

不同发育时期油桐种仁油体蛋白质组分析

张启燕^{1,2}, 高 暝¹, 吴立文¹, 汪阳东¹, 陈益存^{1*}

(1. 中国林业科学研究院亚热带林业研究所, 浙江 杭州 311400; 2. 南京林业大学, 江苏 南京 210037)

摘要: [目的] 对油桐种仁不同生长时期的油体蛋白进行定量蛋白质组研究, 鉴定油桐种仁油体蛋白质的组成, 探讨不同生长时期种仁油体蛋白质组的表达差异。 [方法] 采用同位素标记相对和绝对定量 (iTRAQ) 技术提取 3 个不同生长时期的油桐种仁油体蛋白质组, 质检合格的蛋白经酶解、iTRAQ 标记后进行高效液相、质谱检测。利用 Mascot 软件对转换的二级质谱数据进行蛋白质鉴定, 通过 iTRAQ 定量寻找差异表达蛋白, 并对差异蛋白进行 GO 功能注释、KEGG 代谢通路和聚类分析。 [结果] 在 3 个生长时期成熟油桐种仁油体亚细胞器中共鉴定出 5 632 个蛋白, 其中, 2 639 个具有催化活性, 401 个蛋白质参与脂肪酸及油脂的合成与代谢。在种仁成熟的 3 个生长时期中, 共有 3 430 个油体蛋白的表达量发生了显著变化。油体差异蛋白的富集分析发现, 具有催化活性这一分子功能的差异蛋白主要在油脂积累初期发生富集。在桐油和桐酸积累时期, 参与脂肪酸代谢和油脂贮藏的酰基辅酶 A 氧化酶、DGA 蛋白和油质蛋白 (Oleosin) 表达量显著上升, 表明这些蛋白在调控桐油和桐酸的积累中可能发挥着关键作用。 [结论] 获得油桐种仁成熟过程中动态特异的油体蛋白质组表达谱, 揭示油体蛋白质组的组成、主要调控桐油合成的关键酶和结构蛋白, 为进一步研究油脂的合成与油体的装配等相关机制奠定了基础。

关键词: 油桐; 油体蛋白质组; iTRAQ; 比较蛋白质组

中图分类号: S794.3

文献标识码: A

文章编号: 1001-1498(2019)01-0047-11

油桐 (*Vernicia fordii* (Hemsl.) Airy Shaw) 属于大戟科 (Euphorbiaceae) 油桐属 (*Vernicia* Lour.) 植物, 是世界著名的木本油料树种。油桐原产我国, 对其栽培、利用已有约 2 300 余年历史, 主要分布在长江中下游地区, 四川、贵州、湖南和广西为主要产区。目前, 世界各地栽培的油桐品种均自我国引进。桐油是从油桐种子中提炼的一种优良的干性油, 桐油具有干燥快、密度轻、耐酸碱、防湿、防腐、防锈等特点, 具有广泛的工业用途, 目前广泛应用于制造涂料、高级油墨、增塑剂、医药以及化学试剂等^[1-2]。同时油桐也是我国重要的生物质能源之一, 其产生的桐油是可以直接利用的优质燃料油。

油桐种仁中, 桐油主要以甘油三酯 (Triacylglycerols, TAG) 的形式形成分散的、相对稳定的亚细胞微滴, 这些小的不连续的亚细胞微滴叫“油体 (Oil body)”^[3-4]。几乎在所有以油脂为主要形式进行营养物质储藏的植物组织中均有油体的存在。作为植物体中最小的细胞器, 油体的直径一般为 0.5 ~ 2.5

μm , 主要包括内部的 TAG 液态基质和外部的磷脂单分子层 (PL) 及蛋白质组成的“半单位”膜^[5]。油体膜及膜上附着的蛋白占油体质量的 1% ~ 4%^[6], 这些油体蛋白中约 90% 为油质蛋白 (Oleosin)^[5], 它们与其它少量 Caleosin 蛋白共同参与油体的形成、稳定及油脂代谢, 并在油体结构的建成中发挥关键作用^[7-8]。目前, 对油体的研究主要集中在生物合成^[9]、油体组装^[10]、油体结合蛋白的成分和结构功能^[11], 以及代谢工程^[12]等方面。

油桐种子中含有丰富的脂肪酸, 特有的桐酸具有较高的工业经济价值, 在油桐种子中的含量高达 80%^[13], 这些脂肪酸储存在胚乳细胞的油体当中。因此, 利用同位素标记相对和绝对定量 (isobaric tags for relative and absolute quantitation, iTRAQ) 技术^[14-15], 对不同发育时期油桐种仁的油体蛋白质组进行研究, 有助于了解种仁油体蛋白的组成以及脂肪酸合成的动态变化过程。本文研究结果, 首次从亚细胞器水平揭示调控特殊油脂“桐油”生物合成

收稿日期: 2018-03-06

基金项目: 浙江省自然科学基金重点项目 (LZ13C160001)

* 通讯作者: 陈益存. E-mail: yicun_chen@163.com

的关键调控酶及结构蛋白,为提高桐油产量、改善桐油特性的定向遗传改良提供科学依据。

1 材料与方 法

1.1 材 料

选取浙江杭州市富阳区黄公望森林公园虎山林区的5年生油桐单株,于2016年8月15日、9月10日和9月26日收集油桐果实,分别命名为OB1、OB2和OB3,剥取种仁,液氮速冻,-70℃冰箱保存备用。

1.2 油桐种仁油体蛋白的提取

1.2.1 油桐种仁油体的提取 参照 Vesna 等油体提取方法^[16]。取5g油桐种仁,在15mL预冷的匀浆液GM I (1 mmol · L⁻¹ EDTA, 10 mmol · L⁻¹ KCl, 1 mmol · L⁻¹ MgCl₂, 2 mmol · L⁻¹ DTT, 0.6 mmol · L⁻¹ Sucrose, 0.15 mol · L⁻¹ Teicine-KOH, pH 7.5)中研磨成匀浆;加入15mL FM I (0.4 mmol · L⁻¹ Sucrose,其它成分同GM I)混匀;10 000 g离心30 min,取上层油层;将油层重悬于15 mL GM II (GM I + 2 mol · L⁻¹ NaCl),加入15 mL FM II (FM I + 2 mol · L⁻¹ NaCl)混匀;10 000 g,30 min,取上层油层;将油层再重悬于15 mL GM I,加入15 mL FM I;10 000 g,30 min,取上层油层;室温重悬于15 mL GM II,震荡30 min;10 000 g,30 min,取上层油层;将油层悬浮于15 mL GM I (4℃预冷),加15 mL FM I (4℃预冷);10 000 g,30 min,取上层油层;最终油层悬浮于3 mL GM I,4℃保存备用。

1.2.2 油体蛋白质组的提取 取0.5 mL油体,加入0.75 mL 氯仿/甲醇(2:1,v/v),振荡;13 000 g,5 min;富含蛋白的中间相重悬于0.25 mL水,加入0.75 mL 氯仿/甲醇(2:1,v/v),振荡;13 000 g,5 min;重复加入0.75 mL 氯仿/甲醇(2:1,v/v),振荡洗涤2次以上;最后中间层重悬于0.5 mL水,超声波处理5 min;加入4倍体积的冷丙酮,于-20℃沉淀16 h以上。取样品于10 000 g,4℃离心15 min,加入适量PBS充分混匀,转移沉淀至新的离心管;于5 000 g,4℃离心5 min,弃上清;每管加入适量无SDS L3,终浓度为1 mmol · L⁻¹的PMSF,2 mmol · L⁻¹的EDTA,置于冰上5 min后加入终浓度为10 mmol · L⁻¹的DTT,冰浴超声5 min,25 000 g,4℃离心15 min,取上清;加终浓度10 mmol · L⁻¹ DTT,56℃水浴1 h;加终浓度55 mmol · L⁻¹ IAM,暗室放置45 min;于25 000 g,4℃离心15 min,上清即为蛋白质溶液。

