

张高原, 魏兵强. 西瓜隐性核雄性不育相关基因筛选及定位[J]. 江苏农业学报, 2024, 40(6): 1089-1097.
doi: 10.3969/j.issn.1000-4440.2024.06.016

西瓜隐性核雄性不育相关基因筛选及定位

张高原, 魏兵强

(甘肃农业大学园艺学院, 甘肃 兰州 730070)

摘要: 为挖掘更多的西瓜花药、花粉发育相关基因, 解析其调控网络, 本研究以西瓜雄性不育植株和可育植株为材料, 构建遗传分离群体, 并利用 BSA-seq 和 RNA-seq 进行基因定位和育性相关基因筛选。结果表明, 雄性不育性状由单个隐性细胞核基因控制, 该基因被定位在 6 号染色体的 5 766 829 bp 至 21 366 400 bp 之间。结合 SNP/InDel 变异位点、基因注释及表达分析, 获得 1 个关键候选差异表达基因 *Cla97C06G117840* 以及受该基因表达影响的 20 个其他差异表达基因。生物信息学分析推测 *Cla97C06G117840* 由于部分编码序列缺失导致编码蛋白质提前终止, 进而引起花粉败育。此外 20 个其他差异表达基因被富集在花药壁绒毡层发育和花粉壁组件基因集通路中, 可能形成基因调控网络来影响花药、花粉发育。本研究为后期基因功能验证奠定了基础。

关键词: 西瓜; 雄性不育; 育性相关基因; 基因定位

中图分类号: S651 文献标识码: A 文章编号: 1000-4440(2024)06-1089-09

Screening and mapping of genes related to recessive nuclear male sterility in watermelon

ZHANG Gaoyuan, WEI Bingqiang

(College of Horticulture, Gansu Agricultural University, Lanzhou 730070, China)

Abstract: To explore more genes related to watermelon anther and pollen development and analyze its regulatory network, the genetic segregation population was constructed using watermelon male sterile materials and male fertile materials. And the gene mapping and fertility related gene screening were carried out by using BSA-seq and RNA-seq. The results showed that male sterility trait was controlled by a single recessive nuclear gene, which was mapped between 5 766 829 bp and 21 366 400 bp on chromosome 6. In combination with SNP/InDel variation site, gene annotation and expression analysis, one key candidate differentially expressed gene *Cla97C06G117840* and 20 other differentially expressed genes affected by the expression of this gene were obtained. Bioinformatics analysis suggested that the deletion of some coding sequences of *Cla97C06G117840* led to premature termination of the coding protein, resulting in pollen abortion. In addition, 20 other differentially expressed genes were enriched in the gene pathway of anther wall tapetum development and pollen wall assembly. These genes might form a gene regulatory network to affect anther and pollen development. This study can lay a foundation for later gene function verification.

Key words: watermelon; male sterility; fertility related genes; gene mapping

收稿日期: 2023-05-06

基金项目: 甘肃省教育厅创新基金项目(2022B-102); 甘肃农业大学省部共建干旱生境作物学国家重点实验室开放基金项目(GSCS-2020-09); 甘肃农业大学公开招聘博士科研启动基金项目(2017RCZX-30)

作者简介: 张高原(1987-), 男, 河南漯河人, 博士, 讲师, 研究方向为西瓜育种及分子生物学。(E-mail) zhanggy@gsau.edu.cn

西瓜(*Citrulus lanatas* L.) 是葫芦科的园艺作物, 因其果肉美味、营养丰富, 在世界各地广泛种植。西瓜杂种一代具有明显的杂种优势, 目前其杂交种多采用人工去雄套袋授粉获得, 不仅成本高而且难

以保证种子纯度,影响育种单位和生产者的利益。利用西瓜雄性不育系进行制种可有效解决以上问题,但是西瓜雄性不育品种选育比较滞后。挖掘雄性不育相关基因是西瓜雄性不育分子育种的基础。1962 年 Watts 报道了西瓜雄性不育植株以后,人们开始对西瓜雄性不育进行研究。细胞学研究结果显示西瓜雄性不育植株的绒毡层细胞多层,体积小,过早降解,无法为小孢子发育提供必需的营养物质,影响小孢子四分体的形成^[1-3]。抗氧化酶及激素研究结果显示在花药发育过程中,不育株抗氧化酶活性变化和内源激素含量异常可能与西瓜雄性不育的发生密切相关^[4]。遗传标记研究结果表明,利用 RAPD 技术对西瓜雄性不育株和可育株基因组 DNA 进行比较分析,确定了特异片段 A123k 与育性基因的遗传距离为 8.1 cM^[5]。POD 同工酶分析结果显示雄蕊中不育株较可育株多 2 条活性强的酶带^[6]。利用 AFLP 标记技术,结合 BSA 法,确定了特异片段 E31M50 与西瓜雄性不育基因的遗传距离为 5.0 cM^[7]。利用 mRNA 差异显示技术,在雄性不育花蕾与可育花蕾中获得了与育性相关的差异 cDNA 片段^[8-9]。转录组分析发现 1 259 个基因在西瓜雄性不育花蕾和可育花蕾中差异表达^[10]。基因定位和表达谱研究发现 *Cla006625* 和 *Cla001608* 是影响西瓜花药花粉发育的关键基因^[11-12]。

本研究前期鉴定出 1 个西瓜雄性不育材料,细胞学研究发现绒毡层细胞发育不正常是导致花粉败育的主要原因,并且在可育花蕾和不育花蕾中共发现 4 011 个差异表达基因^[13]。为了明确导致雄性不育的原因,本研究以雄性不育材料为母本,雄性可育材料为父本构建遗传分离群体并进行遗传规律分析,然后结合集团分离分析法(BSA-seq)以及转录组分析法(RNA-seq)进行育性相关基因筛选和关键候选基因定位,为后期基因功能验证奠定基础。

