A comparative meta-analysis of QTL between intraspecific Gossypium hirsutum and interspecific G. hirsutum × G. barbadense populations

Article *in* Molecular Genetics and Genomics · May 2015

Impact Factor: 2.73 · DOI: 10.1007/s00438-014-0963-9

7 authors, including:

Mingzhou Song
New Mexico State University
57 PUBLICATIONS **721** CITATIONS
SEE PROFILE

Hantao Wang
Huazhong Agricultural University
9 PUBLICATIONS **30** CITATIONS
SEE PROFILE

Jinfa Zhang
New Mexico State University
188 PUBLICATIONS **1,737** CITATIONS
SEE PROFILE

Mol Genet Genomics
DOI 10.1007/s00438-014-0963-9

ORIGINAL PAPER

A comparative meta-analysis of QTL between intraspecific *Gossypium hirsutum* and interspecific *G. hirsutum* × *G. barbadense* populations

Joseph I. Said · Mingzhou Song · Hantao Wang · Zhongxu Lin · Xianlong Zhang · David D. Fang · Jinfa Zhang

Received: 27 August 2014 / Accepted: 18 November 2014

Abstract

Key message **Based on 1075 and 1059 QTL from intraspecific Upland and interspecific Upland × Pima populations, respectively, the identification of QTL clusters and hotspots provides a useful resource for cotton breeding.**

Abstract Mapping of quantitative trait loci (QTL) is a prerequisite of marker-assisted selection for crop yield and quality. Recent meta-analysis of QTL in tetraploid cotton (*Gossypium* spp.) has identified regions of the genome with high concentrations of QTL for various traits called clusters and specific trait QTL called hotspots or meta-QTL (mQTL). However, the meta-analysis included all population types of *Gossypium* mixing both intraspecific *G. hirsutum* and interspecific *G. hirsutum* × *G. barbadense* populations. This study used 1,075 QTL from 58 publications on intraspecific *G. hirsutum* and 1,059 QTL from 30 publications on *G. hirsutum* × *G. barbadense* populations to perform a comprehensive comparative analysis of QTL clusters and hotspots between the two populations for yield, fiber and seed quality, and biotic and abiotic stress tolerance. QTL hotspots were further analyzed for mQTL within the hotspots using Biomercator V3 software. The ratio of QTL between the two population types was proportional yet differences in hotspot type and placement were observed between the two population types. However, on some chromosomes QTL clusters and hotspots were similar between the two populations. This shows that there are some universal QTL regions in the cultivated tetraploid cotton which remain consistent and some regions which differ between population types. This study for the first time elucidates the similarities and differences in QTL clusters and hotspots between intraspecific and interspecific populations, providing an important resource to cotton breeding programs in marker-assisted selection .

Communicated by S. Hohmann.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-014-0963-9) contains supplementary material, which is available to authorized users.

J. I. Said · J. Zhang (✉)
Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM, USA
e-mail: jinzhang@nmsu.edu

J. I. Said
e-mail: joesaid@nmsu.edu

M. Song
Department of Computer Science, New Mexico State University, Las Cruces, NM, USA
e-mail: joemsong@nmsu.edu

H. Wang · Z. Lin · X. Zhang
National Key Laboratory of Crop Genetic Improvement and National Centre of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, Hubei, China
e-mail: w.wanghantao@163.com

Z. Lin
e-mail: linzhongxu@mail.hzau.edu.cn

X. Zhang
e-mail: xlzhang@mail.hzau.edu.cn

D. D. Fang
Cotton Fiber Bioscience Research Unit, USDA-ARS-SRRC, New Orleans, LA 70124, USA
e-mail: david.fang@ars.usda.gov

Published online: 12 December 2014

Keywords Cotton · *Gossypium hirsutum* · *Gossypium barbadense* · Meta analysis · Quantitative trait loci

Introduction

There are approximately 50 species of *Gossypium* which originated from a common ancestor 5–15 million years ago (Percival et al. 1999; Wendel and Cronn 2002). Cotton is grown throughout Africa, Asia, Australia, and warm regions of America, and the tropics (Wendel and Cronn 2002). There are five tetraploid AD genome ($2n = 4x = 52$) species of *Gossypium* which evolved from a New World diploid D genome ($2n = 2x = 26$) species (possibly *G. raimondii* Ulbr.) and an Old World diploid A genome ($2n = 2x = 26$) species (presumably *G. herbaceum* L. or *G. arboreum* L.) (Wendel and Cronn 2002; Fryxell 1992). The five species include *G. hirsutum* L., *G. barbadense* L., *G. tomentosum* Nuttall ex Seemann, *G. mustelinum* Watt, and *G. darwinii* Watt. The species *G. hirsutum* and *G. tomentosum* are more related, as did *G. barbadense* and *G. darwinii* (Wendel and Cronn 2002). However, only *G. hirsutum* and *G. barbadense* are cultivated and heavily studied for their agronomic importance. *G. hirsutum* also called Upland cotton is commercially known for producing a high lint yield. *G. barbadense* which is also called Egyptian cotton, Pima cotton, or Sea-island cotton is commercially known for possessing superior fiber strength, length, and fineness traits (Zhang et al. 2014). Extensive crossing of *G. hirsutum* × *G. barbadense* (hereafter *Gh* × *Gb*) to create interspecific populations has been attempted by cotton geneticists and breeders to combine the desirable traits of the two species since the rediscovery of Mendelian genetics, with limited success in developing commercial cultivars (Zhang et al. 2014).

Quantitative trait loci (QTL) are chromosomal regions which contribute cumulatively with varying degrees of phenotypic variance to a phenotypic trait (Mauricio 2001; Miles and Wayne 2008). The identification of QTL and their placement on chromosomes is important to plant breeders. Over the last decade as molecular mapping techniques have improved and become less expensive, there have been numerous QTL studies in cotton. A multitude of fiber quality, yield, biotic and abiotic resistance, morphological, and physiological-related QTL have been identified in cotton. The data from all of these publications are better viewed cumulatively by a meta-analysis to identify regions of the genome which are more heavily populated with QTL than others and will be important in breeding and genetics.

Meta-QTL analysis which uses QTL from previous studies is useful in the identification of regions rich in QTL. A QTL cluster is a densely populated QTL region of the chromosome which contains multiple QTL associated with various traits. A QTL hotspot is a densely populated region of the chromosome which contains multiple QTL for a single specific trait. A meta-QTL (mQTL) identified within a hotspot is the region which the Biomercator V3 meta-analysis software infers is the most likely position of the mQTL within the hotspot. In this study we provide maps to cotton breeders with the positions of QTL clusters, hotspots, and mQTL within hotspots inferred by Biomercator V3 software. Previous meta-QTL analyses in cotton (Lacape et al. 2010; Rong et al. 2007; Said et al. 2013) have identified QTL clusters and hotspots predominantly in interspecific *Gh* × *Gb* populations. The analysis by Said et al. (2013) used QTL data by Rong et al. (2007), Lacape et al. (2010) and 40 other additional intraspecific *G. hirsutum* (hereafter *Gh*) and interspecific QTL studies published to create a comprehensive meta-analysis. However, no meaningful comparison between the two types of populations (i.e., *Gh* and *Gh* × *Gb*) has been performed. As most cotton breeders are focused on Upland cotton only, QTL, and QTL clusters and hotspots from *Gh* × *Gb* would be of little use. Furthermore, important QTL clusters and hotspots have not been recommended to breeders for consideration of marker-assisted selection (MAS). This study provides a comparative map of QTL clusters, hotspots and meta-QTL which can be utilized by breeders for MAS. In addition, many QTL studies were not included in the past reports.

