

锥栗种仁转录组及淀粉和蔗糖代谢相关酶基因的表达分析

张琳, 范晓明, 林青, 向晖, 汪武, 袁德义

(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/经济林培育与利用湖南省 2011 协同创新中心, 长沙 410004)

摘要:本研究采用高通量测序技术对淀粉积累高峰期的锥栗种仁进行转录组测序分析,对得到的 Unigene 进行功能注释、分类及代谢通路分析,并进一步通过实时定量 PCR 方法分析了 7 个与淀粉和蔗糖代谢相关酶基因在锥栗种仁发育过程中的表达特征。结果显示:*de novo* 组装后共获得 53629 条 Unigene 序列,序列平均长度为 746 bp;通过与其他核酸、蛋白质数据库的 Blast 搜索比对,共 26739 条 Unigene 序列获得基因注释,占 All-Unigene 的 49.86%;与 COG 数据库比对后将其注释的 14413 条 Unigene 序列划分成 25 类;GO 功能注释的 33926 条 Unigene 基因共分成细胞组分、分子功能和生物过程 3 大类 58 个分支;与 KEGG 数据库比对结果注释的 5277 条 Unigene 序列划分为 116 条代谢通路;7 个与淀粉和蔗糖代谢相关酶基因表达量变化趋势中,CH.29636(淀粉合成酶,SS),CH.11971(淀粉分支酶,SBE)两个基因随着种仁的发育呈现逐渐上升的趋势,而其余 5 个基因(CH.31302、CH.33690、CH.19238、CH.30128、CH.13088)表达量随着种仁的发育呈先上升后下降的趋势,该结果与锥栗种仁发育期淀粉和蔗糖的变化规律一致。这些信息为锥栗果实品质形成相关功能基因的研究提供了重要依据。

关键词:锥栗;RNA-Seq;种仁;实时荧光定量 PCR;淀粉;蔗糖

Transcriptome Analysis for Developing Kernel and Expression Analysis of Starch and Sucrose Metabolism-related Genes in *Castanea henryi*

ZHANG Lin, FAN Xiao-ming, LIN Qing, XIANG Hui, WANG Wu, YUAN De-yi

(The Key Lab of Cultivation and Protection for Non-wood Forest Tree of Education Ministry, Central South University of Forestry and Technology Changsha / Cooperative Innovation Center of Cultivation and Utilization for Non-Wood Forest Trees of Hunan Province, Changsha 410004)

Abstract: *Castanea henryi* is an important woody grain plants in south of China. The current research on the fruit development of *C. henryi* mainly focuses on fruit component analysis, fruit quality analysis, and fruit processing, whereas the molecular mechanisms involved in fruit development remain to be uncovered. Transcriptome sequencing and analysis were conducted on kernel in the beginning of starch synthesis peak phase of *C. henryi* using Illumina platform-based strategy. Expression patterns of seven starch and sucrose metabolism-related Unigenes were assayed with real-time quantitative PCR (qPCR) method. The results showed that *de novo* assembly generated a total of 53629 Unigenes with average length of 746 bp. Among them, 26739 unigenes were annotated against NR, NCBI and Swiss-Port protein databases, 14413 Unigenes were assigned 25 categories by comparing with COG (Clusters of Orthologous Groups) database, and 33926 Unigenes were assigned 3 categories including cellular component, molecular function and biological process and 58 branches with GO (Gene Ontology) database. Functional annotation against KEGG (Kyoto Encyclopedia of Genes and Genomes) database identified 5277 Unigenes which were mapped to 116 metabolic pathways. Moreover, the qPCR analysis showed that two Unigenes

收稿日期:2013-08-14 修回日期:2014-10-23 网络出版日期:2015-04-10

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20150410.1619.012.html>

基金项目:国家“十二五”科技支撑计划课题(2013BAD14B04)

第一作者主要从事经济林遗传育种与分子育种的工作。E-mail: triwoodtim918@126.com

通信作者:袁德义,主要从事经济林育种与栽培研究。E-mail: yuan-deyi@163.com

CH. 29636 and CH. 11971 had a gradual increase, while the other 5 genes (CH. 31302, CH. 33690, CH. 19238, CH. 30128, and CH. 13088) increased at the early stage and then decreased. The expression profiles of the 7 genes were in accordance with the accumulation of starch and sucrose in the developing fruits of *C. henryi*. This study provided a foundation for further characterizing the functional genes involved in the formation of fruit quality in *C. henryi*.

Key words: *Castanea henryi*; RNA-seq; kernel; real-time RT-PCR; starch; sucrose

锥栗 (*Castanea henryi*) 果实锥形, 其坚果淀粉含量高, 同时富含糖分、蛋白质、氨基酸、脂肪、维生素以及矿质元素等多种营养成分, 是中国南方重要的木本粮食树种^[1]。大力发展锥栗产业, 可以极大促进我国木本粮食产业的发展, 加快我国南方山区经济发展。目前用于栽培锥栗良种少, 通过生物技术手段进行品种遗传改良具有重要理论意义和应用前景。前人对锥栗果实品质分析及果品加工等方面进行了大量研究^[2-4], 而锥栗果实发育相关分子生物学研究则鲜有报道^[5]。目前, 锥栗尚未完成全基因组测序, 遗传背景不清楚, 相关基因尤其是与果实发育有关的重要基因种类及其功能不明确, 阻碍了锥栗果实品质形成机理等相关基础研究的开展。

新一代高通量测序技术具有数据量大、准确度和灵敏度高优点, 目前为植物群体遗传规律、基因发掘等方面的研究提供了强有力的技术支撑^[6-8]。前期工作中, 阐明了锥栗种仁发育的动态过程和淀粉等积累的规律^[9], 在此基础上, 本研究采用 Illumina HiSeq™ 2000 高通量测序平台对锥栗种仁中淀粉合成的关键时期进行了转录测序, 对组装后的 Unigene 序列进行了 COG、GO 和 KEGG 分类统计、功能注释和代谢通路分析, 并进一步通过实时定量 PCR 技术研究了 7 个与淀粉和蔗糖代谢相关酶基因在锥栗种仁发育过程中的表达规律。本研究为了解锥栗遗传信息, 特别是挖掘锥栗果实发育过程中与淀粉、蔗糖代谢相关基因, 对锥栗果实品质的定向遗传改良具有十分重要的意义。

