

黄秋葵嫩果多糖含量性状全基因组关联分析

戴志刚¹, 孙健^{1,2}, 程玉静³, 陆旭鹏², 李世玉², 谢冬微^{2*}, 粟建光^{1*}

¹中国农业科学院麻类研究所, 湖南 长沙

²南通大学生命科学学院, 江苏 南通

³江苏沿江地区农业科学研究所, 江苏 南通

收稿日期: 2021年8月27日; 录用日期: 2021年10月26日; 发布日期: 2021年11月4日

摘要

植物多糖是具有保健功能的生物活性成分, 黄秋葵鲜嫩果具有很高的多糖含量。对黄秋葵多糖含量进行全基因组关联分析(Genome-wide Association Study, GWAS), 挖掘相关基因位点, 可以为利用分子育种手段培育高多糖黄秋葵新品种奠定理论和材料基础。对180份黄秋葵种质进行GBS建库测序(Genotyping by sequencing, GBS), 结合多环境下黄秋葵多糖含量的表型数据进行全基因组范围的关联分析, 以明确与多糖含量显著关联的SNP位点, 挖掘位点附近可能与多糖含量相关的基因。通过4个环境共检测到46个显著关联的SNPs位点, 其中Cluster1671-67622和Cluster86-210279位点至少在两个环境下稳定出现。在这2个显著关联的位点附近的100 Mb区域内发掘出黄秋葵多糖含量性状关系密切的候选基因, 经过分析筛选确定UGT-糖基转移酶和MYB转录因子为候选基因。UGT-糖基转移酶和MYB转录因子与黄秋葵嫩果中多糖的合成、积累和分解紧密相关。本研究结果为阐明多糖的遗传调控机制提供了新的视角, 亦为黄秋葵嫩果多糖含量的遗传改良奠定了一定的理论基础。

关键词

黄秋葵, 多糖含量, GWAS, SNP, 候选基因

Genome-Wide Association Analysis of Polysaccharide Content in Young Fruits of Okra

Zhigang Dai¹, Jian Sun^{1,2}, Yujing Cheng³, Xupeng Lu², Shiyu Li², Dongwei Xie^{2*}, Jianguang Su^{1*}

¹Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences, Changsha Hunan

²School of Life Science, Nantong University, Nantong Jiangsu

*通讯作者。

文章引用: 戴志刚, 孙健, 程玉静, 陆旭鹏, 李世玉, 谢冬微, 粟建光. 黄秋葵嫩果多糖含量性状全基因组关联分析[J]. 植物学研究, 2021, 10(6): 749-760. DOI: 10.12677/br.2021.106094

Abstract

Plant polysaccharide is a biological active component with health care function. The fresh and tender fruit of okra has a high content of polysaccharide. Genome-wide Association Study (GWAS) on the polysaccharide content of okra and the mining of relevant gene loci can lay a theoretical and material foundation for the cultivation of new varieties of okra with high polysaccharide by molecular breeding methods. Based on the accurate determination of polysaccharide content of 180 okra Germplasms in multi environment, the whole genome sequencing (Genotyping by sequencing, GBS) of the tested materials was carried out, and the genome-wide association analysis was carried out combined with the phenotypic data of polysaccharide content in multi environment, to identify the SNP sites significantly associated with polysaccharide content. Meanwhile, excellent genes which may be related to the content of polysaccharide near the sites were mined. A total of 46 SNPs were detected in four environments, and two of them were selected as candidate sites. In the 100 Mb region near these two significant loci, candidate genes closely related to the polysaccharide content traits of okra were found. UGT-glycosyltransferase and MYB transcription factor were finally identified as candidate genes. UGT-glycosyltransferase and MYB transcription factor are closely related to the synthesis, accumulation, and decomposition of polysaccharides in okra tender fruit. The results of this study provided a new perspective to clarify the genetic regulation mechanism of polysaccharides, and laid a theoretical basis for the genetic improvement of polysaccharide content in okra tender fruit.

Keywords

Okra, Polysaccharide Content, GWAS, SNP, Candidate Gene

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

黄秋葵(*Abelmoschus esculentus* L.)为锦葵科(*Malvaceae*)秋葵属(*Abelmoschus Medicus*)一年生草本植物。其用途十分广泛,嫩果、花均可供食用,种子可榨油,丰富的茎秆纤维也具有较高利用价值[1]。黄秋葵的果荚中含有丰富的糖类物质,主要是半乳聚糖、阿拉伯聚糖、果胶类多糖、及少量的糖蛋白混合物。多糖作为天然的生物大分子,具有多种生物活性,如提高免疫力、增强身体耐力、抗疲劳、降低血糖,以及抗肿瘤等[2]。因此,黄秋葵因其较高的营养价值和显著的保健和食疗效果越来越受到人们的关注和喜爱[3] [4]。培育营养价值高和保健功效好的新品种是目前黄秋葵的育种目标和方向,当前可供黄秋葵育种利用的多糖功能性主效基因匮乏,严重影响了高多糖黄秋葵品种的培育的发展。目前国内外对黄秋葵多糖的提取及加工工艺报道较多,但对黄秋葵高多糖种质资源的鉴定评价与关键基因挖掘并无深入研究。

近些年,随着黄秋葵所含功能营养保健成分不断被发掘,黄秋葵作为一种保健型蔬菜已经在欧美、

东南亚及非洲等地大面积种植, 相关基础研究也引起了学者们的广泛关注。对黄秋葵的基础研究主要集中在遗传育种[5] [6]、栽培生理[7] [8]、营养成分[9]和药用功效[10] [11]等方面, 而有关分子生物学的研究特别是多糖含量相关功能基因研究方面的研究甚少。王旭等(2014)利用克隆和 RT-PCR 技术, 首次获得了黄秋葵查尔酮合成酶基因(CHS)cDNA 的全长序列[12]。张少平等(2017)通过高通量测序掌握了紫色黄秋葵叶片基因表达的情况, 对紫色黄秋葵叶片的代谢途径进行了初探[13]。李和平等(2018)利用 RNA-seq 技术分析黄秋葵果实基因表达情况, 以及了解各代谢途径关键基因的表达模式, 为黄秋葵遗传分析提供了一定的依据[14]。对于植物多糖含量相关基因的挖掘已有一些报道, 韩文韬(2019)以海带全基因组和转录组数据为基础, 利用生物信息学方法, 在整个生物界范围内对岩藻多糖合成途径中的重要基因家族进行比较基因组学研究, 推测 *FucXylTs* 可能参与了褐藻中岩藻多糖的侧链的生物合成[15]。范华(2017)发现在盐胁迫条件下, 发菜多糖分泌量增加。对盐胁迫及正常培养条件下的发菜样品进行转录组测序分析, 筛选出部分可能参与盐胁迫下发菜多糖代谢相关的差异表达基因[16]。耿腊(2021)等以前期收集的全球 119 份大麦基因型为材料, 对不同大麦材料籽粒 β -葡聚糖含量进行 GWAS 分析, 共鉴定到 10 个与糖合成、转运及分解相关的酶类基因[17]。