In this current most comprehensive meta-QTL analysis, QTL results from a total of 88 publications including 58 QTL studies based on *Gh* populations (most of which were from China and inaccessible to the international community) and 30 studies using *Gh* × *Gb* populations were used to create a comparison of placement and variation of QTL clusters and hotspots between the two types of populations. In the *Gh* population, 38 papers were from F2 and/or F2:3 populations, while 20 were from recombinant inbred line (RIL) populations. In the *Gh* × *Gb* population, 13 papers were from F2/F2:3 populations, seven from BC1, BC2, or BC3 generations but not inbred lines, four each from RIL and backcross inbred line (BIL) populations, and two from chromosome segment substitution lines (CSSL) or chromosome segment introgression lines (CSIL). Therefore, most of the mapping populations were early segregating generations, so they were not immortal and could not be studied in replicated tests in multiple environments. In fact, only one plant (F2) or limited individuals (F2:3) were tested in one test. False-positive QTL were undoubtedly declared. Furthermore, different studies used different markers and different segregating populations which were tested in different environmental conditions, which make a direct comparison of QTL across studies difficult. Thus, a meta-analysis of QTL is extremely important. This study elucidates differences and similarities in QTL cluster and hotspot placements and distribution over the tetraploid cotton genome between the two populations and places

all QTL on the ["Guazuncho2" (*G. hirsutum*) × "VH8-4602" (*G. barbadense*)] map developed by Lacape et al. (2009) and obtained from the Cotton Marker Database (Blenda et al. 2006). The QTL clusters and hotspots identified in each population are displayed on (Fig. 1). The QTL found within hotspots were subjected to statistical analysis using the Biomercator V3 meta-analysis tool. Such information will be useful to cotton breeders since the markers delineating these regions can be used to select traits of interests in cotton breeding via a marker-assisted selection (MAS) strategy. While this study does not contain QTL studies in cotton from unpublished theses or dissertations due to the lack of peer reviews, the data from this study can be compiled into a functional database which can be updated to keep the data from this study current and useful to the cotton community.

Materials and methods

The previous QTL data from Said et al. (2013) containing 1,223 QTL, from 42 publications on *Gh* and interspecific *Gh* × *Gb* populations were included in this study, and a few of the studies from other interspecific populations were excluded. When confidence interval (CI) or QTL positions were unavailable in a publication flanking markers given in the publication were used as CI to place the QTL on the chromosome, and the average of the CI was used as the position of the QTL. Otherwise QTL positions and CI provided in the publication were used. The CottonGen (Yu et al. 2013a) was used to find marker positions in older studies which did not include exact CI or QTL positions in centimorgans. Using previous QTL data from Said et al. (2013) and using newly published data comparisons between QTL cluster and hotspot regions were made between intraspecific and interspecific populations. Additional QTL from publications since Said et al. (2013) from both *Gh* and *Gh* × *Gb* populations were added to the study for a total of 2,134 QTL (Supplementary Figures 1 and 2). Supplementary Table 1 contains a more detailed summary with the number of each trait QTL on each chromosome for the *Gh* population, while Supplementary Table 2 contains similar data for the interspecific population. Supplementary Tables 3 and 4 contain mQTL data for each trait type identified as a hotspot on each chromosome in both populations. The mQTL data consist of the top four models of mQTL placement and their corresponding Akaike information criterion (AIC) values. The AIC value is calculated by $2k - 2\ln(L)$ where k is the number of parameters in the model and L is the maximized likelihood for the model. Biomercator V3 meta-analysis software uses AIC values to provide the user with four models each with an AIC value. The model with the lowest AIC value and, therefore, best score was chosen to represent the most probable mQTL placement on the chromosome in Supplementary Figures 3 and 4. Supplementary Table 5 contains a description of each trait used in the study. Table 6a contains a summary of the publications used in the *G. hirsutum* QTL used in the study (An et al. 2010; Chen et al. 2008, 2009; Fang et al. 2014; Kong et al. 2011; Jia et al. 2011; Gore et al. 2014; Guo et al. 2008; Gutierrez et al. 2010; Ge et al. 2008; Jiang et al. 2009; Yao et al. 2010; Kumar et al. 2012; Li et al. 2006, 2008, 2010, 2012; Liang et al. 2013; Liu et al. 2010, 2012a; b, 2013a, b, c; Lopez-Lavalle et al. 2012; Mei et al. 2014; Ning et al. 2014; Nusurat et al. 2012; Qin et al. 2008, 2009; Romano et al. 2009; Shen et al. 2005, 2006a, b; Sun et al. 2012; Ulloa et al. 2009; Wang et al. 2006; 2007a, b, c, 2009, 2011, 2012a, b; Hu et al. 2008; Wu et al. 2009; Yang et al. 2007, 2009; Xu et al. 2010; Yin et al. 2002; Wang et al. 2010; Zhang et al. 2005, 2006, 2009, 2010, 2011a; b, c, 2012, 2013; Zhao et al. 2014). Supplementary Table 6b contains a summary of the publications used in the interspecific *Gh* x *Gb* QTL used in the study (Chee et al. 2005a, b; Draye et al. 2005; Gutierrez et al. 2011; Jiang et al. 1998, 2000; Lacape et al. 2005, 2010, 2013; Mei et al. 2004; Paterson et al. 2003; Saranga et al. 2004; Shen et al. 2006a, b, 2010 ; Su et al. 2013; Wang et al. 2008, 2012a, b, c, 2013; Wright et al. 1998, 1999; Yang et al. 2008; Yu et al. 2012, 2013b, c; Ulloa et al. 2009, 2011, 2013, Fang et al. 2013).

Both populations were mapped using Biomercator V3 software (Arcade et al. 2004) to the concensus ["Guazuncho2" (*G. hirsutum*) × "VH8-4602" (*G. barbadense*)] map developed by Lacape et al. (2009) and obtained from the Cotton Marker Database (Blenda et al. 2006). The Biomercator V3 software requires two input files, a map file and a QTL file. The map file contains distances between markers on each chromosome. The QTL file requires ten values for each QTL, the map name, QTL name, chromosome number, trait, LOD score, phenotypic variance (R^2), yes or no if single interval mapping (SIM) was used, position of QTL, and right and left CIs. The map file and QTL file are loaded into the software in tab delimited format. Each population's QTL were mapped to the ["Guazuncho2" (*G. hirsutum*) × "VH8-4602" (*G. barbadense*)] map separately. Each of the 26 chromosomes from the two types of populations was then compared for cluster and hotspot number, placement and hotspot trait type.

It was observed that aggregates of QTL on the genome usually existed within approximately 20 cM regions. In this study, a QTL cluster was defined as a QTL-rich region which contained four or more QTL of various trait types, and a hotspot was defined as four or more QTL of the same trait type within a 20 cM region. For this reason, clusters and hotspots were manually inferred based on the presence of four (with a false-positive rate of 6.25 %) or more QTL within approximately a 20 cM region on each chromosome with a few exceptions. The presence of just single QTL in a region would have a false-positive rate of 50 %; however,

Fig. 1 Clusters and hotspots identified from *G. hirsutum* intraspecific and interspecific *G. hirsutum* x *G. barbadense* populations. Below are the chromosomes from both the *G. hirsutum* intraspecific and interspecific populations with QTL clusters and hotspots. The similarities and differences in QTL cluster and hotspot placement can be seen in the figures below. The trait abbreviations used in the legend above are as follows: *AA* amino acid, *BN* boll number, *BW* boll weight, *FBNode* fruiting branch node, *FBNum* fruiting branch number, *FE* fiber elongation, *FL* fiber length, *FS* fiber strength, *Fusarium* Fusarium wilt, *LP* lint percent, *Micronaire*, micronaire/fiber fineness, *PH* plant height, *FU* fiber uniformity, *HairL* leaf hair, *HairS* stem hair, *LeafM* leaf morphology, *Color* fiber color, *PF* percent fiber; *PetCol* petal color; *PetSpot* petal spots, *PolCol* pollen color, *SI* seed index, *SW* seed weight, *VW* Verticillium wilt, *RN-Nematode* root-knot nematode

with each additional QTL this rate is reduced by 50 %. At times the region was <20 cM as all QTL in the cluster or hotspot were confined to a smaller region. In a few other cases, some chromosomes had regions densely packed with many QTL in that the range was extended beyond 20 cM up to 29 cM to include all QTL in the region. The existence of clusters was manually inferred using BiomercatorV3 from Supplementary Figures 1 and 2. QTL hotspots were also manually inferred; however, they were confirmed using Biomercator V3 meta-analysis feature. The Biomercator V3 software utilizes a maximum likelihood method to calculate the likely position of a cluster

