

花生异型雄蕊转录组分析

王海霞^{1a,1b,2}, 王铭伦², 丁雨龙^{1a,1c,①}

(1. 南京林业大学: a. 南方现代林业协同创新中心, b. 生物与环境学院, c. 竹类研究所, 江苏 南京 210037;
2. 青岛农业大学, 山东 青岛 266109)

摘要: 以花生 (*Arachis hypogaea* Linn.) 品种 ‘Florunner’ 为研究对象, 利用 RNA-seq 技术对花生异型雄蕊进行转录组测序; 在此基础上, 运用 Nr、KOG、GO 和 KEGG 数据库进行同源序列比对和基因功能分析。结果显示: 花生异型雄蕊转录组测序共获得数据量约 31.46 G, 其中, clean reads 所占比例均在 96.0% 以上; 经过组装和拼接后得到 197 193 个 contigs, 片段总长度 240 463 900 bp, 最长片段为 15 810 bp, 非冗余 unigenes 片段的平均长度为 1 219 bp, GC 含量为 39.1%, 片段长度 N50 为 1 966 bp; 长度大于等于 1 000 bp 的 unigenes 占 unigenes 总数的 43.4%, 表明组装完整性较高, 可用于注释分析。与 Nr 数据库的比对结果显示: 共有 129 223 个 unigenes 注释到 Nr 数据库, 占 unigenes 总数的 65.5%; 其中, $1 \times 10^{-100} < E \leq 1 \times 10^{-5}$ 的 unigenes 占 54.2%, 说明注释的 unigenes 序列与 Nr 数据库的同源序列具有高度相似性; 共有 106 900 个 unigenes 注释到已探明的同源序列中, 占注释 unigenes 总数的 82.7%, 且与大豆 [*Glycine max* (Linn.) Merr.] 和蒺藜苜蓿 (*Medicago truncatula* Gaertn.) 等同科种类的序列同源性较高。基因功能分析结果显示: 共有 133 387 个 unigenes 注释到 25 个 KOG 功能类别上, 且功能明确的 unigenes 均与配子体发育功能相关, 其中, 注释到信号转导机制功能的 unigenes 数量最多 (17 135)。共有 48 497 个 unigenes 注释到生物过程、细胞组分和分子功能 3 大类 55 个亚类的 GO 功能类别上, 其中, 与核酸代谢过程、磷酸化、转录调节、核、核糖体、ATP 结合、转录因子和钙离子结合相关的亚类中 unigenes 高度富集。共有 30 428 个 unigenes 注释到 2 708 个酶功能, 映射了 354 个 KEGG 代谢通路, 可分为遗传信息处理、细胞过程、生物系统、环境信息处理和代谢 5 大类 32 个亚类, 其中, 与代谢相关通路的 unigenes 数量最多 (14 231)。综合分析结果表明: 在花生异型雄蕊转录组中, 多数 unigenes 的功能与雄蕊产生雄配子的功能有关, 且部分基因可参与不同的代谢途径, 对花生异型雄蕊的器官发生和花药形成具有重要作用。

关键词: 花生; 异型雄蕊; 转录组; 同源性; 功能注释

中图分类号: Q945.4; Q946-33; S565.2 文献标志码: A 文章编号: 1674-7895(2020)05-0021-09

DOI: 10.3969/j.issn.1674-7895.2020.05.03

Analysis on transcriptome from heteromorphic stamens of *Arachis hypogaea* WANG Haixia^{1a,1b,2}, WANG Minglun², DING Yulong^{1a,1c,①} (1. Nanjing Forestry University; a. Co-Innovation Center for Sustainable Forestry in Southern China, b. College of Biology and the Environment, c. Bamboo Research Institute, Nanjing 210037, China; 2. Qingdao Agricultural University, Qingdao 266109, China), *J. Plant Resour. & Environ.*, 2020, 29(5): 21-29

Abstract: Taking cultivar ‘Florunner’ of *Arachis hypogaea* Linn. as research object, transcriptome sequencing was conducted on its heteromorphic stamens by using RNA-seq technology; on the basis, homologous sequence alignment and gene function analysis were conducted by using Nr, KOG, GO and KEGG databases. The results show that about 31.46 G of data are obtained by transcriptome sequencing of heteromorphic stamens of *A. hypogaea*, in which, clean reads account for above 96.0%; 197 193 contigs are obtained after assembling and splicing, and the total length of fragments is 240 463 900 bp, the

收稿日期: 2019-08-29

基金项目: 国家花生产业技术体系建设项目 (CARS-14); 江苏高校优势学科建设工程 (PAPD)

作者简介: 王海霞 (1979—), 女, 山东济宁人, 博士, 副教授, 主要从事植物发育生物学方面的研究。

①通信作者 E-mail: ylding@vip.163.com

longest fragment is 15 810 bp; the average length of non-redundant unigenes fragment is 1 219 bp, the GC content is 39.1%, and the fragment length N50 is 1 966 bp; unigenes with a length greater than or equal to 1 000 bp account for 43.4% of total unigenes number, indicating that the assembly integrity is relatively high, and can be used for annotation analysis. The alignment result with Nr database shows that there are 129 223 unigenes annotated in Nr database, accounting for 65.5% of total unigenes number; in which, unigenes of $1 \times 10^{-100} < E \leq 1 \times 10^{-5}$ account for 54.2%, indicating that annotated unigenes sequences have high similarity with homologous sequences in Nr database; there are 106 900 unigenes annotated in known homologous sequences, accounting for 82.7% of total annotated unigenes number, and have high sequence homology with plants of the same family, such as *Glycine max* (Linn.) Merr. and *Medicago truncatula* Gaertn., etc. The gene function analysis result shows that there are 133 387 unigenes annotated in 25 KOG functional categories, and functional unigenes are all associated with gametophyte development function, in which, number of unigenes annotated in signal transduction mechanism function is the largest (17 135). There are 48 497 unigenes annotated in 55 subcategories belonging to 3 GO functional categories including biological process, cellular component and molecular function, in which, unigenes are highly enriched in subcategories related to nucleic acid metabolic process, phosphorylation, transcriptional regulation, nucleus, ribosome, ATP binding, transcription factor and calcium binding. There are 30 428 unigenes annotated in 2 708 enzyme function mapping to 354 KEGG metabolic pathways, which can be divided into 32 subcategories belonging to 5 categories including genetic information processing, cellular process, organismal system, environmental information processing and metabolism, in which, number of unigenes in pathway related to metabolism is the largest (14 231). The comprehensive analysis result shows that in the transcriptome from heteromorphic stamens of *A. hypogaea*, function of most unigenes is associated with the function of stamens producing male gametes, and some genes are involved in different metabolic pathways, and play an important role in the organogenesis of heteromorphic stamens and formation of anther of *A. hypogaea*.

