Genome-wide association analysis of fiber fineness and yield in ramie (Boehmeria nivea) using SLAF-seq

Kunyong Huang · Yaliang Shi · Gen Pan · Yicheng Zhong · Zhimin Sun · Juan Niu · Jing Chen · Jianhua Chen · Mingbao Luan

Received: 30 June 2020 / Accepted: 19 December 2020 / Published online: 20 January 2021
© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

Abstract Ramie (Boehmeria nivea L. Gaudich) is a major natural fiber crop cultivated in East Asia. The improvement of its fiber yield and fineness are important breeding goals. Fiber yield is a complex quantitative trait comprising ramet number, stem diameter, plant height, skin thickness, and fiber percentage. The fiber fineness is a crucial trait for ramie fiber quality. However, there are few association studies for fiber yield traits and fiber fineness on ramie, and lack high-density SNP maps in natural ramie population. Here, a panel of 112 core ramie germplasms were genotyped by 215,376 consistent single-nucleotide polymorphisms (SNPs) from specific-locus amplified-fragment sequencing (SLAF-seq), and used for genome-wide association study of fiber fineness and yield. Subsequently, the genetic diversity, linkage disequilibrium (LD), and population structure was conducted based on 215,376 SNPs. Population cluster analysis disclosed five subpopulations. Neighbor-joining (NJ) analysis revealed three major clusters. No obvious relationships were identified between them and their geographic origins. The genome-wide LD decayed to $r^2 = 0.1$ was $\sim 11.75$ kbp in physical distance. One, seven, one, seven, and twenty-seven significant SNP marker associations were detected for fiber fineness (third season), stem diameter (third season), stem skin thickness (third season), fiber percentage (second season), and fiber percentage (third season), respectively. Two promising candidate genes, whole_GLEAN_10029622 and whole_GLEAN_10029638 resided in the significant trait-SNP association for fiber fineness (third season), which annotated as a cotton fiber-expressed protein and an Arabidopsis thaliana homebox protein ATH1, respectively and validated by qPCR. The identified loci or genes for fiber fineness and yield may provide the basis for future research on fiber fineness and yield and marker-assisted selection breeding for ramie.

Keywords Ramie · Genome-wide association study · Fiber fineness · Fiber yield · Single nucleotide polymorphism (SNPs)

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10681-020-02757-w) contains supplementary material, which is available to authorized users.

Kunyong Huang and Yaliang Shi have contributed equally to this work.

K. Huang · Y. Shi · G. Pan · Y. Zhong · Z. Sun · J. Niu · J. Chen · J. Chen · M. Luan
Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences/Key Laboratory of Stem-Fiber Biomass and Engineering Microbiology, Ministry of Agriculture, Changsha 410205, China
e-mail: luanmingbao@caas.cn
J. Chen
e-mail: chenjianhua@caas.cn
Introduction

Ramie (*Boehmeria nivea* L. Gaudich) is a perennial fiber crop and native to eastern Asia (Luan et al. 2015). About 87,620 ha of ramie were cultivated in China in 2016 (http://data.stats.gov.cn). Ramie is the second largest fiber crop in China (Luan et al. 2017). Fiber yield and quality limit ramie cultivation in China (Chen et al. 2016; Luan et al. 2009). While, fiber yield is a complex quantitative trait comprising ramet number, stem diameter, plant height, skin thickness, and fiber percentage (Luan et al. 2017). Qualitative fiber characteristics include length, strength, and fineness. The length and strength of both cotton and ramie fiber are suitable for textile manufacture. Unlike cotton, however, the fineness of ramie fiber is too lower for fabric production. Thus, it’s a crucial trait for fiber quality. Ramie fiber traits may be quantitative and regulated by numerous genes with small phenotypic effects (Xiong et al. 1998; Lu et al. 2016). Thus, understanding the genetic basis of these traits were critical for molecular designing breeding of the traits.

Owing to the next-generation sequencing (NGS), large number of molecular markers detected and used for marker-assisted selection. Recently, 76 SSR markers were developed from 320 expressed sequence tags (ESTs) of ramie from the National Center for Biotechnology Information (NCBI), of these, twenty-seven SSR loci show polymorphisms among 62 ramie individuals (Chen et al. 2011). Liu et al. (2013) identified 1827 EST-SSR markers. Chen et al. (2016) identified 4230 SSR loci and developed 2431 SSR markers from the ramie transcriptome at different developmental stages. Using SSR markers, Liu et al. (2014) and Luan et al. (2017) identified several agronomic trait-related quantitative trait loci (QTLs) in ramie. Nevertheless, it is difficult to use them for cloning or breeding programs due to these QTLs with large intervals. Compared with SSR markers, SNPs are more abundant and stable in plant genomes and their frequency is 1 in 100–300 bp (Edwards et al. 2007). With the application of the NGS technologies, it is possible to generated thousands of genome-wide SNPs markers for high-density genetic maps (Davey et al. 2011; Qi et al. 2014; Zhang et al. 2016a). Several sequencing-based technologies were developed for SNP discovery and high-throughput genotyping, such as two-enzyme genotyping-by-sequencing (GBS) (Poland et al. 2012), restriction site-associated sequencing (RADseq) (Miller et al. 2007), and double-digest RADseq (Peterson et al. 2012). Recently, another high-throughput DNA sequencing technology, specific-locus amplified-fragment sequencing (SLAF-seq) was developed Sun et al. (2013) SLAF-seq had successfully create a high-density SNP maps for soybean (*Glycine max* L.) (Li et al. 2014), cucumber (*Cucumis sativus* L.) (Wei et al. 2014), sesame (*Sesamum indicum* L.) (Zhang et al. 2013) and upland cotton (*Gossypium hirsutum* L.) (Zhang et al. 2016b). However, there are few reports of high-density SNP maps for ramie.

