、KAS和KAS。KAS催化乙酰ACP与丙二酸单酰ACP的起始缩合反应, 而KAS和KAS则催化之后酰基链延长的缩合反应^[9]。RNA-seq检测到在种子发育 3 个时期表达量有明显变化的为型和型。KAS基因在期的表达量明显上调, 而在期的表达量与期相比无明显变化。KAS基因在期上调表达, 之后表达量下降, 在期时表达量回复到期的水平(表 1)。3-氧酰基-酰基载体蛋白还原酶(3-oxoacyl-[acyl-carrier protein] reductase)在NADPH存在的条件下催化底物生成 β -羟烷基ACP(β -hydroxyacyl-ACP)^[10], 其基因表达量在期明显上调, 之后逐渐下降, 但直到期时仍高于期的表达量(表 1)。酰基-ACP去饱和酶(Acyl-[acyl-carrier-protein] desaturase)催化羟丁酰ACP脱去 1 分子水生成丁烯酰ACP^[11]。酰基-ACP去饱和酶基因的表达量在种子发育的 3 个时期内持续上调表达(表 1)。烯酰-ACP还原酶(Enoyl-[acyl-carrier protein] reductase)催化脂肪酸生物合成第一个循环的最后一步, 也是总反应中的第二次还原反应, 其在NADPH的参与下催化丁烯酰ACP生成丁酰ACP^[12]。烯酰-ACP还原酶基因的表达量在期上调, 而期时的表达量与期相比无明显变化(表 1)。脂肪酸的生物合成途径在完成一系列的循环反应使酰基碳链达到特定长度后, 脂肪酰-ACP硫酯酶(Fatty acyl-ACP thioesterase)催化酰基碳链从ACP上解离并从脂肪酸合酶(Fatty acid synthase)复合体中释放, 因此脂肪酰-ACP硫酯酶的活性决定了脂肪酸碳链的长度^[13,14]。脂肪酰-ACP硫酯酶基因在期时的表达量与期相比明显上调并保持这一表达量至期(表 1)。

2.3 油桐种子中甘油磷脂代谢途径关键基因表达模式分析

对 3 个时期的转录组数据进行 Pathway 富集性分析发现共有 428 个 Unigene 可归类于甘油磷脂代谢(Glycerophospholipid metabolism)途径, 占有非

表 1 油桐种子脂肪酸生物合成途径关键酶基因在油脂合成期的表达模式

KEGG 编号	EC 编号	基因名	蛋白名	基因表达模式		
				vs	vs	vs
K00059	1.1.1.100	<i>fabG</i>	3-oxoacyl-ACP reductase	上调	下调	上调
K10782	3.1.2.14	<i>FATA</i>	Fatty acyl-ACP thioesterase A	上调	—	上调
K11262	6.3.4.14	<i>ACAC</i>	Acetyl-CoA carboxylase / biotin carboxylase	下调	下调	下调
K01963	6.4.1.2	<i>accD</i>	Acetyl-CoA carboxylase carboxyl transferase subunit beta	下调	—	下调
K00208	1.3.1.9	<i>fabI</i>	Enoyl-ACP reductase	上调	—	上调
K01961	6.4.1.2	<i>accC</i>	Acetyl-CoA carboxylase, biotin carboxylase subunit	上调	—	上调
K01962	6.4.1.2	<i>accA</i>	Acetyl-CoA carboxylase carboxyl transferase subunit alpha	上调	下调	下调
K01716	4.2.1.59	<i>fabA</i>	β-hydroxyacyl-ACP dehydratase	上调	上调	上调
K00645	2.3.1.39	<i>fabD</i>	ACP S-malonyltransferase	上调	—	上调
K09458	2.3.1.179	<i>fabF</i>	β-Ketoacyl-ACP synthase	上调	—	上调
K02160	NE	<i>accB</i>	Acetyl-CoA carboxylase biotin carboxyl carrier protein	上调	上调	上调
K02372	4.2.1.59	<i>fabZ</i>	β-hydroxyacyl-ACP dehydratases	上调	下调	—
K00648	2.3.1.180	<i>fabH</i>	β-Ketoacyl-ACP synthase	上调	上调	上调
K10781	3.1.2.14	<i>FATB</i>	Fatty acyl-ACP thioesterase B	—	上调	上调