1 材料与方法

1.1 试验材料

以雄性不育材料 M0012 为母本(P_1),雄性可育材料 WF-2018 为父本(P_2)构建遗传群体(F_1 和 F_2)。2021 年 4 月 23 日,将 35 粒 P_1 、35 粒 P_2 、35 粒 F_1 和 400 粒 F_2 种子直播在甘肃省兰州市甘肃农业大学试验田中,田间常规管理,待植株雄花显现时进行育性调查。利用 SPSS17.0 中的卡方检验对西瓜雄性不育

性状的遗传分析进行适合度检验。从 F_2 分离群体中分别采集 40 株雄性不育单株叶片和 40 株雄性可育单株叶片构建子代雄性不育混合池(MS-pool)和雄性可育混合池(MF-pool),亲本 P_1 和 P_2 分别挑选 1 株并采集叶片构建亲本池,用于雄性不育相关基因定位。

1.2 试验方法

1.2.1 文库构建和测序 由上海欧易生物科技有限公司完成样本 DNA 的提取、建库和测序。采用 CTAB 法对 P_1 、 P_2 、MS-pool 和 MF-pool 的叶片进行 DNA 提取。利用 NanoDrop 核酸定量分析仪和 1% 琼脂糖凝胶电泳进行 DNA 质量检测。检测合格的 DNA 样品先经过 Covaris 超声波 DNA 破碎仪随机打断成 350~500 bp 的片段,然后利用 TruSeq DNA LT Sample Prep kit 试剂盒进行文库构建,文库构建合格后利用测序仪进行双端测序。

1.2.2 生物信息学分析 利用 Fastp 软件^[14]对测序获得的原始数据(Raw data)进行质量控制,然后利用 BWA 软件^[15]将过滤后数据(Clean reads)比对到西瓜参考基因组(<http://cucurbitgenomics.org/ftp/genome/watermelon/97103/v2/>)上,比对结果经过 SAMtools 软件^[16]转换格式后再利用 Picard 软件(<http://broadinstitute.github.io/picard>)去除冗余,并用 Qualimap 软件对比对结果进行分析。使用 GATK4 软件^[17]的 Haplotypecaller 模块进行 SNP 和 InDel 检测。为降低 SNP 和 InDel 检测的错误率,选用 QD (经过深度校正的质量值) ≥ 2.0 的标准进行过滤,只保留满足该条件的突变位点。利用 Annotvar 软件^[18]对 SNP 和 InDel 检测结果进行注释。

1.2.3 SNP-index 计算及候选基因筛选 筛选亲本间纯合差异的 SNP 位点。根据亲本的基因型,对子代池进行 SNP 检测确定突变亲本来源的 Read,计算子代池的 SNP-index。同时将 2 个子代池的 SNP-index 相减获得 Δ SNP-index。利用二项分布检验 2 个子代池的覆盖深度,考虑其混池数量以及群体类型,算出其统计学上 95% 的和 99% 的置信线,当拟合线超出置信线的染色体区域为可能的表型连锁区域。然后在超出置信线的染色体区域内筛选候选基因。

1.2.4 候选基因表达及功能注释 利用前期研究发表的转录组数据,分析候选基因在不育和可育花蕾中的表达情况,筛选定位区间内的差异表达基因,并利用基迪奥云平台在线分析软件(<https://www.>

omicsshare.com/tools) 进行基因本体 (Gene Ontology; GO) 功能注释。利用 Orthofinder 软件进行同源基因分析^[19]。利用 Exon-Intron Graphic Maker 在线软件 (<http://wormweb.org/exonintron>) 进行基因结构绘制。基因表达由 R 语言中 Pheatmap 和 Ggpolt2 包进行绘制。*Clu97C06G117840* 基因上游引物为 5'-ATGTTTATTTGTTTCTTTGCTCTTTTC-3', 下游引物为 5'-TTAATAGGTATTGGATGTGCTGCTG-3', 基因测序由擎科生物技术有限公司完成。

1.2.5 GSEA 富集分析及蛋白质互作预测 利用前期转录组数据,以定位基因表达模式为筛选标准线,计算其他基因与该基因的表达相关性,以西瓜基因 GO 功能注释为背景文件,然后利用 R 语言进行基因集富集分析 (Gene Set Enrichment Analysis, GSEA),筛选重要富集通路中的调控基因。蛋白质互作预测由 STRING 在线软件 (<https://cn.string-db.org/>) 完成。利用 Cytoscape 软件^[20] 进行蛋白质互作网络图绘制。

2 结果与分析

2.1 雄性不育表型和遗传规律分析

前期表型研究发现雄性可育花蕾花药饱满,开花散粉正常,而不育花蕾花药瘦小,呈空瘪状,并且细胞

学研究结果显示花药绒毡层细胞发育异常是导致花粉败育的主要原因^[13]。遗传群体育性调查发现 F₁ 全为可育株,说明不育性状是由隐性基因控制。F₂ 分离群体中可育单株 290 个,不育单株 101 个,经卡方检验,该比例符合 3 : 1 的遗传规律,说明雄性不育性状是由单个隐性细胞核基因控制 (表 1)。

表 1 表型数据统计和遗传规律分析

Table 1 Phenotypic data statistics and genetic law analysis						
群体	可育	不育	期望比例	实际比例	卡方(χ^2)	P 值
P ₁	-	30	-	-		
P ₂	30	-	-	-		
F ₁	30	-	-	-		
F ₂	290	101	3.0 : 1.0	2.9 : 1.0	0.123	0.726

2.2 测序质量评估

本试验利用高通量测序技术对 2 个亲本池和 2 个子代混合池进行重测序。结果如表 2 所示,一共产生 71.79 G 的原始数据量,通过质量控制过滤后获得 71.15 G 的高质量 Clean read 数据,有效数据比率达 99.80% 以上,G+C 含量为 34.79%~34.89%,Q20 含量>95.70%,Q30 含量>87.44%,说明测序数据可靠合格,可用于下一步分析。

