

DOI:10.13718/j.cnki.xsxb.2020.11.022

怀玉山高山马铃薯及其平原种植退化块茎的转录组分析^①

洪森荣¹, 刘洁¹, 曾欣怡¹, 蔡红², 陈荣华²

1. 上饶师范学院生命科学学院, 江西 上饶 334001; 2. 江西省上饶市红日农业开发有限公司, 江西 上饶 334700

摘要: 为探究怀玉山高山马铃薯平原生境种植出现种性退化导致形态和品质变异的机制, 本研究以怀玉山高山马铃薯(麻籽洋芋)正常种及其平原种植退化种的采后块茎为试验材料进行转录组分析。结果表明: 怀玉山高山马铃薯正常种及其平原种植退化种块茎差异表达基因数目为 3 604 个, 其中上调基因的数目为 1 331 个, 下调基因的数目 2 273 个。GO 富集分析显示, 差异基因主要注释到镉离子响应、根毛伸长、对缺水的反应、病原菌防御反应、盐胁迫反应、脱落酸反应、细胞组分、质膜、细胞内包膜细胞器、叶绿体、高尔基体、光系统 II、光系统 I、蛋白质结合、电子载体活性、血红素结合、叶绿素结合、谷胱甘肽结合等功能。KEGG 富集分析显示, 差异基因主要注释到苯丙氨酸代谢、植物-病原菌相互作用、苯丙素生物合成、谷胱甘肽代谢、半乳糖代谢、光合作用、氮代谢等生物功能。qRT-PCR 检测了 10 个差异基因在怀玉山高山马铃薯正常种及其平原种植退化种块茎的表达量, 其变化趋势均与 RNA-seq 一致。研究结果将为进行怀玉山高山马铃薯平原生境种植出现形态变大、品质下降等现象的基因表达分析提供参考。

关键词: 怀玉山高山马铃薯; 正常种; 平原种植退化种; 块茎; 转录组分析

中图分类号: S532

文献标志码: A

文章编号: 1000-5471(2020)11-0157-10

马铃薯(*Solanum tuberosum* L.)是茄科、茄属双子叶植物, 是中国乃至全球次于小麦、水稻和玉米的第 4 大重要粮食作物, 是支撑中国农村经济发展的重要组成部分^[1]。马铃薯高产、耐瘠薄、分布广泛、生育期短、适应性强, 富含碳水化合物、蛋白质、膳食纤维、矿物质及维生素等营养物质, 而且粮用、菜用、饲用兼优^[2]。2015 年起国家农业部开始大力推广马铃薯主粮化工程^[3]。怀玉山高山马铃薯又称麻籽洋芋, 2013 年 4 月核准为国家地理标志农产品, 主要种植于江西省上饶市怀玉山山区; 其块茎呈卵圆形或长圆形, 皮糙色黄肉白, 淀粉含量 10.3%~14.4%, 蛋白质含量 2.56%~2.60%, 干物质含量 16.8%~18.0%, 维生素 C 0.131~0.133 mg/g, 未检测出还原糖, 富含膳食纤维; 有和中养胃、健脾利湿、降糖降脂、美容养颜、防治胃癌等功效^[4]。农业生产实践表明, 怀玉山高山马铃薯不能在平原地区种植, 否则将会出现形态变大、品质下降, 且这些生产农艺性状移回高海拔地区也不可逆^[5], 究其原因, 尚无相关报道。

转录组学已逐渐成为基因功能研究的重要手段之一^[6]。近年来, 基于高通量方法的转录组测序正成为植物生理生态研究领域的有力工具^[7]。转录组测序是对某一物种的 mRNA 进行高通量测序, 而转录组测序的结果反映了特定条件和特定时间点的所有转录本序列信息^[8], 进而得到基因功能注释、蛋白质编码区序列、基因的表达量、代谢途径等大量信息, 为进一步研究提供基础数据和重要参考^[9]。该方法不仅能够用于有参考基因组序列的物种研究, 也能用于无参考基因组序列的物种研究, 已经被广泛应用于各个研究领域^[10-11]。

① 收稿日期: 2020-03-29

基金项目: 江西省 2019 年度重点研发计划项目(20192BBF60006); 江西省 2020 年度重点研发计划项目(20202BBF60001)。

作者简介: 洪森荣(1974—), 男, 教授, 主要从事植物生物技术研究。

本研究基于边合成边测序(Sequencing by synthesis, SBS)技术,利用 Illumina Hiseq 4 000 高通量测序平台对怀玉山高山马铃薯正常种及其平原种植退化种块茎进行转录组测序分析、基因表达量分析和差异表达分析、新基因预测与功能注释分析,以及差异表达基因功能注释分析,对怀玉山高山马铃薯平原种植出现形态变大、品质下降的机理进行研究,以期解析怀玉山高山马铃薯平原种植退化种的内在机理,并为其适应高海拔生境相关基因的挖掘与应用提供参考。

1 材料与方法

1.1 试验材料

怀玉山高山马铃薯正常种(麻籽洋芋,编号 CK)及其平原种植退化种(编号 BZ)采后未萌芽的块茎,由江西省上饶市红日农业开发有限公司提供。

1.2 试验方法

1.2.1 RNA 提取

参照说明书用 Trizol 试剂提取怀玉山高山马铃薯正常种(CK)及其平原种植退化种(BZ)采后未萌芽块茎总 RNA。所提 RNA 的质量由 Nanodrop 2000 和琼脂糖凝胶电泳检测,Agilent 2100 测定 RIN 值。若 RNA 总量 5 μg ,浓度 ≥ 200 ng/ μL ,OD260/280 介于 1.8 和 2.2 之间即为可用。

1.2.2 Illumina Hiseq 测序

RNA 测序送由上海美吉生物科技股份有限公司完成。利用带有 Oligo(dT)的磁珠,将 3'端带有 PolyA 尾结构的 mRNA 从总 RNA 中分离出来。然后利用 TruSeq™ RNA sample prep kit 试剂盒构建 cDNA 文库。利用 Illumina hiseq 4000 进行测序和分析(3 次技术重复)。使用 Seq Prep 工具(<https://github.com/jstjohn/seqprep>)和 Sickle 工具(<https://github.com/najoshi/sjckle>)将原始的 FASTQ 文件中包含一些带接头、低质量、N 率较高序列及长度过短序列的 reads 进行过滤、修剪或删除,得到高质量的测序数据(clean reads),以保证后续信息分析的质量。