1 材料与方法

1.1 材料

试验材料为湖南郴州汝城县土桥镇中南林业科技大学锥栗实验基地生长 6 年的芒子锥栗。于锥栗开花后 109 d (淀粉积累高峰期) 采集其果实, 快速去除刺苞、果壳, 取出种仁用锡箔纸包好后置于液氮中保存备用。

1.2 方法

1.2.1 RNA 提取 锥栗种仁总 RNA 的提取采用 Ambion 试剂盒。

1.2.2 cDNA 文库构建 锥栗种仁 cDNA 文库的构建使用 TruSeq RNA Sample Preparation 试剂盒法。

1.2.3 转录组测序与数据整理分析 锥栗种仁样本转录组测序工作委托北京百迈克公司采用 Illumina HiSeq™ 2000 测序仪完成。参照江香梅等^[10]对转录组测序数据进行整理与分析。

1.2.4 实时荧光定量 RT-PCR 根据范晓明等^[9]对锥栗种仁营养物质变化规律的研究, 提取锥栗开花后 95 d、102 d、109 d、123 d 及 137 d 等 5 个种仁发育关键时期的总 RNA。使用 Rever Aid First Strand cDNA Synthesis Kit 进行逆转录 PCR 扩增, 以 *CsAtubulin* 为内参基因, 选取 7 条与锥栗种仁淀粉和蔗糖代谢相关酶基因进行实时定量 PCR 试验, 基因及引物信息见表 1。实时定量 PCR 方法参照周俊琴等^[11], 每个样本重复 3 次, 表达量的显著性分析用 Duncan (SPSS 19.0) 进行检验。

2 结果与分析

2.1 测序结果与组装

锥栗种仁转录组测序后获得约 5.4 G Clean reads, 其 Q20 值为 96.24%, GC 含量为 43.57%, 符合转录组测序要求。对 Clean Reads 进行组装拼接获得 1837366 个重叠群 (Contig) 序列, 序列长度主要分布在 0 ~ 3000 nt 之间, 平均长度为 69 nt, 其中在 0 ~ 300 nt 序列较多为 1796825 条, 约占总 Contig 序列的 97.79% (图 1A)。

组装获得 53629 条 Unigene, 平均长度为 746 bp (图 1B)。Unigene 序列长度主要分布在 200 ~ 2000 bp 范围内, 其中, 200 ~ 500 bp 范围内序列较多为 32296 条, 占全部 Unigene 序列的 60.22%, 500 ~ 2000 bp 的 Unigene 序列有 16996 条, 占全部 Unigene 序列的 31.69%, 大于 2000 bp 的 Unigene 序列有 4337 条, 占全部 Unigene 序列的 8.09%。

表 1 引物序列

Table 1 Primers used in this study

引物编号	基因名称	引物序列(5'-3')	退火温度(℃)	GC 含量(%)	扩增长度 (bp)
Primers code	Gene name	Primer sequence (5'-3')	Annealing temperature	GC content	Amplification length
CH. 29636-F	淀粉合成酶	AGATGTCCTGTTATTGGATTCA	58.1	57.1	132
CH. 29636-R	Starch synthase	AAGTCTTTTCTGCCAGTGCC	57.1	50.0	
CH. 11971-F	淀粉分支酶 II	AGAGTCAAGTTCCGATTC	45.7	44.1	96
CH. 11971-R	Starch branch enzyme II	TCCAAAAGTTGTAGGGTC	47.4	44.4	
CH. 12978-F	蔗糖合成酶	GAACGATAAGTCAAAGCC	47.2	44.4	81
CH. 12978-R	Sucrose synthase	TACATCTCTACCAACCCAG	47.2	47.4	
CH. 33690-F	蔗糖磷酸合成酶	TTGGCATCTCGCCCGTA	58.5	58.8	137
CH. 33690-R	Sucrose phosphate synthase	TTTCCTTCTCCCGCTCG	55.1	58.8	
CH. 19238-F	果糖激酶	GTCTCCTCTCTTATGACCCAAA	55.4	45.5	125
CH. 19238-R	Fructose kinase	TTCAACTTCTTCCAAACTTACCT	55.2	34.8	
CH. 30128-F	己糖激酶	GGCTAAGGCTATGGCAAT	52.8	50.0	129
CH. 30128-R	Hexokinase	TTACTTCCACCCTCCGATG	55.4	52.6	
CH. 13088-F	糖基转移酶	TGTCGTTTAGACTATTGGCT	52.3	40.0	96
CH. 13088-R	Glycosyl transferase	CTTGAGTTTCCGAGCATAG	53.3	50.0	
CsAtubulin-F	内参基因	GCCACCATCAAGACCAAGAG	62.0	55.0	108
CsAtubulin-R		CAAGTCACCACCAGGAACCA	62.0	55.0	