1.3 油桐种仁油体蛋白质组 iTRAQ 实验

1.3.1 蛋白质提取质控 Bradford 定量:在96孔酶

标板 A1 至 A10 位置依次加入标准蛋白(0.2 μg · μL⁻¹ BSA)0、2、4、6、8、10、12、14、16、18 μL,再依次加纯水 20、18、16、14、12、10、8、6、4、2 μL,之后各孔加考马斯亮蓝 G-250 定量工作液 180 μL。用 SpectraMax iD3 酶标仪(美谷分子仪器(上海)公司)测量 OD₅₉₅,依据 OD₅₉₅与蛋白浓度制作线性标准曲线。稀释待测蛋白质溶液若干倍,在20 μL蛋白溶液中加入180 μL定量工作液,读取 OD₅₉₅。依据标准曲线以及样品 OD₅₉₅计算样品蛋白浓度。

SDS-PAGE:每个样品取10 μg蛋白溶液,加入适量 Loading buffer 混匀后95℃加热5 min,25 000 g离心5 min,取上清点入12% SDS 聚丙烯酰胺凝胶的点样孔中。120 V恒压电泳2 h;电泳结束后,考马斯亮蓝染色2 h,之后加入适量脱色液(40%乙醇,10%醋酸)置于摇床脱色液3~5次,每次30 min。

1.3.2 蛋白酶解与肽段标记 每个样品取100 μg蛋白溶液;按蛋白:酶=40:1的比例加 Trypsin 酶2.5 μg,37℃酶解4 h;按上述比例再补加 Trypsin 1次,37℃继续酶解8 h;酶解的肽段利用 Strata X 柱进行除盐,真空抽干。

根据样品数量,取出一定量 ITRAQ 标签试剂;待试剂恢复至室温后,每管试剂加50 μL异丙醇,涡旋震荡后低速离心;用0.5 mol · L⁻¹ TEAB 溶解肽段样品,并加到对应 ITRAQ 标签试剂中。不同样品肽段选用不同的 ITRAQ 标签;室温静置2 h。

1.3.3 肽段分离 采用 LC-20AB 液相色谱仪(日本岛津公司),分离柱为5 μm 4.6 × 250 mm Gemini C18 柱对样品进行液相分离。用2 mL流动相 A (5% CAN, pH9.8)复溶抽干的肽段样品并进样,以1 mL · min⁻¹的流速梯度洗脱:5% 流动相 B (95% CAN, pH9.8)10 min,5%至35% 流动相 B 40 min,35%至95% 流动相 B 1 min,流动相 B 持续3 min,5% 流动相 B 平衡10 min。在214 nm波长下监测洗脱峰并每分钟收集组分,结合色谱洗脱峰图合并样品得到20个组分,然后冷冻抽干。

1.3.4 高效液相 将肽段样品用流动相 A (2% ACN,0.1% FA)复溶,20 000 g离心10 min后,取上清进样。通过 LC-20AD 纳升液相色谱仪(日本岛津公司)进行分离。样品首先进入 trap 柱富集并除盐,随后与自装 C18 柱(75 μm 内径,3.6 μm 柱料粒径,15 cm 柱长)串联,以300 NL · min⁻¹流速通过如下有效梯度进行分离:0~8 min,5% 流动相 B (98% ACN,0.1% FA);8~43 min,流动相 B 从8%线性

升至 35%; 43 ~ 48 min, 流动相 B 从 35% 升至 60%; 48 ~ 50 min, 流动相 B 从 60% 升至 80%; 50 ~ 55 min, 80% 流动相 B; 55 ~ 65 min, 5% 流动相 B。纳升液相分离末端直接连接质谱仪。

1.3.5 质谱检测 经过液相分离的肽段进入到 ESI 串联质谱仪: TripleTOF 5600 (SCIEX, Framingham, MA, USA), 离子源为 Nanospray III source (SCIEX, Framingham, MA, USA), 放射器为石英材料控制的喷针 (New Objectives, Woburn, MA, USA)。数据采集时, 质谱仪的参数设置如下: 离子源喷雾电压 2 300 V, 氮气压力为 30 psi, 喷雾气流量为 15, 喷雾接口温度 150℃。采用高灵敏度模式进行扫描, 一级质谱扫描累积时间为 250 ms, 扫描质量范围为 350 ~ 1 500 Da。基于扫描信息, 按照一级谱图中的离子强度从高到低, 选择强度超过 150 cps 的前 30 个进行碎裂并扫描二级信息, 筛选标准如下: (1) m/z 扫描范围为 350 ~ 1 250 Da; (2) 电荷数目为 2 ~ 5 个电荷; (3) 母离子动态排除设置为: 在一半的出峰时间内 (约 12 s), 相同母离子的碎裂不超过 2 次。二级质谱的扫描累积时间为 100 ms。针对 ITRAQ 类型的数据采集, 碎裂能量选择根据 ITRAQ 试剂调整, 第 4 级杆 Q2 在 100 Da 时的离传输效率为 100%。

1.4 油桐种仁油体蛋白质组 iTRAQ 数据分析

1.4.1 数据转换和蛋白质鉴定 原始质谱数据经过 ProteoWizard 工具 msConvert 转换成包含二级质谱 (MS/MS) 谱图信息的 mgf 格式文件后, 用蛋白质鉴定软件 Mascot 与 NCBI nr、SwissProt/UniProt、TrEMBL 数据库进行比对搜索得到最终的蛋白鉴定结果^[17]。最终选定的可信蛋白必须包含至少一个可信的特异性 (Unique) 肽段, 同时进行质控分析以判断本次数据是否合格。

1.4.2 iTRAQ 定量分析 利用 IQuant 软件进行 ITRAQ 数据的定量分析^[18]。IQuant 的工作流程主要包括以下几个步骤: 蛋白质过滤, 报告基团标签纯度校正, 定量值归一化, 缺失值补全, 蛋白定量值计算和统计检验分析。首先在谱图/肽段水平进行 1% FDR 的过滤 (PSM-level FDR \leq 0.01), 从而获得显著性鉴定的谱图和肽段列表。接着基于“简约原则”利用肽段进行蛋白组装并产生一系列蛋白组。为了控制蛋白的假阳性率, 在蛋白水平上以 FDR 1% 再次进行过滤 (Protein-level FDR \leq 0.01)。同时, 利用定量重复性分析对本次定量数据的质量进行判断。

1.4.3 差异蛋白 GO 富集分析 对鉴定蛋白进行

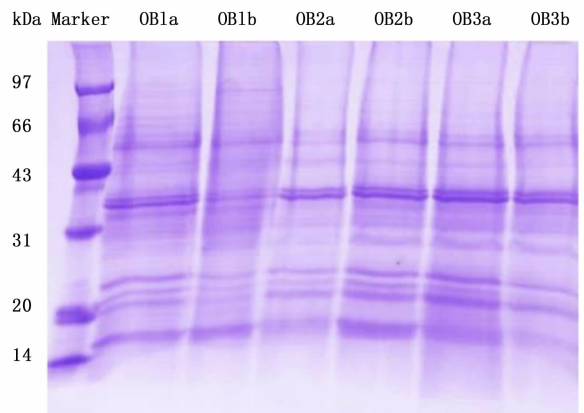
GO、Pathway、COG 等功能注释, 差异蛋白进行 GO、Pathway 富集分析。以 P -value $<$ 0.05 和 FDR $<$ 0.01 两个条件对差异蛋白进行筛选, 从定量结果中筛选出更为关心的显著差异蛋白。在差异蛋白的 GO 富集分析中, 通过显著差异蛋白和作为背景的全体鉴定蛋白相比, 利用超几何检验找出显著富集的 GO 条目; 差异蛋白的 Pathway 富集分析原理类似于此。

