Theoretical and Applied Genetics (2019) 132:2663–2676
https://doi.org/10.1007/s00122-019-03380-7

ORIGINAL ARTICLE

QTL analysis and candidate gene identification for plant height in cotton based on an interspecific backcross inbred line population of *Gossypium hirsutum* × *Gossypium barbadense*

Jianjiang Ma[1,3] · Wenfeng Pei[4] · Qifeng Ma[1] · Yanhui Geng[1] · Guoyuan Liu[1] · Ji Liu[1] · Yupeng Cui[1] · Xia Zhang[1] · Man Wu[1] · Xingli Li[1] · Dan Li[1] · XinShan Zang[1] · Jikun Song[1] · Shurong Tang[1] · Jinfa Zhang[2] · Shuxun Yu[1,3] · Jiwen Yu[1]

Received: 22 January 2019 / Accepted: 14 June 2019 / Published online: 24 June 2019

Abstract

Key message **We constructed the first high-quality and high-density genetic linkage map for an interspecific BIL population in cotton by specific-locus amplified fragment sequencing for QTL mapping. A novel gene *GhPIN3* for plant height was identified in cotton.**

Abstract Ideal plant height (PH) is important for improving lint yield and mechanized harvesting in cotton. Most published genetic studies on cotton have focused on fibre yield and quality traits rather than PH. To facilitate the understanding of the genetic basis in PH, an interspecific backcross inbred line (BIL) population of 250 lines derived from upland cotton (*Gossypium hirsutum* L.) CRI36 and Egyptian cotton (*G. barbadense* L.) Hai7124 was used to construct a high-density genetic linkage map for quantitative trait locus (QTL) mapping. The high-density genetic map harboured 7,709 genotyping-by-sequencing (GBS)-based single nucleotide polymorphism (SNP) markers that covered 3,433.24 cM with a mean marker interval of 0.67 cM. In total, ten PH QTLs were identified and each explained 4.27–14.92% of the phenotypic variation, four of which were stable as they were mapped in at least two tests or based on best linear unbiased prediction in seven field tests. Based on functional annotation of orthologues in Arabidopsis and transcriptome data for the genes within the stable QTL regions, *GhPIN3* encoding for the hormone auxin efflux carrier protein was identified as a candidate gene located in the stable QTL qPH-Dt1-1 region. A qRT-PCR analysis showed that the expression level of *GhPIN3* in apical tissues was significantly higher in four short-statured cotton genotypes than that in four tall-statured cotton genotypes. Virus-induced gene silencing cotton has significantly increased PH when the expression of the *GhPIN3* gene was suppressed.

Introduction

Cotton is a worldwide cash crop that supplies indispensable natural fibre for the textile industry. Upland cotton (*Gossypium hirsutum* L.) has high lint yield, accounting for more than 95% of world cotton production (Li et al. 2016b). Compared with upland cotton, Extra-long staple, Pima, Egyptian or Sea-Island cotton (*G. barbadense* L.) has high fibre quality, but its low yield limits the areas where it is grown (Said et al. 2015b). Plant height (PH) is one of the main factors affecting light interception and photosynthesis in the canopy, playing a key role in overall photosynthate partitioning and harvesting index in several crops (Teng et al. 2013). In the process of domestication of crops, plant architecture modification is increasingly the focus of breeding efforts, i.e. 'ideal-type breeding', which has been widely employed to enhance crop resistance to lodging in various environments and to improve photosynthesis in several crop species

Communicated by Diane E. Mather.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03380-7) contains supplementary material, which is available to authorized users.

Jianjiang Ma and Wenfeng Pei have contributed equally to this work.

✉ Jinfa Zhang
jinzhang@nmsu.edu

✉ Shuxun Yu
ysx195311@163.com

✉ Jiwen Yu
yujw666@hotmail.com

Extended author information available on the last page of the article

(Chen et al. 2018; Liu et al. 2018). For example, via hybrid breeding, the adoption of compact and moderately short cultivars that are better suited to a high planting density significantly improves maize productivity (Du et al. 2018b; Peiffer et al. 2014). Similarly, decreased height in rice and wheat strengthens plant resistance to lodging due to heavy rain and high wind damage, which contributes to a high harvesting index and high yield in 'green revolution' (Peng et al. 1999; Sasaki et al. 2002). However, in cotton, short plants do not always equate to high yielding in that some environments require taller or shorter plants depending on environment and cultural practices. For example, in the US, the Texas High Plains requires a short plant due to a windy environment and the harvest method of using a cotton stripper; however, in the Southeastern USA, where the climate is humid with a great deal of rain, the cotton crop is harvested with mechanical pickers requiring tall plants. Currently, increased demand in harvest mechanization in China makes it necessary for cotton researchers to breed cotton plants with a suitable height (Su et al. 2018). Hence, understanding the genetic basis of PH will be useful for effectively improving plant architecture adapted to mechanical harvest.

The restricted recombination events and low genetic diversity in designed populations add to the difficulty of breeding by hybridization. To improve accuracy and shorten breeding cycles, marker-assisted selection (MAS) is increasingly applied in cotton crossbreeding (Cao et al. 2014; Huang et al. 2018; Jia et al. 2016; Wen et al. 2018). For MAS, a genetic map of high quality with a high marker density is very important as it lays the foundation for future studies, such as mapping reliable quantitative trait loci (QTLs) of important agronomic traits and further identifying candidate genes within these QTL regions. Following the construction of the first genetic map of allotetraploid cotton, numerous studies have focused on genetic map construction and QTL mapping (e.g., Shang et al. 2016; Sun et al. 2017; Yu et al. 2007, 2013; Zhang et al. 2015c; Zhu et al. 2018). To date, many genetic maps have been constructed using simple sequence repeat (SSR) markers. However, the low polymorphic level of SSR markers is a bottleneck for constructing a fine SSR-based genetic map on the genome scale. With the rapid development of DNA sequencing technology, single nucleotide polymorphism (SNP) markers with high genetic diversity at the genome-wide level have been widely used to construct genetic linkage maps. Accordingly, the use of SNP markers is an effective way to construct a high-quality and high-density genetic linkage map (HDGM) for detecting QTLs. Specific-locus amplified fragment sequencing (SLAF-seq) represents one of the genotyping-by-sequencing (GBS) methods and can be used to develop large quantities of genome-wide SNP markers, as detailed by Sun et al. (2013). To date, the HDGMs of various species have been successfully constructed using the SLAF-seq technique (e.g., Ji et al. 2017; Keerio et al. 2018; Tao et al. 2017; Zhang et al. 2015a, 2017).

PH is a quantitative trait, and it is affected by genotype, environment, genotype by environment interaction, and development during plant growth. Currently, although PH QTLs have been mapped onto chromosomes in cotton, such as At5 (4 QTLs), At9 (4 QTLs), At11 (6 QTLs), Dt1 (5 QTLs), Dt3 (8 QTLs), Dt5 (4 QTLs) and Dt12 (5 QTLs) (Jia et al. 2016; Said et al. 2015a; Su et al. 2018; Sun et al. 2017), few candidate genes associated with PH have been mined within these QTL regions. As is well known, plant hormones such as auxin (IAA) and gibberellic acid (GA) play a key role in plant growth and developmental processes (Friml et al. 2003; Kimura and Kagawa 2006; Palme et al. 2006; Sasaki et al. 2002). The GA biosynthesis and GA signalling have been identified to influence PH. For example, the recessive allele of the *sd1* gene, encoding a mutant oxidase enzyme involved in the biosynthesis of gibberellin, leads to the short stature in rice (Monna et al. 2002; Sakamoto et al. 2004). IAA is mainly synthesized in young leaves and apical organs and promotes root elongation, stem growth and flower differentiation. At the cellular level, the uneven distribution of three IAA membrane carriers, including a number of p-glycoprotein ATP-binding cassette transporters, IAA efflux transporter (PIN) family members, and the influx carrier AUX1 protein family, leads to the directional transport of IAA (Wang et al. 2018a, b). The enzyme of PINs reportedly plays a rate-limiting role in catalysing the transport of IAA and participates in PH developmental process (Petrásek et al. 2006). In recent studies, some members of the PIN family were found to influence PH in some species, such as ZmPIN1a overexpression dwarfing PH in maize (Du et al. 2018b) and overexpression of OsPIN2 leading to a shorter PH in rice (Chen et al. 2012). However, there is no study on whether any plant hormone biosynthetic genes are associated with PH in cotton.

To provide additional information for breeding programmes, the genetic basis of PH must be further investigated in cotton. In this study, we used an interspecific backcross inbred line (BIL) population containing 250 individual lines was derived from a cross between upland cotton (CRI36) and Egyptian cotton (Hai7124). The objectives of this study were to construct a fine interspecific genetic linkage map in cotton based on 7,709 high-quality SNP markers, and to detect QTLs for PH in seven environments and across the tests using the best linear unbiased prediction (BLUP) method. As a result, a stable QTL for PH, qPH-Dt1-1, was mapped in multiple environments, which was also reported in previous studies (Jia et al. 2016; Said et al. 2015a). Through gene annotation and differential gene expression analysis in different tissues and genotypes differing in PH, one candidate gene in the qPH-Dt1-1 region encoding an IAA efflux carrier family protein, namely *GhPIN3*, was

identified for functional verification using virus-induced gene silencing (VIGS) in the two parental lines (CRI36 and Hai7124) used to develop the BIL population. *GhPIN3* represents the first candidate gene for PH identified through QTL mapping in cotton.