1.2.3 数据分析

使用 Trinity 软件(<https://github.com/trinityrnaseq/trinityrnaseq/wiki>)对 clean reads 进行 *de novo* 拼接,形成一个不能在两端扩展的 Unigene 序列。利用 Trinity 软件组装得到转录本,并与参考蛋白质数据库 NR(Non-redundant protein, <http://www.ncbi.nlm.nih.gov>)进行 BLAST 比对,保留最优比对结果,挑选一条最长的代表序列作为 Unigene。使用软件 Trans Decode 对组装结果进行 CDS 预测。再将序列注释到 Swiss-Protein(<http://www.expasy.ch/sprot>),KEGG(Kyoto Encyclopedia of Genes and Genomes, <http://www.genome.jp/kegg>)和 GO(Gene Ontology Consortium, <http://geneontology.org>)数据库中获得基因功能信息。对所有单基因进行 GO 功能分类,在宏观层面上了解基因功能的分布。基因的表达水平由转录本的丰度表示,利用 RPKM(每百万个 reads 映射到外显子区的每千个碱基上的 reads 数)对 Unigene 的 read count 值进行标准化。P 值来识别差异表达的基因,在多次检验和分析中,用错误发现率来识别 P 值的阈值。根据错误发现率(false discovery rate, $\text{FDR} \leq 0.001$)和表达量倍数变化($|\text{foldchange}| > 1.5$)的阈值筛选出 2 个样本间差异表达的基因(Degs)。对显著性差异表达基因进行 KEGG Pathway 显著性富集分析。

1.2.4 相关差异表达基因的 qRT-PCR

根据转录组试验结果,采用 Primer BLAST 在线工具对相关差异表达基因(GSC0003DMG400011811, PGSC0003DMG402005881, PGSC0003DMG400021727, PGSC0003DMG400000926, PGSC0003DMG400022022, PGSC0003DMG400013684, PGSC0003DMG400027631, PGSC0003DMG400004595, PGSC0003DMG400006448, PGSC0003DMG400003563)进行引物设计(表 1),每个样品以 GAPDH(GenBank 登录号为 NM_001288415.1)为内参,3 个技术重复,采用 1.2.1 的方法提取 2 种材料的总 RNA,按照 PrimeScript™ RT reagent Kit with gDNA Eraser(Perfect Real Time)试剂盒说明书合成 cDNA 的第 1 条链,在荧光定量 PCR 仪中进行扩增反应,每个反应重复 3 次。数据采用 SPSS 19.0 版本进行统计分析。

表 1 qRT-PCR 引物序列

| 引物名称 | 引物序列 | 位置 | TM 值 | 产物长度 |
|------------------------|------------------------|-------|------|------|
| GAPDH_1F | GCTCCCATCACTTAACGG | 732 | 62.6 | 145 |
| GAPDH_1R | ACCCTCTGATTCCTCCTTG | 876 | 63.2 | |
| GAPDH_2F | GGCTTGATTGATGCTGCT | 1 071 | 63.2 | 127 |
| GAPDH_2R | GTTCCGTTCTCTCTGGTT | 1 197 | 64.0 | |
| GSC0003DMG400011811_F | AAGGAACCATACACAGCGA | 465 | 63.7 | 141 |
| GSC0003DMG400011811_R | AGGAGGAACGACACCATAAC | 605 | 64.1 | |
| PGSC0003DMG402005881_F | ATCAGCATCAGCAGTTTACAAG | 683 | 63.7 | 130 |
| PGSC0003DMG402005881_R | GCAAGAATAAGGGAGTTCGAC | 812 | 63.5 | |
| PGSC0003DMG400021727_F | TGAGGCAGTCCAGAGAGTTA | 456 | 64.7 | 154 |
| PGSC0003DMG400021727_R | TGGCTTAGCAGAGATGACAG | 609 | 64.1 | |
| PGSC0003DMG400000926_F | CGATTGGCTCTTACTCTCCT | 156 | 63.6 | 113 |
| PGSC0003DMG400000926_R | GTGTTCTCCTTTGGCTTACC | 268 | 63.4 | |
| PGSC0003DMG400022022_F | TCAGTGGATGGAGCAATGT | 51 | 63.8 | 106 |
| PGSC0003DMG400022022_R | GGAGCAAAAAGCAGAAGTAGG | 156 | 63.4 | |
| PGSC0003DMG400013684_F | CCTTTGCTTAGACCATTCTTG | 753 | 61.7 | 115 |
| PGSC0003DMG400013684_R | CAGCCATTAGTTTCCTTCTTTT | 867 | 61.6 | |
| PGSC0003DMG400027631_F | CCAGAAAACATACGGTAAGACA | 464 | 62.8 | 126 |
| PGSC0003DMG400027631_R | AGCCCTGCAATAACAATAACA | 589 | 62.8 | |
| PGSC0003DMG400004595_F | CACAATCAGAGCCAATATAGCA | 58 | 63.0 | 140 |
| PGSC0003DMG400004595_R | CCCTTTCACCAACTACTGTTC | 197 | 63.9 | |
| PGSC0003DMG400006448_F | CCAATTTCTTTCAATGCTTCT | 226 | 59.8 | 124 |
| PGSC0003DMG400006448_R | TGCTATTACTTTCCCATCTTCA | 349 | 61.7 | |
| PGSC0003DMG400003563_F | ATAGCCACTTTTCAGAATCCA | 951 | 61.9 | 128 |
| PGSC0003DMG400003563_R | TCCTTACTCATTTTCCTCCTGT | 1 078 | 63.4 | |

2 结果与分析

2.1 测序数据产出统计

经过测序质量控制, 共得到 14.83 Gb Clean Data. 其中 BZ 组 Clean Data 中 pair-end Reads 总数为 25 461 192, 总碱基数为 7 689 279 984, G 和 C 2 种碱基占总碱基的百分比为 44.46%, 质量值大于或等于 30 的碱基所占的百分比为 93.76%; CK 组 Clean Data 中 pair-end Reads 总数为 23 643 517, 总碱基数为 7 140 342 134, G 和 C 2 种碱基占总碱基的百分比为 47.31%, 质量值大于或等于 30 的碱基所占的百分比为 93.53%. 2 组样品 Q30 碱基百分比均不小于 93.53%. 结果表明, 此次转录组测序数据质量较好, 满足转录组分析的基本要求.