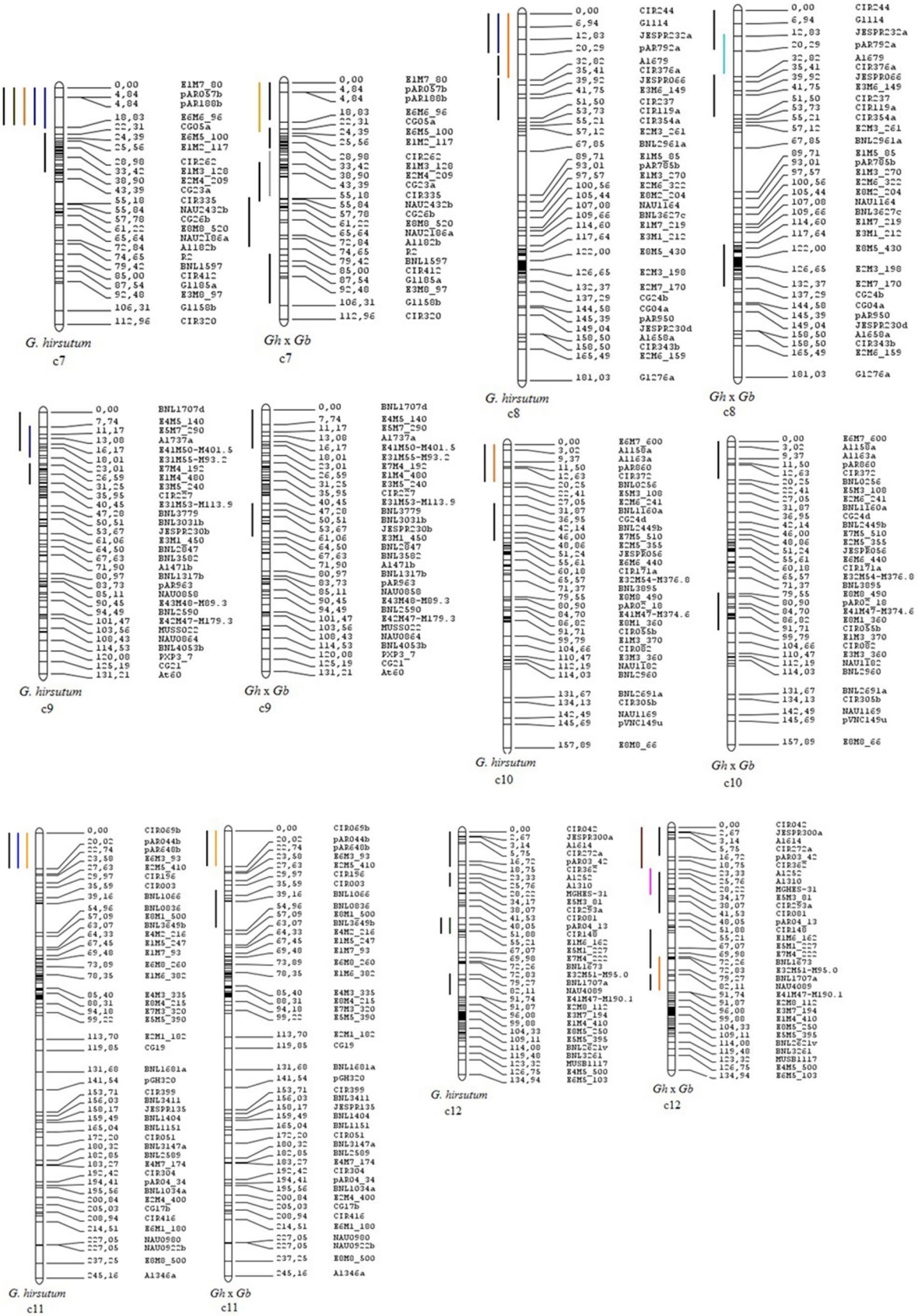

Fig. 1 continued

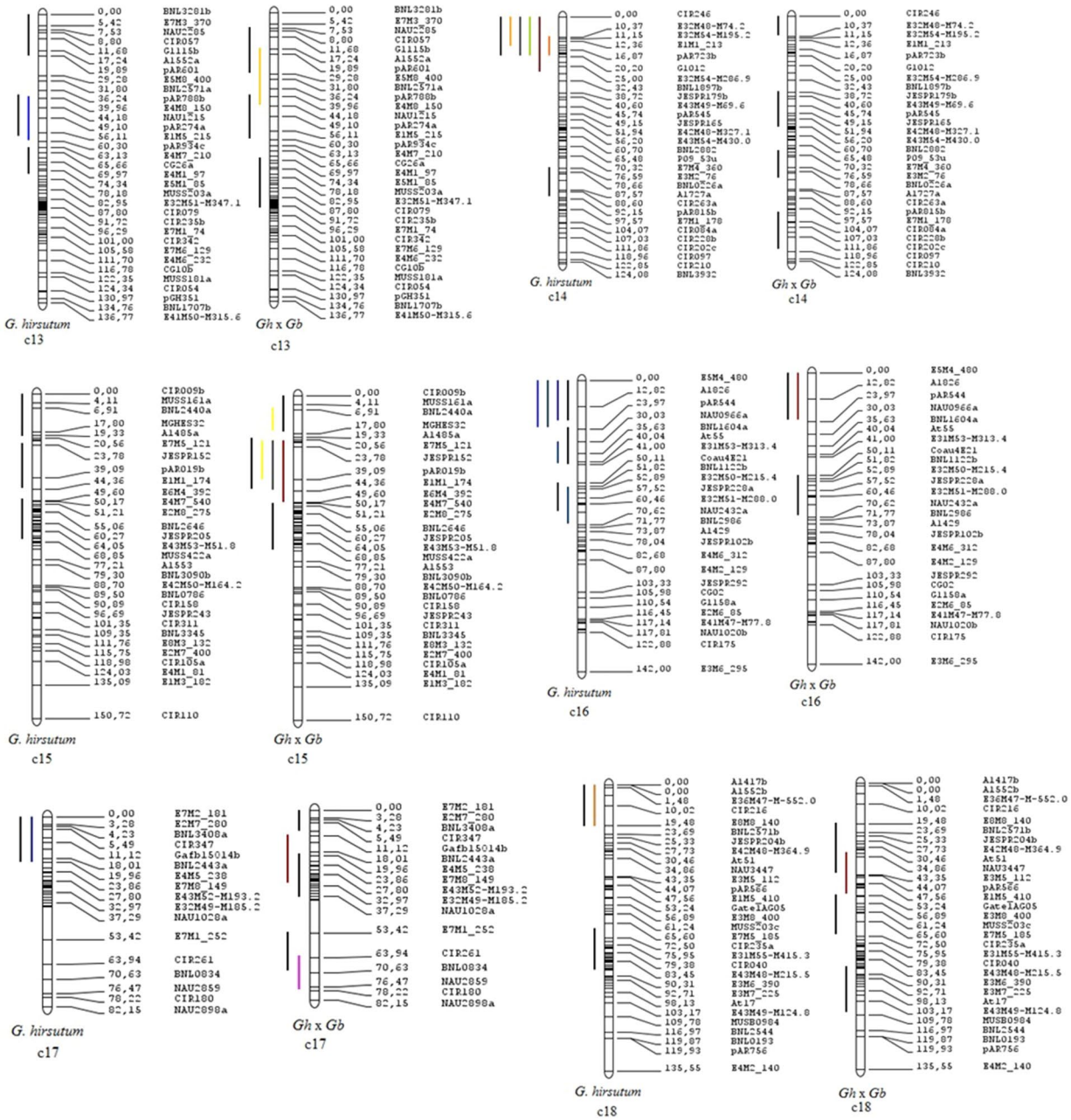

Fig. 1 continued

(Arcade et al. 2004). Supplementary Figures 3 and 4 contain the results of the Biomercator V3 meta-analysis feature for the mQTL hotspots identified. The meta-analysis feature analyzes each chromosome only taking one trait type into consideration and then places mQTL regions based on the confidence intervals and placement of the QTL in the region. The software then assigns AIC values to the top four models which mQTL can be placed on the chromosome according to the QTL in the analysis. If only one QTL is present in a region the software will assign an mQTL region on the basis of only that QTL. For this reason only mQTL inferred by the software which overlapped with previously defined QTL hotspots were considered. The QTL hotspot regions defined by the manual inference have a lower probability of being a false positive than regions containing fewer than four QTL. At times the software infers more than one mQTL within a hotspot region narrowing confidence intervals from the original 20 cM to narrower regions within the hotspot. The mQTL seen in Supplementary Figures 3 and 4 serve as a more in-depth

