中国油料作物学报
Chinese Journal of Oil Crop Sciences
2026,48(5):114-122
doi:10.19802/j.issn.1007-9084.2026026

基于BSA-seq技术定位花生早花相关基因位点

王芳[1],吕娇艳[1],邢晓宁[1],王海莉[1],任锐[2],芦连勇[1],王基国[1],王艳[1],张少聪[1],沈希华[1*]
(1.安阳市农业科学院,河南 安阳,455000;
2.河南农业大学农学院,河南 郑州,450046)

摘要:为解析花生开花期调控的分子机制并筛选相关的候选基因,助力花生早花的分子育种,本研究以花生早花与晚花表型F_2分离群体为材料,构建极端性状混池,采用BSA-seq技术开展早花性状关联位点定位。结果显示,在4条染色体的相应区间内共鉴定到20个与早花性状显著关联的区间。这些染色体和区间包括A06染色体的23.50~98.50 Mb、A07染色体的62.00~63.00 Mb、B05染色体的14.00~14.10 Mb,和B08染色体的13.05~13.25 Mb区间。通过GO功能注释、KEGG通路富集分析及数据库多重比对,初步筛选出10个潜在开花相关候选基因;结合转录组表达数据进一步验证,最终确定2个候选基因*Arahy.M2THPA*与*Arahy.UD4HAV*。

关键词:花生;初花期;BSA-seq;基因定位
中图分类号:　　**文献标识码**:A　　**文章编号**:1007-9084(2026)05-0114-09

Localization of peanut genes associated with early flowering based on BSA-seq technology

WANG Fang[1], LYU Jiao-yan[1], XING Xiao-ning[1], WANG Hai-li[1], REN Rui[2], LU Lian-yong[1], WANG Ji-guo[1], WANG Yan[1], ZHANG Shao-cong[1], SHEN Xi-hua[1*]
(1. *Anyang Academy of Agriculture Sciences, Anyang* 455000, *China*; 2. *College of Agronomy, Henan Agricultural University, Zhengzhou* 450046, *China*)

Abstract: To elucidate molecular mechanism underlying flowering time regulation in peanut (*Arachis hypogaea* L.), and to screen candidate genes associated with early flowering for breeding on early-flowering peanut, this study took an F_2 segregating population with contrasting early and late flowering phenotypes in peanut as material, and constructed extreme trait bulks. Authors adopted Bulked Segregant Analysis sequencing (BSA-seq) to map the loci associated with early-flowering trait. Results showed that a total of 20 intervals significantly associated with early flowering were identified in 4 genomic regions, including 23.50-98.50 Mb on chromosome A06, 62.00-63.00 Mb on chromosome A07, 14.00-14.10 Mb on chromosome B05, and 13.05-13.25 Mb on chromosome B08. 10 potential flowering-related candidate genes were preliminarily screened *via* Gene Ontology (GO) functional annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and multiple database alignment. Further validation combined with transcriptomic expression data ultimately identified 2 candidate genes: *Arahy.M2THPA* and *Arahy.UD4HAV*.

Key words: peanut; initial flowering days; BSA-seq; gene mapping

花生(*Arachis hypogaea* L.)作为我国重要的油料、经济和植物蛋白来源作物,在农业生产体系与国民经济中占据举足轻重的地位。我国花生种植面积广泛,常年种植面积稳定在470万公顷以上,总产量突破1900万吨,不仅是食用植物油的核心原料之一,其加工制品远销国内外,创造了显著的经济收益[1]。近年来极端气候事件频发,麦田套作、玉米-花生间作等花生种植模式的应用面积持续扩增;而在我国东北、辽宁及新疆北部等无霜期较短的区域,花生生育前期易遭遇低温冷害,后期常受

收稿日期:2026-01-28
基金项目:2025年河南省中央引导地方科技发展资金(Z20251831174)
作者简介:王芳(1988-),女,助理研究员,硕士,主要从事花生育种与栽培研究,E-mail: wfwr515107@163.com
* 通讯作者:沈希华(1971-),男,副研究员,主要从事花生新品种选育研究及技术推广工作,E-mail: aysxh@163.com

高温干旱与早霜天气胁迫,上述生产问题均迫切需要早熟花生品种的支撑,以保障花生生产的优质与高产[2]。

花生早熟的调控机制复杂,受遗传因素与环境条件的共同调控[3,4]。花生早熟性的鉴定可根据初花期、开花集中程度、籽仁灌浆速度、生育后期植株落叶率、荚果内外壳性状(内果壳色泽、外果壳网纹明晰度)等指标判断[5]。其中,初花期直接决定花生生育期长短,并与产量密切相关,是评价花生早熟性的一个重要指标[6]。

花生基因组测序工作的完成与持续完善,为花生早花相关基因的挖掘提供了高效技术支撑。已有研究通过数量性状位点(QTL)定位技术,初步定位到与花生初花期[7,8]、开花天数[9]、开花集中度[10]及开花量[11]相关的QTL位点;另外,通过对野生花生与栽培花生开展全基因组PEBP基因家族分析,筛选出多个与花生开花相关的候选基因[12,13]。但目前关于花生早花分子遗传机制的研究仍较为薄弱,难以满足分子标记辅助育种在花生早花育种中的实际应用需求,因此揭示花生早花的分子遗传机制,对花生生育期改良和早熟育种具有重要意义。

本研究选取安花5号和豫航花1号为亲本材料。安花5号是安阳市农业科学院自主培育的早熟品种,兼具早花表型显著、适应性广等优势,是研究花生早花性状的理想材料;豫航花1号为常规品种,农艺性状优良,可与安花5号形成明显表型差异,为分离群体构建提供可靠对照。利用这两种材料构建F_2分离群体,并对群体中的极端材料进行集群分离重测序分析(BSA-seq),检测与花生初花期紧密关联的基因区间,系统挖掘调控早花性状的候选基因,研究结果可为解析花生早花分子调控机制、加快早花优质花生品种选育进程提供重要理论支撑与技术参考。