Key words: *Arachis hypogaea* Linn.; heteromorphic stamen; transcriptome; homology; functional annotation

花生(*Arachis hypogaea* Linn.)为严格的闭花受精植物,每朵花内有10枚雄蕊,其中,具4枚长椭圆球状花药的雄蕊为对萼长药雄蕊(anteseptal stamen, Sts),药室较大,花丝较长;具4枚圆球状花药的雄蕊为对瓣圆药雄蕊(antepetal stamen, Stp),药室较小,花丝较短;另2枚雄蕊为退化雄蕊(staminode, Sta),仅存花丝,未见花药。花生具有的这类雄蕊为异型雄蕊^[1]。异型雄蕊现象从18世纪开始就引起了科学家们的关注^[2-3],一些研究者针对异型雄蕊的种类分布^[4-7]以及花粉活性、花粉组织化学成分和结实率等^[8-11]方面开展了广泛的研究,然而,异型雄蕊的遗传基础和分子调控机制目前仍未知。

花生的基因组较大、遗传基础狭窄,且其花器官的性状稳定、遗传多态性极低,因而,花生的花器官突变体很少,限制了花生的花发育分子生物学研究。雄蕊发育在被子植物的花器官进化中起关键作用,但是与其他花器官(如花瓣)相比,其分子调控机制尚处于探索阶段,对花生异型雄蕊发育的分子机制进行探讨可为雄蕊分化的分子机制研究提供思路。

作者采用RNA-seq技术对花生异型雄蕊进行转

录组分析,挖掘和探究与花生雄蕊分化相关的候选基因和生物学途径,以期解释花生异型雄蕊的形成机制,为花生新品种选育提供理论基础。

1 材料和方法

1.1 材料

供试花生品种为青岛农业大学花生研究中心王铭伦教授提供的品种‘Florunner’,该品种由品种‘Early Runner’与‘Florispan’杂交获得,为普通型蔓生品种,花多且花型较大^[12]。于2017年4月末将品种‘Florunner’的种子覆膜种植于青岛农业大学试验田中,并于同年6月中旬(盛花期)选取100株发育良好的植株,每天约17:00时于各植株同一位置采集发育成熟、待开放的花蕾(即花瓣已撑破萼片、微露黄色花瓣),共采集约400枚花蕾。

将采集的花蕾置于冰盒内,在解剖镜下将雄蕊分为对萼长药雄蕊、对瓣圆药雄蕊和退化雄蕊3组,每纽约800枚雄蕊(每组设置3个生物学重复);将雄蕊分别置于液氮中速冻,-80℃保存,送至上海伯豪

生物技术有限公司进行 RNA 提取及测序分析。

1.2 方法

1.2.1 总 RNA 的提取和检测 采用 RNAqueous™ phenol-free total RNA 提取试剂盒(美国 Ambion 公司)分别提取 3 组雄蕊的总 RNA,获得的总 RNA 经 Agilent Bioanalyzer 2100 生物分析仪(美国 Agilent Technologies 公司)电泳检测合格,使用 RNAClean XP Kit 试剂盒(美国 Beckman Coulter 公司)和 RNase-Free DNase Set 试剂盒(德国 Qiagen 公司)进行纯化。

1.2.2 文库制备和测序 对纯化的总 RNA 进行 mRNA 的富集和片段化,采用 SuperScript II Reverse Transcriptase 试剂盒(美国 Invitrogen 公司)合成 cDNA 第 1 链,然后合成 cDNA 第 2 链,随之进行末端修复,在 3'末端加腺苷酸并连接测序接头,富集、纯化后构建 cDNA 文库;采用 Illumina HiSeq™ 2000 高通量测序平台(美国 Illumina 公司)对 cDNA 文库进行双端测序。

1.2.3 原始数据统计和功能注释 原始测序文件经碱基识别,应用 Seqtk 软件进行误差过滤;将 3 组雄蕊样本的测序数据合并形成 pool reads,然后应用 Trinity 软件^[13-14]进行 *de novo* 拼接;将拼接得到的 final unigenes 序列与 NCBI 官方非冗余蛋白质序列数据库(Nr)和 UniProt 数据库进行 BLASTx 比对($E \leq 1 \times 10^{-5}$)^[15]。

应用 rpstblastn 程序(<http://www.biomedcentral.com/content/supplementary/1471-2105-13-42-s1/Cloud-BioLinux-Package-Documentation/docs/rpstblastn.html>)将 unigenes 序列与真核生物直系同源序列数据库(KOG)进行比对($E \leq 1 \times 10^{-5}$),取排名前 5 的比对结果进行 KOG 功能分析^[16-17]。将 unigenes 序列与基因功能分类体系数据库(GO)进行 BLASTx 比对($E \leq 1 \times 10^{-5}$),取最佳比对结果进行 GO 功能分析^[18]。

应用 KEGG KAAS 在线 pathway 比对分析工具(https://www.genome.jp/kaas-bin/kaas_main)对 unigenes 进行 KEGG 代谢途径分析^[19]。

