

庞欣,陈军,李林芝,等. 基于转录组测序的不同辣度辣椒差异表达基因分析[J]. 江苏农业科学,2025,53(6):24-33.

doi:10.15889/j.issn.1002-1302.2025.06.003

基于转录组测序的不同辣度辣椒差异表达基因分析

庞欣¹, 陈军¹, 李林芝¹, 刘佳², 黄文娟²

(1. 苏州农业职业技术学院园艺科技学院, 江苏苏州 215008; 2. 乌兰察布市农林科学研究所, 内蒙古乌兰察布 012000)

摘要:为了研究不同辣度辣椒的基因表达差异,探究参与或调控辣椒素生物合成的基因或转录因子。试验以 2 种不同辣度的辣椒 B1-2(高辣)、D50(微辣)为材料,通过高通量 RNA 测序(RNA-seq)对 35 d 胎座进行转录组分析,共筛选出 1 570 个差异表达基因(DEG),包括 673 个上调基因、897 个下调基因。经 GO 富集分析,DEG 主要集中在氧化还原反应、光合作用、光反应等生物过程;经 KEGG 富集分析,DEG 主要富集在光合作用-天线蛋白、次级代谢产物的生物合成、代谢途径等信号通路。鉴定出 7 个 CapCyc 模型基因、8 个 MYB 转录因子在 2 份辣椒材料中的差异表达;在果实不同发育时期进一步利用 qRT-PCR 进行分析,发现 4CL、COMT、KAS 可能具有调控辣椒素合成的作用。研究结果可为探索辣椒素生物合成的调控网络提供更多的候选基因,期待可为下一步关键基因的功能分析和辣椒育种提供理论基础。

关键词:辣椒;转录组测序;差异表达基因;辣椒素类物质;生物合成;转录因子

中图分类号:S641.301 **文献标志码:**A **文章编号:**1002-1302(2025)06-0024-09

辣椒属于茄科(Solanaceae)一年生或多年生草本植物,是我国种植面积最大的蔬菜,年种植面积稳定在 210 万 hm^2 以上,也是全球消费量最大的辛辣调味品^[1-2]。辣椒中含有一种特殊的物质——辣椒素,又称辣椒碱,主要在辣椒胎座中合成,常在食品和临床医学上被用于调味或镇痛。不同栽培品种辣椒的辣椒素类物质含量差别较大,史高维尔指数(scoville heat unit, SHU)一般在 $0 \sim 2.5 \times 10^6$ 不等^[3]。因此,辣椒育种主要的目标之一是培育产生大量辣椒素的优质辣椒品种。

辣味的有无由显性遗传位点控制。目前已报道的辣椒辣味基因有 *Pun1*、*Pun2*、*Pun3*、*pAMT*[假定氨基转移酶(putative aminotransferase)基因]、*CaKRI*[假定酮脂酰 ACP 还原酶(putative ketoacyl-ACP reductase)基因]^[4-10]等。*Pun1* 基因(AT3)的

2 个拷贝可能起到协同作用,共同合成并积累含量不同的辣椒素,决定不同品种辣椒的辣度^[11]。*Pun2* 是野生辣椒 *Capsicum chacoense* 的辣味遗传位点^[5]。Yi 等利用 *C. chacoense*、*C. annuum* 种间 F_2 代遗传群体定位,获得 *Pun2* 候选基因 *pAMT*,发现在无辣味 *C. chacoense* 材料中,*pAMT* 第 8 外显子发生 7 bp 的插入突变^[6]。*Pun3* 是从栽培种一年生辣椒中发现的,属 MYB 转录因子^[12]。Han 等将 *Pun3* 定位在 7 号染色体上,并预测 *CaMYB31* 为 *Pun3* 候选基因^[7]。Zhu 等定位了与辣椒素含量相关的主效 QTL (*Capsaicinoid 1*, *Cap1*)^[8],*Cap1* 与已报道的辣椒素含量相关数量遗传位点 *cap*、*cap7.2*、*Pun3* 位置重叠,其候选基因 *MYB31* 的核苷酸和蛋白质序列与 *CaMYB31*、*Pun3*/*CaMYB31* 高度相似^[7,13-15]。*CaKRI* 是从中国辣椒经图位克隆发现的,无辣味隐性遗传位点,获得候选基因 *CA10g18840*^[10]。辣椒素含量的多少决定辣味的高低,呈现数量遗传特征^[16-18],如 *Pun3* 就是从辣椒素含量主效 QTL 发现的。

辣椒素类物质是通过香草基胺和支链脂肪酸,由辣椒素合成酶(CS)合成的酰胺化合物。苯丙氨酸经过一系列酶的催化作用,生成香草基胺,其中关键酶有 PAL、C4H、C3H、COMT、KAS 等;8-甲基-6-壬烯辅酶 A 的支链脂肪酸由缬氨酸经过系列酶催化生成,BCAT、IVD、ACL、FatA、ACS 等是参与该途径的关键酶;两者在辣椒素合成酶的催化下

收稿日期:2024-03-01

基金项目:国家自然科学基金(编号:32100302、32260769、31960601);2022 年江苏省高校“青蓝工程”中青年学术带头人项目;内蒙古农牧业青年创新基金(编号:2023QNJJN03);乌兰察布市基础科研项目(编号:2021JC202);内蒙古自然科学基金(编号:2021MS03106、2021MS03071)。

作者简介:庞欣(1985—),女,黑龙江哈尔滨人,博士,副教授,主要从事辣椒分子生物学研究。E-mail:pxtracy916@163.com。

通信作者:黄文娟,副研究员,主要从事辣椒种质资源的鉴定及基因挖掘研究。E-mail:huangwenjuan_013@163.com。

形成了辣椒素^[19-22]。Mazourek 等对辣椒素生物合成途径建立了 CapCyc Model 模型,但目前仅了解该模型中的少量基因,模型中其他基因的具体功能还需进一步研究^[23]。

得益于辣椒基因组序列的公布和转录组学的快速发展,辣椒 MYB 转录因子功能的研究也取得了进展,如调控辣椒素等次生代谢产物的合成。Liu 等发现, *CaMYB37* 基因与 AT3 的启动子结合并激活其转录,从而调节辣椒素的生物合成^[24]。Sun 等还发现, *Cap1*、*MYB31* 这 2 个基因通过调控辣椒素合成途径相关基因的表达,来调控辣椒素的积累^[25]。Sun 等发现, *CaMYB108* 受到茉莉酸诱导,参与调控辣椒素类物质的合成^[26]。

近年来,转录组测序(RNA sequencing, RNA-seq)技术已大量应用于基础研究和分子育种等领域。一些研究者利用该技术对辣椒进行研究,如雄性不育^[27]、果实形态^[28]、果实着色^[29]、辣椒疫病^[30-31]、黄瓜花叶病毒病^[32]、辣度^[33]等。本研究选取花后 35 d 的高辣、低辣 2 种辣椒的胎座,利用 RNA-seq 技术对差异表达基因进行 GO、KEGG 等分析,旨在为以后辣椒辣度相关基因的功能鉴定、品种选育等提供一定的理论基础。

