作物学报 ACTA AGRONOMICA SINICA 2026, 52(3): 780−789 http://zwxb.chinacrops.org/
ISSN 0496-3490; CN 11-1809/S; CODEN TSHPA9 E-mail: zwxb301@caas.cn

DOI: 10.3724/SP.J.1006.2026.53051

基于 BSA-seq 技术定位玉米籽粒花青素关联基因

张 超[1,2,*] 郭 欢[1] 李忠玲[1] 岳淑宁[1] 赵 娜[1]

[1] 陕西植物线虫重点实验室 / 陕西省生物农业研究所, 陕西西安 710043; [2] 青海省农林科学院 / 青海大学, 青海西宁 810016

摘 要: 花青素作为鲜食玉米的表型之一, 使其具有更高的营养价值和经济价值。为了研究玉米籽粒中花青素的关联基因, 本研究利用连续自交多代的黑糯和白糯玉米作为亲本材料构建获得 F_2 代分离群体。于 $F_{2:3}$ 统计玉米籽粒颜色表明, 黑白粒占比表现为极显著差异。利用 BSA-seq 技术对黑粒和白粒的表型混池样本进行基因组测序, 挖掘花青素积累的关联基因。通过表型和基因测序联合分析, 玉米籽粒中花青素合成的关联基因片段主要位于 6 号染色体的 107.9 ~113.2 Mb 区域内。候选区域内共注释到 116 个基因, 其中包括 53 个非同义突变基因, 17 个移码突变基因。通过基因注释信息分析, 推测出花青素结构基因 *Zm00001eb276450* (*ANR*)、转录因子基因 *Zm00001eb276870* (*WD40*) 和 *Zm00001eb276640* (*MYB*)可能为玉米籽粒中花青素关联的候选基因。该研究结果为进一步完善玉米籽粒花青素关联基因的功能解析奠定了重要基础, 也为快速筛选花青素鲜食玉米新品种提供了参考基因。

关键词: 鲜食玉米; 花青素; BSA-seq; 表型-基因型关联分析; 序列差异

Mapping genes associated with anthocyanin in maize kernels using BSA-seq technology

Zhang Chao[1,2,*], Guo Huan[1], Li Zhong-Ling[1], Yue Shu-Ning[1], and Zhao Na[1]

[1] Shaanxi Key Laboratory of Plant Nematology / Bio-Agriculture Institute of Shaanxi, Xi'an 710043, Shaanxi, China; [2] Qinghai Academy of Agricultural and Forestry Sciences / Qinghai University, Xining 810016, Qinghai, China

Abstract: Anthocyanin is a phenotypic trait in fresh corn that enhances both its nutritional and economic value. To identify genes regulating anthocyanin accumulation in maize kernels, we developed an F_2 segregating population by crossing black and white waxy maize lines, followed by self-pollination. Statistical analysis of kernel coloration in the $F_{2:3}$ population revealed a highly significant segregation between black and white phenotypes, prompting bulked segregant analysis sequencing (BSA-seq) of phenotype-based pools. Integrating phenotypic data with genome sequencing, we mapped the anthocyanin-associated locus to a 107.9–113.2 Mb region on chromosome 6. Within this interval, 116 annotated genes were identified, including 53 genes with nonsynonymous mutations and 17 with frameshift mutations. Gene annotation highlighted three key candidates potentially involved in anthocyanin biosynthesis: *Zm00001eb276450* (*ANR*), *Zm00001eb276870* (*WD40*), and *Zm00001eb276640* (*MYB*). These findings provide a valuable foundation for elucidating the genetic mechanisms underlying anthocyanin accumulation in maize kernels and offer promising candidate genes for the rapid development of fresh-eating maize varieties with enhanced anthocyanin content.

Keywords: fresh corn; anthocyanin; BSA-seq; phenotype-genotype association analysis; sequence differences

本研究由陕西省科技厅青年基金项目(2024JC-YBQN-0194), 陕西省科学院种子基金项目(2022K-18), 西安市科技局农业技术攻关一般项目(2024JH-NYYBN0175), 陕西省生物农业研究所青年项目(2023SY01), 陕西省科学院基础研究面上项目(2023K-13)和陕西省科学院应用技术研发项目(2025K-05)资助。

This study was supported by the Shaanxi Provincial Natural Science Foundation Youth Project (2024JC-YBQN-0194), the Seed Fund of Shaanxi Academy of Sciences (2022K-18), the General Project of Agricultural Technology Research of the Xi'an Municipal Science and Technology Bureau (2024JH-NYYBN0175), the Youth Basic Research Project of Bio-Agriculture Institute of Shaanxi (2023SY01), the General Program for Basic Research Project of Academy of Sciences (2023K-13), and the Applied Technology Research and Development Project of Shaanxi Academy of Sciences (2025K-05).

* 通信作者(Corresponding author): 张超: E-mail: zhangc@xab.ac.cn

Received (收稿日期): 2025-07-16; Accepted (接受日期): 2025-11-18; Published online (网络出版日期): 2025-12-10.
URL: https://link.cnki.net/urlid/11.1809.S.20251210.1348.002

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

花青素(anthocyanin)属于黄酮类化合物, 在模式植物中, 其合成途径已被证实始于苯丙氨酸, 并经过一系列酶促反应完成。其中涉及多种酶活性[1-2]。花青素合成途径的解析表明, 结构基因是优化稳定性和颜色的主要焦点。在玉米调控系统中, Ming 等[3]对紫色玉米 963 进行了转录组测序, 对该品种不同生育期的样本进行差异分析, 最终初步形成了花青素合成过程中差异基因的表达模式, 并明确了在花青素合成途径中各差异所调控的位置。主要包括 *PAL*、*CHS*、*CHI*、*F3H*、*DFR* 和 *ANS* 基因。

玉米作为花青素的载体, 其籽粒、穗轴、茎段等部位均可含有花青素, 因此也作为花青素提取的原料之一。黑糯玉米作为富含花青素的鲜食玉米受到了消费者的追捧, 因此对影响玉米籽粒颜色差异相关的研究十分重要。玉米中花青素主要为矢车菊素、天竺葵素和芍药素, 其中矢车菊素和天竺葵素占比较高[4]。转录因子是调控结构基因的主要方式, 其可以促进花青素积累。其中, MBW 复合物是研究最为广泛的调控因子, 由 R2R3-MYB 蛋白、bHLH 蛋白和 WD40 重复序列结构域蛋白组成。WD40 结构域蛋白在多种细胞过程中发挥功能, 尤其是其家族中的TTG蛋白参与调控果实中花青素的积累并影响下游花青素合成结构基因的表达[5-6]。MYB 蛋白已在多个植物中参与调控花青素的合成[7], 其可诱导 *PAL*、*CHS*、*ANS* 和 *UFGT* 等结构基因的表达[8]。MYB 的 *C1* 和 *P1* 编码基因的表达分别决定了果皮和糊粉层花青素合成的不同发育时间[9]。bHLH 蛋白常与 MYB 相互作用从而调控花青素的合成。其中bHLH蛋白的*R1/B1*基因的表达具有组织特异性, 决定了色素的组织分布[10-12]。