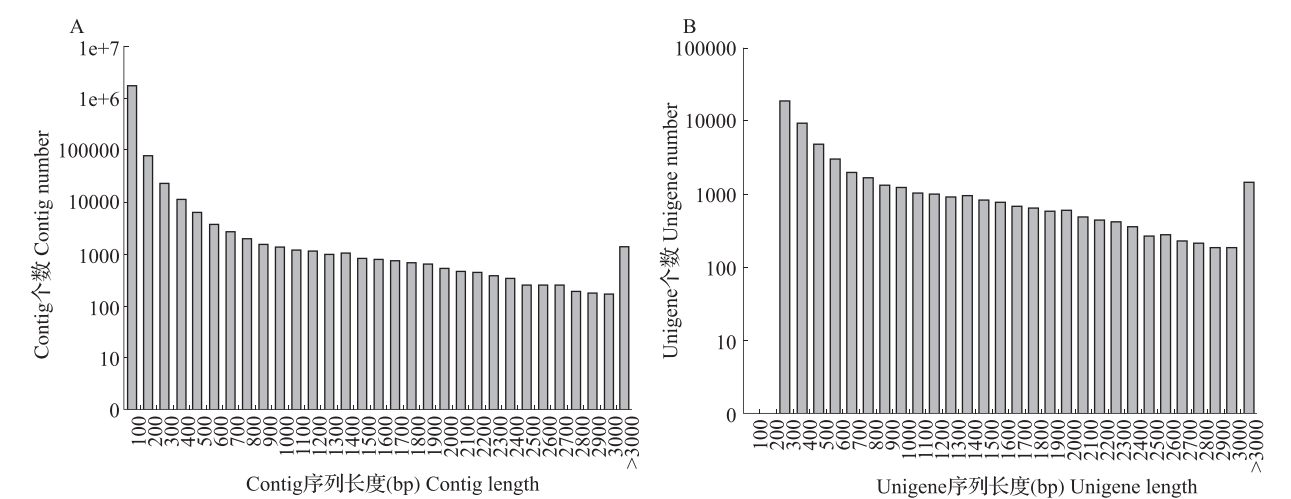


图 1 锥栗 Contig 和 Unigene 的长度分布

Fig. 1 The length distribution of Contig and Unigene for *C. henry*

2.2 Unigene 的功能注释

分别将 Unigene 序列注释到 NR、Nt (E-value < 1e-5)、Swiss-Prot、COG (E-value < 1e-5)、GO 和 KEGG 库,并分别统计每个库中注释上 Unigene 的

数目(表 2)。通过 Blast 搜索比对,共 26739 条 Unigene 序列获得基因注释,占 All-Unigene 的 49.86%。注释的数据库中以 NR 数据库注释最多(26564 条),以 KEGG 数据库注释信息最少,为 5277 条。

表 2 Unigene 注释结果统计表
Table 2 The statistical results of functional annotations in *C. henry*

数据库 Database	注释基因数 No. of annotated genes	占注释基因 百分数 (%) Percentage of annotated genes	占 All-Unigene 百分比 (%) Percentage of All-Unigene
NR	26564	99.35	49.53
SwissProt	17254	64.53	32.17
COG	11447	42.81	21.34
GO	22066	82.52	41.15
KEGG	5277	19.74	9.84
总计 Total	26739	—	49.86

将组装得到的 53629 条锥栗 Unigene 与 NR 数据库比对 (E 值 $\leq 1e-5$, 图 2A), 经 NR 数据库注释的 Unigene 序列与其匹配序列相似性在 60% 以上的占 82.74%, 相似性在 80% 以上的占 45.3% (图 2B)。与蓖麻 (*Ricinus communis*)、毛果杨 (*Populus trichocarpa*) 和可可 (*Vitis vinifera*) 等 9 个物种已知序列相似性的 Unigene 序列数量占 NR 数据库注释总 Unigene 序列数量的 92.17%, 其中蓖麻 (*R. communis*) 所占的比例最高, 为 57.86% (图 2C)。

