

文章编号: 1674-7054(2020)04-0426-11

## 海南油茶 2 个优良单株的比较转录组学分析

代佳妮<sup>1</sup>, 于靖<sup>1,2</sup>, 郑蔚<sup>1</sup>, 王健<sup>1,2</sup>, 吴友根<sup>1,2</sup>, 胡新文<sup>1,2</sup>, 赖杭桂<sup>1,2</sup>

(1. 海南大学园艺学院, 海口 570228; 2. 海南大学热带作物新品种选育教育部工程研究中心, 海口 570228)

**摘要:** 选取“万海 1 号”和“万海 4 号”2 个海南产油茶 (*Camellia vietnamensis*) 优良单株为试验材料, 采用新一代高通量 Illumina 转录组测序技术, 对二者进行转录组测序分析, 并进行差异比较。结果表明, “万海 1 号”的特异性基因中, 有 651 条被注释到与植株抗病性相关的途径, 约占 5.26%, 有 510 条被注释到脂肪酸代谢相关的途径, 约占 4.12%; “万海 4 号”中有 169 条特异性基因被注释到倍半萜和三萜生物合成、油菜素类固醇生物合成、类固醇生物合成和单萜类生物合成等植物次生代谢相关途径。2 个样本的单拷贝同源基因多被富集到与抗氧化能力相关的 GO 条目, 以及与氨基酸合成代谢相关的 KEGG 条目, 且 *P* 值均小于 0.05, 显示出显著的相关性。本试验可进一步探究海南本地油茶的生理特性及代谢过程, 作为海南本地油茶的优良品种选育的科学依据。

**关键词:** 海南油茶; 比较转录组; 基因表达

**中图分类号:** S 794.4

**文献标志码:** A

**DOI:** 10.15886/j.cnki.rdsxb.2020.04.005

油茶是山茶科 (Theaceae) 山茶属 (*Camellia*) 木本油料树种, 目前在国内主要的栽培种除了普通油茶 (*Camellia oleifera* Abel.), 还包括小果油茶 (*Camellia meocarpa* Hu)、越南油茶 (*Camellia vietnamensis* T. C. Huang ex Hu) 和博白大果油茶 (*Camellia gigantocarpa* Hu et Huang) 等。其分布区域主要集中在湖南、福建等南方省份, 由于油茶及其各类产品具有独特的食用、药用价值, 尤其是籽油中含有脂肪酸、植物甾醇、角鲨烯和维生素 E 等多种活性成分, 因此备受生产者及消费者的欢迎<sup>[1]</sup>。油茶与椰子 (*Cocos nucifera*)、橄榄 (*Olea europaea*)、油棕 (*Elaeis guineensis*) 并称为世界 4 大木本油料树种<sup>[2]</sup>。自 20 世纪 60 年代起, 我国通过油茶育种工作已选育出近 200 个优良无性系, 另外, 还选育出众多的农家良种、优良家系, 为油茶的高品质和高产量奠定了基础<sup>[3]</sup>。近年来, 海南省不少单位和种植户积极尝试, 引进外省品种种植于海南, 但由于海南省气候具有高温高湿的特点, 外来苗木在此不仅适应性和抗逆性差, 而且易受病虫害侵害, 导致成活率低, 幼苗即使成活也表现出长势不佳、长速缓慢的生长趋势<sup>[4]</sup>。目前, 海南省已初步选取了 20 个海南本地油茶优良单株作为应急品种在全省推广种植。因此, 比较和研究这些优良单株的遗传和生理特性, 特别是其抗病性和抗逆性, 早日选出适合海南种植的品质优良的油茶品种, 是海南油茶产业发展的当务之急。

转录组是指某个物种或者特定的细胞类型产生的所有转录本的集合<sup>[5]</sup>, 包括转录本的数量、特定发育阶段的表达动态、转录后的修饰以及非编码 RNA 的调控表达情况等<sup>[6]</sup>, 以转录组分析为基础, 能够更加全面深入地研究特定物种的基因表达、结构以及功能, 从而进一步了解特定的生理特性背后的分子机理<sup>[7]</sup>。转录组测序 (RNA-Seq) 是在深度测序的基础上进行转录组分析的技术, 2008 年首次面世, 随后在多个领域被广泛投入应用<sup>[8]</sup>, 该技术可用于确定基因的表达模式、发现新转录本、非编码 RNA 的调控机制和单细胞转录组等多个方面的研究<sup>[9]</sup>, 相比基因组, 它更具有时间性和空间性, 目前已应用于人参<sup>[10]</sup>、毛

收稿日期: 2020-05-18

修回日期: 2020-08-22

基金项目: 海南省自然科学基金高层次人才项目 (820RC585)

第一作者: 代佳妮 (1996-), 女, 海南大学园艺学院 2018 级硕士研究生. E-mail: [djn1212@163.com](mailto:djn1212@163.com)

通信作者: 于靖 (1985-), 女, 讲师. 研究方向: 南药的种质资源与开发利用. E-mail: [yujinghxy@163.com](mailto:yujinghxy@163.com)

地黄<sup>[11]</sup>、重楼<sup>[12]</sup>、罗勒<sup>[13]</sup>、沙棘<sup>[14]</sup>等多种药用植物的研究。转录组测序技术可为油茶的基因功能及生理特性研究提供新的手段。

目前,海南省认定的油茶优良单株所处的地理位置、树龄、水肥环境等有较大差异,其芽砧苗也尚未大规模推广种植并结实。因此,笔者选取了2个海南油茶优良单株为材料,采用新一代高通量 Illumina 转录组测序技术,分别对这2个优良单株的两年生芽砧苗的健康成熟叶片进行转录组分析,研究和比较2个优良单株的无性系基因表达的差异与共同点,为油茶的品种改良及优良品种选育奠定基础。

## 1 材料与方法

**1.1 实验材料** 以海南省林木品种审定委员会选取认定的“万海1号”和“万海4号”这2个海南油茶单株为试验材料,于2018年12月在海南大学儋州校区的油茶育苗圃进行样品采集,选择2个优良单株的两年生长势一致的健康芽砧苗各20株,每株随机采集新鲜成熟叶片,混合后迅速放入液氮保存,用于转录组测序。