1 材料与方法

1.1 试验材料

供试辣椒材料为 B1-2(高辣)、D50(低辣),由浙江省农业科学院蔬菜研究所提供。试验于 2023 年 3 月在苏州农业职业技术学院试验基地进行。田间试验采取常规田间管理,在辣椒开花后 35 d 采集辣椒果实中的胎座,分别命名为 WIF(B1-2)、GIF(D50),立即放入液氮速冻,然后转移到 -80℃ 超低温冰箱中保存。试验设置 3 次生物学重复。

1.2 转录组测序

采用 Trizol 试剂法提取辣椒胎座中的全部 RNA,按照试剂盒说明书完成操作步骤。由南京诺唯赞生物科技有限公司完成转录组测序工作。通过 3 种方法检测 RNA 样品, RNA 的纯度由 Nanodrop 检测($D_{260\text{ nm}/280\text{ nm}}$ 比值), RNA 的浓度由 Qubit 进行定量, RNA 的完整性由 Agilent 2100 生物分析仪精确检测(RIN 值)。质检合格后,采用 mRNA 捕获磁珠富集真核生物 mRNA,以加热的方式达到 mRNA 的片段化。以 mRNA 为模板,进行一链 cDNA 的合成、二链 cDNA 的合成、双链 cDNA 的

纯化。纯化后的双链 cDNA 进行末端修复,加 A 尾并连接测序接头,进行片段大小分选;最后进行 PCR 扩增,纯化 PCR 产物,得到最终的文库。采用 Agilent 2100 生物分析仪对构建好的文库进行质检。文库质检合格后,根据不同需求 pooling 后,再进行 Illumina HiSeq 测序。

1.3 差异表达基因的筛选与分析

从 NCBI、UCSC、Ensembl 等数据库下载参考序列,对相应样品的原始测序数据进行生物信息分析。将原始测序数据进行过滤[去除含接头、含 N 过多(>5%)或含大量低质量碱基的数据],进而得到高质量的测序数据,用于后续信息分析。使用 Tophat 2,利用参考序列进行序列比对分析,根据比对结果进行质量评估,从而完成第 2 次数据质检。利用 Cufflinks 软件完成基因表达定量分析,基因表达计算方法为 FPKM(每百万映射片段中每千碱基转录本的片段数)。显著性差异表达基因筛选的标准为: $|\log_2(\text{fold change})| \geq 1$ 且 q 值 ≤ 0.05 (fold change 为差异倍数, q 值为 P 值经过多重检验校正后的结果)。将差异表达基因与 Gene Ontology、KEGG Pathway 数据库对比,利用超几何检验找出显著性富集的 GO、Pathway 条目。

1.4 实时荧光定量 PCR 检测

为了探究辣椒素类物质生物合成途径和转录因子 MYB 的差异表达基因在 2 份材料间的表达量,挑选 16 个差异表达基因,采用实时荧光定量 PCR (qRT-PCR) 进行检测。qRT-PCR 反应体系为: 10 μL SYBR Green I, 1 μL cDNA 模板,各 1 μL 的上、下游引物, 7 μL ddH₂O, 总反应体积共 20 μL 。反应程序为: 95℃ 预变性 10 min; 95℃ 变性 30 s, 58℃ 退火 30 s, 72℃ 延伸 20 s, 35 个循环; 3 次重复。以 *GAPDH* 为内参基因,利用 Primer 5.0 软件设计特异性引物(表 1)。

2 结果与分析

2.1 转录组测序数据质量分析

为了分析 2 种不同辣度辣椒的转录组水平差异,将 B1-2、D50 开花后 35d 的胎座进行转录组测序。WIF1、2、3 和 GIR1、2、3 分别代表 B1-2 和 D50 的 3 次生物学重复。经原始测序数据质检、过滤,结果如图 1 显示。6 个样本获得的 clean reads 在 49 898 206 ~ 58 602 230 之间,样本产生的 clean bases 在 7 407 290 700 ~ 8 790 334 500 之间,共得到

表 1 qRT-PCR 引物序列

基因名称	正向引物(5'→3')	反向引物(5'→3')
LOC107849600	TTATTGCCGGAAGGCTACCA	GGGTAATTGGGCTTGGGTTG
LOC107851710	TCATGCTGAAGGTCGTTGGA	TTCCACGTCGAACATCTGCT
LOC107852202	GCTGGGCTACAAAGATGTGG	TTCCGGGTAAATGACTTGCG
LOC107858775	ATGGGTAGACTCCCTTGCTG	TATCCCTGGCCTCAGGTAGT
LOC107862231	AGGACCATGGACACCAGAAG	AGTCCATCTGAGCCTACAGC
LOC107864180	TGTCGTCTAGCGAGGTATCG	TCCATCTGATCACCACCGTT
LOC107871221	TCATGCTGAAGGTCGTTGGA	TTCCCACGTCGAACATCTGG
LOC107877223	GGCAATAAATGGGCTCGCA	GGCTTTGTGGGCTTGGTTA
LOC107851037	TCAAGGAAGGAACAGCACCA	GCCTTTGCTTTCTTCATCGC
LOC107863937	GGTTCTTGAACGCTGGAGTC	CCTTATGCCGGAGTACCCCTT
LOC107845437	ATGCTTTGTGTGTTGCCCTT	TGGCCCTATTGTGACCTTGT
LOC107869755	TGGATTGGATCACCGTTCA	GCAATATCACCGCTGTGGAG
LOC107840262	CCGCTTTGCCTGTTCTTGAT	CATTGGTGCATCTGCTGGAG
LOC107848771	CGGTATAGCAACGCGTTAG	TAAGGCCAGCTCCAAATCCA
LOC107864113	CGAGGGACAGAGTGAATCA	GGCCAAGGTTAGCATCTTCG
LOC107859692	CGCACAAAGATTGGTGATGGT	GTACGCACTGGTTGAGATCG

49.7 Gb 的高质量 clean 数据。Q20 碱基百分比为 97.22% ~ 97.43%, Q30 碱基百分比为 93.38% ~ 93.74%, GC 含量为 41.64% ~ 42.45% (表 2)。以上结果表明,6 组样品转录组测序质量较好,指标符合要求,可用于后续分析。

将质控后的 clean reads 进行序列比对分析,比对到参考基因组唯一位置的 reads 数在测序有效数

据中的比例为 86.7% ~ 87.8% (表 2)。统计比对到参考序列不同位置的 reads,结果发现,6 个样品比对到外显子区域的比例平均为 65.16%,比对到内含子区域的比例平均为 11.26%,比对到基因间区的比例平均为 23.59% (图 1),说明比对结果具有可靠性。

表 2 转录组测序数据统计结果

样品名称	原始测序	高质量测序	GC 含量 (%)	Q20 比例 (%)	Q30 比例 (%)	比对到参考基因组的序列数及所占比例 (%)
WIF1	53 893 080	53 793 790	41.64	97.22	93.38	46 720 834 (86.9)
WIF2	55 927 852	55 777 654	42.45	97.31	93.53	48 991 783 (87.8)
WIF3	49 996 424	49 898 206	41.78	97.30	93.47	43 242 336 (86.7)
GIF1	58 723 612	58 602 230	41.80	97.31	93.50	50 992 181 (87.0)
GIF2	56 626 674	56 427 290	41.73	97.30	93.52	49 060 236 (86.9)
GIF3	56 932 874	56 810 652	41.76	97.43	93.74	49 437 468 (87.0)