2.3 Unigene 的功能分类

2.3.1 Unigene 的 COG 分类 与 COGs 数据库比对, 11447 条锥栗 Unigene 找到了同源序列, 分布于

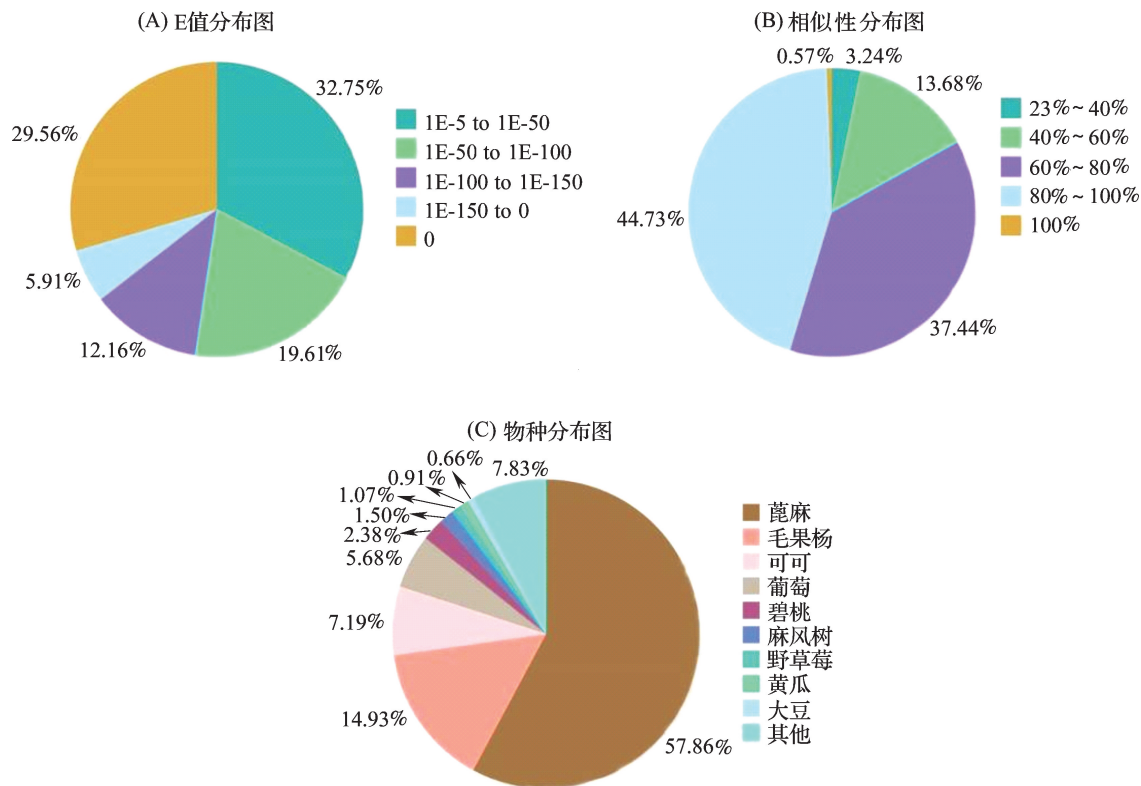


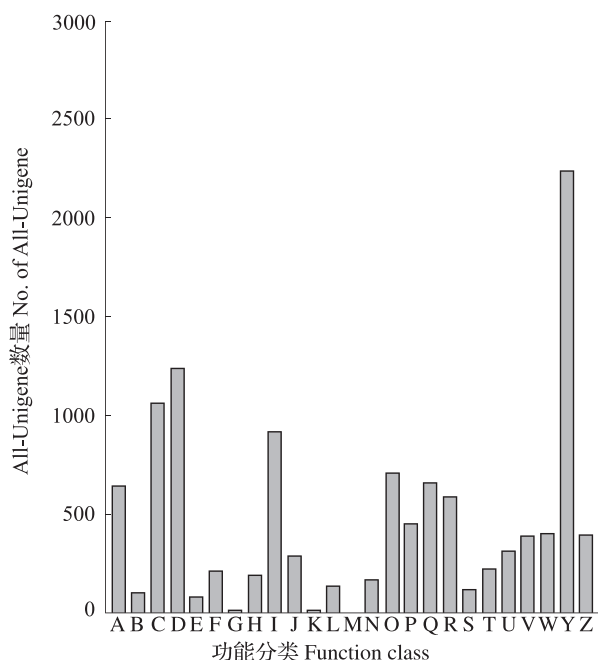
图 2 锥栗种仁转录组数据的 NR 分类图

Fig. 2 NR classification of *C. henry* developing kernel in transcriptome

25 个类别 (图 3)。其中, 与生物过程相关的 Unigene 数量最多, 包括复制、重组和修饰 (1235, 10.79%), 转录 (1056, 9.23%), 翻译后修饰、蛋白翻转、分子伴侣 (702, 6.13%); 与代谢相关的 Unigene 基因数量为 2661 条, 约占总体 Unigene 数量的 23.25%; 此外, 与功能预测类别及功能未知类别的 Unigene 基因数量有 2238 及 391 条, 分别占 19.55% 和 3.42%。

2.3.2 Unigene 的 GO 分类 锥栗 Unigene 的 GO

注释分类结果表明, 22066 条 Unigene 被注释分成细胞组分、分子功能和生物学过程 3 大类别 (图 4), 其 Unigene 序列数量分别占总体 Unigene 序列数量的 13.68%、34.63% 和 51.69%。根据其细胞的组分, 已注释 Unigene 可分为 13 类, 其中细胞 (17431)、细胞要素 (17605)、细胞器 (15717) 及膜功能 (8696) 相关的 Unigene 序列数量占细胞组分分类整体的 77.37%; 根据其分子功能, 可以将 Unigene 分为 15 类, 其中与结合、催化相关的 Unigene



A: 翻译、核糖体结构和生物合成; B: RNA 加工和修饰; C: 转录; D: 复制、重组和修复; E: 染色质结构和动力学; F: 细胞周期控制、细胞分裂和染色体分区; G: 核结构; H: 防御系统; I: 信号转导机制; J: 细胞壁/膜/包膜生物合成; K: 细胞迁移; L: 细胞骨架; M: 细胞外结构; N: 细胞内运输、分泌和膜泡运输; O: 翻译后修饰、蛋白翻转、分子伴侣; P: 能量产生和转化; Q: 碳水化合物转运和代谢; R: 氨基酸转运和代谢; S: 核苷酸转运和代谢; T: 辅酶转运和代谢; U: 脂类转运和代谢; V: 无机离子转运和代谢; W: 次生代谢物生物合成、转运和代谢; Y: 一般功能预测; Z: 功能未知

A: Translation, ribosomal structure and biogenesis, B: RNA processing and modification, C: Transcription, D: Replication, recombination and repair, E: Chromatin structure and dynamics, F: Cell cycle control, cell division and chromosome partitioning, G: Nuclear structure, H: Defense mechanisms, I: Signal transduction mechanisms, J: Cell wall/membrane/envelope biogenesis, K: Cell motility, L: Cytoskeleton, M: Extracellular structures, N: Intracellular trafficking, secretion, and vesicular transport, O: Posttranslational modification, protein turnover and chaperones, P: Energy production and conversion, Q: Carbohydrate transport and metabolism, R: Amino acid transport and metabolism, S: Nucleotide transport and metabolism, T: Coenzyme transport and metabolism, U: Lipid transport and metabolism, V: Inorganic ion transport and metabolism, W: Secondary metabolites biosynthesis, transport and catabolism, Y: General function prediction only, Z: Function unknown

图3 锥栗 Unigene 的 COG 分类

Fig. 3 COG classification for *C. henry* Unigenes

序列数量占主导地位,其比例分别为 43.04% 和 38.62%;根据其参与的生物学过程,可以将其分为 24 类,其中与细胞进程功能相关的序列数量最多 (16703),其次是与代谢过程相关的 Unigene 序列数量

(16569)。生物过程中的代谢过程 (16569)、生物调节 (10024)、发育过程 (8274) 等以及种仁发育相关功能的预测在未来锥栗种仁发育的研究中将会提供基础数据。

2.4 KEGG pathways 分析

根据 KEGG 注释信息,共识别出 5277 个 Unigene 涉及 116 条代谢通路与之关联,选取注释基因比例大于 2% (占有注释基因) 的 Pathway 列于表 3 所示。包括核糖体 (ko03010) 途径、植物激素信号转导 (ko04075) 途径、氧化磷酸化 (ko00190) 途径、淀粉和蔗糖代谢 (ko00500) 途径 (13 条,2.63%) 等。