注： 、 、 分别代表油桐种子油脂合成的 3 个不同时期；符号“--”表示基因表达水平无显著差异；NE 为 not enzyme 的缩写。

冗余 Unigene 的 1.93%。甘油磷脂是生物体油脂的主要成分之一，甘油磷脂的代谢通过甘油磷脂代谢途径进行，因此对甘油磷脂代谢途径基因的研究有助于阐述甘油磷脂代谢的分子机理。对每组两两比较获得的差异表达 Unigene 分别进行 Pathway 富集性分析发现，每组差异表达 Unigene 中都有一些被归类于甘油磷脂代谢途径(图 5)。将 428 个 Unigene 序列在 KEGG 数据库中比对后获得 31 个与其他物种同源的甘油磷脂代谢相关基因并对这些基因在油桐油脂合成期的表达模式进行了分析(表 2)。

甘油磷脂合成代谢中两个关键酶分别是甘油 3 磷酸酰基转移酶(Glycerol-3-phosphate acyltransferase, GPAT)和磷脂酸胞苷转移酶(Phosphatidate-cytidylyltransferase, CDS1)。GPAT催化酰基辅酶A的酰基转移到甘油 3 磷酸的sn-1 位，酯酰化甘油 3 磷酸从而产生合成膜脂与贮存脂质时的重要中间产物溶血磷脂酸(Lysophosphatidic acid, LPA)^[36]。植物体内有 3 种不同类型的甘油 3 磷酸酰基转移酶，根据其催化反应发生的不同位置，这些酶分别存在于质体、内质网和线粒体中^[37]。RNA-seq分析发现，甘油 3 磷酸酰基转移酶基因(*GPAT*)的表达量在种子油脂合成期持续上调表达，与油脂快速合成呈正相关的关系(表 2)。CDS1 也称 CDP-二脂酰甘油合酶(CDP-diacylglycerol synthase)，它催化底物胞苷三磷

酸(CTP)和二酰基磷脂酸生成 CDP-二脂酰甘油(CDP-diacylglycerol)^[38,39]。CDP-二脂酰甘油是甘油磷脂合成途径的关键中间产物，在它的基础上通过不同的后续步骤可合成磷脂酰丝氨酸、磷脂酰甘油及磷脂酰乙醇胺等各类型甘油磷脂。磷脂酸胞苷转移酶基因(*CDS1*)的表达量在 期和 期时无明显变化，之后则开始上调表达(表 2)。磷脂分解代谢的关键酶是磷脂酶(Phospholipase, PL)，根据裂解酯键位置的不同可将其分为磷脂酶A₁、A₂、B、C和D 5 大类(PLA₁、PLA₂、PLB、PLC和PLD)，它们在生物体内行使细胞膜结构的维护与修复、细胞信号转导及磷脂的消化等功能^[40]。PLA₁基因(*pldA*)和PLA₂基因(*TGL4*)在种子油脂合成的 3 个时期持续下调表达(表 2)。PLB基因(*PLB*)在种子油脂快速合成期不表达或少量表达以至于RNA-Seq检测无信号(表 2)。PLC基因(*plcC*)在种子油脂合成的 3 个时期也是持续下调表达(表 2)。PLD基因(*PLD*)在 期时上调表达，随后表达量下降，在 期时的表达量低于 期(表 2)。