**1.2 样品处理及转录组数据库分析流程** 从样品中提取总RNA后,转录组测序过程委托赛哲生物公司完成。用带有 Oligo (dT)的磁珠富集 mRNA,使用 fragmentation buffer 将 mRNA 分解成短片段,反转录后经 QiaQuick PCR 试剂盒纯化回收,建立 cDNA 文库。使用 Illumina HiSeqTMPE150 对构建的测序文库进行测序。使用 Trinity 对原始数据进行组装,随后注释到 NR、KOG、KEGG、Swiss-Prot 数据库,根据最佳比对结果获得蛋白质功能注释。使用 blastp 与 OrthoMCL 联用的方法来找出物种间同源的 Unigene,在得到物种间同源与特有基因家族后,将各类基因家族在两物种中分别包含的基因数(即基因家族拷贝数)分别统计,找出各物种特有的基因家族,分别对其进行功能注释、KEGG 富集分析、GO 富集分析。使用软件 MISA(<http://pgrc.ipk-gatersleben.de/misa/>)对各个物种的单拷贝同源基因 Unigene 进行搜索,进行 SSR 分析。

## 2 结果与分析

**2.1 转录组基本信息分析** 使用 Trinity 对2个油茶样本进行组装,得到其转录组参考序列。如表1所示,共对185 698条基因进行了测序。“万海1号”和“万海4号”2个样本的 N50 长度分别为 809 bp 和 912 bp,“万海1号”的最大基因长度为 21 213 bp,最小为 201 bp;“万海4号”的最大基因长度为 21 867 bp,最小为 201 bp。“万海1号”和“万海4号”的 GC 值均较高,分别为 41.66% 和 43.57%,表明测序数据质量良好。对2个样本中各类基因和基因家族的数目进行统计,结果(图1)表明,“万海1号”和“万海4号”同源的基因及基因家族共有 16 918 个,“万海1号”特有的基因有 92 705 个,“万海4号”特有的基因数则为 52 799 个。

表1 “万海1号”和“万海4号”转录组分析数据统计

Tab. 1 Transcriptome analysis of 2 clones derived from 2 elite individual plants of *Camellia vietnamensis*

样品 Sample	基因数/个 Genes	N50	GC/%GC percentage	最大长度/bp Maximum length	最小长度/bp Minimum length	CDS	蛋白数 Proteins
“万海1号”	113 846	809	41.66	21 213	201	52 549	52 549
“万海4号”	71 852	912	43.57	21 867	201	36 930	36 930

### 2.2 特有基因分析

**2.2.1 特有基因 KEGG 富集分析** 将2个样本的 Unigenes 注释到 KEGG 数据库进行比对,得到它们的特有基因 KEGG 富集分析气泡图,从而找出物种特有基因中显著性富集的 pathway。“万海4号”与“万海1号”油茶的比较结果(图2)表明,“万海4号”以核糖体途径的基因丰度最高,约为 600 条,其次是碳代谢和氧化磷酸化途径,三者的 P 值均小于 0.03。在“万海4号”所有的代谢途径中,富集因子最大的是单萜类生物合成途径,富集因子约为 0.79;其次是油菜素的生物合成途径,富集因子约为 0.73。值得注意

的是,在“万海 4 号”排名前 20 的富集途径中,出现了倍半萜和三萜生物合成、油菜素类固醇生物合成、类固醇生物合成和单萜类生物合成等与植物次生代谢相关的途径。“万海 1 号”的 KEGG 富集结果中,以植物病原体互作途径基因丰度最高,富集了 651 条基因,占总数的 5.26%,这可能与植物的抗病性相关;其次是 RNA 转运途径与内吞作用途径。在所有的代谢途径中,富集因子最大的是酮体的合成与降解途径,富集因子约为 0.90;其次是丙酸酯代谢和醚脂质代谢途径(图 3)。在万海 1 号的前 20 条 KEGG 途径中,出现了脂肪酸生物合成及脂肪酸代谢等与脂肪酸相关的途径,二者分别富集了 136 和 239 条基因,这可能与茶籽及茶油品质相关。

