

## 研究报告

## Research Report

# 人参种胚不同后熟发育阶段比较转录组学分析

冷容<sup>1,2</sup> 胡彦婷<sup>1,2</sup> 周艳丽<sup>1</sup> 金桂花<sup>1,2</sup> 李玲<sup>1,2</sup> 石勇<sup>3</sup> 杨红<sup>1,2</sup> 桂大萍<sup>1</sup> 薛润光<sup>4</sup> 章成君<sup>1\*</sup>

1 中国科学院昆明植物研究所, 中国西南野生生物种质资源库, 昆明, 650201; 2 中国科学院大学生命科学学院, 北京, 100049; 3 中国科学院华南植物园, 广州, 510650; 4 云南省农业科学院高山经济植物研究所, 丽江, 674100

\* 通信作者, zhangchengjun@mail.kib.ac.cn

**摘要** 胚后熟发育对人参种子休眠解除具有重要影响。为探索人参胚后熟发育的分子机制, 本研究利用 Illumina HiSeq X-ten 平台对人参种子五个后熟发育时期(心形胚, 鱼雷胚, 子叶胚, 成熟胚和芽胚)的胚组织进行转录组测序, 共获得 412 604 390 条高质量 reads。以心形胚为对照, 在鱼雷胚、子叶胚、成熟胚和芽胚时期分别鉴定到 2 359 个、3 048 个、6 180 个、12 067 个差异表达基因(DEGs), 去重复后四个时期共获得 13 724 个 DEGs。GO 富集分析显示, 四个时期的和总的 DEGs 主要共同富集到催化活性、细胞和发育过程等 GO 条目。与种子休眠结束中胚发育相关的 DEGs 编码蛋白构成的蛋白互作网络分析显示, 指导 mRNA 合成的 RNA 聚合酶 II 第二大亚基(NRPB2)、脯氨酸合成的关键酶  $\Delta^1$ -吡咯啉-5-羧酸合成酶(P5CS1 和 P5CS2)为核心节点蛋白, 它们可以通过调控 mRNA 合成、脯氨酸合成和与其互作蛋白的表达来影响人参种胚的发育。KEGG 富集分析发现, 四个发育时期的 DEGs 主要共同富集到代谢过程、次生代谢合成和植物激素信号转导通路。其中, 在植物激素信号转导通路中参与生长素和脱落酸信号转导的 DEGs 最多。因此, 我们认为 NRPB2、P5CS1、P5CS2 蛋白互作和植物激素信号转导可能对人参胚的后熟发育至关重要。本研究结果对揭示人参胚的发育和种子休眠解除的分子调控机制有一定的参考意义。

**关键词** 人参种胚, 后熟发育, 转录组

## Comparative Transcriptome Analysis of *Panax ginseng* C. A. Mey. Embryos in Different After-Ripening Development Stages

Leng Rong<sup>1,2</sup> Hu Yanting<sup>1,2</sup> Zhou Yanli<sup>1</sup> Jin Guihua<sup>1,2</sup> Li Ling<sup>1,2</sup> Shi Yong<sup>3</sup> Yang Hong<sup>1,2</sup> Gui Daping<sup>1</sup> Xue Runguang<sup>4</sup> Zhang Chengjun<sup>1\*</sup>

1 The Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, 650201; 2 College of Life Sciences, University of Chinese Academy of Sciences, Beijing, 100049; 3 South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, 510650; 4 Alpine Economic Plant Research Institute, Yunnan Academy of Agricultural Sciences, Lijiang, 674100

\* Corresponding author, zhangchengjun@mail.kib.ac.cn

DOI: 10.13271/j.mpb.018.005638

**Abstract** The after-ripening development of the embryo has an essential influence on the dormancy release of ginseng (*Panax ginseng* C. A. Mey.) seeds. In order to explore the molecular mechanism of the after-ripening development of ginseng embryos, in this study we used the Illumina HiSeq X-ten platform to sequence the transcriptome of ginseng embryos at five different after-ripening development stages: heart-shaped embryo, torpedo-shape embryo, cotyledonary embryo, mature embryo and bud embryo, which generate 412 604 390 clean reads in total. Comp-

基金项目: 本研究由中国科学院百人计划自然科学基金面上项目(31571311)和云南省第三批博士后定向培养基金(39Y73241261)共同资助

引用格式: Leng R., Hu Y.T., Zhou Y.L., Jin G.H., Li L., Shi Y., Yang H., Gui D.P., Xue R.G., and Zhang C.J., 2020, Comparative transcriptome analysis of *Panax ginseng* C. A. Mey. embryos in different after-ripening development stages, *Fenzi Zhiwu Yuzhong* (Molecular Plant Breeding), 18(17): 5638-5649 (冷容, 胡彦婷, 周艳丽, 金桂花, 李玲, 石勇, 杨红, 桂大萍, 薛润光, 章成君, 2020, 人参种胚不同后熟发育阶段比较转录组学分析, 分子植物育种, 18(17): 5638-5649)

aring to heart-shaped embryos, 2 359, 3 048, 6 180 and 12 067 differentially expressed genes (DEGs) are identified in torpedo-shaped embryos, cotyledonary embryos, mature embryos, and bud embryos, respectively. A total of 13 724 DEGs are obtained after removing the duplicates. GO (Gene ontology) enrichment analysis showed that these genes (DEGs in four stages and total) were all significantly enriched into GO terms related to catalytic activity, cell and developmental process. Protein-protein interaction analysis showed that the second largest subunit of RNA polymerase II (NRPB2),  $\Delta^1$ -pyrroline-5-carboxylate synthetase (P5CS1 and P5CS2, key enzyme of proline synthesis) are core nodal proteins in the protein interaction network created by DEGs related to embryo development in seed dormancy ending, and they might affect the development of ginseng seed embryos by regulating mRNA synthesis, proline synthesis and expression of other protein they interact with. KEGG (Kyoto encyclopedia of genes and genomes) enrichment analysis showed that DEGs in four developmental stages are mainly enriched into metabolic processes, biosynthesis of secondary metabolites, and plant hormone signal transduction pathways. In terms of DEGs in the plant hormone signal transduction pathway, most of them involve auxin and abscisic acid signal transduction. Therefore, we infer that NRPB2, P5CS1, P5CS2 protein and plant hormone signal transduction may have an important role in after-ripening development of ginseng embryos. The results of this study have certain reference meaning for revealing the regulation mechanism of ginseng embryo development and seed dormancy release.

**Keywords** Ginseng embryos, After-ripening development, Transcriptome

人参(*Panax ginseng* C. A. Mey.)是五加科(Araliaceae)一种名贵的草本药用植物,在中国已经有几千年的药用历史(Yun, 2001)。人参的根、茎、叶可药食两用,特别是根部组织含有大量的人参皂苷、多糖、氨基酸等化学成分,具有很好的补气、抗衰老、抗癌、降血压和提高免疫力等功效(Qi et al., 2011; Sun, 2011; Rajabian et al., 2019)。因此,一直以来人参都深受人们的喜爱,也是世界上应用最为广泛的中药材之一。

人参主要以种子进行繁殖。刚采收时(果实成熟期)人参种子的胚发育不全,胚长仅 0.2~0.4 mm,只能看到 2 个呈心形的子叶原基,没有完整的胚根、胚轴、胚芽和子叶结构,并且种子具有深度休眠特性(李方元和马淑琴, 1991, 人参研究, (1): 20-23)。根据 Baskin 和 Baskin 等(2004)种子休眠分类理论,人参种子属于复杂的形态生理休眠(Morphophysiological dormancy)类型,即种子同时存在形态和生理发育不全而使种子深度休眠的现象(付长珍和郭宝, 2013, 中国现代中药, 15(10): 856-859)。该类型种子的休眠期较长,给人参种子的萌发造成了很大的阻碍。研究表明,成熟的人参种子在常温条件下播种并不能萌发,种子必须要经过暖温(15℃~25℃)和冷温(0℃~5℃)条件下层积或历经秋冬两季变温催芽 9~22 个月后,胚才能够完成后熟发育(脱离母体的种子,继续发育直至形态和生理完全成熟),种子才能解除休眠而发芽(王化武, 1979, 特产科学实验, (2): 1-3; Chen et al., 1984)。