Methods and materials

Plant materials

An interspecific BIL population of 250 lines was utilized in this study. The BILs were produced by a cross between *G. barbadense* Hai7124 and *G. hirsutum* CRI36, using CRI36 as the recurrent parent for backcrossing with the F_1 to produce BC_1F_1, followed by seven generations of selfing. The seeds of Hai7124 were obtained from the National Cotton Germplasm Collections of the Low-temperature Germplasm Gene Bank, Institute of Cotton Research, Chinese Academy of Agricultural Sciences (CRI-CAAS); and the detailed information of Hai7124 was described by Zhao et al. (2013). The CRI36 was bred by CRI-CAAS. The two parents and 250 BC_1F_7 lines were planted in seven environments in four locations: the experimental farm, CRI-CAAS, Anyang (Henan Province, 36.06°N, 114.49°E) with four tests including one each in 2015 and 2016 and two in 2017 (one in south farm and another in east farm); Sanya (Hainan Province, 18.41°N, 109.20°E) in 2016; Alaer (Xinjiang Uyghur Autonomous Region, 40.55°N, 81.28°E) in 2016; and Weixian (Hebei Province, 36.95°N, 115.46°E) in 2017. The cotton seeds were hill-sown by hand with two replications in a randomized complete block design in each environment and covered with plastic mulch applied directly by a machine during the growing season. In Anyang, Weixian and Sanya, approximately 16 plants per 4-m-long row were retained, and the row spacing was 0.80 m. In Alaer, where a high-density seeding rate was used, approximately 44 plants per 5-m-long row were retained, and the row spacing was 0.38 m. Crop management practices followed local recommendations for the production area. The use of the two cotton production systems (i.e. normal and high plant density) allowed detection of consistent QTLs for PH between the two systems.

Phenotypic measurement and analysis

In all the environments, PH was measured as the distance in cm from the soil line of the plant to the top of the apical bud at reproductive maturity. Six normally growing plants in each replication were randomly selected for height testing. The R software package 'lme4' was used to calculate the BLUPs for PH across the seven environments (Liu et al. 2016a). A statistical analysis including the analysis of variance (ANOVA) of PH for the BIL population across environments was performed using SPSS 23.0 (Li et al. 2016a). QTL IciMapping 4.1.0.0 software was used to estimate broad-sense heritability (H^2) of PH across environments (Meng et al. 2015).

DNA and SLAF library preparation

The young leaves of the two parents and 250-BIL population were sampled in July 2015 and stored in a refrigerator at − 70 °C. Using a miniprep method, the genomic DNA of the two parents and each BIL was extracted. Several modifications of the SLAF-seq strategy were adopted in the library construction. First, the reference genomic data of TM-1 were used to perform a pilot experiment, and the quantities of markers produced by different endonucleases were assessed with an in-silico simulation software (Zhang et al. 2015b). According to the result of the SLAF pilot experiment, the SLAF library was constructed, and the *Hae*III endonuclease was selected to digest the genomic DNA of the two parents and 250-BIL population. Detailed information on the SLAF-seq strategy was provided in a previous study (Zhang et al. 2015a).

Genotyping of SLAF markers

High-quality SNP markers were detected and genotyped through detailed processes as described in previous studies (Liu et al. 2016b; Sun et al. 2013). First, low-quality reads (quality score < 30e) were excluded, and the remaining reads were assigned to each offspring through duplex barcode sequences. Then, five nucleotide bases at the terminal position of each remaining read were trimmed off. Finally, 100-bp paired-end clean reads, acquired from each sample, were aligned to the reference genome of TM-1 (Zhang et al. 2015b) using Burrows–Wheeler Alignment (BWA) software (Li and Durbin 2009). Sequences from each offspring located to a common physical position with identities greater than 95% were regarded as one SNP locus (Zhang et al. 2015a). Using the software GATK, SNP markers in all of the SLAF loci between the two parents were detected. As the length of each SLAF was approximately 200 bp, three or more SNPs in one SLAF showing high heterozygosity would lead to a low accuracy of genotyping in cotton. Accordingly, SLAFs with high heterozygosity were discarded. The detailed procedures of SLAF genotyping were described by Zhang et al. (2016). The criteria used to define the SLAF repetitiveness and polymorphism were reported in previous studies (Zhang et al. 2015a, 2016). SLAF repeats were excluded, and polymorphic SNPs were consequently used. Ultimately, only common SLAFs in both the BIL population and the parents were used.

Linkage map construction and quality assessment

The TM-1 reference genome database was used to construct the linkage map of the BIL population, as described in detail by Zhang et al. (2015a). Correcting genotyping errors and ordering the SLAF sequences along the chromosomes were conducted by HighMap software (Van Ooijen 2011; Liu et al. 2014). In addition, SMOOTH was used to correct errors based on the contribution of parental genotypes (Van et al. 2005), and incorrect genotypes were filtered with a K-nearest neighbour algorithm (Van et al. 2005). Skewed markers as defined below were supplemented into the linkage map by a multipoint maximum likelihood (MML) method. The map distances were calculated by the Kosambi mapping function. The Chi-square test was used to identify the markers with segregation distortion. As described in a previous study, the segregation distortion markers (SDMs) ($0.001 \leq p \leq 0.05$) were acceptable in the HDGM (Zhang et al. 2016). The SNP markers in the linkage map were aligned to the genome with the local BLAST method. The collinear coefficient of markers in each chromosome was analysed by CIRCOS 0.66 software. A recombination hotspot (RH) was defined as a genetic distance between two adjoining markers greater than 20 cM per Mb (Zhang et al. 2015a).

QTL analysis

PH from the seven individual environments and its BLUP across the tests were used for QTL analysis, and QTLs were identified with the inclusive composite interval mapping (ICIM) method in IciMapping 4.1.0.0 software (Meng et al. 2015). A stringent logarithm of odds (LOD) threshold was calculated by a permutation test, and the parameters of the QTL mapping were set as follows. The time of calculation was 1,000; the *p* value of the type I error was 0.05; the PIN was 0.001; and the mapping step was 1.0 cM. The same QTLs in two or more environments with LOD threshold of > 2.5 were also considered as significant QTLs (Shang et al. 2015a). A QTL confidence interval (95%) was set as a mapping distance interval corresponding to 1 LOD decline on either side of the peak (Liu et al. 2019; Yu et al. 2012). QTLs for PH detected in two or more environments were considered as 'stable' when their confidence intervals overlapped (Shi et al. 2016; Yu et al. 2012; Zhang et al. 2015c). QTLs were named according to Sun et al. (2013). The positive additive effect of a QTL suggests that the allele derives from the male parent Hai7124; otherwise, the allele derives from the female parent CRI36.

Candidate gene identification

To obtain potential candidate genes, gene sequences within stable QTLs were extracted. The TAIR database was employed for the annotation of gene functions (https://www.arabidopsis.org/index.jsp). To reveal the general pattern of expression of the candidate genes, the transcriptome sequencing data for different tissues (ovule, fibre, root and stem) of TM-1 were employed as a reference (Zhang et al. 2015b). The apical bud, young stems, roots and leaves of the short-statured recurrent parent CRI36 and the tall-statured parent Hai7124 at the three-true leaf stage and apical organ mixture samples (young stems, apical buds and young leaves) of eight BILs (including four short-statured and four tall-statured lines) at the flowering stage were sampled to determine the expression level of candidate genes by qRT-PCR.

Functional analysis of *GhPIN3* in cotton

For the virus-induced gene silencing (VIG) assay, primers VIGsGhPIN3-SpeI-F (5′- ATGCCTGCAGACTAGTGT CGCCATATTTGCCGTTCC -3′) and VIGsGhPIN3-AscI-R (5′- TAGACCTAGGGGCGCGCCCCGATTTCAGCGTCG GTTTC -3′) were used for VIG vector construction. A 455-bp specific fragment from *GhPIN3* was cloned and inserted into the pCLCrVA vector. The resulting pCLCrVA:*GhPIN3* was co-infiltrated with pCLCrVB via *Agrobacterium tumefaciens* into cotyledons of cotton seedlings as described in a previous study (Su et al. 2018). Plants co-inoculated with pLCrVB and empty pLCrVA were used as the negative control (CK). A well-managed phytotron (16 h light, 8 h dark and 25 °C) was used to grow the cotton plants (Gu et al. 2014). To detect the silencing power on *GhPIN3*, qRT-PCR was further performed to examine the expression level of *GhPIN3* in young leaves of each plant (pCLCrVA:*GhPIN3* and CK) when virus had infected the cotton plants for 30 d. The PH of CK and pCLCrVA:*GhPIN3* plants was measured at 30, 40, 50 and 60 d.