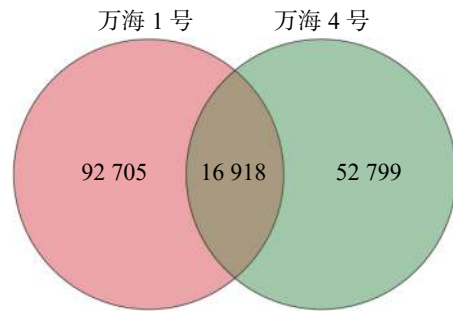


图 1 样品间基因家族 venn 图

Fig. 1 Venn diagram showing gene families between the samples

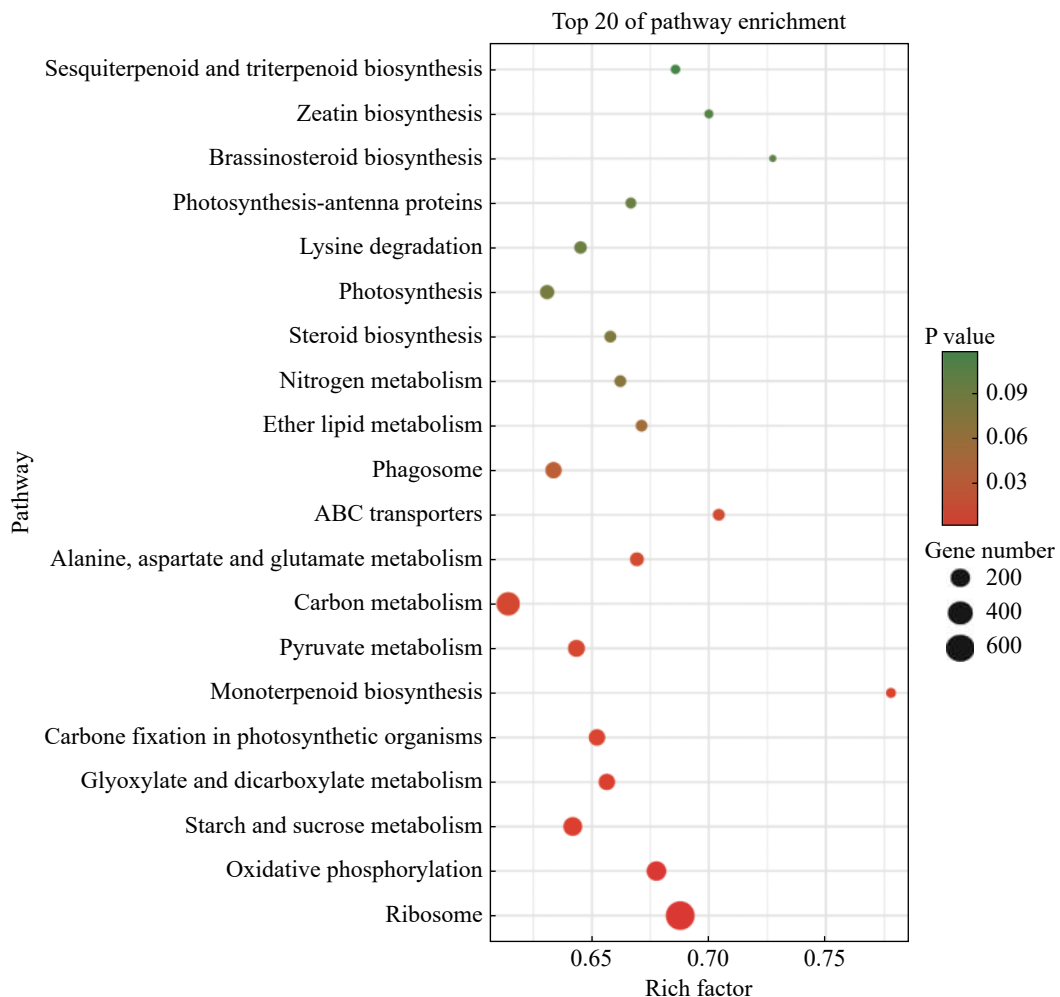


图 2 “万海 4 号”特有基因 KEGG 富集分析气泡图

Fig. 2 Dotplot of KEGG enrichment analysis results of specific genes of Clone ‘Wanhai4’

在此基础上,对 2 个样本的共有基因进行表达谱分析,结果(图 4)表明,“万海 4 号”油茶在油菜素类固醇生物合成、倍半萜和三萜生物合成、单萜类生物合成和类固醇生物合成等与植物次生代谢相关的 4 个途径的表达量都明显高于“万海 1 号”,而“万海 1 号”油茶在植物病原体互作、脂肪酸降解与脂肪酸生物合成等途径的表达量都高于“万海 4 号”,这和 KEGG 富集的结果相一致。

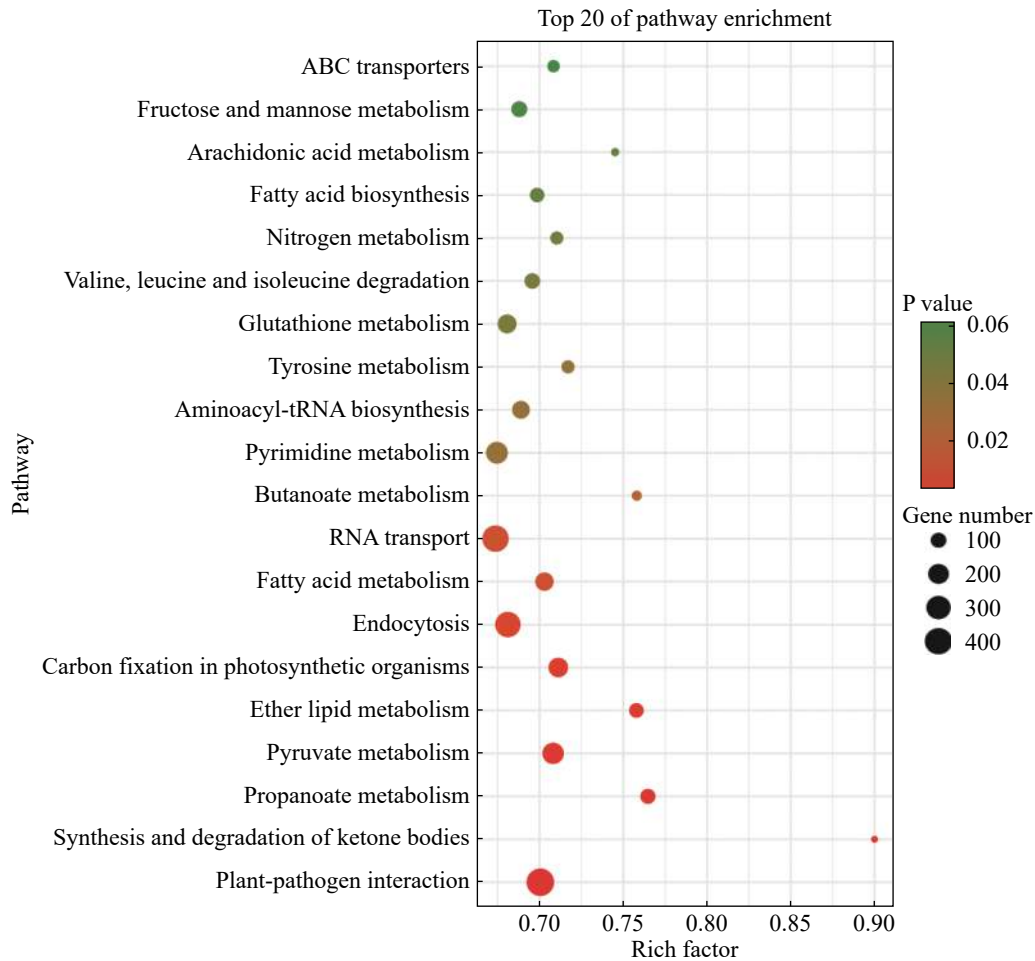


图 3 “万海 1 号”特有基因 KEGG 富集分析气泡图

Fig. 3 Dotplot of KEGG enrichment analysis results of specific genes of Clone ‘Wanhai1’