花青素还原酶(anthocyanin reductase, ANR)作为花青素合成途径中的关键限速酶。与无色花青素还原酶(leucoantho-cyanidin reductase, LAR)均能将花青素还原成为无色原花青素[13]。同时也有大量研究表明, ANR 参与多种植物的原花青素的合成并能降低植物花青素的含量[14]。另一方面, 原花青素对人体健康具有显著益处, 其中包括抗氧化、抗菌、抗过敏和降血压等作用。

本研究通过构建黑糯和白糯玉米籽粒颜色表型分离 F_2 群体的极端表型混池, 应用 BSA-seq 全基因组测序技术及生物信息学方法, 快速分析调控玉米籽粒花青素合成相关候选基因, 为进一步开发分子标记提供良好的理论依据。

1　材料与方法

1.1　群体的构建与花青素含量检测

本试验以实验室保留的自交系黑糯玉米“N227”为父本、白糯玉米“3R1872”为母本, 于 2022 年在海南省进行繁种、留种。2023 年 5 月于陕西省西安市临潼区进行杂交, 获得 F_1 代杂交种; 2024 年 5 月于陕西省渭南市蒲城县进行套袋自交, 获得 F_2 代分离群体种; 2024 年 12 月在海南省三亚市崖州区对 F_2 代种进行播种、套袋自交, 观察 $F_{2:3}$ 代植株玉米籽粒颜色。共种植 600 余株, 最终可统计籽粒颜色单株共 594 株。

对收获的玉米籽粒进行颜色分类, 随机选择 6 穗玉米籽粒(3 穗黑粒和 3 穗白粒)进行真空冷冻干燥, 研磨至粉末状。称取 50 mg 粉末于 0.1 mL 50%甲醇(含 0.1%盐酸)提取液中, 涡旋 5 min, 超声 5 min 后, 4℃, 13,400×*g* 离心 3 min 取上清液并进行过滤(0.22 μm 微孔过滤膜), 超高效液相色谱(UPLC-MS)对 19 种花青素类物质进行定量检测。检测结果利用 SPSS 计算平均值和标准差, 并进行单因素方差分析获得其相关显著性。利用 GraphPad Prism 8 进行数据可视化。

超高效液相色谱(ExionLC AD)条件为: 选择色谱柱为 ACQUITY BEH C18 1.7 μm (2.1 mm × 100 mm); 流动相 A 相为超纯水(加入 0.5%的甲酸), B 相为甲醇(加入 0.5%的甲酸); 洗脱梯度为 0 min, B 相比例为 5%, 6.00 min 增至 50%, 12.00 min 增至 95%, 保持 2 min, 14 min 降至 5%, 并平衡 2 min; 流速 0.35 mL min^{-1}; 柱温 40℃; 进样量 2 μL。

质谱(QTRAP 6500+)条件为: 电喷雾离子源(electrospray ionization, ESI)温度 550℃, 正离子模式下质谱电压 5500 V, 气帘气(curtain gas, CUR) 35 psi。在 Q-Trap 6500+中, 每个离子对根据优化的去簇电压(declustering potential, DP)和碰撞能(collision energy, CE)进行扫描检测。

1.2　差异基因池构建及 BSA-seq 重测序

在 $F_{2:3}$ 代植株籽粒着色后期筛选长势相对一致的黑粒和白粒单株各 30 株, 分别对应编号并采收约 5~8 cm^2 的叶片于网袋中, 加入硅胶干燥剂进行干燥处理。采用 CTAB 法(cetyltrimethylammonium bromide)对每株样本进行总 DNA 的提取, 并检测浓度进行相应比例的混合, 建立 F_2 代表型分离群体的基

因池。

样品基因组 DNA 检测合格后, 用超声波打断法将 DNA 片段化, 然后对片段化的 DNA 进行片段纯化、末端修复、3′端加 A、连接测序接头, 利用琼脂糖凝胶电泳进行片段大小选择, 进行 PCR 扩增形成测序文库, 建好的文库先进行文库质检, 质检合格的文库用 Illumina 进行测序。

1.3 表型-基因型关联分析

将测序数据与玉米 B73_RefGen_v5.055 参考基因组进行比对, 利用 GATK 软件工具包对单核苷酸多态性(single nucleotide polymorphism, SNP)和插入缺失标记(insertion-deletion, InDel)差异结果进行检测分析。通过欧式距离算法(euclidean distance, ED) [15]和 SNP-Index / InDel-Index 算法分别对差异混池基因组进行表型的关联计算[16], 将 2 种方法获得的结果取交集, 利用 TBtools 对交集进行可视化。

1.4 候选基因注释及 qRT-PCR 分析

将候选蛋白于蛋白质序列数据库(Swiss-Prot protein knowledgebase, SwissProt)、基因本体论数据库(gene ontology database, GO)、京都基因与基因组百科全书(kyoto encyclopedia of genes and genomes, KEGG)和蛋白质直系同源簇数据库(clusters of orthologous groups of proteins database, COG)中进行比对[16], 快速定位与花青素合成或调控相关的候选基因。

在玉米果穗成熟中后期随机选择黑穗和白穗各 3 穗, 将新鲜取下的玉米籽粒置于液氮中保存带回实验室, 于液氮中进行研磨, 并选择 RNA 提取试剂盒(DR441, 天根)提取黑粒和白粒玉米总 RNA。第一链反转录使用 Hifair Ⅲ 1st Strand cDNA Synthesis SuperMix for qPCR (gDNA digester plus); 荧光定量检测使用 Hieff qPCR SYBR Green Master Mix (Low Rox Plus)试剂盒。将 3 份 cDNA 混样后进行表达量分析, 技术重复设置为 3 个。利用 Primer primer 5 软件设计相关候选基因的引物(表 1), 采用 AB-7500 荧光定量 PCR 仪进行检测。