1 材料与方法

1.1 材料

2022年,以早花型花生品种安花5号为母本、普通花生品种豫航花1号为父本进行杂交,获得F_1杂交种子。2023年,在河南省安阳市崔家桥镇牛庄村完成F_1的田间种植,2024年进一步种植获得由686个单株组成的F_2遗传分离群体。F_2种植设置为行长2m、行距40 cm、穴距16.7 cm,每穴播2粒种子,整个生育期的田间栽培管理均按照花生育种试验的常规技术要求实施。

1.2 方法

1.2.1 *表型数据处理* 对F_2分离群体的初花期性状开展调查,初花期定义为从播种至植株首次开花的天数。采用SPSS 27.0软件对调查数据进行统计分析,计算初花期天数的均值、标准差、峰度及偏度指标。

1.2.2 *DNA混池构建与测序* 从F_2分离群体中,分别筛选初花期极短和极长的单株各30株,组建极早花和极晚花两个极端性状集合。采集两个集合中各单株的新鲜花生叶片,提取基因组DNA后按等浓度比例混合,分别构建子代早花极端混池与晚花极端混池;同时采用单株取样策略,采集2个亲本新鲜叶片提取DNA。样本混池的文库构建及测序工作由杭州联川生物技术股份有限公司完成,采用NovaSeq 6000测序平台(Illumina公司)对4个混池DNA进行高通量测序,子代混池测序深度设定为30×。为保障后续分析的准确性,对测序获得的原始序列进行质量过滤,得到高质量有效测序数据,并将其比对至花生(*Arachis hypogaea*)参考基因组Tifrunner.gnm1.KYV3进行后续分析。

1.2.3 *欧氏距离(Euclidean distance, ED)算法和SNP-Index方法* ED算法通过分析测序数据筛选混池间存在显著差异的分子标记,进而定位与目标性状相关联的基因组区域,其ED值越高,表明对应标记在两个混池间的差异越显著。该方法的计算公式如下:

$$ED=\sqrt{(A_{mut}-A_{wt})^2+(C_{mut}-C_{wt})^2+(G_{mut}-G_{wt})^2+(T_{mut}-T_{wt})^2}$$

式中,A_{mut}、C_{mut}、G_{mut}、T_{mut}分别代表早花混池中SNP位点上碱基A、C、G、T的检测频率;A_{wt}、C_{mut}、G_{mut}、T_{mut}分别代表晚花混池中对应碱基的检测频率。基于两混池间基因型的SNP位点,统计各碱基在不同混池中的测序深度,计算每个位点的原始ED值;为剔除背景噪音干扰,对原始ED值进行3次方转换,再采用滑窗降噪法,依据窗口内均值拟合得到降噪后的ED值。筛选亲本间纯合差异SNP位点,以早花亲本安花5号为参考对照,计算两个子代混池在亲本差异标记位点的SNP-index(即SNP频率)及ΔSNP-index,具体计算公式如下:

$$SNP\text{-}index(a/b)=M_{a/b}/(M_{a/b}+P_{a/b})$$

$$\Delta SNP\text{-}index=SNP\text{-}index(a)-SNP\text{-}index(b)$$

式中,$M_{a/b}$表示a/b混池中来源于早花亲本的测序深度,$P_{a/b}$表示a/b混池中来源于晚花亲本的测序深度;a代表早花性状混池,b代表晚花性状混池。以1 Mb为窗口、1 kb为步长,对ΔSNP-index在各染

色体上的分布进行平滑作图,设定99%置信水平作为筛选阈值,将该阈值以上的窗口区域界定为与初花期性状相关联的候选区域。

1.2.4 候选基因筛选 以花生栽培种Tifrunner参考基因组2.0版本的注释信息为依据,借助联川生物OmicStudio云平台(https://www.omicstudio.cn/tool.),对落在候选关联区域内的基因进行GO功能富集分析、KEGG通路分析及可视化绘图,明确候选基因的功能及参与的代谢通路。花生转录组数据来源于PeanutBase数据库(https://www.peanutbase.org/expression/expr_tissue_Hyp.html),数据中各组织的采集时期描述和荚果发育时期依据来自前人的研究,其中花生籽仁成熟度划分为15个等级,该数据采集了15个等级中前10个发育时期的荚果[14]。对转录组数据进行$\log_2$转换及Z-score标准化处理后,利用联川生物云平台OmicStudio绘制基因表达量热图,直观呈现候选基因的表达差异特征。

2 结果与分析

2.1 表型数据分析

对F_2分离群体686株花生的初花期天数表型数据进行统计分析,结果显示(表1),该性状均值为28.66 d,表型变异范围为23~34 d;其分布的峰度为-0.141、偏度为-0.155,二者绝对值均小于1,分布形态接近正态,结合频次分布特征分析(图1),表明花生初花期天数为典型的数量性状,且表型数据符合正态分布特征。

2.2 全基因组测序数据统计及多态性位点检测

对早花和晚花的2个亲本池和2个子代极端混池进行测序,得到原始数据量分别为458 874 652、471 299 456、512 778 702、493 892 692 reads,过滤后得到的数据量为455 663 252、467 321 524、508 162 974、489 638 072 reads。原始数据质量大于Q20的比例在98%以上,大于Q30的比例在94%,表明测序质量高,GC含量在37.38%~38.10%,GC分布正常,建库测序质量合格。比对到参考基因组上的总reads数目占比98.78%~99.90%,两个亲本平均覆盖深度分别为26.44、28.44,早花基因池和晚花基因池分别为29.46、28.44。