2.5 转录组中与淀粉和蔗糖代谢相关酶基因分析

将 Unigene 数据注释结果与 Tair (<http://www.arabidopsis.org/>) 数据库中种仁发育相关的基因比对分析,检索到 224 条与种仁发育相关的基因,筛选出拟南芥中 139 条与种仁发育相关基因位点,并以拟南芥的种仁发育基因为参考,鉴定出锥栗种仁发育转录组中 114 个与淀粉和蔗糖代谢相关的基因,其中 36 条为重要淀粉与蔗糖代谢相关酶基因 (表 4)。鉴定出的 36 条重要淀粉与蔗糖代谢基因中,与淀粉合酶 (SS) 及淀粉分支酶 II (SBE) 相关基因各 2 条,与蔗糖合成酶 (SuSy) 相关基因 8 条,与蔗糖磷酸合成酶 (SPS) 相关基因 5 条,与果糖激酶 (FRK) 相关基因 3 条,与己糖激酶 (HXK) 和糖基转移酶 (glycosyltransferase) 相关的基因分别为 6 条和 10 条。

2.6 锥栗种仁发育过程中淀粉和蔗糖代谢相关酶基因表达量分析

对于表 3 中淀粉和蔗糖合成有关的 7 类基因,每一类中选择 PRKM 值最高的进行实时荧光定量 PCR 试验,结果如图 5 所示。与淀粉代谢相关酶 (SS、SBE) 基因 (CH. 29636、CH. 11971) 表达量随着种仁的发育呈逐渐增加的趋势;与蔗糖代谢相关酶 (SuSy、SPS、FPK、HXK、Glycosyltransferase) 基因 (CH. 31302、CH. 33690、CH. 19238、CH. 30128、CH. 13088) 表达量随着种仁的发育呈先上升后下降的趋势;7 类基因表达量均在开花后 109 d 呈现显著性变化,说明开花后 109 d 是锥栗种仁中淀粉和蔗糖相关酶活性变化的活跃时期。