对 2 个样品中富集特异性基因较多的 KEGG 通路进行统计, 所得结果具有一定的相似性, 2 个样品共同的基因高富集途径见表 2。其中, 有 2 条与代谢相关的途径: 碳代谢和氨基酸的生物合成; 有 4 条与遗传信息处理相关的途径, 分别是内质网中的蛋白质加工、核糖体、RNA 转运和剪接体; 还有 1 条胞吞作用途径, 则与细胞过程相关。富集到这些共同途径的基因对于越南油茶维持正常的生理功能可能起到了一定的调节作用。

**2.2.2 特有基因 GO 富集分析** 对“万海 4 号”和“万海 1 号”特有的差异表达基因进行 GO 富集分析, 结果(图 5)表明, “万海 4 号”的特有基因被聚集到生物过程、细胞成分和分子功能 3 大类, 在生物过程大类中, 代谢过程、细胞内过程、单组织过程所富集的基因较多; 基因在细胞成分大类富集最多的 3 个途径分别是膜、细胞和细胞组分; 而在分子功能大类中, 催化活性和结合 2 个分支富集的差异基因较多。“万海 1 号”和“万海 4 号”GO 富集分析的结果相似(图 6)。

**2.3 Ka/Ks 分析** 编码基因中 Ka 与 Ks 的比值可以用来估计选择压力。Ka/Ks 值大于或小于 1 的基因表明该基因分别受到阳性或纯化选择。估算“万海 4 号”油茶和“万海 1 号”油茶中所有 8 835 对同源 unigene 对的 Ka 和 Ks 值, 计算所有 unigene 对的 Ka/Ks 比值, 结果(图 7)表明, Ka 和 Ks 值之间显著正相关, 共有 1 980 个 unigenes 的 Ka/Ks 比值大于 1, 其余 6 855 个 unigenes 的 Ka/Ks 比值小于或等于 1。

对“万海 1 号”与“万海 4 号”单拷贝同源基因中受到强烈正向选择的基因(Strong Positive Select, SP)进行功能富集分析, GO 注释的结果表明, 基因被富集到生物过程、分子功能和细胞成分 3 个大类, 在生物过程大类中, 某些基因富集到与抗氧化能力相关的 GO 条目, 如“GO:0006801-超氧化物代谢过程”, 且

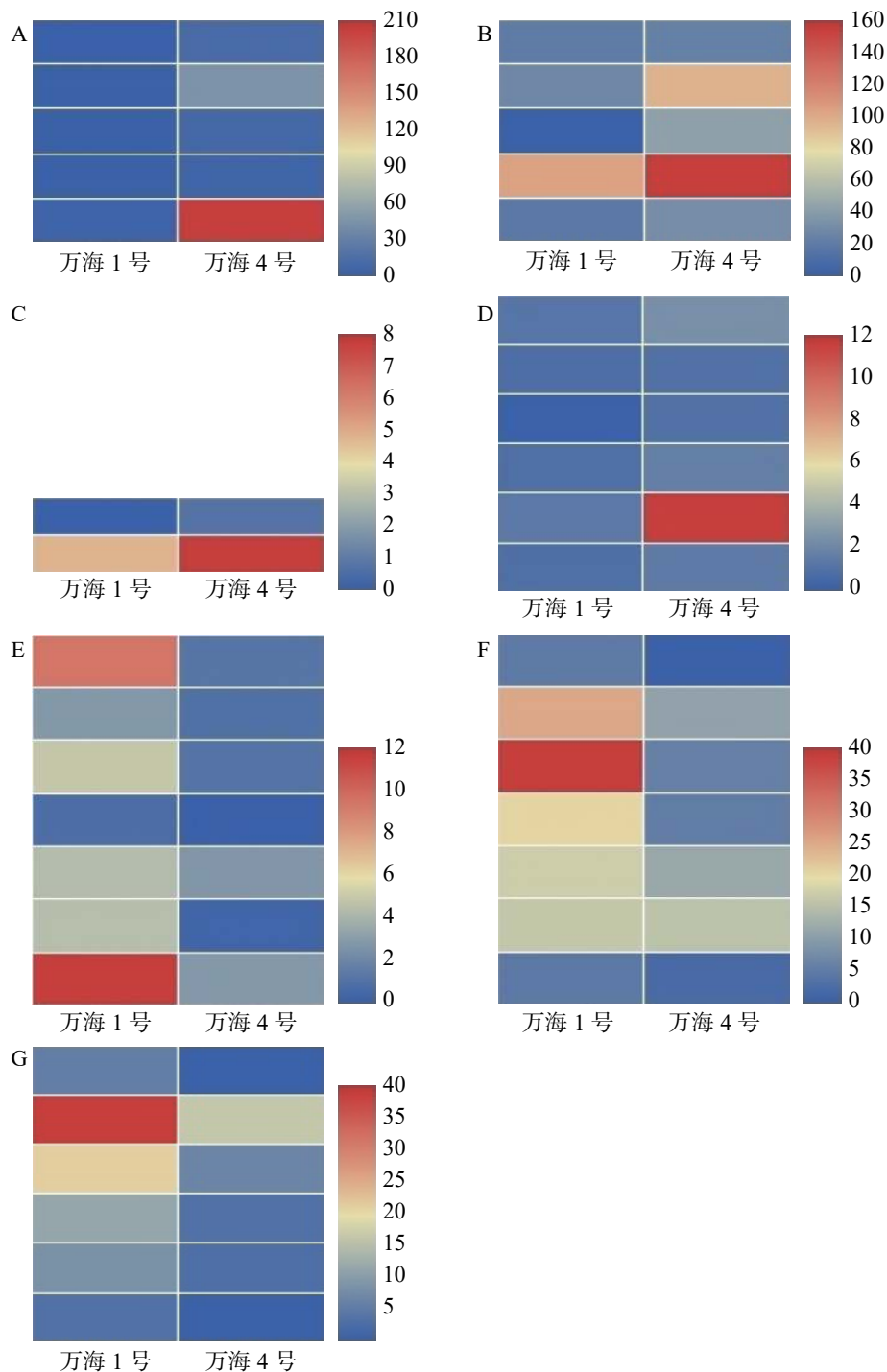


图4 共有基因表达谱

A.油菜素类固醇生物合成; B.单萜类生物合成; C.倍半萜和三萜生物合成; D.类固醇生物合成; E.植物病原体互作; F.脂肪酸降解; G.脂肪酸合成。

Fig. 4 Gene expression profile of shared genes

A. Brassinosteroid biosynthesis; B. Monoterpenoid biosynthesis; C. Sesquiterpenoid and triterpenoid biosynthesis; D. Steroid biosynthesis; E. Interaction of plant pathogens; F. Degradation of fatty acids; G. Biosynthesis of fatty acids.

