

油茶甘油二酯激酶基因(*CoDGK3*)的克隆与表达分析江 南^{1,2}, 张维利¹, 李捷宇², 刘奇龙¹, 刘 敏¹

(1. 湖南工业大学 包装与材料工程学院, 湖南 株洲 412008; 2. 中南林业科技大学 经济林培育与保护教育部重点实验室/经济林育种与栽培国家林业局重点实验室, 湖南 长沙 410004)

摘要: 以油茶转录组数据为基础, 采用 RT-PCR 和 RACE 技术, 根据 Unigene 序列设计油茶甘油二酯激酶基因(*CoDGK*) RACE 引物, 分离克隆了 DGK 基因全长 cDNA 序列, 命名为 *CoDGK3* (GeneBank 登录号: KM068056), 基因全长 1 568 bp, 开放阅读框 1 455 bp, 编码 485 个氨基酸; 氨基酸同源比对及序列分析表明油茶与其它物种的 DGK 氨基酸序列具有较高的相似性, 其中与麻风树和毛果杨相似性最高, 达到 78%; 利用生物学软件对蛋白质进行生物信息学分析, 获得其结构特点及理化性质; 分析转录组数据中 DGK 相关的 Unigene 序列 RTKM 值和对 *CoDGK3* 基因实施实时荧光定量 PCR, 结果表明 *CoDGK3* 在油茶种子发育各时期表达量变化差异不大, 并且发现转录组数据分析和实时荧光定量分析的结果变化规律一致, 证明了转录组测序的有效性。

关键词: 油茶; 甘油二酯激酶基因; 克隆; 生物信息学分析; 基因表达

中图分类号: Q 785 **文献标志码:** A **文章编号:** 1673—1689(2014)10—1090—06

Diacylglycerol Kinase 3(*CoDGK 3*) Gene Cloning and Expression Analysis of *Camellia Oleifera*JIANG Nan^{1,2}, ZHANG Weili¹, LI Jieyu², LIU Qilong¹, LIU Ming¹

(1. School of Packing and Material Engineering, Hunan University of Technology, Zhuzhou 412008, China; 2. Key Laboratory of Cultivation and Protection for Non-Wood Forest of Ministry of Education/Key Laboratory of Non-Wood Forest Product of Forestry Ministry, Central-South University of Forestry and Technology, Changsha 410004, China)

Abstract: Based on transcriptome Unigene sequence the diacylglycerol Kinase gene named *CoDGK3* (GeneBank accession number: KM068056) was isolated from *Camellia oleifera* with the method of RT-PCR and RACE. Bioinformatics approaches are used to analyze and predict the structural characters and physicochemical property. The bioinformatics analysis showed that the CDS of *CoDGK3* is 1568 bp in length, OFR 1455 bp and it codes predicted protein of 485 amino acids. Homologous comparison and sequence analysis of *CoDGK3* amino acids suggested that the *CoDGK3* is highly conserved and showed the highest similarity (78%) with the DGK of *Jatropha*

收稿日期: 2014-04-01

基金项目: 湖南省科技计划项目(2012NK3065)。

作者简介: 江 南(1968—), 女, 湖南醴陵人, 副教授, 主要从事农林生物技术研究。Email: namijiangnan@126.com

curcas and *Populus trichocarpa*. The gene expression changes by qPCR were similar to those of RTKM in transcriptome data indicate that the *CoDGK3* expresses stably in different developmental stage of *Camellia oleifera* seeds and the transcriptome data of *Camellia oleifera* seeds is valid and reliable.

Keywords: *Camellia oleifera*, diacylglycerol kinase, cloning, bioinformatics analysis, gene expression

油茶(*Camellia oleifera*)是我国南方重要的木本食用油料树种和良好的生态树种,油茶种仁所含油脂含有90%以上的不饱和脂肪酸和丰富的 α -生育酚、甾醇、维生素等。油茶的抗性与油茶信号转导逆境胁迫应答相关基因的表达密切关联。目前对油茶抗炭疽病及油茶环境胁迫引起的植物形态和生理变化研究较多^[1-8],而对引发这些变化的分子机制研究报道极少。油茶转录组数据分析显示,*DGK*是油茶磷脂酰肌醇(PI)信号系统下游反应过程中的关键基因且在油茶种子发育过程中表达并呈现出多基因家族特征^[9],作者在油茶种子转录组数据分析的基础上,筛选出*DGK*的Unigene序列,根据Unigene序列设计引物克隆出1条油茶*DGK*全长序列,命名*CoDGK3*;并通过qPCR结合不同时期表达谱数据分析对*CoDGK3*在油茶种子发育过程中的表达模式进行分析,以期研究油茶抗性机制和提高油茶品种单产量的遗传改良提供依据。