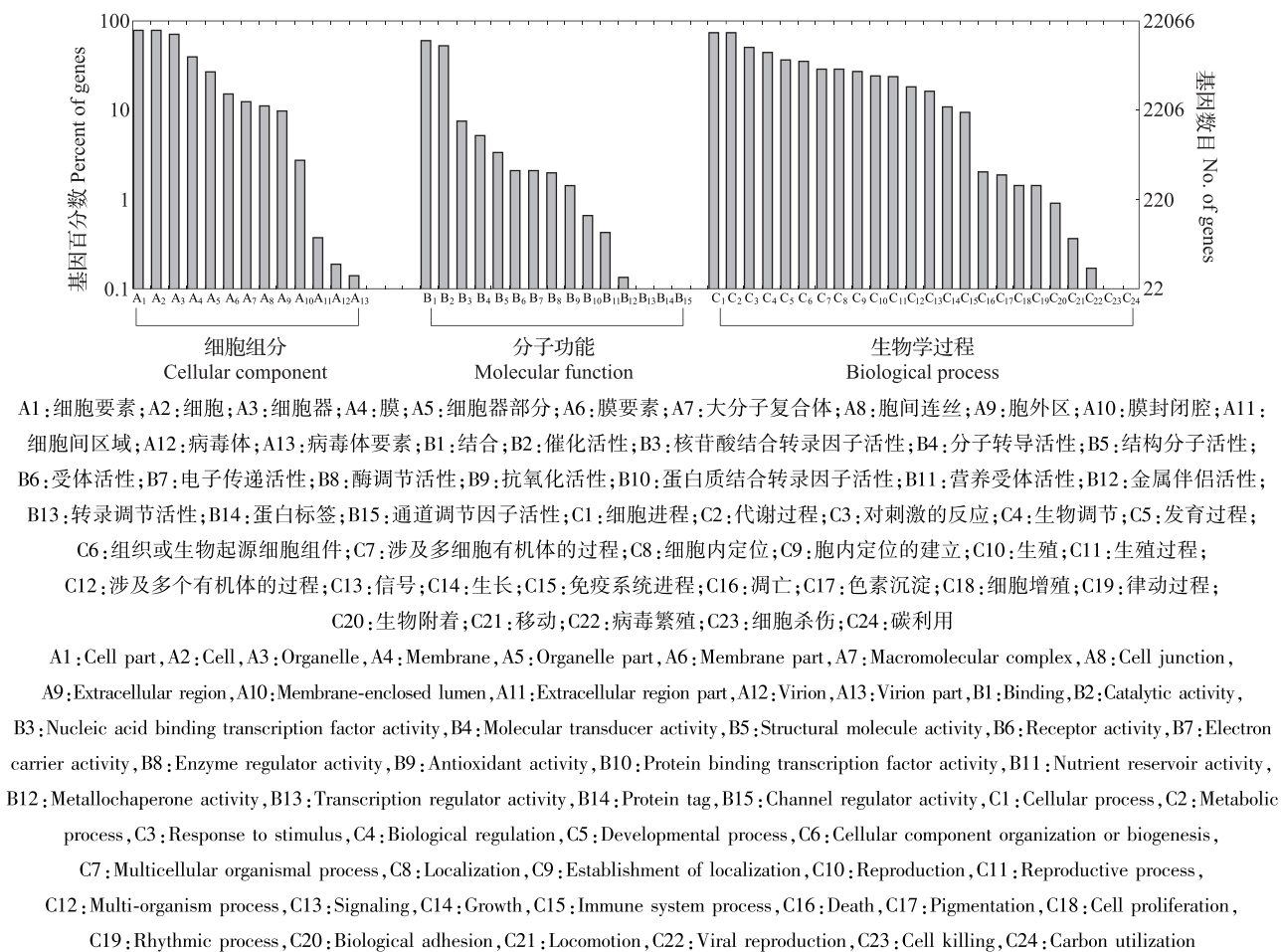


图 4 锥栗 Unigene 的 GO 分类图

Fig. 4 GO classification for *C. henry* Unigenes

表 3 部分锥栗种仁转录组的 KEGG 代谢途径分类

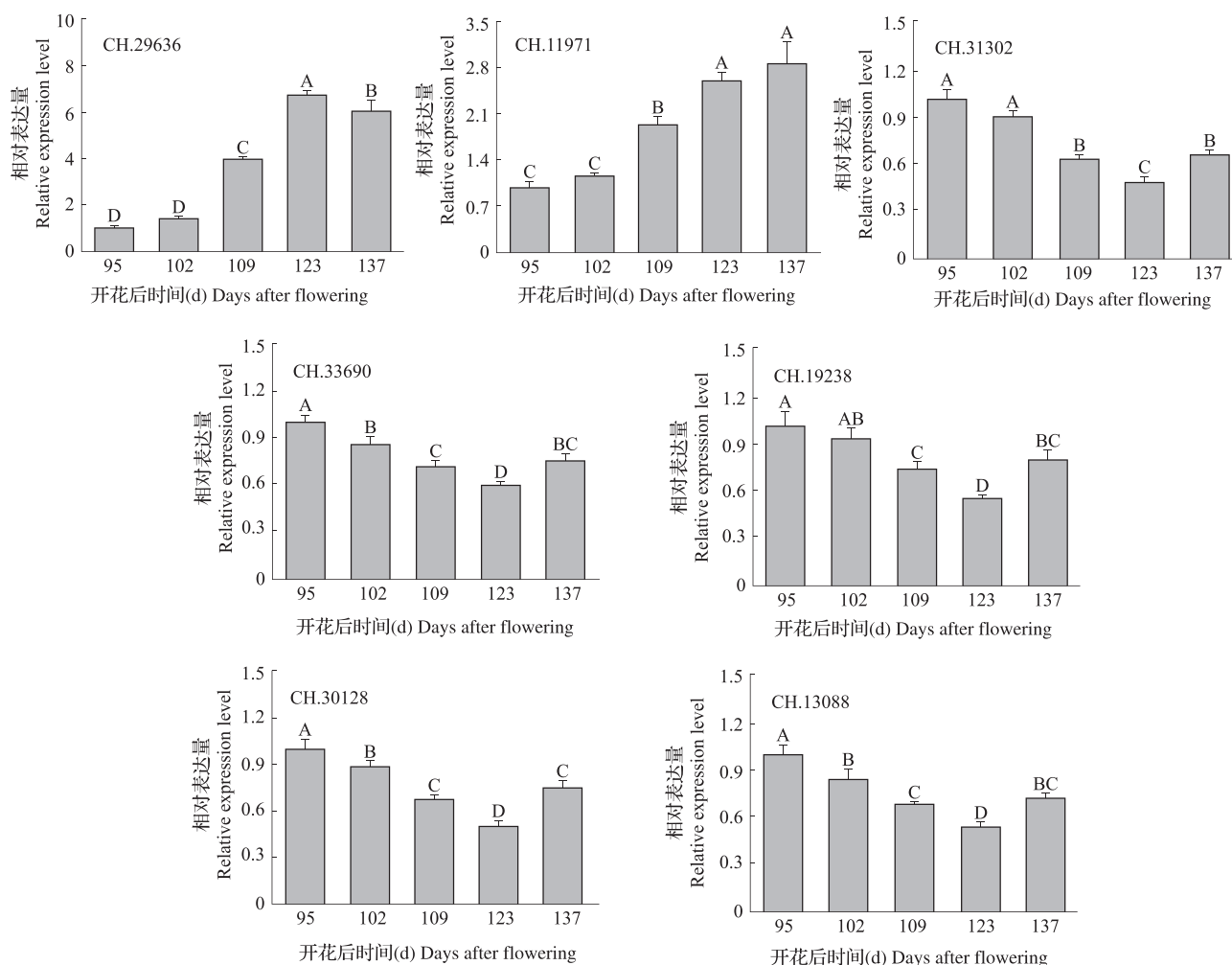
Table 3 Part of KEGG classification of developing kernel in *C. henry*

编号 Code	代谢通路 ID	代谢通路名称 Pathway name	Unigene 数目 No. of Unigenes	所占百分比(%) Percentage of Unigenes
1	ko03010	核糖体 Ribosome	216	4.09
2	ko04075	植物激素信号传导 Plant hormone signal transduction	165	3.13
3	ko03040	剪接体 Spliceosome	149	2.82
4	ko00190	氧化磷酸化 Oxidative phosphorylation	147	2.79
5	ko04141	蛋白质合成 Protein processing in endoplasmic reticulum	143	2.71
6	ko00230	淀粉与蔗糖代谢 Starch and sucrose metabolism	139	2.63
7	ko00500	嘌呤代谢 Purine metabolism	139	2.63
8	ko03013	RNA 转运 RNA transport	138	2.62
9	ko00010	糖酵解/糖原异生 Glycolysis / Gluconeogenesis	131	2.48
10	ko04626	植物—真菌互作 Plant-pathogen interaction	129	2.44
11	ko00240	嘧啶代谢 Pyrimidine metabolism	114	2.16
12	ko03008	核糖体生成 Ribosome biogenesis in eukaryotes	109	2.07
13	ko03015	mRNA 监控通路 mRNA surveillance pathway	108	2.05

表 4 锥栗种仁转录组中与淀粉和蔗糖代谢相关基因的基本信息

Table 4 The basic information of the starch and sugar metabolism related genes for developing kernel in *C. henry*