表 2 测序数据质量统计分析

Table 2 Statistical analysis of sequencing data quality

样本	原始序列数	过滤后序列数	序列有效率 (%)	原始测序数据量 (G)	过滤后数据量 (G)	测序数据有效率 (%)	G+C 含量 (%)	大于 Q20 含量 (%)	大于 Q30 含量 (%)
P ₁	112 176 306	111 975 900	99.82	16.83	16.62	98.80	34.82	95.70	87.44
P ₂	105 566 046	105 358 236	99.80	15.83	15.69	99.11	34.89	96.52	89.69
MS-pool	126 307 544	126 111 734	99.84	18.95	18.79	99.19	34.79	96.39	89.27
MF-pool	134 529 570	134 359 072	99.87	20.18	20.05	99.38	34.84	96.34	89.21

MS-pool:子代雄性不育混合池;MF-pool:子代雄性可育混合池。

与参考基因组比对结果如表 3 所示,4 个样本比率为 99.35%~99.48%,平均覆盖深度为 39.515 0~49.183 2,平均比对质量>47,1X 覆盖度>98.09%,10X 覆盖度>96.36%,说明过滤后的数据与参考基因组的比率和覆盖度高,可以用于下一步 SNP 和 InDel 变异检测。

2.3 SNP 和 InDel 变异位点检测和注释

SNP 注释结果显示 4 个样本池中平均检测注释到 26 521 个 SNP 变异位点,其中子代池 MS-pool 最多,为 30 900 个,亲本池 P₁ 最少,为 16 063 个;共注释

了 10 种 SNP 类型,其中编码基因区域内非同义 SNP 突变位点最多,4 个样本平均检测注释到 2 428 个,且子代池 MF-pool 中最多,为 2 835,亲本池 P₁ 中最少,为 1 471 (表 4)。InDel 注释结果显示 4 个样本池中平均检测注释到 26 052 个 InDel 变异位点,其中子代池 MS-pool 中最多,为 28 384 个,亲本池 P₁ 中最少,为 20 715 个。共注释了 14 种 InDel 类型,其中编码基因区域内移码缺失突变最多,4 个样本平均检测注释到 155 个,且亲本池 P₂ 中最多,为 172 个,亲本池 P₁ 中最少,为 109 个 (表 5)。

表 3 质量控制后测序数据与参考基因组比对结果

Table 3 Comparison results of sequencing data and reference genome after quality control

样本	质量控制后 总序列数	比对上参考基因 组上的 Reads 数	比对率 (%)	平均覆盖深度	平均比对质量	1X 覆盖度 (%)	10X 覆盖度 (%)
P ₁	111 975 900	102 979 422	99.35	41.181 4	47.495 1	98.19	96.59
P ₂	105 358 236	98 291 528	99.45	39.515 0	47.888 7	98.09	96.36
MS-pool	126 111 734	116 875 633	99.48	47.041 7	48.177 5	98.23	97.07
MF-pool	134 359 072	122 068 930	99.37	49.183 2	47.979 6	98.24	97.10

MS-pool:子代雄性不育混合池;MF-pool:子代雄性可育混合池;1X 覆盖度:参考基因组中至少有 1 条 Reads 覆盖的位点数占基因组全长的百分比;10X 覆盖度:参考基因组中至少有 10 条 Reads 覆盖的位点数占基因组全长的百分比。

表 4 SNP 变异位点注释统计

Table 4 Annotation statistics of SNP mutation sites

基因区域	突变类型	P ₁	P ₂	MS-pool	MF-pool	平均
基因下游		4 464	7 980	8 738	8 706	7 472
编码区	非同义突变	1 471	2 575	2 829	2 835	2 428
	终止子获得	40	76	71	73	65
	终止子丢失	12	25	25	24	22
	同义突变	1 257	2 321	2 568	2 567	2 178
	未知	3	10	11	11	9
其他		2 783	5 007	5 504	5 510	4 701
内含子		3 556	5 963	6 587	6 585	5 673
剪切位点		18	31	31	30	28
基因上游		5 242	9 247	10 040	10 065	8 649
总数		16 063	28 228	30 900	30 896	26 522