3 讨 论

到目前为止，油桐的分子生物学研究基础薄弱，能够参考的背景研究很少，这严重制约了对油桐各生物学性状分子机理的研究。利用基于第二代测序

表 2 油桐种子甘油磷脂代谢途径关键酶基因在油脂合成期的表达模式

KEGG 编号	EC 编号	基因名称	蛋白质定义	基因表达模式		
				VS	VS	VS
K13511	2.3.1	<i>TAZ</i>	Monolysocardiolipinacyltransferase	下调	下调	下调
K00901	2.7.1.107	<i>dgkA</i>	Diacylglycerol kinase	下调	下调	下调
K01115	3.1.4.4	<i>PLD</i>	Phospholipase D	上调	下调	下调
K00981	2.7.7.41	<i>CDS1</i>	Phosphatidatecytidyltransferase	—	上调	上调
K01058	3.1.1.32	<i>pldA</i>	Phospholipase A1	下调	下调	下调
K08730	2.7.8.29	<i>PTDSS2</i>	Phosphatidylserine synthase 2	上调	—	上调
K13510	2.3.1.23	<i>LPCAT1_2</i>	Lysophosphatidylcholineacyltransferase	—	下调	下调
K05929	2.1.1.103	<i>NMT</i>	Phosphoethanolamine N-methyltransferase	下调	下调	下调
K00006	1.1.1.8	<i>GPD1</i>	Glycerol-3-phosphate dehydrogenase	上调	上调	上调
K00630	2.3.1.15	<i>ATS1</i>	Glycerol-3-phosphate O-acyltransferase	—	上调	上调
K06129	3.1.1.5	<i>LYPLA3</i>	Lysophospholipase	下调	下调	下调
K01126	3.1.4.46	<i>glpQ</i>	Glycerophosphoryldiesterphosphodiesterase	—	下调	下调
K01613	4.1.1.65	<i>PISD</i>	Phosphatidylserine decarboxylase	上调	上调	上调
K14156	2.7.1.32	<i>CHK</i>	Choline/ethanolamine kinase	上调	下调	上调
K00967	2.7.7.14	<i>PCYT2</i>	Ethanolamine-phosphate cytidyltransferase	—	下调	下调
K01047	3.1.1.4	<i>PLA2G</i>	Secretory phospholipase A2	上调	下调	下调
K08744	2.7.8	<i>CRLS</i>	Cardiolipin synthase	—	—	—
K01114	3.1.4.3	<i>plcC</i>	Phospholipase C	下调	下调	下调
K00655	2.3.1.51	<i>plsC</i>	1-acyl-sn-glycerol-3-phosphate acyltransferase	—	上调	上调
K00995	2.7.8.5	<i>pgsA</i>	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	—	下调	下调
K13535	3.1.1	<i>CLD1</i>	Cardiolipin-specific phospholipase	下调	下调	下调
K13513	2.3.1	<i>LCLAT1</i>	Lysocardiolipinacyltransferase	—	下调	下调
K13523	2.3.1.51	<i>AGPAT3_4</i>	Lysophosphatidic acid acyltransferase	上调	—	上调
K13508	2.3.1.15	<i>GPAT</i>	Glycerol-3-phosphate acyltransferase	上调	上调	上调
K00111	1.1.5.3	<i>glpA</i>	Glycerol-3-phosphate dehydrogenase	—	下调	下调
K00993	2.7.8.1	<i>EPT1</i>	Ethanolaminephosphotransferase	—	下调	下调
K00968	2.7.7.15	<i>PCYT1</i>	Choline-phosphate cytidyltransferase	—	—	—
K14674	3.1.1.4	<i>TGL4</i>	Phospholipase A2	下调	下调	下调
K13519	2.3.1.51	<i>LPT1</i>	Lysophospholipidacyltransferase	上调	下调	下调
K06130	3.1.1.5	<i>LYPLA2</i>	Lysophospholipase	—	下调	下调
K00894	2.7.1.82	<i>ETN</i>	Ethanolamine kinase	—	下调	下调

技术的RNA-seq进行转录组研究时不需要知道物种的遗传背景且灵敏度高,可以检测到低至几个拷贝的转录本,因此在油桐转录组研究中相对于其它常用的转录组研究方法(cDNA芯片、LongSAGE、MPSS)具有明显的优越性^[22-25]。本研究采用RNA-seq转录组研究技术对油桐种子油脂合成期的转录组进行了比较研究,将油桐种子油脂合成期差异表达的基因进行了GO和Pathway信号途径分类,初步明确了各个基因编码的蛋白质的功能,这为进一步研究油桐

种子油脂合成的分子机理提供了必要的数据平台,为研究基因表达变化和油脂积累之间的关系奠定了研究基础。

油桐的脂肪酸生物合成途径中,乙酰辅酶A羧化酶 4 个亚基中的生物素羧基载体蛋白基因在油桐种子发育过程中持续上调表达,而生物素羧化酶和转羧酶β亚基基因则持续下调表达,转羧酶α亚基基因先上调再下调表达且在 期时的表达量低于 期,这一结果暗示乙酰辅酶A羧化酶的 4 个亚基在油桐