功能基因名称 Name of functional genes	Unigene 编号 Unigene ID	长度 (bp) Length	CDS 长度 (bp) CDS length	起始碱基位置 Start	终止碱基位置 End	总读序 Total reads	表达量 RPKM	E 期望值 E-value
淀粉合成酶	CH. 29636	1567	1176	142	1317	692	20. 14	0
Starch synthase	CH. 5056	1442	81	128	208	191	6. 04	7. 1E-140
淀粉分支酶 II	CH. 11971	2797	2583	99	2681	2855	46. 56	0
Starch branch enzyme II	CH. 33608	3278	2595	411	3005	2267	31. 54	0
蔗糖合成酶	CH. 12978	3135	2433	289	2721	9054	131. 73	0
Sucrose synthase	CH. 2443	588	42	149	190	6	0. 47	6E-123
	CH. 2506	235	300	118	417	1	0. 19	3. 08E-25
	CH. 31302	1063	846	207	1052	10380	445. 38	1. 6E-176
	CH. 31303	442	249	193	441	360	37. 15	3. 12E-31
	CH. 52037	1077	999	9	1007	7	0. 30	2E-167
	CH. 9333	792	759	34	792	3	0. 17	1. 1E-151
	CH. 9594	2063	1761	280	2040	70	1. 55	0
蔗糖磷酸合成酶	CH. 30899	1566	1452	1	1452	40	1. 17	0
Sucrose phosphate synthase	CH. 33690	3791	3066	90	3155	1980	23. 82	0
	CH. 35364	786	561	224	784	13	0. 75	2. 7E-157
	CH. 35365	1272	1191	81	1271	52	1. 86	0
	CH. 41953	1284	1086	180	1265	47	1. 67	0
果糖激酶	CH. 19238	698	507	171	677	29	1. 90	1. 72E-94
Fructose kinase	CH. 19239	771	522	1	522	20	1. 18	7. 14E-68
	CH. 28321	278	117	138	254	9	1. 48	1. 12E-45
己糖激酶	CH. 18306	1900	1524	23	1546	529	12. 70	0
Hexokinase	CH. 26652	809	654	156	809	51	2. 88	4. 5E-131
	CH. 26653	230	216	15	230	3	0. 59	8. 38E-42
	CH. 30128	2070	1494	386	1879	1356	29. 88	0
	CH. 40518	1957	1485	74	1558	264	6. 15	0
	CH. 5272	1027	588	340	927	98	4. 35	1. 2E-123
糖基转移酶	CH. 20302	2232	1602	380	1981	431	8. 81	0
Glycosyl transferase	CH. 13088	2393	1671	697	2367	2410	45. 93	0
	CH. 19415	2175	1890	18	1907	260	5. 45	0
	CH. 27384	2502	1935	370	2304	1238	22. 57	0
	CH. 28874	1867	1605	29	1633	323	7. 89	0
	CH. 32528	2169	1686	212	1897	426	8. 96	0
	CH. 32865	2643	1605	278	1882	542	9. 35	0
	CH. 33371	2616	1827	365	2191	747	13. 02	0
	CH. 36120	2726	1440	517	1956	489	8. 18	0
	CH. 37470	2030	1599	258	1856	245	5. 50	0



不同字母表示差异达到极显著水平 ($P = 0.01$) Different letters indicate significant difference at $P = 0.01$ level

图5 锥栗种仁发育过程中7个与淀粉和蔗糖代谢相关酶基因的表达量分析

Fig. 5 Expression analysis of seven starch and sucrose metabolism related enzyme gene for developing kernel in *C. henry*

3 讨论

新一代高通量 RNA-seq 测序技术具有高效、高精度、无需内参基因组信息等特点,为非模式植物分子基础平台的建立提供了便捷的途径^[12],大量非模式植物借助此技术进行了转录组研究^[13-18]。果实品质形成机理的研究是调控果实品质的前提^[19]。果实发育是一个高度协调的复杂生物学过程,尽管锥栗果实和成熟过程的研究得到了重视,但是对其分子调控机制的诠释还远远不够。因此,对锥栗果实进行高通量测序分析意义重大。本研究首次利用 Illumina 高通量测序技术对锥栗种仁转录组进行了测序分析,获得 Unigene 序列 53629 条,为进一步挖掘并鉴定新的功能基因提供了数据信息。共注释 Unigene 序列 26739 条,占 All-Unigene 的 49.86%,为壳斗科隶属其他植物种提供注释信息参考。

22066 条得到 GO 注释的锥栗 Unigene 序列中,其中归入细胞 (17431 条 Unigene, 7.86%) 和细胞组分 (17605 条 Unigene, 7.94%) 类别在 58 个小类别中所占比例较高,说明有较多的基因与细胞和细胞组分的代谢相关。KEGG 结果表明,共识别出 5277 个 Unigene 涉及 116 条代谢通路,其中参与锥栗种仁淀粉和蔗糖代谢的 Unigene 共 13 条,鉴定出了与淀粉和蔗糖代谢相关关键酶基因 36 条,分属于 7 类酶。这些有代表性的注释为研究锥栗果实营养物质积累和转换过程、淀粉合成途径及其他特殊生物学进程提供了重要依据。

淀粉和蔗糖的积累和转换不仅与果实发育物质基础的形成有关,而且也与坚果产量和品质密切相关^[20]。范晓明等^[9]对锥栗种仁发育期种仁中的淀粉及蔗糖含量进行了研究,结果表明锥栗种仁发育期,种仁中淀粉含量呈现先上升后微弱下降的变化

规律,而蔗糖含量则呈先下降后上升的趋势,为锥栗种仁发育过程中淀粉和蔗糖的变化规律提供了参考。前人研究表明果实发育期间淀粉和蔗糖的代谢与其相对应酶的活性密切相关^[21-22]。谢鹏等^[23]对板栗果实糖和淀粉积累及相关酶活性的关系进行了研究,发现水溶性淀粉合成酶活性与淀粉含量呈显著正相关,蔗糖合成酶和蔗糖磷酸合成酶活性与蔗糖存在极显著正相关。本研究结果表明:(1)锥栗种仁在开花后 109 d 以前,SS 基因(CH. 29636)相对低水平表达,而与催化蔗糖合成的两种酶 SuSy (CH. 31302)及 SPS (CH. 33690)基因表达水平相对较高,这与范晓明等^[9]研究发现锥栗种仁发育在开花后 109 d 以前,种仁中可溶性糖的积累含量显著高于淀粉含量的结论相吻合。这是因为高活性的催化蔗糖合成的两种酶促进了蔗糖的形成,而在淀粉合成酶活性较低条件下,淀粉积累较少;(2)锥栗种仁中 SS (CH. 29636)和 (CH. 11971)基因表达量呈现逐渐上升的趋势。淀粉分支酶是合成支链淀粉的一类关键水解酶,其产生分支的主要作用机理为切开 α -1,4-糖苷键(α -1,4-葡聚糖直链供体),催化所切下的短链与受体链间 α -1,6-糖苷键的形成^[24-25]。随着锥栗种仁的发育,Amy 与 SS 活力逐渐增强,淀粉总量不断增加。(3)锥栗种仁中 SuSy (CH. 31302)、SPS (CH. 33690)、FPK (CH. 19238)、HXK (CH. 30128)、Glycosyltransferase (CH. 13088)5 种酶基因均呈现先上升后下降的趋势。锥栗种仁发育前期,随着 SuSy 和 SPS 活性不断增加,蔗糖积累量迅速增多;当种仁中的蔗糖积累到一定量时,蔗糖将被转化酶分解为葡萄糖和果糖,而这些果糖在进一步代谢前必须由 FPK 或 HXK 进行磷酸化^[26],因此 FPK 和 HXK 酶基因的表达量增加,即酶的活力增大。这些研究为了解锥栗果实品质形成的机理提供了科学依据,通过系统的植物生理学和分子生物学进一步揭示锥栗种仁中淀粉和蔗糖的代谢规律将是后续研究的重要内容。