Using the genotyping-by-sequencing technique, Liu et al. (2017) and Chen et al. (2018) discovered numerous SNPs and detected QTLs associated with ramie fiber yield. However, the QTLs of Liu et al. (2017) were based on the F2 agamous line population of the varieties Qingdaye and Zhongzhu 1. Chen et al. (2018) focused only on the ramet number and made no reference to the ramie genome. There are still lack studies assess the genetic variation and GWAS for fiber yield and fineness in natural ramie accessions. Therefore, research is needed to assess the genetic variation and association analysis of fiber yield and fineness on natural ramie accessions using the *Boehmeria nivea*.

Genome-wide association studies (GWAS) are an efficient strategy for detect genomic regions controlling complex traits within unstructured crop germplasms (Kang et al. 2015). In the present study, we used SLAF-seq technology perform a GWAS of 112 ramie accessions (107 core and 5 cultivar accessions). The aims of this study were to (1) develop high-density SNP map and genetic variants; (2) assess the genetic diversity of the *B. nivea* germplasm pool; (3) investigate the population structure and estimate the linkage disequilibrium (LD) across the accessions, and (4) identify single-nucleotide polymorphisms (SNPs) and candidate genes associated with fiber fineness and yield.

Materials and methods

Plant materials and phenotyping

A set of 112 ramie accessions from the National Infrastructure for Ramie Germplasm Resources of China were selected for this study (Table S1),
including 106 accessions originally collected from 11 provinces of China (96 landrace accessions and 10 cultivated varieties), one landrace accession from Brazil, one landrace accession from India, one landrace accession from Cuba, and three landrace accessions from Indonesia. The germplasm was selected to represent variations in fiber fineness and yield. Briefly, the accessions were vegetatively propagated from the National Infrastructure for Ramie Germplasm Resources of China and planted in our scientific research field (Wangcheng Experimental Station of the Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences, Changsha, China) at 2014, with two replicates, randomized block design. Each plot consisted of one accession and five clones planted 1 m apart in a single row (3 m long × 1 m wide). Standard cultural practices were performed according to Luan et al. (2017). For plant height and stem diameter were measured in three successive seasons of 2015 from March to June (environment 1), June to August (environment 2), and August to October (environment 3). While, ramie skin thickness and fiber percentage were measured in two successive seasons of 2015 from June to August (environment 2), and August to October (environment 3) in 2015. For the fiber fineness, three successive harvest seasons were August to October in 2014 (environment 1), June to August (environment 2) and August to October (environment 3) in 2015. All ramie plants were planted under the same field conditions and were under conventional agricultural management.

The three harvest periods represented different phases of phenological development, so they were analyzed separately. Thirteen plant traits were measured and the data were applied to the association mapping (Table 1). Measurements were made according to the methods of Luan et al. (2017). Briefly, 15 ramets per plot were randomly selected to measure plant height, stem diameter, skin thickness, and fiber percentage using tape measure, Vernier caliper, and screw micrometer, respectively. Fiber fineness was examined with optical fiber diameter analyzer systems (OFDA 100 and OFDA 2000; HORNIK FIBER-TECH, Wald, Switzerland) according to Wang and Wang (2004). Fiber percentage was calculated as follows:

\[
\text{Fiber }\% = \frac{\text{FW (after bast decortication)}; \text{dry bast fiber weight)}}{\text{BW (fresh bast weight)}} \times 100
\]

SLAF sequencing and data evaluation

Young healthy leaves were excised from plants growing in the Wangcheng Experimental Station in Changsha, China (Wangcheng) for DNA extraction with a DNeasy plant mini kit (Qiagen, Germany). DNA quality and quantity were assessed with a NanoDropTM 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

Ramie genomic DNA was analyzed by SLAF-seq (Sun et al. 2013) as described in Zhou et al. (2017). Briefly, genomic DNA was digested with restriction enzyme Rsai and HaeIII (New England Biolabs Inc., Ipswich, MA, USA), followed by fragment-end repair, duplex tag-labeled adapter ligation, PCR amplification, and size-selected for SLAF libraries. Finally, the SLAF libraries were sequencing using an Illumina HiseqTM 2500 (Illumina, San Diego, CA, USA) at Biomarker Technologies Corporation in Beijing, China. The average sequencing depth was 10.89 × . After sequencing, the low-quality reads (quality score < 20) were filtered out using dual-index software (Kozich et al. 2013). All high-quality reads were mapped onto the Boehmeria nivea reference genome (Luan et al. 2018) using the Burrows-Wheeler alignment tool (BWA) (Li and Durbin 2009) with default parameters.

SNP calling

The software GATK (McKenna et al. 2010) and SAMtools (Li et al. 2009a) were used to discover SNPs, and SNPs detected by both methods were treated as reliable. After SNPs with minor allele frequency (MAF) of less than 5% and missing data of more than 20% were filtered out, a total of 215,376 consistent SNPs remained and were used for analysis.

Phylogenetic tree and principal component analysis

Based on the 215,376 consistent SNPs identified from the 112 ramie samples, a phylogenetic tree was constructed in MEGA 5 (Saitou and Nei 1987) using
the neighbor-joining algorithm (Eickmeyer et al. 2008). The genetic structure of the panel was assessed using principal components analysis (PCA) with GAPIT software (Lipka et al. 2012).

Deducing population structure and estimating linkage disequilibrium (LD)

Population structure was investigated using Admixture software (Alexander et al. 2009) based on those consistent SNPs. The number of subgroups (K) was varied from 1 to 10, which represented the assumed groups of simulated population in ancient times. The best value of K was determined by the cross-validation (CV). The PLINK 1.9 BETA (Purcell et al. 2007; Shaun et al. 2007) was used to obtain the required data files for ADMIXTURE. LD was estimated between SNP pairs located on the same scaffold based on the alignment of reference genome *Boehmeria nivea* (Luan et al. 2018). The ramie chromosome sequence was not publicly available at that time. The 215,376 consistent SNPs were used to infer LD in TASSEL (Bradbury et al. 2007). The $r^2$ were obtained for all SNP pairs using Plink (Purcell et al. 2007).