Results

PH variation of parents and the BIL population

The PH values of the two parents and the BIL population were measured in seven environments. Among the seven environments, PH in the BILs ranged from 31.67 to 125.00 cm, with an average of 74.14 cm; and the mean PH of the parents CIR36 and Hai7124 was 72.23 and 85.74 cm, respectively (Table 1). The PH of Hai7124 was significantly ($p < 0.01$) greater than that of CIR36, and the value of skewness in each environment showed that PH followed a normal

Table 1 Plant height (PH) of backcross inbred lines (BILs) of Hai7124 × CRI36 hybrids and their parents

Trait	Environment	Parent		Diff.	BILs			SD	Skewness	Kurtosis	CV (%)
		Hai7124	CRI36		Min	Max	Mean				
PH	Ay-nc (2015)	90.60	79.50	**	33.00	125.00	66.56	14.23	0.63	1.40	21.38
	Ay-nc (2016)	84.00	71.00		34.42	94.20	67.21	11.74	−0.19	−0.25	17.47
	Hn-sy (2016)	94.50	85.20		37.20	113.00	83.29	12.68	−0.12	−0.03	15.22
	Xj-al (2016)	76.60	67.40		34.80	103.50	66.26	14.19	0.21	−0.58	21.42
	Ay-dc (2017)	79.00	67.00		46.33	118.83	83.07	13.74	−0.06	−0.33	16.54
	Ay-nc (2017)	82.00	60.50		42.00	100.40	75.47	9.83	−0.35	0.11	13.03
	Hb-wx (2017)	93.50	75.00		31.67	110.00	77.09	12.37	−0.17	0.57	16.05

**Indicates significant difference (Diff.) at $p = 0.01$, SD and CV represent the standard deviation and coefficient of variation, respectively

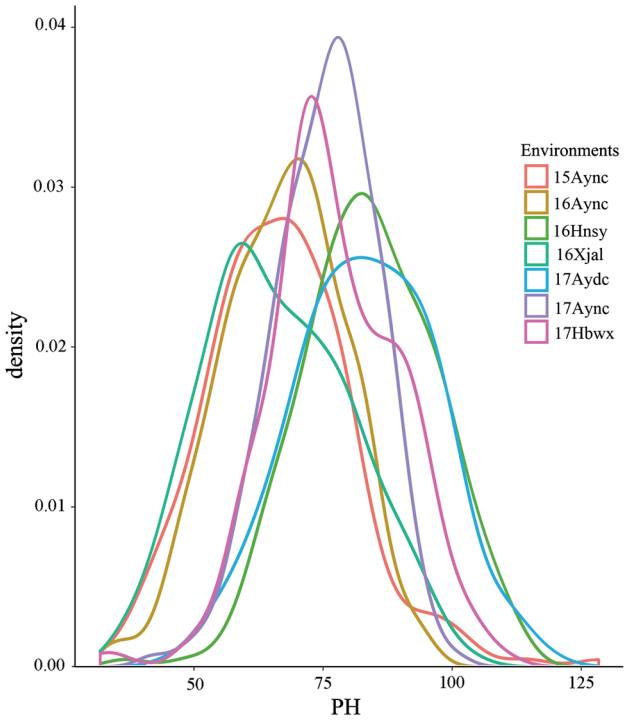


Fig. 1 Frequency distribution of plant height of 250 BILs in different environments. 15Aync, 16Aync and 17Aync represent the environment of Anyang south farm in 2015, 2016 and 2017, respectively; 16Hnsy and 16Xjal represent the environment of Sanya and Alaer in 2016; 17Aydc and 17Hbwx represent the environment of Anyang east farm and Weixian in 2017

distribution in the BIL population (Fig. 1, Table 1). Furthermore, there was a transgressive segregation of PH in the BIL population when compared with the parents Hai7124 and CRI36. The result of variance analysis suggested significant variations in PH ($p < 0.01$) due to environment, genotype and genotype × environment (Supplemental Table S1). However, the broad-sense heritability estimate (H^2) for PH was 88.01% (i.e. the percentage of the total phenotypic variance accounted by the genotypic variance), suggesting that PH was highly heritable (Supplemental Table S1).

Statistics of the SLAF-seq data

Via SLAF sequencing, we acquired approximately 152 GB of initial DNA sequence data containing 1,136.19 M paired-end reads, and the size of each read ranged from 264 to 364 bp. The SLAF numbers of CRI36 and Hai7124 were 468,850 and 500,626 (Supplemental Figure S1a), and their average sequencing depths were 25.03 × and 21.20 × with 5.46% and 5.81% of genome coverage ratio, respectively (Supplemental Figure S1b). In the BIL population, the number of SLAFs ranged from 275,045 to 527,319, with an average of 352,920 (Supplemental Figure S1a). The average sequencing depth was 11.12 ×, with 4.14% of the average genome coverage ratio (Supplemental Figure S1b). The average Q30 ratio was 91.23%, indicating that the sequences were of high quality, and the guanine-cytosine (GC) content was 39.70%.

Information of the SNP markers

From the SLAF-seq data, we detected 255,795 SNP markers between the two parents CRI36 and Hai7124, and all of the markers could be grouped into four genotypes: nn × np, lm × ll, hk × hk and aa × bb (Supplemental Table S2). Ultimately, the aa × bb type of markers with an average sequence depth ≥ 20 × in the parents and ≥ 5 × in the offspring were used for the further analysis. After the aa × bb markers with 20% missing data in the BILs were filtered, 10,694 markers were retained. Following a further filtering of the SDMs markers, 7,709 markers were retained and used to construct the genetic map (Fig. 2). The 7,709 SNP markers could be divided into six mutation genotypes: 2,728 and 2,730 SNP markers were G/A and T/C mutation genotypes, respectively; and only 351 SNP markers were G/C mutation genotypes (Supplemental Table S3). These results showed that in cotton, the SNPs from G to A or T to C replacements were more common, while that from G to C was rare.

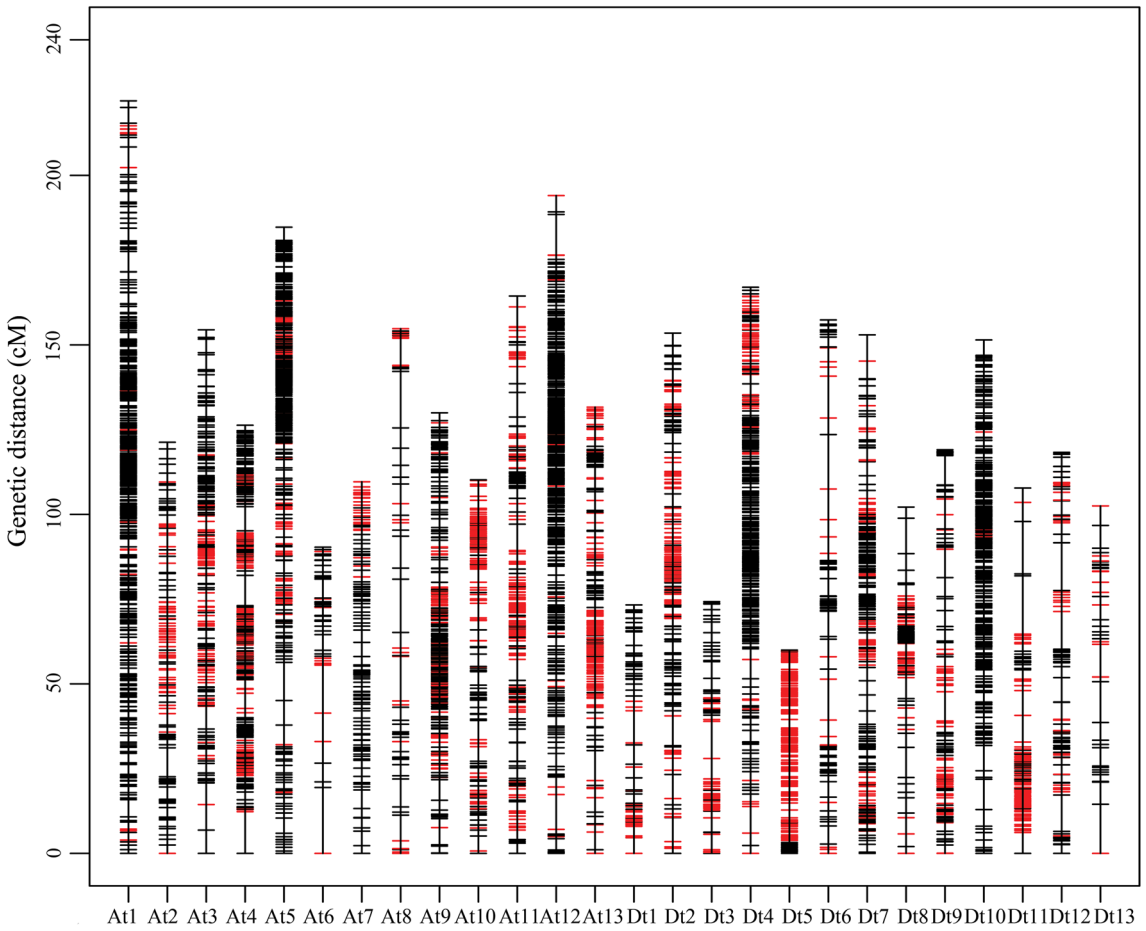


Fig. 2 The distribution of SNP markers in each of the 26 linkage groups/chromosomes. Acceptable segregation distortion markers ($0.001 < p < 0.05$) on the map are highlighted in red