因此,胚的后熟发育过程对人参种子的萌发具有重要影响。目前,有关人参种子休眠或种子萌发的研究虽然有大量文献报道,但是多集中在植物激素处理、变温处理和生理指标检测等方面(Kim et al., 2014; Rajametov et al., 2014; Lee et al., 2018),对人参种子萌发的分子机制的研究相对较少。特别是与人参种子后熟过程相关的基因、基因调控模式等还尚不清楚。转录组测序是对生物体细胞或组织中所有 RNA 进行高通量测序的方法,基于整个转录水平我们可以了解到不同环境条件下生物体中的基因表达情况并对它们进行分析(Costa et al., 2010)。如 Gao 等(2019)对拟南芥种子发育过程中 7 个不同发育阶段的胚组织进行转录组测序和分析发现,胚发育的早期/中期、晚期和成熟阶段的基因表达分别与甲基化、光合作用启动、能量储存相关蛋白活性调控有关。Qi 等(2015)对不同休眠解除阶段的西洋参种子进行转录组测序,通过分析发现种子休眠解除过程中大量差异表达基因的功能与植物激素特别是 GA 和 ABA 相关调控基因(如 *GA2ox1*, *GA3ox1*, *CYP707A* 等)的功能相似。虽然近年来有关人参的转录组分析也有大量文献报道,但是主要集中在对根部组织的研究,如人参皂苷合成、萜类生物合成等相关基因的挖掘(Fan et al., 2019; Zhao et al., 2019),对人参种子特别是胚后熟发育过程的研究尚未有文献报道。因此,本研究采用前人的变温层积方法使人参种子完成后熟发育,并对不同发育阶段的胚组织进行转录测序和

分析,解析人参胚后熟发育过程中基因调控情况,以期揭示胚发育的分子调控机制,为人参种子解除休眠、快速萌发和长期保存提供一定帮助。

## 1 结果与分析

### 1.1 人参种子后熟过程和测序数据分析

通过观察人参种子后熟发育过程的形态,我们发现刚采收的人参种子的胚乳已经发育完全,而胚大部分处于心形胚阶段,此时种子已经停止了发育,并进入休眠状态。当把种子放入 16℃ 暖温条件下层积后,子叶原基会继续生长,大约 45 d 胚发育成鱼雷形。随后胚体进一步发育,大约 90 d 胚体分化出了胚根、胚轴、胚芽和子叶结构,此时大部分种子呈现于裂口状态,大约 180 d 胚体几乎充满整个胚腔,胚形态发育成熟。随后我们将完成形态后熟的种子再转入 4℃ 冷温条件下层积使胚继续发育,直至第 250 天胚的形态、结构和大小相对 180 d 而言几乎没有变化,层积第 270 天左右胚根开始突破胚乳,种子逐渐发芽(图 1)。然而,在整个胚后熟发育过程中,没有经过冷温处理过的种子几乎不会发芽。

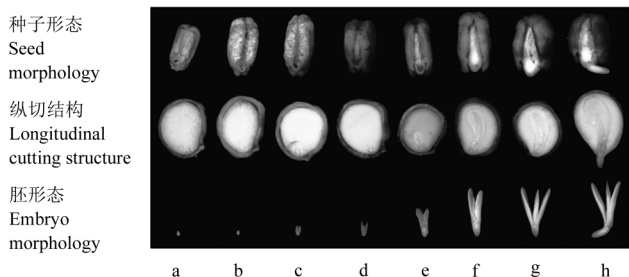


图 1 人参种子后熟发育过程

注: a: 心形胚早期(刚采收时成熟果实); b: 心形胚中期(暖温层积 7 d); c: 鱼雷胚阶段(暖温层积 45 d); d: 子叶胚早期(暖温层积 90 d); e: 子叶胚中期(暖温层积 135 d); f: 子叶胚末期或成熟胚阶段(暖温层积 180 d); g: 成熟胚阶段(冷温层积 250 d); h: 芽胚/发芽阶段(冷温层积 270 d)

Figure 1 The after-ripening development of ginseng (*Panax ginseng* C. A. Mey.) seeds

Note: a: Early heart-shaped embryo (The fully ripe fruit just harvested); b: Mid-term heart-shaped embryo (Warm stratified seeds for 7 days); c: Torpedo-shaped embryo stage (Warm stratified seeds for 45 days); d: Early cotyledonary embryo (Warm stratified seeds for 90 days); e: Mid-term cotyledonary embryo (Warm stratified seeds for 135 days); f: Late stage of cotyledonary embryo or mature embryo (Warm stratified seeds for 180 days); g: Mature embryo stage (Cold stratified seeds for 250 days); h: Bud embryo (Cold stratified seeds for 270 days)

对刚采收时种子的心形胚、层积第 45 天的鱼雷胚、层积第 90 天的子叶胚、层积第 180 天的成熟胚和层积第 270 天芽形胚,分别标记为 T1、T2、T3、T4 和 T5 五个后熟发育时期的胚组织进行高通量测序。共获得 412 604 390 条原始 reads,去除低质量序列后获得 404 850 531 高质量 reads (clean reads),且 Q30 达到 99%以上。将 clean reads 与人参基因组进行比对,总比对率达到 94%以上(表 1)。说明我们测序获得的 reads 十分可靠。

### 1.2 差异表达基因(DEGs)鉴定与分析

以 T1 时期为对照分别与 T2、T3、T4 和 T5 四个时期进行比较基因差异表达分析( $\log_2\text{FoldChange}>1$  和 P-value 值( $\text{padj}<0.01$ )),其中在 T1 vs T2 组鉴定到 2 359 个 DEGs,包含 886 个上调和 1 473 个下调基因;在 T1 vs T3 组鉴定到 3 048 个 DEGs,包含 1 068 上调和 1 980 下调基因;在 T1 vs T4 组鉴定到 6 180 个 DEGs,包含 2 627 个上调和 3 553 个下调基因;在 T1 vs T5 组鉴定到 12 067 个 DEGs,包含 4 714 个上调和 7 353 个下调基因;四个比较组中去掉重复后共获得 13 724 个 DEGs,记为 All DEGs。我们发现 456 个和 994 个基因分别在四个组中共同上调和下调表达,这些基因可能在整个后熟发育过程比较重要。此外,随着胚的不断发育参与人参胚后熟发育调控的差异表达基因数呈现递增趋势(图 2)。

### 1.3 差异表达基因 GO 富集分析

通过序列相似性比对((BlastP, E-value  $\leq 1e-5$ ), 13 724 个人参 DEGs 中有 13 418 (97.77%)个与拟南芥(*Arabidopsis thaliana*)的基因具有同源性。基于拟南芥基因,将 T1 vs T2、T1 vs T3、T1 vs T4、T1 vs T5 和 All DEGs 五个组的 DEGs 进行 GO 富集分析,分别显著富集到 617、726、898、866、894 和 GO term。五个组中属于分子功能(Molecular function)类的 GO 条目(占比为 14.70%~16.86%)主要与催化活性(Catalytic activity)和结合(Binding)有关,并且 GO 层级关系终节点(功能更具体)为 ATP 结合(ATP binding)、蛋白结合(Protein binding)和阳离子转运 ATP 酶活性(Cation-transporting ATPase activity)条目(图 3)。

属于细胞组分(Cellular component)类的 GO 条目(占比为 14.19%~18.94%)主要与细胞(Cell)和膜(Membrane)有关(图 3)。

属于生物学过程(Biological process)类的 GO 条

表 1 不同后熟发育阶段胚 RNA-Seq 数据

Table 1 RNA-Seq data for embryos at different after-ripening developmental stages