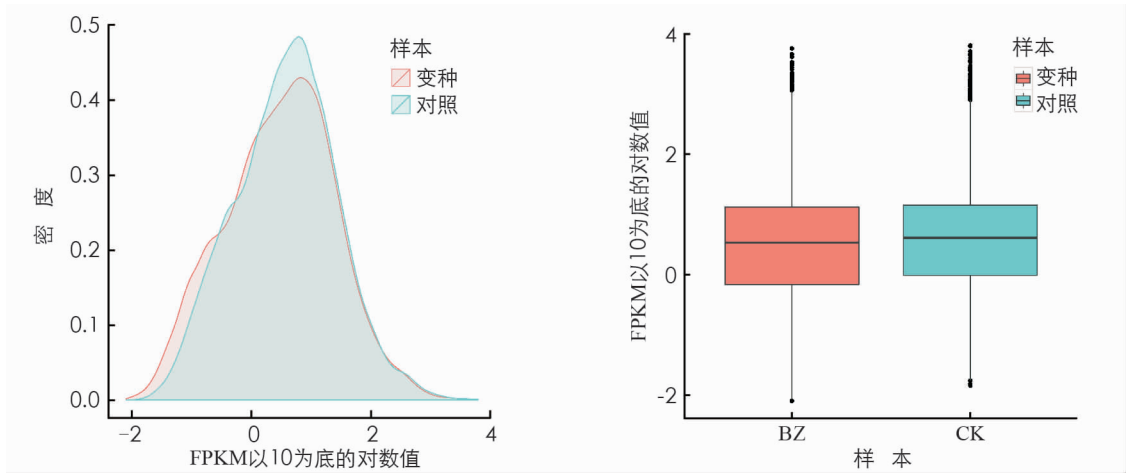
2.2 转录组数据与参考基因组序列的比对效率

比对效率指 Mapped Reads 占 Clean Reads 的百分比, 是转录组数据利用率的最直接体现. 经统计, BZ 组 Total Reads(即 Clean Reads 数目, 按单端计)为 50 922 384, 比对到参考基因组上的 Reads 数(Mapped Reads)为 37 005 516, 在 Clean Reads 中占的百分比为 72.67%; 比对到参考基因组唯一位置的 Reads 数(Uniq Mapped Reads)为 35 067 236, 在 Clean Reads 中占的百分比为 68.86%; 比对到参考基因组多处位置的 Reads 数(Multiple Mapped Reads)为 1 938 280, 在 Clean Reads 中占的百分比为 3.81%; 比对到参考基因组正链的 Reads 数为 17 764 983, 在 Clean Reads 中占的百分比为 34.89%; 比对到参考基因组负链的 Reads 数为 18 197 367, 在 Clean Reads 中占的百分比为 35.74%. CK 组 Total Reads(即 Clean Reads 数目, 按单端计)为 47 287 034, 比对到参考基因组上的 Reads 数(Mapped Reads)为 17 179 940, 在 Clean Reads 中占的百分比为 36.33%; 比对到参考基因组唯一位置的 Reads 数(Uniq Mapped Reads)为 16 150 402, 在 Clean Reads 中占的百分比为 34.15%; 比对到参考基因组多处位置的 Reads 数(Multiple Mapped Reads)为 1 029 538, 在 Clean Reads 中占的百分比为 2.18%; 比对到参考基因组正链的 Reads 数为 8 093 153, 在 Clean Reads 中占的百分比为 17.11%; 比对到参考基因组负链的 Reads 数为 8 399 367,

在 Clean Reads 中占的百分比为 17.76%。比对效率除了受测序质量影响外，还与指定的参考基因组组装的优劣、参考基因组与测序样品的生物学分类关系远近(亚种)有关。通过本次试验的比对效率可知，所选参考基因组组装能满足后续信息分析的需求。

2.3 基因表达量分析

从图 1 可知，CK 组和 BZ 组的密度分布范围在 0~0.5。CK 组和 BZ 组表达量 FPKM 的对数值基本为 0~2。

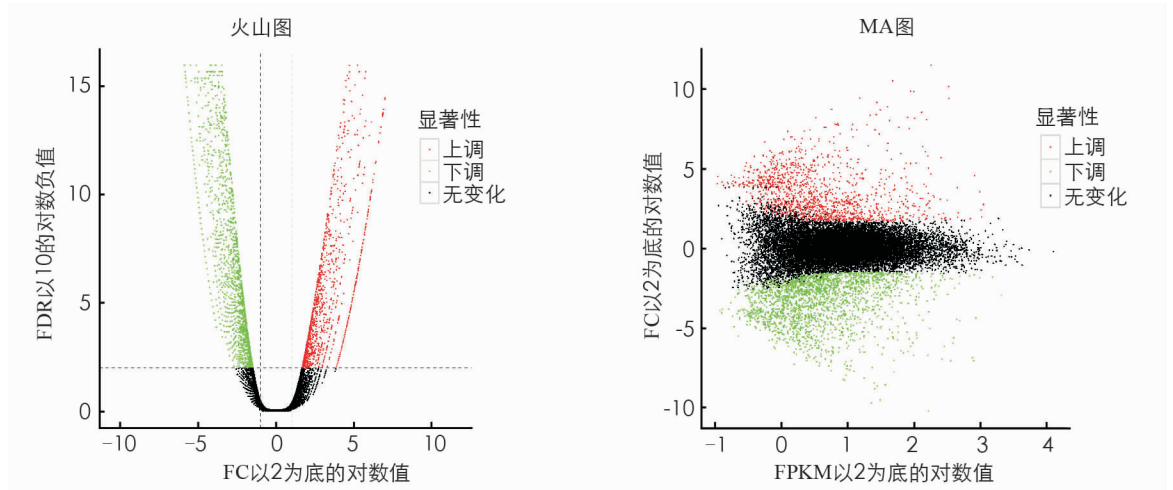


左图：红色表示 BZ 组，蓝色表示 CK 组，横坐标表示对应样品 FPKM 的对数值，纵坐标表示概率密度；
右图：横坐标代表 CK 组和 BZ 组；纵坐标表示 CK 组和 BZ 组表达量 FPKM 的对数值。

图 1 CK 组和 BZ 组 FPKM 密度分布对比图(左)和 FPKM 箱线图(右)