但是, 对黄秋葵果实多糖含量相关基因的全基因组范围挖掘至今还未见报道。本研究拟利用高通量测序技术对 180 份黄秋葵核心种质进行全基因组测序, 对其群体结构和遗传进化进行深入分析, 并结合多环境下多糖表型数据进行候选基因挖掘, 并开发功能标记, 获得高多糖黄秋葵新材料。为揭示黄秋葵多糖合成的分子机制提供重要线索, 为分子育种培育高多糖黄秋葵新品种奠定理论和材料基础, 对加快开发利用黄秋葵产品、丰富区域性菜篮子工程、提高黄秋葵产品的市场竞争力具有重要意义。

2. 材料与方法

2.1. 试验材料

通过国内考察收集和国外引种, 收集保存黄秋葵 180 份核心种质, 其中来自国内 8 个省(自治区)的 108 份, 来自国外 10 个国家和地区的 72 份。分别于 2019 年 5 月和 2020 年 5 月播种于湖南省望城白箬铺基地(28°15'N, 112°43'E) (该环境以下简称为 2019 白箬铺), 江苏省南通市南通大学试验基地(31°41'N, 120°12'E) (该环境以下简称为 2020 南通), 湖南省长沙市中国农业科学院麻类研究所沅江实验基地(28°50'N, 112°20'E) (该环境以下简称为 2019 沅江和 2020 沅江)试验基地。每份材料按双行区种植, 3 次重复, 行长 10 m, 行宽 1 m, 行距 30 cm, 穴播, 穴距 50 cm, 田间管理同大田生产。取花后 6 d 同一结位的黄秋葵嫩果烘干测定多糖含量。

2.2. 试验方法

2.2.1. 黄秋葵嫩果多糖含量红外光谱测定

分别称取冻干的黄秋葵果实样品 1 mg, 加入 100 mg 烘干平衡重的 KBr, 置于研钵中研磨至极细, 用压片机压片, 用红外光谱仪在 4000~400 cm^{-1} 范围内进行红外扫描。

2.2.2. 基因组 DNA 提取

采集 180 份黄秋葵核心种质材料的幼嫩叶片于液氮中速冻, 用于基因组总 DNA 的提取。DNA 浓度稀释至 50~100 ng/ μL , NanoDrop 2000 紫外可见光度计检测 DNA 浓度和质量, 确保符合测序分型的要求。

2.2.3. 基因组修饰

利用限制性内切酶 Hae III 对黄秋葵基因组 DNA 进行酶切。酶切体系: 黄秋葵基因组 DNA 500 ng, NEB Buffer 41 μL , 限制性内切酶 0.12 μL , 补水至 50 μL , 37°C 酶切反应 15 h。酶切结束后, 用纯化试剂

盒纯化,并用 50 μL EB 回溶。对酶切产物进行平端化修复,并且对 5'末端进行磷酸化的修饰。在 5'磷酸化的平端片段 3'末端添加一个“A”,以便后续与 5'端突出的“T”Solexa 接头进行互补的连接,从而提高了接头的连接效率,同时阻止了 Solexa 接头的自连。反应结束后再用纯化试剂盒纯化,并于 10 μL EB 回溶。最后连接 Solexa 的接头,从而使不同的 DNA 片段具有了相同的末端序列,从而可以在后续顺利的杂交到流路池上,进行扩增、测序。

2.2.4. PCR 扩增与序列测序

根据酶切预测软件 PProC (<http://www.paproc.de/>)的分析结果,切胶回收 500~580 bp 的 DNA 片段,进行 PCR 扩增。扩增体系如下:纯化获得的 DNA 产物 8 μL , Forward-引物 1.5 μL , Reverse-引物 1.5 μL , DNA 聚合酶 20 μL , 补充 dd 水至 40 μL 。扩增程序如下:95 $^{\circ}\text{C}$ 预变性 30 s;95 $^{\circ}\text{C}$ 变性 10 s,62 $^{\circ}\text{C}$ 复性 30 s,72 $^{\circ}\text{C}$ 延伸 40 s,20 个循环;最后再经 72 $^{\circ}\text{C}$ 延伸 6 min。反应结束后使用纯化试剂盒纯化,再于 30 μL EB 回溶。最后将纯化获得的 PCR 产物进行精细定量,然后在芯片表面进行桥式 PCR,使 DNA 片段扩增成为单分子的 DNA 簇。单分子的 DNA 簇扩增完成后,将所得产物移入 IlluminaHiSeq2000 中,进行测序。

2.2.5. 高通量测序

根据 DNA 条形码序列,将原始的 reads 按照其个体分类,并剔除低质量的 reads,两端均比对到基因组上的为可靠的 reads,其可以用以定义 SNP 标签。根据序列比对和错误校正的结果,选取样本平均深度达到 4 以上的组别,用来定义 SNP 标记。统计基因组上的每 100 Kb 范围内 SNP 标记个数,得到 SNP 标记在 scaffolds 上的分布,进行基于 180 个样本的群体内部 SNP 检测,统计每 100 Kb 范围内 SNP 个数。

2.2.6. 群体结构分析

为了估计样本的群体结构,根据 LD 分析结果,选取相互间不连锁的 SNP 标记,采用 STRUCTURE2.2 对黄秋葵材料进行了基于数学模型的亚群划分,并获得每个材料的 Q 值。