2 结果与分析

2.1 油体蛋白质定量

经作者对 7—10 月份的油桐种仁中桐油和桐酸的变化趋势分析, 发现 8 月中旬到 9 月下旬是桐油的快速积累时期, 且在 8 月 15 日、9 月 10 日和 9 月 26 日 3 个时期中, 桐油和桐酸的含量具有明显差别, 因此, 选取生长于这 3 个时期的油桐果实进行油体蛋白质组的提取, 每个时期进行 2 次重复实验, 共 6 组样品, 分别命名为: OB1a、OB1b、OB2a、OB2b、OB3a、OB3b。根据 Bradford 定量方法, 得到的定量标准曲线为 $y = 2.7467x + 0.0316$, $R^2 = 0.9902$ 。根据此标准曲线, 提取的油体蛋白质组的浓度分别为 0.72、1.19、0.31、0.53、0.61 和 0.63 $\mu\text{g} \cdot \mu\text{L}^{-1}$ 。

经过 SDS-PAGE 分离后, 油体蛋白质组样品表现出优良的完整性, 电泳检测结果显示: 分子量大小在 14 ~ 97 kDa 之间的蛋白质均有分布 (图 1)。且 3 个时期的 2 个重复样品之间具有良好的重复性, 说明提取的油体蛋白质组的质量较高, 满足 iTRAQ 定量的实验条件。



注: 分离胶的浓度为 12%, Marker 上样量为 10 μg , 样品上样量为 10 μg 。

Note: The concentration of separating gel is 12%, the loading quantity of marker is 10 μg , the loading quantity of samples is 10 μg .

图 1 不同成熟时期油桐种仁油体蛋白 SDS-PAGE 图谱
Fig. 1 SDS-PAGE of the proteins of kernel oil body in different stages of *V. fordii*

2.2 蛋白质鉴定

检测的6组样品共生成382 390张二级谱图。在“1% FDR”的过滤标准下,共有24 909条肽段和5 632个蛋白得到鉴定(表1),其中,油体蛋白质的分子量大小为3.42~490.53 kDa,蛋白质分子量分布图显示分子量大小在10~60 kDa之间的蛋白质较多,占有检测蛋白的59%(图2a)。肽段长度

分布图显示:鉴定到的肽段长度主要集中在7~17 aa之间(图2b)。特异性肽段数目分布图显示:在鉴定到的5 632个蛋白中,特异性肽段数目大于1的蛋白有3 265个,约占总蛋白的58%(图2c)。相应地,特异性肽段的值越大,对应的特异性谱图的值也较大(图2d)。这些结果表明检测到的蛋白可信度较高。

表1 油桐种仁油体蛋白质鉴定结果

Table 1 Overview of the identification results of oil body proteins in seed kernels of *V. fordii*

| 样品名称 Sample name | 总谱图数目 Total spectra | 谱图数目 Spectra | 特异谱图数目 Unique spectra | 总肽段数目 Peptide | 特异肽段数目 Unique peptide | 蛋白数目 Protein |
|---------------------|------------------------|-----------------|--------------------------|------------------|--------------------------|-----------------|
| 油桐 <i>V. fordii</i> | 382 390 | 58 045 | 46 383 | 24 909 | 21 665 | 5 632 |

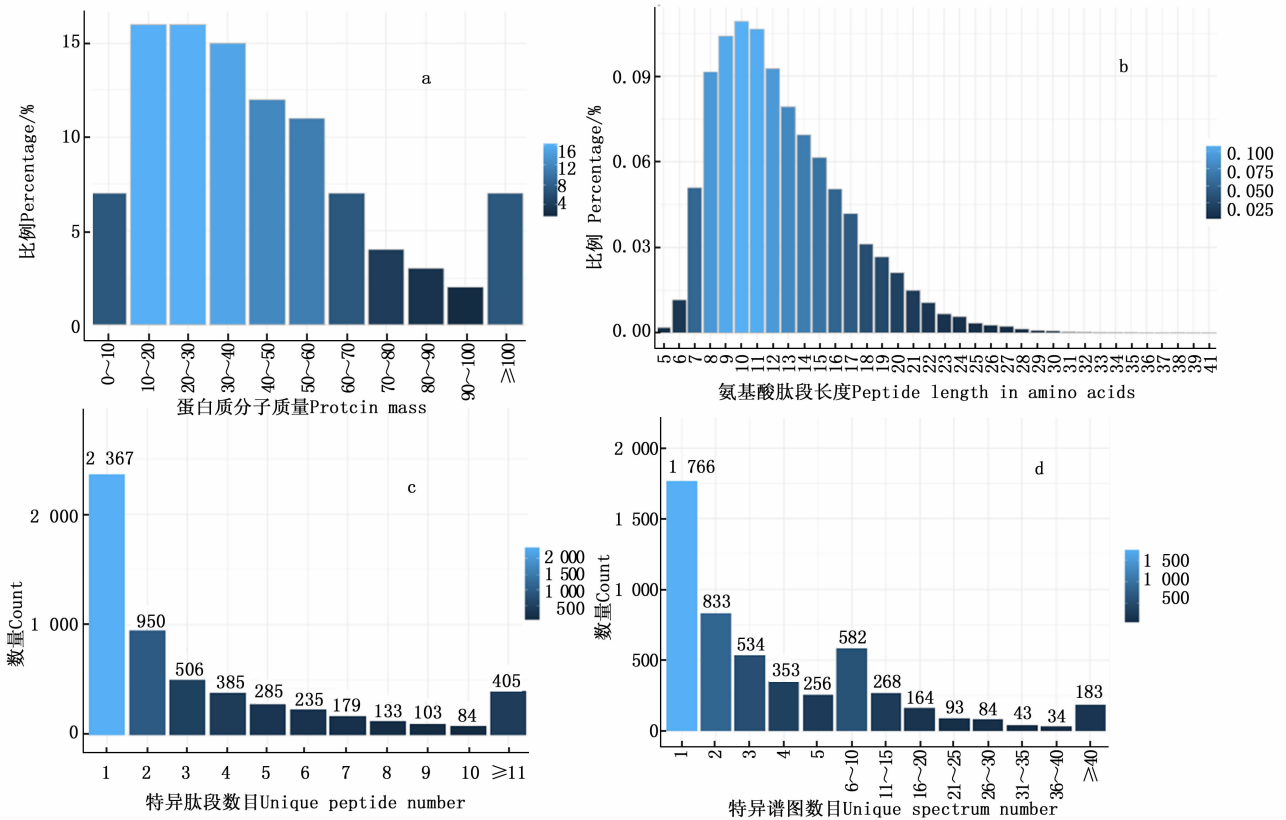


图2 油桐种仁油体蛋白的基本情况

Fig. 2 The basic information of the identified proteins of the kernel oil body of *V. fordii*