样品	原始 reads 数	高质量 reads 数	质量分数(%)	Reads 比对率(%)
Samples	Total raw reads	Total clean reads	Q30 (%)	Total mapped reads (%)
T1_1	23 341 585	22 913 775	99.37	94.36
T1_2	22 596 675	22 136 027	99.40	94.63
T1_3	24 011 977	23 594 240	99.46	94.65
T2_1	24 077 203	23 656 699	99.29	95.10
T2_2	25 696 113	25 250 945	99.43	94.71
T2_3	40 537 010	39 809 141	99.68	95.75
T3_1	22 771 544	22 338 186	99.42	94.92
T3_2	24 229 338	23 810 515	99.67	94.84
T3_3	36 084 375	35 516 134	99.43	95.52
T4_1	24 549 981	23 865 672	99.19	94.46
T4_2	22 472 765	22 060 758	99.37	94.97
T4_3	34 040 376	33 438 747	99.39	94.90
T5_1	29 254 597	28 541 295	99.29	94.45
T5_2	22 810 949	22 389 386	99.23	94.58
T5_3	36 129 902	35 529 011	99.41	95.11

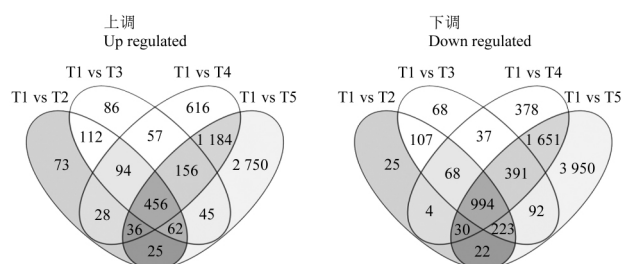


图 2 人参胚五个后熟发育时期显著上调(左边)和下调(右边)差异表达基因分析

Figure 2 Analysis of significantly up-regulated (left) and down-regulated differentially expressed genes (right) in five after-ripening development stages of ginseng embryos

目(占比为 64.20%~69.88%)主要与细胞过程(Cellular process)、代谢过程(Metabolic process)和发育过程(Developmental process)有关(图 3)。其中,细胞过程和代谢过程的 GO 层级关系终节点分别为染色质沉默(Chromatin silencing)、糖酵解过程(Glycolytic process)条目;发育过程的 GO 层级关系终节点为种子休眠结束中胚发育(Embryo development in seed dormancy ending)、花发育调控(Regulation of flower development)和种子萌发调控(Regulation of seed germination),其中以“种子休眠结束中胚发育”条目显著性最高(图 4)。因此,ATP 结合、蛋白结合、糖酵解过程和种子休眠结束中胚发育等对人参胚后熟发育有重要影响。尤其是种子休眠结束中胚发育条目中的 DEGs 对人参胚的后熟发育至关重要。

#### 1.4 种子休眠结束中胚发育相关差异表达基因蛋白互作(PPI)分析

STRING 是一个收录蛋白和蛋白与蛋白间相互作用信息的数据库,可以从蛋白水平分析人参胚发育相关差异表达基因的生物学功能和调控作用(Szklarczyk et al., 2018)。根据 GO 富集结果,我们将种子休眠结束中胚发育 GO 条目(GO: 0009793)中,共 196 个拟南芥基因对应的 381 个人参 DEGs 导入 STRING 数据库进行 PPI 分析(图 4)。共注释到有 97 个蛋白之间存在相互作用。获得的 PPI 网络分别由 97 个节点(蛋白)和 200 条边(互作关系)组成,combined score>0.4,平均节点度为 4.19,其中 RNA 聚合酶 II 第二大亚基(NRPB2)、 $\Delta^1$ -吡咯啉-5-羧酸合成酶 1 (P5CS1)和  $\Delta^1$ -吡咯啉-5-羧酸合成酶 2 (P5CS2)为网络的核心节点,节点度分别为 14、18 和 18,三个蛋白与其它蛋白有较强的相互作用关系(图 5)。并且,以 T1 时期为对照,编码 NRPB2 (1 个基因)、P5CS1 (2 个基因)和 P5CS2 (2 个基因)蛋白的 DEGs 在 T2、T3、T4 和 T5 时期均上调表达。因此,我们认为编码这三个蛋白的 DEGs 可能为人参胚后熟发育相关的关键基因。

#### 1.5 差异表达基因 KEGG 富集分析

将 T1 vs T2、T1 vs T3、T1 vs T4、T1 vs T5 和 All DEGs 五个组的差异表达基因进行 KEGG 富集分析。五个组的 DGEs 分别显著富集(FDR<0.05)到了 38、50、65、42 和 42 条通路,在前五个显著富集通路

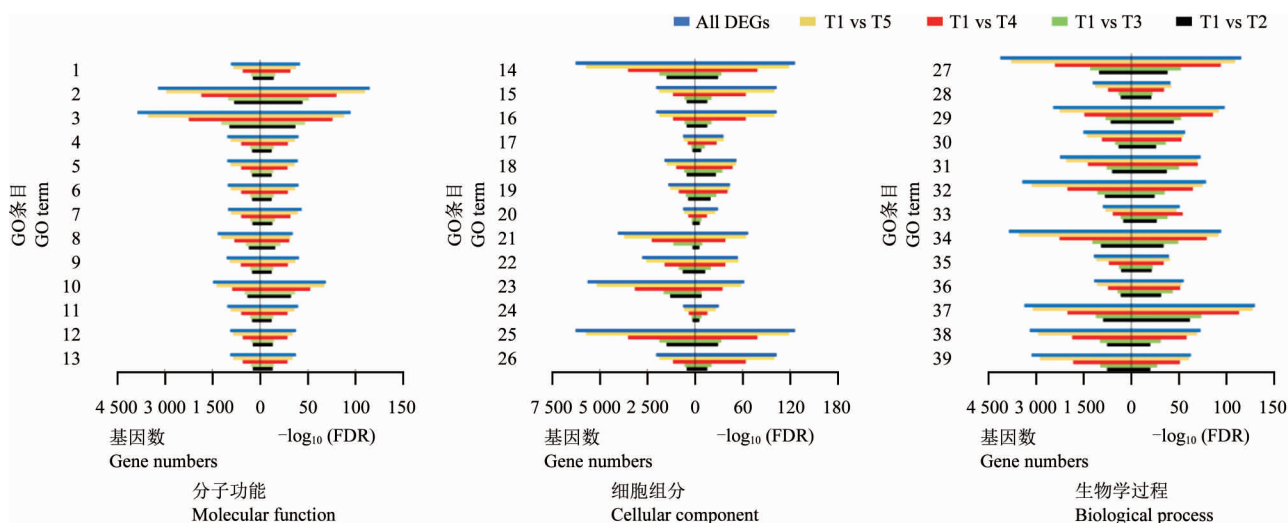


图3 五个组中差异表达基因前13个共同显著富集GO条目

注: 1: ATP 结合; 2: 催化活性; 3: 结合; 4: 核苷酸结合; 5: 嘌呤核苷酸结合; 6: 核糖核苷酸结合; 7: 嘌呤核糖核苷三磷酸结合; 8: 离子结合; 9: 碳水化合物衍生物结合; 10: 蛋白结合; 11: 嘌呤核糖核苷酸结合; 12: 腺苷核糖核苷酸结合; 13: 腺苷酸核苷酸结合; 14: 细胞部分; 15: 细胞器部分; 16: 细胞内; 17: 液泡; 18: 细胞周边; 19: 质膜; 20: 质体部分; 21: 细胞质; 22: 膜; 23: 细胞内部分; 24: 叶绿体部分; 25: 细胞; 26: 细胞内细胞器部分; 27: 细胞过程; 28: 发育过程; 29: 单个有机体细胞过程; 30: 单个有机体代谢过程; 31: 刺激反应; 32: 有机物代谢过程; 33: 非生物刺激的反应; 34: 代谢过程; 35: 组织结构发育; 36: 化学反应; 37: 发育过程; 38: 细胞代谢过程; 39: 初级代谢过程

Figure 3 The top 13 common significantly enriched GO terms in the five groups of differentially expressed genes