2.2.7. 全基因组关联分析

利用 Plink 对获得的 SNP 进行质量控制,标准是剔除最小检出率 < 85 % 和最小等位基因频率 < 5 % 的 SNP。使用 SPSS19.0 对多糖含量数据进行统计分析,利用 Admixture [18]分析群体结构,使用 TASSEL3.0 [19]的一般线性模型(GLM)和混合线性模型(MLM)进行全基因组关联分析,通过 SPAGeDi 1.4 软件计算样品间亲缘关系 K。使用 TASSEL 3.0 软件生成各性状全基因组关联分析的 Quantile-Quantile (QQ)图和曼哈顿图。

3. 结果与分析

3.1. 不同环境下多糖含量的描述性统计

在表型鉴定数据的基础上,结合地理来源等信息,采取逐级取样策略,构建了 180 份不同地理来源的黄秋葵自然群体。为解决目前育成品种遗传基础狭窄和满足多用途应用提供了材料保障。在 4 个不同环境条件下,黄秋葵嫩果中的多糖含量的最小值为 0.9834 mg/g,最大值为 3.3315 mg/g,变异系数最大为 16.62 %。以上结果表明供试黄秋葵种质资源的变异非常丰富,所得结果见表 1。

3.2. 群体结构和进化分析

分析 180 份黄秋葵资源的群体聚类和群体结构(图 1 和图 2),分别假设研究群体的分群数(K 值)为 0~10,进行聚类,并对聚类结果进行交叉验证,根据交叉验证错误率的谷值确定最优分群数,发现当 K = 3 时 ΔK 值最大,所以该群体可以分为 3 个亚群,表明所研究群体最可能来自于 3 个原始祖先。基于群体

SNP, 通过 MEGA 5.0 软件的 neighbor-joining 算法, 对 180 份核心种质进行了进化分析(图 3), 结果显示, 180 份材料可以分为 3 大类, 第 1 类主要为日本引进品种及含有其血缘的种质(红色表示); 第 2 类主要为印度及东南亚引进品种及含有其血缘的种质(蓝色表示); 第 3 类主要为国内野生种及相关种质(黄色表示)。

Table 1. Variation of polysaccharide content in okra core germplasm under different environments

表 1. 不同环境下黄秋葵核心种质多糖含量的变异情况

年份 Year	地点 Location	最小值(mg/g) Minimum value (mg/g)	最大值 (mg/g) Maximum value (mg/g)	极差(mg/g) Range (mg/g)	平均数 (mg/g) Average value (mg/g)	标准差 Standard deviation	变异系数(%) Coefficient of variation (%)
2019	沅江	1.0834	3.2369	2.1535	2.2952	0.3814	16.62
2019	望城	1.0216	3.3315	2.3099	2.3015	0.3992	16.34
2020	沅江	0.9834	3.0396	2.0562	2.0115	0.3574	16.42
	南通	1.0126	3.2153	2.2027	2.1139	0.3862	16.61

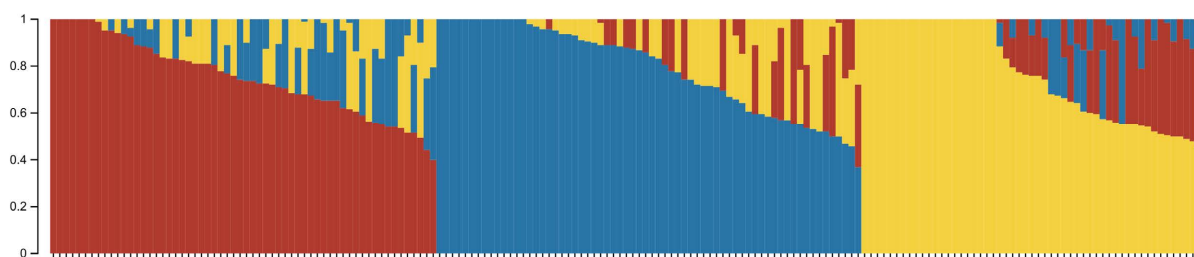


Figure 1. Population structure of 180 core germplasm of okra

图 1. 180 份黄秋葵核心种质的群体结构

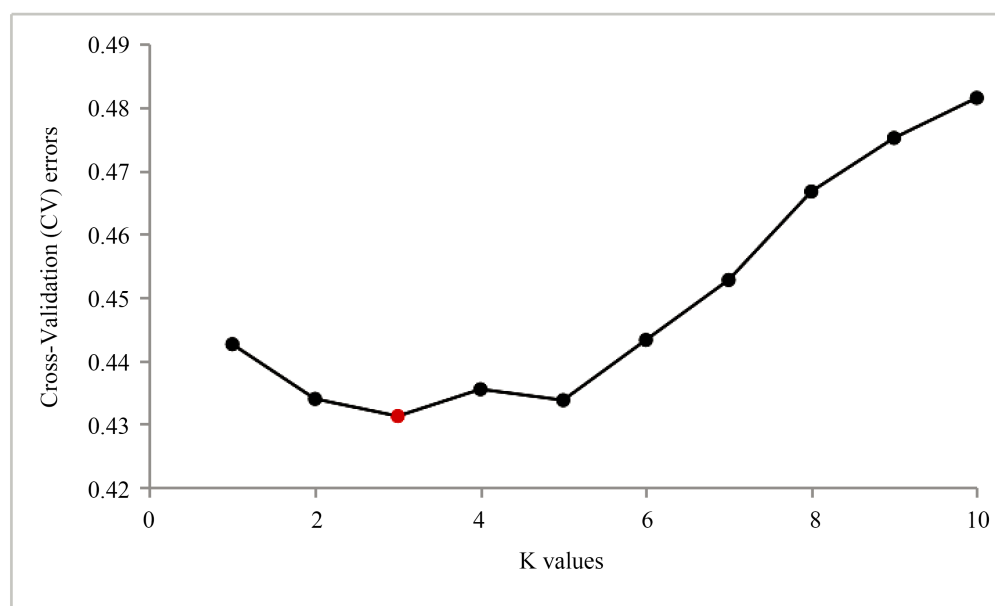


Figure 2. Selection of K value in population structure of 180 okra core germplasm