Association mapping

Three Genome-wide association studies (GWAS) models were used to association analysis. The models GLM-Q and compressed MLM-Q + K were analysis in TASSEL (Bradbury et al. 2007), kinship (K) matrix was estimated using SPAGe Di v. 1.4b (Hardy and Vekemans 2002) and the best admixture results (Q) representing population membership served as covariates in the model. The EMMAX model were analysis using identify-by-state (IBS) relationship matrix according to Kang et al. (2015).

Based on the number of SNPs analyzed ($n = 215,376$), the significance threshold was estimated use the Bonferroni correction method, it was $P = 4.64 \times 10^{-7}$ (0.1/215,376). SAMtools (Li et al. 2009a) was used for the manual verification of regions of the aligned sequencing reads with significant association relative to the reference genome of *Boehmeria nivea* (Luan et al. 2018).

Identification of candidate genes and quantitative PCR (qRT-PCR) analysis

Functionally characterized genes harboring the detected SNPs or within a 100-kbp interval were sought in the reference genome of *Boehmeria nivea*. Then qRT-PCR was used to determine differential expression of the genes putatively controlling fiber fineness, stem diameter, and skin thickness. Based on our phenotypic data of field experiment, we selected six accessions (1–26 (kk), 2–34 (ct), 2–35 (bo), 1–43 (aw), 2–41 (kr), and 2–7 (ba)) for RT-qPCR. Bast samples were collected from those six accessions at the vegetative stage (~ 30 cm height) and freezing in liquid nitrogen immediately. Total RNA was extracted with a TaKaRa MinBEST plant RNA extraction kit (TaKaRa Bio Inc., Shiga, Japan) and reverse-transcribed with a Thermo Scientific RevertAid first-

| Table 1 | Summary statistics of phenotypes used in GWAS |
|---------|---------------------------------------------|
| Trait   | Unit | Min     | Max     | Mean    | SD      | CV   |
| FF (environment 1) | m/g | 1062.00 | 2869.00 | 1616.00 | 324.16  | 0.20 |
| FF (environment 2) | m/g | 916.00  | 2412.50 | 1429.00 | 292.50  | 0.20 |
| FF (environment 3) | m/g | 965.50  | 2815.50 | 1796.00 | 390.13  | 0.22 |
| RH (environment 1) | cm  | 93.83   | 213.83  | 152.30  | 19.53   | 0.13 |
| RH (environment 2) | cm  | 61.82   | 145.55  | 109.80  | 14.92   | 0.14 |
| RH (environment 3) | cm  | 46.00   | 129.50  | 96.00   | 16.75   | 0.17 |
| RSD (environment 1) | mm | 6.68    | 11.78   | 9.20    | 1.25    | 0.14 |
| RSD (environment 2) | mm | 5.03    | 10.59   | 7.83    | 1.08    | 0.14 |
| RSD (environment 3) | mm | 3.54    | 21.57   | 7.02    | 2.35    | 0.33 |
| SKT (environment 2) | mm | 0.48    | 1.19    | 0.77    | 0.13    | 0.18 |
| SKT (environment 3) | mm | 0.30    | 0.73    | 0.49    | 0.09    | 0.18 |
| FP (environment 2) | g/g | 0.03    | 0.25    | 0.12    | 0.03    | 0.29 |
| FP (environment 3) | g/g | 0.02    | 0.25    | 0.11    | 0.04    | 0.35 |

*FF fiber fineness, RH plant height, RSD stem diameter, SKT ramie skin thickness, FP fiber percentage, SD standard deviation, CV coefficient of variance.
strand cDNA synthesis kit (Thermo Scientific, Vilnius, Lithuania) according to the manufacturer’s instructions. The qRT-PCR was performed in a Monad selected q225 real-time PCR system (Monad Biotech Co. Ltd., Shanghai, China) in a 10 μL reaction volume containing 0.5 μL cDNA sample, 5 μL MonAmpTM SYBR Green qPCR mix (Monad Biotech Co. Ltd., Shanghai, China), 0.5 μL each forward and reverse primers (10 μM each), and 3.5 μL nuclease-free water. The PCR parameters were 95 °C for 1 min followed by 40 cycles of 95 °C for 5 s and 55 °C for 30 s. Each sample was analyzed in triplicate. Relative expression were calculated as previously described (Livak and Schmittgen 2001). All qRT-PCR primers (Table S2) were designed by the Primer3 online tool (http://primer3.ut.ee/) and adjusted with Oligo v. 7.56 (Rychlik 2007). The internal control of choice was 18S based on preliminary experiments (data not shown). The qRT-PCR data were analyzed in Microsoft Excel and reported as mean ± SD.

Validation SNP marker of fiber fineness through polymerase chain reaction (PCR)

In order to verified the obtained SNP (PHNS01005842.1:2705313) can be used for molecular marker research, we selected another 133 ramie accessions for genomic PCR amplification. The fiber fineness of those 133 ramie accessions was shows in Table S3, which from our field experiment of our laboratory. We design a pair of PCR primers (F-primer: AAGCTTGAATAACGATCTCCGA; R-primer: CAAATACGCTTCTGAATGAACCTGT; the 3’ end of F-primer was on the location of SNP PHNS01005842.1:2705313) and followed by a base that manually introduced mismatch using the Primer3 online tool (http://primer3.ut.ee/). The PCR was performed in a 25 μL reaction volume containing 1 μL genome DNA, 12.5 μL PCR mix (Beijing GreatOcean Biotech Co. Ltd., Beijing, China), 0.25 μL each forward and reverse primers (10 μM each), and 11 μL nuclease-free water. The PCR parameters were 94 °C for 3 min followed by 32 cycles of 94 °C for 30 s, 58 °C for 30 s and 72 °C for 15 s, and terminal elongation at 72 °C for 5 min. The PCR products (8 μL/sample) were analyzed on 1.5% agarose gel and visualized after ethidium bromide staining under UV light. One fragment of 199 bp will be amplified by the accessions with A allele. Correlation analysis using R software with “psych” package was performed to determine if the fiber fineness had any correlation with the SNPs.