表 1 候选基因定量引物设计
Table 1 Design of quantitative primers for candidate genes

基因名称 Gene name	引物序列 Primer sequence (5′–3′)
ACTIN	F: CCAAGGCCAACAGAGAGAAA
	R: TCACCGGAATCCATCACAATAC
Zm00001eb276450	F: CCTGCCGCTGGTTCACAT
	R: GGCGATGGTCGTGTTGAG
Zm00001eb276870	F: GGTGACTGGATTGGGACT
	R: AGACGCCTAAGCGACCAC
Zm00001eb276640	F: CTGGACCAACCACCTTAG
	R: GTTTCAGAACAACCCTCA
Zm00001eb276630	F: CCCAACTGCCTTCAAAGT
	R: GACGCTATCGGTCTCCTC
Zm00001eb276620	F: GTATTTGAGCACCTTGACT
	R: CAATGGCACTACCAACAG

2 结果与分析

2.1 玉米籽粒花青素含量分析

将黑糯玉米与白糯玉米进行杂交, F_1 代籽粒颜色均呈现紫黑色。F_2 代玉米籽粒颜色出现分离, 分别有黑粒、白粒和红粒的表型差异(附图 1)。根据表型差异进行分组统计, 共获得 F_2 代群体 594 株, 其中白粒占比不足 10%, 而黑粒占比超过 57%。根据孟德尔遗传定律和卡方检验(χ^2), 黑粒、红粒和白粒的表型统计 χ^2 值为 10.606, 比对卡方分布表发现, 自由度为 2 的条件下, P 值介于 0.01 至 0.001 之间, 表明实测值与理论值存在显著差异(附表 1)。此外, 对黑色和白色籽粒花青素类物质进行检测表明, 黑色籽粒的花青素类含量主要为矢车菊素类(矢车菊素-3-O-(6-O-丙二酰-β-D-葡萄糖苷)和矢车菊素-3-O-葡萄糖苷), 以及部分的芍药花素类(芍药花素-3-O-葡萄糖苷)、天竺葵素类(天竺葵素-3-O-(6-O-丙二酰-β-D-葡萄糖苷)和天竺葵素-3-O-葡萄糖苷); 而

白色籽粒中, 仅天竺葵素-3-O-(6-O-丙二酰-β-D-葡萄糖苷)具有较少的含量被检测出(图 1)。

2.2　BSA-seq 测序质量分析

分别选择 $F_{2:3}$ 群体中籽粒表型为黑粒和白粒的单株各 30 株, 构建黑粒和白粒的混池, 利用 BSA-seq 基因组测序技术进行差异表型相关基因的定位, 并将测序基因与玉米参考基因组进行比对。对 2 个样本进行测序, 均获得了超过 74 Gb 的有效碱基数, 测序质量保持在 95.7% (Q30)以上, 测序结果在参考基因组的覆盖率超过 78.6%, 平均覆盖深度为 34, 基因组覆盖率(10 ×)超过 84.2% (表 2)。根据这些结果, 所有样品的数据量充足, 测序质量得到确认, GC 分布正常。测序数据与玉米参考基因组的比对结果正常, 可用于性状的次突变检测和基因定位。

通过 SNP 检测, 2 组杂交群体样本共获 12,590,579 个 SNP, 转换型 SNP 与颠换型 SNP 数量的比率为 2.57, 且各群体中杂合类型 SNP 占比高于 66.63%, 其中黑粒群体的杂合型 SNP 占比最高。通过 Small InDel 检测, 共获得 36,036 个 CDS InDels (17,593 个插入突变和 18,443 个缺失突变) (表 3)。大多数的插入和缺失碱基数为 3~6 个。

2.3　SNP 和 InDel 连锁分析

为了更好地对 SNP 进行筛选和关联分析, 本研究分别利用 Euclidean Distance (ED)算法和 SNP-Index 算法测序数据寻找混池间存在显著差异的标

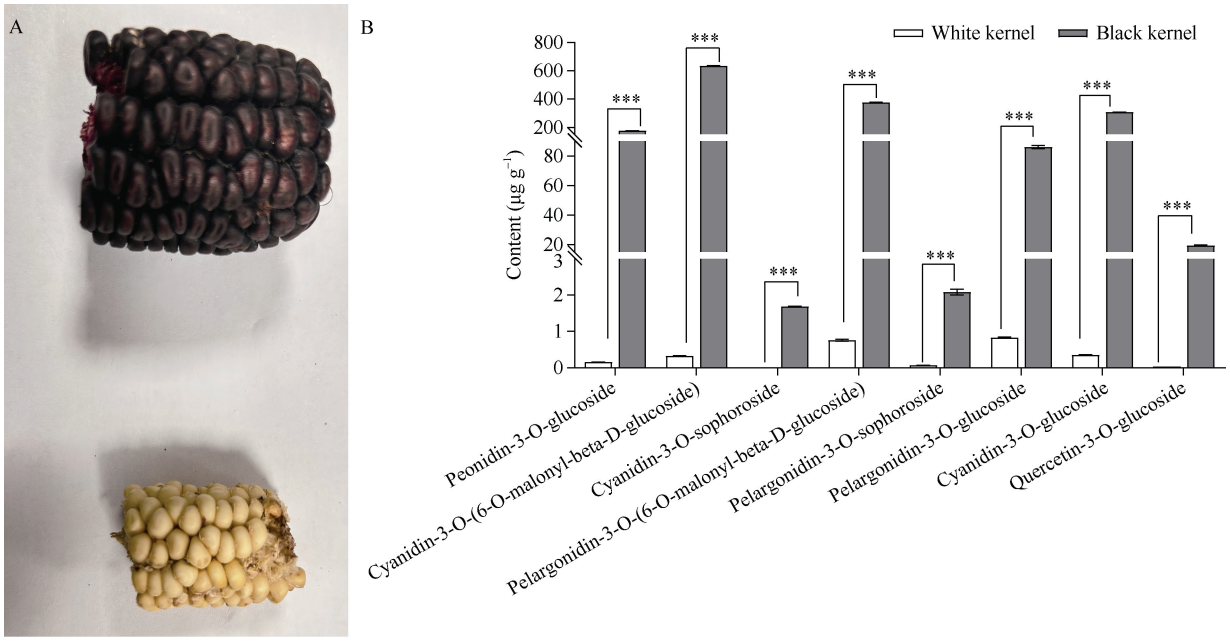


图 1　黑籽粒和白籽粒群体花青素类物质含量检测分析

Fig. 1　Analysis of anthocyanin content in the black and white kernels population

A: 黑、白粒玉米籽粒照片; B: 黑、白粒玉米籽粒花青素类物质含量。

A: photos of black and white kernels; B: anthocyanin content in black and white kernels. ***: $P < 0.001$.