图 2. 180 份黄秋葵核心种质群体结构的 K 值选择

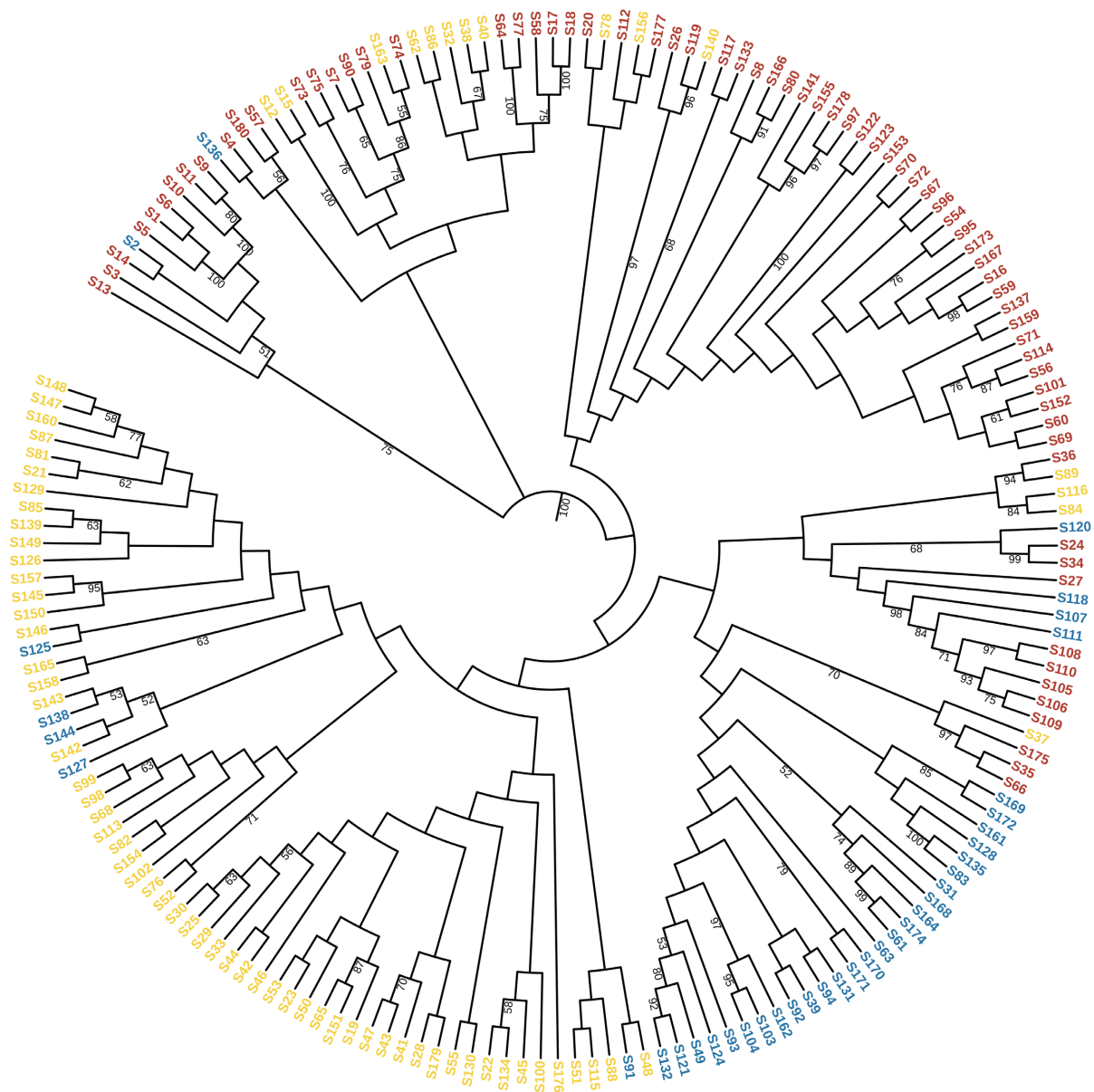


Figure 3. Evolutionary analysis of germplasm resources of okra

图 3. 黄秋葵种质资源的进化分析

3.3. 黄秋葵嫩果多糖含量的全基因组关联分析

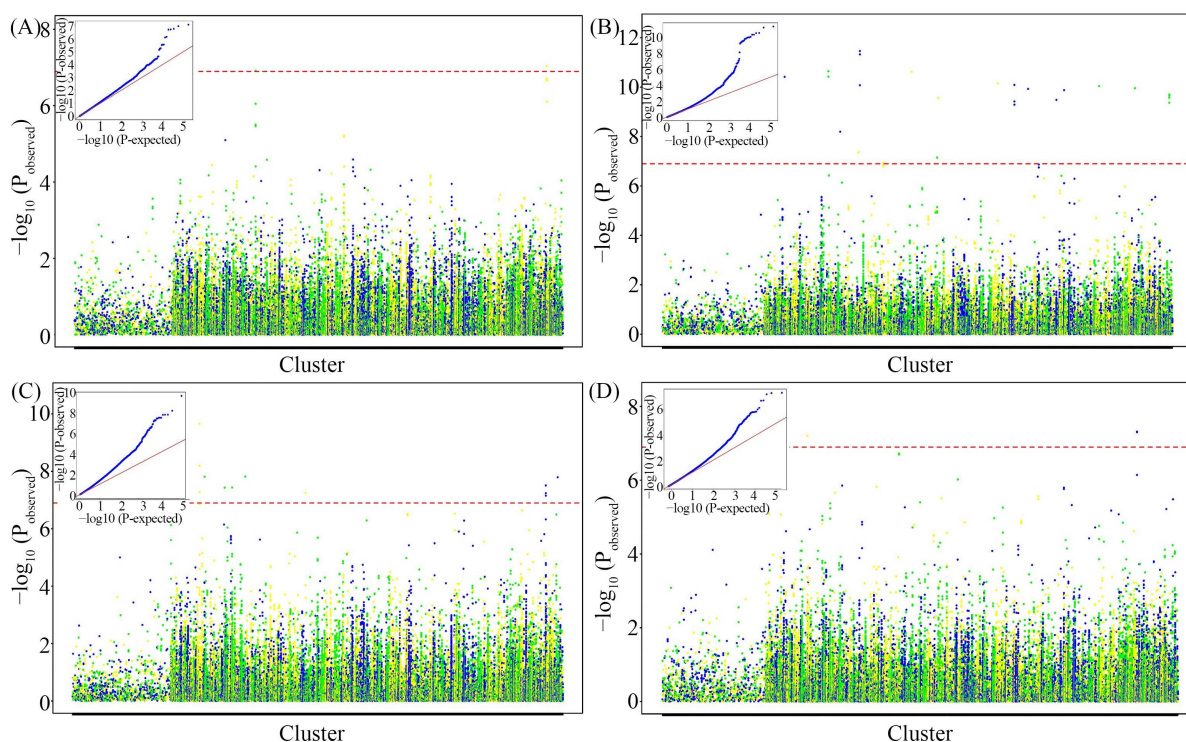
3.3.1. 阈值确定

利用 GEC 软件计算 579,654 个 SNPs 的有效标记数(En), 结果表明, En 为 261,342, 软件建议的 $P = 1.16 \times 10^{-7}$ ($1/En$), 该值作为关联分析的显著阈值。

