

陆地棉纤维长度和强度的优异位点 挖掘及其候选基因预测

史春辉¹, 张 爱¹, 马 麒², 谢晓宇¹, 刘娟娟¹, 李美丽¹, 李朝周¹, 王彩香¹, 宿俊吉¹

(¹ 甘肃农业大学生命科学技术学院, 兰州 730070, ² 新疆农垦科学院棉花研究所, 石河子 832000)

摘要: 纤维长度和强度是陆地棉 (*Gossypium hirsutum* L.) 纤维品质性状中的 2 个关键性状, 其遗传基础的解析对优质棉品种培育具有重要意义。本研究以 315 个陆地棉品种 (系) 为关联分析群体, 利用混合线性模型 (MLM, mixed linear model) 对来自 5 个环境的纤维长度和强度进行全基因组关联分析 (GWAS, genome-wide association study)。结果表明, 5 个环境下纤维长度和强度的表型值均呈现出一定的差异, 且广义遗传力较高, 纤维长度变异系数为 3.97%~8.44%, 纤维强度变异系数为 7.85%~11.26%。方差分析表明, 基因型、环境和基因型 × 环境互作对纤维长度和强度均有极显著影响 ($P < 0.001$)。聚类分析和群体结构分析表明, 315 份材料可分为 2 个类群。GWAS 共检测到 5 个与纤维长度和强度显著关联的 SNP, 其中位点 D12_57032285 与纤维长度和强度均显著关联。与纤维长度显著关联 3 个 SNP 位点, 分别位于 A05、D11 和 D12 染色体上, 解释 8.05%、12.47% 和 8.79% 的表型变异, 优异等位变异类型分别为 A05_15144433 (AA)、D11_24483544 (TT) 和 D12_57032285 (CC); 与纤维强度显著关联 3 个 SNP 位点, 分别位于 A08、D09 和 D12 染色体上, 解释 9.03%、7.94% 和 7.90% 的表型变异, 优异等位变异类型分别为 A08_84604654 (TT)、D09_43463271 (TT) 和 D12_57032285 (CC)。通过两组不同转录组数据的基因表达模式分析, 筛选出 30 个可能与纤维发育相关的候选基因。通过 GO 富集分析和 KEGG 代谢途径分析发现, 候选基因主要参与蛋白质或蛋白质复合物及 5' - 三磷酸腺苷 (ATP) 选择性且非共价地相互作用, 代谢途径主要为核糖体代谢途径。本研究结果可为棉花纤维品质性状的分子遗传改良提供理论依据。

关键词: 陆地棉; 纤维长度; 纤维强度; 全基因组关联分析; 优异等位变异; 候选基因

Exploration of Elite Loci for Fiber Length and Strength in Upland Cotton and Prediction of Their Candidate Genes

SHI Chun-hui¹, ZHANG Ai¹, MA Qi², XIE Xiao-yu¹, LIU Juan-juan¹,

LI Mei-li¹, LI Chao-zhou¹, WANG Cai-xiang¹, SU Jun-ji¹

(¹ College of Life Science and Technology, Gansu Agricultural University, Lanzhou 730070;

² Cotton Research Institute, Xinjiang Academy of Agricultural and Reclamation Science, Shihezi, 832000)

Abstract: Fiber length and strength are the two most important traits in fiber quality of upland cotton (*Gossypium hirsutum* L.), and the understanding of their genetic basis is significant for breeding cultivars with high-quality cotton. We performed the genome-wide association study (GWAS) for fiber length and strength of association analysis group which comprises 315 upland cotton accessions grown in five different environments,

收稿日期: 2021-01-18 修回日期: 2021-02-22 网络出版日期: 2021-03-12

URL: <http://doi.org/10.13430/j.cnki.jpgr.20210118002>

第一作者研究方向为棉花分子育种应用, E-mail: shichunhui5117@163.com

通信作者: 宿俊吉, 研究方向为棉花数量性状遗传及分子育种, E-mail: sujuncotton@126.com

王彩香, 研究方向为棉花基因功能及分子调控机制, E-mail: wangcx8050@163.com

李朝周, 研究方向为植物生态生理, E-mail: lichaozhou@eyou.com

基金项目: 国家自然科学基金项目 (31971986); 甘肃省科技计划项目 (20JR10RA520); 甘肃省陇原青年创新创业人才项目 (2020RCXM182); 新疆石河子市现代农业科技攻关与成果转化计划 (2017HZ02)

Foundation projects: National Natural Science Foundation of China (31971986), Science and Technology Plan Project of Gansu Province (20JR10RA520), Longyuan Young Talents for Innovation and Entrepreneurship (2020RCXM182), Modern Agricultural Science and Technology Research and Achievement Transformation Plan of Shihezi in Xinjiang (2017HZ02)

through the mixed linear model (MLM). The results showed the presence of some differences in the phenotypic values of fiber length and strength, of high generalized heritability, with the fiber length variation coefficient ranging from 3.97% to 8.44%, and the fiber strength variation coefficient ranging from 7.85% to 11.26%. The analysis of variance for fiber length and strength showed highly significant effects of the genotype, environment, and genotype-environment interaction ($P < 0.001$). Cluster analysis and population structure analysis showed that the 315 accessions could be divided into 2 groups. A total of 5 SNPs significantly associated with fiber length and strength were detected by GWAS, among which the locus D12_57032285 was significantly associated with both the fiber length and strength. The three loci significantly associated with fiber length were located on chromosomes A05, D11, and D12, respectively, which could explain 8.05%, 12.47% and 8.79% of the phenotypic variation, the elite allele types being A05_15144433 (AA), D11_24483544 (TT) and D12_57032285 (CC). The three loci significantly associated with fiber strength were located on chromosomes A08, D09 and D12, respectively, which could explain 9.03%, 7.94% and 7.90% of the phenotypic variation, the elite allele types being A08_84604654 (TT), D09_43463271 (TT) and D12_57032285 (CC). Through the analysis of gene expression patterns of two sets of different transcriptome data, 30 candidate genes that might be related to fiber development were selected. Through GO enrichment analysis and KEGG metabolic pathway analysis, it was found that the candidate genes mainly involved proteins or protein complexes and selectively and non-covalently interact with adenosine 5'-triphosphate (ATP), and the metabolic pathway was mainly the ribosomal metabolic pathway. The results can provide a theoretical basis for molecular genetic improvement of cotton fiber quality traits.

Key words: upland cotton; fiber length; fiber strength; genome-wide association study; elite allelic variation; candidate genes

陆地棉 (*Gossypium hirsutum* L.) 是世界上最重要的天然纤维作物^[1], 具有高产特性和广泛的适应能力, 其种植面积和产量均占全球的 95% 以上^[2-4]。近年来, 随着我国纺织工业对优质棉的需求不断增加, 纤维品质性状的改良成为我国棉花育种最主要的目标和研究者关注的重点^[5-7]。纤维长度和纤维强度是纤维品质的 2 个最关键性状, 其受一系列数量性状基因位点 (QTL, quantitative trait locus) 控制。利用现代遗传学的研究手段, 挖掘纤维长度和强度性状的优异位点, 并预测其候选基因, 对培育优质纤维的棉花品种具有重要的指导意义。

随着基因组测序技术的快速发展和高密度 SNP (SNP, single nucleotide polymorphism) 标记的开发, 全基因组关联分析 (GWAS, genome-wide association study) 方法已成为研究农作物复杂性状的强有力工具^[8], 目前, 通过连锁作图方法已在棉花上进行大量研究, 鉴定出众多与纤维品质相关的 QTL^[9]。GWAS 作为 QTL 定位有益的补充手段, 近年来在棉花纤维品质性状 QTL 检测上得到了广泛的应用^[7, 10-18]。一些控制纤维品质性状的关键区段或位点相继被发现, 如与纤维长度显著关联的区段 A13: 75.10~75.70 Mb^[11]、A10: 65.68~65.72 Mb^[12]、A05: 14.64~15.64 Mb^[13] 和 D11: 24.02~24.09 Mb^[7, 10, 12-15] 等, 与纤维