表 2　测序样本 DNA 测序质量分析

Table 2　DNA sequencing quality assessment of the sequenced samples

样本 Sample	过滤读取数 Clean reads	过滤碱基数 Clean base	Q30 (%)	靶定参考基因组 Properly mapped (%)	覆盖深度 Coverage depth	10 ×覆盖深度比例 Cov_ratio_10 × (%)
黑粒 Black kernels	535,674,944	79,865,813,648	95.73	82.81	35	85.41
白粒 White kernels	517,409,140	76,429,320,409	96.02	78.68	34	84.22

Q30: 质量值大于等于 30 的碱基占总碱基数的百分比; Properly mapped: 双端测序序列均定位到参考基因组上且距离符合测序片段的长度分布; Coverage depth: 样品平均覆盖深度; Cov_ratio_10 × (%): 覆盖深度在给定深度及以上的碱基数占参考基因组总碱基数 10 × 以上比例。

Clean reads: number of clean reads; Clean base: number of clean bases; Q30: percentage of bases with a quality score ≥ 30. Properly mapped: percentage of read pairs where both ends are aligned to the reference genome with an insert size conforming to the expected library fragment length distribution; Coverage depth: average sequencing depth of the sample; Cov_ratio_10 ×: percentage of the reference genome covered at a depth of 10 × or greater.

表 3 样本间 SNP 和 InDel 类型和数量分析
Table 3 Analysis of SNP and InDel types and counts among samples

样本 Sample	SNP 数量 SNP number	转换型 SNP 数量 Number of transition SNPs	颠换型 SNP 数量 Number of transversion SNPs	杂合类型比率 Heterozygous genotype ratio	CDS 插入变异 Number of CDS-insertion	CDS 缺失变异 Number of CDS-deletion
黑粒 Black kernels	10,355,335	7,446,721	2,908,614	0.7731	14,864	15,989
白粒 White kernels	9,551,131	6,860,013	2,691,118	0.6663	13,832	14,885
总计 Total	12,590,579	9,065,232	3,525,347	—	17,593	18,443

记, 从而评估与性状关联区域。通过数据过滤分析, 黑粒和白粒混池共获得 9,895,682 个高质量 SNP。

在玉米黑粒和白粒混池中, 通过分别计算 ED 和 SNP-Index 的关联值, 将这些数据映射到玉米染色体上, 从而直接反映 SNP 指数在染色体上的分布情况(图 2)。针对 ED 值高于 0.7 的数据进行筛选和

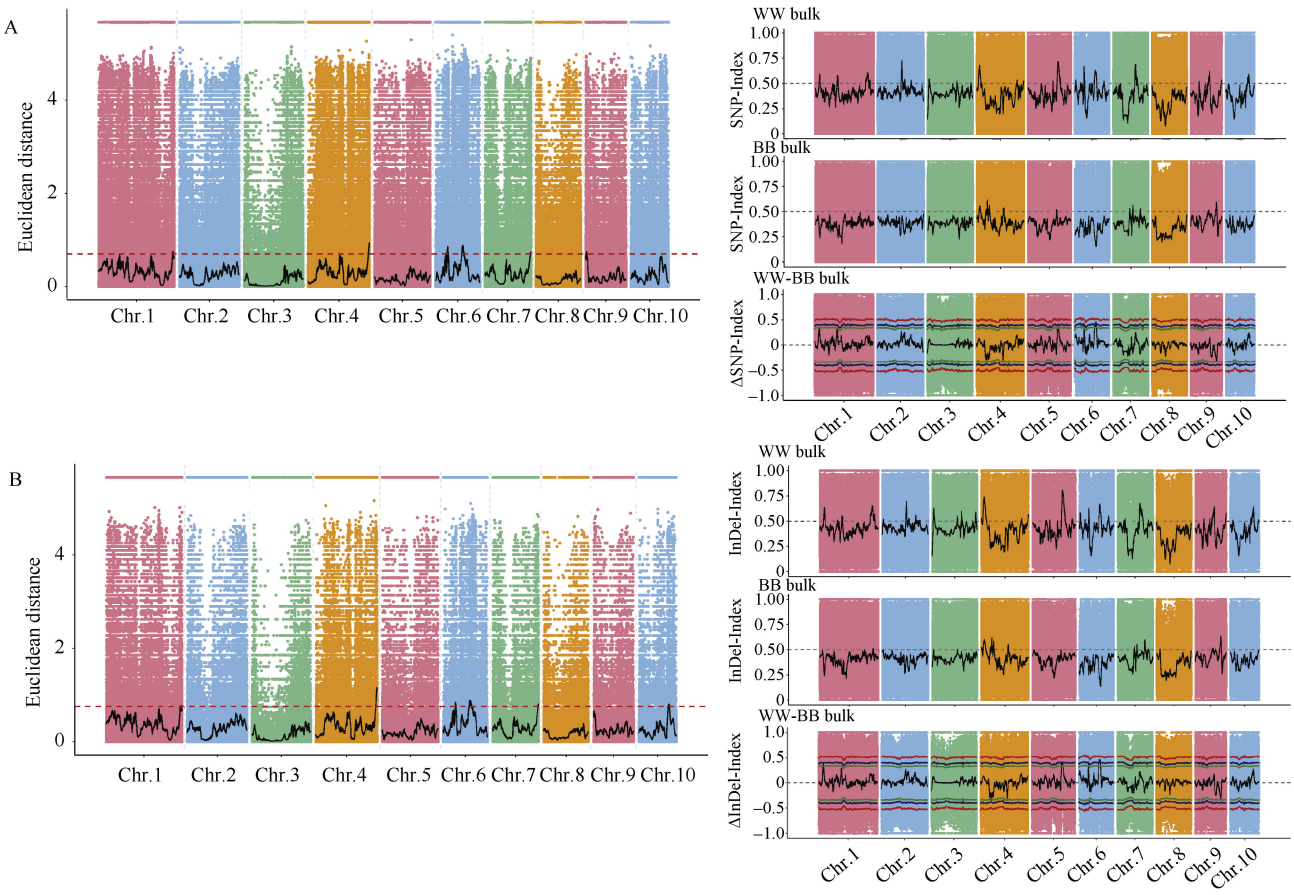


图 2 黑粒和白粒基因混池分析
Fig. 2 BSA analysis of black and white kernels pools

A: SNP 位点与表型关联分析: 左图为 ED 关联值在染色体上的分布, 右图为 SNP-Index 关联值在染色体上的分布; B: InDel 位点与表型关联分析: 左图为 ED 关联值在染色体上的分布, 右图为 InDel-Index 关联值在染色体上的分布(上图是隐性混池的 InDel-Index 值的分布图, 中图是显性混池的 InDel-Index 值的分布图, 下图是 ΔInDel-Index 值的分布图, 其中红色的线代表置信度为 0.99 的阈值线, 蓝色的线代表置信度为 0.95 的阈值线, 绿色的线代表置信度为 0.90 的阈值线)。