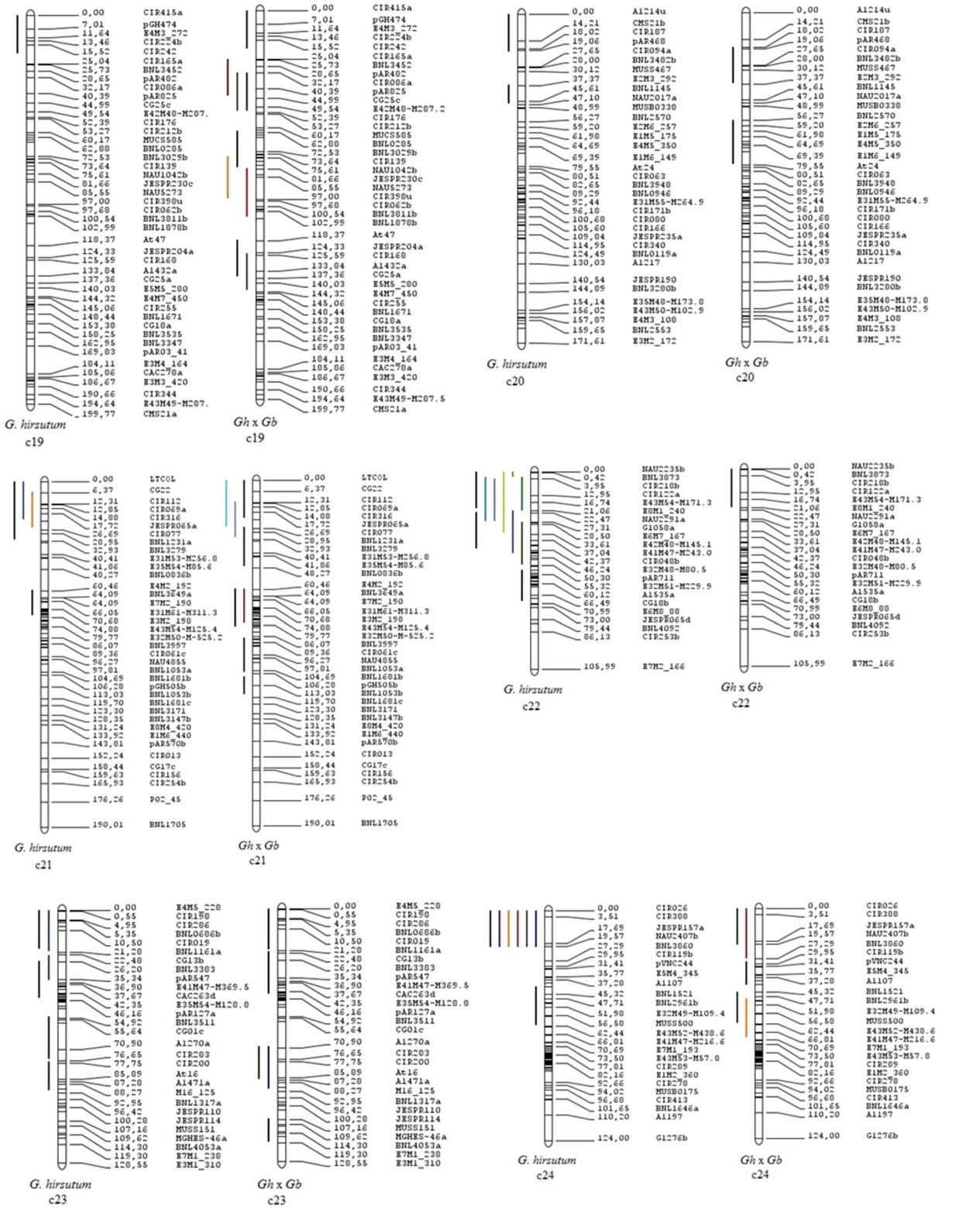

Fig. 1 continued

analysis of QTL hotspot regions which may be of interest to cotton breeders looking to utilize these data for specific traits in MAS. During manual inference, QTL with CIs which spanned two clusters were present; they were excluded from either cluster QTL counts as it was unclear as to their placement. When QTL CIs placed the majority

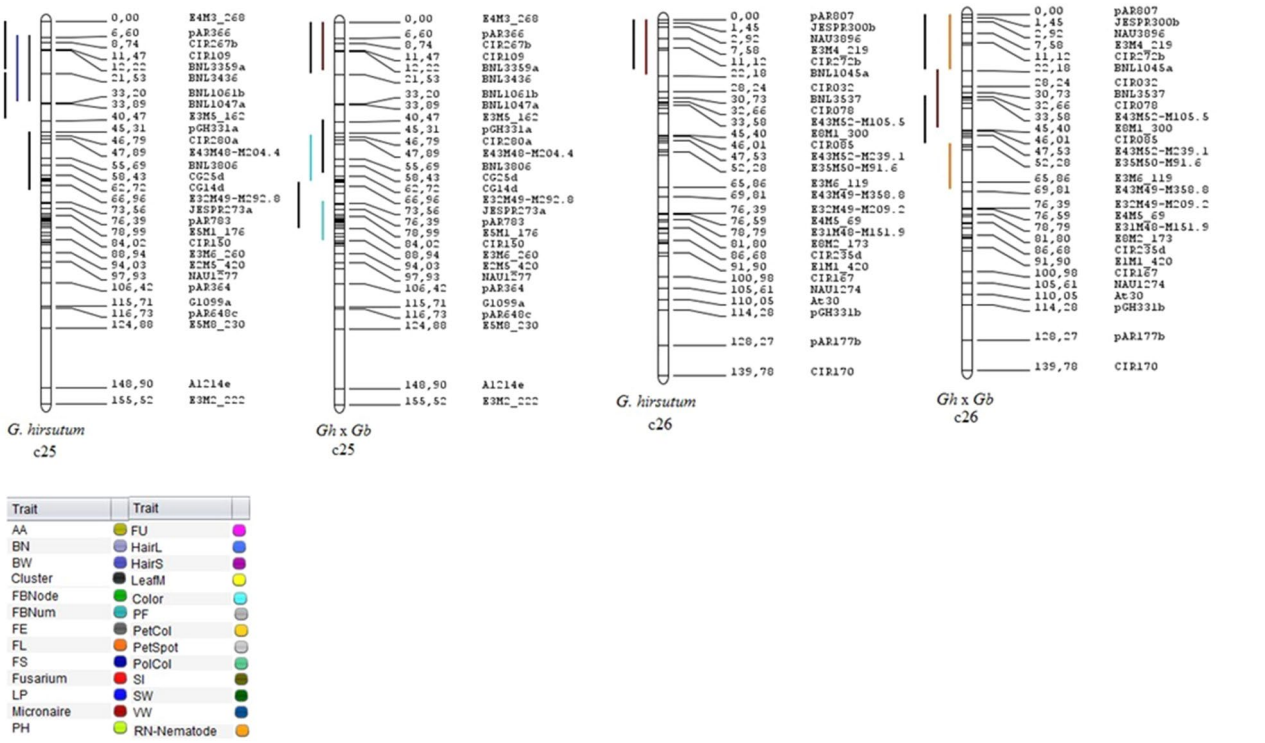

Fig. 1 continued

of one side of the CI in a cluster or hotspot region it was included in the QTL count for the cluster or hotspot.

The identification of putative gene sequences within hotspot regions was accomplished using sequence data from the CottonGen database (Yu et al. 2013a). Centimorgan (cM) positions were translated into nucleotide positions on the genome by equating 1 cM to 500,000 nucleotides. Using the cM intervals of hotspot regions the nucleotide positions were identified and selected on the chromosome. The CottonGen database (Yu et al. 2013a) maps predicted gene sequences from genome data on to the chromosome and displays it graphically. The A2 and D5 genomes were used to investigate the presence of gene within hotspot regions identified in the *Gh* and interspecific *Gh* × *Gb* populations. The CottonGen database provided genomic data on hotspot regions of the A2 and D5 genomes in Genbank file format which was viewed using SnapGene software. SnapGene provides sequence data including gene positions and predicted ORFs from the Genbank files provided by the CottonGen database. The number of genes and positions within hotspot regions was determined using data from the CottonGen database and analyzed with SnapGene software. To investigate conserved domains in the region the conserved domain database (CDD) was downloaded from NCBI. Using BLASTx on the command line the CDD file containing CD protein sequences were used as a protein database against the translated nucleotide query of the hotspot region. Both the fiber length and amino acid hotspot regions were used as queries against the CDD using BLASTx. The results were generated in tabular format using the command line.

Results

Supplementary Figures 1 and 2 display QTL on all 26 tetraploid cotton chromosomes for the intraspecific and interspecific populations, respectively. The figures can help researchers identify QTL regions for traits of their interest. A detailed summary of QTL clusters and hotspots is presented in Tables 1 and 2, showing the chromosome, number of QTL carried, and range of each cluster and hotspot identified for the intraspecific *Gh* and interspecific *Gh* × *Gb* populations, respectively. Supplementary Figures 3 and 4 investigate QTL hotspot regions for mQTL for specific traits. Supplementary Tables 3 and 4 shows a summary of the meta-analysis run by Biomercator V3 of hotspot regions with AIC values for each model. Table 3 contains a brief summary of the number of QTL, clusters, and hotspots found on the genome of each population.