强度显著关联的区段 A07: 72.17~72.23 Mb^[12] 和 D09: 42.96~43.96 Mb^[13] 等。同时也在这些 QTL 区间内, 鉴定到了控制目标性状的候选基因, 例如: 控制纤维长度的关键基因有 *GhFL1* (*Gh_A10G1256*)、*GhFL2* (*Gh_D11G1929*)、*GhXI-K* (*Gh_A13G1707*) 和 *GH_D02G0143* 等, 控制纤维强度的关键基因有 *GH_A07G1769*、*GH_D13G2200* 等, 而同时控制纤维长度和强度的关键基因有 *GH_A07G1758*、*GH_D11G1929* 等。Li 等^[14] 研究发现 1 个包括 962 个基因的 eQTL 热点 (Hot216), 这些基因是通过调节细胞从快速伸长到次生细胞壁合成的发育转变来促进纤维生长。可以看出, 一些与棉花纤维品质性状相关的主效 QTL 区段及候选基因已被发现, 这些 QTL 及其区间内的候选基因应该是今后我们关注的重点, 值得后续深入研究。

尽管目前已在陆地棉纤维品质性状 QTL 和候选基因的挖掘方面取得了较大进展, 但所挖掘到的相关 QTL 和候选基因的数量仍然有限, 难以从全基因组水平上开展纤维品质分子设计育种。纤维长度和强度作为棉花纤维品质最重要的 2 个性状, 其 QTL 定位及关键调控基因挖掘的研究仍需进一步加强。本研究对 315 个陆地棉品种 (系) 在 5 个环境下进行纤维长度和强度的表型鉴定, 并主要针对这 2 个目标性状

进行 GWAS 分析,进一步挖掘与目标性状显著关联的 SNP 位点及其优异等位变异类型。并结合已发表的棉花不同组织转录组数据,预测分布在关联位点附近的与纤维长度和强度相关的候选基因。该研究中挖掘的优异等位变异类型及候选基因,将为后期培育优良纤维品质的棉花品种提供理论依据。

1 材料与方法

1.1 试验材料

本研究选用 315 个陆地棉品种(系)为试验材料,其中 290 个来自中国不同时期选育的品种(系),25 个是从国外引进的品种。中国 290 个品种(系)中,123 个来自于黄河流域棉区、51 个来自于长江流域棉区、96 个来自于西北内陆棉区、20 个来自于北部特早熟棉区;25 个国外引进品种包括美国 21 个和前苏联 4 个。试验材料来源及分类与我们前期发表报道的相一致^[13](详见 <http://doi.org/10.13430/j.cnki.jpgr.20210118002>,附表 1)。

1.2 田间试验

于 2014 年和 2015 年在河南安阳(AY)、新疆石河子(SHZ)以及 2019 年在新疆库尔勒(KEL)田间种植关联分析群体,总共 5 个环境,分别命名为 AY-14、AY-15、SHZ-14、SHZ-15 和 KEL-19。所有试验点田间试验均采用随机区组试验设计,每个环境设置 3 个重复。安阳试验点:每个品种(系)单行种植,行长 5.00 m,行距 0.80 m,株距 0.20 m,每个品种(系)有 18~23 个单株;石河子和库尔勒试验点:每个品种(系)均种植 2 行,行长 2.00 m,行距 0.45 m,株距 0.10 m,每个品种(系)约 40 个单株。5 个种植环境均以当地常规方法进行田间管理。

1.3 取样、纤维品质检测与数据分析

在每年的 9 月下旬棉花收获期,对每个品种(系)的每个重复,从植株的中部手工采摘正常吐絮的棉铃 20 个。轧花后取其纤维样品约 12 g 送农业农村部棉花纤维品质监督检验测试中心(河南安阳),使用 HVI-MF 100 仪器检测纤维长度和纤维强度等纤维品质性状指标。利用 SPSS26.0 软件对各环境下的纤维长度和强度表型值进行统计描述和多环境方差分析,并计算纤维长度和强度的广义遗传力。

1.4 聚类分析、群体结构分析与连锁不平衡分析

实验室前期对包括 315 个品种(系)的关联分析群体利用 CTAB 法从幼叶组织中提取基因组 DNA^[19],使用 SLAF-seq 方法对该群体进行测序^[20],获得该群体的基因组序列,再与新发表的陆

地棉 TM-1 参考基因组序列^[21]进行比对,淘汰缺失率 >10% 和最小等位基因频率(MAF, minor allele frequency) <0.05 的标记,共筛选得到 13391 个高质量 SNP 标记^[13]。基于这些高质量的 SNP 标记,利用 TASSEL 5.2.50 软件对该群体进行聚类分析、主成分分析(PCs),并计算亲缘关系系数(K 值)和连锁不平衡(LD, linkage disequilibrium)的平均衰减距离^[22]。

1.5 全基因组关联分析(GAWS)

利用 TASSEL 5.2.50 软件混合线性模型(MLM, mixed linear model),以 PCs 和亲缘关系系数 K 为协变量进行 GWAS 分析,鉴定纤维长度和纤维强度的显著 SNP 位点^[23]。GWAS 显著位点的阈值(Bonferroni) P 为 7.5×10^{-6} ($0.5/\text{总标记数 } n$, 即 $-\lg(P) = 4.43$)。 $-\lg(P) \geq 4.43$ 的位点被认为是与目标性状显著相关联的 SNP 位点,然后利用 EXCEL 绘制出关联位点的曼哈顿图(Manhattan Plot)。根据全基因组的 LD 衰减距离,最终确定显著 SNP 位点所在的 QTL 区间。

1.6 优异等位变异鉴定

利用 GWAS 方法鉴定出目标性状的显著 SNP 位点,计算显著位点的每个等位基因在 5 个环境的表型平均值,用 Omicshare 网站(<https://www.omicshare.com>)绘制这些位点的等位基因对应表型值的箱线图,对每个等位变异对应的表型值进行双尾 T 检验,检验两组间的显著性差异,并确定优异的等位变异类型。

1.7 候选基因预测

根据 SNP 位点在陆地棉基因组上的物理位置,结合 LD 的衰减距离,利用棉花功能基因组学数据库网站(<http://cottonfgd.org/>)获得 QTL 区间内注释基因。然后利用南京农业大学(NAU, Nanjing Agricultural University)和中国农业科学院棉花研究所(CRI, Cotton Research Institute of Chinese Academy of Agricultural Sciences)完成的两组转录组数据,分别计算上述每个基因的 FPKM (fragments per kilobase per million)值,分析区段内所有基因在棉花 12 个组织(包括开花后 4 个不同时期的纤维和根、茎、叶、花托、雌蕊、雄蕊、花萼和花瓣)中的表达模式,筛选出在 4 个时期的纤维组织中高量表达的基因作为候选基因^[24]。利用 Omicshare 网站(<https://www.omicshare.com>)绘制基因表达模式热图。对潜在的候选基因,利用棉花功能基因组学数据库网站(<http://cottonfgd.org/>),进行 GO 富集分析和 KEGG 代谢途径分析,推断候

选基因的功能及其参与的生物学调控通路。

2 结果与分析

2.1 纤维长度和强度性状的统计分析

通过对陆地棉关联分析群体 5 个环境下纤维长度和强度的数据分析,发现 2 个性状均具有广泛的变异(表 1、图 1)。纤维长度分布在 20.14~34.59 mm 之间,平均值为 28.08 mm;标准差范围为 1.12~2.17;变异系数范围为 3.97%~8.44%,平均变异系数为 5.60%。纤维强度分布在 21.17~40.65 cN/tex 之间,平均值为

