作物学报 ACTA AGRONOMICA SINICA 2026, 52(5): 1341−1351 http://zwxb.chinacrops.org/
ISSN 0496-3490; CN 11-1809/S; CODEN TSHPA9 E-mail: zwxb301@caas.cn

DOI: 10.3724/SP.J.1006.2026.53070

基于 RNA-Seq 和 BSA-Seq 联合分析挖掘玉米侧根密度基因资源

韩亚鑫[1,2] 何冠华[2,*] 张小琼[1,2] 张登峰[2] 李永祥[2] 刘旭洋[2] 王天宇[2] 黎 裕[2] 邹华文[1,*] 李春辉[2,*]

[1] 长江大学农学院, 湖北荆州 434025; [2] 作物基因资源与育种全国重点实验室 / 中国农业科学院作物科学研究所, 北京 100081

摘 要: 根系作为玉米植株吸收水分和营养成分的重要器官, 对玉米生长发育及产量形成至关重要。因此研究玉米根系结构的遗传机制对玉米高产育种实践具有重要意义。本研究利用玉米自交系丹黄 02 (侧根密度高)和 HRB16041 (侧根密度低)构建了一个 F_2 群体, 利用极端性状混池测序法(BSA-seq)获得控制侧根密度(LRD)的显著关联位点, 结合 8 个玉米自交系侧根组织的转录组测序, 挖掘控制玉米苗期侧根密度的候选基因。基于侧根密度的 BSA-seq 分析, 研究共检测到 240 个显著关联的 SNP 和 2420 个显著关联的 InDel 位点, 其中, 位于基因上下游区域或影响可变剪接位点和基因编码蛋白的变异位点共涉及 235 个候选基因。通过侧根组织 RNA-seq 分析, 共鉴定出与侧根密度相关的差异表达基因 1979 个, 其中, 954 个基因上调表达, 1025 个基因下调表达。BSA-seq 和 RNA-seq 分析共同鉴定到了 12 个基因, 这些基因被确定为控制侧根密度的候选基因。通过对这 12 个候选基因进行 qRT-PCR 验证, 确定 *Zm00001d033708* 和 *Zm00001d045720* 等候选基因的基因表达和玉米侧根密度呈正相关, 而 *Zm00001d040375* 基因表达与玉米侧根密度呈负相关。本研究为玉米侧根密度研究提供了候选基因资源。

关键词: 玉米; 侧根密度; 混池测序; 转录组分析; 表达量

Identification of maize lateral root density genes resources through integrated RNA-seq and BSA-seq analyses

Han Ya-Xin[1,2], He Guan-Hua[2,*], Zhang Xiao-Qiong[1,2], Zhang Deng-Feng[2], Li Yong-Xiang[2], Liu Xu-Yang[2], Wang Tian-Yu[2], Li Yu[2], Zou Hua-Wen[1,*], and Li Chun-Hui[2,*]

[1] College of Agriculture, Yangtze University, Jingzhou 434025, Hubei, China; [2] State Key Laboratory of Crop Gene Resources and Breeding / Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract: As the primary organ responsible for water and nutrient uptake, the root system plays a vital role in maize growth and yield formation. Understanding the genetic basis of root architectural traits—particularly lateral root density (LRD)—is crucial for advancing high-yield breeding strategies. In this study, an F_2 population was developed from a cross between Danhuang 02 (high LRD) and HRB16041 (low LRD). Bulked segregant analysis coupled with high-throughput sequencing (BSA-seq) was conducted to identify genomic loci significantly associated with LRD. In parallel, transcriptome sequencing of lateral root tissues from eight diverse maize inbred lines was performed to explore candidate genes regulating LRD at the seedling stage. BSA-seq identified 240 significantly associated SNPs and 2420 InDels, involving 235 candidate genes located in regulatory regions, alternative splicing sites, or coding sequences. RNA-seq analysis revealed 1979 differentially expressed genes (DEGs) associated with LRD, including 954 up-regulated and 1025 down-regulated genes. Integrated analysis of BSA-seq and RNA-seq data identified 12 overlapping genes as strong candidates for regulating LRD. qRT-PCR validation suggested that *Zm00001d033708* and

本研究由国家自然科学基金项目(32201751), 财政部和农业农村部国家现代农业产业技术体系建设专项(CARS-0204)和中国农业科学院创新工程项目资助。

This study was supported by the National Natural Science Foundation of China (32201751), the China Agriculture Research System of MOF and MARA (CARS-0204), and the Innovation Program of Chinese Academy of Agricultural Sciences.

* 通信作者(Corresponding authors): 何冠华, E-mail: heguanhua@caas.cn; 邹华文, E-mail: zouhuawen@yangtzeu.edu.cn; 李春辉, E-mail: lichunhui@caas.cn

第一作者联系方式: E-mail: 1412147019@qq.com

Received (收稿日期): 2025-09-12; Accepted (接受日期): 2026-01-22; Published online (网络出版日期): 2026-02-11.

URL: https://link.cnki.net/urlid/11.1809.S.20260211.1022.002

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Zm00001d045720 may act as positive regulators of lateral root density, while *Zm00001d040375* may function as a negative regulator. This study provides valuable genetic resources and candidate genes for understanding and improving lateral root development in maize.

Keywords: maize; lateral root density; bulked segregant analysis; transcriptome analysis; expression level

玉米是全球最重要的粮食作物之一[1-2], 在保障粮食安全和农业可持续发展中占据核心地位。根系作为玉米吸收水分和养分的主要器官, 其形态与功能直接影响植株的生长潜力和产量形成[3-6]。玉米根系由胚根和萌发后的胚后根构成[7]。胚根由一条主根和不定数目的种子根组成。胚后根包括冠根、气生根和侧根[8]。胚根在玉米早期发育中起到非常重要的作用, 胚后根影响玉米成熟期根系的形态结构。田间玉米的地下根系构型主要由冠根及其上面的侧根组成, 侧根的基部直径变化很大, 基部直径决定着侧根的长度[9]。侧根通过增加土壤勘查[10]、增大根系对养分的吸收面积发挥重要作用[11]。双子叶植物拟南芥的侧根发生不同于单子叶植物玉米的侧根[12-13], 其起始位点定位于皮质组织深层的特定细胞群, 具体包括中柱鞘细胞与内胚层细胞[14], 这一细胞来源的特异性是玉米根系发育区别于拟南芥的关键特征之一。从发育进程来看, 玉米侧根的形成可划分为 8 个连续阶段: 发育起始于韧皮部中柱鞘细胞的背斜向分裂(定义为阶段 I); 进入阶段 II 后, 该类分裂细胞进一步发生周斜向分裂[15]; 演进至阶段 III 时, 中柱鞘细胞持续向外周方向分裂, 同时内胚层细胞启动背斜向分裂程序; 后续通过背斜分裂与周斜分裂的协同作用, 最终形成侧根原基[16]。随着侧根原基的持续生长与分化, 其逐渐突破主根的表皮组织, 完成侧根的形态建成过程。