2.3 蛋白表达功能聚类

GO注释主要是对蛋白质的生物过程(Biological process)、细胞组分(Cellular component)和分子功能(Molecular function)进行分析。利用Blast2GO软件将所有鉴定到的蛋白质与NR数据库进行比对,得到相应的GO功能注释(图3)。在生物过程方面,注释结果显示5 632个蛋白主要在催化活性(Catalytic activity)和结合活性(Binding)中富集,可以通

过与ATP、GTP、DNA、RNA和金属离子的结合来发挥功能;在细胞组分方面,蛋白主要富集在细胞(Cell)、膜(Membrane)和细胞器(Organelle)中;在分子功能方面,鉴定到的蛋白主要参与代谢过程(Metabolic process)、细胞过程(Cellular process)和单细胞过程(Signal-organism process),它们的占比分别为46.5%、41.3%和20.1%。

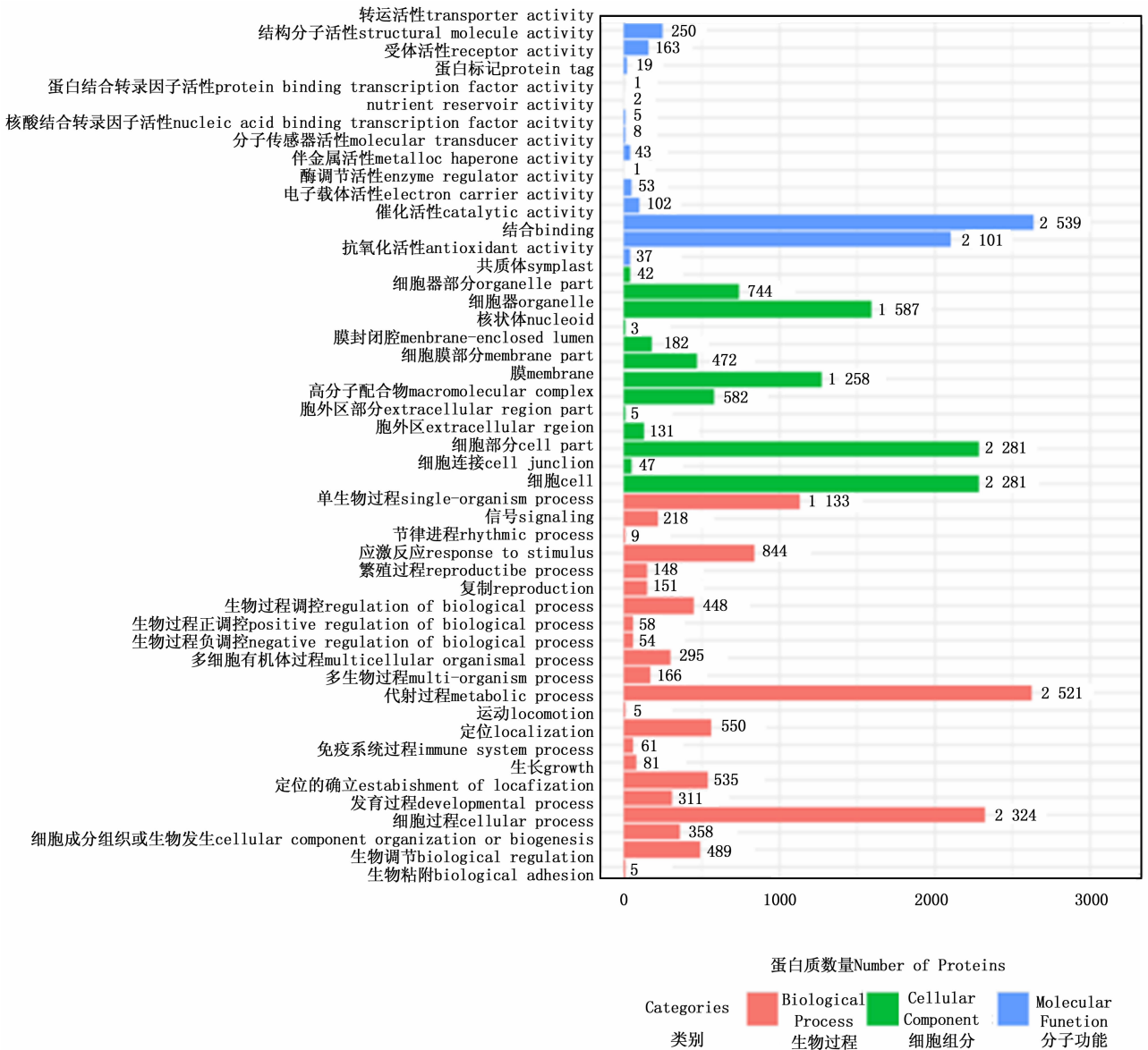


图 3 油桐种仁油体蛋白的 GO 功能注释

Fig. 3 The GO function annotation analysis of the identified proteins in kernel oil body of *V. fordii*

在鉴定到的 5 632 个蛋白中,有催化活性的蛋白 2 639 个,其中,包括脂肪酸合成和油脂合成的一系列酶,包括乙酰-CoA 羧化酶、油 3-磷酸酰基转移酶(GPAT)、酰甘油酰基转移酶(DGAT)、基-CoA 合成酶(FATP1)、脂酰-ACP 还原酶(KAR)、脂二酰甘油酰基转移酶 2(PDAT)、脂酰-ACP 去饱和酶、酮脂酰-ACP 合成酶、基-ACP 硫酯酶、4-二烯酰-CoA 还原酶、3-羟基异丁酸酯脱氢酶、3-羟酰-CoA 脱氢酶、丙二酰-CoA 脱羧酶、β-羟酰基酰基载体蛋白质脱水酶、类固醇脱氢酶、烯酰-CoA 水合酶、长链脂肪酸 CoA 连接酶、质体 1-酰基甘油磷酸

酯酰基转移酶等;其次,还包括脂转运和代谢相关的蛋白,包括磷脂酶、酰基载体蛋白、脂肪酸脱氢酶、鲨烯合酶、羧酸酯酶、3-氧酰基-酰基载体蛋白]合酶、鞘氨醇激酶、酰基-CoA 氧化酶、酰基-CoA 脱氢酶、短链脱氢酶、乙酰-CoA 羧化酶羧基转移酶、1-脱氧-D-木酮糖 5-磷酸合酶等,其中,乙酰-CoA 羧化酶是脂肪酸合成途径的限速酶,甘油 3-磷酸酰基转移酶(GPAT)和二酰甘油酰基转移酶(DGAT)是油脂合成途径中的关键酶,它们的调控可能增强油脂的积累;同时,在 5 632 个蛋白中共鉴定出 4 个油质蛋白(Oleosin)(Unigene27814_All、

Unigene22642 _ All、CL11909. Contig2 _ All、Unigene116694 _ All, 分子量分别为 16.66、14.01、17.81、9.07 kD)。鉴定到的蛋白质涉及了桐油脂生物合成、代谢和油体组装过程的主要酶,为进一步研究林木特殊油脂的合成奠定了基础。

KEGG 注释结果显示:4 541 个油体蛋白质在

KEGG 数据库中有功能注释,共参与了 135 条代谢通路,其中,最主要的代谢通路是代谢途径和次级代谢物生物合成,分别有 1 292 (28.45%) 和 752 (16.56%) 个蛋白参与。另外,有 401 个非冗余(598 个冗余)的蛋白质参与了脂质的合成、代谢和转运等 27 条代谢途径(表 2)。