Note: 1: ATP binding; 2: Catalytic activity; 3: Binding; 4: Nucleoside binding; 5: Purine nucleotide binding; 6: Ribonucleoside binding; 7: Purine ribonucleoside triphosphate binding; 8: Ion binding; 9: Carbohydrate derivative binding; 10: Protein binding; 11: Purine ribonucleotide binding; 12: Adenyl ribonucleotide binding; 13: Adenyl nucleotide binding; 14: Cell part; 15: Organelle part; 16: Intracellular; 17: Vacuole; 18: Cell periphery; 19: Plasma membrane; 20: Plastid part; 21: Cytoplasm; 22: Membrane; 23: Intracellular part; 24: Chloroplast part; 25: Cell; 26: Intracellular organelle part; 27: Cellular process; 28: Developmental process; 29: Single-organism cellular process; 30: Single-organism metabolic process; 31: Response to stimulus; 32: Organic substance metabolic process; 33: Response to abiotic stimulus; 34: Metabolic process; 35: Anatomical structure development; 36: Response to chemical; 37: Developmental process; 38: Cellular metabolic process; 39: Primary metabolic process

中,主要共同富集到了代谢通路(Metabolic pathways)、次生代谢生物合成(Biosynthesis of secondary metabolites)和植物激素信号转导(Plant hormone signal transduction)通路(图6)。五个时期共计1788和970个DEGs分别富集到代谢通路和次生代谢生物合成通路,参与到碳代谢(Carbohydrate metabolism)、核苷酸代谢(Nucleotide metabolism)和氨基酸代谢(Amino acid metabolism)等过程中,并且NR-PB2基因也富集到了核苷酸代谢中的嘧啶和嘌呤代谢过程,P5CS1和P5CS2基因则富集到了氨基酸代谢中的脯氨酸代谢过程。

205个DEGs富集到植物信号激素转导通路,参与生长素(auxin, IAA)、赤霉素(gibberellin, GA)、细胞分裂素(cytokinin)、脱落酸(abscisic acid, ABA)、乙烯(Ethylene)、油菜素内酯(Brassinosteroid)、茉莉酸(Jasmonic acid)和水杨酸(Salicylic acid)八大植物激素的

信号转导调控,其中以IAA和ABA信号转导途径富集基因数最多,分别为67个(32.68%)和53个(25.85%)DEGs(表2)。且在IAA信号转导中,以生长素抑制因子AUX/IAA家族基因最多(23个DEGs),分别在四个发育时期下调表达;ABA信号转导中以脱落酸受体PYR/PYL/RCAR家族基因最多(18个DEGs),也分别在四个发育时期下调表达,并且,随着胚的不断发育参与代谢过程、次生代谢合成个植物激素信号转导调控的DGEs呈现逐渐递增趋势。

### 1.6 差异表达基因qRT-PCR分析

为验证转录组数据的可靠性,我们选取P5CS1、NEPB2、GID1、CRE1、BSK六个上调DEGs以及AUX/IAA、PYR/PYL/RCAR两个下调DEGs,利用qRT-PCR对基因表达模式进行验证,以T1为对照(各基因表达量为1),计算T2、T3、T4和T5时期基因的相对表

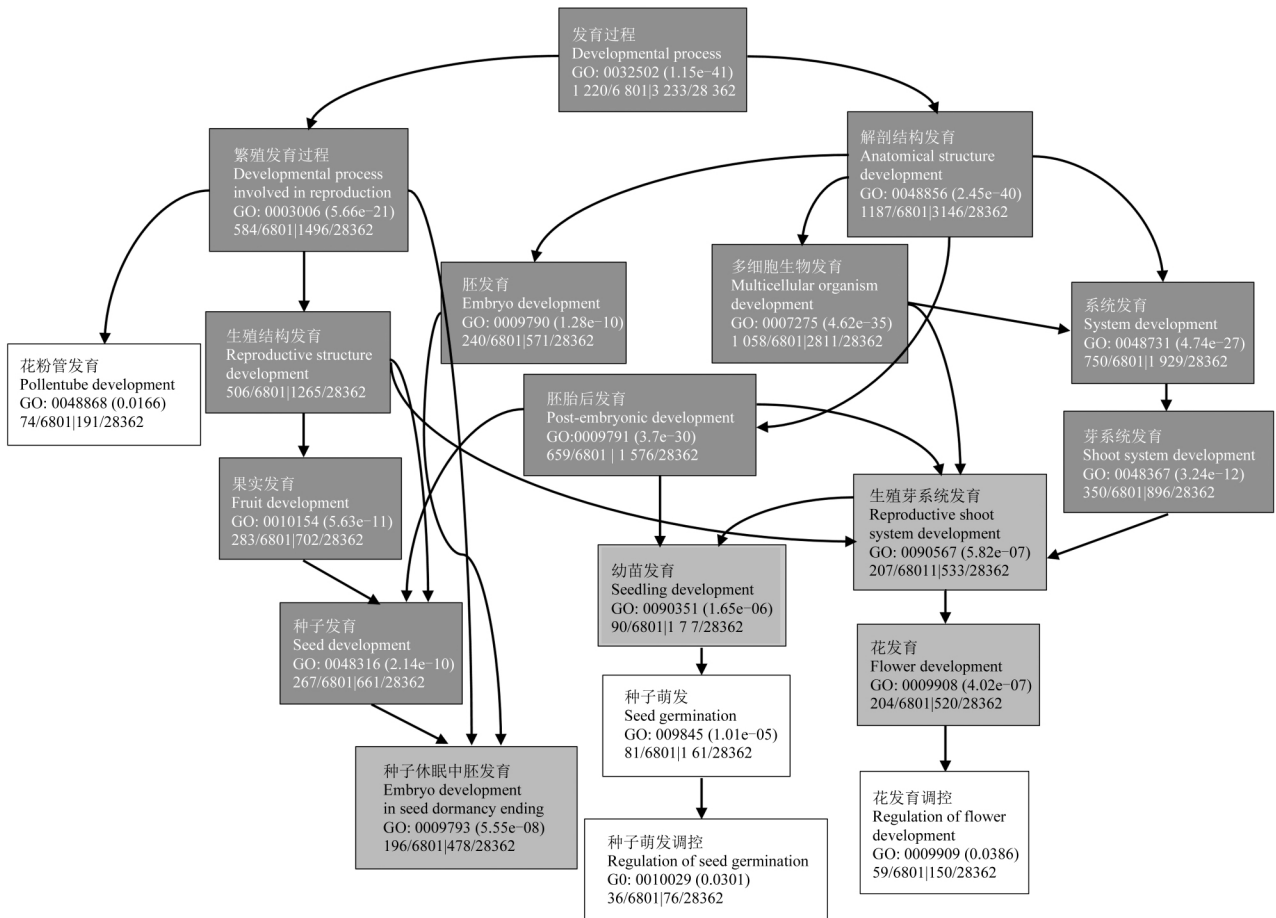


图4 发育过程 GO 层级关系

Figure 4 The GO level relationship of developmental processes

达量。结果显示, qRT-PCR 中 7 个基因的表达变化趋势与转录组测序数据基本一致, 由此说明我们的转录组分析结果是可靠的(图 7)。

## 2 讨论

胚后熟特性对人参种子的萌发具有重要影响。本研究对人参种子后熟发育过程形态观察分析发现, 16℃ 暖温层积可以促进人参种胚完成形态后熟和种子裂口, 而 4℃ 冷温层积可以进种子完成生理后熟和种子发芽, 该结果与 Chen 等(1984)研究结果相符合。

对人参种胚五个后熟发育阶段转录数据分析发现, GO 富集揭示人参胚发育转录特性主要与催化活性、细胞和发育过程等相关。进一步对发育过程中与种子休眠结束中胚发育相关 DEGs 进行 PPI 网络分析发现, NRPB2、P5CS1 和 P5CS2 为核心节点蛋白, 与其它蛋白有较强的相互作用关系, 可能是人参胚后熟发育过程中的关键蛋白。且相关研究表明, NRPB2 是指导 mRNA 合成的 RNA 聚合酶 II 第二大