28.56 cN/tex;标准差范围为 2.36~3.07;变异系数范围为 7.85%~11.26%,平均变异系数为 9.42%。统计描述结果表明,纤维长度和强度性状表型值在 5 个环境下分布趋势基本一致,且呈现出一定的差异。纤维长度和强度的广义遗传力较高,分别为 60.75% 和 62.52%,表明纤维长度和强度性状的遗传相对稳定(表 2)。但是多环境方差分析表明,基因型、环境和基因型 × 环境互作对纤维长度和强度都有极显著影响 ($P<0.001$),表明目标性状受到试验环境因素的显著影响,需要后续分单个环境进行 GWAS 分析。

表 1 5 个不同环境下纤维长度和强度的统计学描述

Table 1 Statistical description of fiber length and strength in five different environments

性状 Trait	环境 Environment	最小值 Minimum	最大值 Maximum	平均值 ± 标准差 Mean ± SD	变异系数(%) CV
纤维长度 (mm) Fiber length	AY-14	24.55	34.59	29.60 ± 1.37	4.62
	AY-15	23.25	34.40	29.06 ± 1.70	5.85
	SHZ-14	25.16	32.09	28.15 ± 1.12	3.97
	SHZ-15	24.27	32.27	27.83 ± 1.50	5.38
	KEL-19	20.14	32.55	25.75 ± 2.17	8.44
纤维强度 (cN/tex) Fiber strength	AY-14	23.87	39.93	30.66 ± 2.46	8.03
	AY-15	23.05	40.65	29.61 ± 3.07	10.38
	SHZ-14	25.64	40.54	30.13 ± 2.36	7.85
	SHZ-15	22.90	36.40	27.06 ± 2.60	9.60
	KEL-19	21.17	36.46	25.36 ± 2.86	11.26

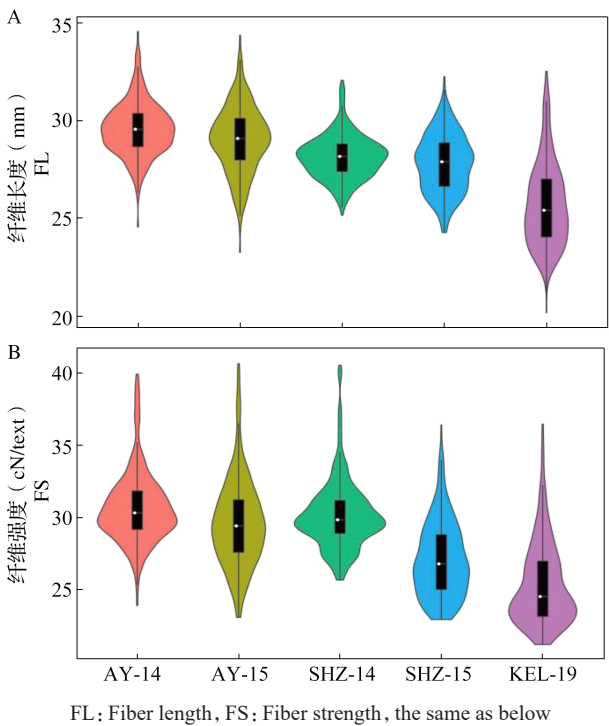


图 1 315 个陆地棉品种(系)中纤维长度和纤维强度的表型分布情况

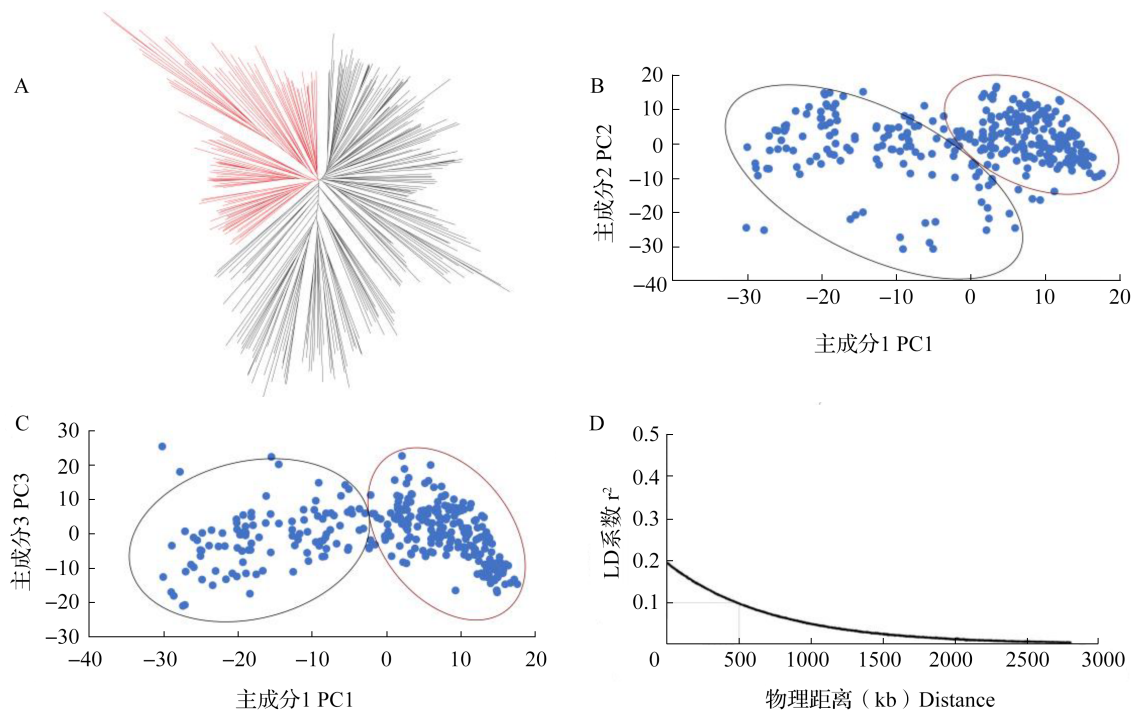
Fig.1 The phenotype distribution of fiber length and fiber strength in 315 upland cotton accessions

2.2 聚类分析、群体结构分析与连锁不平衡分析

本实验室前期利用 SLAF-seq 方法对包括 315 个陆地棉品种(系)的关联分析群体进行简化测序,筛选得到 13391 个高质量 SNP 标记^[13]。本研究基于这些 SNP 基因型数据,利用 TASSEL 软件对其进行聚类分析、群体结构分析及连锁不平衡分析。系统进化树显示,这些材料大致可以分为两大类群,其中红色标记部分的 114 个大部分是国外品种和中国早期品种(系),黑色标记部分的 201 个大部分是中国近期选育的品种(系)(图 2A)。使用过滤后的 SNP 标记进行主成分分析(PCs),并根据 PC1 和 PC2、PC1 和 PC3 的结果也将该群体分为 2 个亚群(图 2B、C)。因此,将 315 个陆地棉品种(系)分为 2 个类群。最后利用过滤后的全部 SNP 标记进行全基因组 LD 分析,当 LD 系数 r^2 衰减到其最大值一半时,即 r^2 值从最大值 0.20 降低到 0.10 时,全基因组上 SNP 标记间所需要物理距离约为 500 kb(图 2D)。结果表明该陆地棉关联分析群体的全基因组平均 LD 值约为 500 kb。

表 2 多环境下纤维长度和强度性状的方差分析
Table 2 Analysis of variance of fiber length and strength in multiple environments

性状 Trait	变异来源 Source of variance	自由度 DF	平方和 SS	均方 MS	F 值 F-value	P 值 P-value	遗传力 (%) H^2
纤维长度 Fiber length	基因型	314	7841.01	24.97	29.59	0	60.75
	环境	4	3425.41	856.35	1014.85	0	
	基因型 × 环境	1256	1795.76	1.43	1.69	3.41E-31	
纤维强度 Fiber strength	基因型	314	26438.54	84.20	28.29	0	62.52
	环境	4	9513.64	2378.41	799.18	0	
	基因型 × 环境	1256	4753.37	3.78	1.27	1.10E-07	