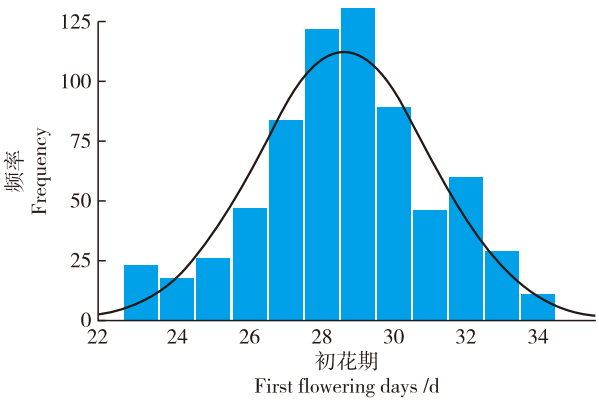


图1 F_2群体初花期分布特征

Fig. 1 Distribution characteristics of the number of initial flowering days in F_2 population

对原始数据的SNP和InDel进行检测和注释,检测到554,734个的SNP,其中20249个位于基因上游区域,13914个位于外显子区域,外显子区域中引起非同义突变的有8670个。检测到141089个InDel,其中12775个位于基因上游区域,1798个位于外显子区域,外显子区域中引起移码突变的有1325个。

2.3 BSA-seq关联分析

在BSA-seq关联分析中常采用ED和SNP-index两种分析方法,这两种方法关联到的区间的交集作为早花相关的候选区间。ED计算方法是基于两混池SNP位点,统计各碱基混池深度并计算原始ED值;为消除背景噪声,对原始ED值取3次方后,经滑窗均值拟合获得降噪后ED值,关联值分布如图2A所示,在99%置信区间内,A04、A07、B05、B07染色体上各关联到1个候选区域,在B01、B08染色体上各关联到2个候选区域,在A06染色体上关联到20个候选区间,共计28个区域,覆盖基因组长度39.5 Mb。SNP-index方法分析结果如图2B,通过执行1000次置换检验,在99%置信区间内鉴定出33个候选区间,分布在A06、A07、A09、B05、B06、B08染色体,覆盖基因组长度41.5 Mb。取ED和SNP-index分析所得候选区间的交集,获得花生初花期相

表1 F_2群体初花期天数分布特征

Table 1 Distribution characteristics of initial flowering days in F_2 population

亲本均值 Parental mean /d		F_2群体初花天数 F_2 population /d						
安花5号 Anhua 5	豫航花1号 Yuhuanghua 1	平均值 Mean	标准差 SD	方差 Variance	偏度 kewness	峰度 Kurtosis	范围 Range	变异系数 CV /%
25	31	28.66	2.43	5.91	-0.155	-0.141	23-34	8.48

表 2 双亲及混池全基因组测序数据统计

Table 2 Statistical data of parental and mixed pool whole genome sequencing data

样本 Sample	测序原始数据 Raw reads	过滤后数据 Valid reads	GC 含量 GC count /%	Q20 /%	Q30 /%	对比率 Mapped /%	测序深度(×) Mean coverage
安花5号 Anhua 5	458 874 652	455 663 252	38.10	98.29	95.08	99.90%	26.44
豫航花1号 Yuhanghua 1	471 299 456	467 321 524	37.44	98.22	94.97	99.90%	27.19
早花基因池 Early-flowering gene pool	512 778 702	508 162 974	37.38	98.14	94.72	99.78%	29.46
晚花基因池 Late-flowering gene pool	493 892 692	489 638 072	37.51	98.22	94.95	99.82%	28.44

关的候选区间(表3)。这些候选区间分布在4条染色体上,A06染色体23.50~98.50 Mb区域包含16个候选区间,A07染色体62.00~63.00 Mb与B05染色体14.00~14.10 Mb区域各含1个候选区间,B08染色体13.05~13.25 Mb区域包含2个候选区间;共计20个候选区间,总长度为27 Mb,包含439个基因。