亚基, NRPB2 基因的表达对 mRNA 的合成至关重要 (Zheng et al., 2009)。P5CS ( $\Delta^1$ -吡咯啉-5-羧酸合成酶) 是植物体内脯氨酸合成的关键酶, 主要由 P5CS 基因家族编码, P5CS1 和 P5CS2 是两个同源蛋白。对烟草和甘蔗的研究表明 P5CS1 基因的过表达可以促进脯氨酸积累和生物量合成, 从而增加植物的抗逆性(如抗盐, 抗旱性)和促进根系生长 (Iskandar et al., 2014; Ibragimova et al., 2015); 在拟南芥中 P5CS1 和 P5CS2 基因的过表达可以促进开花, 反之则延迟开花, 当出现 P5CS2 突变体时胚胎停止发育 (Mattioli et al., 2009)。此外, 本研究 KEGG 分析也显示, NRPB2 基因通过调控嘌呤和嘧啶代谢, 影响 mRNA 的合成; P5CS1 和 P5CS2 基因参与脯氨酸代谢过程。因此, 我们推测 NRPB2 基因的表达可以影响与之互作蛋白的表达和调控 mRNA 的合成来影响胚发育过程中蛋白合成; 而 P5CS1 和 P5CS2 基因的表达量增加可以促进脯氨酸的合成, 脯氨酸含量的增加可以加快器官生长和分生组织形成, 并刺激互作蛋白的表达, 从而促进人参种胚的发育。

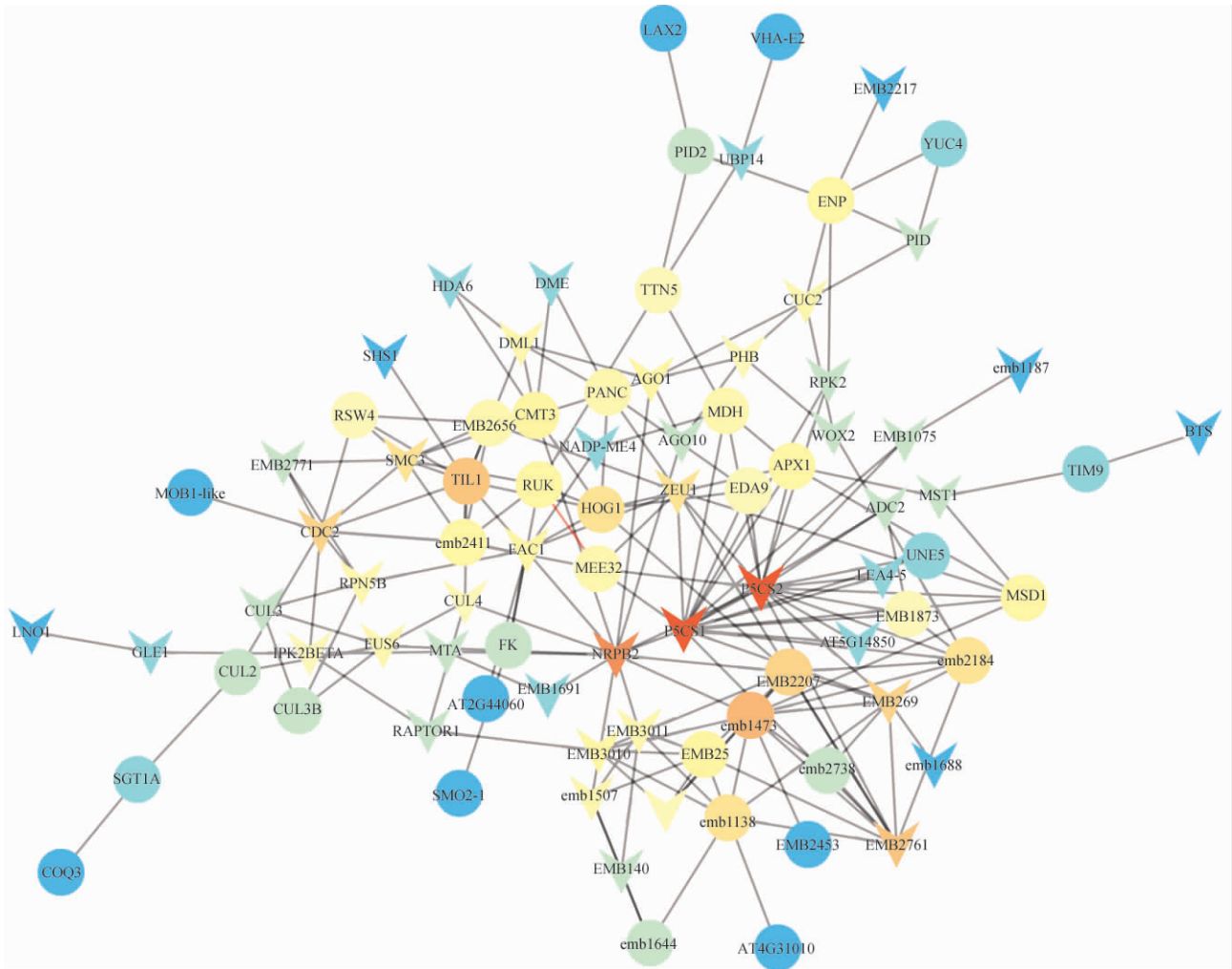


图 5 差异表达基因 PPI 网络

注: 线条代表蛋白之间互作关系; 圆形表示编码该蛋白的基因调表达, 三角形表示编码该蛋白的基因上调表达(红色越深节点度越大, 蓝色越深节点度越小)

Figure 5 Protein-protein interaction network of differentially expressed genes

Note: The lines represent interactions between protein and protein; The circle indicates the gene expression of the gene encoding the protein, and the triangle indicates the up-regulated expression of the gene encoding the protein (The deeper the red, the bigger the node degree is, and the deeper the blue color is, the smaller the node degree is)

植物激素, 一直以来对种子的生长、发育和萌发有着非常重要的影响(Miransari and Smith, 2014)。本研究通过 KEGG 分析发现, 植物激素信号转导也是人参胚后熟过程中重要代谢途径, 并且大量 DEGs 主要富集到了 IAA 和 ABA 信号转导过程。相关研究表明, IAA 是调控植物器官生长、发育、形态建成和组织分化等的主要植物激素(Villalobos et al., 2012), 生长素受体 TIR1 和同源受体 AFBs 与生长素抑制因子 AUX/IAA 结合后, 使 AUX/IAA 蛋白迅速泛素化降解, 并释放生长素响应因子 ARFs, 从而介导生长素信号传递, 促进植物生长(Peer, 2013); ABA 是调控种子休眠的主要植物激素, 脱落酸受体 PYR/PYL/

RCAR 与 ABA 结合后, 抑制下游 PP2C 蛋白磷酸化, 并激活 SnRKs 促进 ABA 合成, ABA 含量增加可以抑制种子的生长和萌发(Nakashima and Yamagu-chi-Shinozaki, 2013; 熊孟连等, 2019, 私人通信)。在对人参体细胞胚胎发生过程中内源激素变化的研究中也发现, 在早期胚中 IAA 含量较高, 在胚发育成熟期 ABA 含量较高, IAA 是维持早期胚发育的重要条件(王义等, 2008, 中草药, 39(7): 1084-1089)。而本次研究鉴定到的, 23 个 IAA 抑制因子 AUX/IAA 基因下调表达和 18 个 ABA 受体 PYR/PYL/RCAR 基因下调表达, 也印证了 DEGs 通过调控 IAA 合成和 ABA 代谢, 来促进人参胚的发育。此外, 其它植物激素如赤