玉米侧根性状具有较大的遗传多样性和较高的遗传力水平。目前, 连锁分析和关联分析是挖掘自然变异中复杂数量性状基因的 2 种主要方式。基于连锁分析和关联分析, 针对玉米侧根发育相关性状已定位了一些基因。优良自交系综 3 和 87-1 分别具有侧根密度大和侧根密度小的特点。Zhang 等[13]利用来源于综 3 和 87-1 的杂交种豫玉 22 构成的 RIL 群体, 图位克隆了控制侧根数量的主要数量性状位点(QTL) *qLRT 5-1*; 精细定位和功能分析确定, *qLRT 5-1* 的候选基因为 *ZmLRT*; *ZmLRT* 的敲除和过表达分别增加和减少侧根数量。Li 等[17]鉴定得到一个 *ZmNAC1* 基因; 在综 3 和 87-1 组成的 RIL 群体中, *ZmNAC1* 的表达量与群体材料的侧根密度表型显著相关; 拟南芥中过表达 *ZmNAC1* 基因可提高植物的侧根密度。Yu 等[18]利用 1604 份玉米自交系组成的自然群体, 结合苗期根系表型数据, 通过 GWAS 技术挖掘到一个调控玉米根系构型的关键基因 *ZmHb77*。功能验证发现, *ZmHb77* 通过控制种子根和侧根密度, 进而影响根系整体的吸水能力, 可提高玉米苗期的抗旱性, 进一步挖掘到该基因控制根系构型及其抗旱性的有利等位变异。Li 等[19]以 357 份玉米自交系为材料, 整合了全转录组关联分析和单细胞 RNA 测序(scRNA-seq)技术, 精准定位到在根系皮层和表皮细胞中高表达并与侧根长度显著关联的关键转录因子 ZmbZIP89; ZmbZIP89 能够直接结合并激活下游过氧化物酶基因 *ZmPRX47* 的表达。ZmbZIP89-ZmPRX47 分子模块通过精细调控侧根原基周围组织的活性氧(ROS)稳态, 促进侧根原基顺利穿透皮层和表皮细胞, 从而显著增加侧根长度, 扩大根系吸收面积。

尽管目前已对玉米侧根发育遗传基础进行了部分解析, 但是对玉米侧根密度相关基因克隆的研究还比较少。本研究利用丹黄 02 (侧根密度大)和 HRB16041 (侧根密度小)构建了一个 F_2 群体, 利用 BSA 极端性状混池测序法(BSA-seq)获得控制侧根密度(lateral root density, LRD)的显著关联位点, 结合 8 个玉米自交系侧根组织的转录组测序, 鉴定出与玉米侧根密度相关的候选基因。这些基因的发现将为深入理解玉米根系发育调控网络提供重要依据。

1 材料与方法

1.1 试验材料与侧根密度表型鉴定

BSA-seq 关联群体: 通过对玉米自交系初生根侧根密度的鉴定, 筛选出具有极端表型的 2 个自交系丹黄 02 和 HRB16041, 并以此构建了 F_2 群体。

转录组测序材料: 通过实验室前期对初生根上侧根密度的表型鉴定结果, 本试验共选用 8 个玉米自交系 LH206、PHN46、LH181、PB80、MC5834、T34907、郑 58 和 XH_1。

玉米发芽培养 9 d 后, 选取正常幼苗, 沿初生根基部剪取 10 cm 根段, 置于透明托盘水中, 经 Epson Scan 扫描后, 人工计数侧根数量, 取平均值并计算单位长度侧根密度(根 cm^{-1})。

1.2　BSA-seq 关联定位

在丹黄 02 与 HRB16041 构建的 F_2 群体中, 对约 7000 个单株进行侧根密度鉴定, 从中各选取 100 株极端表型个体, 分别取叶片混池后提取 DNA。经质量检测合格后, 将 DNA 片段化至 350 bp, 使用 TruSeq 建库试剂盒构建文库, 并在 Illumina HiSeq PE150 平台上完成全基因组重测序, 该测序工作由北京诺禾致源科技股份有限公司承担。

经过数据质量控制和过滤后得到 418.76 G 的 clean reads, 各样本的原始数据在 12,123.17~197,574.02 M 之间, 表明所有样本的数据量足够用于进一步研究。数据质量检测结果显示, 测序错误率低(error rate ≤ 0.03%), 测序数据质量高(Q20 ≥ 93.06%、Q30 ≥ 86.03%), GC 含量为 46.11%~47.27%, GC 含量正常分布, 表明建库测序成功。下载的玉米参考基因组大小为 2135.08 G, 所有样本的比对率在 97.27%~98.87%之间, 2 个亲本对参考基因组(排除 N 区)的平均覆盖深度为 6.09×和 7.29×, 高池和低池的平均覆盖深度是 72.70×和 73.47×, 4 个样本的 1×覆盖度(至少有一个碱基的覆盖)均在 82.16%以上。测序数据质量检测结果显示, 过滤数据可进行后续的变异检测及相关分析。

1.3　侧根组织的转录组分析

种子发芽后更换为 50% Hoagland's No.2 营养液, 在 28℃/25℃、16 h 光照/8 h 黑暗条件下培养 2 d, 收集正常幼苗根系 1~2 cm (每个样品 3 次生物学重复), 选用 8 个自交系样品进行 RNA 提取与测序。转录组建库测序部分由北京诺禾致源生物信息科技有限公司完成。

经过数据过滤后, 共得到 421.80 million 的 clean reads, 平均每个样品 52.73 million。各样品的错误率均为 0.03%, Q20 均达 96.50%以上, Q30 值范围在 91.44%~94.10%之间, 表明测序结果可靠。将 clean reads 与玉米参考基因组(B73_RefGen_V4)进行序列比对, 共比对到的 reads 有 370.09 million, 占 87.74%, 每个样品约有 85.94%的序列在参考基因组上只有唯一比对位置, 1.80%的序列有多个比对位置。

采用 HTSeq v0.6.1 对定位至基因组或外显子区域的测序数据进行分析, 以 FPKM 值估算基因表达水平, 该指标综合考虑测序深度与基因长度的影响。基于各样本基因 FPKM 矩阵, 进一步开展样本间相关性分析与层次聚类, 使用 R 语言计算 Pearson 相关系数, hclust 函数完成聚类。运用 DESeq2 软件对高、低侧根密度 2 组样本进行差异表达分析: 首先对 read count 数据进行标准化, 再基于负二项分布模型计算 *P* 值, 并经 Benjamini-Hochberg 法校正得到 FDR。差异基因筛选标准为$|\log_2(\text{Fold change})| > 1$ 且 $P_{adj} < 0.05$。将基因的 CDS 序列通过 blast 比对到 Swiss-prot 数据库 (http://www.uniprot.org/) 和 NCBI NR 数据库(https://www.ncbi.nlm.nih.gov/)中, 对基因的功能进行注释。基于差异表达基因, 使用 agriGO v2.0 软件进行 GO (gene ontology, http://www.geneontology.org/)富集分析[20], 使用 KOBAS v2.0 进行 KEGG (kyoto encyclopedia of genes and genomes) 通路富集分析[21]。