1 材料与方法

1.1 试验材料

经济林育种与栽培国家林业局重点实验室构建的油茶果实膨大期和油脂合成高峰期转录组数据;国家审定油茶品种‘华硕’果实采自中南林业科技大学油茶种植资源圃,分别于2013年6月、8月和10月下旬采集,果实采摘后立即剥取种仁,液氮保存,回实验室后-80℃保存。

1.2 试验方法

1.2.1 油茶种仁 RNA 提取及 cDNA 合成 Invitrogen 公司的 PureLink™ RNA Mini Kit 试剂盒结合 CTAB 裂解、氯仿抽提,分别提取油茶各时期种仁 RNA,琼脂糖电泳检测 RNA 提取效果, Nanodrop ND-1000 测定 RNA 浓度,选取浓度 ≥ 200 ng/ μ L, 28S:18S 约为 2:1 的 RNA 为模板, Fermentas 公司 RevertAid™ First Strand cDNA Synthesis Kit 逆

转录获得 cDNA 用于 *CoDGK3* 全长 PCR 和荧光定量 PCR, Invitrogen 公司 3'RACE 和 5'RACE 试剂盒进行第一链反转录获得的 cDNA 用于 *CoDGK3* 克隆。

1.2.2 甘油二酯激酶基因的克隆 油茶转录组数据分析共获得 *CoDGK* Unigene 序列 19 条, 序列长度 269~988 bp 不等, 对 Unigene68639 序列长度为 944 bp 的 cDNA 片段设计 3'RACE 和 5'RACE 引物(如表 1), 分别以 3'RACE 和 5'RACE 试剂盒反转录获得的 cDNA 为模板进行降落 PCR (TD-PCR), 反应程序: 94℃ 2 min; 94℃ 30 s, 退火温度 40 s, 72℃ 延伸 1.5 min, 70、68、65℃ 各 5 个循环, 61℃ 15 个循环, 57℃ 10 个循环, 72℃ 总延伸 10 min。获得 *CoDGK3* 3' 和 5' 端序列; 用 Vector NTI 10.3.0 将 5' 端序列、944 bp 的 Unigene 序列和 3' 端序列进行拼接, 获得 *CoDGK3* cDNA 全长拼接序列; 分别在全长 cDNA 拼接序列的 3' 端和 5' 端设计引物(如表 1), Takara 公司 Prime STAR HS DNA Polymerase 高保真 DNA 聚合酶进行全长扩增, TA 克隆, 挑选阳性克隆送铂尚生物技术公司测序。引物均由北京六合华大基因公司合成。

1.2.3 甘油二酯激酶基因同源性分析 将克隆的 *CoDGK3* cDNA 编码区翻译成氨基酸序列, 利用 NCBI 数据库中的 BLAST 功能, 搜索 *CoDGK3* 的同源核苷酸序列和氨基酸序列, MEGA5.03 程序构建系统进化树, 分析 *CoDGK3* 与其他各物种 *DGK* 的同源性大小。

1.2.4 甘油二酯激酶基因生物信息学分析 用生物学软件结合网络数据库对 *CoDGK3* cDNA 全长序列进行分析。在线软件 ProtParam 对基因编码的蛋白质作结构特点和理化分析、SignalP 4.1 Server 分析信号肽识别位点、ProtScale 分析疏水性位点、TMpred 预测蛋白质的跨膜结构、Predictprotein 完成蛋白质二级结构及功能分析。

表 1 *CoDGK3* 克隆相关引物Table 1 Primer used in *CoDGK3* cloning

引物名称	引物序列	引物用途
3'-GSP1	GAAGATGCGTTATAGGGACTTGACTCCT	3'RACE PCR
3'-GSP2	GATGACGAGGAGGATGACAGTAATGAAG	3'RACE 巢式 PCR
3'-GSP3	GCTTAGTTAGCTCTCCAGATTTGTTC	3'RACE 巢式 PCR
5'-GSP2	GTCTTCATTACTGTCATCCTCCTCGTCAT	5'RACE PCR
5'-GSP3	CATGTCCAGTTGGAGCAAACATAACGAG	5'RACE 巢式 PCR
5'-GSP4	GAGGACTCAAGTCCCTATAACGCATCT	5'RACE 巢式 PCR
PF1	CCGTTACCGCGATTTCGATTCTC	<i>CoDGK3</i> 全长扩增
PR1	GATGCCAATGCTGACAGACGAAC	<i>CoDGK3</i> 全长扩增
PF2	CATCAGTTGAGTAGTAATGGGTAACC	<i>CoDGK3</i> 巢式扩增
PR2	GAGAGGCTAACTAAGCTGAGAAATGTC	<i>CoDGK3</i> 巢式扩增
AUAP	GGCCACGCGTCGACTAGTAC	通用引物