MS-pool:子代雄性不育混合池;MF-pool:子代雄性可育混合池。

表 5 InDel 变异位点注释统计

Table 5 Annotation statistics of InDel mutation sites

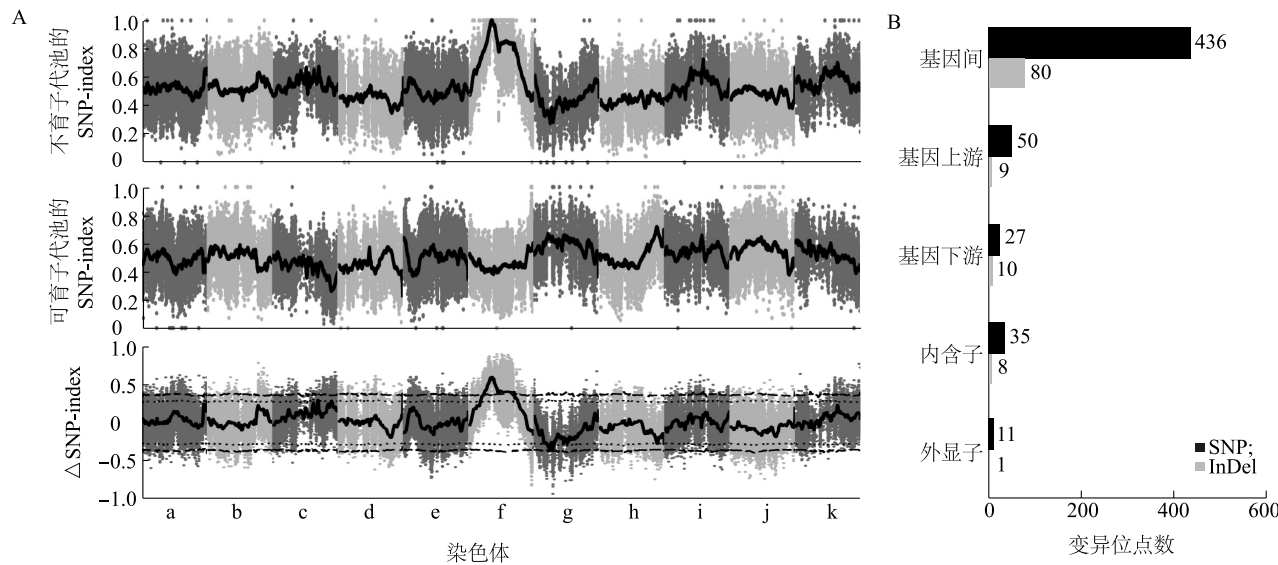
基因区域	突变类型	P ₁	P ₂	MS-pool	MF-pool	平均
5'非翻译区		2	3	2	3	3
基因下游		6 727	8 885	9 363	9 288	8 566
编码区	移码缺失	109	172	169	169	155
	移码插入	100	160	155	154	142
	移码替换	2	2	1	3	2
	非移码缺失	74	142	142	136	124
	非移码插入	91	160	160	160	143
	终止子获得	26	36	33	37	33
	终止子丢失	5	6	5	8	6
	未知	0	1	1	1	1
其他		407	679	666	668	605
内含子		5 389	6 634	7 024	7 077	6 531
剪切位点		26	38	39	42	36
基因上游		8 164	10 574	11 290	11 219	10 312
总数	-	20 715	26 813	28 384	28 297	26 052

MS-pool:子代雄性不育混合池;MF-pool:子代雄性可育混合池。

2.4 连锁区域分析

利用 Δ SNP-index进行基因定位,结果显示与雄性不育表型紧密连锁区域位于6号染色体的5 766 829~21 366 400 bp(图1A)。该区域在4个样本池中共发现559个SNP变异位点和108个InDel变异位点,其中436个SNP和80个InDel变异

位点位于基因间,50个SNP和9个InDel变异位点位于基因上游,27个SNP和10个InDel变异位点位于基因下游,35个SNP和8个InDel变异位点位于内含子处,只有11个SNP和1个InDel变异位点位于外显子(图1B)。另外,位于SNP和InDel变异位点附近的候选基因共有136个。



a:1号染色体;b:2号染色体;c:3号染色体;d:4号染色体;e:5号染色体;f:6号染色体;g:7号染色体;h:8号染色体;i:9号染色体;j:10号染色体;k:11号染色体。

图1 作图群体连锁定位图和 SNP/InDel 统计分析
Fig.1 Mapping population linkage mapping and SNP/InDel statistical analysis