2.2 基因表达分析

基因表达的箱线图可以显示各个样本基因表达水平的离散程度。如图 2 所示, WIF1、WIF2、WIF3 和 GIF1、GIF2、GIF3 组内样本的表达水平分布相近,而 WIF、GIF 组的基因量分布较远,说明不同辣椒材料在基因表达水平差异显著。

样品间基因表达水平相关性可以检验试验可靠性和样本选择是否合理,本试验采用 3 次生物学重复,从表 3 可以看出,2 个材料重复处理间的皮尔逊 (Pearson) 相关系数的平方 (r^2) 都大于 0.9,说明试验数

据可靠。

2.3 差异表达基因分析

利用 DESeq2 软件对辣椒各样本的转录组数据进行分析,以 $|\log_2(\text{fold change})| \geq 1$ 且 q 值 ≤ 0.05 作为显著性差异表达基因筛选的标准。2 个不同辣度辣椒的差异基因分布用火山图来表示 (图 3),其中红色代表显著上调的基因,绿色代表显著下调的基因,蓝色代表无显著性差异的基因。结果表明, WIF 组与 GIF 组之间鉴定出 1 570 个差异表达基因,其中 673 个 DEG 表达上调,897 个 DEG 表达下调。

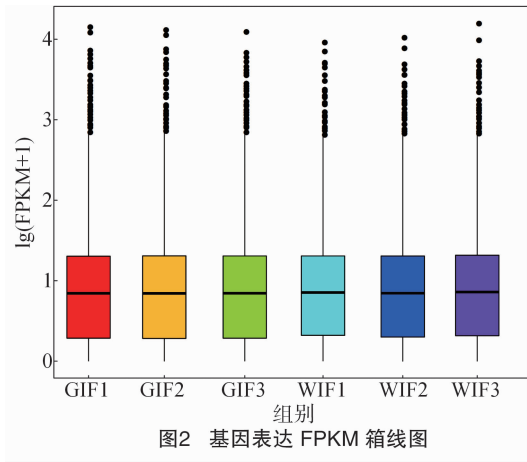
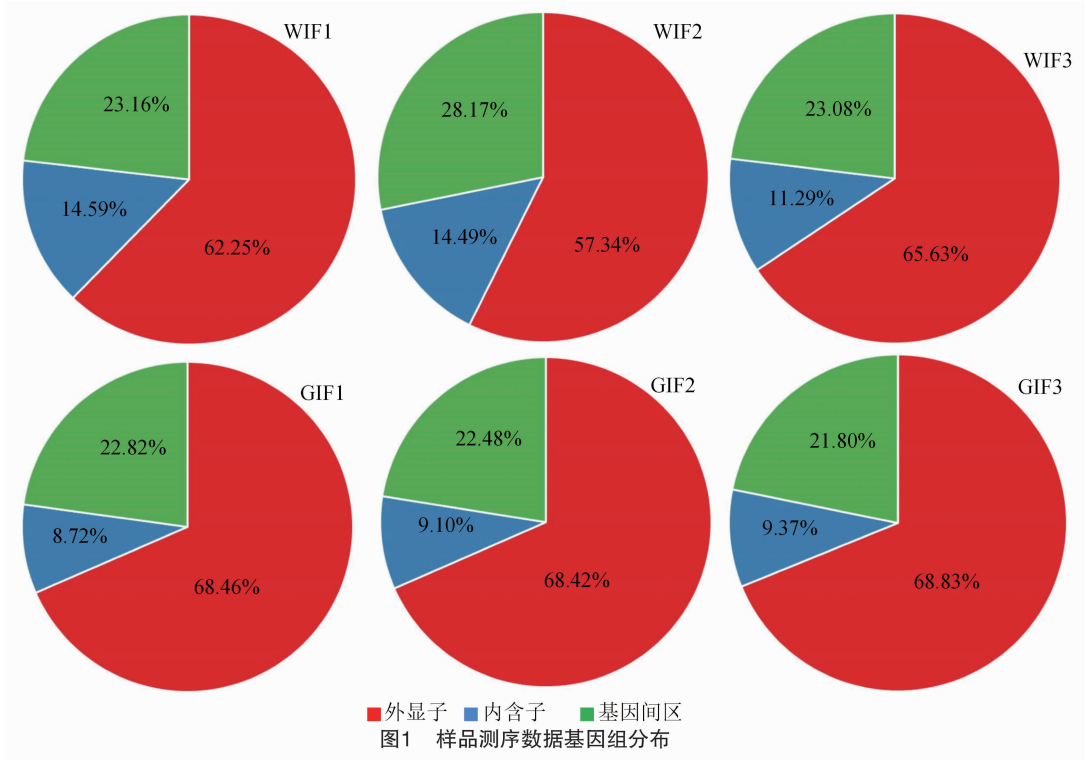


表 3 样品间基因表达相关性分析

处理组	r^2					
	WIF1	WIF2	WIF3	GIF1	GIF2	GIF3
WIF1	1.000 0					
WIF2	0.902 8	1.000 0				
WIF3	0.912 1	0.946 6	1.000 0			
GIF1	0.923 7	0.839 4	0.889 3	1.000 0		
GIF2	0.926 4	0.844 9	0.881 5	0.993 2	1.000 0	
GIF3	0.929 5	0.844 6	0.885 1	0.988 5	0.993 0	1.000 0

2.4 差异表达基因的 GO 功能富集分析

对差异表达基因进行 GO 功能富集分析,得知其可分为生物过程、细胞成分、分子功能 3 类(图 4)。生物过程中主要富集了氧化还原反应、光合作

用、光反应等,基因数分别是 154、44、30 个。分子功能过程主要富集了氧化还原酶活性;氧化还原酶作用于成对供体,伴随分子氧的掺入或还原、四吡咯结合、单加氧酶活性等,基因数分别为 195、63、54、53 个。细胞成分过程主要富集了类囊体、类囊体部分、光合膜、叶绿体类囊体膜等,基因数分别为 67、56、53、47 个。

2.5 差异表达基因 KEGG 富集分析

对差异表达基因进行 KEGG 富集分析,其中显著($P < 0.05$)富集的通路有 28 条,共 1 032 个差异表达基因;图 5 展示的是丰度排名前 20 位的 KEGG 富集通路。涉及到的通路按照显著性差异水平由高到低依次是:光合作用-天线蛋白,次级代谢产物的生物合成,类黄酮生物合成,异黄酮生物合成,芪类、二芳基庚烷和姜辣素生物合成,代谢途径,二萜生物合成,光合作用,柠檬烯和蒎烯降解,苯丙烷代谢,类胡萝卜素合成,激素信号转导,角质、木栓质和蜡质生物合成,不饱和脂肪酸的生物合成,黄酮和黄酮醇的生物合成,淀粉和蔗糖代谢,亚油酸代谢, α -亚油酸代谢,卟啉和叶绿素代谢,其他聚糖降解。