2 结果和分析

2.1 测序结果和数据组装结果分析

测序结果显示:测序样品的 Q20 值均大于 96%,数据量约 31.46 G, clean reads 所占比例均在 96.0%以

上。说明转录组测序质量较高,可以用于 *de novo* 拼接。

对 clean reads 进行转录组装和拼接,得到 197 193 个 contigs,片段总长度为 240 463 900 bp,最长片段为 15 810 bp,非冗余 unigenes 片段的平均长度为 1 219 bp,GC 含量为 39.1%,片段长度 N50 为 1 966 bp。统计结果显示:长度大于等于 2 000 bp 的 unigenes 有 38 740 个,占 unigenes 总数的 19.6%;长度小于 2 000 bp 但大于等于 1 000 bp 的 unigenes 有 46 857 个,占 unigenes 总数的 23.8%;长度小于 1 000 bp 但大于等于 200 bp 的 unigenes 有 111 596 个,占 unigenes 总数的 56.6%。总体上长度大于等于 1 000 bp 的 unigenes 占 unigenes 总数的 43.4%,表明组装完整性较高,可用于注释分析。

2.2 unigenes 功能注释分析

2.2.1 序列相似性分析 花生异型雄蕊转录组 unigenes 序列与 Nr 数据库的比对结果(图 1)显示:在花生异型雄蕊转录组中共有 129 223 个 unigenes 注释到 Nr 数据库,占 unigenes 总数的 65.5%;其中,与 Nr 数据库中的同源序列相比, E 值较低($1 \times 10^{-30} < E \leq 1 \times 10^{-5}$)的 unigenes 占 unigenes 总数的 19.5%, E 值极低($1 \times 10^{-100} < E \leq 1 \times 10^{-30}$)的 unigenes 占 unigenes 总数的 34.7%,表明注释的 unigenes 序列与 Nr 数据库的同源序列具有高度相似性。

在 Nr 数据库中对花生异型雄蕊转录组 unigenes 与其他植物种类转录组的注释 unigenes 进行同源序列匹配,结果见图 2。结果显示:在花生异型雄蕊转录组中共有 106 900 个 unigenes 注释到已探明的同源序列中,占注释 unigenes 总数的 82.7%。其中,与大豆 [*Glycine max* (Linn.) Merr.] 的 32 466 个 unigenes 序列相似,占注释 unigenes 总数的 25.1%,同源性最高;与蒺藜苜蓿 (*Medicago truncatula* Gaertn.) 的 17 343 个 unigenes 序列相似,占注释 unigenes 总数的 13.4%;与菜豆 (*Phaseolus vulgaris* Linn.) 的 12 710 个 unigenes 序列相似,占注释 unigenes 总数的 9.8%;与小豆 [*Vigna angularis* var. *angularis* (Willd.) Ohwi et H. Ohashi] 的 9 804 个 unigenes 序列相似,占注释 unigenes 总数的 7.6%;与赤豆 [*Phaseolus angularis* (Willd.) Ohwi et H. Ohashi] 的 3 958 个 unigenes 序列相似,占注释 unigenes 总数的 3.1%;与花生(Nr 数据库)的 3 771 个 unigenes 序列相似,占注释 unigenes 总数的 2.9%;与百脉根 (*Lotus corniculatus* Linn.) 的

3 273个 unigenes 序列相似,占注释 unigenes 总数的 2.5%;与葡萄 (*Vitis vinifera* Linn.)、桃 (*Amygdalus persica* Linn.)和可可 (*Theobroma cacao* Linn.)等 13 种植物的 unigenes 序列同源性均较低,在已注释 unigenes 总数中所占比例均在 2.0%以下。

2.2.2 KOG 功能分析 在 KOG 数据库中对花生异型雄蕊转录组 unigenes 序列进行功能比对,结果见图

3。结果显示:在花生异型雄蕊转录组 unigenes 序列中,共有 133 387 个 unigenes 注释到 25 个 KOG 功能类别上,其中有 8 776 个 unigenes 注释到未知功能 (function unknown),占注释 unigenes 总数的 6.6%,这些 unigenes 的功能仍不明确。

功能明确的 unigenes 均与配子体发育功能有关。其中,有 17 135 个 unigenes 注释到信号转导机制

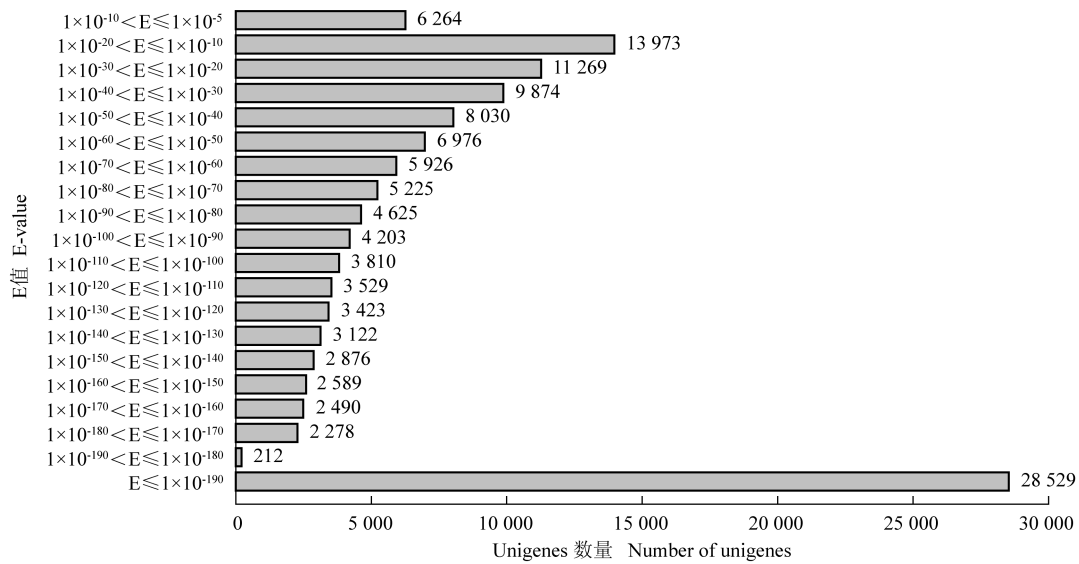


图 1 花生异型雄蕊转录组 unigenes 序列的 E 值分布图
 Fig. 1 E-value distribution diagram of unigenes sequences in transcriptome from heteromorphic stamens of *Arachis hypogaea* Linn.