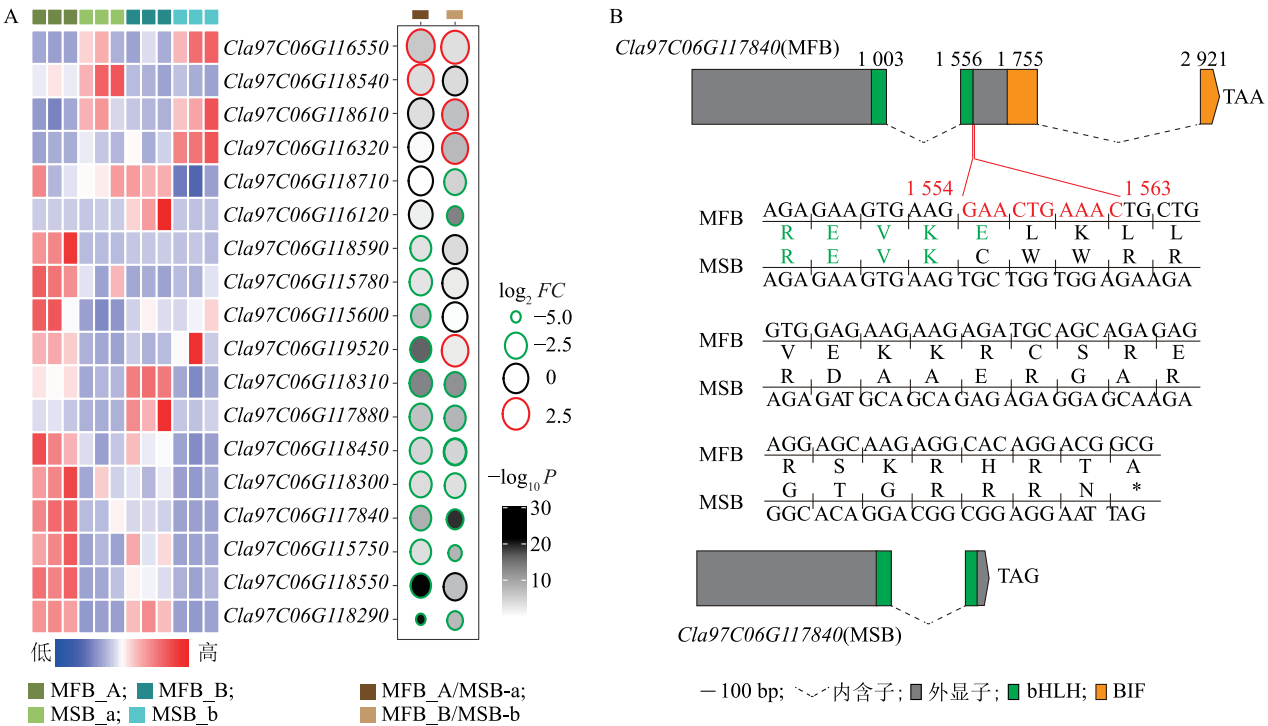
2.5 候选基因筛选

利用前期已发表的转录组数据,对136个基因进行表达分析,发现只有18个基因在可育花蕾(MFB_A和MFB_B)和不育花蕾(MSB_a和MSB_b)之间存在差异表达。在对比组(MFB_A/MSB_a)中,2个基因在不育花蕾中明显上调表达,而12个基因显著下调表达。在对比组(MFB_B/MSB_b)中,4个基因在不育花蕾中明显上调表达,而9个基因显著下调表达(图2A)。SNP、InDel变异位点分析发现,17个差异表达基因的SNP或InDel变异位点分别位于基因间、基因上游、基因下游和内含子处,只有1个差异表达基因(*Cla97C06G117840*,编码bHLH转录因子)的InDel变异位点位于外显子处,而且该基因被注释到花药壁绒毡层发育(Anther wall tapetum development)、程序性细胞死亡的正调控(Positive regulation of programmed cell death)、细胞壁胼胝质沉积(Callose deposition in cell wall)等代谢调控通路上(表6)。与可育花蕾对比,

*Cla97C06G117840*基因在不育花蕾中下调表达,其编码序列在1 554~1 563 bp处缺失10个碱基(GAACTGAAAC),其中GAA位于bHLH结构域区域;由于碱基缺失导致编码基因发生移码,提早出现终止子(TAG),造成无法编码BIF结构域序列(图3B)。另外,该基因是拟南芥*bHLH91*(*AT2G31210*)同源基因,研究结果表明该基因参与调控拟南芥花药及花粉发育^[21]。因此,*Cla97C06G117840*基因被推测为导致西瓜雄性不育的关键基因。

2.6 关联基因表达及富集通路分析

GSEA分析结果如图3A所示,与*Cla97C06G117840*基因表达正相关的基因集的前5个富集通路主要为花粉壁组件(GO:0010208,富集分数为0.75)、细胞外基质组件(GO:0085029,富集分数为0.75)、乙酰辅酶A代谢过程(GO:0006084,富集分数为0.73)、花药壁绒毡层发育(GO:0048658,富集分数为0.72)和花粉外壁形成(GO:0010584,富集分数为0.71)。



MFB: 雄性可育花蕾; MSB: 雄性不育花蕾; MFB_A: 3~5 mm 雄性可育花蕾; MSB_a: 3~5 mm 雄性不育花蕾; MFB_B: 7~9 mm 雄性可育花蕾; MSB_b: 7~9 mm 雄性不育花蕾。

图 2 候选基因表达趋势及筛选鉴定
Fig.2 Expression trend and screening identification of candidate genes