$P < 0.01$ , 呈极显著相关, 而在分子功能大类中, 部分基因富集到“GO:0016655-氧化还原酶活性, 作为NAD(P)H, 醌或类似化合物的受体”条目, 也显示出与抗氧化能力的相关性(表3)。KEGG富集分析结果(表4)表明, 单拷贝同源基因中受到强烈正向选择的基因多被富集到与氨基酸合成代谢相关的条目, 如“精氨酸和脯氨酸代谢”、“嘌呤代谢”, 以及与生物碱合成代谢相关的条目, 如“莨菪烷、哌啶和吡啶生物



表 2 “万海 1 号”“万海 4 号”KEGG 高富集途径统计  
Tab. 2 Statistics of KEGG enrichment pathway of 2 elite clones of *Camellia oleifera*

KEGG_A_class	KEGG_B_class	代谢通路 Pathway	富集基因数/个 Enriched genes	
			“万海1号”	“万海4号”
代谢 metabolism	全局和概述 Global and Overview	碳代谢 Carbon metabolism	519	422
遗传信息处理 Genetic Information Processing	折叠, 分类和降解 Folding, sorting and degradation	内质网中的蛋白质加工 Protein processing in endoplasmic reticulum	416	291
代谢 metabolism	全局和概述 Global and Overview	氨基酸的生物合成 Biosynthesis of amino acids	401	280
遗传信息处理 Genetic Information Processing	翻译 Translation	核糖体 Ribosome	393	729
细胞过程 Cellular Processes	运输和分解代谢 Transport and catabolism	胞吞作用 Endocytosis	378	230
遗传信息处理 Genetic Information Processing	翻译 Translation	RNA转运 RNA transport	415	242
遗传信息处理 Genetic Information Processing	转录 Transcription	剪接体 Spliceosome	347	220

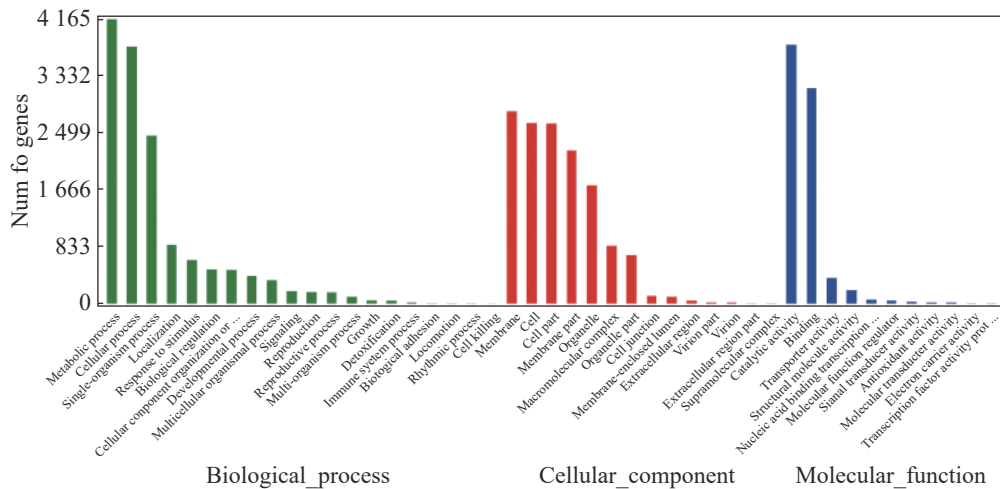


图 5 “万海 4 号”特有基因 GO 分类图

Fig. 5 GO gene classification of specific genes in Clone ‘Wanhai4’

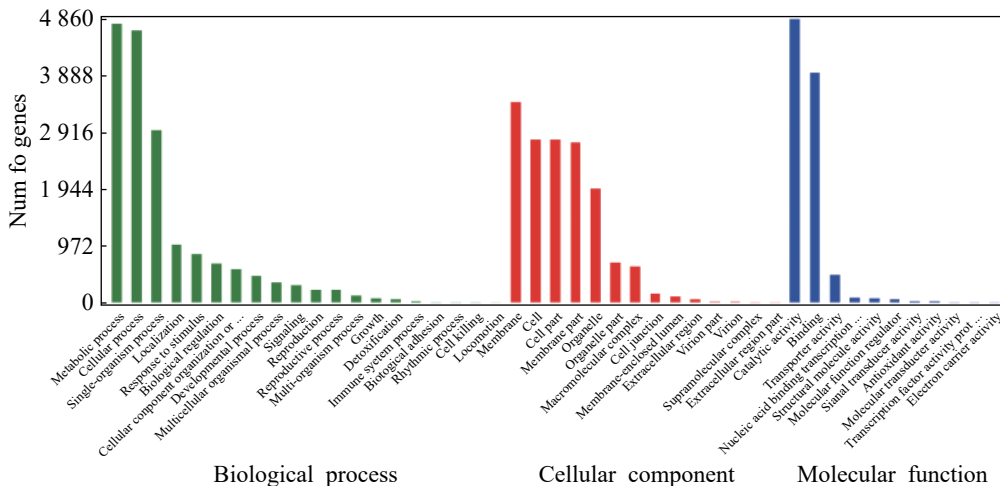


图 6 “万海 1 号”特有基因 GO 分类图

Fig. 6 GO gene classification of specific genes in Clone ‘Wanhai1’