2.6 辣椒素类物质生物合成途径的差异表达基因

辣椒素的生物合成模型中共鉴定出 11 个差异表达基因(表 4),包含 2 个 BCAT(支链氨基酸转移

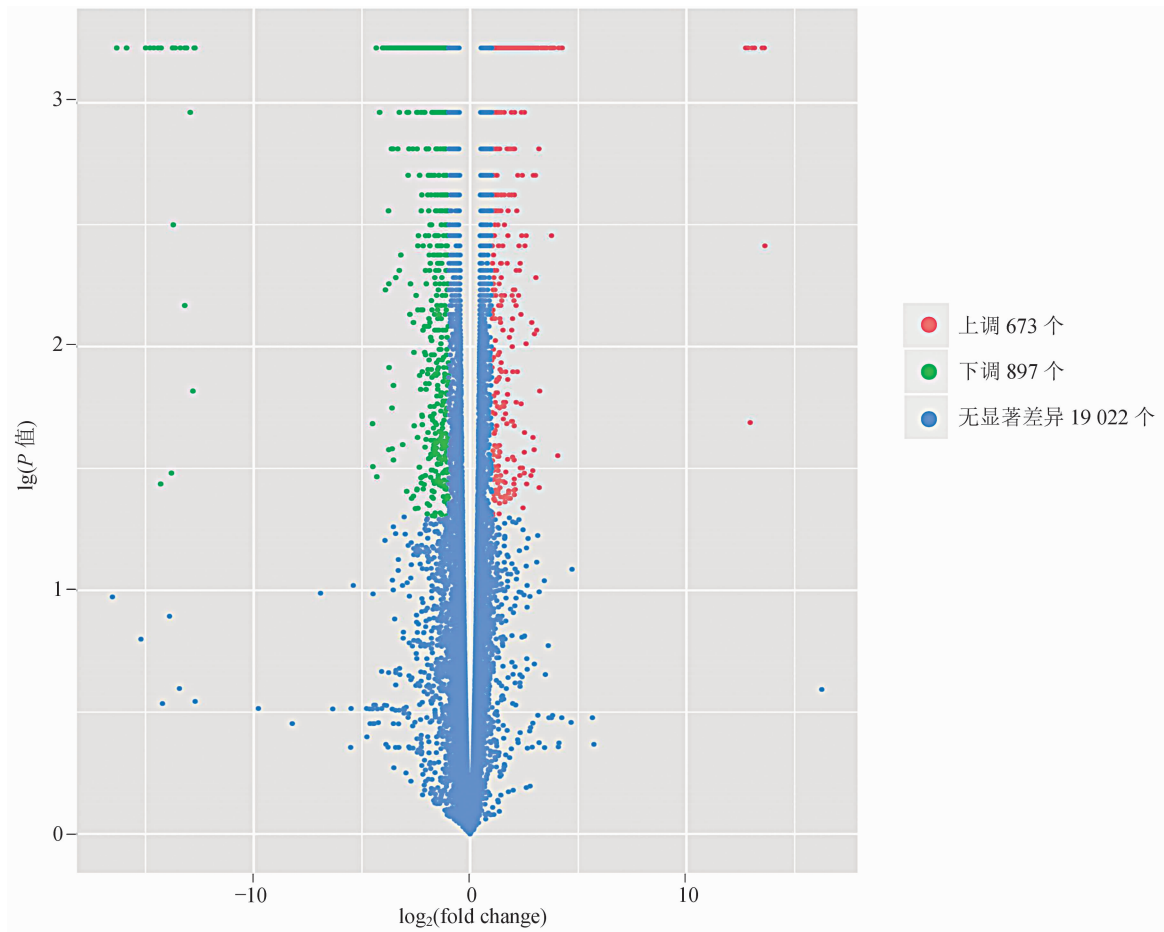


图3 2 种辣椒的差异表达基因火山图

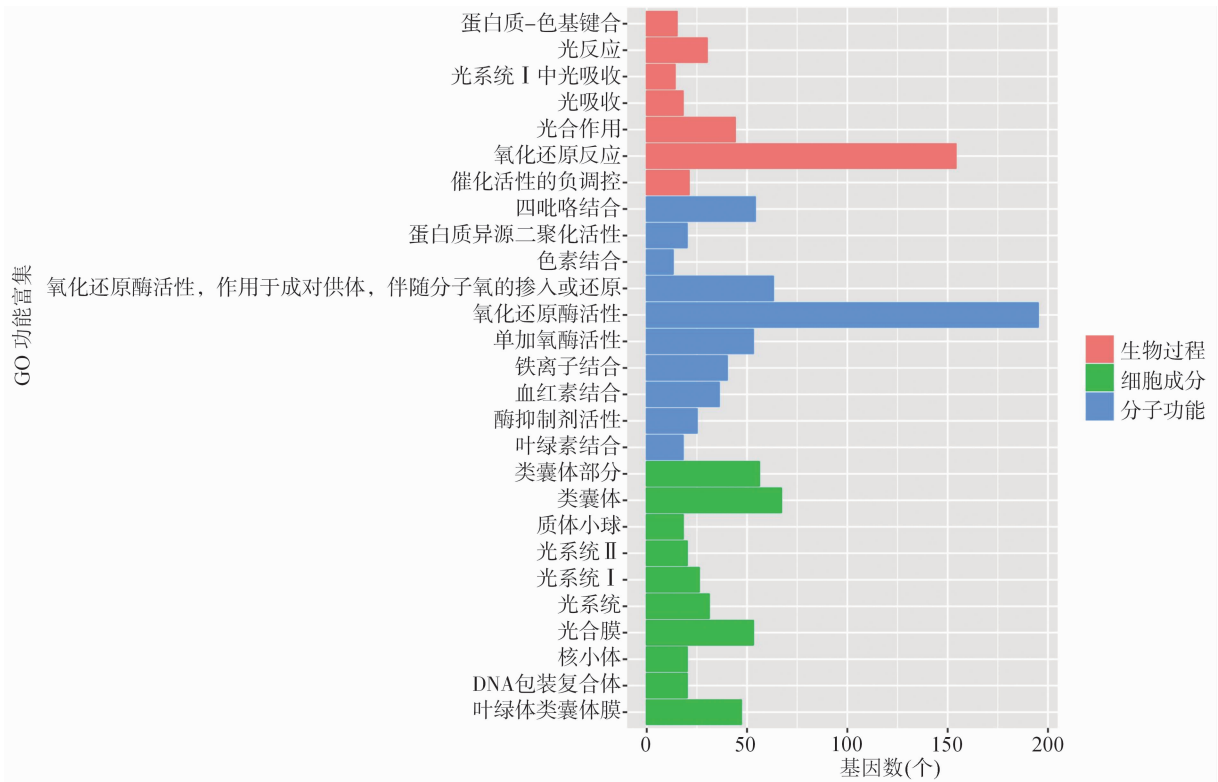


图4 显著性差异表达基因 GO 功能富集分析

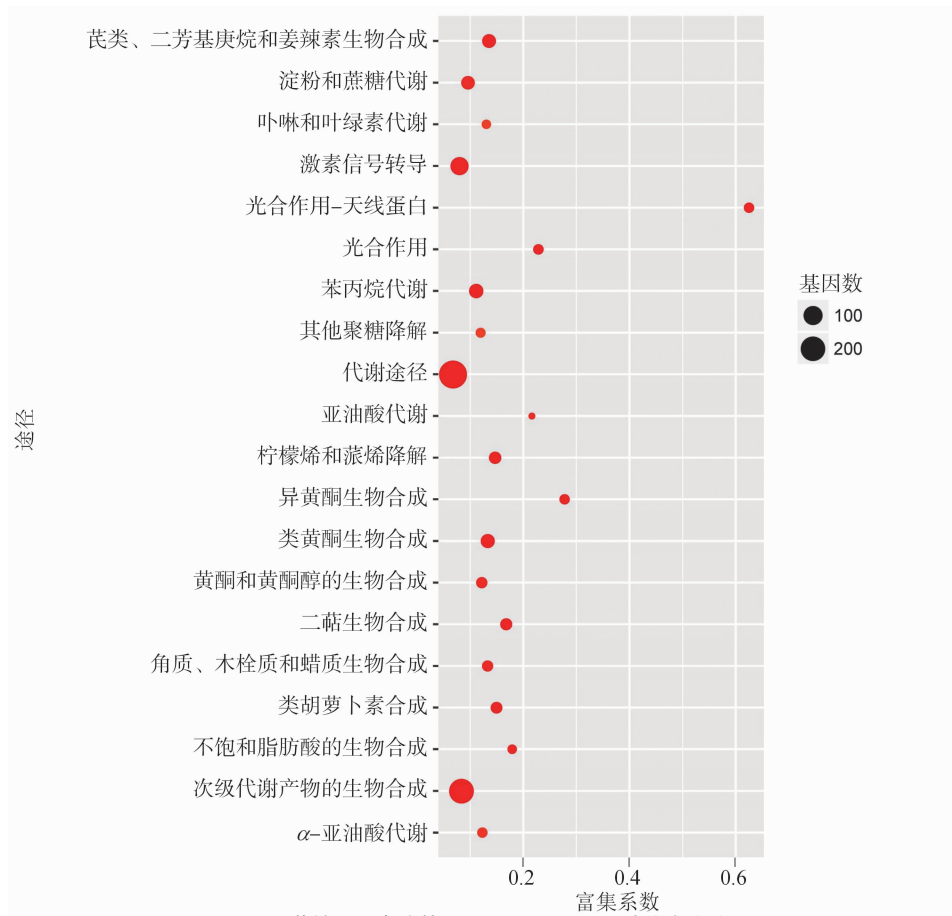


图5 显著性差异表达基因 KEGG Pathway 功能富集分析

酶)、2 个 PAL(苯丙氨酸解氨酶)、1 个 C4H(肉桂酸-4-羟化酶)、2 个 4CL(4-香豆酰-CoA 连接酶)、1 个 COMT(咖啡酸转甲氧基酶)、2 个 KAS(β -酮脂酰-ACP 合成酶)、1 个 AT(脂肪酰转移酶)。在 11 个差异表达基因中,有 1 个 BCAT 和 1 个 KAS 基因上调表达,其余基因皆为下调表达。

2.7 转录因子 MYB 的差异表达基因

在 2 个辣椒样本差异表达基因库的基础上,对参与的 MYB 转录因子进行筛选,共鉴定出 8 个差异表达基因(表 5)。其中 3 个 DEG 表达上调,5 个 DEG 表达下调。*LOC107849600* 差异倍数最高,为 1.958 84 倍,*LOC107864180* 差异倍数最低,仅为 1.004 03 倍。

表 4 辣椒素类物质生物合成途径相关的差异表达基因

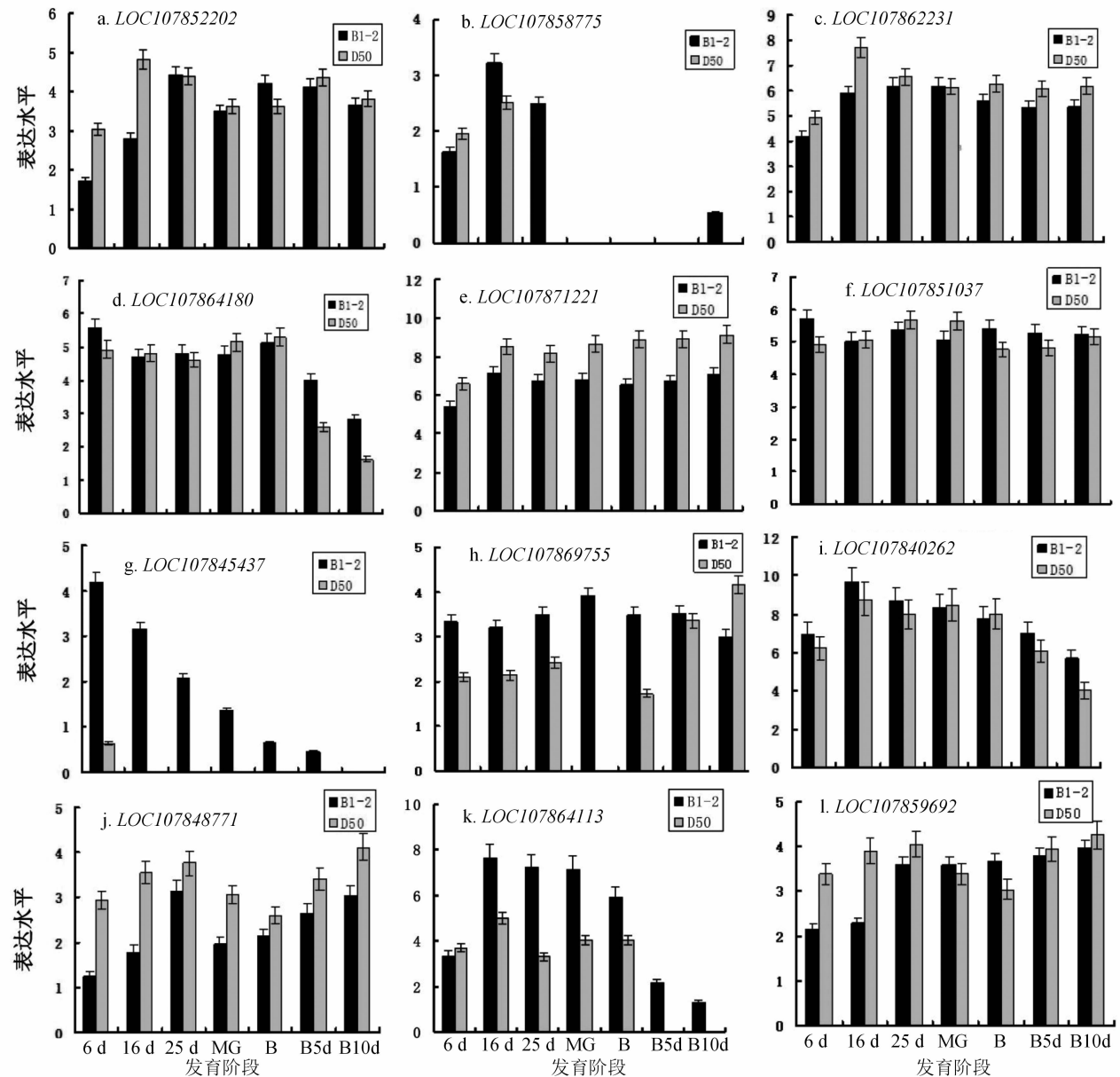