Genetic map construction in the BIL population

A total of 7,709 high-quality SNP markers were mapped to 26 chromosome-corresponding linkage groups which spanned a total genetic distance of 3,433.24-cM. The 13 chromosomes in the A subgenome contained 4,641 markers with a 1,893.68-cM genetic distance, and the other 13 chromosomes in the D subgenome included 3,068 markers with a 1,539.56-cM genetic distance. The genetic distance and marker number varied among chromosomes. The longest linkage group, At1, spanned a 221.99-cM genetic length with 622 markers. In contrast, the shortest linkage group, Dt5, spanned only a 59.97-cM genetic length with 489 markers. Moreover, the linkage group At5 harboured the most markers (658), while At6 had the fewest SNP markers (52). Among the 26 linkage groups, the maximum gap between two adjacent markers varied from 2.52 cM on Dt5 to 19.43 cM on At6 (Table 2).

The quality assessment of the genetic map

Collinearity results showed that the SNP markers on 26 linkage groups had high levels of coverage across the cotton genome (Supplemental Figure S2). Moreover, the Spearman correlation coefficient for each of the 26 linkage groups ranged from 0.72 to 1.00, and the average coefficient was 0.92 (Supplemental Table S6). This result suggested that the order of most of the markers in the linkage groups has a high consistency with those in the TM-1 reference genome. Several chromosomes, namely, At4, At5, At6, At9 and Dt2, with some deviation, were identified by the collinearity analysis (Supplemental Table S6). A total of 22 linkage groups were identified with RHs (Table 2). Dt5 had the largest number (80) of RHs, while Dt8 harboured only two RHs (Supplemental Table S7). Chromosomes At8, At9, Dt12 and Dt13 did not harbour any RHs (Table 2).

Table 2 Detailed information of the SNP markers in the genetic map

LG	Number of SNPs	Length (cM)	Average distance (cM)	Gaps < 5 cM (%)	Max gap	RH	SDM	SDM %
At1	622	221.99	0.36	99.52	6.08	16	23	3.70
At2	138	121.28	0.88	99.27	8.54	3	58	42.03
At3	345	154.45	0.45	98.84	7.52	27	110	31.88
At4	542	126.3	0.23	99.45	12.29	4	263	48.52
At5	658	184.72	0.28	99.54	11.24	9	86	13.07
At6	52	90.32	1.74	88.24	19.43	3	12	23.08
At7	167	109.61	0.66	99.40	5.47	25	34	20.36
At8	79	154.76	1.96	87.18	16.66	0	26	32.91
At9	500	129.96	0.26	99.60	6.06	0	134	26.80
At10	244	110.19	0.45	99.59	5.68	7	161	65.98
At11	361	164.41	0.46	98.33	7.77	3	221	61.22
At12	611	194.05	0.32	99.67	11.97	33	17	2.78
At13	322	131.64	0.41	99.38	8.86	41	230	71.43
Dt1	118	73.3	0.62	98.29	9.49	10	53	44.92
Dt2	360	153.48	0.43	99.16	8.78	10	223	61.94
Dt3	116	74.25	0.64	98.26	10.96	10	57	49.14
Dt4	343	166.99	0.49	99.42	7.85	17	63	18.37
Dt5	489	59.97	0.12	100.0	2.52	80	462	94.48
Dt6	109	157.36	1.44	91.67	16.04	7	30	27.52
Dt7	357	152.96	0.43	99.44	7.74	3	100	28.01
Dt8	141	102.16	0.72	97.86	10.46	2	57	40.43
Dt9	175	119.07	0.68	97.70	8.68	14	81	46.29
Dt10	435	151.46	0.35	99.54	8.96	7	4	0.92
Dt11	206	107.78	0.52	97.07	17.20	0	158	76.70
Dt12	162	118.29	0.73	96.89	14.07	0	39	24.07
Dt13	57	102.49	1.80	85.71	14.48	5	23	40.35
Total	7,709	3433.24	0.67	97.27	10.18	336	2725	35.00

LG, RH and SDM represent the linkage group, recombination hotspot and segregation distortion marker, respectively

QTLs for PH in the BILs and previous studies

In total, ten QTLs for PH were detected on nine chromosomes across seven environments and with BLUP. Among the ten QTLs, four and six QTLs were identified on the A subgenome and D subgenome, respectively. Five QTLs (qPH-At3-2, qPH-At5-1, qPH-At10-1, qPH-Dt11-1 and qPH-Dt13-1) had positive additive effects from Hai7124, while another five QTLs (qPH-At3-1, qPH-Dt1-1, qPH-Dt3-1, qPH-Dt9-1 and qPH-Dt12-1) had negative additive effects from Hai7124 (Table 3). Importantly, four QTLs, qPH-At3-2, qPH-At5-1, qPH-Dt1-1 and qPH-Dt3-1, were stable as they were identified in at least two tests or one test and based on BLUP of the seven tests (Fig. 3, Table 3); another six QTLs were detected in one environment or with BLUP only and explained 6.84–14.92% of the PV (Supplemental Figure S3, Table 3). We further compared these ten QTLs with PH-associated QTLs reported in previous studies. As a result, two stable QTLs were reported in previous QTL linkage or association studies (Table 3), such as SNP markers (Marker22291 and Marker22588) were found to be associated with PH (Jia et al. 2016), and their physical positions ranged from 26.90 to 32.30 Mb on the Dt1 chromosome, which was contained in the interval of the stable QTL qPH-Dt1-1 (Supplemental Table S5, Table 3). The SNP markers (Marker25959 and Marker26500) and (D03_31584163 and D03_32132408) were reported to be associated with PH in a 137-RIL population and 355 upland cotton accessions, respectively (Jia et al. 2016; Su et al. 2018), and were anchored to 31.58 to 38.69 Mb on the Dt3 chromosome and overlapped with another stable QTL qPH-Dt3-1 (Supplemental Table S5, Table 3). The stable QTLs qPH-Dt1-1 and qPH-Dt3-1 were detected in both normal and high planting density regimes (Aync and Xjal) (Table 3).

Prediction of candidate genes in stable QTLs

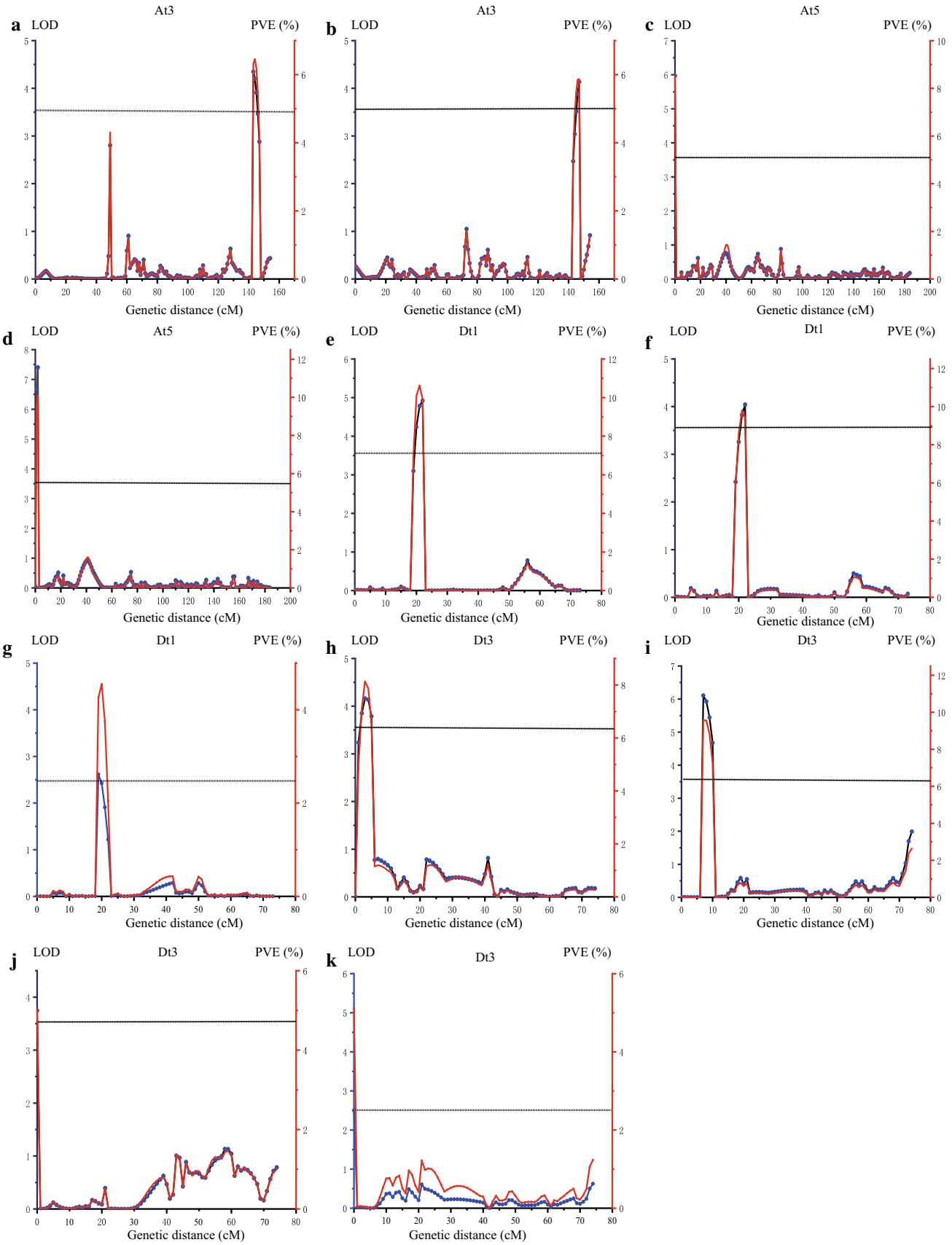
To identify candidate genes associated with PH, we adopted a method combining the genes within the stable QTL regions with functional annotation of orthologues in Arabidopsis and transcriptome data. In total, 1,220 genes were included in the regions of the four stable QTLs. Of these candidate genes, 956 had annotation information (Supplemental Table S8). As PH is mainly determined by stem growth, we focused on the genes that are preferentially expressed in young stems with significantly higher expression levels than in other important tissues (ovule, fibre and root) in cotton. This resulted in the selection of 19 genes (Fig. 4a). Among these 19 genes, *Gh_D01G1471* was located in qPH-Dt1-1(Supplemental Table S8). Amino acid residues of Gh_D01G1471 had the highest (78%) identity with AtPIN3 encoding the IAA efflux carrier protein in Arabidopsis. Hence, we named *Gh_D01G1471 GhPIN3*. To verify that *GhPIN3* was associated with PH, the qRT-PCR analysis revealed that *GhPIN3* exhibited a significantly higher expression level in apical buds, young stems and leaves of CRI36 than in those of Hai7124, and the expression level was not significantly different in roots of CRI36 and Hai7124 (Fig. 4b). Moreover, we randomly selected eight lines (four short and four tall) from the 250