碱的生物合成”、“异喹啉生物碱的生物合成”,且  $P$  值均小于 0.05, 显示出显著的相关性。

**2.4 单拷贝同源基因 SSR 分析** 在“万海 1 号”和“万海 4 号”2 个样品的单拷贝同源基因中共检索到 11 949 个 SSR 候选位点(表 5)。“万海 1 号”序列总大小为 18 736 662 bp, 共检索到 6 746 个 SSR 位点, 包含 SSR 的序列数为 4 937 条, 其中有 1 364 条序列包含 1 个以上 SSR, 化合物形成中存在的 SSR 数为 848; “万海 4 号”序列总大小为 16 440 557 bp, 共检索到 5 203 个 SSR 位点, 包含 SSR 的序列数为 3 990 条, 其中有 956 条序列包含 1 个以上 SSR, 化合物形成中存在的 SSR 数为 604。

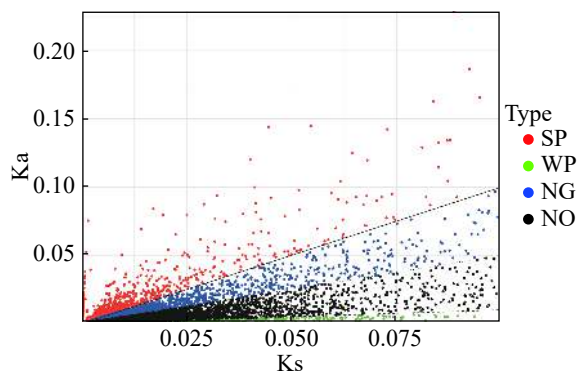


图 7 “万海 4 号”与“万海 1 号”比较所得  $K_a$  与  $K_s$  散点图  
Fig. 7 Scatter plots of  $K_a$  and  $K_s$  in Clones ‘Wanhai4’ and ‘Wanhai1’

表 3 “万海 1 号”和“万海 4 号”中 SP 基因的 GO 富集

Tab. 3 GO enrichment for SP genes in Clones ‘Wanhai1’ and ‘Wanhai4’

GO 功能分类 GO classification	GO ID	功能描述 Function description	注释基因数(占比) Gene Number(Ratio)	P-value	Q-value
生物过程 Biological Process	GO:0006801	超氧化物代谢过程 Superoxide metabolic process	3 (1.06%)	0.001 007	0.521 483
	GO:0000338	蛋白质树突化 Protein deneddylation	2 (0.71%)	0.005 381	0.521 483
	GO:0004457	乳酸脱氢酶活性 lactate dehydrogenase activity	2 (0.74%)	0.001 579	0.290 456
分子功能 Molecular Function	GO:0016655	氧化还原酶活性, 作为 NAD(P)H, 醌或类似化合物的受体 Oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor	3 (1.11%)	0.005 501	0.327 781
	GO:0070546	L-苯丙氨酸氨基转移酶活性 L-phenylalanine aminotransferase activity	2 (0.74%)	0.007 537	0.327 781
	GO:0044435	质体部分 Plastid part	27 (10.63%)	0.000 816	0.074 252
细胞成分 Cellular Component	GO:0009526	质体包膜 Plastid envelope	16 (6.3%)	0.001 150	0.074 252
	GO:0031967	细胞器包膜 Organelle envelope	20 (7.87%)	0.002 151	0.074 252
	GO:0031975	包膜 Envelope	20 (7.87%)	0.002 216	0.074 252
	GO:0009536	质体 Plastid	38 (14.96%)	0.009 439	0.217 881

表 4 “万海 1 号”和“万海 4 号”中 SP 基因的 KEGG 富集

Tab. 4 KEGG enrichment for SP gene in Clones ‘Wanhai1’ and ‘Wanhai 4’

途径 Pathway	注释基因数(占比)	P-value	Q-value	Pathway ID
莨菪烷、哌啶和吡啶生物碱的生物合成 Tropane, piperidine and pyridine alkaloid biosynthesis	5 (1.58%)	0.018 213	0.623 606	ko00960
精氨酸和脯氨酸代谢 Arginine and proline metabolism	8 (2.53%)	0.022 31	0.623 606	ko00330
基础切除修复 Base excision repair	7 (2.22%)	0.023 441	0.623 606	ko03410
RNA 聚合酶 RNA polymerase	7 (2.22%)	0.027 78	0.623 606	ko03020
嘌呤代谢 Purine metabolism	17 (5.38%)	0.039 612	0.623 606	ko00230
异喹啉生物碱的生物合成 Isoquinoline alkaloid biosynthesis	4 (1.27%)	0.048 535	0.623 606	ko00950

表 5 SSR 类型统计  
Tab. 5 Statistics of SSR types

项目 Statistical item	“万海1号”	“万海4号”
评估的序列总数 Total number of sequences examined	17 377	16 405
评估序列总碱基量 Total size of examined sequences/bp	18 736 662	16 440 557
识别的SSRs总数 Total number of identified SSRs	6 746	5 203
包含SSR的序列数目 Number of SSR containing sequences	4 937	3 990
包含1个以上 SSR 的序列数目 Number of sequences containing more than 1 SSR	1 364	956
以复合物存在的 SSR 数目 Number of SSRs present in compound formation	848	604
二核苷酸 Di-nucleotide	3 855	2 882
三核苷酸 Tri-nucleotide	2 134	1 711
四核苷酸 Tetra-nucleotide	253	195
五核苷酸 Penta-nucleotide	158	131
六核苷酸 Hexa-nucleotide	346	284

在 SSR 位点各碱基重复数量中, 2 个样品的占比大小相似, 占比最大的是二碱基重复(“万海 1 号”中占 56.9%, “万海 4 号”中占 55.2%), 其次是三碱基重复(“万海 1 号”中占 31.8%, “万海 4 号”中占 33%)。与“万海 1 号”相比, “万海 4 号”中各碱基重复的类型数量均较为接近, 且二者在二碱基、三碱基重复的优势基元类型上完全一致; 不同之处在于“万海 1 号”的四碱基重复优势基元类型为 AAAG/CTTT 和 AAAT/ATTT, 而“万海 4 号”的四碱基重复优势基元类型为 AAAC/GTTT 和 AAAG/CTTT (图 8, 9)。