2.4 基因差异表达分析

基因差异表达火山图(图 2)中的每一个点表示一个基因，横坐标表示某一个基因在 2 个样品中表达量差异倍数的对数值；纵坐标表示基因表达量变化的统计学显著性的负对数值。横坐标绝对值越大，说明表达量在 2 个样品间的表达量倍数差异越大；纵坐标值越大，表明差异表达越显著，筛选得到的差异表达基因越可靠。差异表达基因 MA 图(图 2)中每一个点代表一个基因。横坐标为 A 值： $\log_2(\text{FPKM})$ ，即 2 个样品中表达量均值的对数值；纵坐标为 M 值： $\log_2(\text{FC})$ ，即 2 个样品间基因表达量差异倍数的对数值，用于衡量表达量差异的大小。

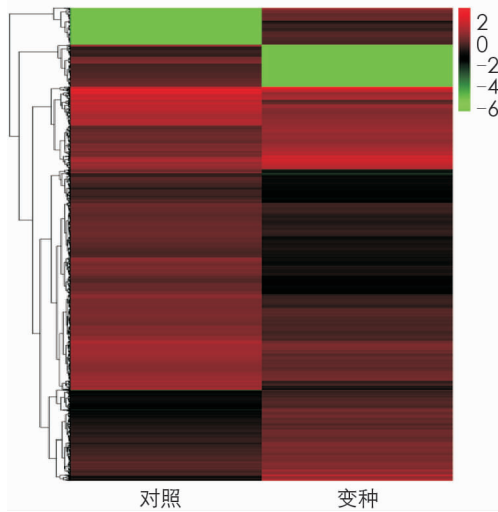


绿色的点代表下调差异表达基因，红色的点代表上调差异表达基因，黑色的点代表非差异表达基因。

图 2 差异表达火山图(左)和差异表达 MA 图(右)

对筛选出的差异表达基因做层次聚类分析，将具有相同或相似表达模式的基因进行聚类，差异表达基因聚类结果如图 3。经分析，BZ 组和 CK 组差异表达基因数目为 3 604 个；其中，上调基因的数目为 1 331 个，下调基因的数目 2 273 个。BZ 组和 CK 组差异最显著的 10 个基因分别为 PGSC0003DMG400000028，

PGSC0003DMG400000046, PGSC0003DMG400000048, PGSC0003DMG400000049, PGSC0003DMG400000064, PGSC0003DMG400000066, PGSC0003DMG400000069, PGSC0003DMG400000070, PGSC0003DMG400000075, PGSC0003DMG400000079. 其中相比 CK 组, BZ 组双功能 L-3-氰基丙氨酸合酶/半胱氨酸合酶 2、类 CCR4-非转录复合物亚单位 1 等 2 个基因上调, LOC101261174、类甘油-3-磷酸 2-O-酰基转移酶 4、类 WRKY 转录因子 22、乙烯反应性晚期胚胎发生丰富蛋白、植物细胞壁蛋白 SITFR88、非特征蛋白 LOC101265683、非特征蛋白 LOC101245497、类 BAG 家族分子伴侣调节因子等 8 个基因下调。



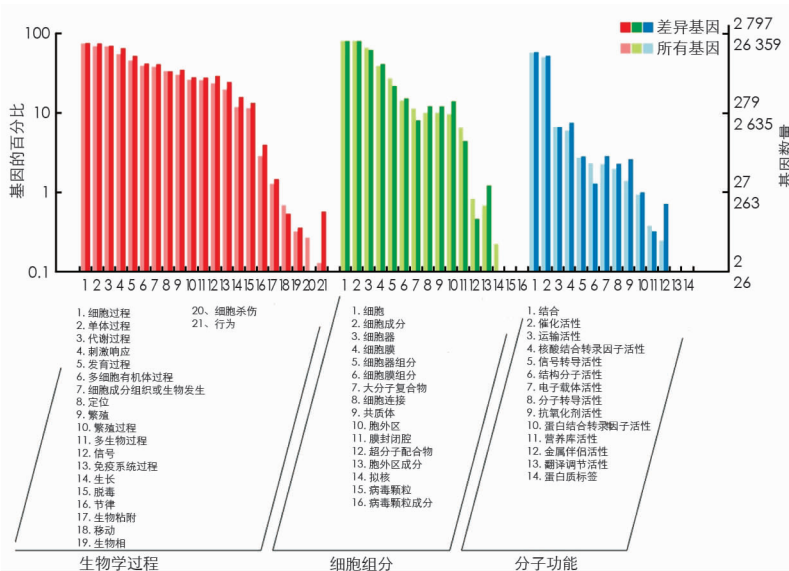
横坐标代表样品名称及样品的聚类结果, 纵坐标代表的差异基因及基因的聚类结果。

图中不同的列代表不同的样品, 不同的行代表不同的基因。颜色代表了基因在样品中的表达量水平 $\lg(\text{FPKM}+0.000001)$ 。

图 3 差异表达基因聚类图

2.5 差异表达基因 GO 分类

图 4 展示的是在差异表达基因背景和全部基因背景下 GO 各二级功能的基因富集情况, 体现 2 个背景下各二级功能的地位, 具有明显比例差异的二级功能说明差异表达基因与全部基因的富集趋势不同。通过富集, 镉离子响应、根毛伸长、对缺水的反应、真菌防御反应、盐胁迫反应、脱落酸反应、单元部分、质膜、细胞内膜结合细胞器、叶绿体、高尔基体、光系统 II、光系统 I、蛋白质结合、电子载体活性、血红素结合、叶绿素结合、谷胱甘肽结合等 GO term 富集显著(表 2)。这些功能类别可能在怀玉山高山马铃薯的平原种植出现形态变大、品质下降中起着重要作用。