表 2 油桐种仁油体蛋白的 KEGG 通路分析

Table 2 KEGG pathway annotation analysis of the identified proteins of the kernel oil body

| 路径 Pathway | 路径 Pathway | 有路径注释的蛋白 Proteins with pathway annotation (4 541) | 路径 ID Pathway ID |
|---------------------|--|---|---------------------|
| 代谢途径 | Metabolic pathways | 1 292 (28.45%) | ko01100 |
| 次级代谢物生物合成 | Biosynthesis of secondary metabolites | 752 (16.56%) | ko01110 |
| 脂肪酸代谢 | Fatty acid metabolism | 59 (1.30%) | ko01212 |
| 脂肪酸合成 | Fatty acid biosynthesis | 41 (0.90%) | ko00061 |
| 脂肪酸降解 | Fatty acid degradation | 37 (0.81%) | ko00071 |
| α -亚麻酸代谢 | alpha-Linolenic acid metabolism | 31 (0.68%) | ko00592 |
| 不饱和脂肪酸生物合成 | Biosynthesis of unsaturated fatty acids | 23 (0.51%) | ko01040 |
| 亚油酸代谢 | Linoleic acid metabolism | 11 (0.24%) | ko00591 |
| 脂肪酸伸长 | Fatty acid elongation | 8 (0.18%) | ko00062 |
| 甘油酯代谢 | Glycerolipid metabolism | 58 (1.28%) | ko00561 |
| 脂质转移 | ABC transporters | 32 (0.70%) | ko02010 |
| 萜类骨架生物合成 | Terpenoid backbone biosynthesis | 35 (0.77%) | ko00900 |
| 单萜生物合成 | Monoterpenoid biosynthesis | 3 (0.07%) | ko00902 |
| 倍半萜和三萜生物合成 | Sesquiterpenoid and triterpenoid biosynthesis | 5 (0.11%) | ko00909 |
| 二萜类生物合成 | Diterpenoid biosynthesis | 16 (0.35%) | ko00904 |
| 花生四烯酸代谢 | Arachidonic acid metabolism | 28 (0.62%) | ko00590 |
| 柠檬烯和蒎烯降解 | Limonene and pinene degradation | 15 (0.33%) | ko00903 |
| 泛醌和其他萜类醌生物合成 | Ubiquinone and other terpenoid-quinone biosynthesis | 24 (0.53%) | ko00130 |
| 甘油磷脂代谢 | Glycerophospholipid metabolism | 54 (1.19%) | ko00564 |
| 鞘脂代谢 | Sphingolipid metabolism | 25 (0.55%) | ko00600 |
| 类胡萝卜素生物合成 | Carotenoid biosynthesis | 17 (0.37%) | ko00906 |
| 醚脂代谢 | Ether lipid metabolism | 17 (0.37%) | ko00565 |
| 糖基磷脂酰肌醇(GPI)-锚定生物合成 | Glycosylphosphatidylinositol (GPI)-anchor biosynthesis | 11 (0.24%) | ko00563 |
| 鞘糖脂生物合成 | Glycosphingolipid biosynthesis-ganglio series | 8 (0.18%) | ko00604 |
| 糖脂合成 | Glycosphingolipid biosynthesis-globo and isoglobo series | 6 (0.13%) | ko00603 |
| 角质和蜡生物合成 | Cutin, suberine and wax biosynthesis | 5 (0.11%) | ko00073 |
| 油菜甾体生物合成 | Brassinosteroid biosynthesis | 4 (0.09%) | ko00905 |
| 甾体生物合成 | Steroid biosynthesis | 22 (0.48%) | ko00100 |
| 酮体的合成与降解 | Synthesis and degradation of ketone bodies | 3 (0.07%) | ko00072 |

2.4 蛋白质 iTRAQ 定量

将 OB1a/OB2a、OB1a/OB2b、OB1b/OB2a、OB1b/OB2b、OB2a/OB3a、OB2a/OB3b、OB2b/OB3a、OB2b/OB3b、OB1a/OB3a、OB1a/OB3b、OB1b/OB3a、OB1b/OB3b、OB1b/OB1a、OB2b/OB2a、OB3b/OB3a

这 15 个组合以及 OB1/OB2、OB1/OB3、OB2/OB3 这 3 个组合设置为比较组,利用 IQuant 软件对鉴定到的蛋白质进行定量。2 次重复实验的最终显著差异蛋白以 Fold change > 1.2(单次实验比较组比值的平均值)和 P-value < 0.05 两个条件筛选,最终得到的

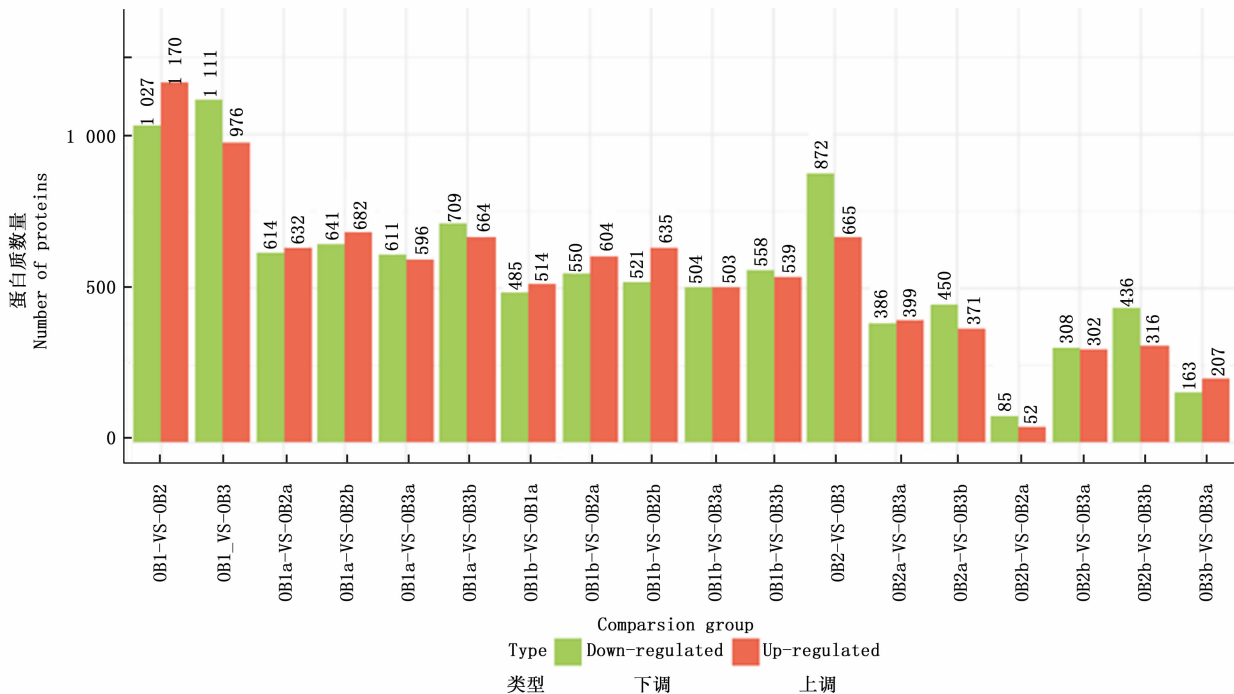


图4 油桐种仁油体蛋白不同比较组间的显著差异蛋白

Fig. 4 The differential expressed proteins of the kernel oil body in different comparison groups