1.4　侧根密度候选基因的 qRT-PCR 分析

取正常幼苗根系 1~2 cm, 使用南京诺唯赞生物科技股份有限公司 Fastpure Universal Plant Total RNA Isolation Kit 试剂盒(RC441)提取总 RNA, 反转录为 cDNA。qRT-PCR 反应程序: 95℃预变性 30 s; 95℃ 10 s, 60℃ 30 s, 40 个循环。使用 $2^{-\Delta\Delta Ct}$ 法计算相对表达量。内参基因为 *GAPDH*, 本研究使用的引物序列见表 1。

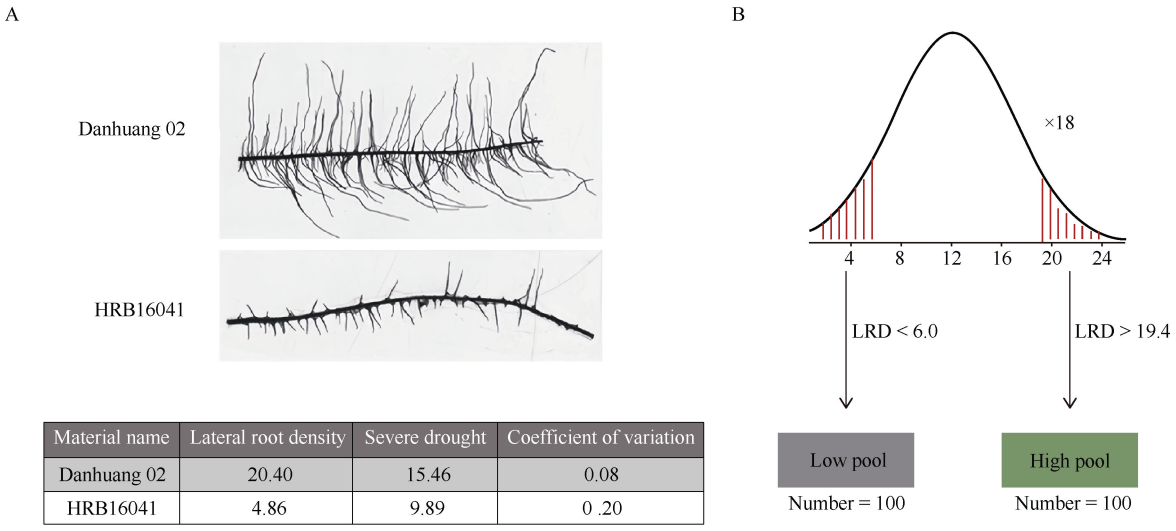
表 1　本研究所使用的引物
Table 1　Primers used in this study

引物名称 Primer name	引物序列 Primer sequence (5′–3′)
Zm00001d033708-F	AATTCCCTGCATCCAGACCG
Zm00001d033708-R	GGAACAACCCATTCTGGGCA
Zm00001d034546-F	CCTCCATGACGATGCTACACT
Zm00001d034546-R	GAAACGAAAACCTCCACAGGC
Zm00001d040315-F	AGGAAGAAGCTCGATGACGC
Zm00001d040315-R	TGGACCACCACCGCTATTTC
Zm00001d040375-F	TGGTAAAGTTGCCACCGTAG
Zm00001d040375-R	TTGCAGTTCGCGGCAGTTA
Zm00001d042901-F	CCGACCTCCCAAAATGCTGA
Zm00001d042901-R	GCAGGAAGCAGGAGAACCTT
Zm00001d045720-F	GCGATACAAGCGGAGGAGTT
Zm00001d045720-R	GGCTGCCACAGACGATACTT

2　结果与分析

2.1　BSA-seq 关联群体侧根密度表型鉴定

对 2 个亲本和 F_2 群体进行侧根密度表型鉴定, 丹黄 02 的侧根密度为 20.40 根 cm^{-1}, HRB16041 的侧根密度为 4.86 根 cm^{-1}, 2 个亲本的侧根密度表型差异较大(图 1-A)。共鉴定了 18 个 F_1 自交后收获的果穗和约 7000 个 F_2 单株的侧根密度表型, F_2 群体的



Material name	Lateral root density	Severe drought	Coefficient of variation
Danhuang 02	20.40	15.46	0.08
HRB16041	4.86	9.89	0 .20

图 1 亲本表型及混池构建方法

Fig. 1 Parental phenotype and construction of bulks

A: 丹黄 02 和 HRB16041 的侧根密度表型; B: F_2 群体的侧根密度表型正态分布图。LRD: 侧根密度。

A: phenotype of lateral root density in Danhuang 02 and HRB16041 under normal moisture conditions; B: normal distribution of lateral root density phenotype in the F_2 population. LRD: lateral root density.

侧根密度表型符合正态分布。在 7000 个单株的侧根密度表型中分别选取侧根密度大于 19.4 根 cm^{-1} 和小于 6.0 根 cm^{-1} 的单株进行取样, 构建侧根密度的低混池和高混池, 每个池各由 100 个极端性状的单株叶片混样构成(图 1-B), 2 个混池的构建可以用于 BSA 分析。对 2 个混池和 2 个亲本的 DNA 样品进行全基因组重测序, 各样本的原始数据在 12,123.17~197,574.02 M 之间, 表明所有样本的数据量足够用于进一步研究。

2.2 基于 BSA-seq 侧根密度遗传位点的挖掘

本试验以亲本为参考, 统计子代混池和亲本在某位点的变异, 利用 SNP-index 和 InDel-index 在全基因组范围内进行关联分析并挑选候选 SNP 和 InDel, 选择在 95%置信水平下大于阈值的窗口作为控制侧根密度性状的候选区间。由混池的 SNP 和 InDel 频率合并后的分布结果可知, 第 3、9、10 号染色体为主要的候选区域(图 2)。基于 SNP 标记的侧根密度性状区域定位, 根据阈值挑选 SNP 位点, 共挑选出 240 个显著关联 SNP, 其中, 5 个为同义突变、2 个为非同义突变、0 个位于基因上游 1 kb 范围内、9 个位于基因下游 1 kb 区域内, 其余的 SNP 位于基因内含子或基因间隔区。根据阈值共得到 2420 个显著关联的 InDel 位点, 其中, 8 个引起移码插入、15 个引起移码缺失、1 个 InDel 位于剪接位点、21 个变异未造成移码插入或缺失、263 个变异位于基因上游、16 个变异位于一个基因的上游区域及另一个基因的下游区域。

挑选 SNP 变异位于基因上游或非同义突变的位点所在的基因作为候选基因, 挑选 InDel 变异引起移码突变和位于剪接位点及基因上、下游的 InDel 所在的基因作为候选基因, 对变异位点进行注释, 共查找到 235 个候选基因。优先查看候选基因内的变异为移码插入或非同义突变或位于剪接位点的候选基因(表 2), 共得到 21 个候选基因, 1 个候选基因

表 2 基于 BSA 分析的候选基因及其功能注释

Table 2 Candidate genes identified by BSA analysis and their functional annotations