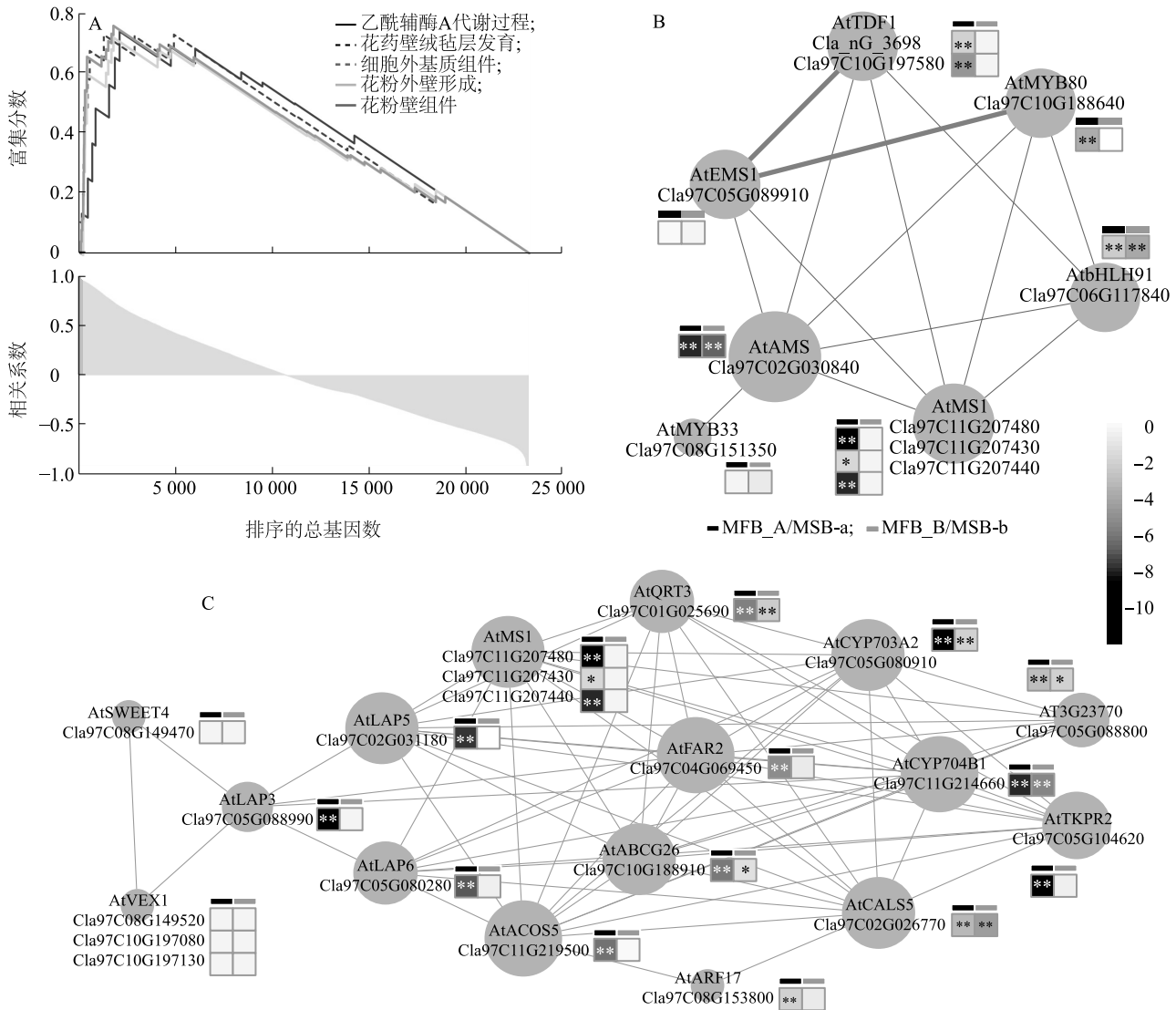
表 6 候选基因变异位点及注释

Table 6 Variation sites and annotations of candidate genes

基因	变异位置	位点类型	基因注释	基因本体注释
Cla97C06G116550	基因间	SNP	乙烯响应转录因子	芳香族化合物生物合成过程; 有机环状化合物生物合成过程; 转录调控
Cla97C06G118540	内含子	SNP	Rac 家族 GTP 结合蛋白	无
Cla97C06G118610	基因间	SNP/InDel	bHLH 转录因子	无
Cla97C06G116320	基因上游、内含子	SNP/InDel	含 <i>p</i> -环核苷三磷酸水解酶超家族蛋白	无
Cla97C06G118710	基因间	SNP/InDel	类蛋白激酶	无
Cla97C06G116120	基因间	SNP	丝氨酸/苏氨酸蛋白激酶	细胞内信号传导; 细胞通讯; 低温响应
Cla97C06G118590	基因下游	SNP	树皮贮藏蛋白 A	无
Cla97C06G115780	基因间	SNP/InDel	热休克蛋白	无
Cla97C06G115600	基因间	SNP/InDel	类卤酸脱卤酶家族蛋白	无
Cla97C06G119520	基因间	SNP/InDel	乙醛脱氢酶	芳香族化合物生物合成过程; 有机环状化合物生物合成过程; 细胞生物合成过程
Cla97C06G118310	基因间、基因下游	SNP/InDel	糖转运蛋白	<i>D</i> -核糖转运; 山梨醇转运; 戊糖转运
Cla97C06G117880	内含子	SNP	非特异性磷脂酶 C	宿主细胞成分共生体破坏; 宿主结构共生体修饰; 发病机制
Cla97C06G118450	基因间、基因上游	SNP/InDel	动力蛋白轻链	无
Cla97C06G118300	基因间、基因下游、基因上游	SNP/InDel	转录抑制因子 OFP8	转录负调控; RNA 生物合成过程负调控; DNA 转录负调控
Cla97C06G117840	外显子	InDel	bHLH 转录因子	花药壁绒毡层发育; 正调控程序性细胞死亡; 细胞壁胼胝质沉积
Cla97C06G115750	基因间、基因下游	SNP	转录抑制因子 OFP12	转录负调控; RNA 生物合成过程负调控; DNA 转录负调控
Cla97C06G118550	基因间、基因下游	SNP	过氧化物酶 2B	镉离子响应; 金属离子响应; 胁迫响应
Cla97C06G118290	基因间、基因下游	SNP/InDel	ATP 酶活性依赖的 RNA 解旋酶	无

花药壁绒毡层发育基因集通路中包含 14 个基因,其中 11 个基因包含在领头亚基中(图 3A)。蛋白质互作预测结果显示 10 个基因可能存在蛋白质互作,其中 8 个基因在 MFB-A/MSB-a 对比组均差异表达。例如 *Cla97C02G030840* 与拟南芥 *AMS* (*AT2G16910*) 同源,下调 88.7%,与 *bHLH91*、*MYB33*、*EMS1*、*MYB80*、*TDF1* 和 *MS1* 可能存在蛋白质互作;*Cla97C06G117840* 与拟南芥

bHLH91 同源,下调 67.7%,与 *MYB80*、*TDF1* 和 *MS1* 可能存在蛋白质互作;*Cla97C11G207430*、*Cla97C11G207440* 和 *Cla97C11G207480* 与拟南芥 *MS1* (*AT5G22260*) 同源,分别下调 61.5%、88.6%和 92.1%,与 *TDF1* 和 *MYB80* 可能存在蛋白质互作(图 3B)。因此,推测这 8 个基因在花药绒毡层细胞发育中起着重要的调控作用。



MFB:雄性可育花蕾;MSB:雄性不育花蕾;MFB_A:3~5 mm 雄性可育花蕾;MSB_a:3~5 mm 雄性不育花蕾;MFB_B:7~9 mm 雄性可育花蕾;MSB_b:7~9 mm 雄性不育花蕾。

图3 GSEA 富集分析及蛋白质互作预测
Fig.3 GSEA enrichment analysis and protein interaction prediction

花粉壁组件基因集通路中包含 38 个基因,其中 22 个基因包含在领头亚基中(图 3A)。蛋白质互作预测显示 20 个基因可能存在蛋白质互作,其中 16 个基因在 MFB-A/MSB-a 对比组均差异表达。例如

Cla97C10G188910 与拟南芥 *ABCG26* (*AT3G13220*) 同源,下调 85.3%,与 *CYP703A2*、*LAP6*、*ACOS5*、*TKPR2*、*CYP704B1*、*CALS5*、*FAR2*、*QRT3*、*LAP5*、*AT3G23770* 和 *MS1* 可能存在蛋白质互作;*Cla97C02G031180* 与拟南芥