3.3.2. 关联分析结果

对 4 个环境下的嫩果多糖含量的 GWAS 结果进行分析, 针对不同模型对假阳性的控制来看, Q 模型最差, K 模型和 Q + K 模型优于 Q 模型。从对假阴性的控制效果看, Q + K 模型比 K 模型稍过严格。因此, 选用表现相对好的 K 模型的 GWAS 结果用于后续的位点挖掘和基因预测。基于 K 模型, 在 $P \leq 1.16$

$\times 10^{-7}$ 下, 4 个环境一共检测到 46 个 SNPs 与嫩果多糖含量显著关联, 涉及到 28 个位点(图 4, 表 2)。在 2019 年望城环境下, 共鉴定到 2 个显著 SNP, 其曼哈顿图和 QQ 图见图 4(A)。距离这 2 个 SNPs 位点最近的基因分别是油菜黄质脱氢酶和 GTP 结合蛋白(表 2)。在 2019 年沅江环境下, 共鉴定到 27 个显著 SNP 其曼哈顿图和 QQ 图见图 4(B), 距离这 27 个位点最近的基因功能分别是 GRAS 家族基因(1 个)、胞质烯醇化酶(2 个)、UGT-糖基转移酶(4 个)、蛋白二硫异构体酶(1 个)、植物细胞器识别域(3 个)、P450 蛋白(1 个)、转运蛋白(3 个)、信号转导蛋白(1 个)、核糖体蛋白(1 个)、DNA 结合蛋白(5 个)及未知功能基因(5 个)(表 2)。在 2020 年沅江环境下, 共鉴定到 14 个显著 SNP 其曼哈顿图和 QQ 图见图 4(C), 距离这 14 个位点最近的基因功能分别是 UGT-糖基转移酶(1 个)、翻译后修饰蛋白(1 个)、光系统 I 反应蛋白(1 个)、DSBA 氧化还原酶(1 个)、信号转导蛋白(1 个)、GDSL 脂肪酶(1 个)、MYB 转录因子(5 个)及未知功能基因(3 个)(表 2)。在 2020 年南通环境下, 共鉴定到 3 个显著 SNP 其曼哈顿图和 QQ 图见图 4(D), 距离这 3 个位点最近的基因功能分别是 GABA 转录因子(1 个), MYB 转录因子(1 个)及细胞分裂蛋白(1 个)。



A: 2019 望城; B: 2019 沅江; C: 2020 沅江; D: 2020 南通。

Figure 4. Manhattan and Quantile-Quantile diagrams of different environments

图 4. 不同环境条件下的曼哈顿图和 Quantile-Quantile 图

3.3.3. 候选基因确定

由于该群体的连锁不平衡衰减距离是 50 Mb ($r^2 = 0.1$), 因此, 在峰值 SNP 的上下游各 50 Mb, 即 100 Mb 范围内搜索候选基因。在每个关联显著的 SNPs 位点周围 100 Mb 的区域内寻找可能的候选基因, 并且结合候选基因在多个环境下的出现情况, 最终确定了 2 个候选基因。Cluster1671-67622 位点在 2019 年沅江和 2020 年沅江均被检测到, 且其最近的基因功能被注释为 UGT-糖基转移酶。另外 Cluster86-210279 在 2020 年沅江和 2020 年南通均被检测到, 该位点最近的基因被注释为 MYB 转录因子。因此初步确定 UGT-糖基转移酶和 MYB 转录因子为候选基因(表 2)。

Table 2. Significantly associated locus information and genes in different environments
表 2. 不同环境下显著关联位点及基因功能注释信息

环境 Environment	SNP 位点 SNP position	基因簇 Cluster	位置 Location (bp)	P 值 P-value	基因功能注释 Gene function
2019 望城	Cluster179-179593	Cluster179	179593	1.22E-07	油菜黄质脱氢酶
	Cluster866-116645	Cluster866	116645	9.18E-08	GTP 结合蛋白
2019 沅江	Cluster116-30201	Cluster116	30201	3.86E-11	GRAS 家族
	Cluster156-1203677	Cluster156	1203677	2.29E-11	胞质烯醇化酶
	Cluster156-1203690	Cluster156	1203690	3.78E-11	胞质烯醇化酶
	Cluster1671-67434	Cluster1671	67434	6.39E-09	UGT-糖基转移酶
	Cluster185-48910	Cluster185	48910	4.35E-08	蛋白二硫异构体酶
	Cluster1863-545	Cluster1863	545	8.39E-11	植物细胞器识别域
	Cluster1863-546	Cluster1863	546	3.48E-12	植物细胞器识别域
	Cluster1863-554	Cluster1863	554	4.75E-12	植物细胞器识别域
	Cluster212-601232	Cluster212	601232	1.16E-07	P450 蛋白
	Cluster2420-96	Cluster2420	96	2.41E-11	未知
	Cluster272-413070	Cluster272	413070	7.14E-08	转运蛋白
	Cluster272-413095	Cluster272	413095	7.14E-08	转运蛋白
	Cluster273-435350	Cluster273	435350	2.74E-10	转运蛋白
	Cluster353-773806	Cluster353	773806	7.04E-11	信号转导蛋白
	Cluster380-827904	Cluster380	827904	5.14E-10	UGT-糖基转移酶
	Cluster380-828062	Cluster380	808062	8.18E-11	UGT-糖基转移酶
	Cluster380-827881	Cluster380	827881	3.82E-10	UGT-糖基转移酶
	Cluster402-586	Cluster402	586	1.19E-10	未知
	Cluster45-814385	Cluster45	814385	3.29E-10	未知
	Cluster464-1169775	Cluster464	1169775	1.33E-10	核糖体蛋白
Cluster5262-149	Cluster5262	149	9.09E-11	未知	
Cluster707-522	Cluster707	522	1.10E-10	未知	
Cluster977-784147	Cluster977	784147	2.69E-10	DNA 结合蛋白	
Cluster977-784164	Cluster977	784164	2.21E-10	DNA 结合蛋白	
Cluster977-784189	Cluster977	784189	2.00E-10	DNA 结合蛋白	
Cluster977-784191	Cluster977	784191	4.30E-10	DNA 结合蛋白	
Cluster977-784225	Cluster977	784225	2.65E-10	DNA 结合蛋白	