参考文献

- [1] 靳智. 锥栗抗性淀粉的制备及其性质研究[D]. 福建:福建农林大学,2012
- [2] 宋爱云,陈钦,雷日平,等. 锥栗栽培品种果实营养成分差异的分析[J]. 经济林研究,2001,19(4):5-7
- [3] 景芸,梁一池,杨华. 不同锥栗无性系果实营养成分的比较分析[J]. 浙江林学院学报,2004,21(2):176-179
- [4] 陈钦,邹双全. 锥栗干制研究[J]. 经济林研究,2002,20(2):84-86
- [5] 雷日平,陈辉,谢利国. 锥栗不同品种遗传距离的 RAPD 分析[J]. 浙江林学院学报,2002,19(3):240-243
- [6] Tuskan G A, Difazio S, Jansson S, et al. The genome of black cotton wood, *Populus trichocarpa* (Torr. & Gray) [J]. Science, 2006,313(5793):1596-1604
- [7] Wilhelm B T, Landry J R. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing [J]. Methods, 2009,48(3):249-257
- [8] Chen Y H, Chen J H, Luo Y M. Complementary biodiesel combination from tung and medium-chain fatty acid oils [J]. Renew Energ, 2012,44:305-310
- [9] 范晓明,袁德义,段经华,等. 锥栗种仁发育期叶片与果实矿物质元素含量变化[J]. 园艺学报,2014,2014,41(1):44-52
- [10] 江香梅,伍艳芳,肖复明,等. 樟树 5 种化学类型叶片转录组分析[J]. 遗传,2014,36(1):58-68
- [11] 周俊琴,谭晓风,袁军,等. 油茶 *Phl1*; 1 基因克隆及其表达分析[J]. 植物遗传资源学报,2013,14(3):512-517
- [12] Kuang Q, Li L F, Peng J Z, et al. Transcriptome analysis of *Gerbera hybrida* ray florets; Putative genes associated with gibberellin metabolism and signal transduction [J]. PLoS ONE, 2013,8(3):e57715
- [13] Novaes E, Drost D R, Farmerie W G, et al. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome [J]. BMC Genomics, 2008,9(1):312
- [14] Barakat A, Dilorieto D S, Zhang Y, et al. Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection [J]. BMC Plant Biol, 2009,9(1):51
- [15] Parchman T L, Geist K S, Grahnen J A, et al. Transcriptome sequencing in an ecologically important tree species; assembly, annotation, and marker discovery [J]. BMC Genomics, 2010,11(1):180
- [16] Cabrera A R, Donohue K V, Khalil S M S, et al. New approach for the study of mite reproduction; The first transcriptome analysis of a mite, *Phytoseiulus persimilis* (Acari:Phytoseiidae) [J]. J Insect Physiol, 2011,57(1):52-61
- [17] Bai X D, Rivera-Vega L, Mamidala P, et al. Transcriptomic signatures of ash (*Fraxinus* spp.) phloem [J]. PLoS ONE, 2011,6(1):e16368
- [18] Torales S L, Rivarola M, Pomponio M F, et al. Transcriptome survey of Patagonian southern beech *Nothofagus nervosa* (= *N. Alpina*); assembly, annotation and molecular marker discovery [J]. BMC Genomics, 2012,13(1):291
- [19] 余克琴. 甜橙果实发育与成熟过程中转录组变化及 *CsASR* 基因的功能分析[D]. 福州:福建农林大学,2012
- [20] 柴叶茂,贾海峰,李春丽,等. 草莓果实发育过程中糖代谢相关基因的表达分析[J]. 园艺学报,2011,38(4):637-643
- [21] Vizzolo G, Pinton R, Varanmin Z, et al. Sucrose accumulation in developing peach fruit [J]. Physiol Plant, 1996,96(2):225-230
- [22] Wanraya S, Toshiyuki M, Kazuhide K, et al. Seasonal fluctuations of sucrose metabolizing enzymes activities and sugar contents in lettuce (*Lactuca sativa* L.) [J]. J Biol Sci, 2007,7(5):752-760
- [23] 谢鹏,郭素娟,熊欢,等. 板栗果实糖和淀粉积累及相关酶活性关系的研究[J]. 园艺学报,2012,39(12):2369-2376
- [24] Visser R G F, Jacobsen E. Towards modifying plants for altered starch content and composition [J]. Trends Biotechnol, 1993,11(2):63-68
- [25] 郭新梅,张晓东,梁荣奇,等. 玉米淀粉分支酶基因的克隆与 RNAi 表达载体的构建[J]. 西北植物学报,2007,27(11):217-2180
- [26] Yamaki S. Physiology and metabolism of fruit development biochemistry of sugar metabolism and compartmentation in fruits [J]. Postharv Physiol Fruits, 1994,398:109-120