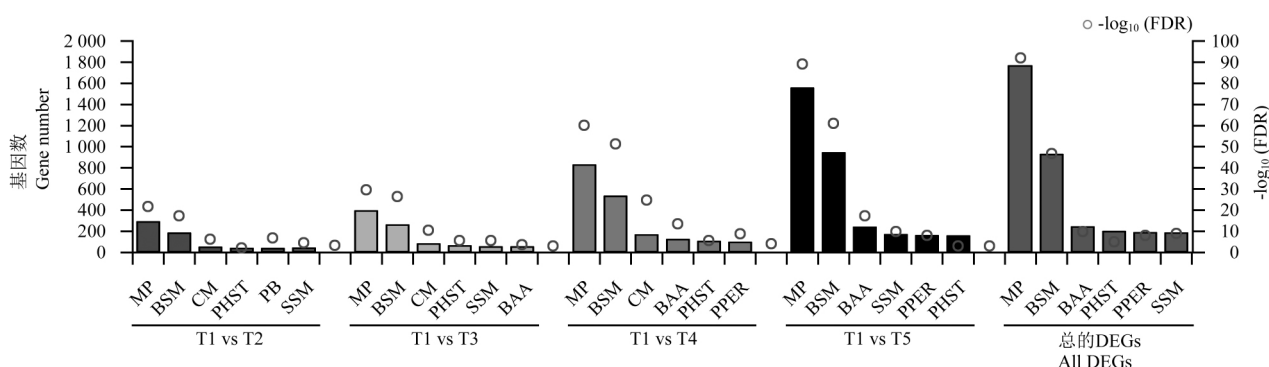


图6 五个组中差异表基因前6个显著富集KEGG通路

注: MP: 代谢通路; BSM: 次生代谢生物合成; CM: 碳代谢; PHST: 植物激素信号转导; PB: 苯丙烷类合成; SSM: 淀粉和蔗糖代谢; BAA: 氨基酸生物合成; PPER: 内质网中蛋白加工过程

Figure 6 The top 5 significantly enriched KEGG pathways in the five groups of differentially expressed genes

Note: MP: Metabolic pathways; BSM: Biosynthesis of secondary metabolites; CM: Carbon metabolism; PHST: Plant hormone signal transduction; PB: Phenylpropanoid biosynthesis; SSM: Starch and sucrose metabolism; BAA: Biosynthesis of amino acids; PPER: Protein processing in endoplasmic reticulum

霉素和细胞分裂素等对胚的发育也有一定促进作用(Guan et al., 2014; Kim et al., 2014), 本研究中也鉴定到GID1、GID2和CRE1等相关基因, 与前人的研究相一致。近年来, 在同属植物西洋参种子休眠解除和三七种子后熟转录组研究中也发现差异表达基因参与植物激素信号转导调控(Qi et al., 2015; Yang et al., 2018)。由此可见, 以IAA和ABA主导与其它植物激素协同调控的植物激素信号转导过程对人参胚后熟发育具十分重要。

虽然近年来人们通过植物激素处理、变温层积等方法来促使人参种子完成后熟发育(Rajametov et al., 2014; Lee et al., 2018), 但对种子发育过程的分子机制尚不清楚。本研究借助转录组测序技术, 通过分析人参种胚不同后熟发育阶段的基因表达调控情况, 为今后揭示人参胚的发育和种子休眠解除过程的分子调控机制的深入研究提供一定帮助, 同时对人参的育种和长期保存种子也有较为重要的参考价值。

### 3 材料与方法

#### 3.1 植物材料

实验材料来源于吉林省国华参业有限公司2017年8月10日采收的林下参种子。去皮后的鲜人参种子用清水浸泡48h后, 用小型尼龙网分装成若干袋(每袋约500粒), 然后再放入经高温灭菌并伴有菌灵的湿沙中催芽层积。由于普通的恒温播种方法不能使人参胚完成后熟发育过程(王化武, 1979, 特产科学实验, (2): 1-3)。因此, 本实验采用Chen等(1984)变温催芽方法使人参种子完成后熟发育, 先在16℃

条件下进行暖温层积(6个月)到大多数种子胚形态大小发育完全, 再转入4℃条件下进行冷温层积到种子发芽(4~5个月), 每周取10粒种子解剖观察胚的形态。本实验分别对刚采收未层积时的心形胚、层积(16℃)第45天的鱼雷形胚、层积(16℃)第90天的子叶形胚、层积(16℃)第180天的成熟胚(胚充满胚腔)和层积(4℃)第270天的芽状胚(胚根突破胚乳2~4mm)五个时期的人参胚组织进行取样。每个时期挑选大部分胚形态大小一致的种子, 在解剖镜下迅速取下胚组织放入离心管中(离心管下部浸于液氮中), 每个样品3个重复。5个时期样品分别标记为: T1、T2、T3、T4和T5。取好的样品材料放置-80℃冰箱保存备用。

#### 3.2 RNA提取与高通量测序

采用天根生物公司多糖多酚植物总RNA提取试剂盒(DP441)提取样品总RNA, 然后采用Nanodrop、Qubit 2.0、Aglient 2100方法分别检测RNA样品的纯度、浓度和完整性。检测合格的样品, 由北京百迈客生物科技有限公司进行RNA文库构建, 并用Illumina HiSeq X-ten进行高通量测序, 测序读长为150bp。

#### 3.3 reads与参考基因组比对

得到的原始测序数据(raw reads)先用Fastqc(v0.11.8)软件进行质量检测, 再用Trimmatic(v0.4.5)软件(Bolger et al., 2014)去掉每条reads中前13个偏离较大的碱基和低质量reads(Q≤15的碱基数占整条read的50%以上的reads), 最终获得高质量的序列



表 2 植物激素信号转导通路差异表达基因

Table 2 The differentially expressed genes in plant hormone signal transduction pathways

类别 Type	基因家族 Gene family	基因数 Gene numbers	调控 Regulated
生长素 Auxin	AUX1 LAX/生长素流入载体	8	↑
	AUX1 LAX/auxin influx carrier		
	TIR1/转运抑制响应 1	4	↓
	TIR1/transport inhibitor response 1		
	AUX/IAA/生长素抑制	23	↓
	AUX/IAA/auxin inhibitor factor		
	ARF/生长素响应因子	7	↑
	ARF/auxin response factor		
	GH3 生长素响应因子	11	↑
	GH3/auxin response factor		
细胞分裂素 Cytokinin	SAUR 家族蛋白	14	↓
	SAUR family protein		
	CRE1/拟南芥组氨酸激酶(细胞分裂素受体)	3	↑
	CRE1/arabidopsis histidine kinase (cytokinin receptor)		
	AHP/含组氨酸的磷酸转移蛋白	5	↓
	AHP/histidine-containing phosphotransfer peotein		
	B-ARR /双组分响应调控 ARR-B 家族	6	↑
B-ARR/two-component response regulator ARR-B family			
赤霉素 Gibberellin	A-ARR /双组分响应调控 ARR-A 家族	5	↓
	A-ARR/two-component response regulator ARR-A family		
	赤霉素受体 GID1	6	↑
脱落酸 Abscisic acid	GID1/gibberellin receptor GID1		
	F-box 蛋白 GID2	1	↑
乙烯 Ethylene	GID2/F-box protein GID2		
	PYL/PYL/RCAR 家族/脱落酸受体	18	↓
	PYL/PYL/RCAR family/abscisic acid receptor		
	PP2C/蛋白磷酸酶 2C	14	↑
	PP2C/protein phosphatase 2C		
油菜素内酯 Brassinosteroid	SNRK2/丝氨酸/苏氨酸-蛋白激酶 SRK2	10	↓
	SNRK2/serine/threonine-protein kinase SRK2		
	ABF/ABA 反应元件结合因子	11	↑
	ABF/ABA responsive element binding factor		
	ETR/乙烯受体	4	↓
	ETR/ethylene receptor		
油菜素内酯 Brassinosteroid	EIN3/乙烯不敏感蛋白 3	1	↑
	EIN3/ethylene-insensitive protein 3		
	EBF1_2/EIN 结合 F-box 蛋白	2	↓
	EBF1_2/EIN3-binding F-box protein		
油菜素内酯 Brassinosteroid	ERF1/乙烯响应转录因子 1	5	↓
	ERF1/ethylene-responsive transcription factor 1		
	BSK/BR 信号激酶(油菜素内酯受体)	14	↑
油菜素内酯 Brassinosteroid	BSK/BR-signaling kinase (brassinosteroid receptor)		
	BIN2/油菜素类内酯不敏感蛋白 2	1	↓
	BIN2/protein brassinosteroid insensitive 2		