A: 系统进化树; B、C: 315 个品种 (系) 的主成分分析; D: 连锁不平衡分析
A: Phylogenetic tree, B and C: Principal component analysis of 315 accessions, D: Linkage disequilibrium analysis

图 2 基于 SNP 的 315 个品种 (系) 群体结构和 LD 衰减研究

Fig.2 SNP-based research on population structure and LD attenuation of 315 accessions

2.3 陆地棉纤维长度和强度的全基因组关联分析 (GWAS)

利用 TASSEL 软件的 MLM (PCs + K) 模型, 分别对 5 个试验环境的纤维长度和强度进行 GWAS 分析, 以 $-\lg(P) \geq 4.43$ 的 SNP 被认为是显著位点, 共检测到 5 个 (D11_24483544、D12_57032285、A05_15144433、D09_43463271 和 A08_84604654) 与纤维长度和强度显著关联的 SNP 位点 (表 3)。在这些 SNP 位点中, 3 个位点 (D11_24483544、D12_57032285

和 A05_15144433) 与纤维长度显著关联, 它们分别分布于染色体 D11、D12 和 A05 上, 分别解释 12.47%、8.79% 和 8.05% 的表型变异 (表 3, 图 3A、B、C)。在 AY-14 和 AY-15 两个试验环境中均能检测到 SNP 位点 D11_24483544, 且具有最高的 $-\lg(P)$ 值 (6.40) 和最大表型变异解释率 (12.47%), 而 D12_57032285 和 A05_15144433 仅在单一环境中被检测到。另外 3 个位点 (D09_43463271、D12_57032285 和 A08_84604654) 与纤维强度显著关联, 它们均在单一环境中被检测

到,分别分布于 D09、D12 和 A08 染色体上,可分别解释 7.94%、7.90% 和 9.03% 的表型变异(表 3,图 4A、

B、C)。此外,SNP 位点 D12_57032285 与纤维长度和纤维强度 2 个性状的关联均达到显著水平。

表 3 与纤维长度和强度性状显著关联 SNP 位点

Table 3 The SNP loci significantly associated with fiber length and/or strength

性状 Trait	SNP 位点 SNP loci	染色体 Chr.	位置 Position	$-\lg(P)$	表型变异解释率(%) R^2	环境 Environment	已报道 QTL Reported QTL
纤维长度 Fiber length	D11_24483544	D11	24483544	6.40	12.47	AY-14/AY-15	[7, 10, 12-15]
	D12_57032285	D12	57032285	5.54	8.79	SHZ-15	
	A05_15144433	A05	15144433	5.24	8.05	SHZ-15	[13, 18]
纤维强度 Fiber strength	D09_43463271	D09	43463271	5.10	7.94	AY-15	[13, 25]
	D12_57032285	D12	57032285	4.91	7.90	SHZ-15	
	A08_84604654	A08	84604654	5.00	9.03	KEL-19	

2.4 显著 SNP 位点优异等位变异的鉴定

对上述 5 个 SNP 位点进行有利等位变异鉴定。在 3 个与纤维长度显著关联的位点中, A05_15144433 存在 AA (n=59)、AC (n=50) 和 CC (n=203) 3 种等位变异类型, AA 型的纤维长度平均值(28.56 mm)极显著高于 CC 型(27.91 mm); D11_24483544 存在 TT (n=176)、CT (n=15) 和 CC (n=71) 3 种等位变异类型, TT 型的纤维长度平均值(28.35 mm)极显著高于 CC 型(27.25 mm); D12_57032285 存在 CC (n=271)、CT (n=11) 和 TT (n=23) 3 种等位变异类型, CC 型的纤维长度平均值(28.18 mm)极显著高于 TT 型(27.35 mm)(表 4,图 3D、E、F)。根据棉花纤维长度的育种目标,可确定 D11_24483544(TT)、D12_57032285(CC) 和 A05_15144433(AA) 为纤维长度的优异等

位变异类型。在 3 个与纤维强度显著关联的位点中, A08_84604654 存在 TT (n=25)、AA (n=235) 和 AT (n=6) 3 种等位变异类型, TT 型的纤维强度平均值(29.94 cN/tex)高于 AA 型(28.36 cN/tex),但未达到显著水平; D09_43463271 存在 TT (n=95)、CC (n=148) 和 CT (n=67) 3 种等位变异类型, TT 型的纤维强度平均值(29.51 cN/tex)极显著高于 CC 型(28.01 cN/tex); D12_57032285 的 3 种等位变异分别为 CC (n=271)、CT (n=11) 和 TT (n=23), CC 型的纤维强度平均值(28.68 cN/tex)高于 TT 型(27.91 cN/tex),但未达到显著水平(表 4,图 4D、E、F)。参照棉花纤维强度的育种目标, A08_84604654(TT)、D09_43463271(TT) 和 D12_57032285(CC) 是与纤维强度性状相关的优异等位变异。

表 4 显著 SNP 标记的等位基因组成情况

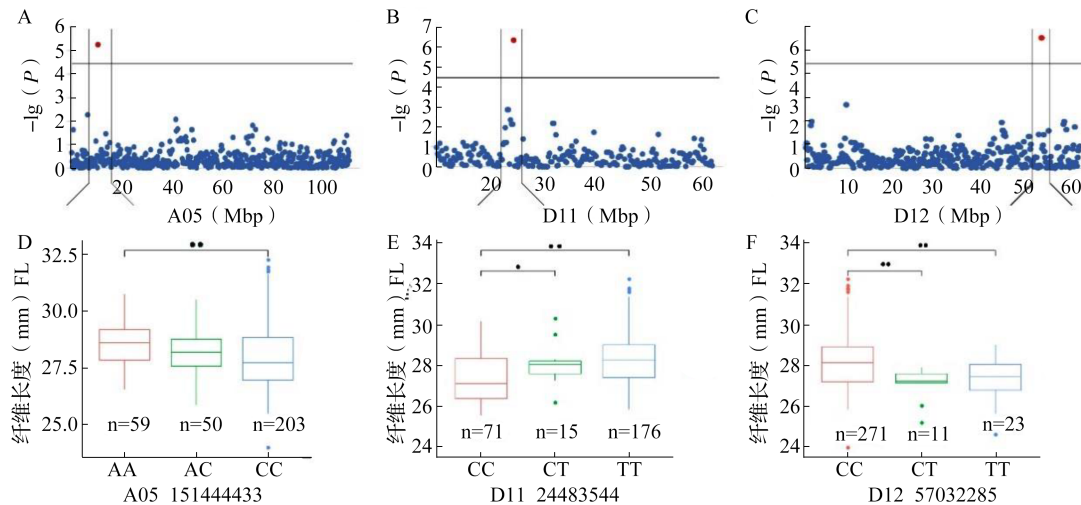
Table 4 Allelic variation composition of the significant SNP markers

性状 Trait	SNP 位点 SNP locus	等位变异类型 1/ 品种数 Allele 1/No. of cultivars	等位变异类型 2/ 品种数 Allele 2/No. of cultivars	等位变异类型 3/ 品种数 Allele 3/No. of cultivars
纤维长度 Fiber length	A05_15144433	AA/59	AC/50	CC/203
	D11_24483544	CC/71	CT/15	TT/176
	D12_57032285	CC/271	CT/11	TT/23
纤维强度 Fiber strength	A08_84604654	AA/235	AT/6	TT/25
	D09_43463271	CC/148	CT/67	TT/95
	D12_57032285	CC/271	CT/11	TT/23

2.5 候选基因的预测

根据陆地棉参考基因组信息,将上述与目标性状显著关联的 SNP 位点比对到陆地棉参考基因组上。按照 LD 为 500 kb 衰减距离,在这些 SNP 位点侧翼 1 Mb 的范围内提取基因,共在 5 个 SNP 位点