染色体 Chr.	基因 Gene	变异 [a] Variant [a]	基因功能注释 Gene functional annotation
1	*Zm00001d028400*	U, FI[(2)]	Pathogenesis-related thaumatin superfamily protein
3	*Zm00001d039542*	FI	Phospholipase A1-igamma1 chloroplastic
3	*Zm00001d040252*	FI	Heavy metal-associated isoprenylated plant protein 39
3	*Zm00001d041745*	FD	Unknown
3	*Zm00001d042347*	FD	Transcription elongation factor 1 homolog

(续表 2)

染色体 Chr.	基因 Gene	变异 [a] Variant [a]	基因功能注释 Gene functional annotation
3	*Zm00001d042393*	FD	Protein arginine methyltransferase NDUFAF7
3	*Zm00001d042753*	N	Transcription repressor OFP13
3	*Zm00001d042964*	U (3), FD (2)	Unknown
3	*Zm00001d043126*	FD	Ulp1 protease family C-terminal catalytic domain containing protein expressed
3	*Zm00001d043420*	FD	Basic leucine zipper 34
5	*Zm00001d013405*	S	Protein kinase superfamily protein
5	*Zm00001d013819*	FI	Galactose mutarotase-like superfamily protein
9	*Zm00001d045721*	FD	TRICHOME BIREFRINGENCE-LIKE 20
9	*Zm00001d047495*	FD	Premnaspirodiene oxygenase
10	*Zm00001d024359*	FD	Alpha-humulene/(−)-(E)-beta-caryophyllene synthase
10	*Zm00001d025255*	FD	Putative DUF26-domain protein kinase
10	*Zm00001d025522*	FD (2)	Putative FKBP-type peptidyl-prolyl *cis*-trans isomerase family protein
10	*Zm00001d025821*	N	Putative DUF231 domain containing family protein
10	*Zm00001d025952*	FI	NAD(P)H-quinone oxidoreductase subunit M chloroplastic
10	*Zm00001d026021*	FD (2), FI (2)	Putative RING zinc finger domain superfamily protein
10	*Zm00001d026120*	FI	Putative HLH DNA-binding domain superfamily protein

a: 括号内为变异数目; U: 基因上游; FI: 移码插入; FD: 移码缺失; N: 非同义突变; S: 可变剪接。

a: numbers in parentheses indicate the number of variations; U: upstream; FI: frameshift insertion; FD: frameshift deletion; N: nonsynonymous; S: splicing.

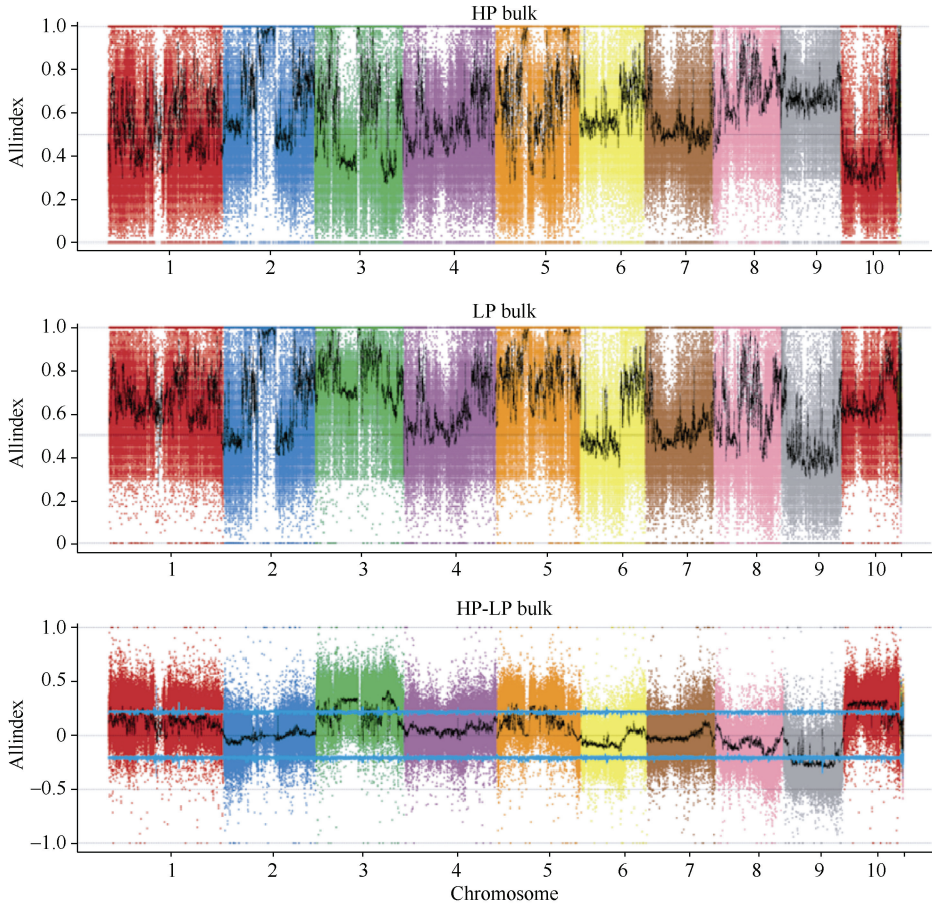


图 2　SNP-index 和 InDel-index 在染色体上的分布

Fig. 2　Distribution of SNP-index and InDel-index across chromosomes

位于第 1 号染色体上, 9 个候选基因位于第 3 号染色体上, 2 个候选基因位于第 5 号染色体上, 2 个候选基因位于第 9 号染色体上, 7 个候选基因位于第 10 号染色体上。其中, 有 4 个候选基因在多个位置存在 SNP 或 InDel 差异, *Zm00001d028400* 的启动子区存在一个 InDel 差异, 基因内存在 2 个移码插入; *Zm00001d042964* 在启动子区域检测到 3 个 InDel 差异, 同时还有 2 个移码缺失; *Zm00001d025522* 检测到 2 个移码缺失; *Zm00001d026021* 基因内存在移码缺失和移码插入各 2 个。这 21 个候选基因中, 有 2 个基因功能未知, 大多数基因的功能为蛋白激酶。

2.3 侧根密度表型分析及 RNA-seq 数据分析

对 8 个玉米自交系的侧根密度性状表型进行鉴定发现, LH206、PHN46、LH181、PB80、MC5834、T34907、郑 58 和 XH_1 这 8 个自交系的侧根密度分别为 14.48、14.10、14.08、17.88、9.93、8.98、7.10 和 9.25 根 cm^{-1}。LH206、PHN46、LH181 和 PB80 这 4 个自交系的侧根密度显著大于 MC5834、T34907、郑 58 和 XH_1 这 4 份自交系(图 3-A)。结果表明, 这 8 个自交系由 4 个侧根密度大和 4 个侧根密度小的自交系组成, 并且 2 组材料在侧根密度表型上存在显著差异, 为有效开展 RNA-seq 分析奠定了基础。为评估转录组的基因表达水平, 首先对 49,725 个基因表达进行筛选, 其中, 至少在 6 个玉米自交系中有表达(FPKM≥1)的基因共 26,741 个。根据其是否编码蛋白, 将其分为非编码蛋白的基因和编码蛋白的基因, 编码蛋白的基因数目明显多于非编码蛋白(图 3-B, C)。