基因家族	基因编码酶	基因名称	差异倍数	P 值	q 值
BCAT	支链氨基酸转移酶	<i>LOC107851037</i>	-1.063 71	0.000 05	0.000 595
		<i>LOC107867296</i>	1.200 57	0.000 05	0.000 595
PAL	苯丙氨酸解氨酶	<i>LOC107843099</i>	-1.387 38	0.000 05	0.000 595
		<i>LOC107863937</i>	-1.573 65	0.000 05	0.000 595
C4H	肉桂酸-4-羟化酶	<i>LOC107875407</i>	-1.321 67	0.000 05	0.000 595
4CL	4-香豆酰-CoA 连接酶	<i>LOC107845437</i>	-2.279 71	0.002 40	0.014 981
		<i>LOC107869755</i>	-1.014 79	0.000 05	0.000 595
COMT	咖啡酸转甲氧基酶	<i>LOC107840262</i>	-1.242 17	0.000 05	0.000 595
KAS	β -酮脂酰-ACP 合成酶	<i>LOC107848771</i>	-1.280 55	0.000 05	0.000 595
		<i>LOC107864113</i>	1.162 64	0.000 05	0.000 595
AT	脂肪酰转移酶	<i>LOC107859692</i>	-1.196 68	0.000 05	0.000 595