QTL cluster and hotspot analysis by chromosome

Of the 1,075 QTL reported in the *Gh* population, chromosomes c14, c16, c23, and c24 carried more QTL (66–84, accounting for 6–7 % of the QTL detected), followed by

Table 1 QTL cluster and hotspot positions and QTL number in *G. hirsutum* intraspecific population

Cluster name and range	No. of QTL	Hotspot name and range	No. of QTL
c1-cluster-Gh-1: 0–20 cM	15	c1-mQTL-SI-Gh-1: 0–20 cM	4
c1-cluster-Gh-2: 28–37 cM	4	c1-mQTL-FL-Gh-1: 26–45 cM	4
c2-cluster-Gh-1: 0–20 cM	9	c3-mQTL-LP-Gh-1: 0–20 cM	4
c3-cluster-Gh-1: 0–20 cM	15	c3-mQTL-LP-Gh-2: 25–45 cM	4
c3-cluster-Gh-2: 34–55 cM	19	c5-mQTL-Mic-Gh-1: 0–20 cM	5
c4-cluster-Gh-1: 0–20 cM	8	c5-mQTL-FE-Gh-1: 0–20 cM	5
c5-cluster-Gh-1: 0–20 cM	33	c5-mQTL-SI-Gh-2: 0–20 cM	5
c5-cluster-Gh-2: 22–44 cM	17	c6-mQTL-Fus-Gh-1: 0–24 cM	4
c5-cluster-Gh-3: 57–74 cM	4	c7- mQTL -SI-Gh-3: 0–18 cM	12
c5-cluster-Gh-4: 100–122 cM	4	c7-mQTL-LP-Gh-3: 0–20 cM	7
c6-cluster-Gh-1: 0–17 cM	18	c7- mQTL -FL-Gh-2: 0–18 cM	6
c6-cluster-Gh-2: 37–44 cM	6	c7-mQTL-FS-Gh-1: 0–18 cM	7
c7-cluster-Gh-1: 0–18 cM	43	c8-mQTL-FL-Gh-3: 0–32 cM	13
c7-cluster-Gh-2: 22–41 cM	15	c8-mQTL-FS-Gh-2: 0–20 cM	5
c8-cluster-Gh-1: 0–20 cM	19	c9-mQTL-FS-Gh-3: 7–23 cM	4
c8-cluster-Gh-2: 21–31 cM	7	c10-mQTL-FL-Gh-4: 0–20 cM	4
c8-cluster-Gh-3: 32–53 cM	6	c11-mQTL-LP-Gh-4: 0–20 cM	4
c9-cluster-Gh-1: 0–20 cM	30	c11-mQTL-RN-Nem-Gh-1: 0–20 cM	7
c9-cluster-Gh-2: 26–37 cM	6	c12-mQTL-SW-Gh-1: 48–57 cM	5
c10-cluster-Gh-1: 0-20 cM	19	c13-mQTL-LP-Gh-5: 39–60 cM	12
c10-cluster-Gh-2: 31–51 cM	4	c14-mQTL-FE-Gh-2: 0-20 cM	5
c11-cluster-Gh-1: 0–20 cM	31	c14-mQTL-PH-Gh-1: 0–20 cM	4
c12-cluster-Gh-1: 0-20 cM	13	c14-mQTL-FL-Gh-5: 10-20 cM	4
c12-cluster-Gh-2: 23–31 cM	4	c14-mQTL-Mic-Gh-2: 0–28 cM	5
c12-cluster-Gh-3: 48-57 cM	8	c14-mQTL-RN-Nem-Gh-2: 0-15 cM	6
c12-cluster-Gh-4: 79–91 cM	4	c16-mQTL-FS-Gh-4: 0–20 cM	6
c13-cluster-Gh-1: 0–20 cM	10	c16-mQTL-LP-Gh-6: 0–23 cM	11
c13-cluster-Gh-2: 38–58 cM	15	c16-mQTL-VW-Gh-1: 0–23 cM	8
c13-cluster-Gh-3: 63–76 cM	5	c16-mQTL-VW-Gh-2: 30–41 cM	5
c14-cluster-Gh-1: 0-20 cM	44	c16-mQTL-VW-Gh-3: 52-70 cM	5
c14-cluster-Gh-2: 76–91 cM	5	c17-mQTL-FS-Gh-5: 0–20 cM	9
c15-cluster-Gh-1: 0–20 cM	11	c18-mQTL-FL-Gh-6: 0–20 cM	4
c15-cluster-Gh-2: 23–44 cM	11	c21-mQTL-BW-Gh-1: 0–21 cM	5
c15-cluster-Gh-3: 49–68 cM	9	c21-mQTL-FL-Gh-7: 6–26 cM	8
c16-cluster-Gh-1: 0–20 cM	41	c22-mQTL-BN-Gh-1: 4–27 cM	5
c16-cluster-Gh-2: 23–41 cM	13	c22-mQTL-SI-Gh-4: 27–50 cM	5
c16-cluster-Gh-3: 50–64 cM	14	c22-mQTL-BW-Gh-2: 21–44 cM	4
c17-cluster-Gh-1: 0–20 cM	32	c22-mQTL-PH-Gh-2: 0–33 cM	5
c18-cluster-Gh-1: 0–20 cM	18	c22-mQTL-FBNum-Gh-1: 3–27 cM	4
c18-cluster-Gh-2: 68–88 cM	6	c22-mQTL-FBNode-Gh-1: 3–21 cM	6
c19-cluster-Gh-1: 0–20 cM	18	c22-mQTL-AA-Gh-1: 0–3 cM	4
c20-cluster-Gh-1: 0–20 cM	10	c23-mQTL-VW-Gh-4: 0–21 cM	14
c20-cluster-Gh-2: 37–47 cM	4	c23-mQTL-VW-Gh-5: 23–36 cM	9
c21-cluster-Gh-1: 0–32 cM	18	c24-mQTL-LP-Gh-7: 0–20 cM	5
c21-cluster-Gh-2: 60–74 cM	5	c24-mQTL-FL-Gh-8: 0–20 cM	5
c22-cluster-Gh-1: 0–20 cM	29	c24-mQTL-Mic-Gh-3: 0–20 cM	4
c22-cluster-Gh-2: 22–42 cM	19	c24-mQTL-FE-Gh-3: 0–20 cM	5
c22-cluster-Gh-3: 53–70 cM	5	c24-mQTL-FS-Gh-6: 0–20 cM	27

Table 1 continued

Cluster name and range	No. of QTL	Hotspot name and range	No. of QTL
c23-cluster-Gh-1: 0–20 cM	32	c25-mQTL-BW-Gh-3: 6–33 cM	12
c23-cluster-Gh-2: 26–45 cM	15	c25-mQTL-FE-Gh-4: 6–33 cM	5
c23-cluster-Gh-3: 54–76 cM	9	c26-mQTL-Mic-Gh-4: 0–22 cM	4
c23-cluster-Gh-4: 77–92 cM	5		
c24-cluster-Gh-1: 0–20 cM	56		
c24-cluster-Gh-2: 41–62 cM	9		
c25-cluster-Gh-1: 0–20 cM	22		
c25-cluster-Gh-2: 21–40 cM	18		
c25-cluster-Gh-3: 45–69 cM	8		
c26-cluster-Gh-1: 0–20 cM	15		

Above are the clusters and hotspots identified in the *G. hirsutum* intraspecific population. The chromosome number, QTL trait, and position of each QTL cluster and hotspot are listed below along with the number of QTL associated with each cluster and hotspot

The trait abbreviations used in the legend above are as follows: *AA* amino acid, *BN* boll number, *BW* boll weight, *FBNode* fruiting branch node, *FBNum*, fruiting branch number; *FE* fiber elongation, *FL* fiber length, *FS* fiber strength, *Fusarium* Fusarium wilt, *LP* lint percent, *Micronaire*, micronaire/fiber fineness, *PH* plant height, *FU* fiber uniformity, *HairL* leaf hair, *HairS* stem hair, *LeafM* leaf morphology, *Color* fiber color, *PF* percent fiber, *PetCol* petal color, *PetSpot* petal spots, *PolCol* pollen color, *SI* seed index, *SW* seed weight, *VW* Verticillium wilt, *RN-Nematode* root-knot nematode

c5, c7, c22, and c25 with 52–63 QTL (4–5 %) each. Chromosomes c3, c8, c9, c10, c11, c12, c13, c15, c17, c18, c19, and c21 carried fewer QTL (31–47 each, 2–4 %). Chromosomes c1, c2, c4, c6, c20, and c26 carried the fewest (26 or fewer, 1–2.4 %). Of a total of 1,059 QTL detected in the *Gh* × *Gb* population, c19 carried most of the QTL (78, accounting for 7.3 % of the QTL detected), followed by c5, c6, and c21 with 56–62 QTL (~5 %). Chromosome c4, c10, c13, c16, and c17 also carried fewer QTL (21–30, <3 %), and c22 carried the least QTL (16, 1.5 %).