A: SNP-trait association: left panel shows distribution of euclidean distance (ED) values across chromosomes, and right panel shows SNP-index profiles. B: InDel-trait association: left panel presents ED distribution across chromosomes; right panel shows InDel-Index profiles (upper panel: InDel-Index values for the recessive pool; middle panel: InDel-Index values for the dominant pool; lower panel: ΔInDel-Index values. In the lower panel, the red line indicates the 99% confidence threshold, the blue line the 95% threshold, and the green line the 90% threshold).

计算, 结果显示, 在 1、4、6、7 和 9 号染色体上存在共计 11 个区域的 22.2 Mb 区域与表型关联效果更好。此外, SNP-Index 计算结果显示, 置信度高于 95%的区域共得到 2 个, 且均定位于 6 号染色体上的 6.1 Mb 区域中。对 2 种关联分析方法得到的 SNP 的关联区域结果取交集, 最终在 6 号染色体上的 6.2 Mb 长度中的基因与表型差异关联性更高。

再次执行 InDel 关联分析, 并识别出 1,507,379 个高质量 InDel。对 ED 值高于 0.75 的进行筛选和计算, 显示在 1、4、6、7 和 10 号染色体上共有 17 个区域的 23.2 Mb 与表型关联效果更好, InDel-Index 关联计算结果共得到 1 个区域, 均位于 6 号染色体上的 7.9 Mb 区域内(图 2-B)。对 2 种关联分析方法得到的 InDel 的关联区域结果取交集, 最终在 6 号染色体 5.3 Mb 长度中的基因与表型差异关联性更高。

为了定位与玉米籽粒颜色相关的基因组区域, 本研究结合 ED 和 Index 两种算法对基因组中的 SNP 和 InDel 进行表型关联分析, 最终筛选获得表型相关差异基因 5.3 Mb, 位于玉米 6 号染色体 107.9~113.2 Mb 区域内。4 种计算方法中各结果包含的基因数量分别为 SNP-ED 关联到 563 个基因、SNP-Index 关联到 139 个基因、InDel-ED 关联到 581 个基因、InDel-Index 关联到 173 个基因。对这些关联基因进行交集分析, 最终获得 116 个关联基因(图 3), 其中注释到 53 个非同义突变基因和 17 个移码突变基因。

2.4　玉米籽粒颜色相关调控基因定位分析及 qRT-PCR 分析

对候选区内非同义突变和移码突变基因注释信息进行分析, 结合基因的 GO 和 KEGG 结果, 最终仅发现与类黄酮生物合成途径(flavonoid biosynthesis, Ko00941)相关的基因 *Zm00001eb276450* (*ANR*)序列存在非同义突变。而花青素调控相关的转录因子编码基因 *Zm00001eb276870* (*WD40*)、*Zm00001eb276620* (*MYB*)、*Zm00001eb276630* (*MYB*)和 *Zm00001eb276640* (*MYB*)也存在非同义突变; 同时, *Zm00001eb276630* (*MYB*)和 *Zm00001eb276640* (*MYB*)基因还存在移码突变。以上 5 个基因在不同颜色籽粒中定量分析结果表明, *Zm00001eb276450* 在黑玉米籽粒中显著下调, *Zm00001eb276870* 和 *Zm00001eb276640* 基因在黑玉米籽粒中表达上调, 而其余 2 个基因表达差异不大(图 4)。

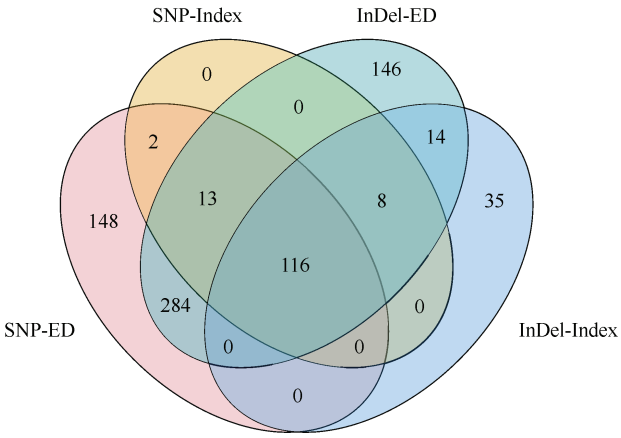


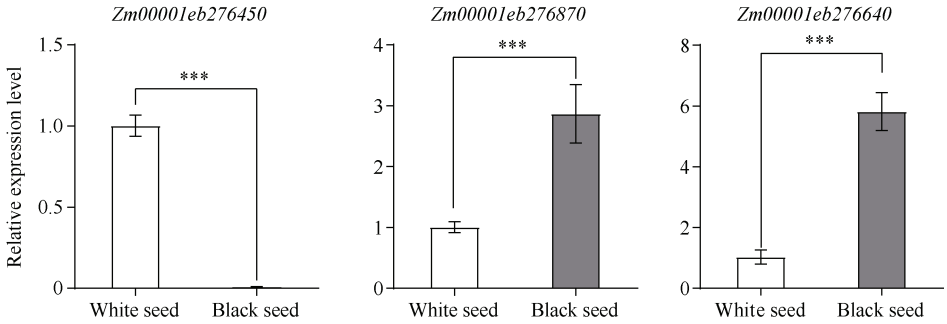
图 3　4 种计算方法连锁关联共有基因韦恩图

Fig. 3　Venn diagram of common genes associated with four calculation methods

SNP: 单核苷酸多态性; InDel: 插入缺失标记; ED: 欧式距离算法。

SNP: single nucleotide polymorphism; InDel: insertion-deletion; ED: euclidean distance.