表 5 转录因子 MYB 的差异表达基因

基因名称	基因描述	差异倍数	P 值
<i>LOC107849600</i>	Myb 相似蛋白 308	1.958 84	0.000 05
<i>LOC107851710</i>	转录因子 MYB 相似蛋白 108	-1.640 14	0.001 00
<i>LOC107852202</i>	转录因子 MYB86	-1.351 40	0.003 95
<i>LOC107858775</i>	Myb 相似蛋白 306	1.110 76	0.000 05
<i>LOC107862231</i>	Myb 相似蛋白 306	-1.610 47	0.000 05
<i>LOC107864180</i>	Myb 相似蛋白 Myb4	-1.004 03	0.000 05
<i>LOC107871221</i>	转录因子 MYB108	-1.071 58	0.002 50
<i>LOC107877223</i>	Myb 相似蛋白 306	1.163 59	0.002 80

2.8 差异表达基因的 qRT-PCR 分析

为了进一步分析差异表达基因在 2 个材料果实发育不同阶段的表达情况,从辣椒素类物质生物合成途径、MYB 转录因子差异表达基因中,选择 16 个进行 qRT-PCR 分析。其中,*LOC107849600*、*LOC107851710*、*LOC107877223*、*LOC107863937* 这 4 个基因表达量较低,几乎不表达。由图 6 可见,在 7 个发育阶段,B1-2 的 *LOC107845437* 表达量都大于 D50,该基因在 B1-2 胎座第 6 天的表达量最大,随后逐渐减小。除了在转色期第 10 天阶段,在其余 6



6、16、25 d分别代表第 6 天、第 16 天、第 25 天的胎座, MG 代表成熟绿果胎座, B 代表转色期绿果胎座, B5d 代表转色期第 5 天胎座, B10d 代表转色期第 10 天胎座

图6 差异表达基因 qRT-PCR 分析

个阶段, B1-2 的 *LOC107869755* 基因表达量都大于 D50, 在成熟绿果期达到最大值。 *LOC107840262* 基因整体表达量都较高, 除了成熟绿果、转色期绿果阶段, B1-2 表达量都大于 D50, 并且 B1-2 在胎座第 16 天的表达量最高, 随后表达量逐渐降低。从第 16 天胎座开始的果实发育阶段, B1-2 的 *LOC107864113* 基因表达量都比 D50 高; *LOC107840262* 基因, B1-2 表达量的峰值出现在胎座第 16 天, 之后随着果实发育表达量逐步减少。几乎在所有的果实发育阶段, *LOC107852202*、*LOC107862231*、*LOC107871221*、*LOC107848771* 这 4 个基因在 B1-2 的表达量都小于 D50。

3 讨论与结论

本研究利用 RNA-seq 测序技术, 对不同辣度的辣椒材料 B1-2、D50 开花后 35 d 果实的胎座进行转录组分析。Illumina Hiseq 测序结果得到的数据量较大, 试验效率较高, 同时成本低, 因此可用于辣椒辣度的转录组测序研究。通过测序, 本研究共获得 331 309 822 条的 reads, 总碱基数为 49.7 Gb, 每个样品均获得 7.4 Gb 以上的碱基。有 86.7% ~ 87.8% 的 clean 数据比对到唯一的基因组位点, 表明本研究的高通量测序结果可靠, 可以进一步探究 2 个材料间辣椒素合成过程中的差异表达基因。对高辣、低辣 2 个辣椒材料进行转录组测序分析, 通过 GO、KEGG 显著性富集分析, 共得到 1 570 个差异表达基因, 主要集中在氧化还原反应、光合作用、光反应等生物过程; 氧化还原酶、氧化还原酶活性等分子功能过程; 类囊体、光合膜等细胞成分过程。富集在光合作用-天线蛋白、次级代谢产物的生物合成、代谢途径、光合作用、苯丙烷代谢、类胡萝卜素合成、激素信号转导等信号通路, 注释到 MYB 类型的转录因子。王红娟等对不同辣度的 2 个辣椒品种及其 2 个不同发育时期的果实进行转录组分析, 结果发现, 在辣椒绿熟期阶段, 2 个辣度品种的差异表达基因主要富集在光合作用、光捕获、蛋白质-生色团连锁中^[33]。

在 Mazourek 等建立的 CapCyc Model 中, 本研究鉴定出 7 个模型基因在 2 份辣椒材料中的差异表达, 共 11 个差异表达基因^[23]。王红娟等鉴定到 18 个 CapCyc 模型基因, 包括本研究中鉴定出的 7 个模型基因 (BCAT、PAL、C4H、4CL、COM、KAS、AT)^[33]。其中 BCAT (*LOC107867296*)、PAL (*LOC107843099*)、

COMT (*LOC107840262*)、KAS (*LOC107864113*)、AT (*LOC107859692*) 是 2 个研究中共同的差异表达基因。为进一步分析模型中差异表达基因在 2 个辣椒材料不同时期、不同组织的表达情况, 利用 qRT-PCR 进行定量分析。*LOC107840262* 基因在 B1-2、D50 不同时期胎座的表达量都较大, 最大出现在 B1-2 第 16 天的胎座, 除成熟绿果、转色期绿果 2 个阶段外, 其他 5 个时期, B1-2 的表达量都大于 D50。在胎座的不同发育时期, *LOC107845437*、*LOC107869755*、*LOC107864113* 这 3 个基因, B1-2 的表达量总体上均显著大于 D50。因此, 综合转录组的差异表达和 qRT-PCR 定量分析, 4CL (*LOC107845437*、*LOC107869755*)、COMT (*LOC107840262*)、KAS (*LOC107864113*) 基因可能具有调控辣椒素合成的作用。