Fig. 3 The information of the stable QTLs mapping. **a** and **b** show the stable QTL in At3 detected in 16Hnsy and 17Aync, respectively; **c** and **d** show the stable QTL in At5 detected in 17Aync and with BLUP, respectively; **e, f** and **g** show the stable QTL in Dt1 detected in 15Aync, 17Aydc and 16Xjal, respectively; **h**, **i**, **j** and **k** show the stable QTL in Dt3 detected in 16Xjal, 16Aync, BLUP and 17Hbwx, respectively; the black dotted line represents LOD = 3.50 or 2.50; the red line represents the percentage of phenotypic variance explained; and the blue line represents the LOD value

Table 3 Summary of plant height (PH) QTLs identified in different environments and using the best linear unbiased prediction (BLUP)

QTL name	Environment	Left marker	Right marker	LOD	Add	PVE (%)	95% Confidence interval (cM)	Physical interval/bp	Reported previously
qPH-At3-1	BLUP	Marker22727	Marker22716	5.06	−2.52	6.84	59.50–60.50	15,005,700–15,100,195	
qPH-At3-2	16Hn	Marker34590	Marker34600	4.35	3.82	7.66	142.50–146.50	98,060,423–98,605,789	
	17Aync	Marker34590	Marker34600	4.14	2.99	6.84	144.50–147.50	98,341,286–98,743,494	
qPH-At5-1	17Aync	Marker41997	Marker42006	5.97	3.62	10.01	0.00–0.50	319,006–595,105	
	BLUP	Marker42013	Marker42017	6.69	2.94	9.47	0.00–2.50	319,006–1,632,832	
qPH-At10-1	16Hn	Marker94549	Marker94428	5.29	4.52	9.26	102.50–104.50	20,994,947–21,939,789	
qPH-Dt1-1	15Aync	Marker137433	Marker140073	4.92	−5.37	10.30	19.50–22.50	30,733,785–53,028,196	(Jia et al. 2016)
	16Xjal	Marker137433	Marker140073	2.61	−3.56	4.27	18.50–21.50	23,277,460–45,571,871	
	17Aydc	Marker137433	Marker140073	4.04	−4.70	9.06	19.50–22.50	30,733,785–53,028,196	
qPH-Dt3-1	BLUP	Marker151158	Marker151090	3.75	−1.88	4.91	0.00–0.50	37,185,781–38,706,260	(Jia et al. 2016; Su et al. 2018)
	17Hbwx	Marker151158	Marker151090	2.51	−3.06	5.11	0.00–0.50	37,185,781–38,706,260	
	16Xjal	Marker150850	Marker149780	4.16	−4.67	10.50	0.00–5.50	25,858,483–38,706,260	
	16Aync	Marker149953	Marker150605	6.11	−4.92	8.73	6.50–9.50	28,033,988–31,799,845	
qPH-Dt9-1	17Aync	Marker179530	Marker179562	3.92	−3.15	6.52	58.50–61.50	34,267,094–35,396,291	
qPH-Dt11-1	16Aync	Marker188702	Marker188604	5.83	4.38	7.70	27.50–28.50	42,846,459–44,363,009	
qPH-Dt12-1	16Aync	Marker193969	Marker193208	10.18	−5.75	14.92	68.50–72.50	27,731,241–39,277,172	
qPH-Dt13-1	17Aync	Marker195491	Marker195480	5.30	3.43	9.15	91.50–99.50	0–976,211	

Add and PVE represent the additive effect and phenotypic variation explained by the QTL, respectively


a
At3
LOD
PVE (%)
Genetic distance (cM)
b
At3
LOD
PVE (%)
Genetic distance (cM)
c
At5
LOD
PVE (%)
Genetic distance (cM)
d
LOD
At5
PVE (%)
Genetic distance (cM)
e
LOD
Dt1
PVE (%)
Genetic distance (cM)
f
LOD
Dt1
PVE (%)
Genetic distance (cM)
g
LOD
Dt1
PVE (%)
Genetic distance (cM)
h
LOD
Dt3
PVE (%)
Genetic distance (cM)
i
LOD
Dt3
PVE (%)
Genetic distance (cM)
j
LOD
Dt3
PVE (%)
Genetic distance (cM)
k
LOD
Dt3
PVE (%)
Genetic distance (cM)