体内并不是等量存在的,复合酶的活性随着亚基之间比例的改变而改变。脂肪酸合酶的非催化亚基酰基载体蛋白的表达量无明显变化,这可能是由于脂肪酸大量合成的负反馈调节导致ACP蛋白的积累从而抑制了ACP基因的过量表达^[41]。脂肪酸合酶其他催化亚基的表达量在 期都明显上调,在 期时的表达量与 期相比有的持续上调,有的保持不变,有的下调,但所有亚基基因在 期时的表达量都不低于 期,这表明这些基因的表达量与种子发育过程中油脂快速合成和积累这一生物学现象是正相关的。关于这些亚基在油桐种子油脂合成过程中的表达模式到目前为止还没有报道,它们行使功能的分子机理还有待进一步研究。

RNA-seq检测到的与甘油磷脂代谢相关的重要酶基因中,与分解代谢相关的基因,尤其是磷脂酶类基因如*PLA2G*、*PLD*、*pldA*、*plcC*等在油脂合成期下调表达,与合成代谢相关的基因如*GPAT*、*CDS1*、*PTDSS2*等则上调表达,这在整体上符合油脂快速合成时的基因表达规律。桐油在 6 月(期)开始缓慢合成,而在 8 月(期)开始快速积累^[42,43],根据本研究的数据分析,这一表型很可能与*CDS1*的表达模式相关。*CDS1*是甘油磷脂合成的限速酶之一,起着“总开关”的作用,转录组数据表明其基因在 期和 期之间的表达量无明显变化,这可能导致其下游合成相关基因的上调表达由于得不到来自*CDS1*的底物供给而作用有限。值得注意的是,除了与分解代谢相关的磷脂酶类基因下调表达外,多个酰基转移酶基因也呈现下调的趋势,这使得转录组测序检测到的甘油磷脂代谢途径基因中下调基因数多于上调基因数。例如在小鼠细胞中,溶血磷脂酰胆碱酰基转移酶(LPCAT, EC 2.3.1.23)催化酰基转移到磷脂酰胆碱的sn-2 位上,合成二酰基磷脂酰胆碱^[44],但在油桐中,序列同源的溶血磷脂酰胆碱酰基转移酶基因(*LPCAT1_2*)的表达与油脂积累负相关(表 2)。这一现象可能暗示了植物尤其是木本油料植物中,一些酰基转移酶基因演化出了哺乳动物中的同源基因所不具备的功能,对这些基因更深入的研究有望揭示它们在植物体内独特的作用机理。

桐油具有很高的经济利用价值,研究油桐种子中油脂转化和积累的分子机理有助于对油桐进行定向的遗传改良,同时也为植物油脂合成机理的解析

提供理论参考。本研究富集得到与油桐油脂合成相关的关键酶基因并初步分析了它们在油桐种子油脂合成期的表达量变化规律,研究结果在为油桐乃至植物油脂合成的机理研究提供大量参考数据的同时,也为油桐的遗传改良提供了潜在的基因资源和理论指导。

参考文献(References):

- [1] 谭晓风, 蒋桂雄, 谭方友, 周伟国, 吕平会, 罗克明, 孙汉洲, 王承南, 马锦林, 何佳林, 梁文汇, 黄艳. 我国油桐产业化发展战略调查研究报告. 经济林研究, 2011, 29(3): 1-7. [\[DOI\]](#)
- [2] 刘金龙, 郑小江, 郑威, 田国政, 杨驰. 油桐品种五爪桐含油量及桐油质量研究. 湖北农业科学, 2011, 50(10): 2031-2035. [\[DOI\]](#)
- [3] Rock CO, Cronan JE. *Escherichia coli* as a model for the regulation of dissociable (type) fatty acid biosynthesis. *Biochim Biophys Acta*, 1996, 1302(1): 1-16. [\[DOI\]](#)
- [4] Ohlrogge J, Browse J. Lipid biosynthesis. *Plant Cell*, 1995, 7(7): 957-970. [\[DOI\]](#)
- [5] Shanklin J, Cahoon EB. Desaturation and related modifications of fatty acids. *Annu Rev Plant Physiol Plant Mol Biol*, 1998, 49: 611-641. [\[DOI\]](#)
- [6] Nikolau BJ, Ohlrogge JB, Wurtele ES. Plant biotin-containing carboxylases. *Arch Biochem Biophys*, 2003, 414(2): 211-222. [\[DOI\]](#)
- [7] Konishi T, Shinohara K, Yamada K, Sasaki Y. Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant Cell Physiol*, 1996, 37(2): 117-122. [\[DOI\]](#)
- [8] Ruch FE, Vagelos PR. The isolation and general properties of *Escherichia coli* malonyl coenzyme A-acyl carrier protein transacylase. *J Biol Chem*, 1973, 248(23): 8086-8094. [\[DOI\]](#)
- [9] White SW, Zheng J, Zhang YM, Rock CO. The structural biology of type fatty acid biosynthesis. *Annu Rev Biochem*, 2005, 74(1): 791-831. [\[DOI\]](#)
- [10] Fisher M, Kroon JTM, Martindale W, Stuitje AR, Slabas AR, Rafferty JB. The X-ray structure of *Brassica napus* β -keto acyl carrier protein reductase and its implications for substrate binding and catalysis. *Structure*, 2000, 8(4): 339-347. [\[DOI\]](#)
- [11] Kimber MS, Martin F, Lu Y, Houston S, Vedadi M, Dharamsi A, Fiebig KM, Schmid M, Rock CO. The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase

- (FabZ) from *Pseudomonas aeruginosa*. *J Biol Chem*, 2004, 279(50): 52593–52602. [\[DOI\]](#)
- [12] Heath RJ, Rock CO. Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*. *J Biol Chem*, 1995, 270(44): 26538–26542. [\[DOI\]](#)
- [13] Voelker T. Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis. *Genet Eng*, 1996, 18: 111–133. [\[DOI\]](#)
- [14] Yuan L, Voelker TA, Hawkins DJ. Modification of the substrate specificity of an acyl-acyl carrier protein thioesterase by protein engineering. *Proc Natl Acad Sci USA*, 1995, 92(23): 10639–10643. [\[DOI\]](#)
- [15] 周奕华, 陈正华. 植物种子中脂肪酸代谢途径的遗传调控与基因工程. *植物学通报*, 1998, 15(5): 16–23. [\[DOI\]](#)
- [16] Jako C, Kumar A, Wei YD, Zou JT, Barton DL, Giblin EM, Covello PS, Taylor DC. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyl-transferase enhances seed oil content and seed weight. *Plant Physiol*, 2001, 126(2): 861–874. [\[DOI\]](#)
- [17] Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM. Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell*, 2006, 18(9): 2294–2313. [\[DOI\]](#)
- [18] 汪阳东, 李元, 李鹏. 油桐桐酸合成酶基因克隆和植物表达载体构建. *浙江林业科技*, 2007, 27(2): 1–5. [\[DOI\]](#)
- [19] 李建安, 孙颖, 陈鸿鹏, 刘丽娜, 郭文丹. 油桐 LEAFY 同源基因片段的克隆与分析. *中南林业科技大学学报*, 2008, 28(4): 21–26. [\[DOI\]](#)
- [20] 李元, 汪阳东, 李鹏, 魏建民. 油桐种子 FADX 基因的克隆和序列分析. *安徽农业科学*, 2008, 36(11): 4753–4755. [\[DOI\]](#)
- [21] 龙洪旭, 谭晓风, 张琳, 谢禄山, 陈洪. 油桐亲环素基因全长 cDNA 的克隆及序列分析. *中南林业科技大学学报*, 2011, 31(12): 111–117. [\[DOI\]](#)
- [22] Alagna F, D'Agostino N, Torchia L, Servili M, Rao R, Pietrella M, Giuliano G, Chiusano ML, Baldoni L, Perrotta G. Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *BMC Genomics*, 2009, 10(1): 399. [\[DOI\]](#)
- [23] Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. *Plant Physiol*, 2010, 152(4): 1787–1795. [\[DOI\]](#)
- [24] González-Ballester D, Casero D, Cokus S, Pellegrini M, Merchant SS, Grossman AR. RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. *Plant Cell*, 2010, 22(6): 2058–2084. [\[DOI\]](#)
- [25] Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang XW, Wu XM, Mitros T, Triplett J, Yang XH, Ye CY, Mauro-Herrera M, Wang L, Li PH, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM. Reference genome sequence of the model plant *Setaria*. *Nat Biotechnol*, 2012, 30(6): 555–561. [\[DOI\]](#)
- [26] O'Loughlin A, Lynn DJ, McGee M, Doyle S, McCabe M, Earley B. Transcriptomic analysis of the stress response to weaning at housing in bovine leukocytes using RNA-seq technology. *BMC Genomics*, 2012, 13: 250. [\[DOI\]](#)
- [27] Shao W, Zhao QY, Wang XY, Xu XY, Tang Q, Li M, Li X, Xu YZ. Alternative splicing and trans-splicing events revealed by analysis of the *Bombyx mori* transcriptome. *RNA*, 2012, 18(7): 1395–1407. [\[DOI\]](#)
- [28] Yang DH, Liu Q, Yang MJ, Wu HZ, Wang QY, Xiao JF, Zhang YX. RNA-seq liver transcriptome analysis reveals an activated MHC-I pathway and an inhibited MHC-pathway at the early stage of vaccine immunization in zebrafish. *BMC Genomics*, 2012, 13(1): 319. [\[DOI\]](#)
- [29] Torales SL, Rivarola M, Pomponio MF, Fernández P, Acuña CV, Marchelli P, Gonzalez S, Azpilicueta MM, Hopp HE, Gallo LA, Paniego NB, Marcucci Poltri SN. Transcriptome survey of Patagonian southern beech *Nothofagus nervosa* (= *N. alpina*): assembly, annotation and molecular marker discovery. *BMC Genomics*, 2012, 13: 291. [\[DOI\]](#)
- [30] Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*, 2008, 18(9): 1509–1517. [\[DOI\]](#)
- [31] Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*, 2011, 29(7): 644–652. [\[DOI\]](#)
- [32] Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation,