LAP5(*AT4G34850*)同源,下调 86.2%,与 *TKPR2*、*QRT3* 和 *MS1* 可能存在蛋白质互作;*Cla97C04G069450* 与拟南芥 *FAR2*(*AT3G11980*)同源,下调 83.6%,与 *LAP6*、*TKPR2*、*LAP3*、*QRT3*、*LAP5* 和 *MS1* 可能存在蛋白质互作;*Cla97C05G088990* 与拟南芥 *LAP3*(*AT3G59530*)同源,下调 92.2%倍,与 *LAP5*、*LAP6* 和 *SWEET4* 可能存在蛋白质互作(图 3C)。因此,推测这 16 个基因在花粉壁发育方面起着重要的调控作用。

3 讨论

虽然 *Cla006625* 和 *Cla001608* 基因已被报道影响西瓜花药、花粉发育^[11-12],但其他的相关基因仍需挖掘。本研究利用 BSA-seq 方法在西瓜 6 号染色体的 5 766 829~21 366 400 bp 初步定位到 1 个雄性不育性状相关的关键候选基因(*Cla97C06G117840*),并结合 RNA-seq 筛选出受该基因表达影响的 20 个其他差异表达基因。通过生物信息学分析推测这些基因在西瓜花药、花粉发育过程中可能具有重要的调控作用。

拟南芥 *AtbHLH91* 与 *AtbHLH10* (*AT2G31220*) 和 *AtbHLH89* (*AT1G06170*) 存在功能冗余,可以通过 BIF 基序与 DYT1 互作,是 DYT1 激活花药相关基因正常表达所必需的^[21]。*Cla97C06G117840* 是 *AtbHLH91* 同源基因,但该基因在西瓜中无旁系同源基因,无功能冗余现象,在雄性不育植株中发现该基因的编码序列缺失 10 bp,发生移码突变,提早出现终止子,导致 BIF 基序缺失。因此,推测该基因由于 BIF 基序缺失,无法激活下游花药相关基因的正常表达,最终导致花药发育异常和花粉败育。

在其他差异表达基因中,7 个差异表达基因被富集在花药壁绒毡层发育基因集通路中,分别与拟南芥 *TDF1*、*AMS*、*MYB80* 和 *MS1* 同源。研究结果显示这些基因可构成调控网络(*TDF1*-*AMS*-*MYB80*-*MS1*),影响绒毡层细胞发育。其中 *TDF1* 编码 R2R3-MYB 转录因子,*tdfl* 突变株中绒毡层细胞异常增大并液泡化,导致花粉败育^[22]。*AMS* 编码 bHLH 转录因子,其突变体中绒毡层发育不正常而导致小孢子母细胞减数分裂失败,随后在四分体时期小孢子降解^[23]。*MYB80* 编码 R2R3-MYB 转录因子,其突变株出现绒毡层细胞过早降解,导致小孢子败育^[24]。*MS1* 编码 PHD-finger 核蛋白,突变后导致绒毡层细胞分泌物合成及分泌异常,花粉外壁结构

改变,引起植物雄性不育^[25]。因此,推测这 7 个差异表达基因可能构成调控网络,在西瓜花药绒毡层细胞发育过程中起着重要调控作用。

此外,13 个差异表达基因被富集在花粉壁组件基因集通路中,分别与拟南芥 *CALS5*、*ARF17*、*LAP3*、*ACOS5*、*LAP5*、*FAR2*、*LAP6*、*CYP703A2*、*CYP704B1*、*TKPR2*、*ABCG26*、*QRT3* 和 *AT3G23770* 同源。研究结果显示这些基因在胼胝质合成、孢粉素前体形成等方面发挥着重要作用,基因突变后造成花粉壁发育异常,导致花粉败育^[26]。其中 *CALS5* 编码胼胝质合成酶,其突变后导致花粉外壁外层缺失^[27]。*ARF17* 为生长素反应因子,可以调控 *CALS5* 基因表达,影响胼胝质合成^[28]。*LAP3* 属于钙依赖性磷酸三酯酶家族成员,其突变后导致花粉外壁形成异常^[29]。8 个基因可构成调控网络(*FAR2*-*ACOS5*-*CYP703A2*/*CYP704B1*-*LAP5*/*LAP6*-*TKPR2*-*ABCG26*),在孢粉素前体形成过程中发挥着关键的调控作用,基因突变后会导致孢粉素沉积受阻,进而引起花粉败育^[30]。*QRT3* 编码聚半乳糖醛酸酶,基因突变后导致花粉母细胞壁降解缺陷,造成花粉发育后期小孢子分离失败^[31]。因此推测这 13 个基因在西瓜花粉壁发育过程中可能起着重要调控作用。

4 结论

通过 BSA-seq 方法在西瓜 6 号染色体的 5 766 829~21 366 400 bp 初步定位到 1 个雄性不育性状相关的关键候选基因(*Cla97C06G117840*),推测该基因由于编码序列部分缺失,提前出现终止子,造成 BIF 基序缺失,导致下游花药相关基因表达异常,最终引起花粉败育。此外结合 RNA-seq 方法筛选出受该基因表达影响的 20 个其他差异表达基因,通过生物信息学分析发现这些基因可能存在互作关系,形成基因调控网络,在西瓜花药、花粉发育中可能起着重要调控作用。本研究结果将为后期西瓜雄性不育相关基因的功能研究奠定了基础。

参考文献:

- [1] 张显,王鸣,马建祥,等. 西瓜隐性核型雄性不育两用系的细胞学研究[J]. 西北农林科技大学学报(自然科学版),2005,33(1):71-74.
- [2] 陈雨,张建农,刘炬. 西瓜雄性不育花芽显微结构观察[J]. 甘肃农业大学学报,2008,43(5):67-70.
- [3] 马建祥,张显,张勇,等. 西瓜雄性不育两用系雄蕊电镜扫