显著差异蛋白见图4。

在油桐种仁成熟的OB1、OB2和OB3这3个生长时期中,共有3430个油体蛋白的表达量发生了显著变化。OB1-vs-OB2中有2197个差异蛋白,其中,1170个蛋白质上调,略多于下调蛋白质的数目。与OB1-vs-OB2相比,OB2-vs-OB3的差异蛋白数目(1537)相对较少,其中,56.7%(872)的蛋白下调(图4)。OB1-vs-OB2和OB2-vs-OB3的共同差异蛋白有712个,在种仁成熟的这3个生长时期中,表达量连续升高的蛋白有125个,表达量连续降低的蛋白有48个(图中未列出)。

2.5 蛋白定量重复性分析

利用变异系数(CV)来评估定量的重复性。当越小的CV值覆盖到越多的定量蛋白时,认为其重复性越好。图5横坐标为CV值,纵坐标为小于CV值的蛋白数量占整个定量蛋白数量的百分比。如图5a所示:当 $CV < 0.4$ 时,OB1-vs-OB2可覆盖到95%的定量蛋白。如图5b所示:当 $CV < 0.3$ 时,OB2-vs-OB3可覆盖到94%的定量蛋白。如图5c所示:当 $CV < 0.4$ 时,OB1-vs-OB3可覆盖到95%的定量蛋白。说明2次重复实验的重复性比较好。

2.6 差异蛋白的富集分析

OB1-vs-OB2、OB1-vs-OB3和OB2-vs-OB3中的

油体差异蛋白在细胞组分和生物过程中主要的GO富集结果是相同的:在细胞组分中,油体差异蛋白主要在细胞、细胞内以及内质网中富集。在生物过程中,差异蛋白主要在代谢过程和细胞转化中富集。但在分子功能方面,3个时期的油体差异蛋白的富集结果却不尽相同。OB1-vs-OB2、OB1-vs-OB3中的差异蛋白主要在催化活性和结合活性中富集;但在OB2-vs-OB3中,差异蛋白却不在催化活性这一分子功能上富集(表3)。

油体差异蛋白的Pathway富集分析发现,OB1-vs-OB2差异蛋白主要的富集通路是RNA转运(RNA transport)、核蛋白体(Ribosome)和吞噬小体(Phagosome)的代谢(图6a)。OB1-vs-OB3差异蛋白主要的富集通路为次生代谢物的生物合成(Biosynthesis of secondary metabolites)、核蛋白体和剪接体(Spliceosome)的代谢(图6b)。OB2 vs. OB3差异蛋白富集的通路较多,主要包括内质网蛋白质加工(Protein processing in endoplasmic reticulum)、丙酮酸代谢(Pyruvate metabolism)、柠檬酸循环(TCA cycle)、苯丙醇生物合成(Phenylpropanol biosynthesis)和不饱和脂肪酸的生物合成(Biosynthesis of unsaturated fatty acids)等(图6c)。

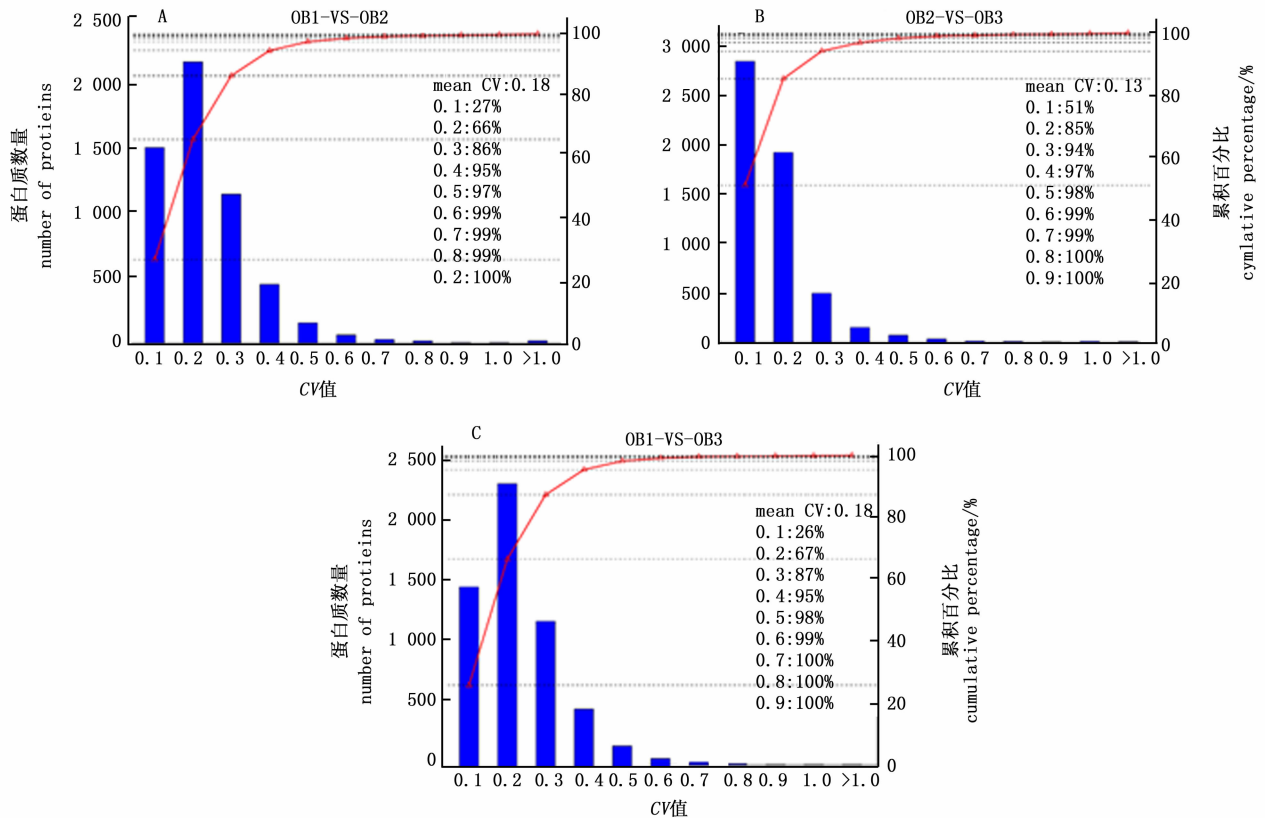


图5 蛋白定量重复实验的变异系数分析

Fig.5 The CV distribution analysis of the repeatability

表3 油体差异蛋白在分子功能方面主要的GO富集条目

Table 3 The main items of molecular function in GO enrichment analysis of differentially expressed proteins

| OB1-vs-OB2 | OB1-vs-OB3 | OB2-vs-OB3 |
|---|---|--|
| 催化活性 Catalytic activity | 催化活性 Catalytic activity | 结合活性 binding activity |
| 结合活性 binding activity | 结合活性 binding activity | 离子结合 Ion binding |
| 杂环化合物结合 Heterocyclic compound binding | 杂环化合物结合 Heterocyclic compound binding | 有机化合物结合 Organic cyclic compound binding |
| 有机环化合物结合 Organic cyclic compound binding | 有机环化合物结合 Organic cyclic compound binding | 杂环化合物结合 Heterocyclic compound binding |
| 离子结合 Ion binding | 离子结合 Ion binding | 小分子结合 Small molecule binding |

2.7 差异蛋白的表达聚类分析

在油桐种仁的3个生长时期中,有401个非冗余蛋白质参与了脂质的合成及代谢途径(表2)。在OB1-vs-OB2比较组中,有135个脂质合成与代谢相关的蛋白质的表达量发生了显著性变化,其中有44个具有功能注释的蛋白质参与了脂肪酸的合成与代谢。对这44个蛋白质和其它4个油质蛋白(Unigene27817_All、Unigene22642_All、Uingene116694_All、CL11909.Contig2_All)进行了聚类分析(图7)。