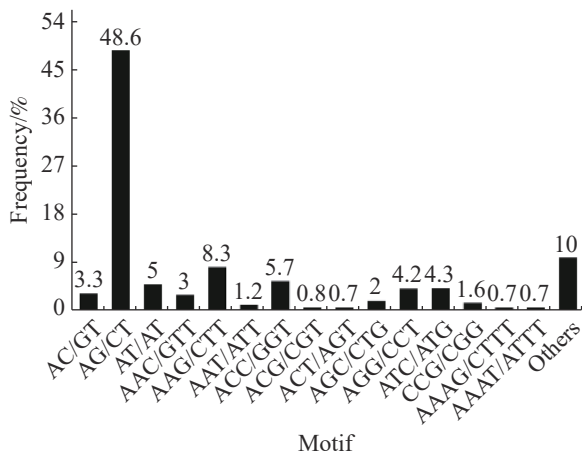


图 8 “万海 1 号”单拷贝同源基因 SSR 分布图

Fig. 8 SSR distribution of single copy orthologs of Clone ‘Wanhail’

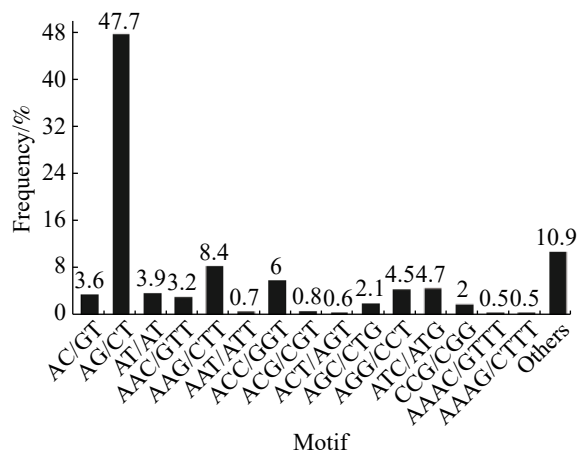


图 9 “万海 4 号”单拷贝同源基因 SSR 分布图

Fig. 9 SSR distribution of single copy orthologs of Clone ‘Wanhai4’



### 3 讨论

转录组测序技术通过在单核苷酸水平进行检测,能够准确高效地获得某物种在不同时期最为全面的转录本信息。作为生物信息学常用的研究方法,它具有通量高、灵敏度高、分辨率高、限制性小等特点<sup>[15]</sup>,近年来,其在油茶上的应用也逐渐增多。林萍等<sup>[16]</sup>初次在油茶研究中采用 solexa 测序技术进行分析,初步展示了油茶种子发育过程中基因组的表达图谱,并阐明了部分编码蛋白的代谢通路。江南等<sup>[17]</sup>以普通油茶的种子为材料,利用转录组测序技术挖掘出 12 个油茶发育过程中调控  $\alpha$ -亚麻酸合成的关键酶基因。吕晓杰等<sup>[18]</sup>通过对 3 个不同时期的油茶籽进行转录组测序,研究了与脂肪酸和甘油三酯代谢相关的 4 个关键酶基因。曾艳玲等<sup>[19]</sup>通过 RNA-Seq 技术对油茶中 WRKY 转录因子及植物萜类物质合成进行了分析。卿卓等<sup>[20]</sup>采用高通量 Illumina Hiseq 测序手段对红花油茶(*Camellia reticulata*)花蕾期和泌蜜高峰期的蜜腺进行转录组测序,得到了与花蜜化学组成调控过程有关的差异表达基因。

本研究通过对“万海 1 号”和“万海 4 号”2 个海南本地油茶优良品种进行转录组测序分析,丰富了其遗传资源。对“万海 1 号”和“万海 4 号”各自的特异性表达基因的进行 GO 和 KEGG 富集分析,结果表明,“万海 4 号”的基因富集结果中,出现了倍半萜和三萜生物合成、油菜素类固醇生物合成、类固醇生物合成和单萜类生物合成等与植物次生代谢相关的途径。而在“万海 1 号”的富集结果中则出现了脂肪酸生物合成及脂肪酸代谢等与脂肪酸相关的途径,这与油茶的茶籽、茶油品质息息相关。另外,在“万海 1 号”的 KEGG 富集结果中,富集基因数最多的 1 条分支为植物病原体互作途径,表明“万海 1 号”特异性基因与油茶抗病性的关联,2 个样品共同基因的表达谱分析也验证了此结果。对 2 个样品中富集特异性基因较多的 KEGG 通路进行统计,总结出部分二者共有的富集途径,其中,有 2 条与代谢相关的途径:碳代谢和氨基酸的生物合成;还有 4 条与遗传信息处理相关的途径,分别是内质网中的蛋白质加工、核糖体、RNA 转运和剪接体;还有 1 条胞吞作用途径,与细胞过程相关。

在此基础上,笔者通过 Ka/Ks 筛选出 2 个样品单拷贝同源基因中受到强烈正向选择的基因,进一步进行 GO 和 KEGG 注释,结果表明,某些单拷贝同源基因被富集到与抗氧化能力相关的 GO 条目,如“GO:0006801-超氧化物代谢过程”、“GO:0016655-氧化还原酶活性,作为 NAD(P)H,醌或类似化合物的受体”等条目,显示出与抗氧化能力及油茶抗病性的相关性。KEGG 富集分析显示:基因多被富集到与氨基酸合成代谢相关的条目,如“精氨酸和脯氨酸代谢”、“嘌呤代谢”,以及与生物碱合成代谢相关的条目,如“莨菪烷、哌啶和吡啶生物碱的生物合成”、“异喹啉生物碱的生物合成”,且  $P$  值均小于 0.05,显示出显著的相关性。

本研究结果表明,“万海 1 号”特异性基因中,有相当一部分与其抗病性及脂肪酸代谢相关,而“万海 4 号”的特异性基因则表现出与植物次生代谢的相关性,2 个样本的单拷贝同源基因多显示出与抗氧化能力及氨基酸合成代谢的相关性。因此,“万海 1 号”油茶可能具有较强抗逆性,而“万海 4 号”油茶可能积累了更多的次生代谢产物。在此基础上,笔者建议可设计进一步实验,验证“万海 1 号”和“万海 4 号”油茶中与抗逆、抗病、次生代谢合成等品质有关的基因,从而对海南本地油茶的生理及代谢过程进行分子层面的调控,为海南本地油茶的优良品种选育和改良提供可靠的科学依据。

### 参考文献:

- [1] 张震,许彦明,陈永忠,等.油茶转录组测序与 SSR 特征分析[J].西南林业大学学报(自然科学版),2018,38(6):63-68.
- [2] 庄瑞林.中国油茶[M].2版.北京:中国林业出版社,2008.
- [3] 庄瑞林.我国油茶良种选育工作的历史回顾与展望[J].林业科技开发,2010,24(6):1-5.
- [4] 贾效成,陈良秋,余凤玉,等.海南本地油茶优良品系遗传及经济性状研究初报[J].热带农业科学,2018,38(6):56-60.
- [5] 邓亦麒,何兴金,余岩.叶子花花瓣状苞片和叶片的比较转录组学研究[J].西北植物学报,2018,38(6):1039-1047.