辣椒素的合成也受到 MYB 转录因子的调控作用。周欣根据辣椒基因组序列, 克隆了 4 个 R2R3-MYB 转录因子 (*CcMYB86*、*CcMYB2*、*CcMYB48*、*CcMYB59*), 结果发现 *CcMYB86*、*CcMYB2*、*CcMYB59* 这 3 个转录因子参与了辣椒素的合成调控^[34]。张威等从海南黄灯笼辣椒中克隆并分离 *CcMYB4-2* (PHU22609.1)、*CcMYB4-12* (PHT97629.1) 2 个典型的 R2R3-MYB 转录因子, 推测 *CcMYB4-12* 在辣椒中可能通过抑制木质素和类黄酮代谢途径, 从而调节辣椒素的合成^[35]。Sun 等鉴定了 1 个 R2R3-MYB 转录因子 *CaMYB108*, 证明它参与调控辣椒素的生物合成^[26]。本研究中从差异表达基因中共找到 8 个 MYB 家族转录因子, 包括 MYB108、MYB86、MYB306 等。注释 MYB306 相关蛋白的 *LOC107858775* 基因, B1-2 的表达量从第 16 天开始都大于 D50, 但在成熟绿果、转色期绿果、转色期第 5 天, 在 2 个材料中都几乎检测不到表达量。注释 MYB4 相关蛋白的 *LOC107864180* 基因, B1-2 在胎座第 6 天、第 25 天、转色后 5 d、转色后 10 d 的表达量均大于 D50, 在第 6 天的表达量最高。因此, *LOC107864180*、*LOC107858775* 基因可能在积累辣椒素的个别时期发挥作用。

本研究对高辣辣椒 B1-2、低辣辣椒 D50 第 35 天胎座的进行转录组测序和分析, 共检测到 1 570 个差异表达基因, 其中有 673 个上调, 897 个下调。通过 GO、KEGG 富集分析发现氧化还原反应、光合作用、光合作用-天线蛋白、次级代谢产物的生物合成等途径出现大量差异表达基因。辣椒素类物

质生物合成途径中的 4CL、COMT、KAS 可能具有调控 B1-2 辣椒素合成的作用。期待本试验结果可为后期更深入地研究辣椒素合成的分子机制提供参考,并为辣椒分子辅助育种和培育新品种提供帮助。

参考文献:

- [1] 邹学校, 马艳青, 戴雄泽, 等. 辣椒在中国的传播与产业发展[J]. 园艺学报, 2020, 47(9): 1715-1726.
- [2] 邹学校, 朱凡. 辣椒的起源、进化与栽培历史[J]. 园艺学报, 2022, 49(6): 1371-1381.
- [3] Bosland P W, Coon D, Cooke P H. Novel formation of ectopic (nonplacental) capsaicinoid secreting vesicles on fruit walls explains the morphological mechanism for super-hot Chile peppers[J]. Journal of the American Society for Horticultural Science, 2015, 140(3): 253-256.
- [4] Stewart C Jr, Kang B C, Liu K D, et al. The Pun1 gene for pungency in pepper encodes a putative acyltransferase[J]. The Plant Journal, 2005, 42(5): 675-688.
- [5] Stellari G M, Mazourek M, Jahn M M. Contrasting modes for loss of pungency between cultivated and wild species of *Capsicum* [J]. Heredity, 2010, 104(5): 460-471.
- [6] Yi S, Lee D G, Back S, et al. Genetic mapping revealed that the *Pun2* gene in *Capsicum chacoense* encodes a putative aminotransferase [J]. Frontiers in Plant Science, 2022, 13: 1039393.
- [7] Han K, Jang S, Lee J H, et al. A MYB transcription factor is a candidate to control pungency in *Capsicum annuum* [J]. Theoretical and Applied Genetics, 2019, 132(4): 1235-1246.
- [8] Zhu Z S, Sun B M, Wei J L, et al. Construction of a high density genetic map of an interspecific cross of *Capsicum chinense* and *Capsicum annuum* and QTL analysis of floral traits [J]. Scientific Reports, 2019, 9: 1054.
- [9] Lang Y Q, Kisaka H, Sugiyama R, et al. Functional loss of *pAMT* results in biosynthesis of capsinoids, capsaicinoid analogs, in *Capsicum annuum* cv. CH-19 Sweet [J]. The Plant Journal, 2009, 59(6): 953-961.
- [10] Koeda S, Sato K, Saito H, et al. Mutation in the putative ketoacyl-ACP reductase *CaKR1* induces loss of pungency in *Capsicum* [J]. Theoretical and Applied Genetics, 2019, 132(1): 65-80.
- [11] Qin C, Yu C S, Shen Y O, et al. Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization [J]. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111(14): 5135-5140.
- [12] Zhu Z S, Sun B M, Cai W, et al. Natural variations in the MYB transcription factor MYB31 determine the evolution of extremely pungent peppers [J]. New Phytologist, 2019, 223(2): 922-938.
- [13] Blum E, Mazourek M, O'Connell M, et al. Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum* [J]. Theoretical and Applied Genetics, 2003, 108(1): 79-86.
- [14] Ben-Chaim A, Borovsky Y, Falise M, et al. QTL analysis for capsaicinoid content in *Capsicum* [J]. Theoretical and Applied Genetics, 2006, 113(8): 1481-1490.
- [15] Arce-Rodríguez M L, Ochoa-Alejo N. An R2R3-MYB transcription factor regulates capsaicinoid biosynthesis [J]. Plant Physiology, 2017, 174(3): 1359-1370.
- [16] 张正海, 曹亚从, 于海龙, 等. 辣椒果实主要品质性状遗传和代谢物组成研究进展 [J]. 园艺学报, 2019, 46(9): 1825-1841.
- [17] Lee J, Park S J, Hong S C, et al. QTL mapping for capsaicin and dihydrocapsaicin content in a population of *Capsicum annuum* 'NB1' × *Capsicum chinense* 'Bhut Jolokia' [J]. Plant Breeding, 2016, 135(3): 376-383.
- [18] Han K, Lee H Y, Ro N Y, et al. QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum* [J]. Plant Biotechnology Journal, 2018, 16(9): 1546-1558.
- [19] 张正海, 毛胜利, 王立浩, 等. 辣椒的辣味遗传控制与辣椒素生物合成研究进展 [J]. 园艺学报, 2014, 41(9): 384-400.
- [20] Curry J, Aluru M, Mendoza M, et al. Transcripts for possible capsaicinoid biosynthetic genes are differentially accumulated in pungent and non-pungent *Capsicum* spp. [J]. Plant Science, 1999, 148(1): 47-57.
- [21] Aluru M R, Mazourek M, Landry L G, et al. Differential expression of fatty acid synthase genes, *AcL*, *Fat* and *Kas*, in *Capsicum* fruit [J]. Journal of Experimental Botany, 2003, 54(388): 1655-1664.
- [22] Stewart C Jr, Mazourek M, Stellari G M, et al. Genetic control of pungency in *C. chinense* via the *Pun1* locus [J]. Journal of Experimental Botany, 2007, 58(5): 979-991.
- [23] Mazourek M, Pujar A, Borovsky Y, et al. A dynamic interface for capsaicinoid systems biology [J]. Plant Physiology, 2009, 150(4): 1806-1821.
- [24] Liu Y, Zhang Z S, Fang K, et al. Genome-wide analysis of the MYB-related transcription factor family in pepper and functional studies of *CaMYB37* involvement in capsaicin biosynthesis [J]. International Journal of Molecular Sciences, 2022, 23(19): 11667.
- [25] Sun B M, Chen C M, Song J L, et al. The *Capsicum* MYB31 regulates capsaicinoid biosynthesis in the pepper pericarp [J]. Plant Physiology and Biochemistry, 2022, 176: 21-30.
- [26] Sun B M, Zhu Z S, Chen C J, et al. Jasmonate-inducible R2R3-MYB transcription factor regulates capsaicinoid biosynthesis and stamen development in *Capsicum* [J]. Journal of Agricultural and Food Chemistry, 2019, 67(39): 10891-10903.
- [27] Liu C, Ma N, Wang P Y, et al. Transcriptome sequencing and de novo analysis of a cytoplasmic male sterile line and its near-isogenic restorer line in chili pepper (*Capsicum annuum* L.) [J]. PLoS One, 2013, 8(6): e65209.
- [28] 刘峰, 谢玲玲, 弭宝彬, 等. 辣椒转录组 SNP 挖掘及多态性分析 [J]. 园艺学报, 2014, 41(2): 343-348.
- [29] 秦于玲, 申龙斌, 曹振木. 辣椒不同发育时期果实着色的转录组分析 [J]. 分子植物育种, 2020, 18(17): 5576-5583.
- [30] Wang P Y, Wang L R, Guo J J, et al. Molecular mapping of a gene conferring resistance to *Phytophthora capsici* Leonian race 2 in