横坐标为 GO 分类, 纵坐标左边为基因数目所占百分比, 右边为基因数目。

图 4 差异表达基因 GO 注释分类统计

表 2 部分差异表达基因 GO 富集结果

| 功能注释 | GO term 的 ID | GO 功能 | 所有基因注释到该功能的基因数/个 | DEG 注释到该功能的基因数/个 | 注释到该功能 DEG 数目的期望值 | 富集 Term 的显著性统计 |
|-------|--------------|----------|------------------|------------------|-------------------|----------------|
| 生物学过程 | GO: 0046686 | 镉离子响应 | 1 570 | 121 | 115.09 | 5.40E-14 |
| 生物学过程 | GO: 0048767 | 根毛伸长 | 733 | 67 | 53.73 | 3.40E-13 |
| 生物学过程 | GO: 0009414 | 缺水响应 | 1 992 | 214 | 146.03 | 1.40E-10 |
| 生物学过程 | GO: 0050832 | 真菌防御反应 | 1 502 | 190 | 110.11 | 3.10E-10 |
| 生物学过程 | GO: 0009651 | 盐胁迫响应 | 2 851 | 254 | 209.00 | 5.50E-10 |
| 生物学过程 | GO: 0009737 | 脱落酸响应 | 2 412 | 236 | 176.81 | 5.80E-10 |
| 细胞组分 | GO: 0044464 | 细胞组分 | 21 443 | 1 503 | 1 544.24 | 1.70E-13 |
| 细胞组分 | GO: 0005886 | 质膜 | 5 758 | 441 | 414.67 | 1.30E-12 |
| 细胞组分 | GO: 0043231 | 胞内膜结合细胞器 | 17 478 | 1 155 | 1 258.69 | 2.40E-09 |
| 细胞组分 | GO: 0009507 | 叶绿体 | 5 871 | 370 | 422.80 | 3.40E-09 |
| 细胞组分 | GO: 0005794 | 高尔基体 | 2 069 | 106 | 149.00 | 3.10E-08 |
| 细胞组分 | GO: 0009523 | 光系统 II | 90 | 16 | 6.48 | 1.00E-07 |
| 细胞组分 | GO: 0009522 | 光系统 I | 55 | 11 | 3.96 | 9.80E-07 |
| 分子功能 | GO: 0005515 | 蛋白质结合 | 6 739 | 527 | 473.39 | 1.40E-25 |
| 分子功能 | GO: 0009055 | 电子载体活性 | 595 | 54 | 41.80 | 1.60E-15 |
| 分子功能 | GO: 0020037 | 血红素结合 | 555 | 65 | 38.99 | 1.60E-07 |
| 分子功能 | GO: 0016168 | 叶绿素结合 | 57 | 10 | 4.00 | 7.20E-07 |
| 分子功能 | GO: 0043295 | 谷胱甘肽结合 | 58 | 20 | 4.07 | 7.30E-07 |

2.6 差异表达基因 COG 分类

差异表达基因 COG 分类统计结果如图 5 所示, BZ 组和 CK 组差异表达基因 General function prediction only(一般功能预测), Transcription(转录), Signal transduction mechanisms(信号转导机制), Replication, recombination and repair(复制、重组和修复)和 Secondary metabolites biosynthesis, transport and catabolism(次级代谢产物的生物合成、运输和分解代谢)等 5 大功能的基因数较多。

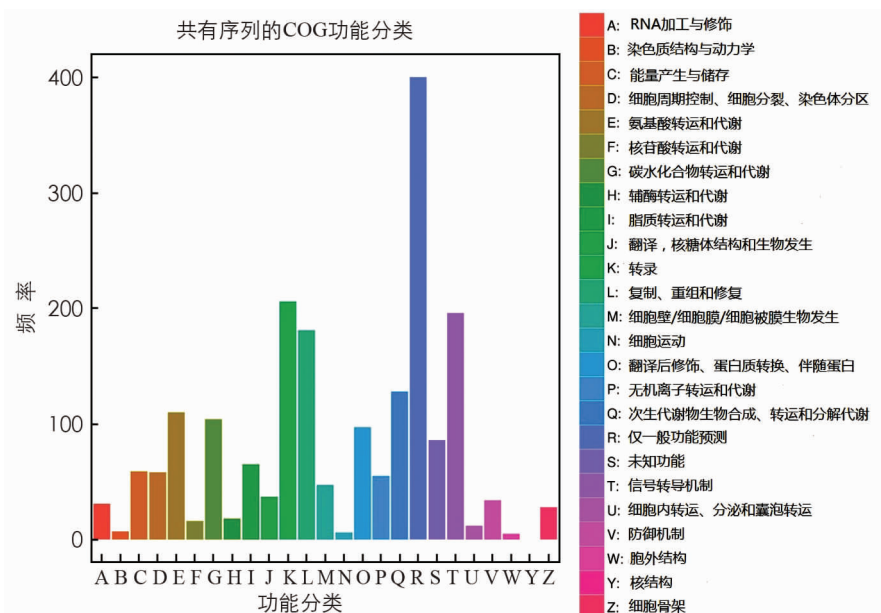
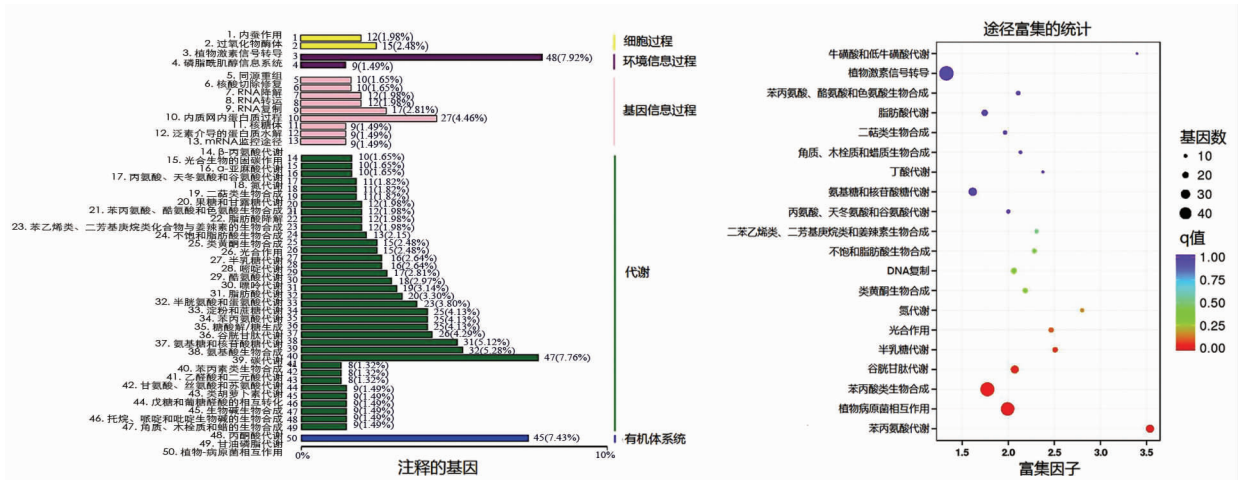