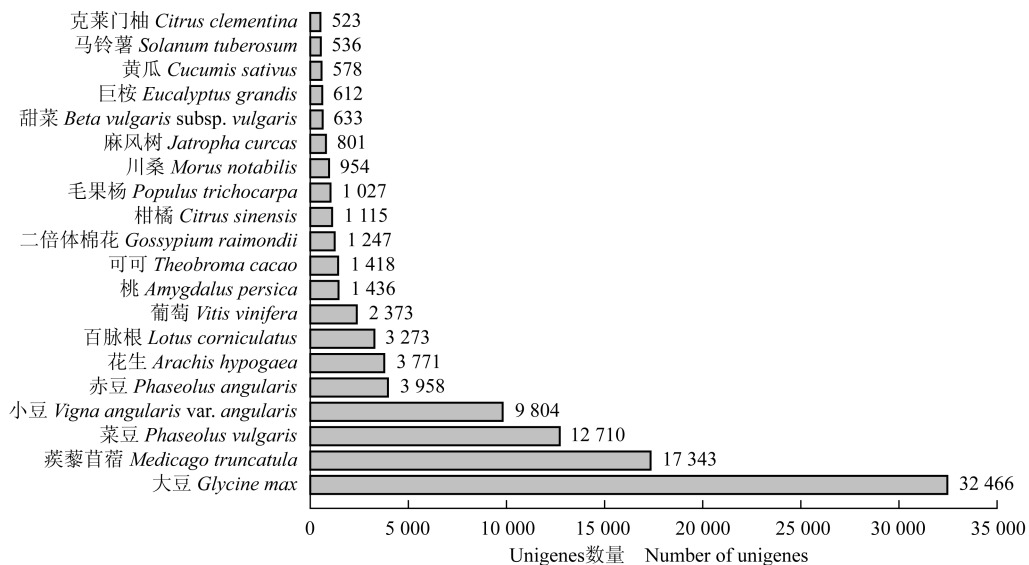


图 2 花生异型雄蕊转录组 unigenes 与其他植物转录组 unigenes 的同源序列数量分布图
 Fig. 2 Distribution diagram of number of homologous sequences of unigenes between the transcriptome from heteromorphic stamens of *Arachis hypogaea* Linn. and those from other plants

(signal transduction mechanism) 功能, 占注释 unigenes 总数的 12.8%; 有 16 407 个 unigenes 注释到一般功能预测 (general function prediction only), 占注释 unigenes 总数的 12.3%; 有 9 893 个 unigenes 注释到翻译后修饰、蛋白质折叠和分子伴侣 (posttranslational modification, protein turnover and chaperone) 功能, 占注释 unigenes 总数的 7.4%; 有 8 870 个 unigenes 注释

到转录 (transcription) 功能, 占注释 unigenes 总数的 6.6%; 有 5 138 个 unigenes 注释到细胞周期调控、细胞分裂和染色体分离 (cell cycle control, cell division and chromosome partitioning) 功能, 占注释 unigenes 总数的 3.9%; 有 2 710 个 unigenes 注释到染色体结构和活力 (chromatin structure and dynamic) 功能, 占注释 unigenes 总数的 2.0%。

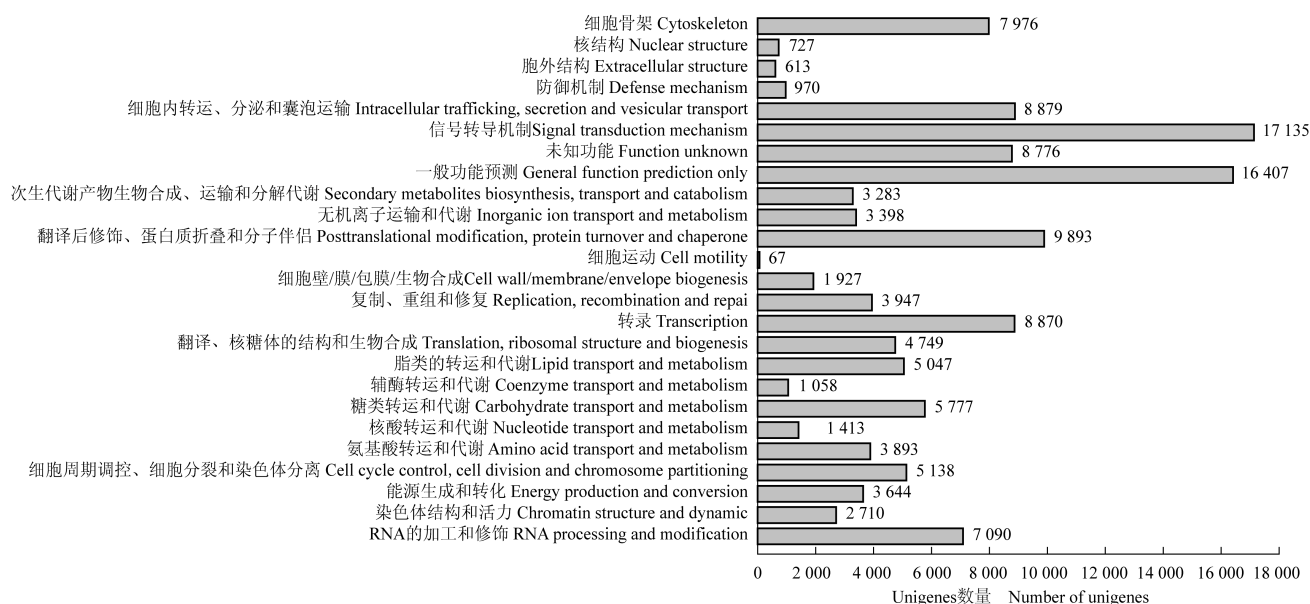


图 3 花生异型雄蕊转录组 unigenes 的 KOG 功能分类

Fig. 3 KOG function classification of unigenes in transcriptome from heteromorphic stamens of *Arachis hypogaea* Linn.