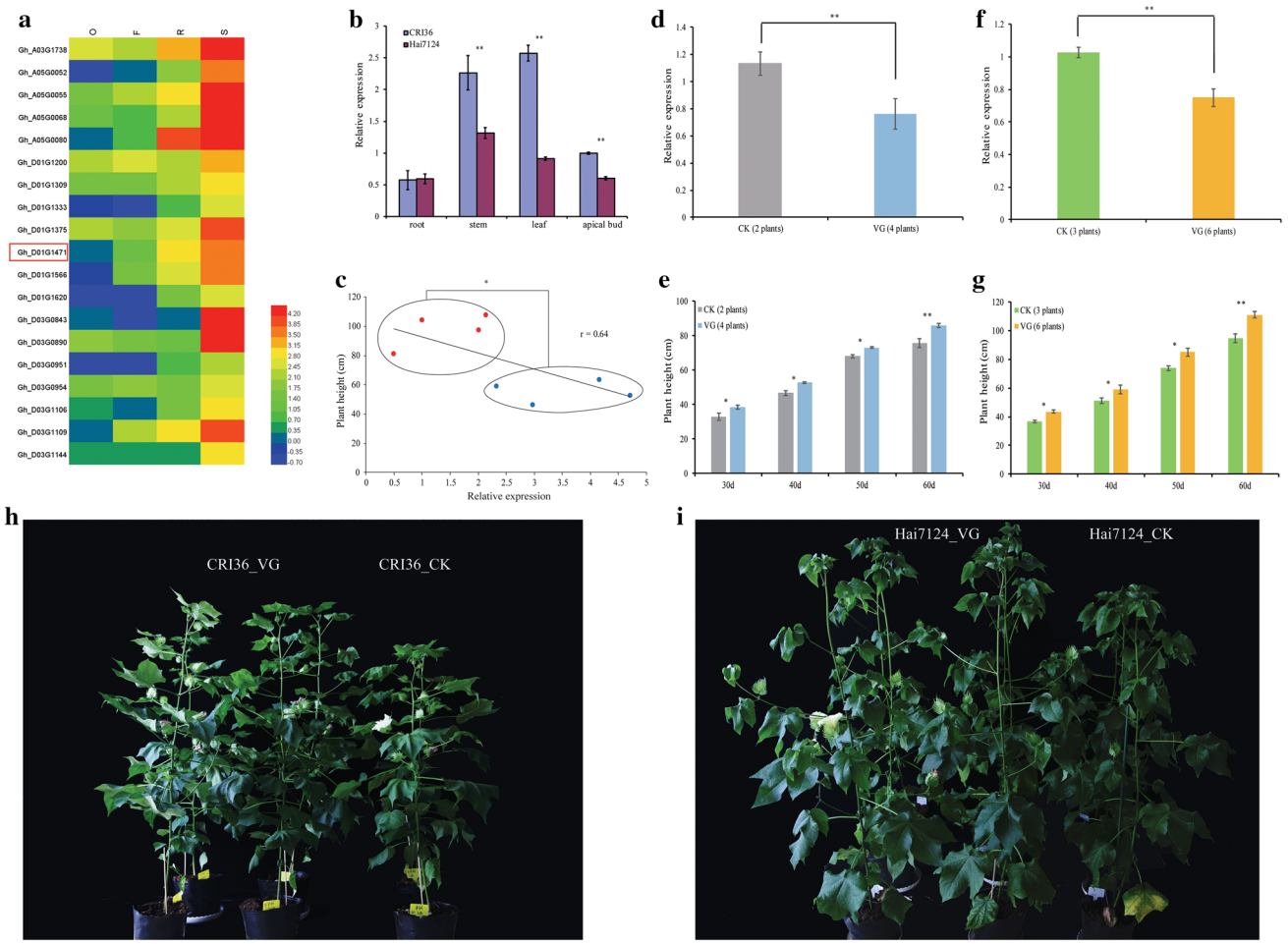


Fig. 4 Expression level and function analysis of *GhPIN3* in cotton. **a** Transcript profiles of promising genes for TM-1 ovule (O), fibre (F), root (R) and stem (S). **b** Expression patterns of *GhPIN3* in different tissues of CRI36 and Hai7124 varieties at three-leaf stage, detected by qRT-PCR. **c** Expression patterns of *GhPIN3* in mixture apical organs of eight varieties of BILs in the flowering period, detected by qRT–PCR; red and blue dots indicate four tall-statured cotton and short-statured cotton, respectively. **d** The mean expression level of *GhPIN3* in control check (CK) and pCLCrVA:*GhPIN3* (VG) plants of CRI36, detected by qRT-PCR. **e** The mean PH of control check (CK) and pCLCrVA:*GhPIN3* (VG) plants in CRI36 cotyledons infected with virus for 30, 40, 50 and 60 days. **f** The mean expression level of *GhPIN3* in control check (CK) and pCLCrVA:*GhPIN3* (VG) plants of Hai7124, detected by qRT–PCR. **g** The mean PH of control check (CK) and pCLCrVA:*GhPIN3* (VG) plants in Hai7124 cotyledons infected with virus for 30, 40, 50 and 60 days. **h** PH phenotypic of CRI36 cotyledons infected with virus for 60 days. **i** PH phenotypic of Hai7124 cotyledons infected with virus for 60 days. * and ** indicate significant differences at $p = 0.05$ and 0.01, respectively

BILs, and the expression level of *GhPIN3* in the four short-statured lines was significantly higher than that in the four tall-statured lines (Fig. 4c). The linear regression equation between *GhPIN3* expression level (X) and PH (Y) of the eight lines was expressed as $Y = -10.91X + 103.64$, with $r = 0.64$ (Fig. 4c). From the data above, we hypothesized that the expression level of GhPIN3 is negatively related to PH.

The function of *GhPIN3* in cotton

To further validate the role of *GhPIN3* in cotton plant height development, we constructed the pCLCrVA:*GhPIN3* recombinant virus vector to suppress the expression level of *GhPIN3* in the recurrent parent CRI36, and the pLCrVA virus vector was used as the control (CK). Four pCLCrVA:*GhPIN3*-silenced CRI36 plants and two CRI36 CK plants were examined, and qRT-PCR was used to confirm the expression level of the silenced *GhPIN3* in these plants. The result indicated that the mean expression levels in the CRI36 pCLCrVA:*GhPIN3*-silenced plants decreased by approximately 32.76%, as compared to the CRI36 CK (Fig. 4d). The mean PH of pCLCrVA:*GhPIN3*-silenced CRI36 plants was significantly higher than that of CRI36 CK plants at four different time points (30, 40, 50 and 60 d), and the difference was approximately 10.33 cm at the

60-d time point (Fig. 4e; Fig. 4h). The results were consistent with the second VIGS experiment using the other parent Hai7124. We examined three Hai7124 CK plants and six pCLCrVA:*GhPIN3*-silenced Hai7124 plants with qRT-PCR. The mean expression level in the Hai7124 pCLCrVA:*GhPIN3*-silenced plants significantly fell by 26.89%, than that in the Hai7124 CK (Fig. 4f). Similar to CRI36, the mean PH of Hai7124 pCLCrVA:*GhPIN3*-silenced plants was significantly higher than that of CK plants at four different time points (30, 40, 50 and 60 d), and the difference was approximately 16.17 cm at the 60-d time point (Fig. 4g, i). From the results above, we inferred that *GhPIN3* contributes to PH in cotton.

Discussion

A fine genetic map plays a crucial role in identifying QTLs of traits of interest. Recently, numerous genetic maps have been constructed by PCR-based markers in cotton, such as the genetic map harbouring 155 SSR markers and spanning a total length of 959.4 cM constructed by Sun et al. (2011). Shang et al. (2015b) used 581 SSR markers to construct a genetic map, and the average distance between two adjacent markers was approximately 6.39 cM. Liu et al. (2015) constructed a genetic map with 1,675 SSR markers that spanned a total genetic distance of 3,338.2 cM. However, the number of SSR markers is not enough to construct a high-density genetic map in cotton. By the use of advanced sequencing methods, the marker density of the genetic map of cotton has undergone a remarkable increase. Jia et al. (2016) used 6,295 SNP markers and 139 SSR markers to construct a genetic map with a total length of 4071.98 cM. Zhang et al. (2016) used 5,521 SNP markers to construct a genetic map encompassing 3,259.37 cM with an average of 0.78 cM between adjacent markers. Sun et al. (2017) used 3,978 SNP markers to construct a genetic map that spanned a total genetic distance of 2,480 cM. Therefore, sequencing methods are more efficient in constructing genetic maps. In the present study, SNP markers were used to construct an HDGM through the SLAF-seq method. In total, 7,709 high-quality markers were identified in 26 linkage groups. Each linkage group contained 52–658 markers, and the length of these groups ranged from 59.97 to 221.99 cM. The average marker density of the 26 linkage groups was approximately 0.67 cM per SNP; moreover, the average marker density on the 22 linkage groups was lower than 0.5 cM per SNP, such as on Dt5, At4 and At9, where the average marker density was 0.12, 0.23 and 0.26 cM per SNP, respectively. The average ratio of gaps < 5 cM on all chromosomes was 97.27%, and the collinearity analysis validated the reliability of the constructed map. These results suggest that this high-density genetic map contributed to the reliable QTL analysis for PH in this study.

China is the largest cotton producer among more than 80 cotton-producing countries, while the planting areas for cotton in China have significantly decreased in the past decade due to the high labour force demand, especially for harvesting seed cotton by hand. To boost harvest efficiency, the mechanized harvesting of cotton is gradually spreading in the main cotton production areas. Variability in the PH of cotton cultivars is one of the main obstacles to expand mechanization. Moreover, PH is one of the main factors influencing the yield in crops, and this trait has been extensively researched in maize, rice, wheat, *Brassica napus* and other crops (Chen et al. 2012; Schiessl et al. 2015; Würschum et al. 2017; Xing et al. 2015). To date, most of the published genetic studies on cotton have focused on yield and fibre quality traits, whereas few studies have focused on PH. In the present study, a total of ten QTLs were identified using the fine genetic map. Among the ten QTLs, eight were identified for the first time, and the other two QTLs overlapped with PH-associated QTLs detected in previous studies. In particular, the gene *Gh_D03G0922*, located in qPH-Dt3-1, has been verified as responsible for PH in upland cotton by Su et al. (2018). These results suggest that more novel and reliable QTLs that contribute to PH were obtained in this study. Combining the genes in stable QTLs with functional annotation of orthologues in Arabidopsis and transcriptome data is an efficient method to mine candidate genes. Using this method, the functions of several previously undescribed causal genes were identified in cotton (Du et al. 2018a; Ma et al. 2018a, b), and the *GhPIN3*, located in qPH-Dt1-1, was mined in this study. When the *GhPIN3* was suppressed in CRI36 and Hai7124, the pCLCrVA:*GhPIN3*-silenced plants were significantly taller than the CK plants. These results showed that *GhPIN3* participates in PH development. This study provides the first HDGM for an interspecific BIL population of cotton by SLAF-seq. The map contains a total of 7,709 high-quality SNP markers covering a 3,433.24-cM genetic distance. Using the fine HDGM, QTLs of PH across seven environments and with BLUP were identified. In total, ten QTLs related to PH were identified: three of these were identified in the previous studies, and the other seven were first found in the present study. Through a combination of transcriptome and qRT-PCR analyses, the candidate gene *GhPIN3*, located in the stable QTL qPH-Dt1-1, was identified. Upon further analysis by VIG, we found that *GhPIN3* influenced PH in cotton. In this study, the detected QTLs and candidate gene *GhPIN3* lay a foundation for cultivating an ideal PH in cotton.