- visualization and analysis in functional genomics research. *Bioinformatics*, 2005, 21(18): 3674–3676. [\[DOI\]](#)
- [33] Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*, 2008, 5(7): 621–628. [\[DOI\]](#)
- [34] Audic S, Claverie JM. The significance of digital gene expression profiles. *Genome Res*, 1997, 7(10): 986–995. [\[DOI\]](#)
- [35] 王伏林, 吴关庭, 郎春秀, 陈锦清. 植物中的乙酰辅酶 A 羧化酶(AACase). *植物生理学通讯*, 2006, 42(1): 10–14. [\[DOI\]](#)
- [36] Zheng ZF, Xia Q, Dauk M, Shen WY, Selvaraj G, Zou JT. *Arabidopsis AtGPAT1*, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell*, 2003, 15(8): 1872–1887. [\[DOI\]](#)
- [37] Chen X, Snyder CL, Truksa M, Shah S, Weselake RJ. sn-Glycerol-3-phosphate acyltransferases in plants. *Plant Signal Behav*, 2011, 6(11): 1695–1699. [\[DOI\]](#)
- [38] Sparrow CP, Raetz CR. Purification and properties of the membrane-bound CDP-diglyceride synthetase from *Escherichia coli*. *J Biol Chem*, 1985, 260(22): 12084–12091. [\[DOI\]](#)
- [39] Shen HF, Heacock PN, Clancey CJ, Dowhan W. The *CDS1* gene encoding CDP-diacylglycerol synthase in *Saccharomyces cerevisiae* is essential for cell growth. *J Biol Chem*, 1996, 271(2): 789–795. [\[DOI\]](#)
- [40] Casado V, Martín D, Torres C, Reglero G. Phospholipases in food industry: a review. *Methods Mol Biol*, 2012, 861: 495–523. [\[DOI\]](#)
- [41] Andre C, Haslam RP, Shanklin J. Feedback regulation of plastidic acetyl-CoA carboxylase by 18: 1-acyl carrier protein in *Brassica napus*. *Proc Natl Acad Sci USA*, 2012, 109(25): 10107–10112. [\[DOI\]](#)
- [42] 蔺定运, 倪善庆. 油桐枝、叶和果实物质代谢的初步研究. *植物生理学通讯*, 1980, (3): 37–40. [\[DOI\]](#)
- [43] 王汉涛, 段聪仁, 徐树华, 曹明远. 油桐种仁与油脂形成规律的研究. *经济林研究*, 1985, 3(2): 29–35. [\[DOI\]](#)
- [44] Nakanishi H, Shindou H, Hishikawa D, Harayama T, Ogasawara R, Suwabe A, Taguchi R, Shimizu T. Cloning and characterization of mouse lung-type acyl-CoA: lyso-phosphatidylcholine acyltransferase 1 (LPCAT1). Expression in alveolar type cells and possible involvement in surfactant production. *J Biol Chem*, 2006, 281(29): 20140–20147. [\[DOI\]](#)