2.2.3 GO 功能分析 在 GO 数据库中对花生异型雄蕊转录组 unigenes 序列进行功能比对, 结果见表 1。结果显示: 在花生异型雄蕊转录组 unigenes 序列中, 注释到生物过程、细胞组分和分子功能 3 大类 55 个亚类的 unigenes 共有 48 497 个。

在生物过程的 23 个亚类中, 分别有 31 817、25 479、16 828、6 327、6 192 和 5 866 个 unigenes 注释到代谢过程 (metabolic process)、细胞过程 (cellular process)、单一生物过程 (single-organism process)、生物调节 (biological regulation)、应激反应 (response to stimulus) 和生物过程调节 (regulation of biological process) 功能, 数量较多, 各占注释 unigenes 总数的 65.6%、52.5%、34.7%、13.0%、12.8% 和 12.1%。其中, 在核酸代谢过程 (GO: 0090304, 8334 个)、磷酸化 (GO: 0016310, 4428 个) 和转录调节 (GO: 0006355, 2322 个) 等与代谢过程相关的 GO 功能类别中

unigenes 高度富集; 而注释到节律过程 (rhythmic process)、运动 (locomotion)、排毒 (detoxification)、生物附着 (biological adhesion) 和行为 (behavior) 功能的 unigenes 数量则较少。

在细胞组分的 18 个亚类中, 分别有 13 334、13 334、9 427 和 8 905 个 unigenes 注释到细胞 (cell)、细胞成分 (cell part)、细胞器 (organelle) 和细胞膜 (cell membrane) 功能, 数量较多, 各占注释 unigenes 总数的 27.5%、27.5%、19.4% 和 18.4%。其中, 在与核 (GO: 0005634, 4 069 个) 和核糖体 (GO: 0005840, 639 个) 等细胞器相关的 GO 功能类别中 unigenes 高度富集; 而注释到类核 (nucleoid)、病毒体部分 (virion part)、病毒 (virion)、膜外区域部分 (extracellular region part)、胞外基质 (extracellular matrix) 和胞外基质组分 (extracellular matrix component) 功能的 unigenes 数量则较少。

表 1 花生异型雄蕊转录组 unigenes 的 GO 功能分类
Table 1 GO function classification of unigenes in transcriptome from heteromorphic stamens of *Arachis hypogaea* Linn.

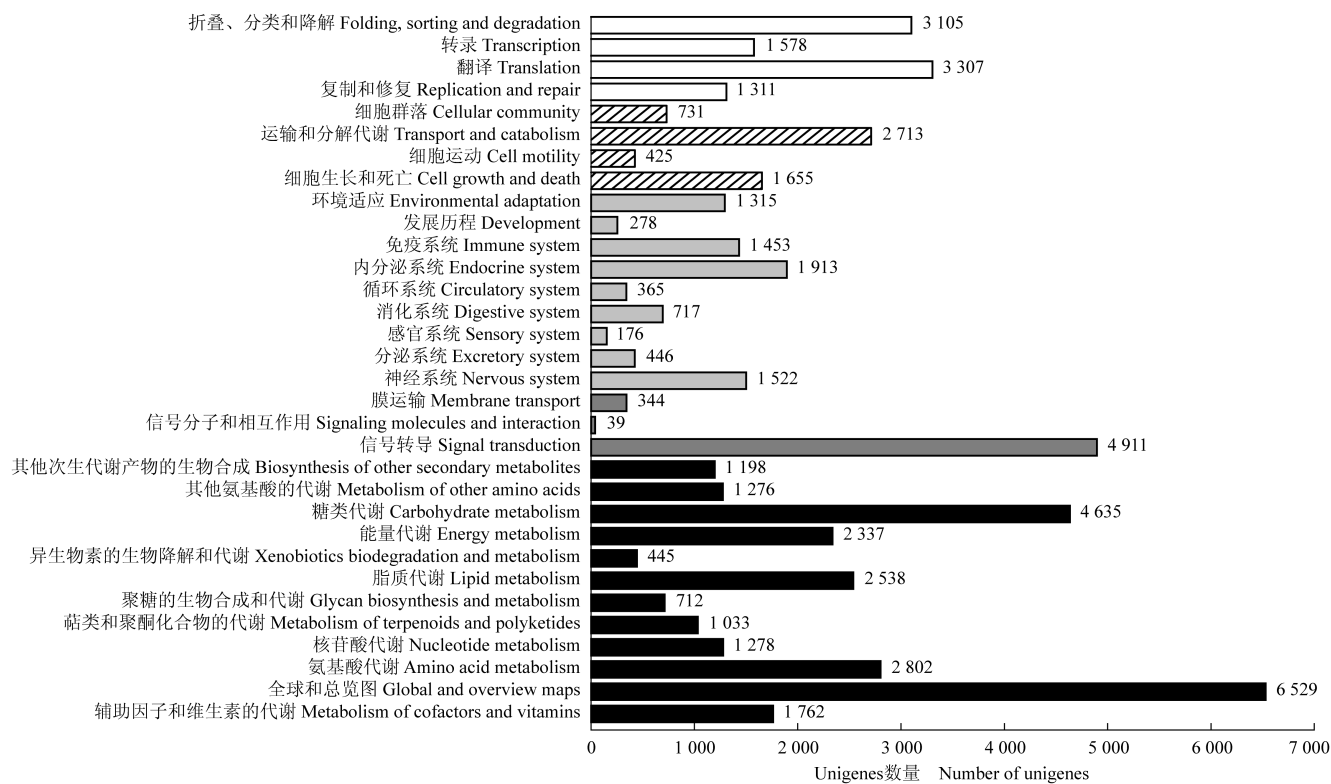
功能分类 Function classification	Unigenes 数量 Number of unigenes	功能分类 Function classification	Unigenes 数量 Number of unigenes
生物过程 Biological process		细胞膜成分 Membrane part	5 668
代谢过程 Metabolic process	31 817	细胞器成分 Organelle part	3 956
细胞过程 Cellular process	25 479	大分子复合物 Acromolecular complex	3 709
单一生物过程 Single-organism process	16 828	细胞连接 Cell junction	760
生物调节 Biological regulation	6 327	共质体 Symplast	757
应激反应 Response to stimulus	6 192	膜封闭的管腔 Membrane-enclosed lumen	595
生物过程调节 Regulation of biological process	5 866	细胞外区域 Extracellular region	498
定位 Localization	5 132	超分子纤维 Supramolecular fiber	452
细胞成分组织或生物合成 Cellular component organization or biogenesis	2 755	类核 Nucleoid	19
信号 Signaling	1 992	病毒体部分 Virion part	14
多细胞生物过程 Multicellular organismal process	1 647	病毒 Virion	14
发育过程 Developmental process	1 602	膜外区域部分 Extracellular region part	14
繁殖 Reproduction	1 096	胞外基质 Extracellular matrix	7
生殖过程 Reproductive process	1 093	胞外基质组分 Extracellular matrix component	1
多生物过程 Multi-organism process	1 011	分子功能 Molecular function	
生物过程的负调控 Negative regulation of biological process	737	结合 Binding	28 066
生物过程的正调控 Positive regulation of biological process	617	催化活性 Catalytic activity	26 091
生长 Growth	350	转运活性 Transporter activity	2 604
免疫系统过程 Immune system process	200	核酸结合转录因子活性 Nucleic acid binding transcription factor activity	1 165
节律过程 Rhythmic process	73	结构分子活性 Structural molecule activity	898
运动 Locomotion	24	分子功能调节 Molecular function regulator	626
排毒 Detoxification	22	信号转导活性 Signal transducer activity	486
生物附着 Biological adhesion	13	电子载体活性 Electron carrier activity	278
行为 Behavior	1	抗氧化活性 Antioxidant activity	259
细胞组分 Cellular component		分子转导活性 Molecular transducer activity	224
细胞 Cell	13 334	转录因子活性, 蛋白结合 Transcription factor activity, protein binding	107
细胞成分 Cell part	13 334	营养储存活性 Nutrient reservoir activity	50
细胞器 Organelle	9 427	转录调控因子活性 Translation regulator activity	6
细胞膜 Cell membrane	8 905	金属伴侣蛋白活性 Metallochaperone activity	4