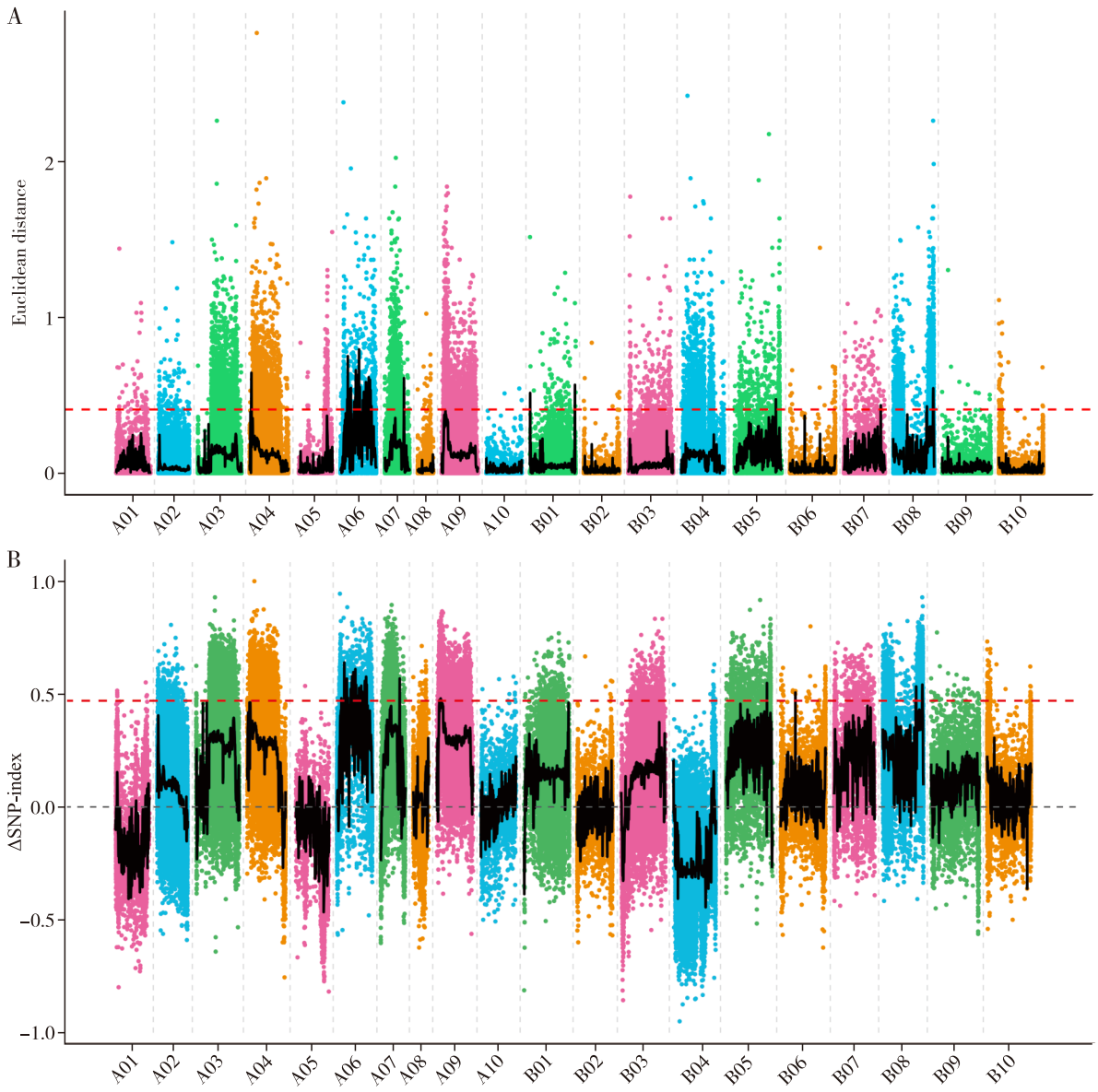
2.4 候选基因筛选

2.4.1 富集分析 通过BSA-seq关联分析获得的27 Mb候选区间中,共有5716个SNP变异,其中上游基因变异有165个位点,下游基因变异的有73个位点,非同义突变的有46个位点,这些变异位点发生在150个候选基因当中。对这150个候选基因进行了GO功能和KEGG通路富集分析(图3)。GO功能富集分析发现,共122个基因被注释到GO功能条目中。在生物过程部分,主要涉及细胞增殖(GO:0006261 DNA复制、GO:0007049 细胞周期调控等)、代谢调控(GO:0006571 酪氨酸生物合成过程、GO:0009813 黄酮类生物合成过程等)、应激响应(GO:0042742 对细菌的防御反应、GO:0009611 对创伤的反应等)、信号传导(GO:0009585 红光、远红光光转

表 3 ED和SNP-index两种算法关联候选区域

Table 3 Correlation of candidate regions between the ED and SNP-index algorithms

序号 No.	染色体 Chromosome	起始位置 Start location	结束位置 End location	区间大小 Region size /Mb	基因数目 Nu Mber of gene
1	A06	23 500 000	24 500 000	1.00	20
2	A06	26 500 000	27 500 000	1.00	8
3	A06	31 500 000	33 000 000	1.50	5
4	A06	44 500 000	45 500 000	1.00	7
5	A06	48 000 000	52 000 000	4.00	19
6	A06	53 500000	54 500 000	1.00	9
7	A06	59 500 000	61 000 000	1.50	4
8	A06	62 000 000	63 500 000	1.50	5
9	A06	72 500 000	73 500 000	1.00	13
10	A06	75 500 000	76 500 000	1.00	8
11	A06	79 000 000	80 500 000	1.50	19
12	A06	82 000 000	83 000 000	1.00	8
13	A06	85 000 000	85 500 000	0.50	18
14	A06	91 000 000	92 500 000	1.50	40
15	A06	94 500 000	96 500 000	2.00	33
16	A06	97 500 000	98 500 000	1.00	25
17	A07	62 000 000	63 000 000	1.00	24
18	B05	140 000 000	141 000 000	1.00	36
19	B08	110 500 000	111 500 000	1.00	18
20	B08	130 500 000	132 500 000	2.00	120
总计Total				27.00	439



注：A:ED 分析法;B:ΔSNP-index 分析法;黑色曲线为拟合值,红线为 99% 置信线

Note: A: ED analysis method; B: ΔSNP-index analysis method; The black curve represents the fitted value, and the red line is the 99% confidence line

图 2 BSA-seq 关联分析图

Fig. 2 BSA-seq association analysis chart

导、GO:0006468 蛋白质磷酸化等)等生物学过程,说明候选基因在细胞增殖、物质代谢等基础生命活动中发挥核心作用。在细胞成分部分,富集到注释基因最多的种类是细胞核(GO:0005634)和细胞膜(GO:0005886 细胞质膜、GO:0016021 膜)。在分子功能部分,主要富集于酶活性(GO:0016301 激酶活性、GO:0004674 蛋白丝氨酸/苏氨酸激酶活性、GO:0004665 预苯酸脱氢酶(NADP+)活性等)、结合功能(如 GO:0005515 蛋白质结合、GO:0003677 DNA 结合、GO:0005524 ATP 结合等),反映候选基因通过催化或结合等分子层面作用参与花生开花的过程。