### Results

Specific-locus, amplified-fragment sequencing and SNP calling

The restriction enzymes RsaI and HaeIII were selected based on predictions from in silico digestion. There were 205,334 predicted SLAF tags 264–394 bp in length. The digestion efficiency was 93.30%. In this experiment, 465.13 Gbp of sequences were generated from the 112 ramie accessions. There were 502,578 SLAF tags identified throughout the genome. The average depth per sample was 12.34 × . There were 369,997 SLAF tags with polymorphisms. All SLAF tags aligned to the reference genome using BWA software (Li and Durbin 2009). The PLINK software (Purcell et al. 2007) performed quality control on the data. A total of 2,955,337 SNP loci were identified by GATK and SAMtools. More than 85.9% of the SNP variant effects were in the intergenic, intron, downstream, and upstream areas of the coding region whereas only 13.58% were in the exons (Table 2).

| Type                          | Count    | Percent (%) |
|-------------------------------|----------|-------------|
| Intergenic                    | 773,264  | 26.17       |
| Intron                        | 438,233  | 14.83       |
| Upstream                      | 686,213  | 23.22       |
| Downstream                    | 640,610  | 21.68       |
| Splice_site_acceptor          | 741      | 0.025       |
| Splice_site_donor             | 661      | 0.022       |
| Splice_site_region            | 12,155   | 0.41        |
| Exon (start lost)             | 440      | 0.015       |
| Exon (synonymous_coding)      | 195,519  | 6.62        |
| Exon (nonsynonymous coding)   | 195,860  | 6.63        |
| Exon (synonymous_stop)        | 309      | 0.010       |
| Exon (stop_gained)            | 5892     | 0.20        |
| Exon (stop_lost)              | 402      | 0.014       |
| Other                         | 5038     | 0.17        |
| Total                         | 2,955,337| 100         |
After filtration for integrity (> 80%) and low minor allele frequency (> 5%), 215,376 highly consistent SNPs were identified and used in the subsequent analyses.

Population genetics analysis

Genetic relationships among the 112 samples were calculated using the 215,376 highly consistent SNPs identified in the present study. Among these 12,432 pairwise combinations, 11,644 (93.662%) with genetic relationship coefficients < 0.05 (Fig. 1a). Only four pairwise combinations (0.032%) (two pairs accessions: cd and br, ce and dr) (Table S1), which with higher kinship value between 0.6 to 0.65, suggesting the two pairs accessions may have closely relationship. Thus, most of accessions were very weak or no relationships in this study.

The population structure analysis indicated that the 112 associations panel derived from five ancestors based on cross-validation (CV) errors (Fig. 2). Of these 112 accessions, 57 were group one, 11 were group two, 15 were group three, 18 were group four, and the remaining 29 originated from Hubei (6), Guangxi (5), Guizhou (5), Sichuan (5), Yunnan (2), Hainan (2), Zhejiang (1), Chongqing (1), India (1), and Shanxi (1). Group two accessions were collected from Sichuan (2), Guangxi (2), Shanxi (2), Indonesia (2), Zhejiang (1), Jiangxi (1), and Hunan (1). Group three accessions were derived from Guizhou (4), Chongqing (3), Hunan (2), Jiangxi (2), Sichuan (1), Hubei (1), Shanxi (1), and Brazil (1). Group four accessions were gathered from Hunan (4), Guizhou (3), Sichuan (3), Hubei (2), Jiangxi (2), Yunnan (1), Hainan (1), Guangxi (1), and Cuba (1). Group five accessions were acquired from Hunan (3), Jiangxi (2), Sichuan (2), Guangxi (1), Shanxi (1), Guizhou (1), and Indonesia (1).

Neighbor-joining (NJ) clustering show this set of germplasm formed three major groups (Fig. 3). Based on the Admixture cross-validation (CV) figure (Fig. 2), the second choice consisted of three subpopulations (K) when the cross-validation error was the second lowest. Another selection comprised three subgroups for the 112 core ramie germplasms. Principal component analysis (PCA) showed genotype dispersal among the various components, and the first three PCs collectively explained only about 11.13% of the total variance, suggesting that the population structure of the 112 accessions was very low (Fig. 1b). No significant correlation was detected between genetic relationships and geographic origins.

![Fig. 1](https://example.com/figure1.png)  
Fig. 1  Kinship and PCA analysis of the 112 ramie accessions. A, Kinship hist of the 112 accessions. B, PCA analysis in 112 accession of ramie

© Springer
Linkage disequilibrium across the whole ramie genome

LD decayed was estimated between SNP pairs along each scaffolds using the 215,376 highly consistent SNPs. The genome-wide LD decayed to $r^2 = 0.1$ was ~11.75 kbp in physical distance (Fig. 4), which is lower than that reported for most other crops. There are major difference among the various scaffolds. LD ranged from 1.09 kbp (Scaffold11 (PHNS01006314.1)) to 20.95 kbp (Scaffold2) (Table S4).

GWAS of phenotypic traits and quantitative real-time PCR analysis of candidate genes

The GWAS models tested included GLM-Q, compressed MLM-Q + K, and EMMAX. According to the quantile–quantile (Q–Q) plot results, the GLM-Q model was strongly skewed towards significance for most traits (Fig. S1). Therefore, the Q matrix did not adequately account for the population structure or cryptic relatedness. The MLM-Q + K used both the Q and kinship matrices overcorrected these confounders (Fig. S2). The EMMAX model use the identity-by-state (IBS) relationship matrix and effectively reduced the false positive rate. It had sufficient statistical component. Each line indicates the subgroup value. Each accession is represented by a vertical line partitioned into $K$ colored components which represent the inferred memberships in $K$ genetic clusters.

Fig. 2 Population structure inferred by Admixture analysis. The accessions were divided into five subgroups. There were minimum cross-validation errors when $K = 5$. Within each subgroup, the accessions were ordered according to their genetic component.
power to identify significant marker-trait associations (Fig. S3). Therefore, the EMMAX-Q model was applied for GWAS in this study.