1.2.5 *CoDGK3* 在油茶种子发育不同时期表达量分析 油茶转录组数据分析显示,DGK 主要参与油茶 PI 信号系统下游反应生理过程,且油茶在油茶种子发育过程不同时期表达量变化不明显^[9]。采用 qPCR 验证 *CoDGK3* 在油茶种子不同发育时期的表达模式。以油茶中表达稳定的 GAPDH 基因 (rzots0_001325.y1.scf) 为内参对照基因 (引物:PF:5'-CTACTGGAGTTTTACCGA 3',PR:5'-TAAGACCC TCA ACAATGCC 3',PCR 产物 231 bp),Primer Primer5.0 设计用于 qPCR 的 *CoDGK3* 引物 (PF:5'-GCTCGTTATGTTTGCTCCA 3',PR:5'-TACTGTCAT CCTCCTCGTC 3'),引物均由北京六合华大基因公司合成。BIO-RAD 荧光定量 PCR 仪上完成 qPCR 过程,反应体系 20 μ L (cDNA 1 μ L,2 $\times$ SYBR Green qPCR Mix 10 μ L,10 μ mol/L 的正反向引物各 1 μ L,超纯水补足至 20 μ L),引物退火温度为 58 $^{\circ}$ C (反应程序:95 $^{\circ}$ C 2 min;95 $^{\circ}$ C 15 s,58 $^{\circ}$ C 30 s,72 $^{\circ}$ C 20 s,35 个循环),*CoDGK3* PCR 产物 275 bp, Bio-Rad CFX Manager 软件进行分析。

2 结果与讨论

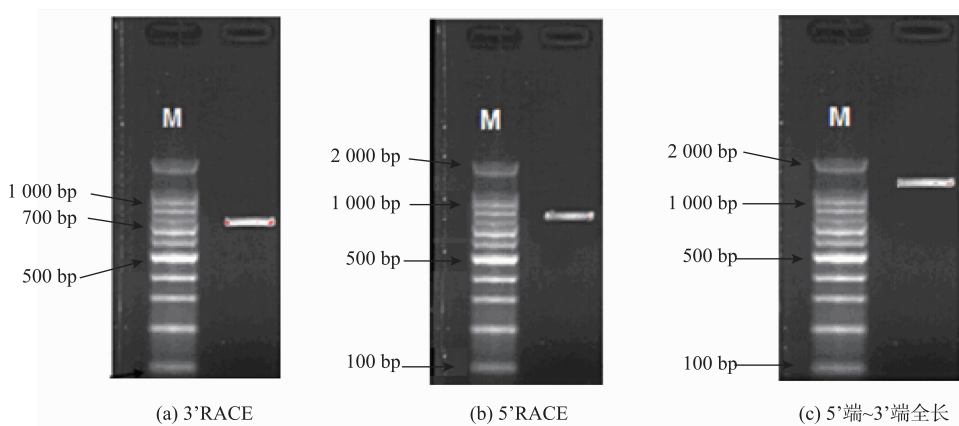
2.1 油茶 *CoDGK3* 基因克隆

BLASTN 显示油茶甘油二酯激酶基因 (*CoDGK*) 19 条 Unigene 序列 (cDNA) 分属于 3 个 DGK 基因;与麻风树 (*Jatropha curcas*)、毛果杨 (*Populus trichocarpa*) 等 DGK 同源性最大。根据 BLASTN 结果选取 Unigene68639 序列 (944 bp) 设计 RACE 引物进行 3'RACE 和 5'RACE,结果如图 1((a)、(b)), 序列拼接后获得长度为 2 120 bp 的完整的油茶甘