图7表明:参与脂肪酸代谢的酰基辅酶A氧化

酶(CLI4770.Contig1_All、CL25395.Contig1_All)和DAG蛋白(CL23974.Contig1_All)以及参与油质贮藏的4个油质蛋白的表达量在油桐种仁3个不同发育时期中均有显著变化。这几个差异蛋白的表达量在这3个生长时期表现出先升后降的趋势。相比于OB1时期,它们的表达量在OB2时期显著升高;而相比于OB2时期,它们的表达量在OB3时期又显著下降。

3 讨论

我国是世界上最大的桐油生产国,年产量达10

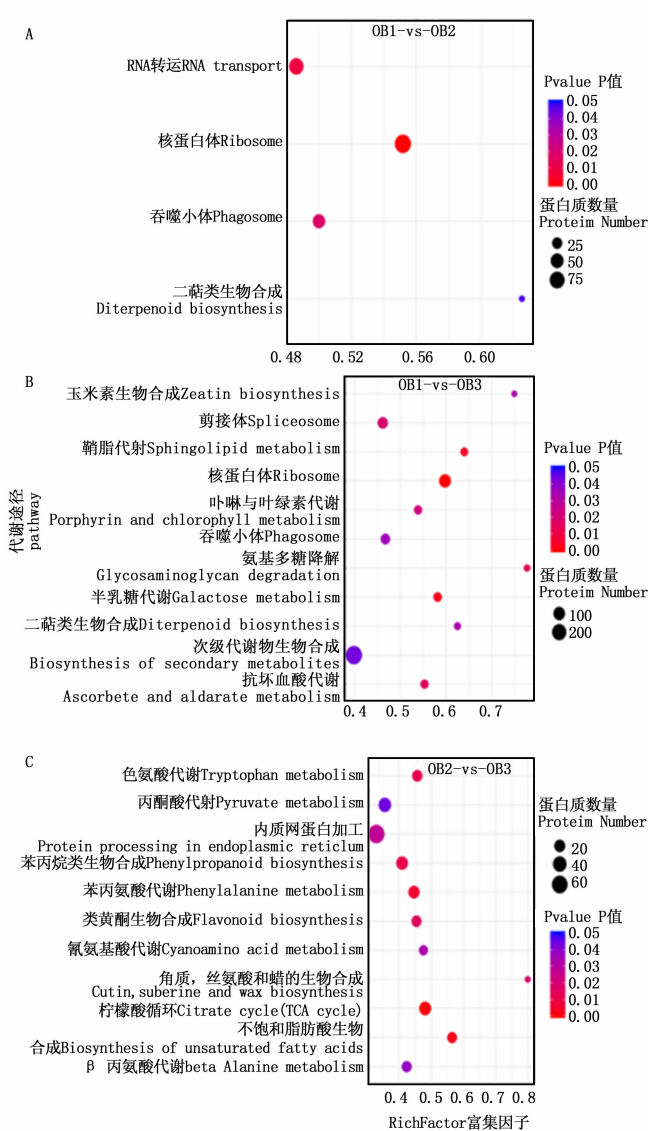


图 6 油桐油体差异蛋白的 KEGG 富集分析

Fig. 6 The KEGG pathway enrichment analysis of the differential expressed proteins of kernel oil body

万 t 以上,占世界桐油产量的 80%^[19]。随着油桐应用领域的不断扩大以及人类对植物油脂需求的日益增长,桐油的产量已经不能满足世界市场的需求。决定桐油性质的桐酸等脂肪酸储存在油体中,因此对不同生长时期种仁油体蛋白质组进行分析可以为油脂合成和油体装配的分子机理研究奠定基础。

作者利用改良的方法对油桐种仁油体蛋白进行了提取,以去除干扰蛋白质提取的多种次生代谢物质,然后采用 iTRAQ 技术,对油桐种仁成熟期 3 个不同生长阶段(8 月初至 9 月末)的油体蛋白进行了定量分析。鉴定得到的 5 632 个油体蛋白的分子量,氨基酸肽段长度和特异性肽段数目的统计分析

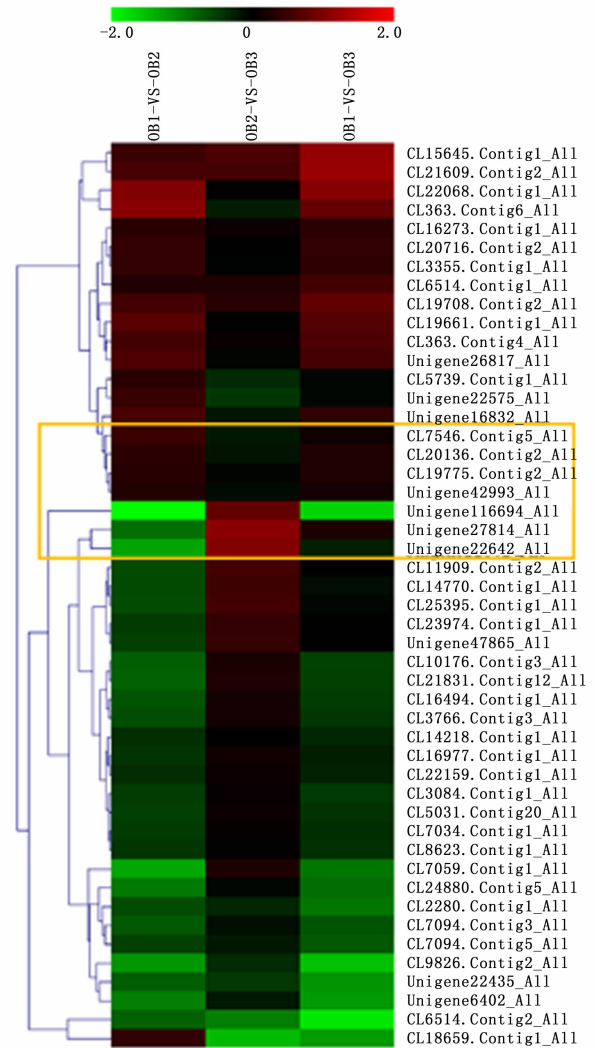


图 7 脂肪酸代谢途径相关差异蛋白的聚类分析

Fig. 7 The cluster analysis of the differential expressed proteins involved in the fatty acids metabolism

以及 iTRAQ 定量的重复性检验均表明测得的数据质量较高。GO 功能注释分析发现,5 632 个油体蛋白中有 2 639 个具有催化活性,这包括催化脂肪酸以及油脂合成与代谢、脂质转运与储存、油体装配的一系列酶。KEGG 通路分析发现,有 401 个非冗余油体蛋白直接参与了与脂质代谢相关的 27 条代谢途径。iTRAQ 定量分析发现,在油桐种仁 3 个不同发育时期中有 3 430 个显著差异表达的蛋白,其中,OB1-vs-OB2 的差异蛋白数目最多,而 OB2-vs-OB3 的差异蛋白数目最少。差异蛋白的 KEGG 富集分析发现,虽然 OB2-vs-OB3 差异蛋白的数目最少,但富集的代谢途径却较多。差异蛋白的 GO 富集分析发现,在 OB1-vs-OB2 中富集的催化活性这一分子功能在 OB2-vs-OB3 中却不富集,这说明第 3 个时期与