在分子功能的 14 个亚类中, 分别有 28 066 和 26 091 个 unigenes 注释到结合 (binding) 和催化活性 (catalytic activity) 功能, 数量较多, 各占注释 unigenes 总数的 57.9% 和 53.8%。其中, 在 ATP 结合 (GO:0005524, 7 229 个)、转录因子 (GO:0003700, 1 161 个) 和钙离子结合 (GO:0005509, 728 个) 等与结合相关的 GO 功能类别中 unigenes 高度富集; 而注释到营养储存活性 (nutrient reservoir activity)、转录调控因子活性 (translation regulator activity) 和金属伴侣蛋白活性 (metallochaperone activity) 功能的 unigenes 数量则较少。

2.2.4 KEGG 代谢通路分析 以 KEGG 代谢库为参

考, 对花生异型雄蕊转录组 unigenes 的代谢通路进行统计和分类, 结果见图 4。结果显示: 在花生异型雄蕊转录组中共有 30 428 个 unigenes 注释到 2 708 个酶功能, 映射了 354 个代谢通路, 可分为遗传信息处理 (genetic information processing)、细胞过程 (cellular process)、生物系统 (organismal system)、环境信息处理 (environmental information processing) 和代谢 (metabolism) 5 大类 32 个亚类, 其中与遗传信息处理、细胞过程和环境信息处理相关的亚类较少, 分别仅有 4、4 和 3 个。

在遗传信息处理代谢通路中, 有 3 307 个 unigenes 注释到翻译 (transcription) 亚类, 有 3 015 个



□: 遗传信息处理 Genetic information processing; ▨: 细胞过程 Cellular process; ▤: 生物系统 Organismal system; ▥: 环境信息处理 Environmental information processing; ■: 代谢 Metabolism.

图 4 花生异型雄蕊转录组 unigenes 的 KEGG 代谢通路

Fig. 4 KEGG metabolic pathway of unigenes in transcriptome from heteromorphic stamens of *Arachis hypogaea* Linn.

unigenes 注释到折叠、分类和降解 (folding, sorting and degradation) 亚类。在细胞过程代谢通路中, 有 2 713 个 unigenes 注释到运输和分解代谢 (transport and catabolism) 亚类。在环境信息处理代谢通路中, 有 4 911 个 unigenes 注释到信号转导 (signal transduction) 亚类, 其中, 有 1 055 个 unigenes 注释到植物激素信号转导通路 (ko04075), 有 285 个 unigenes 注释到钙信号通路 (ko01100)。

共有 5 475 个 unigenes 注释到与生物系统相关的代谢通路中, 占注释 unigenes 总数的 18.0%, 分别注释到 9 个亚类。其中, 有 1 913 个 unigenes 注释到内分泌系统 (endocrine system) 亚类, 有 1 453 个 unigenes 注释到免疫系统 (immune system) 亚类。

注释到与代谢相关通路的 unigenes 数量最多, 有 14 231 个, 占注释 unigenes 总数的 46.8%, 分别注释到 12 个亚类。其中, 有 6 529 个 unigenes 注释到全球和总览图 (global and overview maps) 亚类, 有 4 635 个 unigenes 注释到糖类代谢 (carbohydrate metabolism)

亚类, 有 2 802 个 unigenes 注释到氨基酸代谢 (amino acid metabolism) 亚类, 有 2 538 个 unigenes 注释到脂质代谢 (lipid metabolism) 亚类, 有 2 337 个 unigenes 注释到能量代谢 (energy metabolism) 亚类, 有 1 278 个 unigenes 注释到核苷酸代谢 (nucleotide metabolism) 亚类。