Between the two types of populations, c5, c7 and c25 carried more QTL with similar proportions of QTL (4–5 %), while c2, c4, and c20 carried least QTL (1–3 %). However, the chromosomes c14, c16, c23, and c24 with the most QTL in the intraspecific population had fewer QTL detected at lower proportions in the interspecific population. On the other hand, chromosomes c1, c6, c19, and c21 with the least QTL detected in the former carried higher proportions of QTL in the latter population. A further analysis is detailed in the following (see Supplementary Tables 1 and 2 for more details).

c1

Of the 23 and 47 QTL reported in the *Gh* and *Gh* × *Gb* populations, 2 and 3 clusters were identified, respectively. The clusters in the former were a yield and fiber quality clusters including a seed index (SI) and a fiber length (FL) QTL hotspot, while clusters in the latter were yield and fiber quality clusters including an FL QTL hotspot.

c2

In the *Gh* and *Gh* × *Gb* populations, 15 and 32 QTL have been reported, resulting in 1 (fiber quality) and 3 QTL clusters, respectively. Cluster 1 in *Gh* × *Gb* was similar to the *Gh* cluster containing QTL for FL and micronaire in the same region. However, no hotspots were identified in both populations.

c3

Gh and *Gh* × *Gb* populations carried two clusters each in similar regions from 37 and 41 QTL, respectively. Both clusters in both populations were mainly consisted of fiber quality QTL including a lint percentage (LP) QTL hotspot in each cluster of the *Gh* population, and an FL QTL hotspot in each cluster of the *Gh* × *Gb* population.

c4

The fewest QTL (14 vs. 26) have been reported in the *Gh* and *Gh* × *Gb* populations, with 1 and 2 clusters identified, respectively. The cluster in the former was composed mainly of fiber quality QTL, while the same is true for cluster 2 in the latter population with an FL QTL hotspot.

c5

More QTL were detected in *Gh* (58) and *Gh* × *Gb* (62) populations each with 3 QTL clusters. The first two clusters share similar regions in both populations for similar traits

Table 2 QTL cluster and hotspot positions and QTL number in the interspecific *G. hirsutum* × *G. barbadense* population

Cluster name and range	No. of QTL	Hotspot name and range	No. of QTL
c1-cluster-GhxGb-1: 0–20 cM	8	c1-mQTL-FL-GhxGb-1: 50–76 cM	9
c1-cluster-GhxGb-2: 45–63 cM	13	c3-mQTL-FL-GhxGb-2: 0–20 cM	6
c1-cluster-GhxGb-3: 76–95 cM	8	c3-mQTL-FL-GhxGb-3: 32–55 cM	7
c2-cluster-GhxGb-1: 0–20 cM	5	c3-mQTL-PF-GhxGb-1: 12–23 cM	4
c2-cluster-GhxGb-2: 30–50 cM	9	c4-mQTL-VW-GhxGb-1: 0–20 cM	4
c2-cluster-GhxGb-3: 57–77 cM	12	c4-mQTL-FL-GhxGb-4: 49–62 cM	5
c3-cluster-GhxGb-1: 0–20 cM	9	c5-mQTL-RN-Nem-GhxGb-1: 25–40 cM	6
c3-cluster-GhxGb-2: 40–60 cM	13	c5-mQTL-PolCol-GhxGb-1: 89–103 cM	6
c4-cluster-GhxGb-1: 0–20 cM	9	c5-mQTL-VW-GhxGb-2: 84–103 cM	4
c4-cluster-GhxGb-2: 49–67 cM	7	c6-mQTL-color-GhxGb-1: 0–20 cM	6
c5-cluster-GhxGb-1: 0–20 cM	12	c6-mQTL-HairL-GhxGb-1: 17–37 cM	7
c5-cluster-GhxGb-2: 25–45 cM	14	c6-mQTL-HairS-GhxGb-1: 17–37 cM	4
c5-cluster-GhxGb-3: 84–103 cM	18	c6-mQTL-Mic-GhxGb-1: 0–24 cM	5
c6-cluster-GhxGb-1: 0–21 cM	24	c7-mQTL-RN-Nem-GhxGb-2: 0–24 cM	9
c6-cluster-GhxGb-2: 24–37 cM	13	c7-mQTL-PetSpot-GhxGb-1: 33–55 cM	7
c6-cluster-GhxGb-3: 51–62 cM	8	c8-mQTL-color-GhxGb-2: 12–32 cM	8
c7-cluster-GhxGb-1: 0–18 cM	12	c11-mQTL-RN-Nem-GhxGb-3: 0–20 cM	10
c7-cluster-GhxGb-2: 22–32 cM	8	c12-mQTL-FL-GhxGb-5: 72–91 cM	4
c7-cluster-GhxGb-3: 38–57 cM	12	c12-mQTL-FU-GhxGb-1: 23–38 cM	5
c7-cluster-GhxGb-4: 55–79 cM	7	c12-mQTL-Mic-GhxGb-2: 0–23 cM	5
c7-cluster-GhxGb-5: 82–106 cM	4	c13-mQTL-PetCol-GhxGb-1: 17–44 cM	9
c8-cluster-GhxGb-1: 0–20 cM	4	c15-mQTL-LeafM-GhxGb-1: 6–17 cM	4
c8-cluster-GhxGb-2: 32–53 cM	22	c15-mQTL-LeafM-GhxGb-2: 21–39 cM	7
c8-cluster-GhxGb-3: 116–137 cM	4	c15-mQTL-Mic-GhxGb-3: 21–50 cM	10
c9-cluster-GhxGb-1: 0–20 cM	10	c15-mQTL-FE-GhxGb-1: 21–44 cM	4
c9-cluster-GhxGb-2: 47–64 cM	10	c16-mQTL-Mic-GhxGb-4: 0–23 cM	4
c10-cluster-GhxGb-1: 0–20 cM	8	c17-mQTL-FU-GhxGb-2: 63–78 cM	4
c10-cluster-GhxGb-2: 79–99 cM	6	c17-mQTL-Mic-GhxGb-5: 11–32 cM	5
c11-cluster-GhxGb-1: 0–20 cM	21	c18-mQTL-Mic-GhxGb-6: 33–53 cM	4
c11-cluster-GhxGb-2: 33–54 cM	7	c19-mQTL-Mic-GhxGb-7: 25–44 cM	4
c12-cluster-GhxGb-1: 0–16 cM	12	c19-mQTL-Mic-GhxGb-8: 81–106 cM	4
c12-cluster-GhxGb-2: 25–48 cM	20	c19-mQTL-FE-GhxGb-2: 32–52 cM	4
c12-cluster-GhxGb-3: 57–79 cM	5	c19-mQTL-FE-GhxGb-3: 125–144 cM	4
c12-cluster-GhxGb-4: 82–91 cM	5	c19-mQTL-FL-GhxGb-6: 40–60 cM	6
c13-cluster-GhxGb-1: 7–29 cM	10	c21-mQTL-color-GhxGb-3: 0–26 cM	4
c13-cluster-GhxGb-2: 39–60 cM	7	c21-mQTL-Mic-GhxGb-9: 60–79 cM	4
c13-cluster-GhxGb-3: 69–93 cM	4	c21-mQTL-PF-GhxGb-2: 12–32 cM	4
c14-cluster-GhxGb-1: 0–10 cM	5	c23-mQTL-FS-GhxGb-1: 70–92 cM	7
c14-cluster-GhxGb-2: 38–56 cM	8	c24-mQTL-FL-GhxGb-7: 49–70 cM	7
c14-cluster-GhxGb-3: 68–82 cM	8	c24-mQTL-Mic-GhxGb-10: 0–27 cM	5
c14-cluster-GhxGb-4: 99–118 cM	5	c25-mQTL-Mic-GhxGb-11: 0–20 cM	9
c15-cluster-GhxGb-1: 0–17 cM	5	c25-mQTL-color-GhxGb-4: 46–65 cM	6
c15-cluster-GhxGb-2: 20–44 cM	21	c25-mQTL-color-GhxGb-5: 73–89 cM	4
c15-cluster-GhxGb-3: 50–72 cM	12	c26-mQTL-FL-GhxGb-8: 0–22 cM	6
c16-cluster-GhxGb-1: 0–23 cM	9	c26-mQTL-FL-GhxGb-9: 51–69 cM	5
c16-cluster-GhxGb-2: 50–70 cM	11	c26-mQTL-Mic-GhxGb-12: 22–45 cM	5
c17-cluster-GhxGb-1: 0–9 cM	5		
c17-cluster-GhxGb-2: 19–38 cM	9		
c17-cluster-GhxGb-3: 53–70 cM	4		
c18-cluster-GhxGb-1: 19–43 cM	13		