KEGG 通路富集分析结果如图 3B,图中显示了 Q 值最小的前 20 个通路,显著富集于代谢通路,具体有苯丙氨酸、酪氨酸、色氨酸生物合成(ko00400)、丙酸代谢(ko00640)、N-糖基化生物合成(ko00510)、赖氨酸分解(ko00310)、色氨酸代谢(ko00380)、萜类骨架生物合成(ko00900)、丙酮酸代谢(ko00620)、氮代谢(ko00910)、黄酮类生物合成(ko00941)。此外,与植物的昼夜节律(ko04712)相关的基因也表现出明显富集。这些通路的富集表明候选基因可能通过调控代谢活动和参与植物的昼夜节律,在花生早花性状的形成中发挥关键作用。

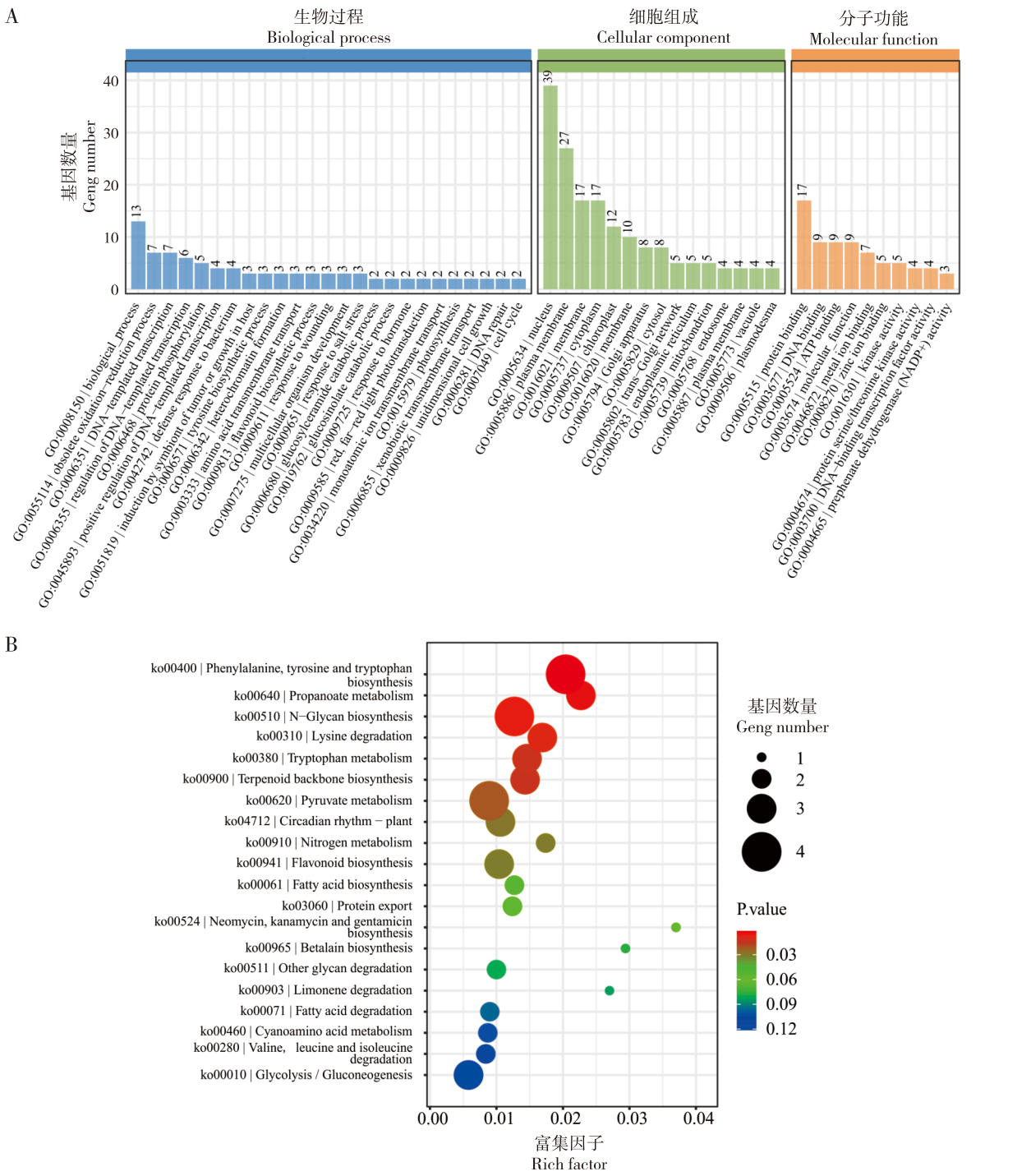


图3 候选基因的GO功能富集分析(A)和KEGG富集分析(B)