第 2 个时期相比,具有催化活性的蛋白的表达量没有发生显著变化。已知具有催化活性的蛋白包含一些调控脂肪酸、油脂合成及代谢的酶,这些酶的表达量的非显著变化也许导致了桐油和桐酸含量积累的减慢。这与油桐桐油及桐酸含量的变化趋势一致,即 8 月中旬至 9 月中旬是油脂积累的高峰期,桐油的含量和脂肪酸中桐酸的含量上升,侧重于脂肪的积累、转化和种胚的成熟^[20];而 9 月下旬,油脂含量几乎不再增加。因此,作者重点关注在 OB1-vs-OB2 中表达量发生显著变化的差异蛋白。在直接参与脂质合成与代谢的 401 个差异蛋白和 4 个油质蛋白中,有 135 个蛋白在这个阶段是显著变化的。脂肪酸代谢途径与脂质相关的 48 个差异蛋白的聚类分析发现,酰基辅酶 A 氧化酶、DAG 蛋白和 4 个油质蛋白的表达量均在桐油和桐酸积累时期显著上升,而在桐油和桐酸稳定时期显著降低,这说明这几个差异蛋白在调控油脂合成、代谢和贮藏中可能发挥着关键作用,进一步证实了油质蛋白是油体结构组成和功能调节最主要的蛋白^[16]。

4 结 论

本试验基于 iTRAQ 定量技术对不同生长时期的油桐种仁油体蛋白质组进行了鉴定及差异分析,研究发现,酰基辅酶 A 氧化酶、DAG 蛋白和 4 个油质蛋白可能在调控油脂合成、代谢与转运以及在油体装配中发挥着关键作用,为阐明木本植物特殊油脂生物合成分子机制和深入实施木本油料植物分子育种提供了新的策略与理论依据。

参 考 文 献:

- [1] 方嘉兴,何 方. 中国油桐[M]. 北京:中国林业出版社,1998.
- [2] 谭晓风,蒋桂雄,谭方友,等. 我国油桐产业化发展战略调查研究报告[J]. 经济林研究,2011,9(3):7-13.
- [3] 吴庆珂,杨素素,汪阳东,等. 油桐油质蛋白 Oleosin 编码基因 5 个 VIOLE 转录本的克隆与表达分析[J]. 林业科学研究,2014,27(2):233-239.
- [4] Murphy D J. Storage lipid bodies in plants and other organisms[J]. Prog Lipid Res, 1990, 29(4): 299-324.
- [5] Tzen J T C, Lie G C, Huang A H C. Characterization of the charged

- components and their topology on the surface of plant seed oil bodies [J]. The Journal of Biological Chemistry, 1992, 267 (22): 15626-15634.
- [6] Anthony H, Huang C. Oil bodies and oleosins in seeds[J]. Annu Rev Plant Physiol Mol Biol, 1992, 43(4): 177-200.
- [7] Frandsen G, Mundy J, Tzen J T, et al. Oil bodies and their associated proteins, oleosin and caleosin[J]. Physiol Plant, 2001, 112(3): 301-307.
- [8] Floriana C, Frederic B, Johnathan A, et al. Properties and exploitation of oleosins[J]. Biotechnol Adv, 2007, 25(2): 203-206.
- [9] Song Y, Wang X D, Rose R J. Oil body biogenesis and biotechnology in legume seeds[J]. Plant Cell Rep, 2017, 36(10): 1519-1532.
- [10] Miquel M, Trigui G, d'Andréa S, et al. Specialization of oleosins in oil body dynamics during seed development in *Arabidopsis* seeds [J]. Plant Physiol, 2014, 164(4): 1866-1878.
- [11] Horn P J, Benning C. The plant lipidome in human and environmental health[J]. Science, 2016, 353(6305): 1228-1232.
- [12] Bates P D. Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis [J]. Biochim Biophys Acta Mol Cell Biol Lipids, 2016, 1861(9): 1214-1225.
- [13] Sonntag N O V. Composition and characteristics of individual fats and oils[M] // Swern D. Bailey's Industrial Oil and Fat Products, Vol. 1. New York: John Wiley & Sons, 1979: 289-477.
- [14] Ross P L, Huang Y N, Marchese J N, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents[J]. Molecular & Cellular Proteomics, 2004, 3(12): 1154-1169.
- [15] Evans C, Noirel J, Ow S Y, et al. An insight into iTRAQ: where do we stand now? [J]. Analytical and Bioanalytical Chemistry, 2012, 404(4): 1011-1027.
- [16] Vesna K, Ganesh K A, Martin H, et al. Protein and lipid composition analysis of oil bodies from two *Brassica napus* cultivars [J]. Proteomics, 2006, 6(16): 4586-4598.
- [17] Brosch M, Yu L, Hubbard T, et al. Accurate and sensitive peptide identification with Mascot percolator. Journal of Proteome Research, 2009, 8(6): 3176-3181.
- [18] Wen B, Zhou R, Feng Q, et al. IQuant: An automated pipeline for quantitative proteomics based upon isobaric tags [J]. Proteomics, 2014, 14(20): 2280-2285.
- [19] 周 冠,汪阳东,陈益存,等. 油桐种仁 cDNA 文库的构建及其油体蛋白 oleosin 基因的生物信息学分析[J]. 林业科学研究, 2009, 22(2): 177-181.
- [20] 汪阳东,陈益存,姚晓华,等. 油桐分子生物学研究[M]. 北京:中国林业出版社,2012.

iTRAQ Quantitative Proteome Analysis of Kernel Oil Body in Different Growth Stages of *Vernicia fordii*

ZHANG Qi-yan^{1,2}, GAO Ming¹, WU Li-wen¹, WANG Yang-dong¹, CHEN Yi-cun¹

(1. Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou 311400, Zhejiang, China;

2. Nanjing Forestry University, Nanjing 210037, Jiangsu, China)

Abstract: [**Objective**] iTRAQ (isobaric tags for relative and absolute quantitation) was performed on the proteome of kernel oil body in different growth stages of *Vernicia fordii* to identify the protein composition of the oil body and to explore the differentially expressed proteins in different growth periods. [**Method**] The proteomes of the oil body were extracted from *V. fordii* seed kernel at three different growth stages. Qualified proteomes were analyzed by LC-MS/MS (Liquid chromatography-mass spectrometry) after protease digestion and iTRAQ labeling. The obtained converted secondary mass spectrometry data were used for protein identification through Mascot software. Differentially expressed proteins were explored through iTRAQ quantification analysis. GO (Gene Ontology) function annotation, KEGG pathway and cluster analysis of differentially expressed proteins identified was further conducted. [**Result**] 5 632 proteins were identified in the proteome of kernel oil body in three growth stages of *V. fordii*. Of them, 2 639 proteins shared catalytic activity and 401 proteins involved in the synthesis and metabolism of fatty acids and fats. In the three growth stages of the kernel oil body, 3 430 proteins were significantly differentially expressed. The GO enrichment analysis showed that the differentially expressed proteins with catalytic activity were mainly enriched in the OB1/OB2 stage rather than in the OB2/OB3 stage. During the accumulation period of tung oil and α -eleostearic acid, the acyl-CoA oxidase, GDA protein and oleosin, which involved in fatty acid metabolism and oil storage, exhibited significantly increased expression. It suggested that these proteins may play a key role in regulating the accumulation of tung oil and α -eleostearic acid. [**Conclusion**] The expression profiles of differentially expressed protein of kernel oil body in different growth stage were analyzed using iTRAQ method. The protein composition of the oil body was identified as well as the metabolic pathway and key functional genes involved fatty acid and lipid metabolism were explored. It is a crucial step for further study on the mechanism of oil synthesis and oil body assembly.

Keywords: *Vernicia fordii*; oil body proteome; iTRAQ; comparative proteome

(责任编辑:金立新)