油二酯激酶基因全长 cDNA 序列,命名 *CoDGK3*,序列提交 GeneBank 编号 KM068056,5' 非编码区 317 bp,3' 非编码区 262 bp (不包括 PolyA),开放阅读框 (ORF) 1 458 bp;以全长 cDNA 5' 端和 3' 端非编码区设计的引物 (如表 1) 和高保真酶对编码区进行特异扩增,琼脂糖电泳检测获得清晰条带 (图 1(c)), 序列经 TA 克隆后测序,结果表明特异扩增所获序列长度为 1 568 bp,扩增条带包含 *CoDGK3* ORF 序列,且特异扩增序列与拼接序列比对一致性为 100%,编码区序列完全一致,可确定 *CoDGK3* 含有 CDS 序列 1 458 bp,编码 485 个氨基酸。

2.2 氨基酸序列同源性分析

用软件 GENDOCE 将 *CoDGK3* ORF 序列翻译为 485 个氨基酸序列如图 2,将 *CoDGK3* 氨基酸序列与其他 11 种植物 DGK 所编码的氨基酸序列比对进行同源性分析,MEGA5.03 程序构建系统进化树。同源性分析显示各物种与 *CoDGK3* 有 72% 以上的相似性,12 个物种长度大于 5 且氨基酸序列完全一致的保守序列有 15 处 (阴影区),1~32、55~105、431~467 位氨基酸序列是高度可变区 (方框标记区),不同物种氨基酸序列差异较大。同时氨基酸序列分析还发现 *CoDGK3* 不含有植物特有的上游基础结构域 (YTVNQLNKNISLSLIRAIKARAK RYKKAKDKVP),根据植物 DGK 结构域分类可判断 *CoDGK3* 不属于第 I 类 DGK 基因。进化系统树显示油茶 *CoDGK3* 与麻风树和毛果杨 DGK 同源性最大,氨基酸序列同源性可达 78%,其次是巨桉 (*Eucalyptus grandis*)、菜豆 (*Phaseolus vulgaris*) 和大豆 (*Glycine max*),相似性分别为 75%。

图1 油茶 *CoDGK3* 琼脂电泳图Fig. 1 Agar gel electrophoresis of *CoDGK3*

阴影区: *CoDGK3* 保守序列, 方框区: 高度变异区, 波浪线: *CoDGK3* 的 ATP 结合位点

图2 *CoDGK3* 氨基酸序列Fig. 2 Amino acid sequence of *CoDGK3*

2.3 油茶甘油二酯激酶(*CoDGK3*)结构分析

2.3.1 *CoDGK3* 理化性质分析 用 ProtParam 对 *CoDGK3* 理化性质进行分析, 结果表明组成 *CoDGK3* 的 485 个氨基酸中 64 个氨基酸带正电荷, 57 个氨基酸带负电荷, 364 个氨基酸为中性。 *CoDGK3* 相对分子质量 54616.2, 分子式 $C_{2453}H_{3790}N_{664}O_{714}S_{19}$, 等电点 PI 为 6.08, 说明 *CoDGK3* 蛋白中酸性氨基酸多于碱性氨基酸, 其中含 Leu 含量在最高达 43 个, 占氨基酸总量的 8.9%, Cys 最少, 只有 8 个占氨基酸总量的 1.6%。 *CoDGK3* 蛋白稳定系数为 41.64, 大于 40 属于不稳定蛋白; 脂肪系数为 79.98, 总平均亲水性 -0.426, 利用软件 protscale 在线分析 *CoDGK3* 的疏水性, 发现 *CoDGK3* 蛋白具有多个较强的亲水区域, 结合其脂肪系数和总平均亲水指数, 可确定 *CoDGK3* 为亲水

性蛋白, 结果见图 3。

2.3.2 *CoDGK3* 酶蛋白二级结构及功能分析 Predictprotein 在线分析 *CoDGK3* 氨基酸序列, 发现 *CoDGK3* 酶蛋白二级结构包含 17.7% 螺旋结构, 20.8% 为直链结构及 61.4% 的 Loop 环结构, 约有 8 个可能形成二硫键的位点。同时发现一个 Low complexity region (DDDEEDDSNEDEDWE, 448~463 位氨基酸), 与功能域分析结果一致。检测出一个核定位信号 ("RKKMRYR" 335~441 位氨基酸), 推测 *CoDGK3* 发挥其功能可能有进入细胞核的过程。共发现 4 种蛋白质功能位点: 起始位分别为 9 (TLK)、156 (TDR)、224 (TFR)、311 (SIR)、472 (TFR) 的 3 个氨基酸为蛋白激酶 C 磷酸化位点 (Protein kinase C phosphorylation site); 起始位分别为 9 (TLKD)、24 (SKVD)、97 (TEIE)、212 (SQDD)、437 (SVED) 和 455