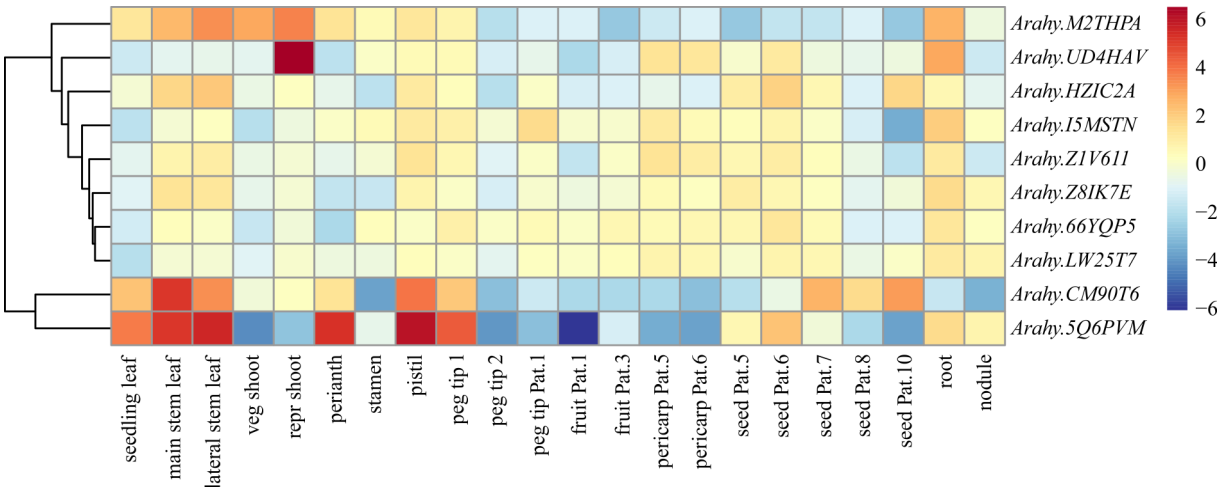
Fig. 3 GO functional enrichment analysis (A) and KEGG enrichment analysis (B) of candidate genes

2.4.2 候选基因注释和转录组数据分析 利用花生基因组数据库对150个候选基因进行注释,结合GO和KEGG富集分析的结果,筛选出10个与开花相关的候选基因(表4)。根据花生转录组数据,获得这10个候选基因在花生不同生长时期和不同部位的表达量数据,并以此绘制表达量热图(图4)。表达量分析发现2个核心候选基因:*Arahy. M2THPA*、*Arahy. UD4HAV*。*Arahy. M2THPA*在主干叶片、第一侧枝叶片、营养枝茎尖、花枝茎尖中均有表达,且表达量呈现逐步升高的趋势,同时在花被、雄蕊、雌蕊和地上果针顶端中也持续表达;*Arahy. UD4HAV*则在花枝茎尖中呈特异性高表达,这些独特的表达模式表明,上述基因可能参与了花生开花相关的生物过程。

表4 候选基因及注释

Table 4 Candidate genes and annotations

序号 No.	基因名 Gene	基因注释 Gene annotation
1	*Arahy.66YQP5*	COP9 信号体复合物亚基 COP9 signalosome complex subunit
2	*Arahy.I5MSTN*	可能的 LRR 受体样丝氨酸/苏氨酸蛋白激酶 Probable LRR receptor-like serine/threonine-protein kinase
3	*Arahy.CM90T6*	查尔酮合酶 Chalcone synthase
4	*Arahy.Z1V611*	可能的无活性亮氨酸重复受体样蛋白激酶 Probable inactive leucine-rich repeat receptor-like protein kinase
5	*Arahy.M2THPA*	开花基因位点T蛋白 Protein FLOWERING LOCUS T
6	*Arahy.UD4HAV*	查尔酮/二苯乙烯合酶 Chalcone/stilbene synthase
7	*Arahy.Z8IK7E*	U结构域含蛋白质 U-box domain-containing protein
8	*Arahy.LW25T7*	细胞周期蛋白依赖性激酶 E-1 Cyclin-dependent kinase E-1
9	*Arahy.5Q6PVM*	UDP-糖基转移酶 UDP-glycosyltransferase
10	*Arahy.HZIC2A*	包含 LOB结构域的蛋白质 12 LOB domain-containing protein 12



注:seeding leaf:幼苗叶片;main stem leaf:初花期主干叶片;lateral stem leaf:初花期侧枝叶;veg shoot:初花期营养枝茎尖;repr shoot:初花期花枝茎尖;perianth:花被;stamen:雄蕊;pistil:雌蕊;peg tip 1:地上果针顶端;peg tip 2:地下果针顶端;peg tip Pat. 1:Pattee 1 时期果针顶端;fruit Pat. 1:Pattee 1 时期荚果;fruit Pat. 3:Pattee 3 时期荚果;pericarp Pat.5:Pattee 5 时期果皮;pericarp Pat.6:Pattee 6 时期果皮;seed Pat. 5:Pattee 5 时期种子;seed Pat. 6:Pattee 6 时期种子;seed Pat. 7:Pattee 7 时期种子;seed Pat. 8:Pattee 8 时期种子;seed Pat. 10:Pattee 10 时期种子;root:根;nodule:根瘤。色阶标注的颜色代表基因的相对表达量,颜色偏红,表达量越高,颜色偏蓝,表达量越低

Note: seeding leaf: seedling leaf; main stem leaf: main stem leaf at early flowering stage; lateral stem leaf: lateral stem leaf at early flowering stage; veg shoot: vegetative shoot apex at early flowering stage; repr shoot: reproductive shoot apex at early flowering stage; perianth: perianth; stamen: stamen; pistil: pistil ; peg tip 1: tip of aerial peg; peg tip: tip of subterranean peg; peg tip Pat. 1: peg tip at Pattee 1; fruit Pat. 1: Pattee 1 pod ; fruit Pat. 3: Pattee 3 fruit; pericarp Pat. 5: Pattee 5 pericarp; pericarp Pat. 6: Pattee 6 pericarp; seed Pat. 5: Pattee 5 seed; seed Pat. 6: Pattee 6 seed; seed Pat. 7: Pattee 7 seed; seed Pat. 8: Pattee 8 seed; seed Pat. 10: Pattee 10 seed; root: root; nodule: nodule. Colors in the scale indicate the relative expression levels of genes. Redder color represents higher expression level, while bluer color represents lower expression level