Author contribution statement SXY, JFZ and JWY conceived and designed the experiments. YHG, GYL, JL, YPC, XZ, XLL, DL, XSZ, JKS and SRT performed the

field cultivation of cotton plants and PH measurements. JFZ, WFP, MW and QFM participated in the experimental design, and drafting and revision of the manuscript. JJM performed the experiments and wrote the manuscript. All the authors read and approved the final manuscript.

Acknowledgements The authors wish to thank New Mexico Agricultural Experiment Station, New Mexico, USA. This research was supported by grants from the National Natural Science Foundation of China (Grant Nos. 31621005 and 31701474), the National Key Research and Development Program of China (Grant Nos. 2018YFD0100300 and 2016YFD0101400).

Compliance with ethical standards

Conflict interests The authors declare that they have no competing interests.

Ethical standards The authors state that all experiments in the study comply with the ethical standards in the USA.

References

Cao Z, Wang P, Zhu X, Chen H, Zhang T (2014) SSR marker-assisted improvement of fiber qualities in Gossypium hirsutum using *G. barbadense* introgression lines. Theor Appl Genet 127:587–594

Chen Y, Fan X, Song W, Zhang Y, Xu G (2012) Over-expression of OsPIN2 leads to increased tiller numbers, angle and shorter plant height through suppression of OsLAZY1. Plant Biotechnol J 10:139–149

Chen W, Yao J, Li Y, Zhao L, Liu J, Guo Y, Wang J, Yuan L, Liu Z, Lu Y, Zhang Y (2018) Nulliplex-branch, a TERMINAL FLOWER 1 ortholog, controls plant growth habit in cotton. Theor Appl Genet 132:97–112

Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M, Jia Y, Pan Z, Gong W, Liu Z, Zhu H, Ma L, Liu F, Yang D, Wang F, Fan W, Gong Q, Peng Z, Wang L, Wang X, Xu S, Shang H, Lu C, Zheng H, Huang S, Lin T, Zhu Y, Li F (2018a) Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. Nat Genet 50:796–802

Du X, Liu S, Sun J, Zhang G, Jia Y, Pan Z, Xiang H, He S, Xia Q, Xiao S, Shi W, Quan Z, Liu J, Ma J, Pang B, Wang L, Sun G, Gong W, Jenkins JN, Lou X, Zhu J, Xu H (2018b) Dissection of complicate genetic architecture and breeding perspective of cottonseed traits by genome-wide association study. BMC Genom 19:451

Friml J, Vieten A, Sauer M, Weijers D, Schwarz H, Hamann T, Offringa R, Jürgens G (2003) Efflux-dependent auxin gradients establish the apical–basal axis of Arabidopsis. Nature 426:147–153

Gu Z, Huang C, Li F, Zhou X (2014) A versatile system for functional analysis of genes and microRNAs in cotton. Plant Biotechnol J 12:638–649

Huang C, Shen C, Wen T, Gao B, Zhu Li X, Ahmed MM, Li D, Lin Z (2018) SSR-based association mapping of fiber quality in upland cotton using an eight-way MAGIC population. Mol Genet Genom 293:793–805

Ji G, Zhang Q, Du R, Lv P, Ma X, Fan S, Li S, Hou S, Han Y, Liu G (2017) Construction of a high-density genetic map using specific-locus amplified fragments in sorghum. BMC Genom 18:51

Jia X, Pang C, Wei H, Wang H, Ma Q, Yang J, Cheng S, Su J, Fan S, Song M (2016) High-density linkage map construction and QTL analysis for earliness-related traits in *Gossypium hirsutum* L. BMC Genom 17:909

Keerio AA, Shen C, Nie Y, Ahmed MM, Zhang X, Lin Z (2018) QTL mapping for fiber quality and yield traits Based on introgression lines derived from *Gossypium hirsutum* × *G. tomentosum*. Int J Mol Sci 19:243

Kimura M, Kagawa T (2006) Phototropin and light-signaling in phototropism. Curr Opin in Plant Biol 9:503–508

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

Li C, Dong Y, Zhao T, Li L, Li C, Yu E, Mei L, Daud MK, He Q, Chen J, Zhu S (2016a) Genome-wide SNP linkage mapping and QTL analysis for fiber quality and yield traits in the upland cotton recombinant inbred lines population. Front Plant Sci 7:1356

Li X, Jin X, Wang H, Zhang X, Lin Z (2016b) Structure, evolution, and comparative genomics of tetraploid cotton based on a high-density genetic linkage map. DNA Res 23:283–293

Liu D, Ma C, Hong W, Huang L, Liu M, Liu H, Zeng H, Deng D, Xin H, Song J, Xu C, Sun X, Hou X, Wang X, Zheng H (2014) Construction and analysis of high-density linkage map using high-throughput sequencing data. PLoS One 9:e98855

Liu D, Liu F, Shan X, Zhang J, Tang S, Fang X, Liu X, Wang W, Tan Z, Teng Z, Zhang Z, Liu D (2015) Construction of a high-density genetic map and lint percentage and cottonseed nutrient trait QTL identification in upland cotton (*Gossypium hirsutum* L.). Mol Genet Genom 290:1683–1700

Liu S, Fan C, Li J, Cai G, Yang Q, Wu J, Yi X, Zhang C, Zhou Y (2016a) A genome-wide association study reveals novel elite allelic variations in seed oil content of *Brassica napus*. Theor Appl Genet 129:1203–1215

Liu T, Guo L, Pan Y, Zhao Q, Wang J, Song Z (2016b) Construction of the first high-density genetic linkage map of Salvia miltiorrhiza using specific length amplified fragment (SLAF) sequencing. Sci Rep 6:24070

Liu D, Teng Z, Kong J, Liu X, Wang W, Zhang X, Zhai T, Deng X, Wang J, Zeng J (2018) Natural variation in a CENTRORADIALIS homolog contributed to cluster fruiting and early maturity in cotton. BMC Plant Biol 18:1

Liu G, Pei W, Li D, Ma J, Cui Y, Wang N, Song J, Wu M, Li L, Zang X, Yu S, Zhang J, Yu J (2019) A targeted QTL analysis for fiber length using a genetic population between two introgressed backcrossed inbred lines in upland cotton (*Gossypium hirsutum*). Crop J 7:273–282

Ma J, Geng Y, Pei W, Wu M, Li X, Liu G, Li D, Ma Q, Zang X, Yu S, Zhang J, Yu J (2018a) Genetic variation of dynamic fiber elongation and developmental quantitative trait locus mapping of fiber length in upland cotton (*Gossypium hirsutum* L.). BMC Genom 19:882

Ma Z, He S, Wang X, Sun J, Zhang Y, Zhang G, Wu L, Li Z, Liu Z, Sun G, Yan Y, Jia Y, Yang J, Pan Z, Gu Q, Li X, Sun Z, Dai P, Liu Z, Gong W, Wu J, Wang M, Liu H, Feng K, Ke H, Wang J, Lan H, Wang G, Peng J, Wang N, Wang L, Pang B, Peng Z, Li R, Tian S, Du X (2018b) Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. Nat Genet 50:803–813

Meng L, Li H, Zhang L, Wang J (2015) QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J 3:269–283

Monna L, Kitazawa N, Yoshino R, Suzuki J, Masuda H, Maehara Y, Tanji M, Sato M, Nasu S, Minobe Y (2002) Positional cloning of rice