Variations in all 13 traits were detected among the genotypes (Table 1). We tested the associations between the 215,376 SNPs and each trait using the EMMAX model. Based on the chosen $P < 4.64 \times 10^{-7}$, we identified 44 trait-SNP associations for five traits, namely, FF (third season), RSD (third season), SKT (third season), FP (second season), and FP (third season). These were designated the top QTL candidates (Table S5). The association pattern across the genome scaffolds visualized with Manhattan plots showed that association trait-SNPs were scattered throughout the genome (Fig. 5, Fig. S4–S6). We arbitrarily searched 100 kbp upstream and downstream of these trait-SNP association loci already known for candidate genes in the *Boehmeria nivea*...
cultivar Zhongzhu No. 1 genome. We predicted the number of candidate genes associated with the five aforementioned traits. We found 30 candidate genes from the significant SNP regions (PHNS01005842.1-2705313) associated with FF of the third season (Table S6). Whole_GLEAN_10029631 and...
whole_GLEAN_10029638 located at 66.7 kbp and 60.4 kbp from the trait-SNP, respectively, were promising candidate genes encoding a cotton fiber-expressed protein of unknown function and an Arabidopsis thaliana homeobox protein ATH, respectively.

In order to validate the obtained SNP (PHNS01005842.1:2705313) can be used for molecular marker research, we design a pair of allele-specific PCR primer and use the primer to identify the different fiber fineness accessions. Correlation analysis shows the fiber fineness was significantly associated ($P < 0.01$) with the amplification results, suggests that the SNP (PHNS01005842.1:2705313) was associated with fiber fineness in ramie (Fig. 6, Table S3).

Seven significant SNP markers were associated with RSD of the third season (Table S7). The most significantly associated SNP marker was at PHNS01002615.1:385693 ($-\log_{10} P = 7.47$), which corresponded to the transcripts of whole_GLEAN_10015232 annotated as AT-rich interactive domain-containing protein 2 whose function is unknown. A subsequent survey identified the following promising candidate genes: a zinc finger protein CONSTANS-LIKE 3 gene (whole_GLEAN_10015241) located 63.3 kbp from the significant SNP (PHNS01002615.1:385693), a cotton fiber-expressed protein of unknown function (whole_GLEAN_10025327) located 9.2 kbp from the significant SNP (PHNS01003751.1:2205139), a gene (whole_GLEAN_10026480) encoding the trihelix transcription factor located 97 kbp from the significant SNP (PHNS01007343.1:2569037), and a gene (whole_GLEAN_10015233) encoding the E3 ubiquitin-protein ligase located 5.2 kbp from the significant SNP (PHNS01002615.1:385693).

One SNP marker (PHNS01006314.1:1911874) was associated with SKT of the third season (Table S8). The transcript of whole_GLEAN_10021907 was annotated as Arabidopsis thaliana G-box-binding factor 1 and was located 31.5 kbp from the marker. The transcript of whole_GLEAN_10021909 was annotated as Arabidopsis thaliana homeobox-leucine zipper protein ATHB-40 and was located 13 kbp from the marker. The transcript of whole_GLEAN_10021913 was annotated as Arabidopsis thaliana BES1/BZR1 homolog protein 2 and was located 22.4 kbp from the marker. All of these were promising candidate genes.

For FP of the second season, seven SNP markers were detected above the threshold ($P = 4.64 \times 10^{-7}$) in the scaffolds PHNS01004650.1, PHNS01011408.1, PHNS01001065.1, PHNS01005672.1, PHNS01004146.1, PHNS01003683.1, and PHNS010007343.1 (Table S9). One Arabidopsis thaliana homeobox-leucine zipper protein HDG5 homolog (whole_GLEAN_10016513) and one Arabidopsis thaliana transcription factor KAN2 homolog (whole_GLEAN_10016511) were located 22.1 kbp and 81.2 kbp, respectively, from the trait-SNP PHNS01003683.1:866079. One Arabidopsis thaliana WAT1-related protein At3g18200 homolog (Whole_GLEAN_10007759) was located 71.9 kbp from SNP PHNS01011408.1:335775 and three Arabidopsis thaliana transcription factor ORG2 homologs (whole_GLEAN_10011656, whole_GLEAN_10011657, and whole_GLEAN_10011658) were located at 74.8 kbp, 83.8 kbp, and 88.8 kbp, respectively, from the significant SNP PHNS01004146.1:755398. These were promising candidate genes.

Twenty-seven significant GWAS loci among 17 scaffolds were identified for FP of the third season (Table S10). There were five SNP markers in scaffold PHNS01000838.1 within whole_GLEAN_10010742 whose function is unknown. The whole_GLEAN_10006587 homolog of Trema orientalis GRAS transcription factor was located 30.4 kbp from the marker PHNS01001992.1:258552. The whole_GLEAN_10025406 homolog of Arabidopsis thaliana Auxin-responsive protein IAA13 was located 0.7 kbp from PHNS01003751.1:2871833. The whole_GLEAN_10029892 ortholog of Trema orientalis GRAS transcription factor was located 48.7 kbp from PHNS01005842.1:4650339. Whole_GLEAN_10017958 was annotated as the NAC domain-containing protein (NAC29) and was located 97 kbp from PHNS01010418.1:933907.

The candidate genes of FF (third season), RSD (third season), SKT (third season), and FP (second season) were validated by qRT-PCR. Whole_GLEAN_10029631 and whole_GLEAN_10029638 were comparatively upregulated in varieties with relatively higher fiber fineness (Fig. 7). Whole_GLEAN_10021913 was comparatively downregulated in thicker SKT varieties. No other candidate
genes were correlated with the RSD (third season), SKT (third season), or FP (second season) for six accessions by qRT-PCR (Fig. S7–S9).

Discussion

Population structure and linkage disequilibrium

The phylogenetic tree constructed using neighbor-joining algorithm based on the genetic similarity coefficients among individuals to reflects genetic relationships. In contrast, the population structure analysis using Admixture software, which estimates
individual ancestries by efficiently computing maximum likelihood in a parametric model. In the present study, three major clusters were defined in a dendrogram. The second lowest cross-validation error was obtained for K = 3. The 112 accessions may be partitioned into three subgroups. No obvious relationship was detected between the accessions and their geographic origins. Indicate these lines might have undergone introgression or gene flow during ramie breeding in China. This finding was consistent with those in previous reports on *Boehmeria nivea* (Luan et al. 2014, 2015; Zhu et al. 2017).