图4 候选基因表达量热图

Fig. 4 Heat map of the expression levels of candidate genes

3 讨论与结论

近年来,国内外学者围绕作物开花期调控机制已开展了大量研究[15-18],但相较于其他作物,花生开花期相关的系统性研究报道仍较为匮乏。BSA-seq可以快速定位控制目标性状基因或QTL位点,具有周期短、精度高、性价比高等优点[19],已在花生的QTL定位和功能基因挖掘上得到广泛应用[8,20-22]。本研究基于BSA-seq技术鉴定了花生早花相关的候选区间,结合GO功能注释、KEGG通路富集分析和转录组数据联合挖掘,最终筛选出2个与花生开花期相关的候选基因,为解析花生开花调控机制提供了新的基因资源。

花生初花期属于复杂的数量性状,其遗传调控受基因型和环境因素的共同影响。目前已有部分花生开花期相关的QTL定位研究报道:Wang等[7]利用花生重组近交系群体开展初花期QTL定位分析,通过构建遗传连锁图谱,鉴定出3个位于B04、B07和B08染色体上的主效QTL,分别命名为qIFDB04.1、qIFDB07.1和qIFDB08.1;Wang等[10]结合SLAF-seq(特异位点扩增片段测序)与SSR(简单序列重复)技术,构建栽培花生的高密度遗传连锁图谱,并鉴定到15个与开花期集中度相关的QTL,其中一个主要QTL位于6号染色体;杨军军[11]通过BSA-seq技术将早花性状初步定位在15号染色体155 200 000 bp~158 900 000 bp之间的3.7 Mb区间内。本研究综合BSA-seq关联分析中ED和SNP-index两种分析方法,在A06、A07、B05、B08染色体上共定位到20个候选区间,总长度为27 Mb。对比上述先前报道的QTL数据,未发现完全匹配的定位结果,推测一方面由于早期研究中部分QTL的物理位置信息尚未明确,缺乏精准的区间比对依据;另一方面可能与所用试验材料的遗传背景差异、遗传群体类型不同有关,导致定位结果存在物种内的遗传分化特征。

对BSA-seq关联区间内的基因进行GO注释与KEGG通路富集分析,富集到与开花相关的过程有红光、远红光光转导(GO:0009585)、蛋白质磷酸化(GO:0006468)、黄酮类生物合成过程(GO:0009813)、植物的昼夜节律(ko04712)、类黄酮生物合成(ko00941)、苯丙氨酸、酪氨酸、色氨酸生物合成(ko00400)。基于基因功能注释信息,对上述通路中的候选基因进行筛选,获得10个潜在与开花相关的基因;结合转录组表达量热图分析,进一步锁定2个核心候选基因:*Arahy. M2THPA*与*Arahy. UD4HAV*。*Arahy. M2THPA*属于栽培花生PE BP(磷脂酰乙醇胺结合蛋白)基因家族的FT亚家族(FLOWERING LOCUS T-like)的关键成员,表达量分析显示,该基因在早花基因型(四粒红、白花5号)和晚花基因型(花育22号、Tifrunner)中的表达水平存在显著差异:早花基因型在开花关键期的表达量是晚花基因型的2~3倍,且表达峰值出现时间更早,直接支持其促进花生开花的功能[13]。将候选基因*Arahy. UD4HAV*的序列提交至InterPro数据库(https://www.ebi.ac.uk/interpro/)进行Blast比对分析,结果显示该基因同时含有查尔酮/二苯乙烯合成酶的N端结构域(PF00195)与C端结构域(PF02797),推测*Arahy. UD4HAV*属于植物特异性III型聚酮合成酶(type III polyketide synthases, PKSs)超家族。该家族包含查尔酮合成酶(CHS)和二苯乙烯合成酶(STS),两类酶共享相同底物,分别是类黄酮与二苯乙烯类化合物生物合成的分支关键酶[23,24]。已有研究在栽培花生中鉴定出6个CHS和46个STS基因[25,26],但*Arahy. UD4HAV*并没有在其中,为尚未报道的*CHS/STS*家族基因。此外,该基因在花生花枝茎尖中特异性高量表达,推测*Arahy. UD4HAV*可能参与花生开花进程调控。

本研究通过构建花生早花与晚花表型F_2分离群体,利用BSA-seq技术精准定位花生早花相关遗传位点,最终在A06、A07、B05、B08染色体上鉴定到20个候选区间,总长度达27 Mb。结合GO功能注释、KEGG通路富集分析及数据库多重比对,初步筛选出10个潜在开花相关候选基因;进一步整合转录组表达数据筛选,最终确定2个核心候选基因,分别为PEBP家族FT亚家族开花促进关键基因*Arahy. M2THPA*,以及III型聚酮合成酶超家族的查尔酮/二苯乙烯合成酶编码基因*Arahy. UD4HAV*。本研究结果为解析花生开花期调控的分子机制提供了关键基因资源,同时为后续花生早花性状的分子育种改良提供了重要理论支撑与技术参考。

参考文献:

[1] 国家统计局. 2025中国统计年鉴[M]. 北京:中国统计出版社,2025.

[2] 王亮,朱金成,王睿,等. 我国花生早熟育种及产量性状基因定位研究进展[J]. 核农学报,2025,39(6):1178-1188.

[3] Rowland D L, Sorensen R B, Butts C L, et al. Determination of maturity and degree day indices and their success in predicting peanut Maturity [J]. Peanut Sci,

2006, 33(2): 125-136. DOI: 10.3146/0095-3679(2006)33[125: domadd]2.0.co;2.

[4] Vishnuprabha R S, Viswanathan P, Manonmani S, et al. Development of early maturing varieties in groundnut (*Arachis hypogaea* L.)[J]. Legume Res Int J, 2022. DOI:10.18805/lr-4784.