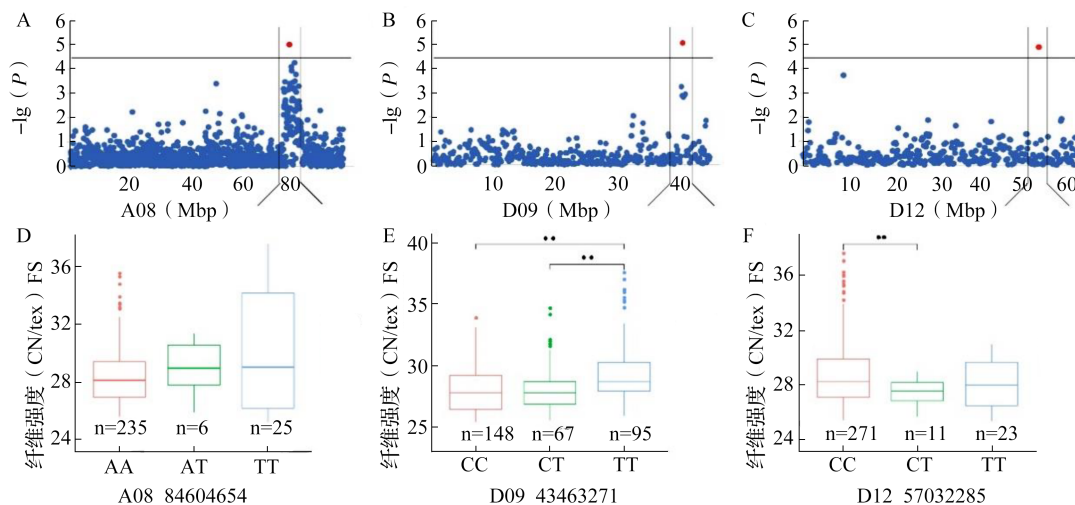
侧翼的目标区域有 356 个基因被注释。利用 NAU 的 RNA-seq 数据中基因表达的 FPKM 值,发现 289 个基因至少在 12 个棉花组织中的一个中有较高表达,67 个基因在所有组织中无表达或表达量很低。在较高表达量的 289 个基因中,63 个基因在棉纤维



A、B 和 C 表示 3 个纤维长度的关联位点; D、E 和 F 表示 3 个位点的箱线图。* 和 ** 分别表示 0.05 和 0.01 的显著性水平; 下同
A, B and C indicate the three loci associated with fiber length, D, E and F represent the boxplot of the three loci. * and ** indicate 5% and 1% significance level, respectively, the same as below

图 3 纤维长度的 GWAS 结果及关联位点的优异等位变异

Fig.3 GWAS results of fiber length and the elite allelic variation of the associated loci



A、B 和 C 表示 3 个纤维强度的关联位点

A, B and C indicate the three loci associated with fiber strength

图 4 纤维强度的 GWAS 结果及关联位点的优异等位变异

Fig.4 GWAS results of fiber strength and the elite allelic variation of the associated loci

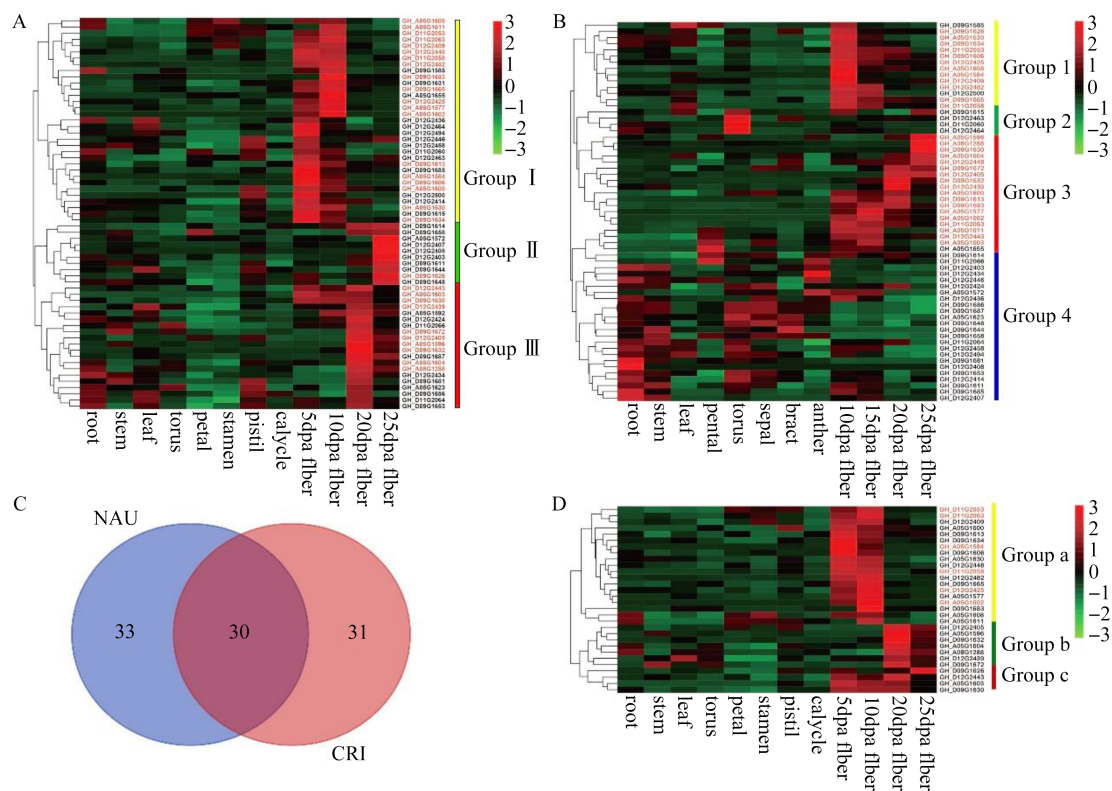
组织中具有高表达量。根据基因表达模式,将上述 63 个基因分为 3 组(I、II、III)。其中, I 组的 33 个基因主要在 5 和 10 DPA(days post anthesis)的纤维组织中高表达; II 组的 10 个基因主要在 25 DPA 纤维组织中高表达; III 组的 20 个基因主要在 20 DPA 纤维组织中高表达(图 5A);通过蛋白质序列,将上述在不同时期纤维组织中高表达的 63 个基因比对到 CRI 数据库中,发现了相对应的 61 个基因。根据 CRI 的 RNA-seq 数据中基因表达的 FPKM 值,可将这 61 个基因分为 1~4 组。第 1 组的 14 个基因主要在 5 DPA 和 10 DPA 的纤维组织

中高表达;第 2 组的 4 个基因主要在花托中高表达;第 3 组的 18 个基因在纤维发育的 4 个时期均有较高表达;第 4 组的 25 个基因主要在根、茎、花托和萼片等组织中高表达(图 5B);综合 NAU 和 CRI 两组转录组数据,发现 30 个基因至少在 4 个时期纤维组织中的一个中高表达(图 5C、D)。初步推测这 30 个基因是调控陆地棉纤维长度和强度的候选基因(表 5)。进一步分析可将这 30 个基因分为 a、b、c 3 个类别。a 类的 19 个基因主要在 5 DPA 和 10 DPA 的纤维组织中高表达,该类基因可能与纤维长度有关;b 类的 7 个基因主要在 20 DPA 和 25 DPA

的纤维组织中高表达,该类基因可能与纤维强度有关;c类的4个基因在4个时期的纤维组织中均有着不同程度的高表达,此类基因可能同时与纤维长度和强度2个性状相关。

对上述30个基因进行GO富集分析和KEGG代谢途径分析,发现28个基因具有一定的生物学功能注释(<http://doi.org/10.13430/j.cnki.jpg.20210118002>,附表2)。GO结果显示,在这30