Continued

2020 沅江	Cluster1671-67622	Cluster1671	67622	2.22E-10	UGT-糖基转移酶
	Cluster1253-31901	Cluster1253	31901	6.40E-09	翻译后修饰蛋白
	Cluster1253-46431	Cluster1253	46431	5.37E-08	光系统 I 反应蛋白
	Cluster1311-53390	Cluster1311	53390	1.55E-08	DSBA 氧化还原酶
	Cluster1491-408169	Cluster1491	408169	3.71E-08	信号转导蛋白
	Cluster156-586707	Cluster156	586707	3.66E-08	未知
	Cluster1692-872	Cluster1692	872	1.53E-08	未知
	Cluster234-21551	Cluster234	21551	5.68E-08	GDSL 脂肪酶
	Cluster86-210246	Cluster86	210246	7.05E-08	MYB 转录因子
	Cluster86-210256	Cluster86	210256	3.10E-08	MYB 转录因子
	Cluster86-210279	Cluster86	210279	5.79E-08	MYB 转录因子
	Cluster86-210291	Cluster86	210291	5.79E-08	MYB 转录因子
	Cluster86-210630	Cluster86	210630	3.10E-08	MYB 转录因子
Cluster958-16256	Cluster958	16256	1.61E-08	未知	
2020 南通	Cluster137-111000	Cluster137	111000	6.28E-08	GABA 转录因子
	Cluster86-210279	Cluster86	210279	5.65E-08	MYB 转录因子
	Cluster687-123908	Cluster687	123908	5.01E-08	细胞分裂蛋白

4. 讨论

GWAS 以遍布整个基因组的单核苷酸多态性分子标记为基础,对复杂性状进行直接的关联分析,被认为是一种确定影响重要性状的分子标记的有效方法[20] [21]。随着测序技术的飞速发展,利用这种普通分子标记进行关联分析研究呈现出诸多缺点。基于 SLAF-seq 的 GWAS 已在玉米、水稻、大豆等多种作物中全面开展[22] [23] [24] [25]。基于 SLAF-seq 测序的 GWAS 挖掘植物优异等位基因技术已经非常成熟,效果十分显著。与水稻、大豆、玉米、棉花等作物相比,黄秋葵 GWAS 及优异等位基因挖掘的研究还远远不够,并且基于全基因组测序的 GWAS 在黄秋葵中还未见报道。本研究首次利用二代测序技术/SLAF-seq 在全基因组范围内筛选与黄秋葵多糖含量性状显著关联的 SNPs 位点,并确定相应的候选基因。

本研究绘制了黄秋葵多糖含量性状不同环境条件下的 QQ 图以检验群体校正的准确性,QQ 图结果显示,观测值与期望值基本相符,仅在末端存在少许偏离位点,说明群体结构校正的结果准确可靠,关联分析结果并没有因为群体分层而产生假阴性。本研究利用 K 模型及 4 个环境下的表型数据对黄秋葵的多糖含量性状进行 GWAS,共筛选得到 46 个效应位点与黄秋葵多糖含量性状显著关联($P < 1.16E-07$)。利用两个环境条件下同时关联到的显著关联的 SNPs 位点进行筛选,最终得到与多糖含量显著关联的 SNPs 位点有 2 个。由于这 2 个 SNPs 位点可以稳定出现,说明它们分别对黄秋葵的多糖含量性状具有重要的影响,可以作为影响黄秋葵多糖含量性状的重要候选遗传标记。对于植物多糖及单糖的全基因组关联分析报道寥寥无几,其中耿腊等(2021)利用 MLM 和 GLM 模型针对大麦籽粒 β -葡聚糖含量进行全基因组关联分析,合并 2 个模型重叠位点后共得到 44 个显著位点,其中 *HORVU5Hr1G022710* 基因在 2 个模型、

2个地点均被鉴定到,故被认为是与 β -葡聚糖含量显著相关的候选基因[17]。该研究结果并未与本研究筛选到相同的候选基因,究其原因主要是由于两者研究的物种和性状都存在差别。

在关联显著的候选 SNPs 标记周围 100Mb 区域内寻找可能的候选基因,结果找到的 *UGT*-糖基转移酶和 *MYB* 转录因子等 2 个基因,初步筛选为黄秋葵多糖含量性状的候选基因。在整个多糖产生的代谢途径中, *UGT*-糖基转移酶参与多糖重复单元的形成,其主要作用于活性糖基,使其从糖基供体转移至糖基受体,最终形成糖苷键,因而对于多糖的生物合成具有重要的调控作用[26] [27],随着分子生物学的快速发展,目前借用基因工程手段,包括基因的超表达及基因敲除等,多个物种都对糖基转移酶的功能进行了深入的研究[28]。本研究恰好筛选到 *UGT*-糖基转移酶基因,因而本研究将糖基转移酶作为重要的候选基因之一。*MYB* 转录因子是植物中数目庞大、功能注释最全面的转录因子之一,在很多生命及代谢过程中扮演着非常重要的角色, *MYB* 转录因子已成为目前植物表达网络调控及基因功能研究的热点。植物 *MYB* 转录因子的主要生物学功能有参与植物的初生及次生代谢过程[29] [30],参与植物细胞形态和模式建成[31] [32],参与植物多个生长发育过程[33] [34] [35],参与植物逆境胁迫应答[36] [37] [38]。而本研究发现 *MYB* 转录因子很有可能参与植物多糖的合成,具体功能有待进一步的研究验证。