[5] 王传堂,张建成. 花生遗传改良[M]. 上海:上海科学技术出版社,2013.

[6] Kaba J S, Kumaga F K, Ofori K. Effect of flower production and time of flowering on pod yield of peanut (*Arachis hypogaea* L.) genotypes[J]. IOSR J Agric Vet Sci, 2014, 7(4): 44-49. DOI:10.9790/2380-07434449.

[7] Wang L, Yang X L, Cui S L, et al. Identification of main effect and epistatic QTLs controlling initial flowering date in cultivated peanut (*Arachis hypogaea* L.)[J]. J Integr Agric, 2020, 19(10): 2383-2393. DOI: 10.1016/s2095-3119(20)63211-7.

[8] Fonceka D, Tossim H A, Rivallan R, et al. Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding[J]. BMC Plant Biol, 2012, 12(1): 26. DOI:10.1186/1471-2229-12-26.

[9] Khedikar Y, Pandey M K, Sujay V, et al. Identification of main effect and epistatic quantitative trait loci for morphological and yield-related traits in peanut (*Arachis hypogaea* L.)[J]. Mol Breed, 2017, 38(1): 7. DOI: 10.1007/s11032-017-0764-z.

[10] Wang L, Yang X L, Cui S L, et al. High-density genetic map development and QTL mapping for concentration degree of floret flowering date in cultivated peanut (*Arachis hypogaea* L.)[J]. Mol Breed, 2020, 40(2): 17. DOI:10.1007/s11032-019-1083-3.

[11] 杨军军. 花生优异突变体创制与熟性相关基因初定位[D]. 吉林:吉林农业大学,2023:40-41.

[12] Jin H Q, Tang X M, Xing M G, et al. Molecular and transcriptional characterization of phosphatidyl ethanolamine-binding proteins in wild peanuts *Arachis duranensis* and *Arachis ipaensis*[J]. BMC Plant Biol, 2019, 19(1): 484. DOI:10.1186/s12870-019-2113-3.

[13] Zhong C, Li Z, Cheng Y L, et al. Comparative genomic and expression analysis insight into evolutionary characteristics of *PEBP* genes in cultivated peanuts and their roles in floral induction[J]. Int J Mol Sci, 2022, 23(20): 12429. DOI:10.3390/ijms232012429.

[14] Clevenger J, Chu Y, Scheffler B, et al. A developmental transcriptome map for allotetraploid *Arachis hypogaea*[J]. Front Plant Sci, 2016, 7: 1446. DOI: 10.3389/fpls.2016.01446.

[15] Srikanth A, Schmid M. Regulation of flowering time: all roads lead to Rome[J]. Cell Mol Life Sci, 2011, 68(12): 2013-2037. DOI:10.1007/s00018-011-0673-y.

[16] Komeda Y. Genetic regulation of time to flower in *Arabidopsis thaliana*[J]. Annu Rev Plant Biol, 2004, 55: 521-535. DOI:10.1146/annurev.arplant.55.031903.141644.

[17] Yoshida H, Nagato Y. Flower development in rice[J]. J Exp Bot, 2011, 62(14): 4719-4730. DOI:10.1093/jxb/err272.

[18] Lin X Y, Liu B H, Weller J L, et al. Molecular mechanisms for the photoperiodic regulation of flowering in soybean[J]. J Integr Plant Biol, 2021, 63(6): 981-994. DOI:10.1111/jipb.13021.

[19] Wang H R, Zhang H Z, Chen Y, et al. Mapping of CaUGT75C1 involved in the formation of pepper (*Capsicum annuum* L.) lateral branches *via* BSA[J]. Plant Physiol Biochem, 2025, 229: 110834. DOI: 10.1016/j.plaphy.2025.110834.

[20] 李玉颖,于海洋,吕玉英,等. 基于BSA重测序定位花生含油量相关基因位点[J]. 山东农业科学,2021,53(1):1-6.

[21] Zhao M B, Sun Z Q, Qi F Y, et al. Identification of two QTLs for web blotch resistance in peanut (*Arachis hypogaea* L.) based on BSA-seq[J]. BMC Plant Biol, 2024, 24(1): 1193. DOI:10.1186/s12870-024-05930-8.

[22] Li L, Cui S L, Dang P, et al. GWAS and bulked segregant analysis reveal the Loci controlling growth habit-related traits in cultivated peanut (*Arachis hypogaea* L.)[J]. BMC Genom, 2022, 23(1): 403. DOI: 10.1186/s12864-022-08640-3.

[23] Schröder J, Raiber S, Berger T, et al. Plant polyketide synthases: a chalcone synthase-type enzyme which performs a condensation reaction with methylmalonyl-CoA in the biosynthesis of C-methylated chalcones, [J]. Biochemistry, 1998, 37(23): 8417-8425. DOI:10.1021/bi980204g.

[24] Okada Y, Sano Y, Kaneko T, et al. Enzymatic reactions by five chalcone synthase homologs from hop (*Humulus lupulus*L.)[J]. Biosci Biotechnol Biochem, 2004, 68(5): 1142-1145. DOI:10.1271/bbb.68.1142.

[25] Brasileiro A C M, Gimenes M A, Pereira B M, et al. The stilbene synthase family in *Arachis*: a genome-wide study and functional characterization in response to stress[J]. Genes, 2023, 14(12): 2181. DOI: 10.3390/genes14122181.

[26] Qayyum Z, Noureen F, Khan M, et al. Identification and expression analysis of stilbene synthase genes in *Arachis hypogaea* in response to methyl jasmonate and salicylic acid induction[J]. Plants, 2022, 11(13): 1776. DOI:10.3390/plants11131776.

(责任编辑:郭学兰)