2.4 差异表达基因的功能富集分析

侧根密度表型值大的一组样品(high lateral root density, MLR)中共有 23,064 个基因表达(FPKM ≥ 1), 表型值小的一组样品(low lateral root density, FLR)中表达的基因有 24,097 个, 其中, 有 21,545 个基因在 2 组中

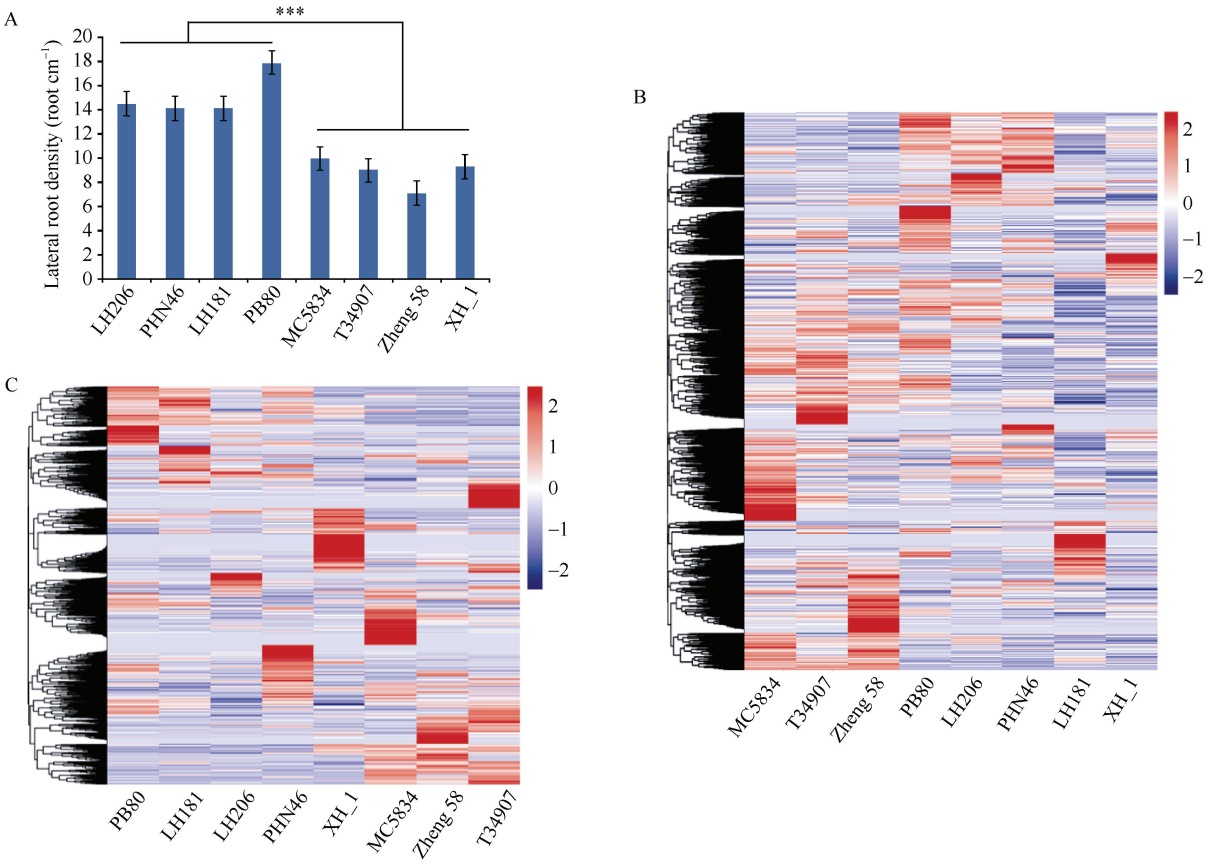


图 3 8 个玉米自交系侧根密度及转录组基因表达分析图

Fig. 3 Lateral root density and transcriptome analysis of eight maize inbred lines

A: 8 个玉米自交系的侧根密度表型变化; B: 转录组差异表达基因中的编码蛋白基因聚类分析; C: 转录组差异表达基因中的非编码蛋白基因聚类分析。Zheng 58: 郑 58。

A: variation in lateral root density among eight maize inbred lines; B: clustering of protein-coding differentially expressed genes; C: clustering of non-coding differentially expressed genes.

均表达, 仅在 MLR 中特异性表达的基因有 1519 个, 仅在 FLR 中表达的基因数目为 2552 个(图 4-A)。为挖掘控制侧根密度的候选基因, 将 read count 数据进行标准化后, 对 2 组自交系的转录组数据进行差异表达分析。以表达水平差异 2 倍以上作为标准筛选差异表达基因, 结果显示, 侧根密度大的自交系相对于侧根密度小的自交系共有 954 个基因上调表达, 1025 个基因下调表达(图 4-B)。对差异表达基因进行 GO 富集分析和 KEGG 通路富集分析。富集结果表明, 差异表达基因的功能大多富集在生物过程和分子功能中, 显著富集且富集基因较多的项目主要是 single-organism metabolic process (单组织代谢过程)、oxidation-reduction process (氧化还原过程)、single-organism process (单组织过程)和 oxidoreductase activity (氧化还原酶活性)、carbohydrate metabolic process (碳水化合物代谢过程)等(图 4-C)。KEGG 富集分析的结果表明, 显著富集的通路包括 inositol phosphate metabolism (磷酸肌醇代谢)、flavonoid biosynthesis (类黄酮生物合成)、phosphatidylinositol signaling system (磷脂酰肌醇信号传导系统)、mismatch repair (错配修复)等(图 4-D)。

2.5　BSA-seq 和 RNA-seq 联合分析鉴定侧根密度的候选基因

将 BSA-seq 分析得到的 235 个候选基因与全基因转录组测序分析鉴定到的 1979 个差异表达基因进行联合分析, 对 2 组候选基因绘制韦恩图, 共有 12 个基因在 2 种组学分析中被同时鉴定到(图 5)。

这些基因极有可能调控玉米侧根的发育, 综合汇总 12 个候选基因的基本信息(表 3)。其中, *Zm00001d040375* 和 *Zm00001d025229* 基因负调控初生根上的侧根密度, 其余 10 个基因显示为正调控; *Zm00001d033708* 和 *Zm00001d040375* 基因在其下游 1 kb 范围内通过 BSA-seq 分析检测到了变异位点,