In the present study, we used the reference genome of *Boehmeria nivea* zhongzhu1 based on SLAF-seq and identified 215,376 highly consistent SNPs which was ~ 2 × the number of SNPs (108,888) reported by Chen et al. (2018). This abundance of SNPs will be useful for constructed a precise haplotype map and LD studies (Buckler and Gore 2007; Gore et al. 2009).

It is crucial to the LD decay of the genome for identifying the number and density of molecular markers for GWAS and their analysis methods (Myles et al. 2009). Ramie is predominantly a cross-pollinated or clonally propagated species. Thus, it should present with lower LD levels than those of self-pollinated crops. Despite several earlier association studies on ramie (Luan et al. 2017; Liu et al. 2017; Chen et al. 2018), none of them reported any LD decay values. Here, the genome-wide LD decayed to $r^2 = 0.1$ was 11.75 kbp in physical distance. This value was substantially lower than those for most other crops such as rice (*Oryza sativa* L.) (167 kbp) (Huang et al. 2010), soybean (420 kbp) (Zhou et al. 2015), and cultivated maize (*Zea mays* L.) (30 kbp) (Hufford et al. 2012), but higher than that of cultivated cassava (*Manihot esculenta* L.) (8 kbp) (Zhang et al. 2018), inbred maize (*Zea mays* L.) (1.5 kbp) (Remington et al. 2001), and inbred potato (*Solanum tuberosum* L.) (1 kbp) (Simko 2004). Therefore, the 112 ramie accessions of the present study should have abundant allelic diversity based on the rapid LD decay in most of the scaffolds. In this way, the search for candidate genes is facilitated through efficient narrowing of the putative QTL regions.

Genome-wide association study (GWAS)

The selection of best statistical models can accurately evaluate the associations between markers and
phenotypes. As the availability of genotype data increases, comprehensive statistical models are required to distinguish true biological associations from the false positives arising from population structure and LD (Wang and Qin 2017). Numerous statistical models are available to calculate the significance of the associations between SNPs and traits (Ogura and Busch 2015). Luan et al. (2017) detected 16 stable molecular markers from an association analysis of fiber yield traits in ramie based on the MLM-Q + K model. However, this model considers the effects of Q for the population structure, analyzes K for the genetic relationships, and Bonferroni correction is required for multiple tests. When marker number is very large, this test is too strict (Hou et al. 2018). In terms of the Q-Q for our results, MLM-Q + K overcorrected the confounders while GLM-Q was strongly skewed towards significance for every trait. EMMAX used the identity-by-state (IBS) relationship matrix, effectively controlled the false positive rate, and provided suitable statistical power to identify significant marker-trait associations. Therefore, the identity-by-state (IBS) relationship matrix was appropriate for ramie GWAS in natural populations.

The discovery of QTL for agronomically and economically important crop traits was of vital importance for marker-assisted breeding. To our best knowledge, this is the first report on the QTL for fiber yield and fineness based on the reference genome of Boehmeria nivea (Luan et al. 2018). We found one, seven, one, seven, and twenty-seven significant SNP marker associations for FF (third season), RSD (third season), RST (third season), FP (second season), and FP (third season), respectively. The associated trait-SNP (PHNS01005842.1:2705313) for fiber fineness was validated by genomic DNA PCR, indicated this SNP can be serve as unique tools for molecular-assisted breeding.

Two candidate genes in the significant trait-SNP associations were identified for FF (third season). The whole_GLEAN_10029622, Pfam annotation encodes a cotton fiber-expressed protein of unknown function. The head of the protein includes the domain DUF4408 and the tail is the cotton fiber-expressed protein domain DUF761. The whole_GLEAN_10029638 is an ortholog of the Arabidopsis thaliana homeobox protein ATH1. In Arabidopsis, the TALE homeobox gene ATH1 positive regulation of the floral repressor FLC, controls floral competency (Proveniers et al. 2007). The qPCR validation disclosed that these genes were upregulated in three fineness accessions compared to their expression levels in three fiber thickness accessions. Therefore, these two may be candidates for fiber fineness. The transcript of the whole_GLEAN_10021913 homolog of Arabidopsis thaliana BES1/BZR1 protein 2 was associated with SKT of the third season. It was a downstream target gene of the brassinosteroid (BR) signal. BRs regulate multiple biological processes including vasculature differentiation, cell elongation, senescence, morphogenetic, and stress response (Yu et al. 2011). In Arabidopsis, AtMYB30 is a direct target of BES1 and co-regulates brassinosteroid-induced gene expression with BES1 (Li et al. 2009b). Chai et al. (2014) reported that PdMYB10 and PdMYB128 in Arabidopsis increased stem fiber cell wall thickness and delayed flowering. Overexpression of PdMYB90, PdMYB167, PdMYB92, and PdMYB125 in Arabidopsis decreased stem fiber and vessel cell wall thickness and promoted flowering. The qPCR validation showed that whole_GLEAN_10021913 expression was relatively lower in thick SKT varieties. Thus, it is a promising candidate gene associated with SKT of the third season. Another candidate gene, whole_GLEAN_10015241, was associated with RSD of the third season. It was annotated as a zinc finger protein CONSTANS-LIKE 3 gene. The Arabidopsis CONSTANS gene promotes flowering and its overexpression results in earlier flowering than the wild type (Putterill et al. 1995). Early flowering reduces the time for vegetative growth and results in comparatively smaller plants. In contrast, late flowering prolongs vegetative growth which culminates in relatively larger plants (Lenser and Theissen 2013). The candidate gene whole_GLEAN_10016511 was associated with PC of the second season. It is a homolog of the Arabidopsis thaliana transcription factor KAN2 gene. In Arabidopsis thaliana, the transcription factor KAN2 regulates lateral organ polarity and determines abaxial cell fate during lateral organ formation (McCabe et al. 2006). KAN genes participate in ovule development. The growth of the outer integument observed is retarded in certain kan mutant combinations (Eshed et al. 2001). The qPCR indicates that the relative expression levels of these two genes did not correlate with the RSD or PC phenotypes for six varieties (Fig. S7–S9). This anomaly merits further
investigation. In the present study, the GWAS results were limited due to the small population size. Comprehensive phenotypic studies are currently in progress, and the future studies on association mapping in a wide range of samples may be performed.