5. 结论

利用 SLAF 测序的方法对 180 份黄秋葵种质资源的多糖含量性状进行 GWAS,通过 4 个环境共检测到 46 个显著关联的 SNPs 位点,其中 2 个位点被筛选为候选位点。在这 2 个显著位点附近的 100 Mb 区域内发掘出黄秋葵多糖含量性状关系密切的候选基因,经过分析最终确定 *UGT*-糖基转移酶和 *MYB* 转录因子为候选基因。以上显著关联的 SNPs 位点和候选基因信息可以作为黄秋葵多糖含量性状改良的理论依据。

基金项目

感谢南通市科技项目(JC2020156);湖南省自然科学基金面上项目(2021JJ30768);中国农业科学院麻类研究所横向课题(IBFC-KYSY-080);国家自然科学基金项目(31801409)的资助。

参考文献

- [1] 卢隆杰, 苏浓, 岳森. 菜药花兼用型植物——黄秋葵[J]. 特种经济动植物, 2004(10): 36.
- [2] Yoshida, T. (2001) Synthesis of Polysaccharides Having Specific Biological Activities. *Progress in Polymer Science*, **26**, 379-441. [https://doi.org/10.1016/S0079-6700\(00\)00045-9](https://doi.org/10.1016/S0079-6700(00)00045-9)
- [3] Adalakun, O.E., Oyelade, O.J., Ade-Omowaye, B.I.O., Adeyemi, I.A. and Van de Venter, M. (2009) Chemical Composition and the Antioxidative Properties of Nigerian Okra Seed (*Abelmoschus esculentus* Moench) Flour. *Food and Chemical Toxicology*, **47**, 1123-1126. <https://doi.org/10.1016/j.fct.2009.01.036>
- [4] Ameena, K., Dilip, C., Saraswathi, R., Krishnan, P.N. and Sankar, C. and Simi, S.P. (2010) Isolation of the Mucilages from Hibiscus Rosasinensis Linn. And Okra (*Abelmoschus esculentus* L.) and Studies of the Binding Effects of the Mucilages. *Asian Pacific Journal of Tropical Medicine*, **3**, 539-543. [https://doi.org/10.1016/S1995-7645\(10\)60130-7](https://doi.org/10.1016/S1995-7645(10)60130-7)
- [5] Mehta, N., Asati, B.S. and Mamidwar, S.R. (2007) Heterosis and Gene Action in Okra. *Bangladesh Journal of Agricultural Research*, **32**, 421-432. <https://doi.org/10.3329/bjar.v32i3.544>
- [6] Wammanda, D.T., Kadams, A.M. and Jonah, P.M. (2010) Combining ability Analysis and Heterosis in a Diallel Cross of Okra (*Abelmoschus esculentus* (L.) Moench). *African Journal of Agricultural Research*, **5**, 2108-2115.
- [7] Olayiwola, M.O. and Ariyo, O.J. (2013) Relative Discriminatory Ability of GGE Biplot and YSi in the Analysis of Genotype×Environment Interaction in Okra (*Abelmoschus esculentus*). *International Journal of Plant Breeding and Genetics*, **7**, 146-158. <https://doi.org/10.3923/ijpb.2013.146.158>
- [8] Islam, M., Hossain, N., Alamgir, M. and Kibria, G., (2014) Comparative Effects of Biogas Plant Residues, Poultry Manure and Inorganic Fertilizer on Growth and Yield of Ladies Finger. *IOSR Journal of Agriculture and Veterinary Science*, **7**, 2319-2372.