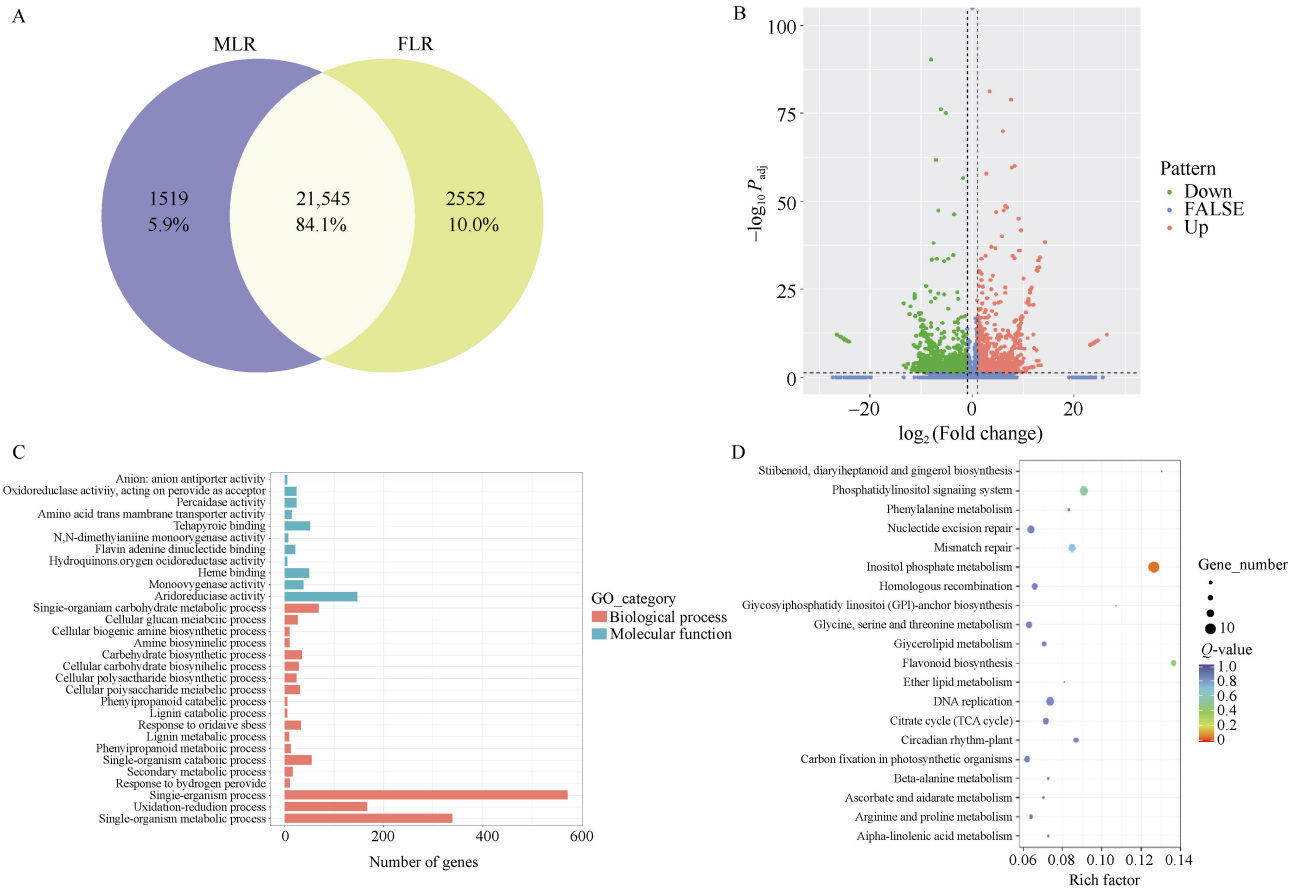


图 4　差异表达基因分析

Fig. 4　Analysis of differentially expressed genes (DEGs)

A: 基因表达韦恩图; B: 簇 4 差异表达基因火山图; C: 差异表达基因的 GO 富集分析; D: 差异表达基因的 KEGG 富集分析。MLR: 侧根密度大; FLR: 侧根密度小。

A: Venn diagram of gene expression; B: volcano plot of DEGs from cluster 4; C: GO enrichment analysis of DEGs; D: KEGG enrichment analysis of DEGs. MLR: high lateral root density; FLR: low lateral root density.

其余 10 个基因均在上游检测到变异位点。*Zm00001d025229* 和 *Zm00001d040315* 基因的功能未被注释。*Zm00001d025325* 编码干燥相关蛋白 PCC13-62; *Zm00001d033708* 编码 mRNA 帽鸟嘌呤-N7 甲基转移酶; *Zm00001d034546* 编码铜氧还蛋白超家族蛋白; *Zm00001d039567* 编码 LRR 蛋白 WM1.10; *Zm00001d040298* 编码 FPF1 蛋白; *Zm00001d040375* 编码磷酸甘油糖异构酶; *Zm00001d042901* 编码 L-抗坏血酸氧化酶; *Zm00001d043149* 编码富亮氨酸重复序列(LRR)家族蛋白; *Zm00001d045139* 编码液泡质子泵同系物 1; *Zm00001d045720* 编码毛状双折射样 20 蛋白。

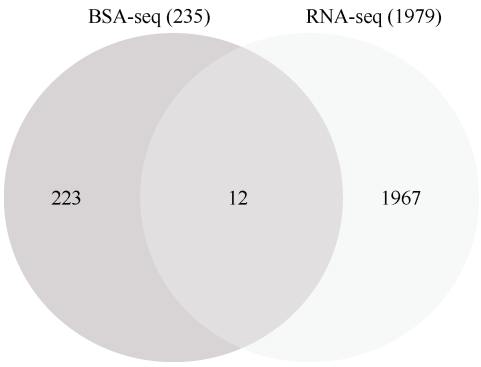


图 5 BSA-seq 和 RNA-seq 候选基因的韦恩图
Fig. 5 Venn diagram of candidate genes identified by BSA and RNA-seq

表 3 候选基因信息及功能注释
Table 3 Candidate gene information and functional annotation

基因 Gene	基因功能注释 Gene functional annotation	RNA-seq [a]	BSA-seq [b]
Zm00001d033708	mRNA cap guanine-N7 methyltransferase	Up	Downstream
Zm00001d034546	Cupredoxin superfamily protein	Up	Upstream
Zm00001d039567	LRR protein WM1.10	Up	Upstream
Zm00001d040298	FPF1-like protein 1	Up	Upstream
Zm00001d040315	Unknown	Up	Upstream (2)
Zm00001d040375	Triosephosphate isomerase cytosolic	Down	Downstream
Zm00001d042901	L-ascorbate oxidase	Up	Upstream
Zm00001d043149	Leucine-rich repeat (LRR) family protein	Up	Upstream
Zm00001d045139	Vacuolar proton pump homolog 1	Up	Upstream
Zm00001d045720	TRICHOME BIREFRINGENCE-LIKE 20	Up	Upstream
Zm00001d025229	Unknown	Down	Upstream
Zm00001d025325	Desiccation-related protein PCC13-62	Up	Upstream

[a]: RNA-seq 分析得到的差异表达基因类型; [b]: BSA-seq 分析得到的显著关联 SNP 或 InDel 变异位点位于候选基因的位置, 上标括号内为变异位点数量。

[a]: types of differentially expressed genes identified by RNA-seq; [b]: significantly associated SNPs or InDel variants identified by BSA-seq located within candidate gene regions, numbers in parentheses indicate the number of mutation sites.