Conclusions

This study provided information for understanding the population structure, LD decay and genetic variation in ramie natural population. The identified SNPs or genes for fiber fineness and yield may provide the basis for future research on fiber fineness and yield and marker-assisted selection breeding for ramie.

Acknowledgements

We would like to thank Jun Liu for the collecting of sample tissues, and Qing Tang for paper revising.

Authors’ contributions

KH, YS, ML, and JC conceived and designed the experiments; KH, YS, GP, YZ, ZS, JN, JC, ML, and JC performed the experiments; and KH, YS, ML, and JC analyzed the data and wrote the paper.

Funding

This study was financially supported by the National Natural Science Foundation of China (Grant Nos. 31671744 and 31471547), the Hunan Provincial Natural Science Foundation of China (Grant No. 2018JJ2465), and the National Modern Agro-industry Technology Research System (Grant No. CARS-19-E04).

Data availability

The genotypic data used in this study are included in the manuscript and supplementary materials.

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

References

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664
Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2632–2635
Buckler E, Gore M (2007) An Arabidopsis haplotype map takes root. Nat Genet 39:1056–1057
Chai G, Wang Z, Tang X, Yu L, Qi G, Wang D, Yan X, Kong Y, Zhou G (2014) R2R3-MYB gene pairs in Populus: evolution and contribution to secondary wall formation and flowering time. J Exp Bot 65:4255–4269
Chen J, Luan M, Song S, Zou Z, Wang X, Xu Y, Sun Z (2011) Isolation and characterization of EST-SSRs in the Ramie. Afr J Microbiol Res 5:3504–3508
Chen J, Yu R, Liu L, Wang B, Peng D (2016) Large-scale developing of simple sequence repeat markers and probing its correlation with ramie (Boehmeria nivea L.) fiber quality. Mol Genet Genom 291:753–761
Chen K, Luan M, Xiong H, Chen P, Chen J, Gao G, Huang K, Zhu A, Yu C (2018) Genome-wide association study discovered favorable single nucleotide polymorphisms and candidate genes associated with ramet number in ramie (Boehmeria nivea L.). BMC Plant Biol 18:345
Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:99–110
Edwards D, Batley J, Cogan N, Forster J, Chagre D (2007) Single nucleotide polymorphism discovery. In: Oraguzie N, Rikkerink E, Gardiner S, DeSilva H (eds) Association mapping in plants. Springer, New York, pp 53–76
Eickmeyer K, Huggens P, Pachter L, Yoshida R (2008) On the optimality of the neighbor-joining algorithm. Algorithm Mol Biol 3:5
Eshed Y, Baum SF, Perea JV, Bowman JL (2001) Establishment of polarity in lateral organs of plants. Curr Biol 11:1251–1260
Gore MA, Chia JM, Elshire RJ, Sun Q, Ersoz ES, Hurwititz BL et al (2009) A first-generation haplotype map of maize. Science 326:1115–1117
Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyses spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620
Hou S, Zhu G, Li Y, Li W, Fu J, Niu E, Li L, Zhang D, Guo W (2018) Genome-wide association studies reveal genetic variation and candidate genes of drought stress related traits in cotton (Gossypium hirsutum L.). Front Plant Sci 9:1276
Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y et al (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961–967
Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA et al (2012) Comparative population genomics of maize domestication and improvement. Nat Genet 44:808–811
Kang Y, Sakiroglu M, Krom N, Stanton-Geddes J, Wang M, Lee YC, Young ND, Udvardi M (2015) Genome-wide association of drought-related and biomass traits with HapMap SNPs in Medicago truncatula. Plant Cell Environ 38:1997–2011
Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. Appl Environ Microb 79:5112–5120
Lenser T, Theissen G (2013) Molecular mechanisms involved in convergent crop domestication. Trends Plant Sci 18:704–714
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009a) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

Li L, Yu X, Thompson A, Guo M, Yoshida S, Asami T, Chory J, Yin Y (2009b) Arabidopsis MYB30 is a direct target of BES1 and cooperates with BES1 to regulate brassinosteroid-induced gene expression. Plant J 58:275–286

Li B, Tian L, Zhang J, Huang L, Han F, Yan S, Wang L, Zheng H, Sun J (2014) Construction of a high-density genetic map based on large-scale markers developed by specific length amplified fragment sequencing (SLAF-seq) and its application to QTL analysis for isoflavone content in Glycine max. BMC Genom 15:1086

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28(18):2397–2399

Liu T, Zhu S, Fu L, Tang Q, Yu Y, Chen P, Lanu M, Wang C, Tang S (2013) Development and characterization of 1,827 expressed sequence tag-derived simple sequence repeat markers for ramie (Boehmeria nivea L. gaud). PLoS ONE 8:e603464

Liu T, Tang S, Zhu S, Tang Q (2014) QTL mapping for fiber yield-related traits by constructing the first genetic linkage map in ramie (Boehmeria nivea L. gaud). Mol Breed 34:883–892

Liu C, Zhu S, Tang S, Wang H, Zheng X, Chen X, Dai Q, Liu T (2017) QTL analysis of four main stem bark traits using a GBS-SNP-based high-density genetic map in ramie. Sci Rep 7:13458

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408

Lu K, Wang T, Xu X (2016) Genome-wide association analysis of height of podding and thickness of pod canopy in Brassica napus. Acta Agron Sin 42:344–352

Luan M, Qin Z, Chen J, Wang X, Xu Y, Sun Z, Mart H, Abecasis G, Durbin R (2009a) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408

Liu T, Zhu S, Fu L, Tang Q, Yu Y, Chen P, Lanu M, Wang C, Tang S (2013) Development and characterization of 1,827 expressed sequence tag-derived simple sequence repeat markers for ramie (Boehmeria nivea L. gaud). PLoS ONE 8:e603464