续表 2

Continuing table 2

类别	基因家族	基因数	调控
Type	Gene family	Gene numbers	Regulated
油菜素内酯 Brassinosteroid	抗油菜素内酯 1/2	2	↓
	BZR1_2/brassinosteroid resistant 1/2		
	CYCD3/细胞周期蛋白 D3	7	↑
茉莉酸 Jasmonic acid	CYCD3/cyclin D3		
	JAR1/茉莉酸-氨基合成酶	3	↑
	JAR1/jasmonic acid-amino synthetase		
	COI-1/冠菌素不敏感蛋白 1	2	↓
	COI-1/coronatine-insensitive protein 1		
水杨酸 Salicylic acid	JAZ/茉莉酸酯 ZIM 结构域蛋白	5	↓
	JAZ/jasmonate ZIM domain-containing protein		
	NPR1/调节蛋白 NPR1(水杨酸受体)	3	↑
	NPR1/regulatory protein NPR1 (salicylic receptor)		
TGA/转录因子 TGA	TGA/转录因子 TGA	8	↑
	TGA/transcription factor TGA		
	PR1/病原体相关蛋白	2	↑
PR1/pathogenesis-related protein 1			

注: ↑ : 上调; ↓ : 下调

Note: ↑ : Up-regulated; ↓ : Down-regulated

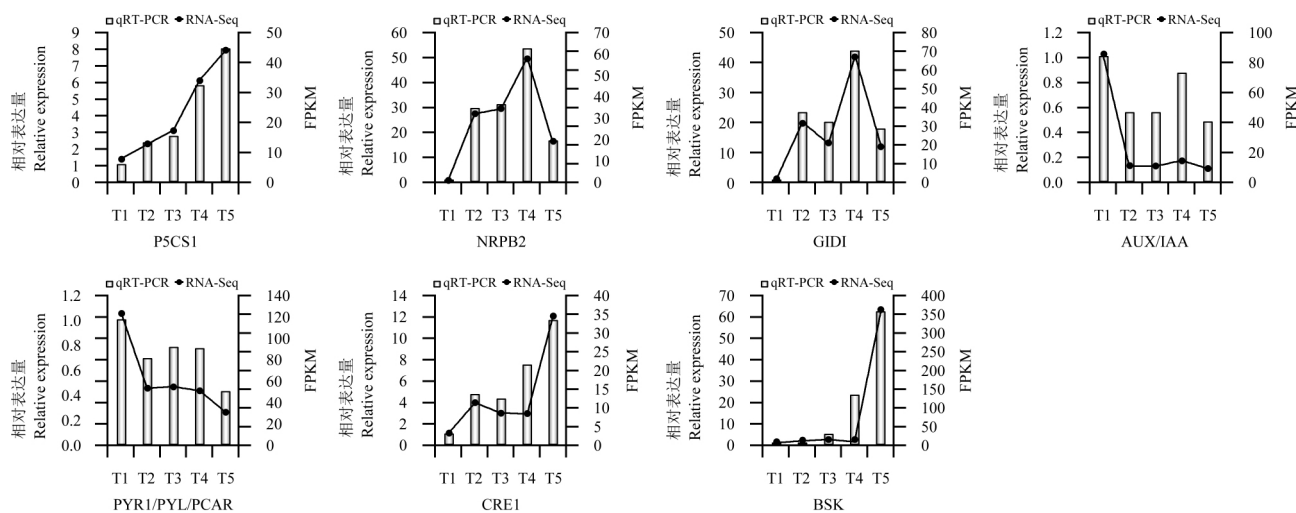


图 7 差异表达基因表达量验证

Figure 7 Verification of relative expression levels of DEGs

(Clean reads)。最后用 Hisat2 (v2.1.0)软件(Kim et al., 2015)将 clean reads 与人参考基因组(Xu et al., 2017)进行比对。

### 3.4 基因差异表达分析

基于比对结果用 HTSeq (v0.11.1)软件(Anders et al., 2015)提取出每个样本中所有基因对应的 reads 数(使用参数: --stranded=no-r pos-i gene\_id), 然后用

R 中的 DESeq2 (v3.8)软件 (Love et al., 2014)以  $|\log_2\text{-FoldChange}| > 1$  和校正后的  $P\text{-value 值}(padj) < 0.01$  为条件筛选出差表达基因。

将差异表达基因进行 GO (Gene ontology)富集和蛋白互作(protein-protein interaction, PPI)网络分析。用 KOBAS 3.0 软件(<http://kobas.cbi.pku.edu.cn/annotate.php>)将差异表达基因与拟南芥(*Arabidopsis thaliana*)基因组进行蛋白序列相似性比对, 获得人参差异

表达基因在拟南芥基因组中的同源基因。基于以上的比对结果,用 AgriGO 软件(<http://systemsbiology.cau.edu.cn/agriGOv2/index.php>)对差异表达基因进行 GO 富集分析,其中以 FDR<0.05 作为显著富集的 GO 条目(GO terms)。STRING 是一个收录蛋白和蛋白与蛋白间互作信息的数据库(Szklarczyk et al., 2018)。根据 GO 富集结果筛选有特殊意义的通路,然后将该通路中差异表达基因与 STRING 蛋白数据库进行比对(<https://string-db.org/>),获得基因对应蛋白之间的互作关系。最后用 Cytoscape 软件对蛋白互作结果进行可视化展示(Shannon et al., 2003)。

对差异表达基因进行 KEGG (Kyoto encyclopedia of genes and genomes)富集分析。基于 KEGG 数据库,用 KOBAS 软件对差异表达基因进行代谢通路富集,并以 FDR<0.05 筛选为显著富集通路,然后对特殊通路进一步分析。

### 3.5 qRT-PCR 验证

根据转录分析结果,挑选出 8 个 DEGs 进行 RT-qPCR 相对表达分析。按照 Prime Script™ II 试剂盒说明书(TaKaRa 公司)将 RNA 反转成 cDNA。用 Primer 5.0 引物(表 3)。qRT-PCR 反应体系 20 μL: EvaGreen 2×qPCR Master Mix 10 μL, cDNA 模板 1 μL,正、反引物各 0.3 μL,补足 ddH<sub>2</sub>O 至 20 μL。将其放置 ABI 荧光定量 PCR 仪中,95℃预变性 10 min,95℃变性 14 s,52℃退火 1 min,72℃延伸 15 s,40 个循环。每个反应重复 3 次,以 *IF3G1* 为内参基因(Fan et al., 2019),采用 2<sup>-ΔΔCt</sup> 法计算基因相对表达量。

### 作者贡献

冷容是本研究的实验设计者和实验研究的执行人,完成数据分析,论文初稿的写作;胡彦婷、周艳

丽、金桂花、李玲、石勇、杨红、桂大萍和薛润光参与实验设计,试验结果分析;章成君是项目的构思者及负责人,指导实验设计、数据分析、论文写作与修改。全体作者都阅读并同意最终的文本。

### 致谢

本研究由中国科学院百人计划自然科学基金面上项目(31571311)和云南省第三批博士后定向培养基金(39Y73241261)共同资助。

### 参考文献

- Anders S., Pyl P.T., and Huber W., 2015, HTSeq-a Python framework to work with high-throughput sequencing data, *Bioinformatics*, 31(2): 166-169
- Baskin J.M., and Baskin C.C., 2004, A classification system for seed dormancy, *Seed Sci. Res.*, 14(1): 1-16
- Bolger A.M., Marc L., and Bjoern U., 2014, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics*, 30(15): 2114-2120
- Chen Y., Sun C.G., and Zheng G.H., 1984, Methods to promote the seed germination of *Panax ginseng*, *Chin. Sci. Bull.*, 29(23): 425-426
- Costa V., Angelini C., De F.I., and Ciccodicola A., 2010, Uncovering the complexity of transcriptomes with RNA-Seq, *J. Biomed. Biotechnol.*, 5757: 853916-853926
- Fan H., Li K., Yao F., Sun L.W., and Liu Y.J., 2019, Comparative transcriptome analyses on terpenoids metabolism in field- and mountain-cultivated ginseng roots, *BMC Plant Biology*, 19: 82
- Gao P., Xiang D., Quilichini T.D., Venglat P., Pandey P.K., Wang E., Gillmor C.S., and Datla R., 2019, Gene expression atlas of embryo development in Arabidopsis, *Plant Reprod.*, 32(1): 93-104
- Guan C.M., Wang X.C., Feng J., Hong S.L., Liang Y., Ren B., and Zuo J.R., 2014, Cytokinin antagonizes abscisic acidme-