针对上述侧根密度候选基因, 本研究选取了 *Zm00001d033708*、*Zm00001d045720*、*Zm00001d040375* 进行 qRT-PCR 验证(图 6)。结果表明, *Zm00001d033708* 和 *Zm00001d045720* 2 个基因在侧根密度大的自交系中的表达量均高于侧根密度小的自交系; *Zm00001d040375* 在侧根密度大的自交系中的表达量低于侧根密度小的自交系。

3 讨论

侧根是植物根系系统的关键组成部分, 对植物的生长发育、资源获取及环境适应均具有重要作用。在玉米中, 侧根能够显著增强植株对水分和养分的吸收, 提高抗倒伏能力与土壤锚定效果, 从而直接促进生物量的积累和产量形成。此外, 侧根的发育状况还直接影响玉米在干旱和营养胁迫等逆境条件下的适应性与生存能力, 是维持其田间稳定性和高产潜力的重要结构基础。因此, 挖掘到的和侧根形成相关的候选基因, 对玉米根系构型的分子育种改良、抗逆性和产量提升提供了关键基因资源。

本研究利用 BSA-seq 共挖掘到 235 个控制侧根密度的候选基因, 通过 RNA-seq 共找到 1979 个和侧根密度性状相关的差异表达基因。通过整合 2 种组学分析结果, 发现共有 12 个基因重叠。无论是 BSA-seq 鉴定到的关联基因, 还是 RNA-seq 筛选出

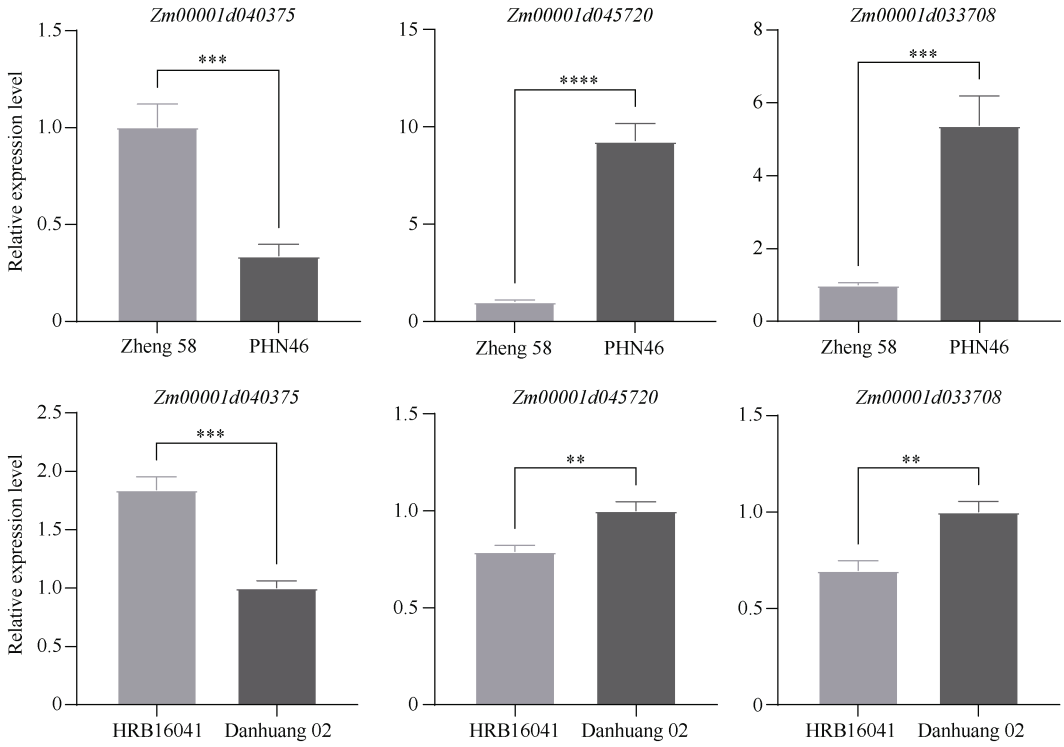


图 6　侧根密度候选基因的 qRT-PCR 分析

Fig. 6　qRT-PCR analysis of candidate genes for lateral root density

表示在 0.01 水平差异显著，*表示在 0.001 水平差异显著，****表示在 0.0001 水平差异显著。Zheng 58: 郑 58, 侧根密度小的自交系; PHN46: 侧根密度大的自交系; HRB16041: 侧根密度小的自交系; Danhuang 02: 丹黄 02, 侧根密度大的自交系。

** indicates a significant difference at $P < 0.01$, *** indicates a significant difference at $P < 0.001$, **** indicates a highly significant difference at $P < 0.0001$. Zheng 58: inbred line exhibiting a high density of lateral roots; PHN46: inbred line exhibiting a low density of lateral roots; HRB16041: inbred line exhibiting a high density of lateral roots; Danhuang 02: inbred line exhibiting a low density of lateral roots.

的差异表达基因, 其数目通常较大, 翻译方法难以获得较小的候选区域或少量的候选基因。相比之下, 通过 BSA-seq 与 RNA-seq 技术联合进行候选基因的挖掘, 显著缩小了候选基因的范围, 对于后续解析目标性状的关键功能基因具有重要推动作用。Park 等[22]进行 QTL 定位和 RNA-seq 分析, 揭示了辣椒果皮中控制辣椒素类生物合成的基因。采用基于全基因组测序的 QTL-seq 分析, 鉴定到第 6 号染色体上存在一个 QTL。同时, 利用同样的群体进行传统的 QTL 定位, 共鉴定了 15 个与辣椒素含量相关的 QTL, 分别位于第 3、6、11 号染色体上。位于第 6 号染色体下臂的 QTL 在 2 种定位方法中均被检测到, 该区域被作为主要的候选区段。利用 RNA-seq 的分析方法对 QTL 中的候选基因在刺激性和非刺激性果皮组织中的差异表达进行分析, 共同阐述辣椒素生物合成的调控机制。

利用 BSA-seq 关联定位检测到的显著变异位点挖掘的候选基因中, *Zm00001d043420* 基因内部存在一个移码缺失变异, 该基因为 bZIP 转录因子家族成员。bZIP 是植物中一个较大的转录因子家族, 参与调节胁迫反应和激素信号转导途径[23]。通过将 *Zm00001d043420* 基因的蛋白序列, 与已知参与根系发育调控的 bZIP 家族转录因子进行同源性比对分析, 结果显示, 该蛋白 N 端与 C 端的 2 个结构域, 与已报道的根系发育相关 bZIP 转录因子 ZmbZIP4 序列存在显著同源性。鉴于 ZmbZIP4 在根系发育过程中的明确调控功能[21], 推测 *Zm00001d043420* 基因可能同样参与根系发育的调控过程。

基于 BSA-seq 和 RNA-seq 联合分析得到的 12 个候选基因, 通过对候选基因功能的查询, 发现 *Zm00001d040298* 功能注释为 FPF1-like protein 1, 在拟南芥和棉花中的同源基因 *AtFPF1* 和 *GhFPF1* 均被研究证实影响拟南芥的开花时间[24]。该基因与水稻中的 *OsRAA1* 基因高度同源, 同源性达 83%, *OsRAA1* 基因编码 12.0 kD 蛋白。通过原位杂交和转基因植株的 GUS 染色, 发现 *OsRAA1* 在顶端分生组