(SNED)的4个氨基酸为酪蛋白激酶Ⅱ磷酸化位点(Casein kinase II phosphorylation site);2 (GNPESD)、46 (GGQLGG)、109 (GGDGTA)、138 (GTGNNL)、227 (GGFWNY)、270 (GCTQGW)、374 (GTRLAQ)、468 (GAADTF)和 478 (GIDISQ)为起始位点的6个氨基酸的豆蔻酰化位点(N-myristoylation site);第149(WGKK)位起4个氨基酸的酰胺化位点(Amidation site),这些功能化位点能使油茶 *CoDGK3* 在油茶种子发育的不同时期发挥其功能维持细胞内的 DAG 和 PA 的平衡。

2.4 油茶 *CoDGK3* 基因表达量分析

油茶转录组数据中 RPKM 值为各 Unigene 在油茶种子发育相应时期的表达量,可直接用于比较不同时期相应基因表达差异^[10]。油茶转录组数据分析获得的 19 条 DGK 基因 Unigene 序列中,其中 3 条能完全比对到 *CoDGK3* cDNA 序列上,属于 *CoDGK3* 完整 cDNA 序列的一部分。转录组中 3 条 *CoDGK3* Unigene 的 RPKM 值在油茶种子发育 6、8、10 月份变化作柱状图如图 4(a),6 月份和 8 月份各 Unigene RPKM 值十分接近,8 月份略有上升,到 10

月份有下降趋势。各时期 RPKM 值虽有变化,但变化不明显,且 3 条 Unigene 的 RPKM 值在油茶发育各时期表达变化规律基本一致。以油茶中表达稳定的 GAPDH 基因为内参对 *CoDGK3* 实施荧光实时定量 PCR(qPCR),采用 Bio-Rad CFX Manager 软件,拟定相对表达丰度最高值为 1.0,计算 *CoDGK3* 的相对表达量,重复实验 3 次,作柱状图如图 4(b),结果表明 *CoDGK3* 8 月份表达量较 6 月份稍有上升,到 10 月份表达水平降低,但总体变化不大, *CoDGK3* 在各时期表达趋于平稳。

qPCR 检测 *CoDGK3* 在油茶各时期的表达规律结果与油茶转录组测序 RPKM 值变化趋于一致,间接地证明了转录组数据可靠性。*CoDGK* 作为油茶 PI 信号系统下游反应过程中的关键酶,在油茶种子发育的各时期表达稳定,不但保证油茶细胞中两个重要的信使分子 DAG 和 PA 含量稳定,同时 *CoDGK* 还参与调控磷脂酰肌醇再生循环过程的关键节点,保证了 PI 信号系统的最初底物供应^[9],推动油茶种子发育过程中油脂积累的顺利进行。