- 描观察[J]. 北方园艺, 2011, (3): 151-152.
- [4] 王永琦, 杨小振, 莫言玲, 等. 西瓜雄性不育系‘Se18’抗氧化酶活性和内源激素含量变化分析[J]. 园艺学报, 2016, 43(11): 2161-2172.
- [5] 刘海河, 侯喜林, 张彦萍. 西瓜核雄性不育育性基因的 RAPD 标记[J]. 果树学报, 2004, 21(5): 491-493.
- [6] 马建祥, 张显, 张勇, 等. 西瓜核型雄性不育两用系 *POD* 同工酶研究[J]. 中国瓜菜, 2005(4): 30-32.
- [7] 郭守鹏, 刘海河, 张彦萍, 等. 西瓜核雄性不育两用系 G17AB 育性基因的 AFLP 分子标记[J]. 园艺学报, 2009, 36(3): 427-430.
- [8] 张彦萍, 刘海河, 谢彬, 等. 西瓜细胞核雄性不育系雄花蕾的 mRNA 差异表达分析[J]. 果树学报, 2010, 27(6): 1037-1041.
- [9] 于远, 张显, 张勇, 等. mRNA 差异显示分离西瓜核雄性不育基因的相关 cDNA 片段[J]. 西北农林科技大学学报(自然科学版), 2008, 36(11): 162-166.
- [10] RHEE S J, SEO M, JANG Y J, et al. Transcriptome profiling of differentially expressed genes in floral buds and flowers of male sterile and fertile lines in watermelon[J]. BMC Genomics, 2015, 16: 914.
- [11] DONG W, WU D, YAN C, et al. Mapping and analysis of a novel genic male sterility gene in watermelon (*Citrullus lanatus*) [J]. Frontiers in Plant Science, 2021, 12: 639431.
- [12] ZHU Y, YUAN G, JIA S, et al. Transcriptomic profiling of watermelon (*Citrullus lanatus*) provides insights into male flowers development[J]. Journal of Integrative Agriculture, 2022, 21(2): 407-421.
- [13] ZHANG G, DING Q, WEI B, et al. Transcriptome analysis of sterile and fertile floral buds from recessive genetic male sterility lines in watermelon (*Citrullus lanatus* L.) [J]. Journal of Plant Growth Regulation, 2023, 42: 2800-2812.
- [14] CHEN S, ZHOU Y, CHEN Y, et al. Fastp: an ultra-fast all-in-one FASTQ preprocessor[J]. Bioinformatics, 2018, 34(17): 884-890.
- [15] LI H, DURBIN R. Fast and accurate short read alignment with Burrows-Wheeler transform [J]. Bioinformatics, 2009, 25(14): 1754-1760.
- [16] LI H, HANDSAKER B, WYSOKER A, et al. The sequence alignment/map format and SAMtools [J]. Bioinformatics, 2009, 25(16): 2078-2079.
- [17] MCKENNA A, HANNA M, BANKS E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data[J]. Genome Research, 2010, 20(9): 1297-1303.
- [18] WANG K, LI M, HAKONARSON H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data [J]. Nucleic Acids Research, 2010, 38(16): e164.
- [19] EMMS D M, KELLY S. OrthoFinder: phylogenetic orthology inference for comparative genomics [J]. Genome Biology, 2019, 20(1): 238.
- [20] SHANNON P, MARKIEL A, OZIER O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks[J]. Genome Research, 2003, 13(11): 2498-2504.
- [21] FU Y, LI M, ZHANG S, et al. Analyses of functional conservation and divergence reveal requirement of bHLH010/089/091 for pollen development at elevated temperature in *Arabidopsis* [J]. Journal of Genetics and Genomics, 2020, 47(8): 477-492.
- [22] ZHU J, CHEN H, LI H, et al. Defective in tapetal development and function 1 is essential for anther development and tapetal function for microspore maturation in *Arabidopsis* [J]. The Plant Journal, 2008, 55(2): 266-277.
- [23] FERGUSON A C, PEARCE S, BAND L R, et al. Biphasic regulation of the transcription factor ABORTED MICROSPORES (AMS) is essential for tapetum and pollen development in *Arabidopsis* [J]. The New Phytologist, 2017, 213(2): 778-790.
- [24] PHAN H A, IACUONE S, LI S F, et al. The MYB80 transcription factor is required for pollen development and the regulation of tapetal programmed cell death in *Arabidopsis thaliana* [J]. Plant Cell, 2011, 23: 2209-2224.
- [25] YANG C, VIZCAY-BARRENA G, CONNER K, et al. MALE STERILITY1 is required for tapetal development and pollen wall biosynthesis [J]. Plant Cell, 2007, 19: 3530-3548.
- [26] 马伯军, 刘程捷, 陈析丰, 等. 花粉壁发育相关基因的研究进展 [J]. 浙江师范大学学报(自然科学版), 2013, 36(4): 438-443.
- [27] DONG X, HONG Z, SIVARAMAKRISHNAN M, et al. Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in *Arabidopsis* [J]. The Plant Journal, 2005, 42(3): 315-328.
- [28] YANG J, TIAN L, SUN M X, et al. AUXIN RESPONSE FACTOR17 is essential for pollen wall pattern formation in *Arabidopsis* [J]. Plant Physiology, 2013, 162(2): 720-731.
- [29] DOBRITSA A A, NISHIKAWA S, PREUSS D, et al. LAP3, a novel plant protein required for pollen development, is essential for proper exine formation [J]. Sex Plant Reproduction, 2009, 22(3): 167-177.
- [30] QUILICHINI T D, GRIENENBERGER E, DOUGLAS C J. The biosynthesis, composition and assembly of the outer pollen wall: a tough case to crack [J]. Phytochemistry, 2015, 113: 170-182.
- [31] RHEE S Y, OSBORNE E, POINDEXTER P D, et al. Microspore separation in the quartet 3 mutants of *Arabidopsis* is impaired by a defect in a developmentally regulated polygalacturonase required for pollen mother cell wall degradation [J]. Plant Physiology, 2003, 133(3): 1170-1180.

(责任编辑: 成纾寒)