3 讨论和结论

花生隶属于豆科 (Fabaceae) 落花生属 (*Arachis* Linn.), 从同源序列的匹配结果看, 其异型雄蕊转录组 unigenes 与同科的大豆和蒺藜苜蓿的同源匹配率较高, 与大豆的 32 466 个 unigenes 序列相似, 与蒺藜苜蓿的 17 343 个 unigenes 序列相似, 这一结果佐证了这些种类亲缘关系。而供试花生品种的异型雄蕊转录组中仅有 3 771 个 unigenes 与 Nr 数据库中花生的同源 unigenes 相匹配, 一方面可能是由于本实验是以花生异型雄蕊为实验材料, 而 Nr 数据库中的花生

基因组注释信息是以果针为研究材料有关;另一方面,花生栽培种是由2个二倍体野生种 *A. duranensis* 和 *A. ipaënsis* 自然杂交后经过染色体加倍形成的异源四倍体^[20-21],导致花生栽培品种的遗传关系复杂。

有花植物的雄蕊和花粉发育是一个复杂的过程,与花被器官相比,生殖器官的特有基因更多^[22]。对花生异型雄蕊转录组测序获得数据量约31.46 G,其中,有129 223个 unigenes 在Nr数据库比对出同源序列信息,且许多 unigenes 功能与雄蕊产生雄配子的功能有关。许多转录因子决定着花器官中萼片、花瓣、雄蕊和心皮^[23]的发育时序与数量,进而参与控制被子植物花发育。在花生异型雄蕊转录组中,有8 870个 unigenes 注释到转录功能,其中5 360个 unigenes 具有转录因子功能,说明转录因子对植物基因表达的调控有至关重要的作用^[24]。值得注意的是,有6.6%的 unigenes 基因功能仍不明确,这与花生的基因组大、重复序列比例高等原因有关。

雄蕊和花粉的发育涉及一系列精密的细胞质和核基因相互作用^[25];Ca²⁺可作为细胞信号转导的信使,通过钙调蛋白(CaM)将胞外信号转换成胞内的生理生化反应,包括细胞分裂、分化和凋亡,对花粉发育和萌发起着非常重要的作用^[26]。在GO数据库中对花生异型雄蕊转录组 unigenes 序列进行功能比对,结果显示其 unigenes 在与核酸代谢过程、磷酸化、转录调节、核、核糖体、ATP结合、转录因子和钙离子结合相关的功能类别中高度富集;而KEGG代谢通路分析结果表明:在花生异型雄蕊转录组中注释到代谢相关通路的 unigenes 数量最多,并以注释到糖类代谢通路、氨基酸代谢通路、脂质代谢通路、能量代谢通路和核苷酸代谢亚类的 unigenes 数量较多,说明有大量涉及孢粉素合成、氨基酸合成、核糖体组装、内质网蛋白质加工和脂质转移等功能的 unigenes 被注释;此外,与钙调蛋白相关的 unigenes 也被注释到无机离子运输与代谢以及信号转导机制这2个类别上。由此可知,在花生异型雄蕊转录组中,一部分基因可参与不同的代谢途径,同时花药发育的精密调控需要多个代谢途径参与,其中,糖类代谢不仅为花药发育提供能量,还可作为信号物质影响花药和花粉的发育^[27];而在花粉发育过程中通常会观察到高呼吸率和高能量需求^[28],主要代谢途径的功能障碍将对花粉粒发育产生不利影响^[29],雄性不育系中与能量供应相关的ATP含量显著降低^[30-31]。因而,对于花生异型雄蕊,

特别是退化雄蕊与可育雄蕊进行差异表达基因功能分析,可为花生雄蕊发育的分子调控提供更多启示。

花生异型雄蕊的发育调控是一个复杂的生物过程,应结合花生基因组测序成果开展大量的基因组分析,以期挖掘更多有价值的遗传信息,从而更全面地了解花生异型雄蕊分化的分子调控机制。

参考文献:

- [1] 王海霞,陶莹芳,丁雨龙.花生花器官形态建成及异型雄蕊发育研究[J].花生学报,2019,48(2):19-26.
- [2] MÜLLER H. Two kinds of stamens with different functions in the same flower[J]. Nature, 1882, 27: 30.
- [3] 罗中莱,张奠湘.异型雄蕊的研究进展[J].热带亚热带植物学报,2005,13(6):536-542.
- [4] VALLEJO-MARÍN M, DA SILVA E M, SARGENT R D, et al. Trait correlates and functional significance of heteranthery in flowering plants[J]. New Phytologist, 2010, 188: 418-425.
- [5] VAN DER NIET T, JOHNSON S D. Phylogenetic evidence for pollinator-driven diversification of angiosperms [J]. Trends in Ecology and Evolution, 2012, 27(6): 353-361.
- [6] 陈珊,任宗昕,王红,等.被子植物退化雄蕊的功能类型及其意义[J].植物分类与资源学报,2015,37(6):683-692.
- [7] WESTERKAMP C. Pollen in bee-flower relations some considerations on melittophily[J]. Botanica Acta, 1996, 109(4): 325-332.
- [8] CARDINAL S, BUCHMANN S L, RUSSELL A L. The evolution of floral sonication, a pollen foraging behavior used by bees (Anthophila)[J]. Evolution, 2018, 72(3): 590-600.
- [9] LUO Z, GU L, ZHANG D. Intrafloral differentiation of stamens in heterantherous flowers [J]. Journal of Systematics and Evolution, 2009, 47(1): 43-56.
- [10] PAULINO J V, DE FREITAS MANSANO V, TEIXEIRA S P. Elucidating the unusual floral features of *Swartzia dipetala* (Fabaceae)[J]. Botanical Journal of the Linnean Society, 2013, 173: 303-320.
- [11] NEPI M, GUARNIERI M, PACINI E. "Real" and feed pollen of *Lagerstroemia indica*: ecophysiological differences [J]. Plant Biology, 2003, 5: 311-314.
- [12] 栾文琪.几个国外特异花生种质资源[J].中国种业,1991(1):46.
- [13] HAAS B J, PAPANICOLAOU A, YASSOUR M, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis [J]. Nature Protocols, 2013, 8: 1494-1512.
- [14] GRABHERR M G, HAAS B J, YASSOUR M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. Nature Biotechnology, 2011, 29: 644-652.
- [15] ALTSCHUL S F, GISH W, MILLER W, et al. Basic local alignment search tool[J]. Journal of Molecular Biology, 1990, 215