semidwarfing gene, sd-1: rice "Green revolution gene" encodes a mutant enzyme involved in gibberellin synthesis. DNA Res 9:11–17
Palme K, Dovzhenko A, Ditengou FA (2006) Auxin transport and gravitational research: perspectives. Protoplasma 229:175
Peiffer JA, Romay MC, Gore MA, Flintgarcia SA, Zhang Z, Millard MJ, Gardner CA, Mcmullen MD, Holland JB, Bradbury PJ (2014) The genetic architecture of maize height. Genetics 196:1337–1356
Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F (1999) 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256
Petrásek J, Mravec J, Bouchard R, Blakeslee JJ, Abas M, Seifertová D, Wisniewska J, Tadele Z, Kubes M, Covanová M (2006) Petráek, J. PIN proteins perform a rate-limiting function in cellular auxin efflux. Science 312:914–918
Said JI, Knapka JA, Song M, Zhang J (2015a) Cotton QTLdb: a cotton QTL database for QTL analysis, visualization, and comparison between *Gossypium hirsutum* and *G. hirsutum*×*G. barbadense* populations. Mol Genet Genom 290:1615–1625
Said JI, Song M, Wang H, Lin Z, Zhang X, Fang DD, Zhang J (2015b) A comparative meta-analysis of QTL between intraspecific Gossypium hirsutum and interspecific *G. hirsutum*×*G. barbadense* populations. Mol Genet Genom 290:1003–1025
Sakamoto T, Miura K, Itoh H, Tatsumi T, Ueguchi-Tanaka M, Ishiyama K, Kobayashi M, Agrawal GK, Takeda S, Abe K, Miyao A, Hirochika H, Kitano H, Ashikari M, Matsuoka M (2004) An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiol 134:1642–1653
Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, Kitano H, Matsuoka M (2002) Green revolution: a mutant gibberellin-synthesis gene in rice - New insight into the rice variant that helped to avert famine over thirty years ago. Nature 416:701–702
Schiessl S, Iniguez-Luy F, Wei Q, Snowdon RJ (2015) Diverse regulatory factors associate with flowering time and yield responses in winter-type *Brassica napus*. BMC Genom 16:737
Shang L, Liang Q, Wang Y, Wang X, Wang K, Abduweli A, Ma L, Cai S, Hua J (2015a) Identification of stable QTLs controlling fiber traits properties in multi-environment using recombinant inbred lines in Upland cotton (*Gossypium hirsutum* L.). Euphytica 205:877–888
Shang L, Liu F, Wang Y, Abduweli A, Cai S, Wang K, Hua J (2015b) Dynamic QTL mapping for plant height in Upland cotton (*Gossypium hirsutum*). Plant Breed 134:703–712
Shang L, Wang Y, Cai S, Ma L, Liu F, Chen Z, Su Y, Wang K, Hua J (2016) Genetic analysis of upland cotton dynamic heterosis for boll number per plant at multiple developmental stages. Sci Rep 6:35515
Shi Y, Zhang B, Liu A, Li W, Li J, Lu Q, Zhang Z, Li S, Gong W, Shang H, Gong J, Chen T, Ge Q, Wang T, Zhu H, Liu Z, Yuan Y (2016) Quantitative trait loci analysis of Verticillium wilt resistance in interspecific backcross populations of *Gossypium hirsutum*×*Gossypium barbadense*. BMC Genom 17:877
Su J, Li L, Zhang C, Wang C, Gu L, Wang H, Wei H, Liu Q, Huang L, Yu S (2018) Genome-wide association study identified genetic variations and candidate genes for plant architecture component traits in Chinese upland cotton. Theor Appl Genet 131:1–16
Sun F-D, Zhang J-H, Wang S-F, Gong W-K, Shi Y-Z, Liu A-Y, Li J-W, Gong J-W, Shang H-H, Yuan Y-L (2011) QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Mol Breed 30:569–582
Sun XW, Liu DY, Zhang XF, Li WB, Liu H, Hong WG, Jiang CB, Guan N, Ma CX, Zeng HP, Xu CH, Song J, Huang L, Wang CM, Shi JJ, Wang R, Zheng XH, Lu CY, Wang XW, Zheng HK (2013) SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PLoS One 8:e58700
Sun Z, Wang X, Liu Z, Gu Q, Zhang Y, Li Z, Ke H, Yang J, Wu J, Wu L (2017) Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L. Plant Biotechnol J 15:982–996
Tao A, Huang L, Wu G, Afshar RK, Qi J, Xu J, Fang P, Lin L, Zhang L, Lin P (2017) High-density genetic map construction and QTLs identification for plant height in white jute(Corchorus capsularis L.) using specific locus amplified fragment (SLAF) sequencing. BMC Genom 18:355
Teng F, Zhai L, Liu R, Bai W, Wang L, Huo D, Tao Y, Zheng Y, Zhang Z (2013) ZmGA3ox2, a candidate gene for a major QTL, qPH3.1, for plant height in maize. Plant J 73:405–416
Van OH, Stam P, Visser RG, van Eck HJ (2005) SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. Theor Appl Genet 112:187–194
Van Ooijen JW (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet Res 93:343–349
Wang L, Guo M, Li Y, Ruan W, Mo X, Wu Z, Sturrock C, Yu H, Lu C, Peng J (2018a) LARGE ROOT ANGLE1, encoding OsPIN2, is involved in root system architecture in rice. J Exp Bot 69:385–397
Wang Y, Zhang T, Wang R, Zhao Y (2018b) Recent advances in auxin research in rice and their implications for crop improvement. J Exp Bot 69:255
Wen T, Wu M, Shen C, Gao B, Zhu Zhang X, You C, Lin Z (2018) Linkage and association mapping reveals the genetic basis of brown fibre (*Gossypium hirsutum*). Plant Biotechnol J 16:1654–1666
Würschum T, Langer SM, Longin C, Tucker MR, Leiser WL (2017) A modern green revolution gene for reduced height in wheat. Plant J 92:892–903
Xing A, Gao Y, Ye L, Zhang W, Cai L, Ching A, Llaca V, Johnson B, Liu L, Yang X, Kang D, Yan J, Li J (2015) A rare SNP mutation in Brachytic2 moderately reduces plant height and increases yield potential in maize. J Exp Bot 66:3791–3802
Yu J, Yu S, Lu C, Wang W, Fan S, Song M, Lin Z, Zhang X, Zhang J (2007) High-density linkage map of cultivated allotetraploid cotton based on SSR, TRAP, SRAP and AFLP markers. J Integr Plant Biol 49:716–724
Yu J, Yu S, Fan S, Song M, Zhai H, Li X, Zhang J (2012) Mapping quantitative trait loci for cottonseed oil, protein and gossypol content in a *Gossypium hirsutum*×*Gossypium barbadense* backcross inbred line population. Euphytica 187:191–201
Yu J, Zhang K, Li S, Yu S, Zhai H, Wu M, Li X, Fan S, Song M, Yang D (2013) Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum*×*Gossypium barbadense* backcross inbred line population. Theor Appl Genet 126:275–287
Zhang J, Zhang Q, Cheng T, Yang W, Pan H, Zhong J, Huang L, Liu E (2015a) High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (*Prunus mume* Sieb. et Zucc). DNA Res 22:183
Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM (2015b) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol 33:531–537
Zhang Z, Li J, Muhammad J, Cai J, Jia F, Shi Y, Gong J, Shang H, Liu A, Chen T, Ge Q, Palanga KK, Lu Q, Deng X, Tan Y, Li W, Sun L, Gong W, Yuan Y (2015c) High resolution consensus mapping of quantitative trait loci for fiber strength, length and micronaire on chromosome 25 of the upland cotton (*Gossypium hirsutum* L.). PLoS One 10:e0135430
Zhang Z, Shang H, Shi Y, Huang L, Li J, Ge Q, Gong J, Liu A, Chen T, Wang D, Wang Y, Palanga KK, Muhammad J, Li W, Lu Q, Deng X, Tan Y, Song W, Cai J, Li P, Rashid H, Gong W, Yuan Y (2016) 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 S, Zhang Z, Bales C, Gu C, DiFonzo C, Li M, Song Q, Cregan P, Yang Z, Wang D (2017) Mapping novel aphid resistance QTL from wild soybean, *Glycine soja* 85-32. Theor Appl Genet 130:1941–1952

Zhao J, Gao Y, Zhang Z, Chen T, Guo W, Zhang T (2013) A receptor-like kinase gene (GbRLK) from Gossypium barbadense enhances salinity and drought-stress tolerance in Arabidopsis. BMC Plant Biol 13:110

Zhu QH, Yuan Y, Stiller W, Jia Y, Wang P, Pan Z, Du X, Llewellyn D, Wilson I (2018) Genetic dissection of the fuzzless seed trait in *Gossypium barbadense*. J Exp Bot 69:997–1009

Affiliations

Jianjiang Ma[1,3] · Wenfeng Pei[4] · Qifeng Ma[1] · Yanhui Geng[1] · Guoyuan Liu[1] · Ji Liu[1] · Yupeng Cui[1] · Xia Zhang[1] · Man Wu[1] · Xingli Li[1] · Dan Li[1] · XinShan Zang[1] · Jikun Song[1] · Shurong Tang[1] · Jinfa Zhang[2] · Shuxun Yu[1,3] · Jiwen Yu[1]

Jianjiang Ma
mjj1699@126.com

Wenfeng Pei
peiwenfeng1988@163.com

Qifeng Ma
maqifeng1985@163.com

Yanhui Geng
13733183665@163.com

Guoyuan Liu
cjqm1989@126.com

Ji Liu
liujicricaas@163.com

Yupeng Cui
yupeng_cui851026@163.com

Xia Zhang
2442935057@qq.com

Man Wu
wuman2004@163.com

Xingli Li
liyd@cricaas.com.cn

Dan Li
18768851389@163.com

XinShan Zang
zangxinshan@163.com

Jikun Song
614622761@qq.com

Shurong Tang
tangshurong2008@126.com

[1] State Key Laboratory of Cotton Biology, Cotton Institute of the Chinese Academy of Agricultural Sciences, Key Laboratory of Cotton Genetic Improvement, Ministry of Agriculture, Anyang 455000, Henan, China

[2] Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces 880033, USA

[3] College of Agriculture, Northwest A&F University, Yangling 712100, Shanxi, China

[4] Xinjiang Research Base, State Key Laboratory of Cotton Biology, Xinjiang Agricultural University, Urumqi 830001, China