图 5 差异表达基因 COG 注释分类统计

2.7 差异表达基因 KEGG 注释

对 BZ 组和 CK 组差异表达基因 KEGG 的注释结果按照 KEGG 中通路类型进行分类, 分类图如图 6 (左) 所示. 结果表明, BZ 组和 CK 组差异表达基因注释到植物激素信号转导(Plant hormone signal transduction)、苯丙酸生物合成(Phenylpropanoid biosynthesis)、植物病原相互作用(Plant-pathogen interaction)、碳代谢(Carbon metabolism)、氨基酸生物合成(Biosynthesis of amino acids)的基因个数及其个数占被注释上的基因总数的比例分别为 48(7.92%), 47(7.76%), 45(7.43%), 32(5.38%) 和 31(5.12%).

差异表达基因 KEGG 通路富集分析结果见图 6(右), 呈现了显著性 Q 值最小的前 20 个通路. 其中, 差异表达基因富集 $q \text{ value} \leq 0.25$ 的 KEGG 通路主要有苯丙氨酸代谢(Phenylalanine metabolism)、植物病原相互作用(Plant-pathogen interaction)、苯丙素生物合成(Phenylpropanoid biosynthesis)、谷胱甘肽代谢(Glutathione metabolism)、半乳糖代谢(Galactose metabolism)、光合作用(Photosynthesis)、氮代谢(Nitrogen metabolism). 这些途径可能对怀玉山高山马铃薯的平原种植出现形态变大、品质下降起着重要调节作用.



纵坐标为 KEGG 代谢通路的名称, 横坐标为注释到该通路下的基因个数及其个数占被注释上的基因总数的比例.

图 6 差异表达基因 KEGG 分类图(左)和富集散点图(右)

2.8 相关差异表达基因的 qRT-PCR 分析验证

从 RNA-seq 数据中挑选 10 个在 KEGG 通路注释数量较多的差异表达基因 PGSC0003DMG400011811 (ferredoxin-NADP, reductase 铁氧还蛋白-NADP 还原酶), PGSC0003DMG402005881 (ferredoxin-3, 铁氧还蛋白-3), PGSC0003DMG400021727 (photosystem II oxygen-evolving complex protein 3, 光系统 II 析氧复合蛋白 3), PGSC0003DMG400000926 (xygen-evolving enhancer protein 2, 产氧增强蛋白 2), PGSC0003DMG400022022 (photosystem I reaction center subunit IV B, 光系统 I 反应中心亚基 IV B), PGSC0003DMG400013684 (trans-cinnamate 4-monooxygenase-like, 反式肉桂酸 4-单加氧酶样), PGSC0003DMG400027631 (gibberellin 2-oxidase, 赤霉素 2-氧化酶), PGSC0003DMG400004595 (ABC transporter B family member 11-like, 类 ABC 转椅子 B 家族成员 11), PGSC0003DMG400006448 (anthocyanin O-methyl transferase, 花青素 O-甲基转移酶), PGSC0003DMG400003563 (flavanone 3 beta-hydroxylase, 黄烷酮 3β 羟化酶), 并利用 qRT-PCR 技术对其表达量进行验证(表 3), 结果显示, 10 个基因表达量的变化趋势均与 RNA-seq 结果一致, 说明本次转录组测序数据可信度高. qRT-PCR 也进一步验证了与怀玉山高山马铃薯平原平原种植退化种有关的 PGSC0003DMG400000926, PGSC0003DMG400003563, PGSC0003DMG400006448, PGSC0003DMG400022022, PGSC0003DMG400021727 基因在怀玉山高山马铃薯平原种植退化种中下调表达, 而 PGSC0003DMG400004595, PGSC0003DMG400011811, PGSC0003DMG400027631, PGSC0003DMG402005881, PGSC0003DMG400013684 基因在怀玉山高山马铃薯正常种中上调表达.

表 3 10 个差异表达基因 qRT-PCR 的 RQ 值

| 基因名称 | 正常种 RQ 值 | 平原种植退化种 RQ 值 | 调节 | RNA-seq 调节 |
|----------------------|----------|--------------|----|------------|
| GAPDH | 1.00a | 1.00a | — | — |
| PGSC0003DMG400000926 | 1.00a | 0.03b | 下调 | 下调 |
| PGSC0003DMG400003563 | 1.00a | 0.54b | 下调 | 下调 |
| PGSC0003DMG400004595 | 1.00b | 1.59a | 上调 | 上调 |
| PGSC0003DMG400006448 | 1.00a | 0.09b | 下调 | 下调 |
| PGSC0003DMG400011811 | 1.00b | 1.85a | 上调 | 上调 |
| PGSC0003DMG400022022 | 1.00a | 0.65b | 下调 | 下调 |
| PGSC0003DMG400027631 | 1.00b | 1.73a | 上调 | 上调 |
| PGSC0003DMG402005881 | 1.00b | 1.53a | 上调 | 上调 |
| PGSC0003DMG400021727 | 1.00a | 0.12b | 下调 | 下调 |
| PGSC0003DMG400013684 | 1.00b | 1.62a | 上调 | 上调 |

注: 同列不同小写字母间表示差异有统计学意义($p \leq 0.05$).