Liu T, Tang S, Zhu S, Tang Q (2014) QTL mapping for fiber yield-related traits by constructing the first genetic linkage map in ramie (Boehmeria nivea L. gaud). Mol Breed 34:883–892

Liu C, Zhu S, Tang S, Wang H, Zheng X, Chen X, Dai Q, Liu T (2017) QTL analysis of four main stem bark traits using a GBS-SNP-based high-density genetic map in ramie. Sci Rep 7:13458

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408

Lu K, Wang T, Xu X (2016) Genome-wide association analysis of height of podding and thickness of pod canopy in Brassica napus. Acta Agron Sin 42:344–352

Luan M, Qin Z, Chen J, Wang X, Xu Y, Sun Z, Mart H, Abecasis G, Durbin R (2009a) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009a) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

Li L, Yu X, Thompson A, Guo M, Yoshida S, Asami T, Chory J, Yin Y (2009b) Arabidopsis MYB30 is a direct target of BES1 and cooperates with BES1 to regulate brassinosteroid-induced gene expression. Plant J 58:275–286

Li B, Tian L, Zhang J, Huang L, Han F, Yan S, Wang L, Zheng H, Sun J (2014) Construction of a high-density genetic map based on large-scale markers developed by specific length amplified fragment sequencing (SLAF-seq) and its application to QTL analysis for isoflavone content in Glycine max. BMC Genom 15:1086

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28(18):2397–2399

Liu T, Zhu S, Fu L, Tang Q, Yu Y, Chen P, Lanu M, Wang C, Tang S (2013) Development and characterization of 1,827 expressed sequence tag-derived simple sequence repeat markers for ramie (Boehmeria nivea L. gaud). PLoS ONE 8:e603464

Lieu C, Zheng X, Chen C, Dai Q, Liu T (2017) QTL analysis of four main stem bark traits using a GBS-SNP-based high-density genetic map in ramie. Sci Rep 7:13458

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408

Lu K, Wang T, Xu X (2016) Genome-wide association analysis of height of podding and thickness of pod canopy in Brassica napus. Acta Agron Sin 42:344–352

Luan M, Qin Z, Chen J, Wang X, Xu Y, Sun Z (2009) Correlation between fiber fineness and expression of bast fiber development related gene FB27 in ramie. Plant Fiber Sci China 31(6):339–343

Luan MB, Zou ZZ, Zhu JJ, Wang XF, Xu Y, Ma QH, Sun ZM, Chen JH (2014) Development of a core collection for ramie by heuristic search based on SSR markers. Biotechnol Biotechnol Equip 28(5):798–804

Luan MB, Zou ZZ, Zhu J, Wang XF, Xu Y, Sun Z, Chen JH (2015) Genetic diversity assessment using simple sequence repeats (SSR) and sequence-related amplified polymorphism (SNAP) markers in ramie. Biotechnol Biotechnol Equip 29:624–630

Luan MB, Liu CC, Wang XF, Xu Y, Sun Z, Chen JH (2017) SSR markers associated with fiber yield traits in ramie (Boehmeria nivea L. Gaudich). Ind Crop Prod 107:439–445

Luan MB, Jian JB, Chen P, Chen JH, Chen JH, Gao Q et al (2018) Draft genome sequence of ramie, Boehmeria nivea (L.) Gaudich. Mol Ecol Resour 18:639–645
Sun X, Liu D, Zhang X, Li W, Liu H, Hong W et al (2013) SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PLoS ONE 8:e58700

Wang H, Qin F (2017) Genome-wide association study reveals natural variations contributing to drought resistance in crops. Front Plant Sci 8:1110

Wang HM, Wang X (2004) Evaluation of the fineness of degummed bast fibers. Fiber Polym 5:171–176

Wei Q, Wang Y, Qin X, Zhang Y, Zhang Z, Wang J, Li J, Lou Q, Chen J (2014) An SNP-based saturated genetic map and QTL analysis of fruit-related traits in cucumber using specific-length amplified fragment (SLAF) sequencing. BMC Genom 15:1158

Xiong HP, Jiang J, Yu C, Guo Y (1998) Relation between yield-related traits and yield in ramie. Acta Agron Sin 24:155–160

Yu X, Li L, Zola J, Aluru M, Ye H, Foudree A et al (2011) A brassinosteroid transcriptional network revealed by genome-wide identification of BES1 target genes in Arabidopsis thaliana. Plant J 65:634–646

Zhang Y, Wang L, Xin H, Li D, Ma C, Ding X, Hong W, Zhang X (2013) Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. BMC Plant Biol 13:141

Zhang Z, Shang H, Shi Y, Huang L, Li J, Ge Q et al (2016a) Construction of a high-density genetic map by specific locus amplified fragment sequencing (SLAF-seq) and its application to Quantitative Trait Loci (QTL) analysis for boll weight in upland cotton (Gossypium hirsutum). BMC Plant Biol 16:79

Zhang Z, Wei T, Zhong Y, Li X, Huang J (2016b) Construction of a high-density genetic map of Ziziphus jujube Mill. using genotyping by sequencing technology. Tree Genet Genomes 12:76

Zhang S, Chen X, Lu C, Ye J, Zou M, Lu K et al (2018) Genome-wide association studies of 11 agronomic traits in Cassava (Manihot esculenta Crantz). Front Plant Sci 9:503

Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W et al (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. Nat Biotechnol 33:125–408

Zhou Q, Zhou C, Zheng W, Mason AS, Fan S, Wu C, Fu D, Huang Y (2017) Genome-Wide SNP markers based on SLAF-Seq uncover breeding traces in rapeseed (Brassica napus L.). Front Plant Sci 8:648

Zhu SY, Liu TM, Dai Q, Wu D, Zheng X, Tang S, Chen J (2017) Genetic structure and relationships of an associated population in ramie (Boehmeria nivea L. Gaud) evaluated by SSR markers. Biotechnol Biotechnol Equip 31:36–44

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.