个基因中总共发现了42个调控途径,其中2个调控途径最为关键,4个基因参与蛋白质或蛋白质复合物选择性且非共价地相互作用(GO:0005515);3个基因参与5'-三磷酸腺苷(ATP)选择性且非共价地相互作用(GO:0005524)。根据KEGG分析结果,共发现了16个调控途径,其中在2个基因(*GH_A05G1600*和*GH_D09G1606*)预测了7条途径,并发现了它们共同的核糖体代谢途径(ko01100)。



A: 南京农业大学 (NAU) RNA-seq 数据在纤维组织中高表达的 63 个基因的表达模式热图; B: 中国农业科学院棉花研究所 (CRI) RNA-seq 数据在纤维组织中高表达的 61 个基因的表达模式热图; C: NAU 和 CRI 中共同基因的韦恩图; D: 30 个候选基因的表达模式热图;

右侧标尺数值范围代表均一化后表达量的变化范围,红色为高表达,绿色为低表达

A: Heat map of the expression pattern of the 63 genes that are highly expressed in fibrous tissue from the RNA-seq data of Nanjing Agricultural University (NAU), B: Heat map of the 61 genes that are highly expressed in fibrous tissues in the RNA-seq data of Cotton Research Institute of Chinese Academy of Agricultural Sciences (CRI), C: Venn diagram of common genes in NAU and CRI, D: Heat map of expression pattern of the 30 candidate genes, values in the scale on the right represent the range of the variation of normalized quantity of expression, in red for high level expressions and in green for low level expressions

图5 纤维长度和强度候选基因在陆地棉不同组织中的表达分析

Fig.5 Expression analysis of candidate genes for fiber length and strength in different upland cotton tissues

3 讨论

棉花纤维长度和强度均是由多基因控制的数量性状,挖掘纤维长度和强度性状相关的QTL和候选基因对棉花纤维品质性状的分子遗传改良具有指导意义。本研究以315个陆地棉品种(系)为关联分析群体,对5个环境下的纤维长度和强度进行GWAS分析,共筛选出5个与目标性状显著

关联的SNP位点(表3)。其中位于D11染色体上的D11_24483544是最重要的一个关联位点,其表现出最高的 $-\lg(P)$ 和最大的表型变异解释率,同时在2个环境下均达到了显著水平。与前期的GWAS结果相比较发现,D11_24483544分布在区段D11:24.02~24.09 Mb附近区域^[7, 10, 12-15],而且该区段已被确定为控制纤维长度性状的热点区域^[14];表明该位点在调控纤维长度方面有着重要作

用。A05_15144433 分布在区段 A05: 14.64~15.64 Mb 内部,且该区段被确定为控制纤维长度的 QTL^[13,18]; D09_43463271 也分布在区段 D09: 42.96~43.96 Mb 内部,且该区段被确定为控制纤维强度的 QTL^[13,25]。上述这些在多个不同研究群体中被检测的 QTL,应当是特别可靠稳定的与纤维品质性状相关 QTL,其

优异等位变异类型及可能的候选基因将是今后需重点研究的对象。其余的 2 个显著关联 SNP 位点在前人的研究中未发现,可能是控制纤维长度和强度性状的新位点。其中位点 D12_57032285 与 2 个目标性状均显著相关联,位点 A08_84604654 可能与纤维强度紧密相关。

表 5 与陆地棉纤维长度和纤维强度相关的候选基因

Table 5 Candidate genes for fiber length and fiber strength of upland cotton

基因 ID Gene ID	基因名称 Gene Name	功能注释 Description	染色体 Chromosome	长度 (bp) Length
<i>GH_A05G1577</i>	<i>IQD1</i>	蛋白质 <i>IQ-DOMAIN 1</i>	A05	2635
<i>GH_A05G1584</i>	<i>At5g45950</i>	GDGL 酯酶 / 脂肪酶	A05	2953
<i>GH_A05G1596</i>	<i>PSKR</i>	植物硫激酶受体 1	A05	1013
<i>GH_A05G1600</i>	<i>glpV</i>	糖原磷酸化酶 1	A05	8224
<i>GH_A05G1602</i>	<i>E6</i>	蛋白 E6	A05	740
<i>GH_A05G1603</i>	<i>Y-3</i>	未表征的蛋白质	A05	1445
<i>GH_A05G1604</i>	<i>At4g28440</i>	未表征的蛋白质	A05	1102
<i>GH_A05G1608</i>	<i>At1g14450</i>	NADH 脱氢酶 [泛醌] 1 beta 亚复合亚基 3-B	A05	221
<i>GH_A05G1611</i>	<i>Stard10</i>	含 START 结构域的蛋白质 10	A05	1921
<i>GH_A05G1630</i>	<i>Piezo1</i>	压电型机械敏感离子通道组件 1	A05	1723
<i>GH_A08G1288</i>	<i>MTP1</i>	金属耐受蛋白 1	A08	758
<i>GH_D09G1606</i>	<i>ADK2</i>	腺苷激酶 2	D09	2954
<i>GH_D09G1613</i>	<i>MYH13</i>	肌球蛋白 13	D09	1466
<i>GH_D09G1626</i>	<i>SIFA2</i>	DNA 结合蛋白 <i>SIFA2</i>	D09	1499
<i>GH_D09G1630</i>	<i>IQD1</i>	蛋白质 <i>IQ-DOMAIN 1</i>	D09	2554
<i>GH_D09G1632</i>	<i>IRX9</i>	可能的 β -1, 4- 木糖基转移酶 <i>IRX9</i>	D09	1725
<i>GH_D09G1634</i>	<i>RPS23</i>	40S 核糖体蛋白 S23	D09	1131
<i>GH_D09G1665</i>	<i>At2g37660</i>	未表征的蛋白,叶绿素	D09	2110
<i>GH_D09G1672</i>	<i>PBL10</i>	可能的丝氨酸 / 苏氨酸蛋白激酶 <i>PBL10</i>	D09	1791
<i>GH_D09G1683</i>	<i>ATL53</i>	推定的 RING-H2 指蛋白 <i>ATL53</i>	D09	548
<i>GH_D11G2053</i>		缺失	D11	1199
<i>GH_D11G2058</i>	<i>RALFL34</i>	蛋白质 RALF-like 34	D11	404
<i>GH_D11G2063</i>	<i>At5g67130</i>	含有 PI-PLC X 结构域的蛋白质	D11	3341
<i>GH_D12G2405</i>	<i>AAK1</i>	AP2 相关蛋白激酶 1	D12	878
<i>GH_D12G2409</i>	<i>At5g47570</i>	NADH 脱氢酶 [泛醌] 1 beta 亚复合亚基 8, 线粒体	D12	1879
<i>GH_D12G2425</i>	<i>CML13</i>	可能的钙结合蛋白 <i>CML13</i>	D12	2165
<i>GH_D12G2439</i>	<i>KINUA</i>	驱动蛋白样蛋白 <i>KINUA</i>	D12	7603
<i>GH_D12G2443</i>	<i>BASS6</i>	可能的钠 / 代谢物共转运蛋白 <i>BASS6</i> , 叶绿体	D12	2553
<i>GH_D12G2448</i>	<i>At1g12500</i>	可能的磷酸糖 / 磷酸盐转运子	D12	1076
<i>GH_D12G2482</i>		缺失	D12	1380