Table 2 continued

Cluster name and range	No. of QTL	Hotspot name and range	No. of QTL
c18-cluster-GhxGb-2: 53–72 cM	8		
c18-cluster-GhxGb-3: 87–109 cM	5		
c19-cluster-GhxGb-1: 0–20 cM	15		
c19-cluster-GhxGb-2: 32–52 cM	15		
c19-cluster-GhxGb-3: 62–81 cM	10		
c19-cluster-GhxGb-4: 118–137 cM	9		
c20-cluster-GhxGb-1: 18–37 cM	13		
c20-cluster-GhxGb-2: 56–79 cM	6		
c21-cluster-GhxGb-1: 0–21 cM	7		
c21-cluster-GhxGb-2: 26–48 cM	9		
c21-cluster-GhxGb-3: 60–81 cM	14		
c21-cluster-GhxGb-4: 87–106 cM	11		
c21-cluster-GhxGb-5: 108–118 cM	6		
c22-cluster-GhxGb-1: 0–21 cM	9		
c23-cluster-GhxGb-1: 0–21 cM	9		
c23-cluster-GhxGb-2: 22–44 cM	10		
c23-cluster-GhxGb-3: 70–87 cM	6		
c23-cluster-GhxGb-4: 107–119 cM	4		
c24-cluster-GhxGb-1: 0–20 cM	9		
c24-cluster-GhxGb-2: 29–42 cM	5		
c24-cluster-GhxGb-3: 45–62 cM	12		
c25-cluster-GhxGb-1: 0–21 cM	14		
c25-cluster-GhxGb-2: 40–62 cM	9		
c25-cluster-GhxGb-3: 65–84 cM	12		
c26-cluster-GhxGb-1: 0–22 cM	6		
c26-cluster-GhxGb-2: 32–51 cM	11		

Above are the clusters and hotspots identified in the interspecific population. The chromosome number, QTL trait, and position of each QTL cluster and hotspot are listed below. The number of QTL associated with each cluster and hotspot are listed below

The trait abbreviations used in the legend above are as follows: *AA* amino acid, *BN* boll number, *BW* boll weight, *FBNode* fruiting branch node, *FBNum* fruiting branch number, *FE* fiber elongation, *FL* fiber length, *FS* fiber strength, *Fusarium* Fusarium wilt, *LP* lint percent, *Micronaire*, micronaire/fiber fineness, *PH* plant height, *FU* fiber uniformity, *HairL* leaf hair, *HairS* stem hair, *LeafM* leaf morphology, *Color* fiber color, *PF* percent fiber, *PetCol* petal color, *PetSpot* petal spots, *PolCol* pollen color, *SI* seed index, *SW* seed weight, *VW* Verticillium wilt, *RN-Nematode* root-knot nematode

in that cluster 1 could be called a cluster for seed quality including SI, a micronaire, and fiber elongation (FE) QTL hotspots in the former, amino acids, protein and oil content in addition to QTL for fiber quality traits, while cluster 2 was mainly consisted of fiber quality QTL in addition to a hotspot for root-knot nematode (RN-nematode hereafter) resistance in the latter. Cluster 3 in the *Gh* × *Gb* population carried a hotspot for pollen color which was overlapped with a hotspot for Verticillium wilt resistance (VW hereafter).

c6

In the *Gh* and *Gh* × *Gb* populations, 2 and 3 clusters were identified from 28 and 56 QTL, respectively. Cluster 1 was mainly consisted of QTL for seed quality and fiber quality in addition to a QTL hotspot for Fusarium wilt resistance (FW hereafter) in the former and a micronaire hotspot and a fiber color hotspot in the latter. Interestingly, a hotspot each for leaf and stem hairiness was also identified in the same region of the *Gh* × *Gb* population. This chromosome or region may carry a QTL or QTL cluster with a pleiotropic effect on plant trichomes (leaf and stem hairs and seed fibers).

c7

Similar QTL have been reported in *Gh* (63) and *Gh* × *Gb* (49) populations, with 2 and 5 clusters, respectively. Cluster 1 in both populations in the same region carried numerous QTL for yield and fiber quality traits including a hotspot for SI, FL, and another for fiber strength (FS), and a fourth for LP in the former and a hotspot for nematode resistance in the latter. In addition, 5 more QTL for RN-nematode resistance were identified in cluster 2 and 3, and a petal spot QTL hotspot in cluster 3 were identified in the *Gh* × *Gb* population.

c8

In the *Gh* population, 36 QTL were detected with three clusters, the first encompassing an FS QTL hotspot, while the first and second comprised an FL QTL hotspot. The first two clusters for fiber quality traits, which were in common

Table 3 QTL cluster and hotspot summary for each chromosome in each type of population

Chromosome	*G. hirsutum*			*G. hirsutum* × *G. barbadense*		
	No. of QTL	No. of clusters	No. of hotspots	No. of QTL	No. of clusters	No. of hotspots
1	23	2	2	47	3	1
2	15	1	0	32	3	0
3	37	2	2	41	2	3
4	14	1	0	26	2	2
5	58	4	3	62	3	3
6	28	2	1	56	3	4
7	63	2	4	49	5	2
8	36	3	2	41	3	1
9	47	2	1	40	2	0
10	31	2	1	21	2	0
11	35	1	2	40	2	1
12	34	4	1	51	4	3
13	37	3	1	30	3	1
14	66	2	5	42	4	0
15	37	3	0	48	3	4
16	71	3	5	27	2	1
17	36	1	1	24	3	2
18	32	2	1	42	3	1
19	34	1	0	78	4	5
20	16	2	0	29	2	0
21	32	2	2	59	5	3
22	62	3	7	16	1	0
23	69	4	2	39	4	1
24	84	2	5	36	3	2
25	52	3	2	46	3	3
26	26	1	1	37	2	3
Total	1,075	58	51	1,059	76	46

in the same region with two of the three clusters detected in the *Gh* × *Gb* population carrying 26 of 41 QTL including a fiber color hotspot.

c9

Of 47 and 40 QTL reported for the *Gh* and *Gh* × *Gb* populations, respectively, but both carried 2 clusters. The first cluster at the same position in both populations carried QTL for yield and fiber quality traits including an FS hotspot from the former. In the second cluster of *Gh* × *Gb* population, QTL for VW and nematode resistance and yield were detected.

c10

Similar number of QTL was detected in *Gh* (31) and *Gh* × *Gb* (21) populations each with 2 clusters for fiber quality traits including an FL hotspot in the former population. However, cluster 2 was located in different regions in the two populations.

c11

Of 35 and 40 QTL in *Gh* and *Gh* × *Gb* populations, 1 and 2 clusters were detected, respectively, including a hotspot for nematode resistance in the same region of cluster 1 in both populations. However, cluster 1 also carried an LP QTL hotspot in the former population which was absent in the latter. Cluster 2 in the latter population also carried 2 QTL for nematode resistance.

c12

In the *Gh* and *Gh* × *Gb* populations, 4 clusters were identified in each from 34 and 51 QTL, respectively. Cluster 1 located in the same region in both populations was consisted of QTL for yield and fiber quality traits, while cluster 2 and 3 in the *Gh* and cluster 3 and 4 in the *Gh* × *Gb* population was consisted of mainly fiber quality QTL. A QTL hotspot for seed weight (SW hereafter) was detected within cluster 3 of the *Gh* population. QTL hotspots for

fiber uniformity (FU) and FL were detected within cluster 2 and another for micronaire within cluster 1 was detected in the *Gh* × *Gb* population.