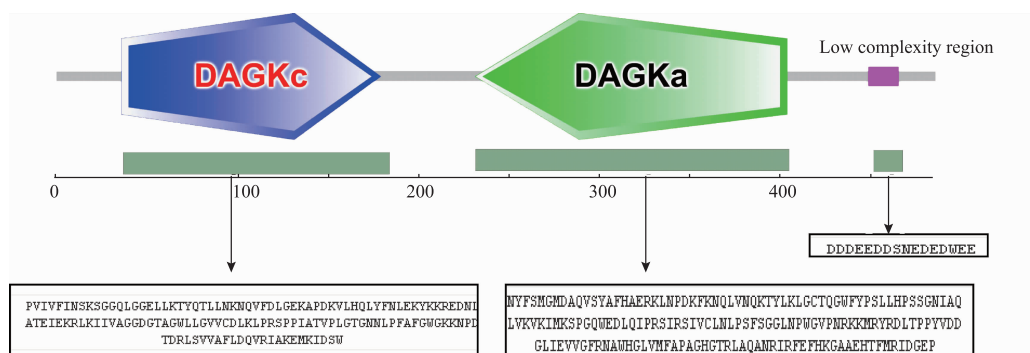


图 3 *CoDGK3* 功能域

Fig. 3 Functional domains of *CoDGK3*

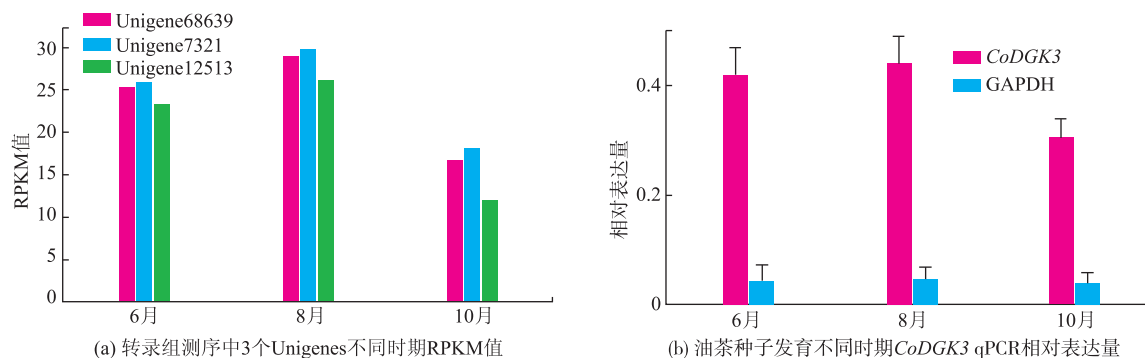


图 4 油茶种子发育各时期 *CoDGK3* 表达变化

Fig. 4 *CoDGK3* expression of *C.oleifera* seeds in different period

3 结 语

在油茶种子各时期转录组数据分析基础上,克隆获得了 *CoDGK3* cDNA 的序列,并对 *CoDGK3* 序列及基因在油茶种子发育过程中不同时期的表达进行了分析,为油茶抗性机制研究及油茶遗传改良提供了依据。油茶转录组数据分析显示油茶中的

DGK 基因表现出多基因家族特征,19 条 Unigene 序列分属于 3 个 *DGK* 基因。后期的工作一方面构建油茶 *CoDGK3* 基因不同的表达载体,利用模式生物进一步确定油茶 *CoDGK3* 基因的功能及表达模式;另一方面继续完成 *DGK* 基因家族其他成员的克隆并对其进行功能及基因表达的研究。

参考文献:

- [1] Meijer H J G, Munnik T. Phospholipid-based signaling in plants[J]. **Annual Review of Plant Biology**, 2003, 54(1): 265–306.
- [2] Laxalt A M, Munnik T. Phospholipid signalling in plant defence[J]. **Current Opinion in Plant Biology**, 2002, 5(4): 332–338.
- [3] Munnik T. Phosphatidic acid: an emerging plant lipid second messenger[J]. **Trends in Plant Science**, 2001, 6(5): 227–233.
- [4] 周长富, 姚小华, 林萍, 等. 基于转录组测序对油茶角鲨烯合酶基因的克隆及分析[J]. 广西植物, 2013, 33(6): 852–858.
ZHOU Changfu, YAO Xiaohua, LIN Ping, et al. Squalene synthase gene cloning and analysis of *Camellia oleifera* based on RNA-Seq[J]. **Guihaia**, 2013, 33(6): 852–858. (in Chinese)
- [5] CAO Z, HU J, SHU Q, et al. Effects of water stress on physical characteristics and survival rate of container seedling in *Camellia oleifera*[J]. **Nonwood Forest Research**, 2011, 4: 13.
- [6] 杨光道. 油茶品种对炭疽病酶活性的生理生化抗性研究[J]. 安徽农学通报, 2011, 17(9): 19–22.
YANG Guangdao. Resistance Mechanism of oil Tea Tree Cultivars to *Colletotrichum Gloeosporioides* [J]. **Anhui Agri Sci Bull**, 2011, 17(9): 19–22. (in Chinese)
- [7] Guoying Z. Resistance to anthracnose in *Camellia oleifera* induced by antagonistic bacteria[J]. **Forest Pest and Disease**, 2010, 3: 2.
- [8] Zhang X, Yang G, Yang J, et al. Physiological mechanism of resistance to anthracnose of different *Camellia* varieties [J]. **African Journal of Biotechnology**, 2014, 11(8): 2026–2031.
- [9] 江南, 谭晓风, 张琳. 基于 RNA-Seq 的油茶种子油脂合成过程磷脂酰肌醇信号系统的基因分析[J]. 食品与生物技术学报, 2014, 33(6): 583–589.
JIANG Nan, TAN Xiaofeng, ZHANG Lin. Gene Analysis of Phosphatidylinositol signaling system in Fatty Acide Developing of *Camellia Oleifera* Seeds Based on RNA-Seq [J]. **Journal of Food Science and Biotechnology**, 2014, 33 (6): 583–589. (in Chinese)
- [10] Mortazavi A, Williams B A, McCue K, et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq [J]. **Nature Methods**, 2008, 5(7): 621–628.