王钰文, 钟 鑫, 范晓庆, 等. 基于 96K SNP 芯片的谷子抽穗期性状全基因组关联分析[J]. 江苏农业科学, 2025, 53(6): 33–47.
doi:10.15889/j.issn.1002-1302.2025.06.004

基于 96K SNP 芯片的谷子抽穗期性状全基因组关联分析

王钰文¹, 钟 鑫¹, 范晓庆¹, 刘颖慧¹, 赵治海², 范光宇², 王小明²

(1. 河北北方学院农林科技学院, 河北张家口 075000; 2. 张家口市农业科学院谷子研究所, 河北张家口 075000)

摘要:抽穗期是衡量农作物重要性的一个关键指标,它直接影响作物对地区和季节的适应能力。了解不同谷子品种的适宜种植区域,对于农业生产和实践具有极其重要的意义。为揭示控制抽穗期性状的遗传机制,为开展理想性状标记辅助选择育种奠定基础,调查了 230 个样本的抽穗期叶姿、刚毛长度、刚毛颜色、护颖颜色、花药颜色、茎秆长度、茎秆粗度、伸长节间数 8 个抽穗期性状并进行基因组重测序,拟以水稻抽穗期相关性状为研究对象,利用 R 软件中 GWASpoly 方法对水稻抽穗前后的相关性状进行多层次的相关性研究,利用基于 1-随机和正双(种群结构+血缘)的杂交 Q+K(种群结构+血缘)方法,分别构建分位数-分位数(QQ-Plot)和曼哈顿图。筛选与表型相关性显著的单核苷酸多态性位点(SNPs),在 10 kb 区间找到与表型相关的基因,并对这些 SNPs 进行功能注释。明确了每个性状与其具有显著关联性的染色体位置以及关联度,与抽穗期叶姿、刚毛长度、刚毛颜色、护颖颜色、花药颜色、茎秆长度、茎秆粗度、伸长节间数关联的 SNP 的关键位点在 9 条染色体上均有分布。

关键词:谷子;全基因组关联分析;SNP;抽穗期性状;候选基因

中图分类号:S188;S515.01 **文献标志码:**A **文章编号:**1002-1302(2025)06-0033-15

谷子(*Setaria italica*)是一种起源于我国的多年生草本植物,小米是谷子去壳后的产物。谷子是我国最主要的农作物之一,也是我国陆地生态系统的核心组成部分^[1],其抗旱节水能力极强,有望成为我国应对未来全球气候变化的战略性后备作物。笔者所在课题组前期已经完成了谷子 SSR 简单重复序列分子标记的检测^[2-3],为开展谷子分子育种研究打下了坚实的基础。

近年来,研究人员在基因组层面上开展了一系列对谷子的研究工作,包括利用重测序技术对重要农艺性状进行全基因组关联分析(GWAS)^[4-5];利用杂交分离群体对重要性状进行数量性状基因座(QTL)作图^[6-8],以及抗逆基因的克隆与功能验证^[9-10]等。谷子是短日喜温的农作物,对光照和温度敏感,具有显著的地域特征,是开展谷子生态适宜度评估的基础。通过对谷子的生态适应性评估,发掘出适应性广的谷子种质,对于培育跨地域适应性广的谷子新品种有重大的实际应用价值。抽穗期是反映作物生态适应能力的一个重要指标,还可以反映作物的生长状况,光照、温度等环境因素对其有显著影响^[11-12]。Gupta 等对 184 个谷子品种的产量性状进行了研究,并利用 50 对 SSR 分子标记对其进行遗传连锁,并将其定位在第 5 条染色体上^[13]。Jia 等对 916 种谷子品种进行重测序,与 47 个性状进行了关联分析^[4]。谷子的全基因组测序工作^[14-15]已全部完成,但其在全基因组范围内的相

收稿日期:2024-04-08

基金项目:国家现代农业产业技术体系建设专项(编号:CARs-06-14.5-A7);河北省重点研发计划(编号:21326302D);河北北方学院校级课题;河北宣化谷子科技小院项目。

作者简介:王钰文(1998—),男,山东潍坊人,硕士研究生,研究方向为作物栽培,E-mail:929660014@qq.com;共同第一作者:钟 鑫(1997—),女,河北承德人,硕士研究生,研究方向为分子植物育种,E-mail:1815708558@qq.com。

通信作者:刘颖慧,博士,教授,研究方向为作物栽培。E-mail:leely519@126.com。

pepper line PI201234 (*Capsicum annuum* L.) [J]. Molecular Breeding, 2016, 36(6): 66.

[31] 雷 阳, 成 妍, 乔 宁, 等. 辣椒苗期抗感疫病比较转录组学分析[J]. 华北农学报, 2019, 34(3): 194–202.

[32] 雷 阳, 成 妍, 乔 宁, 等. 辣椒苗期抗感黄瓜花叶病毒病比较转录组学分析[J]. 中国蔬菜, 2021(7): 24–31.

[33] 王红娟, 唐荣莉, 蒋晓英, 等. 辣椒不同发育时期果实的转录组分析[J]. 西南农业学报, 2022, 35(3): 517–525.

[34] 周 欣. 辣椒素合成结构基因及 MYB 转录因子的克隆和功能验证[D]. 广州: 华南农业大学, 2016: 44–49.

[35] 张 威, 张丽平, 吴 丹, 等. 黄灯笼辣椒 MYB4 转录因子的克隆与序列分析[J]. 分子植物育种, 2023, 21(6): 1801–1810.