织、根尖的伸长区和侧根中均特异性表达。*OsRAA1* 的过表达在转基因水稻植株中导致初生根生长受到抑制、侧根增多, 在烟草中过表达该基因还导致烟草的叶片伸长、产生不育小花。同时证实, OsRAA1 是由生长素介导的, 在水稻根系发育中发挥作用[25]。根据同源基因的功能研究, *Zm00001d040298* 调控玉米开花时间和侧根发育的可能性较高, 但其余候选基因仍不排除其功能与侧根发育相关的可能性。本研究通过整合 BSA-seq 和 RNA-seq, 在侧根密度大和侧根密度小的自交系中鉴定到 12 个与侧根密度相关的候选基因, 并对其表达模式进行了验证, 为玉米根系遗传改良提供了候选基因资源。

4 结论

本研究利用 2 份侧根密度差异较大的自交系丹黄 02 和 HRB16041 构建 F_2 群体, 基于 BSA 的方法进行极端表型混池的构建, 利用全基因组重测序获得基因型, 并检测混池间 SNP 和 InDel 位点的频率差异, 挖掘出和侧根密度关联的候选基因。另外, 选取 8 个玉米自交系, 对侧根组织进行取样, 通过转录组测序研究基因的表达水平。BSA-seq 共检测到 240 个显著关联的 SNP 变异位点, 2420 个显著关联的 InDel 变异位点, 在基因上下游、编码区和可变剪接位点的变异共涉及 235 个候选基因。RNA-seq 分析共鉴定出与侧根密度相关的差异表达基因 1979 个, 其中, 954 个基因的表达和侧根密度呈正相关, 1025 个基因和侧根密度呈负相关。BSA-seq 和 RNA-seq 联合分析, 共有 12 个基因重叠, 可被认为是和侧根密度关联的候选基因。

References

[1] Dong Z Y, Wang Y B, Bao J X, et al. The genetic structures and molecular mechanisms underlying ear traits in maize (*Zea mays* L.). Cells, 2023, 12: 1900.

[2] Feng Y J, Li X L, Qin Y T, et al. Identification of anther thermotolerance genes by the integration of linkage and association analysis in maize. Plant J, 2024, 119: 1953–1966.

[3] Hammer G L, Dong Z S, McLean G, et al. Can changes in canopy and/or root system architecture explain historical maize yield trends in the U.S. corn belt? Crop Sci, 2009, 49: 299–312.

[4] Lynch J P. Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. Ann Bot, 2013, 112: 347–357.

[5] Villordon A Q, Ginzberg I, Firon N. Root architecture and root and tuber crop productivity. Trends Plant Sci, 2014, 19: 419–425.

[6] Ning P, Li S, White P J, et al. Maize varieties released in different eras have similar root length density distributions in the soil, which are negatively correlated with local concentrations of soil mineral nitrogen. PLoS One, 2015, 10: e0121892.

[7] Hochholdinger F, Woll K, Sauer M, et al. Genetic dissection of root formation in maize (*Zea mays*) reveals root-type specific developmental programmes. Ann Bot, 2004, 93: 359–368.

[8] Tai H H, Lu X, Opitz N, et al. Transcriptomic and anatomical complexity of primary, seminal, and crown roots highlight root type-specific functional diversity in maize (*Zea mays* L.). J Exp Bot, 2016, 67: 1123–1135.

[9] Wu Q, Pagès L, Wu J. Relationships between root diameter, root length and root branching along lateral roots in adult, field-grown maize. Ann Bot, 2016, 117: 379–390.

[10] Zhu J M, Kaeppler S M, Lynch J P. Topsoil foraging and phosphorus acquisition efficiency in maize (*Zea mays*). Funct Plant Biol, 2005, 32: 749–762.

[11] Xie Y H, Yu D. The significance of lateral roots in phosphorus (P) acquisition of water hyacinth (*Eichhornia crassipes*). Aquat Bot, 2003, 75: 311–321.

[12] Jansen L, Hollunder J, Roberts I, et al. Comparative transcriptomics as a tool for the identification of root branching genes in maize. Plant Biotechnol J, 2013, 11: 1092–1102.

[13] Zhang M, Chen Y H, Xing H Y, et al. Positional cloning and characterization reveal the role of a miRNA precursor gene *ZmLRT* in the regulation of lateral root number and drought tolerance in maize. J Integr Plant Biol, 2023, 65: 772–790.

[14] 程鹏. 玉米侧根发育缺陷突变体 *zmlrd7* 的基因克隆及功能研究. 河南大学硕士学位论文, 河南开封, 2024.
Cheng P. Map-Based Cloning and Functional Study of Lateral Root Development Deficiency Mutant *zmlrd7* in Maize. MS Thesis of Henan University, Kaifeng, Henan, China, 2024 (in Chinese with English abstract).

[15] Liu Y F, Li S J, Chen Y, et al. snRNA 3' end processing by a CPSF73-containing complex essential for development in *Arabidopsis*. PLoS Biol, 2016, 14: e1002571.

[16] Jansen L, Roberts I, De Rycke R, et al. Phloem-associated auxin response maxima determine radial positioning of lateral roots in maize. Philos Trans R Soc Lond B Biol Sci, 2012, 367: 1525–1533.

[17] Li J, Guo G H, Guo W W, et al. miRNA164-directed cleavage of ZmNAC1 confers lateral root development in maize (*Zea mays* L.). BMC Plant Biol, 2012, 12: 220.

[18] Yu P, Li C H, Li M, et al. Seedling root system adaptation to water availability during maize domestication and global expansion. Nat Genet, 2024, 56: 1245–1256.

[19] Li P C, Zhu T Z, Wang Y Y, et al. Natural variation in a cortex/epidermis-specific transcription factor bZIP89 determines lateral root development and drought resilience in maize. Sci Adv, 2025, 11: eadt1113.

[20] Tian T, Liu Y, Yan H Y, et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res, 2017, 45: W122–W129.

[21] Xie C, Mao X Z, Huang J J, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res, 2011, 39: W316–W322.

[22] Park M, Lee J H, Han K, et al. A major QTL and candidate genes

for capsaicinoid biosynthesis in the pericarp of *Capsicum chinense* revealed using QTL-seq and RNA-seq. Theor Appl Genet, 2019, 132: 515–529.

[23] Ma H Z, Liu C, Li Z X, et al. ZmbZIP4 contributes to stress resistance in maize by regulating ABA synthesis and root development. Plant Physiol, 2018, 178: 753–770.

[24] Wang X Y, Fan S L, Song M Z, et al. Upland cotton gene *GhFPF1* confers promotion of flowering time and shade-avoidance responses in *Arabidopsis thaliana*. PLoS One, 2014, 9: e91869.

[25] Ge L, Chen H, Jiang J F, et al. Overexpression of *OsRAA1* causes pleiotropic phenotypes in transgenic rice plants, including altered leaf, flower, and root development and root response to gravity. Plant Physiol, 2004, 135: 1502–1513.