棉纤维是由胚珠表皮细胞分化、发育而成的单细胞纤维,其发育过程可分为起始、伸长、次生壁增厚和脱水成熟 4 个时期^[26]。在棉纤维发育过程中,15 DPA 是 1 个关键的时期,之前主要为纤维伸长阶段,之后主要为次生壁增厚阶段^[26-27]。研究发现 *GhCaM7*、 Ca^{2+} 、ROS、*GhPSK*、*GhHOX3* 和 *GhTCP4* 等是与纤维伸长和次生壁增厚相关的重要调节因子或基因^[26-30]。本研究通过 GWAS 方法共获得 5 个显著关联位点,并在这 5 个位点的 QTL 区域共预测到 30 个与纤维长度和强度有关的候选基因,通过 GO 富集分析和 KEGG 代谢途径分析发现,候选基因主要参与蛋白质或蛋白质复合物及 5'-三磷酸腺苷(ATP)选择性且非共价地相互作用;代谢途径主要为核糖体代谢途径。本研究预测到 30 个候选基因,其中 2 个基因(*GH_A05G1584* 和 *GH_A05G1602*) 在 5 DPA 和 10 DPA 的纤维中优势表达,且这 2 个基因分布在稳定 QTL 区域 A05:14.64~15.64 Mb 内部,其中 *GH_A05G1584* 功能注释为 GDSL 酯酶/脂肪酶,是参与植物生长发育、器官形态建成和次生代谢有关的一类水解酶;而 *GH_A05G1602* 功能注释为蛋白质 E6,是参与细胞壁结构的一种蛋白;表明这 2 个基因可能在调控纤维长度方面有着关键作用。另外,3 个基因(*GH_D11G2053*、*GH_D11G2058* 和 *GH_D11G2063*) 分布在位点 D11_24483544 的 QTL 区域附近,且该位点与 Li 等^[14]的研究中发现的 D11 染色体上 eQTL 热点相重叠,其中 *GH_D11G2053* 的功能未知;*GH_D11G2058* 负责编码 RALF-like34 蛋白,该蛋白被证明与根毛伸长有关^[31]; *GH_D11G2063* 功能注释为包含 PI-PIC X 域的蛋白(磷脂酰肌醇特异性磷脂酶 C X 域),磷脂酰肌醇特异性磷脂酶 C 是水解磷脂生成两种钙离子信号转导途径中的第二信使二酰甘油和三磷酸肌醇的一种酶^[32]。因此,D11 染色体上的 3 个候选基因可能对纤维长度有重要的调控作用。而本研究在与纤维长度和强度都显著关联的新位点 D12_57032285 上预测到基因 *GH_D12G2425*,其功能注释为钙结合蛋白 CML13,而 Ca^{2+} 是参与早期纤维发育的重要调节因子,推测该基因可能对纤维长度和强度性状均有着调控作用^[33]。其余的候选基因如何对目标性状产生影响还需要进一步去研究。总之,本研究鉴定的纤维长度和强度优异等位变异和预测的候选基因可对后期开展全基因组水平上的纤维品质分子设计育种提供理论基础。

4 结论

本研究通过 GWAS 分析共检测到 5 个与纤维长度和强度显著关联的 SNP 位点,并鉴定了它们的优异等位变异类型,预测到 30 个可能与纤维长度和强度有关的候选基因。

参考文献

- [1] 张素君,李兴河,唐丽媛,王海涛,刘存敬,蔡肖,张香云,张建宏. 陆地棉纤维品质性状关联分析及优异等位基因挖掘. 植物遗传资源学报, 2021, 22(1): 214-228
Zhang S J, Li X H, Tang L Y, Wang H T, Liu C J, Cai X, Zhang X Y, Zhang J H. Exploration of elite alleles related with fiber quality traits in *Gossypium hirsutum* L. by association analysis. Journal of Plant Genetic Resources, 2021, 22(1): 214-228
- [2] 朱国忠,张芳,付洁,李乐晨,牛二利,郭旺珍. 适于陆地棉品种身份鉴定的 SNP 核心位点筛选与评价. 作物学报, 2018, 44(11): 1631-1639
Zhu G Z, Zhang F, Fu J, Li L C, Niu E L, Guo W Z. Genome-wide screening and evaluation of SNP core loci for identification of upland cotton varieties. Acta Agronomica Sinica, 2018, 44(11): 1631-1639
- [3] Zhang J F, Abdelraheem A, Tom W. Genetic variation of waterlogging tolerance in Pima (*Gossypium barbadense*) cotton and glanded and glandless upland cotton (*Gossypium hirsutum*) under field conditions. Industrial Crops & Products, 2019, 129: 169-174
- [4] Sun Z W, Wang X F, Liu Z W, Gu Q S, Zhang Y, Li Z K, Ke H F, Yang J, Wu J H, Wu L Q, Zhang G Y, Zhang C Y, Ma Z Y. A genome-wide association study uncovers novel genomic regions and candidate genes of yield-related traits in upland cotton. Theoretical and Applied Genetics, 2018, 131(11): 2413-2425
- [5] 杨君,马峙英,王省芬. 棉花纤维品质改良相关基因研究进展. 中国农业科学, 2016, 49(22): 4310-4322
Yang J, Ma Z Y, Wang S F. Progress in studies on genes related to fiber quality improvement of cotton. Scientia Agricultura Sinica, 2016, 49(22): 4310-4322
- [6] 唐淑荣,魏守军,郭瑞林,韦京艳,孟俊婷,杨长琴. 不同熟性棉花品种纤维品质特征分析与评价. 中国生态农业学报, 2019, 27(10): 1564-1577
Tang S R, Wei S J, Guo R L, Wei J Y, Meng J T, Yang C Q. Evaluation and analysis of fiber quality characteristics of different maturity cotton varieties. Chinese Journal of Eco-Agriculture, 2019, 27(10): 1564-1577
- [7] Naoumkina M, Thyssenh G N, Fang D D, Jenkins J N, Mcarty J C, Florane C B. Genetic and transcriptomic dissection of the fiber length trait from a cotton (*Gossypium hirsutum* L.) MAGIC population. BMC Genomics, 2019, 20(1): 112
- [8] 姜洪真,马伯军,钱前,高振宇. 全基因组关联分析(GWAS)在作物农艺性状研究中的应用. 农业生物技术学报, 2018, 26(7): 1244-1257
Jiang H Z, Ma B J, Qian Q, Gao Z Y. The application of