对 *Zm00001eb276620* (*MYB*)、*Zm00001eb276630* (*MYB*)和 *Zm00001eb276640* (*MYB*)进行保守基序(conserved domain)分析发现, 仅有 *Zm00001eb276640* 包含 MYB 转录因子必须基序 PLN03091 superfamily, 且 *Zm00001eb276870* (*WD40*)也具有其典型的转录因子基序。因此认定 *Zm00001eb276620* (*MYB*)和 *Zm00001eb276630* (*MYB*)不具有 MYB 转录因子的功能。



(图 4)

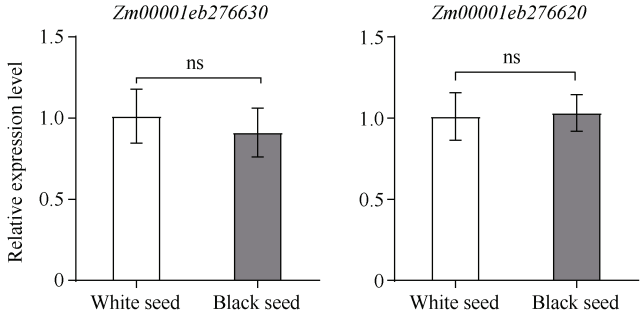


图 4 候选基因 qRT-PCR 分析

Fig. 4 qRT-PCR analysis of candidate genes

表示在 0.001 水平上差异显著; ns 表示差异不显著。 indicates a significant difference at $P < 0.001$; ns indicates no significant difference.

3 讨论

本研究以白糯和黑糯玉米为亲本, 构建了 F_2 代遗传分离群体。利用超高相液相色谱分析表明, 黑玉米籽粒中的花青素主要由矢车菊素、天竺葵素以及少量芍药素类构成, 且白粒玉米中基本无任何花青素类化合物, 这与 Hong 等[17]的研究一致。通过对 F_1 代群体表型分析可以推断出, 黑粒表型为显性遗传表型; 而近 600 份 F_2 分离群体单株籽粒颜色表型分析, 黑白籽粒表型比例超过 6∶1, 根据孟德尔遗传定律分析来看, 籽粒颜色的差异个数无法遵从孟德尔 3∶1 的遗传定律。通过卡方测试明确黑、红和白粒的表型分离比与显性上位的 12∶3∶1 的分离比、相加型变式的 9∶6∶1 以及重叠作用变式的 15∶1 均存在显著性差异(附表 1)。且在性状调查中玉米籽粒颜色会存在较小的颜色差异, 如紫红色、淡红色等过度色系。结合前人的研究表明, 花青素的积累是一个多基因控制的生物学过程, 因此推断其应该是数量性状。而在群体测序中高比例的表型差异则更有利于差异基因和表型的连锁研究[18]。

因此, 利用 BAS-seq 技术定位了玉米颜色差异相关性状的基因区域, 根据 SNP-ED、SNP-Index、InDel-ED 和 InDel-Index 4 种计算方法得出的定位区域进行比较分析, 最终明确位于 6 号染色体上共 5.3 Mb 的区域为 4 个计算方法中共同定位的区域, 因此可以推测该区域的置信度更高。对该区域进行分析也表明, 靶定的 116 个基因中有 53 个非同义突变基因, 17 个移码突变基因。基因注释信息中与花青素或黄酮类合成相关的基因仅有 1 个 *ANR* 基因[2]。而调控花青素合成的相关转录因子, 如 MYB、bHLH、WD40、ERF、WRKY、NAC 等[19], 在该区域内仅有 2 类共 4 个相关转录因子 MYB [20-21]和 *WD40* 基因[22]被预测出来。

Zm00001eb276450 (*ANR*)作为花青素结构基因, 是将花青素还原成原花青素的重要基因之一。张鲲[23]对不同种皮颜色花生进行遗传图谱定位时, 发现 *ANR* 基因为与花生种皮颜色相关的重要候选基因。通过对该基因在黑、白玉米籽粒表达的分析中发现, 在黑粒中 *ANR* 基因的表达反而低于白粒, 这也再次说明该基因可能参与原花青素的合成功能。Delgado 等[24]的研究也发现, 在胁迫诱导草莓花青素积累的过程中, *ANR* 基因表达与花青素含量呈显著负相关。同时, 刘泽[25]对花青素还原酶调控绿豆芽下胚轴花青苷积累的研究表明, *ANR* 可以正调控原花青素的生物合成, 负调控花青苷的生物合成。

本研究针对 *Zm00001eb276450* 在黑粒和白粒玉米中的表达差异情况, 对其基因序列进行了测序分析。由于 *Zm00001eb276450* 基因为反向序列, 因此对序列进行反向互补之后进行比对发现, *ANR* 基因中存在一段较长的差异序列(图 5), 在黑粒中这段差异序列有一段 75 bp 的碱基丢失。这与 BSA 测序与参考基因组比对分析中的 InDel 序列位置基本一致。因此推断可能是该序列的丢失导致 *ANR* 基因功能的减弱, 从而降低了有色花青素还原成原花青素的效率。

通过 *ANR* 的基因结构分析明确了该基因由 4 个外显子和 3 个内含子构成, 该结果与 NCBI 中的基因结构完全一致(https://www.ncbi.nlm.nih.gov/genecmd=search&term=109939362)。对 *ANR* 基因进行保守结构域分析表明, 黑粒中碱基缺失的区域所在第 3 外显子(400~600 bp), 是 *ANR* 基因的重要功能区域 PLN00198 (anthocyanidin reductase, ANR) (https://www.ncbi.nlm.nih.gov/Structure/labs/wrpsb.cgirid=CKBU1288015), 因此在花青素还原为原花青素的

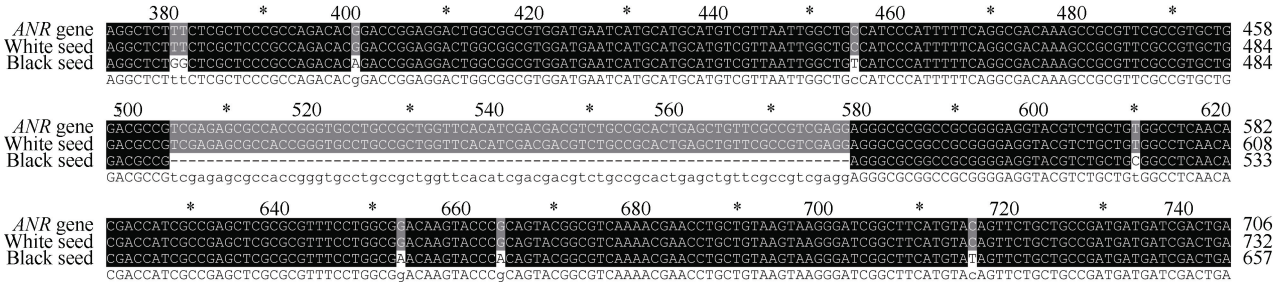


图 5　黑粒、白粒玉米植株中 *Zm00001eb276450* (*ANR*)基因序列比对

Fig. 5　Sequence alignment of *Zm00001eb276450* (*ANR*) gene in black and white kernels