3 讨 论

怀玉山高山马铃薯种植于平原地区即低海拔区域容易产生平原种植退化种, 其主要原因是怀玉山高山马铃薯的高海拔生境发生了改变. 怀玉山高山马铃薯在高山地区的长期栽培已经产生一整套适于高海拔生境的生理机制, 将其移种到低海拔区域, 必然打破其原有的生理适应机制, 产生一整套适于低海拔生境的生理机制. 高山地区对植物生理机制产生影响的因素主要有低温、紫外线、水分等, 这些因素的胁迫对高山地区植物的生理机制将会产生一个综合性影响. 张飞等^[10]对水分胁迫下高粱差异基因 KEGG 分析结果显示, 苯丙烷类生物合成、苯丙氨酸代谢、类黄酮生物合成 3 个途径中差异基因表达较多. 白子或等^[12]的研究也指出水分胁迫与苯丙氨酸代谢密切相关. 此外, 有学者指出水分胁迫将造成光合系统发生变化, 影响光合物质生产^[13]. 王晓娇等^[14]发现“克新 1 号”马铃薯在萌芽出苗期遇到水分胁迫时, 显著富集的有谷胱甘肽代谢、植物激素信号转导、苯丙素生物合成、苯丙氨酸代谢和植物-病原菌相互作用等途径. 本研究与上述研究结果基本一致. 在本研究中, 怀玉山高山马铃薯平原种植退化种差异表达基因富集的 KEGG 通路主要有苯丙氨酸代谢、植物-病原菌相互作用、苯丙素生物合成、谷胱甘肽代谢、半乳糖代谢、光合作用、氮代谢. 这些途径对怀玉山高山马铃薯的平原种植退化起着重要调节作用. 苯丙氨酸代谢途径是植物重要的次生代谢途径, 主要包括苯丙烷代谢和异黄酮合成代谢^[15-16]. 植物的抗病与病原菌的致病是生物界一种十分复杂的相互作用系统. 当植物组织受到病原真菌侵染后, 寄主体内的防御体系会发生一系列的改变, 具有明显的时空性, 且受到环境条件、代谢产物与基因表达的调控^[17]. 怀玉山高山马铃薯移种到低海拔区域, 受低温、紫外线、水分等胁迫减少, 病原菌侵染加强, 植物与病原菌相互作用可能也是怀玉山高山马铃薯平原种植退化的主要原因. 苯丙素是具有 C6-C3 芳香核骨架的一类天然化合物的统称, 包括简单苯丙素类、木质素和木脂素类、香豆素类及黄酮类等许多天然芳香族化合物^[18-19]. 其中香豆素类具有多种生理活性^[20], 如抗菌、抗病毒等; 木质素作为植物细胞壁的主要组成部分, 与纤维素、半纤维素、果胶等物质交叉联接, 其合成能够对植物细胞起重要的保护作用, 增强植物细胞抵抗病害和逆境的能力^[21]. 由于苯丙素生物合成和苯丙氨酸代谢均属于次生代谢产物生物合成, 这些次生代谢物质对植物抵抗逆境胁迫具有重要作用, 说明怀玉山高山马铃薯移种到低海拔区域对马铃薯的生长发育造成一定程度的伤害. 谷胱甘肽是一种含有巯基的三肽化合物, 在植物体内是一种尤为重要的抗氧化物质^[22], 在减缓细胞氧化损伤上具有重要的作用, 可降低怀玉山高山马铃薯移种到低海拔区域对马铃薯造成的毒害作用^[23]. 维生素 C 含量的多少可以作为衡量园艺产品品质的重要指标之一, 在植物中维生素 C 的合成中, L-半乳糖内酯脱氢酶作为关键酶可直接氧化半乳糖内酯形成 L-抗坏血酸(VC)^[24]. 抗坏血酸是植物生长发育所必需的微量营养素, 在生命过程中发挥重要生理功能, 特别是在抵抗逆境胁迫反应中具有重要作用^[25]. 研究表明, 低温胁迫下秋茄叶片上述抗氧化酶活性和抗氧化剂含量表现为抗坏血酸/氧化型抗坏血酸、还原型谷胱甘肽/氧化型谷胱甘肽

比值增加,从而抑制活性氧(ROS)大量产生,减少膜系统伤害^[26]。怀玉山高山马铃薯移种到低海拔区域,降低了抗坏血酸-还原型谷胱甘肽循环中抗氧化酶活性,降低清除 ROS 能力,增强了叶片 PSII 光抑制,从而降低了植株光合作用,这可能也是怀玉山高山马铃薯移种到低海拔区域出现形态变大、品质下降的原因之一。氮在植物生长中起着重要的作用,氮素作为植物蛋白质、核酸、酶及叶绿素的组成成分,又参与多种内源激素及其前体的合成^[27]。研究表明,在逆境条件下,氮的有效性对于植物的光合能力和植株生长至关重要,植株中硝态氮含量呈降低的趋势,但铵态氮含量呈增加的趋势,氮代谢相关酶活性均有不同程度的降低,植物的硝态氮更多的分配到了根部,以增强植物的抗逆性^[28]。怀玉山高山马铃薯移种到低海拔区域出现形态变大、品质下降等现象,可能是降低了氮代谢相关酶活性,促使碳流由光合代谢转向氮代谢,破坏了怀玉山高山马铃薯碳氮代谢的正常进行,减弱了怀玉山高山马铃薯抵御逆境胁迫的能力。

本试验研究了怀玉山高山马铃薯正常种及其平原种植退化种块茎涉及苯丙氨酸代谢、植物-病原菌相互作用、苯丙素生物合成、谷胱甘肽代谢、半乳糖代谢、光合作用、氮代谢等相关基因的差异表达,仅是关于怀玉山高山马铃薯正常种及其平原种植退化种块茎转录组的变化研究;下一步将对怀玉山高山马铃薯正常种及其平原种植退化种块茎试管苗转录组变化进行系统研究,以获得对怀玉山高山马铃薯适应高海拔生境的全面认识。

参考文献:

- [1] 清源,李佩华,方志荣,等. 马铃薯专性寄生真菌内生集壶菌的研究进展 [J]. 西昌学院学报(自然科学版), 2019, 33(4): 16-19, 74.
- [2] 王英,张富仓,王海东,等. 滴灌频率和灌水量对榆林沙土马铃薯产量、品质和水分利用效率的影响 [J]. 应用生态学报, 2019, 30(12): 4159-4168.
- [3] 申光辉,覃小艳,刘海娜,等. 紫色马铃薯全粉加工的汁液腥味物质成分及酶解脱腥效果 [J]. 华南农业大学学报, 2019, 40(4): 83-91.
- [4] 尹明华,叶思雨,宁本松,等. 高山马铃薯脱毒苗 DNA 甲基化的 MSAP 分析 [J]. 核农学报, 2019, 33(6): 1079-1087.
- [5] 尹明华,谭鑫,郑亚娇,等. 高海拔生境下怀玉山高山马铃薯和怀玉山本土农家薯块茎的转录组分析 [J]. 核农学报, 2019, 33(7): 1330-1339.
- [6] CUI D, MA J Y, LIU Y, et al. Analysis of Long Non-Coding RNA Expression Profiles Using RNA Sequencing in Ovarian Endometriosis [J]. Gene, 2018, 673: 140-148.
- [7] 刘志强,朱国琴,陈冉,等. 转录组学研究方法在药用植物银杏中的应用 [J]. 上海中医药大学学报, 2020, 34(1): 88-93.
- [8] 慧芳,刘秀岩,李宗谕,等. 转录组测序技术在药用植物研究中的应用 [J]. 中草药, 2019, 50(24): 6149-6155.
- [9] 宁宇,郑彦超,武高洁,等. 扁秆荆三棱不同组织混合样品的转录组分析 [J]. 华北农学报, 2019, 34(6): 74-81.
- [10] 张飞,王艳秋,朱凯,等. 不同耐盐性高粱在盐逆境下的比较转录组分析 [J]. 中国农业科学, 2019, 52(22): 4002-4015.
- [11] 张纯,唐承晨,王吉永,等. 转录组学在植物应答逆境胁迫中的研究进展 [J]. 生物学杂志, 2017, 34(2): 86-90.
- [12] 白子彧,丁博,李杨,等. 小麦盐应答基因 TaSR1 的生物信息学鉴定及表达验证 [J]. 麦类作物学报, 2017, 37(3): 307-311.
- [13] 赵丽娜,张芙蓉,莫霏,等. 甜瓜盐碱逆境生理响应及相关基因研究进展 [J]. 上海农业学报, 2016, 32(6): 176-180.
- [14] 王晓娇,蒙美莲,曹春梅,等. 水分胁迫下马铃薯萌芽出苗期根系转录组差异表达分析 [J]. 东北师大学报(自然科学版), 2019, 51(4): 109-117.
- [15] 蒋明,曹家树. 查尔酮合成酶基因 [J]. 细胞生物学杂志, 2007, 29(4): 525-529.
- [16] 崔建东, LI Yan, 牟德华. 苯丙氨酸解氨酶(PAL)的研究进展 [J]. 食品工业科技, 2008, 29(7): 306-308.
- [17] 陈璋. 植物-病原真菌相互作用的分子细胞生物学 [J]. 生物技术通报, 1993 (12): 3-6.
- [18] 周荣汉. 药用植物化学分类学 [M]. 上海: 上海科学技术出版社, 1988.
- [19] 姚新生. 天然药物化学 [M]. 北京: 人民卫生出版社, 2001.
- [20] 章维华. 新型香豆素类化合物的合成及其抑菌活性研究 [D]. 南京: 南京农业大学, 2010.

- [21] 张 旭, 王小佳, 黎思辰, 等. 柑橘果实粒化过程中木质素生物合成与调控研究进展 [J]. 浙江农业学报, 2019, 31(12): 2131-2140.
- [22] 范旭杪, 秦 丽, 王吉秀, 等. 植物谷胱甘肽代谢与镉耐性研究进展 [J]. 西部林业科学, 2019, 48(4): 50-56.
- [23] SCHÜTZENDÜBEL A, POLLE A. Plant Responses to Abiotic Stresses: Heavy Metal-Induced Oxidative Stress and Protection by Mycorrhization [J]. Journal of Experimental Botany, 2002, 53(372): 1351-1365.
- [24] 牛歆雨, 雷玉山, 梁 东, 等. 猕猴桃果实 L-半乳糖内酯脱氢酶和脱氢抗坏血酸还原酶 cDNA 片段的克隆与序列分析 [J]. 西北农林科技大学学报(自然科学版), 2007, 35(12): 57-62.
- [25] 张庆田, 范书田, 艾 军, 等. 软枣猕猴桃 L-半乳糖内酯脱氢酶基因的克隆及原核表达分析 [J]. 山东农业科学, 2019, 51(11): 1-7.
- [26] 郑春芳, 刘伟成, 魏 龙, 等. 外施褪黑素对低温胁迫下红树植物秋茄光合作用和抗坏血酸-谷胱甘肽循环的调控 [J]. 植物生理学报, 2019, 55(8): 1211-1221.
- [27] 姚悦梅, 孙德举, 李建斌, 等. 低温对羽衣甘蓝幼苗氮代谢的影响 [J]. 江苏农业学报, 2019, 35(5): 1191-1196.
- [28] ÅGREN G I, WETTERSTEDT J Å M, BILLBERGER M F K. Nutrient Limitation on Terrestrial Plant Growth-Modeling the Interaction between Nitrogen and Phosphorus [J]. New Phytologist, 2012, 194(4): 953-960.

Transcriptome Analysis of Alpine Potato in Huaiyushan Mountains and Its Degraded Tuber Cultivated in Plain Area

HONG Sen-rong¹, LIU Jie¹,
ZENG Xin-yi¹, CAI Hong², CHEN Rong-hua²

1. College of Life Sciences, Shangrao Normal University, Shangrao Jiangxi 334001, China;

2. Red Sun Agricultural Development Co., Ltd., Shangrao Jiangxi 334700, China

Abstract: In order to explore the mechanism of morphological and quality variation caused by species degeneration, the transcriptome analysis was carried out on the tubers of normal potato species grown on high mountains in Huaiyushan and its degraded species cultivated in plain area. The results show that the number of differentially expressed genes in tubers of normal potato and its degenerate species was 3 604, the number of up-regulated genes was 1 331, and the number of down-regulated genes was 2 273. GO enrichment analysis shows that the differential genes were mainly annotated to response to cadmium ion, root hair elongation, response to water deprivation, defense response to fungus, response to salt stress, response to abscisic acid, cell part, plasma membrane, intracellular membrane-bounded organelle, chloroplast, Golgi apparatus, photosystem II, photosystem I, protein binding, electron carrier activity, heme binding, chlorophyll binding, glutathione binding and other functions. The analysis of KEGG enrichment shows that the differential genes were mainly related to phenylalanine metabolism, plant-pathogen interaction, phenylpropanoid biosynthesis, glutathione metabolism, galactose metabolism, photosynthesis, nitrogen metabolism and other biological functions. qRT-PCR was used to detect the expression of 10 different genes in the tubers of normal and degraded potato species in huaiyushan mountains, and the trend was consistent with that of RNA SEQ. The results will provide a reference for the analysis of gene expression of morphological changes and quality decline of the alpine potato in Huaiyushan after it was planted in plain area.

Key words: alpine potato in Huaiyushan; normal potato; variant; tuber; transcriptome analysis