- [6] 唐仕云, 杨丽涛, 李杨瑞. 低温胁迫下不同甘蔗品种的转录组比较分析[J]. 生物技术通报, 2018, 34(12): 116 – 124.
- [7] 李和平, 姚运法, 练冬梅, 等. 黄秋葵果实转录组测序及分析[J]. 生物技术通报, 2018, 34(3): 121 – 127.
- [8] WANG Z, GERSTEIN M, SNYDER M. RNA-Seq: a revolutionary tool for transcriptomics [J]. *Nat Rev Genet*, 2009, 10(1): 57 – 63.
- [9] 崔凯, 吴伟伟, 刁其玉. 转录组测序技术的研究和应用进展[J]. 生物技术通报, 2019, 35(7): 1 – 9.
- [10] LI C F, ZHU Y J, GUO X, et al. Transcriptome analysis reveals ginsenosides biosynthetic genes, microRNAs and simple sequence repeats in *Panax ginseng* C. A. Meyer [J]. *BMC Genomics*, 2013, 14: 245.
- [11] WU B, LI Y, YAN H, et al. Comprehensive transcriptome analysis reveals novel genes involved in cardiac glycoside biosynthesis and miRNAs associated with secondary metabolism and stress response in *Digitalis purpurea* [J]. *BMC Genomics*, 2012, 13: 15.
- [12] QI J, ZHENG N, ZHANG B, et al. Mining genes involved in the stratification of Paris polyphylla seeds using high-throughput embryo transcriptome sequencing [J]. *BMC Genomics*, 2013, 14: 358.
- [13] 刘雷, 赵欢, 冉茂中, 等. 罗勒花和叶的转录组数据组装及基因功能注释[J]. *中草药*, 2017, 48(17): 3612 – 3618.
- [14] FATIMA T, SNYDER C L, SCHROEDER W R, et al. Fatty acid composition of developing sea buckthorn (*Hippophae rhamnoides* L.) berry and the transcriptome of the mature seed [J]. *PLoS One*, 2012, 7(4): e34099.
- [15] 张春兰, 秦孜娟, 王桂芝, 等. 转录组与 RNA-Seq 技术[J]. 生物技术通报, 2012(12): 51 – 56.
- [16] 林萍, 曹永庆, 姚小华, 等. 普通油茶种子 4 个发育时期的转录组分析[J]. *分子植物育种*, 2011, 9(4): 498 – 505.
- [17] 江南, 谭晓风, 张琳, 等. 基于 RNA-Seq 的油茶种子  $\alpha$ -亚麻酸代谢途径及相关基因分析[J]. *林业科学*, 2014, 50(8): 68 – 75.
- [18] 吕晓杰, 潘德灼, 李健, 等. 油茶种仁成熟过程油脂合成代谢的转录组分析[J]. *森林与环境学报*, 2018, 38(2): 156 – 163.
- [19] 曾艳玲, 谭晓风, 颜亚丹, 等. 基于转录组数据的油茶 WRKY 转录因子序列分析[J]. *经济林研究*, 2017, 35(4): 9 – 12.
- [20] 卿卓, 苏睿, 赵文正, 等. 腾冲红花油茶蜜腺 2 个发育时期转录组差异性分析[J]. *云南农业大学学报(自然科学版)*, 2019, 34(2): 185 – 192.

## Comparative Transcriptome Analysis of Two Clones of *Camellia vietnamensis* in Hainan

DAI Jiani<sup>1</sup>, YU Jing<sup>1,2</sup>, ZHENG Wei<sup>1</sup>, WANG Jian<sup>1,2</sup>, WU Yougen<sup>1,2</sup>, HU Xinwen<sup>1,2</sup>, LAI Hanggui<sup>1,2</sup>

(1. Institute of Horticulture, Hainan University, Haikou, Hainan 570228; 2. Engineering and Technology Research Center of Ministry of Education for Selection and Breeding of New Tropical Crop Varieties, Hainan University, Haikou, Hainan 570228)

**Abstract:** In order to further explore the physiological characteristics and metabolic process and also to provide scientific basis for the selection of excellent clones of *Camellia vietnamensis* in Hainan, mature leaves of two years old plants from two clones (*Camellia vietnamensis* ‘Wanghai1’ and ‘Wanghai4’ derived from two elite individuals of *Camellia vietnamensis* plants from Hainan were selected for transcriptome sequencing and analysis by using next-generation high-throughput Illumina transcriptome sequencing technology, and their differences in gene expression were compared. The results showed that among the specific genes of ‘Wanghai1’ there were 651 annotated to the path related to plant disease resistance, accounting for about 5.26%, and 510 annotated to the pathway related to fatty acid metabolism, accounting for about 4.12 %, and that among the specific genes in ‘Wanghai2’ there were 169 annotated to pathways such as sesquiterpenoid and triterpenoid biosynthesis, brassinosteroid biosynthesis, steroid biosynthesis and monoterpene biosynthesis, which are related to plant secondary metabolism. The single-copy homologous genes of the two samples were mostly enriched to GO entries related to antioxidant capacity and KEGG entries related to amino acid anabolic metabolism, and their P levels were less than 0.05, showing a significant correlation between their genes and antioxidant capacity or amino acid anabolic metabolism. This might indicate that Clone ‘Wanghai1’ has higher stress resistance while Clone ‘Wanghai4’ accumulates more secondary metabolites.

**Keywords:** *Camellia vietnamensis* ; comparative transcriptome; gene expression

(责任编辑: 钟云芳)