酶促反应中起到了较强的限速作用, 从而使有色花青素得以保留。

MYB 的 *P1* 编码基因在籽粒中促进联苯合成, 促进了花青素合成相关的一系列基因(*CHS*、*CHI*、*DFR* 和 *FLS*)的表达, 且有些 MYB 可在植物中单独起到调控花青素合成的功能[1,26]。本研究中 3 个 *MYB* 基因中仅有 *Zm00001eb276640* 基因在不同样本中表达存在显著性差异, 也说明其在籽粒花青素合成过程中起到了重要的调控作用。WD40 蛋白多与 MYB 和 bHLH 相互作用形成蛋白复合物促进花青素的合成, 多与这 2 个转录因子共同正向调控花青素相关结构基因的表达[27-28]; 此外, 在苹果中也明确了 WD40 与花青素积累呈正相关[29]。

柳政岑[30]利用 BSA-seq 技术成功定位出玉米紫色叶鞘花青素合成调控基因 *bHLH*, 说明在玉米花青素相关基因的定位中 BSA-seq 技术具有可操作性和高效性。本研究预测出的花青素相关基因 *Zm00001eb276640* 和 *Zm00001eb276870* 可能在鲜食玉米籽粒花青素的合成和调控中起到了特异性的作用。在其他玉米花青素相关研究中, 未有对本研究所筛选出的 3 个基因的功能进行系统性的验证分析的报道。为了进一步深入研究这 3 个基因在玉米籽粒花青素合成中的重要功能, 后续将对基因的功能进行更为细致的试验验证, 并通过 SNP 和 InDel 数据设计相关分子标记, 为高花青素鲜食玉米的选育提供辅助标记育种基础。

4　结论

本研究通过构建黑、白籽粒糯玉米表型 F_2 代差异基因混池, 利用 BSA-seq 技术将与玉米籽粒花青素合成相关的基因定位在 6 号染色体 5.3 Mb 区域内。通过数据库的多重比对和生物信息分析筛选出 3 个候选基因, 为花青素合成基因 *Zm00001eb276450* (*ANR*)、转录因子编码基因 *Zm00001eb276870* (*WD40*)和 *Zm00001eb276640* (*MYB*)。结合 qRT-PCR 的初步验证也明确了这 3 个基因为糯玉米籽粒花青素积累相关性状的关键候选基因。本研究为后续分子标记育种的开发提供了较好的基础。

References

[1] Falcone Ferreyra M L, Rius S, Emiliani J, et al. Cloning and characterization of a UV-B-inducible maize flavonol synthase. Plant J, 2010, 62: 77–91.

[2] Sasaki N, Nishizaki Y, Ozeki Y, et al. The role of acyl-glucose in anthocyanin modifications. Molecules, 2014, 19: 18747–18766.

[3] Ming H N, Wang Q, Wu Y, et al. Transcriptome analysis reveals the mechanism of anthocyanidins biosynthesis during grains development in purple corn (*Zea may*s L.). J Plant Physiol, 2021, 257: 153328.

[4] 胡晓丹. 鲜食玉米籽粒花青素合成规律及其对糖脂代谢的调节作用. 华南理工大学博士学位论文, 广东广州, 2021.
Hu X D. Anthocyanin Accumulation and Synthetic Metabolism of Fresh Corn and Its Regulation of Glucose and Lipid Metabolism. PhD Dissertation of South China University of Technology, Guangzhou, Guangdong, China, 2021 (in Chinese with English abstract).

[5] Brueggemann J, Weisshaar B, Sagasser M. A *WD40*-repeat gene from *Malus* x domestica is a functional homologue of *Arabidopsis thaliana* TRANSPARENT TESTA GLABRA1. Plant Cell Rep, 2010, 29: 285–294.

[6] Ben-Simhon Z, Judeinstein S, Nadler-Hassar T, et al. A pomegranate (*Punica granatum* L.) *WD40*-repeat gene is a functional homologue of *Arabidopsis* TTG1 and is involved in the regulation of anthocyanin biosynthesis during pomegranate fruit development. Planta, 2011, 234: 865–881.

[7] Li D, Zhang X C, Xu Y Q, et al. Effect of exogenous sucrose on anthocyanin synthesis in postharvest strawberry fruit. Food Chem, 2019, 289: 112–120.

[8] Gao H N, Jiang H, Cui J Y, et al. Review: the effects of hormones and environmental factors on anthocyanin biosynthesis in apple. Plant Sci, 2021, 312: 111024.

[9] Procissi A, Dolfini S, Ronchi A, et al. Light-dependent spatial and temporal expression of pigment regulatory genes in developing maize seeds. Plant Cell, 1997, 9: 1547–1557.

[10] Goff S A, Cone K C, Chandler V L. Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev, 1992, 6: 864–875.

[11] Petroni K, Cominelli E, Consonni G, et al. The developmental expression of the maize regulatory gene Hopi determines germination-dependent anthocyanin accumulation. Genetics, 2000, 155: 323–336.

[12] Tonelli C, Consonni G, Dolfini S F, et al. Genetic and molecular analysis of Sn, a light-inducible, tissue specific regulatory gene in maize. Mol Gen Genet, 1991, 225: 401–410.

[13] 李军, 王崇敏, 周涛, 等. 血人参原花青素合成关键酶基因 *IsANR*、*IsLAR* 的鉴定及分析. 分子植物育种, 2024, 22: 8048–8056.
Li J, Wang C M, Zhou T, et al. Identification and analysis of *IsANR* and *IsLAR* as key enzyme gene in biosynthesis of procyanidins in *Indigofera stachyoides*. Mol Plant Breed, 2024, 22: 8048–8056 (in Chinese with English abstract).

[14] 杨捷. 东方百合花青素苷生物合成转录调控机制的研究. 北京林业大学博士学位论文, 北京, 2022.
Yang J. Transcriptional Regulation Mechanism of Anthocyanin Biosynthesis in Oriental Hybrid Lily. PhD Dissertation of Beijing Forestry University, Beijing, China, 2022 (in Chinese with English abstract).

[15] Hill J T, Demarest B L, Bisgrove B W, et al. MMAPPR: mutation mapping analysis pipeline for pooled RNA-seq. Genome Res, 2013, 23: 687–697.