- (3): 403–410.
- [16] MARCHLER-BAUER A, PANCHENKO A R, SHOEMAKER B A, et al. CDD: a database of conserved domain alignments with links to domain three-dimensional structure [J]. *Nucleic Acids Research*, 2002, 30: 281–283.
- [17] MARCHLER-BAUER A, DERBYSHIRE M K, GONZALES N R, et al. CDD: NCBI's conserved domain database [J]. *Nucleic Acids Research*, 2015, 43: 222–226.
- [18] ASHBURNER M, BALL C A, BLAKE J A, et al. Gene ontology: tool for the unification of biology [J]. *Nature Genetics*, 2000, 25: 25–29.
- [19] KANEHISA M, GOTO S. KEGG: Kyoto Encyclopedia of Genes and Genomes [J]. *Nucleic Acids Research*, 2000, 28: 27–30.
- [20] KRAPOVICKAS A, GREGORY W C. Taxonomy of the genus *Arachis* (Leguminosae) [J]. *Bonplandia*, 2007, 16 (Supl.): 7–205.
- [21] BERTIOLI D J, CANNON S B, FROENICKE L, et al. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut [J]. *Nature Genetics*, 2016, 48(4): 438–446.
- [22] ALVES-FERREIRA M, WELLMER F, BANHARA A, et al. Global expression profiling applied to the analysis of Arabidopsis stamen development [J]. *Plant Physiology*, 2007, 145: 747–762.
- [23] THEISSEN G, MELZER R. Molecular mechanisms underlying origin and diversification of the angiosperm flower [J]. *Annals of Botany*, 2007, 100: 603–619.
- [24] YAN X, DONG C, YU J, et al. Transcriptome profile analysis of young floral buds of fertile and sterile plants from the self-pollinated off spring of the hybrid between novel restorer line NR1 and *Nsa* CMS line in *Brassica napus* [J]. *BMC Genomics*, 2013, 14: 26.
- [25] SHEORAN I S, SAWHNEY V K. Proteome analysis of the normal and *Ogura* (*ogu*) CMS anthers of *Brassica napus*, to identify proteins associated with male sterility [J]. *Botanique*, 2010, 88 (3): 217–230.
- [26] ORRENIUS S, ZHIVOTOVSKY B, NICOTERA P. Regulation of cell death: the calcium-apoptosis link [J]. *Nature Reviews Molecular Cell Biology*, 2003, 4: 552–565.
- [27] WU Z, CHENG J, QIN C, et al. Differential proteomic analysis of anthers between cytoplasmic male sterile and maintainer lines in *Capsicum annum* L. [J]. *International Journal of Molecular Sciences*, 2013, 14(11): 22982–22996.
- [28] TADEGE M, KUHLEMEIER C. Aerobic fermentation during tobacco pollen development [J]. *Plant Molecular Biology*, 1997, 35: 343–354.
- [29] DORION S, LALONDE S, SAINI H S. Induction of male sterility in wheat by meiotic-stage water deficit is preceded by a decline in invertase activity and changes in carbohydrate metabolism in anthers [J]. *Plant Physiology*, 1996, 111: 137–145.
- [30] BERGMAN P, EDQVIST J, FARBOS I, et al. Male-sterile tobacco displays abnormal mitochondrial *atp1* transcript accumulation and reduced floral ATP/ADP ratio [J]. *Plant Molecular Biology*, 2000, 42: 531–544.
- [31] TEIXEIRA R T, KNORPP C, GLIMELIUS K. Modified sucrose, starch, and ATP levels in two alloplasmic male-sterile lines of *B. napus* [J]. *Journal of Experimental Botany*, 2005, 56 (414): 1245–1253.

(责任编辑: 郭严冬)

《植物遗传资源学报》2021 年征订启事

《植物遗传资源学报》是中国农业科学院作物科学研究所和中国农学会主办的学术期刊,为中国科技核心期刊、全国中文核心期刊和中国科学引文数据库(CSCD)核心期刊,被国内多家数据库收录,并被 CA 化学文摘(美国)(2014)和 JST 日本科学技术振兴机构数据库(日本)(2018)收录,荣获 2015 年度中国自然资源学会高影响力十佳期刊。据 2019 年版《中国科技期刊引证报告(核心版)》统计:在农艺学类 22 种期刊中,本刊的影响因子、总被引频次和综合评价总分排名均有所提升,核心影响因子为 1.344,比 2018 年提高 15.96%,排名稳居第 2 位。在 2018 年中国科学文献计量评价研究中心发布的《世界学术期刊学术影响力指数(WAJCI)年报》中,本刊在农艺学 102 种期刊中排名第 49 位,入选 Q2 区。

本刊报道内容为有关植物遗传资源基础理论研究、应用

研究方面的研究成果、创新性学术论文和高水平综述或评论。如种质资源的考察、收集、保存、评价、利用、创新、信息学、管理学等;起源、演化、分类等系统学;基因发掘、鉴定、克隆、基因文库建立、遗传多样性研究等。

本刊为双月刊,大 16 开本,每期 320 页,彩色铜版纸印刷。国内统一连续出版物号 CN 11–4996/S,国际标准连续出版物号 ISSN 1672–1810。全国各地邮局均可订阅,邮发代号 82–643,单价 68 元,全年 408 元。编辑部常年办理订阅手续,如需邮挂每期另加 3 元。地址:北京市中关村南大街 12 号《植物遗传资源学报》编辑部(邮编:100081)。电话:010–82105794;010–82105795。E-mail:zwyczyxb2003@163.com;zwyczyxb2003@sina.com。微信 ID:植物遗传资源学报;作者 QQ 群:372958204。网址:http://www.zwyczy.cn。