- genome-wide association study (GWAS) in crop agronomic traits. *Journal of Agricultural Biotechnology*, 2018, 26 (7) : 1244-1257
- [9] Said J I, Song M Z, Wang H T, Lin Z X, Zhang X L, Fang D D, Zhang J F. A comparative meta-analysis of QTL between intraspecific *Gossypium hirsutum* and interspecific *G. hirsutum* $\times$ *G. barbadense* populations. *Molecular Genetics and Genomics*, 2015, 290 (3) : 1003-1025
- [10] Sun Z W, Wang X F, Liu Z W, Gu Q S, Zhang Y, Li Z K, Ke H F, Yang J, Wu J H, Wu L Q, Zhang G Y, Zhang C Y, Ma Z Y. Genome-wide association study discovered genetic variation and candidate genes of fiber quality traits in *Gossypium hirsutum* L. *Plant Biotechnology Journal*, 2017, 15 (8) : 982-996
- [11] Fang L, Wang Q, Hu Y, Jia Y H, Chen J D, Liu B L, Zhang Z Y, Guan X Y, Chen S Q, Zhou B L, Mei G F, Sun J L, Pan Z, He S P, Xiao S H, Shi W J, Gong W F, Liu J G, Ma J, Cai C P, Zhu X F, Guo W Z, Du X M, Zhang T Z. Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits. *Nature Genetics*, 2017, 49 (7) : 1089-1098
- [12] Ma Z Y, He S P, Wang X F, Sun J L, Zhang Y, Zhang G Y, Wu L Q, Li Z K, Liu Z H, Sun G F, Yan Y Y, Jia Y H, Yang J, Pan Z E, Gu Q S, Li X Y, Sun Z W, Dai P H, Liu Z W, Gong W F, Wu J H, Wang M, Liu H M, Feng K Y, Ke H F, Wang J D, Lan H Y, Wang G N, Peng J, Wang N, Wang L R, Pang B Y, Peng Z, Li R Q, Tian S L, Du X M. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nature Genetics*, 2018, 50 (6) : 803-813
- [13] Su J J, Wang C X, Yang D L, Shi C H, Zhang A, Ma Q, Liu J J, Zhang X L, Huang L, Ma X F. Decryption of favourable haplotypes and potential candidate genes for five fibre quality properties using a relatively novel genome-wide association study procedure in upland cotton. *Industrial Crops & Products*, 2020, 158: 113004
- [14] Li Z H, Wang P C, You C Y, Yu J W, Zhang X N, Yan F L, Ye Z X, Shen C, Li B Q, Guo K, Liu N, Thyssen G N, Fang D D, Lindsey Y K, Zhang X L, Wang M J, Tu L L. Combined GWAS and eQTL analysis uncovers a genetic regulatory network orchestrating the initiation of secondary cell wall development in cotton. *New Phytologist*, 2020, 226 (6) : 1738-1752
- [15] Su J J, Li L B, Pang C Y, Wei H L, Wang C X, Song M Z, Wang H T, Zhao S Q, Zhang C, Mao G Z, Huang L, Wang C S, Fan S L, Yu S X. Two genomic regions associated with fiber quality traits in Chinese upland cotton under apparent breeding selection. *Scientific Reports*, 2016, 6 (1) : 38496
- [16] Shen C, Wang N, Huang C, Wang M J, Zhang X L, Lin Z X. Population genomics reveals a fine-scale recombination landscape for genetic improvement of cotton. *Plant Journal*, 2019, 99 (3) : 494-505
- [17] Islam M S, Thyssen G N, Jenking J N, Zeng L, Delhom C D, McCarty J C, Deng D D, Hinchliffe D J, Jones D C, Fang D D. A MAGIC population-based genome-wide association study reveals functional association of *GhRBB1_A07* gene with superior fiber quality in cotton. *BMC Genomics*, 2016, 17 (1) : 903
- [18] Li C Q, Fu Y Z, Sun R R, Wang Y Y, Wang Q L. Single-locus and multi-locus genome-wide association studies in the genetic dissection of fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Frontiers in Plant Science*, 2018, 9: 1083
- [19] Paterson A H, Brubaker C L, Wendel J F. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Kluwer Academic Publishers*, 1993, 11 (2) : 122-127
- [20] Sun X W, Liu D Y, Zhang X F, Li W B, Liu H, Hong W G, Jiang C B, Guan N, Ma C X, Zeng H P, Xu C H, Song J, Huang L, Wang C M, Shi J J, Wang R, Zheng X H, Lu C Y, Wang X W, Zheng H K. SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE*, 2013, 8 (3) : e58700
- [21] Zhang T Z, Hu Y, Jiang W K, Fang L, Guan X Y, Chen J D, Zhang J B, Saski C A, Scheffler B E, Stelly D M, Hulse-kemp A M, Wan Q, Liu C X, Wang S, Pan M Q, Wang Y K, Wang D W, Ye W X, Chang L J, Zhang W P, Song Q X, Kirkbride R C, Chen X Y, Dennis E, Llewellyn D J, Peterson D G, Thaxton P, Jones D C, Wang Q, Xu X Y, Zhang H, Wu H T, Zhou L, Mei G F, Chen S Q, Tian Y. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nature Biotechnology*, 2015, 33 (5) : 531-537
- [22] Su J J, Ma Q, Li M, Hao F S, Wang C X. Multi-locus genome-wide association studies of fiber-quality related traits in Chinese early-maturity upland cotton. *Frontiers in Plant Science*, 2018, 9: 1169
- [23] Bradbury P J, Zhang Z, Kroon D E, Casstevens T M, Ramdoss Y, Buckler E S. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 2007, 23 (19) : 2633 - 2635
- [24] 张爱, 王彩香, 宿俊吉, 张先亮, 史春辉, 刘娟娟, 彭云玲, 马雄凤. 陆地棉 *MADS-box* 家族基因鉴定及组织特异性表达分析. *棉花学报*, 2020, 32 (5) : 404-417
- Zhang A, Wang C X, Su J J, Zhang X L, Shi C H, Liu J J, Peng Y L, Ma X F. Identification of *MADS-box* family and analysis of tissue specific expression in *Gossypium hirsutum* L. *Cotton Science*, 2020, 32 (5) : 404-417
- [25] Wang F R, Zhang J X, Chen Y, Zhang C Y, Gong J W, Song Z Q, Zhou J, Wang J J, Zhao C J, Jiao M J, Liu A Y, Du Z H, Yuan Y L, Fan S J, Zhang J. Identification of candidate genes for key fibre-related QTLs and derivation of favourable alleles in *Gossypium hirsutum* recombinant inbred lines with *G. barbadense* introgressions. *Plant Biotechnology Journal*, 2020, 18 (3) : 707-720
- [26] Shan C M, Shangguan X X, Zhao B, Zhang X F, Chao L M, Yang C Q, Wang L J, Zhu H Y, Zeng Y D, Guo W Z, Zhou B L, Hu G J, Guan X Y, Chen Z J, Wendel J F, Zhang T Z, Chen X Y. Control of cotton fibre elongation by a homeodomain transcription factor GhHOX3. *Nature Communications*, 2014, 5: 5519
- [27] Tang W X, Tu L L, Yang X Y, Tan J F, Deng F L, Hao J, Guo K, Lindsey K, Zhang X L. The calcium sensor *GhCaM7* promotes cotton fiber elongation by modulating reactive oxygen species (ROS) production. *New Phytologist*, 2014, 202 (2) : 509-520

- [28] Zhang F, Jin X X, Wang L K, Li S F, Wu S, Cheng C Z, Zhang T Z, Guo W Z. A cotton annexin affects fiber elongation and secondary cell wall biosynthesis associated with Ca^{2+} influx, ROS homeostasis, and actin filament reorganization. *Plant Physiology*, 2016, 171 (3): 1750-1770
- [29] Han J, Tan J F, Tu L L, Zhang X L. A peptide hormone gene, *GhPSK* promotes fibre elongation and contributes to longer and finer cotton fibre. *Plant Biotechnology Journal*, 2014, 12 (7): 861-871
- [30] Cao J F, Zhao B, Huang C C, Chen Z W, Zhao T, Liu H R, Hu G J, Shangguan X X, Shan C M, Wang L J, Zhang T Z, Wendel J F, Guan X Y, Chen X Y. The miR319-targeted *GhTCP4* promotes the transition from cell elongation to wall thickening in cotton fiber. *Molecular Plant*, 2020, 13 (7): 1063-1077
- [31] Ge Z X, Zhao Y L, Liu M C, Zhou L Z, Wang L L, Zhong S, Hou A Y, Jiang J H, Liu T X, Huang Q P, Xiao J Y, Gu H Y, Wu H M, Dong J, Thomas D, Alice Y C, Qu L J. *LLG2/3* are co-receptors in BUPS/ANX-RALF signaling to regulate *Arabidopsis* pollen tube integrity. *Current Biology*, 2019, 29 (19): 3256-3265
- [32] 孙大千,董金晔,李洋,肖红庆,徐赫韩,李海燕,王法微. 磷脂酰肌醇特异性磷脂酶 C 基因研究进展. *生物技术*, 2017, 27 (1): 92-97
- Sun D Q, Dong J Y, Li Y, Xiao H Q, Xu H H, Li H Y, Wang F W. Advances in the study of the structure and function of phospholipase C. *Biotechnology*, 2017, 27 (1): 92-97
- [33] 肖光辉,梅文倩,朱玉贤. 棉纤维细胞发育过程中钙离子内流和钙依赖蛋白激酶活性的动态变化. *中国科学: 生命科学*, 2013, 43 (10): 886-896
- Xiao G H, Mei W Q, Zhu Y X. Ethylene and VLCFA promote cotton fiber growth by stimulating Ca^{2+} ion influx and activation of several Ca^{2+} -dependent protein kinases. *Scientia Sinica Vitae*, 2013, 43 (10): 886-896