表 3 基因名称及引物序列

Table 3 Gene name and primer sequence

基因名称 Gene name	正向引物(5'-3') Forward primer (5'-3')	反向引物(5'-3') Reverse primer (5'-3')
<i>IF3G1</i>	CTTGAAAAGGTGGTGCTGT	TTGTTGTATCCGAGGTTGGG
<i>P5CS1</i>	GTAGAATTCATGCTCGTGGTC	GTCCCTGTGAGTGTAACAAC
<i>NRPB2</i>	GCAGAAACACCAGAAGGACAA	GATGAATACCAACCCAGCAAC
<i>GIDI</i>	GCACATTCCTCCGCAAACAG	CGCAGTCCAGCCATCTTCAT
<i>AUX/IAA</i>	GTTTATTCTCCCCGAGGTT	ATATTGACCCCAAGGAGCTG
<i>PYL/PYL/RCAR</i>	CAGAGATTATGGATGGGAGGC	AATAGGCTCAGTTCGGTCTTG
<i>CRE1</i>	TGAGGATGGCTGCTTGATAC	TGCATTGTCAAACCTCAGCAC
<i>BSK</i>	TCTGTGACAAACCAATGAAGTG	CATGATATAACGCACGACCTTT

- diated inhibition of cotyledon greening by promoting the degradation of abscisic acid insensitive5 protein in *Arabidopsis*, *Plant Physiol.*, 164(3): 1515-1526
- Ibragimova S.M., Trifonova E.A., Filipenko E.A., and Shymny V.K., 2015, Evaluation of salt tolerance of transgenic tobacco plants bearing the P5CS1 gene of *Arabidopsis thaliana*, *Russian Journal of Genetics*, 51(12): 1181-1188
- Iskandar H.M., Widyaningrum D., and Suhandono S., 2014, Cloning and characterization of P5CS1 and P5CS2 genes from *Saccharum officinarum* L. under drought stress, *Journal of Tropical Crop Science*, 1(1): 23-30
- Kim D., Langmead B., and Salzberg S.L., 2015, HISAT: a fast spliced aligner with low memory requirements, *Nat. Methods*, 12(4): 357-360
- Kim Y.H., Ahn I.O., Khan A.L., Kamran M., Waqas M., Lee J.S., Kim D.H., Jang S.W., and Lee I.J., 2014, Regulation of endogenous gibberellins and abscisic acid levels during different seed collection periods in *Panax ginseng*, *Hortic. Environ. Biotechnol.*, 55(3): 166-174
- Lee J.W., Jo I.H., Kim J.U., Hong C.E., Kim Y.C., Kim D.H., and Park Y.D., 2018, Improvement of seed dehiscence and germination in Ginseng by stratification, gibberellin, and/or kinetin treatments, *Hortic. Environ. Biotechnol.*, 59(3): 293-301
- Love M.I., Huber W., and Anders S., 2014, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.*, 15(12): 550-571
- Mattioli R., Falasca G., Sabatini S., Altamura M.M., Costantino P., and Trovato M., 2009, The proline biosynthetic genes P5CS1 and P5CS2 play overlapping roles in *Arabidopsis* flower transition but not in embryo development, *Physiol. Plant.*, 137(1): 72-85
- Miransari M., and Smith D., 2014, Plant hormones and seed germination, *Environ. Exp. Bot.*, 99(3): 110-121
- Nakashima K., and Yamaguchi-Shinozaki K., 2013, ABA signaling in stress-response and seed development, *Plant Cell Rep.*, 32(7): 959-970
- Peer W.A., 2013, From perception to attenuation: auxin signalling and responses, *Curr. Opin. Plant Biol.*, 16(5): 561-568
- Qi J.J., Sun P., Liao D.Q., Sun T.Y., Zhu J., and Li X.N., 2015, Transcriptomic analysis of American ginseng seeds during the dormancy release process by RNA-Seq, *PLoS One*, 10(3): e0118558
- Qi L.W., Wang C.Z., and Yuan C.S., 2011, Isolation and analysis of Ginseng: advances and challenges, *Nat. Prod. Rep.*, 28(3): 467-495
- Rajabian A., Rameshrad M., and Hosseinzadeh H., 2019, Therapeutic potential of *Panax ginseng* and its constituents, ginsenosides and gintonin, in neurological and neurodegenerative disorders: a patent review, *Expert Opin. Ther. Pat.*, 29(1): 55-72
- Rajametov S., Youngyi L., Youngchang K., Sokyoun L., Yi J.Y., Youngah J., Jungsook S., Gian L., and Jaegyun G., 2014, Response of pre and post treatments for cryopreservation of Korean ginseng seeds on recovering viability, *Korean Journal of Breeding Science*, 46(4): 408-416
- Shannon P., Markiel A., Ozier O., Baliga N.S., Wang J.T., Ramage D., Amin N., Schwikowski B., and Ideker T., 2003, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Research*, 13(11): 2498-2504
- Sun Y., 2011, Structure and biological activities of the polysaccharides from the leaves, roots and fruits of *Panax ginseng* C.A. Meyer: an overview, *Carbohydr. Polym.*, 85(3): 490-499
- Szklarczyk D., Gable A.L., Lyon D., Junge A., Wyder S., Huerta-Cepas J., Simonovic M., Doncheva N.T., Morris J.H., and Bork P., 2018, STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Res.*, 47(D): 607-613
- Villalobos L.I.A.C., Lee S., De Oliveira C., Ivetaç A., Brandt W., Armitage L., Sheard L.B., Tan X., Parry G., and Mao H., 2012, A combinatorial TIR1/AFB-Aux/IAA co-receptor system for differential sensing of auxin, *Nat. Chem. Biol.*, 8(5): 477-485
- Xu J., Chu Y., Liao B., Xiao S., Yin Q., Bai R., Su H., Dong L., Li X., Qian J., Zhang J.J., Zhang Y.J., Zhang X.Y., Wu M. L., Zhang J., Li G.Z., Zhang L., Chang Z.Z., Zhang Y.B., Jia Z.W., Liu Z.X., Afreh D., Nahurira R., Zhang L.J., Chen R. Y., Zhu Y.J., Zhu G.W., Rao W.M., Zhou C., Qiao L.R., Huang Z.H., Chen Y.C., and Chen S.L., 2017, *Panax ginseng* genome examination for ginsenoside biosynthesis, *Giga-Science*, 6(11): 1-15
- Yang K., Yang L., Fan W., Long G.Q., Xie S.Q., Meng Z.G., Zhang G.H., Yang S.C., and Chen J.W., 2018, Illumina-based transcriptomic analysis on recalcitrant seeds of *Panax notoginseng* for the dormancy release during the after-ripening process, *Physiol. Plant.*, 167(4): 597-612
- Yun T.K., 2001, Brief introduction of *Panax ginseng* C. A. Meyer, *J. Korean Med. Sci.*, 16(S): 3-5
- Zhao M.Z., Lin Y.P., Wang Y.F., Li X.Y., Han Y.L., Wang K.Y., Sun C.Y., Wang Y., and Zhang M.P., 2019, Transcriptome analysis identifies strong candidate genes for ginsenoside biosynthesis and reveals its underlying molecular mechanism in *Panax ginseng* CA Meyer, *Sci. Rep.*, 9(1): 615-625
- Zheng B., Wang Z., Li S., Yu B., Liu J.Y., and Chen X., 2009, Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*, *Genes and Development*, 23(24): 2850-2860