- [9] 黄阿根, 陈学好, 高云中, 车婧. 黄秋葵的成分测定和分析[J]. 食品科学, 2007, 28(10): 451-455.
- [10] Mairuae, N., Connor, J.R., Lee, S.Y. and Cheepsunthorn, P. and Tongjaroenbuangam, W. (2015) The Effects of Okra (*Abelmoschus esculentus* Linn.) on the Cellular Events Associated with Alzheimer's Disease in a Stably Expressed HFE Neuroblastoma SH-SY5Y Cell Line. *Neuroscience Letters*, **603**, 6-11. <https://doi.org/10.1016/j.neulet.2015.07.011>
- [11] Monte, L.G., Santi-Gadelha, T., Reis, L.B., Braganhol, E., Prietsch, R.F., Dellagostin, O.A., Lacerda, R.R.E., Gadelha, A.A., Concei, F.R.O. and Pinto, L.S. (2014) Lectin of *Abelmoschus esculentus* (Okra) Promotes Selective Antitumor Effects in Human Breast Cancer Cells. *Biotechnology Letters*, **36**, 461-469. <https://doi.org/10.1007/s10529-013-1382-4>
- [12] 王旭, 韩春乐, 周亚楠, 王春国, 宋文芹, 陈成彬. 黄秋葵查尔酮合成酶基因 AeCHS 的克隆与表达分析[J]. 植物遗传资源学报, 2014, 15(3): 561-567.
- [13] 张少平, 邱珊莲, 郑云云, 张帅, 吴松海, 何炎森, 郑开斌. 紫色黄秋葵转录组功能基因测序及分析[J]. 核农学报, 2017, 31(4): 643-653.
- [14] 李和平, 姚运法, 练冬梅, 赖正锋, 洪建基. 黄秋葵果实转录组测序及分析[J]. 生物技术通报, 2018, 34(3): 121-127.
- [15] 韩文韬. 海带岩藻多糖合成关键基因家族的全基因组鉴定和系统发育研究[D]: [硕士学位论文]. 上海: 上海海洋大学, 2019.
- [16] 范华. 盐胁迫下发菜多糖代谢相关差异表达基因的克隆与原核表达[D]: [硕士学位论文]. 西安: 陕西科技大学, 2017.
- [17] 耿腊, 黄业昌, 李梦迪, 谢尚耿, 叶玲珍, 张国平. 大麦籽粒 β -葡聚糖含量的全基因组关联分析[J]. 作物学报, 2021, 47(7): 1205-1214.
- [18] Alexander, D.H., Novembre, J. and Lange, K. (2009) Fast Model-Based Estimation of Ancestry in Unrelated Individual. *Genome Research*, **19**, 1655-1664. <https://doi.org/10.1101/gr.094052.109>
- [19] Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics*, **23**, 2633-2635. <https://doi.org/10.1093/bioinformatics/btm308>
- [20] Gu, X.R., Feng, C.G., Ma, L., Song, C., Wang, Y., Da, Y., et al. (2011) Genome-Wide Association Study of Body Weight in Chicken F₂ Resource Population. *PLoS ONE*, **6**, Article ID: e21872. <https://doi.org/10.1371/journal.pone.0021872>
- [21] Liu, R.R., Sun, Y.F., Zhao, G.P., Wang, F., Wu, D., Zheng, M., et al. (2013) Genome-Wide Association Study Identifies Loci and Candidate Genes for Body Composition and Meat Quality Traits in Beijing-You Chickens. *PLoS ONE*, **8**, Article ID: e61172. <https://doi.org/10.1371/journal.pone.0061172>
- [22] Zhao, X., Han, Y., Li, Y., Liu, D., Sun, M., Zhao, Y., et al. (2015) Loci and Candidate Gene Identification for Resistance to *Sclerotinia sclerotiorum* in Soybean (*Glycine max* L. Merr.) via Association and Linkage Maps. *The Plant Journal*, **82**, 245-255. <https://doi.org/10.1111/tpj.12810>
- [23] Han, Y., Zhao, X., Cao, G., Wang, Y., Li, Y., Liu, D., et al. (2015) Genetic Characteristics of Soybean Resistance to HG Type 0 and HG Type 1.2.3.5.7 of the Cyst Nematode Analyzed by Genome-Wide Association Mapping. *BMC Genomics*, **16**, Article No. 598. <https://doi.org/10.1186/s12864-015-1800-1>
- [24] Su, J., Pang, C., Wei, H., Li, L., Liang, B., Wang, C., et al. (2016) Identification of Favorable SNP Alleles and Candidate Genes for Traits Related to Early Maturity via GWAS in Upland Cotton. *BMC Genomics*, **17**, Article No. 687. <https://doi.org/10.1186/s12864-016-2875-z>
- [25] Yang, X., Nong, B., Xia, X., Zhang, Z., Zeng, Y., Liu, K., et al. (2016) Rapid Identification of a New Gene Influencing Low Amylose Content in Rice Landraces (*Oryza sativa* L.) Using Genome-Wide Association Study with Specific-Locus Amplified Fragment Sequencing. *Genome*, **60**, 465-472. <https://doi.org/10.1139/gen-2016-0104>
- [26] 田鹏, 刘占林. 糖基转移酶超家族[J]. 生命的化学, 2011, 31(5): 732-736.
- [27] 陈云, 赵丽婷, 顾正华, 李由然, 石贵阳, 丁重阳. 糖基转移酶 *GL24971* 对灵芝多糖合成的影响[J]. 食品与发酵工业, 2021: 1-9. <https://doi.org/10.13995/j.cnki.11-1802/ts.027324>.
- [28] 孙浩然. 铝胁迫下大豆三个 UDP-糖基转移酶基因的功能分析[D]: [硕士学位论文]. 长春: 吉林大学, 2018.
- [29] Stracke, R., Ishihara, H., Hupf, G., Barsch, A., Mehrtens, F., Niehaus, K., et al. (2007) Differential Regulation of Closely Related R2R3-MYB Transcription Factors Controls Flavonol Accumulation in Different Parts of the Arabidopsis Thaliana Seedling. *The Plant Journal*, **50**, 660-677. <https://doi.org/10.1111/j.1365-313X.2007.03078.x>
- [30] Song, S., Qi, T., Huang, H., Ren, Q., Wu, D., Chang, C., et al. (2011) The Jasmonate-ZIM Domain Proteins interact with the R2R3-MYB Transcription Factors MYB21 and MYB24 to Affect Jasmonate-Regulated Stamen Development

- in *Arabidopsis*. *The Plant Cell*, **23**, 1000-1013. <https://doi.org/10.1105/tpc.111.083089>
- [31] Pesch, M. and Huelskamp, M. (2009) One, Two, Three ... Models for Trichome Patterning in *Arabidopsis*. *Current Opinion in Plant Biology*, **12**, 587-592. <https://doi.org/10.1016/j.pbi.2009.07.015>
- [32] Walford, S.-A., Wu, Y., Llewellyn, D.J. and Dennis, E.S. (2011) GhMYB25-Like: A Key Factor in Early Cotton Fiber Development. *The Plant Journal*, **65**, 785-797. <https://doi.org/10.1111/j.1365-313X.2010.04464.x>
- [33] Byrne, M.E., Barley, R., Curtis, M., Arroyo, J.M., Dunham, M., Hudson, A. *et al.* (2000) *Asymmetric leaves1* Mediates Leaf Patterning and Stem Cell Function in *Arabidopsis*. *Nature*, **408**, 967-971. <https://doi.org/10.1038/35050091>
- [34] Shin, R., Burch, A.Y., Huppert, K.A., Tiwari, S.B., Murphy, A.S., Guilfoyle, T.J., *et al.* (2007) The *Arabidopsis* Transcription Factor MYB77 Modulates Auxin Signal Transduction. *The Plant Cell*, **19**, 2440-2453. <https://doi.org/10.1105/tpc.107.050963>
- [35] Mu, R.L., Cao, Y.R., Liu, Y.F., Lei, G., Zou, H.-F., Liao, Y., *et al.* (2009) An R2R3-Type Transcription Factor Gene *AtMYB59* Regulates Root Growth and Cell Cycle Progression in *Arabidopsis*. *Cell Research*, **19**, 1291-304. <https://doi.org/10.1038/cr.2009.83>
- [36] Cominelli, E., Galbiati, M., Vavasseur, A., Conti, L., Sala, T., Vuylsteke, M., *et al.* (2005) A Guard-Cell-specific MYB Transcription Factor Regulates Stomatal Movements and Plant Drought Tolerance. *Current Biology*, **15**, 1196-1200. <https://doi.org/10.1016/j.cub.2005.05.048>
- [37] Jung, C., Seo, J.S., Han, S.W., Koo, Y.J., Kim, C.H., Song, S.I., *et al.* (2008) Overexpression of *AtMYB44* Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic *Arabidopsis*. *Plant Physiology*, **146**, 623-635. <https://doi.org/10.1104/pp.107.110981>
- [38] Park, M.Y., Kang, J.Y. and Kim, S.Y. (2011) Overexpression of *AtMYB52* Confers ABA Hypersensitivity and Drought Tolerance. *Molecular Cells*, **31**, 447-54. <https://doi.org/10.1007/s10059-011-0300-7>