[16] 吴蓓, 田耀加, 叶伟忠, 等. 基于 BSA-Seq 筛选鲜食糯玉米抗南方锈病候选基因. 植物生理学报, 2025, 61(1): 61–70.
Wu B, Tian Y J, Ye W Z, et al. Screening candidate genes for resistance to southern corn rust in fresh waxy maize based on BSA-Seq. Plant Physiol J, 2025, 61(1): 61–70 (in Chinese with English abstract).

[17] Hong H T, Netzel M E, O'Hare T J. Anthocyanin composition and changes during kernel development in purple-pericarp supersweet sweetcorn. Food Chem, 2020, 315: 126284.

[18] 李丹. 基于BSA-seq技术的豇豆株高相关基因的初步分析. 江西农业大学硕士学位论文, 江西南昌, 2024.
Li D. Preliminary Analysis of Genes Related to Cowpea Plant Height Based on BSA-seq Technic. MS Thesis of Jiangxi Agricultural University, Nanchang, Jiangxi, China, 2024 (in Chinese with English abstract).

[19] 史君彦. 蓝莓果实内源激素对花色苷生物合成调控机理的研究. 北京林业大学博士学位论文, 北京, 2023.
Shi J Y. Research on the Regulation Mechanism of Endogenous Hormones on Anthocyanin Biosynthesis in Blueberry Fruit. PhD Dissertation of Beijing Forestry University, Beijing, China, 2023 (in Chinese with English abstract).

[20] Lloyd A, Brockman A, Aguirre L, et al. Advances in the MYB-bHLH-WD repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. Plant Cell Physiol, 2017, 58: 1431–1441.

[21] Quattrocchio F, Wing J, van der Woude K, et al. Molecular analysis of the *anthocyanin2* gene of *Petunia* and its role in the evolution of flower color. Plant Cell, 1999, 11: 1433–1444.

[22] 岳茂兰. R2R3-FaMYB5/10 主导的 MYB-bHLH-WD40 复合体对草莓果实类黄酮代谢的调控. 四川农业大学博士学位论文, 四川雅安, 2023.
Yue M L. Regulation of Flavonoids in Strawberry Fruits by R2R3-FaMYB5/10 Dominated MYB-bHLH-WD40 Ternary Complexes. PhD Dissertation of Sichuan Agricultural University, Ya'an, Sichuan, China, 2023 (in Chinese with English abstract).

[23] 张鲲. 控制花生种皮颜色关键基因的鉴定和功能研究. 海南大学博士学位论文, 海南海口, 2023.
Zhang K. Identification and Functional Analysis of Key Genes Controlling Peanut Testa Color. PhD Dissertation of Hainan University, Haikou, Hainan, China, 2023 (in Chinese with English abstract).

[24] Delgado L D, Zúñiga P E, Figueroa N E, et al. Application of a JA-ile biosynthesis inhibitor to methyl jasmonate-treated strawberry fruit induces upregulation of specific *MBW* complex-related genes and accumulation of proanthocyanidins. Molecules, 2018, 23: 1433.

[25] 刘泽. 花青素还原酶调控绿豆芽下胚轴花青苷积累的机理. 南京农业大学硕士学位论文, 江苏南京, 2022.
Liu Z. Mechanism of Anthocyanin Reductase Regulation of Anthocyanin Accumulation in Mung Bean Hypocotyl. MS Thesis of Nanjing Agricultural University, Nanjing, Jiangsu, China, 2022 (in Chinese with English abstract).

[26] Grotewold E, Drummond B J, Ben B W, et al. The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell, 1994, 76: 543–553.

[27] Liang C M, Guo J Y, Chen M X, et al. VvMYBPA2 regulated the accumulation of flavan-3-ols though forming a trimeric complex in 'Zaoheibao' grape. Agriculture, 2022, 12: 1414.

[28] Xie S, Qiao X L, Chen H W, et al. Coordinated regulation of grape berry flesh color by transcriptional activators and repressors. J Agric Food Chem, 2019, 67: 11815–11824.

[29] An X H, Tian Y, Chen K Q, et al. The apple WD40 protein MdTTG1 interacts with bHLH but not MYB proteins to regulate anthocyanin accumulation. J Plant Physiol, 2012, 169: 710–717.

[30] 柳政岑. "十叶紫"玉米紫色叶鞘基因的定位及候选基因挖掘. 四川农业大学硕士学位论文, 四川雅安, 2024.
Liu Z C. QTL Mapping and Candidate Gene Mining of Purple Leaf Shealth in Maize at V10 Stage. MS Thesis of Sichuan Agricultural University, Ya'an, Sichuan, China, 2024 (in Chinese with English abstract).



附图 1　黑粒、白粒和红粒的表型差异
Fig. S1　Phenotypic variation among black, white, and red kernels

附表 1　籽粒颜色遗传模型分析
Table S1　Genetic model analysis of kernel color

遗传模型 Genetic model	名称 Name	实测值 Observed value (O)	理论值 Expected value (E)	O–E	$(O–E)^2$	$(O–E)^2$ / E	χ^2
相加型变式 (9∶6∶1) Additive Model (9∶6∶1)	黑粒 Black kernels	342	334.125	7.875	62.016	0.186	10.606
	红粒 Red kernels	198	222.750	–24.750	612.563	2.750	
	白粒 White kernels	54	37.125	16.875	284.766	7.670	
	总计 Total	594	594.000	—	—	—	
显性上位变式 (12∶3∶1) Dominant epistatic model (12∶3∶1)	黑粒 Black kernels	342	445.500	–103.500	10,712.250	24.045	99.091
	红粒 Red kernels	198	111.375	86.625	7503.891	67.375	
	白粒 White kernels	54	37.125	16.875	284.766	7.670	
	总计 Total	594	594.000	—	—	—	
重叠作用变式 (15∶1) Duplicate gene model (15∶1)	有色 Colored	540	556.875	–16.875	284.766	0.511	0.511
	无色 Colorless	54	37.125	16.875	284.766	7.670	
	总计 Total	594	594.000	—	—	—	

实测值(O)代表实际统计玉米穗籽粒颜色数量; 理论值(E)代表在不同遗传模型下不同表型的分布比例; χ^2 代表概率分布, 通过卡方对照表判断遗传模型的可信度。

O (observed value): the actual counted number of kernels for each color phenotype on the corn ear; E (expected value): the expected distribution ratio of different phenotypes under a given genetic model; χ^2: a test statistic for which the value is compared against a chi-square distribution table to assess the goodness-of-fit of the genetic model.