c13

The *Gh* population with 37 QTL and the *Gh* × *Gb* population with 30 QTL both carried 3 clusters. Cluster 1 at the same position in both populations was consisted of QTL for yield and fiber quality traits in addition to a petal color hotspot in the *Gh* × *Gb* population while cluster 2 in the *Gh* population carried an LP QTL hotspot.

c14

More QTL (66) in the *Gh* population were detected than in the *Gh* × *Gb* population (42). In the former, only two clusters were detected, but within the first cluster 5 QTL hotspots were detected for FE, plant height (PH hereafter), FL, micronaire, and RN-nematode resistance. In the *Gh* × *Gb* population, 4 clusters were identified, including cluster 1 carrying QTL for nematode resistance and yield, similar to the cluster in the *Gh* population and cluster 2 and 4 for fiber quality traits. No hotspots were identified.

c15

Both the *Gh* population with 37 QTL and the *Gh* × *Gb* population with 48 QTL each carried 3 clusters. Cluster 1 in similar region in both populations was mainly composed of QTL for fiber quality and yield traits. While the *Gh* population carried no QTL hotspots, leaf morphology (LM hereafter) was detected in the first and second clusters in the *Gh* × *Gb* population. Micronaire and FE QTL hotspots were also detected in the second cluster in the *Gh* × *Gb* population.

c16

More QTL (71) with 3 clusters were detected in the *Gh* population than in the *Gh* × *Gb* population (27) with 2 clusters. Cluster 1 carried a number of QTL for fiber quality traits including an FS hotspot in addition to an LP and VW hotspot in the former and a micronaire hotspot in the latter. Cluster 2 and 3 each also carried 1 VW hotspot, making a total of 15 VW QTL detected on c16 in the *Gh* population. Cluster 2 in the *Gh* × *Gb* population was in similar region to cluster 2 in the *Gh* population but carried mostly fiber and seed quality QTL.

c17

In the *Gh* population with 36 QTL 1 cluster was identified while in the *Gh* × *Gb* population 3 clusters were detected. Cluster 1 in the same region carried fiber quality QTL in both populations. An FS QTL hotspot was detected within the first cluster in the former population while a micronaire QTL hotspot was detected in the second cluster of the latter population. The *Gh* × *Gb* population also carried an FU QTL hotspot within the third cluster. Two more clusters for fiber quality traits were detected in the *Gh* × *Gb* population including a hotspot for micronaire in cluster 2 and another for FU in cluster 3.

c18

The *Gh* population with 32 QTL carried 2 clusters and the *Gh* × *Gb* population carried 42 QTL with 3 clusters. Cluster 1 in the *Gh* population carried an FL QTL hotspot, while cluster 1 in a different region in the *Gh* × *Gb* population carried a micronaire QTL hotspot. Clusters 2 and 3 in the *Gh* × *Gb* population were composed of fiber quality QTL. Cluster placement in general differed between the two populations.

c19

More than double the QTL (78) with 4 clusters were detected in the *Gh* × *Gb* population than in the *Gh* population (34) with 1 cluster. Cluster 1 in the former carried seed and fiber quality QTL while it carried yield and fiber quality QTL in the latter, in addition to QTL for FW or VW resistance. Interestingly, cluster 2 and 3 in the *Gh* × *Gb* population each carried QTL for nematode resistance and fiber quality including a hotspot for micronaire in cluster 2. Another micronaire hotspot was also identified between cluster 2 and 3. Two FE hotspots in cluster 2 and 4 and an FL hotspot in the region of cluster 2 were also identified.

c20

Fewer QTL were detected in the *Gh* (16) than in the *Gh* × *Gb* (29) population each with 2 clusters but not overlapped. Both clusters in the former were predominately composed of fiber quality QTL, while two clusters in the latter were composed of fiber and seed quality QTL. In addition, 3 QTL in cluster 2 were identified for nematode resistance in the *Gh* × *Gb* population.

c21

In the *Gh* population with 32 QTL and the *Gh* × *Gb* population with 59 QTL, 2 and 5 clusters were detected, respectively. In the *Gh* population a boll weight (BW) and FL QTL hotspot was detected within the first cluster, while in a similar region between clusters 1 and 2 in the *Gh* × *Gb* population fiber color and percent fiber (PF) QTL hotspots

were detected. A micronaire hotspot and 3 QTL for oil and protein contents were also detected in cluster 3, while cluster 4 and 5 were mainly composed of fiber quality QTL.

c22

The *Gh* population carried 62 QTL with 3 clusters while the *Gh* × *Gb* carried over three times less QTL (16). In the *Gh* population cluster 1 carried a boll number (BN), a PH, a fruiting branch number (FBNum), a fruiting branch node (FBNode), and an amino acid (AA) QTL hotspot and represents an important region of the genome pertaining to many morphological and biochemical characteristics. The second cluster carried an SI and a BW QTL hotspot and represents an important yield-related region of the genome. In the interspecific population, the cluster carried 2 VW and 1 nematode resistance QTL in addition to 4 QTL for fiber quality (2 were also detected in the *Gh* population) and 2 QTL for BW. But no hotspots were identified.

c23

In the *Gh* population, more QTL (69) were detected than in the *Gh* × *Gb* population (39), both with 4 clusters. Clusters 1 and 2 in the *Gh* population both carried VW resistance QTL hotspots, while cluster 3 in the *Gh* × *Gb* population carried an FS QTL hotspot in addition to VW QTL. Interestingly, cluster 1 and 2 in similar regions of the *Gh* × *Gb* population each carried 1 QTL for nematode resistance in addition to a number of QTL for fiber quality traits.

c24

More QTL were detected in the *Gh* population (84) with 2 clusters than in the *Gh* × *Gb* population (36) with 3 clusters. Cluster 1 in the *Gh* population carried an FS hotspot densely populated with 27 QTL, while it in the same region of the *Gh* × *Gb* population was composed of yield and fiber quality QTL including a micronaire hotspot. Cluster 1 in the *Gh* population also carried an LP, FL, micronaire, and FE QTL hotspot and represents an important fiber quality region of the genome. Cluster 2 in the *Gh* population and cluster 3 in the *Gh* × *Gb* population were in the similar region and were composed of fiber quality QTL including an FL hotspot in the *Gh* × *Gb* population.

c25

The *Gh* population carried 52 QTL while the *Gh* × *Gb* population carried similar number (46) of QTL with both populations carrying 3 clusters. Between clusters 1 and 2 in the *Gh* population a BW and an FE QTL hotspot were detected, while in the *Gh* × *Gb* population a micronaire QTL hotspot was detected in cluster 1, and a fiber color hotspot was detected in both clusters 2 and 3.

c26

Fewer QTL were detected in the *Gh* population (26) with 1 cluster than in the *Gh* × *Gb* population (37) with 2 clusters. The cluster in the *Gh* population was composed of yield and fiber quality trait QTL including a micronaire hotspot, while the cluster in the same region of *Gh* × *Gb* population carried an FL hotspot. In the *Gh* × *Gb* population, a micronaire hotspot was detected in cluster 2, while another FL hotspot was detected within the range of this cluster.

In summary, among the clusters identified in the *Gh* and *Gh* × *Gb* populations, many were composed of fiber quality QTL, while a few (e.g., on c1, c7, c9, c12, c13, c14, c19 and c26) contained QTL for both fiber quality and yield traits. The co-localization of the QTL for yield and fiber quality traits may explain close correlations between yield and fiber quality and between fiber quality traits due to pleiotropic effects of same QTL or tightly linked QTL for different traits.

QTL hotspots which we defined as having four or more QTL of the same trait within a 20 cM span were analyzed for mQTL regions. Since four or more QTL were used in the analysis the probability of declaring false-positive mQTL regions based on too few QTL was greatly reduced. The Biomercator V3 software always identified mQTL within declared hotspot regions. At times the software divided the four or more QTL into two or more possible mQTL positions. The software did not always associate all 4 QTL in the hotspot to the same region and often declared more than one mQTL within a narrow interval. As seen in Supplementary Figures 3 and 4 in the absence of multiple QTL the software will declare mQTL regions in the presence of only one QTL. For this reason we focused on previously defined hotspot regions where the probability of a false-positive mQTL region is the lowest. The mQTL regions identified within QTL hotspot regions are intended to provide MAS programs with a better idea of the most probable positions of QTL of interests and to provide MAS programs with maps showing the most concentrated regions of QTL.

It is interesting to note that, the more QTL present on the chromosome, the more QTL clusters and hotspots were detected (Table 3). In both the *Gh* and *Gh* × *Gb* populations, the numbers of QTL clusters and hotspots were significantly or highly significantly positively correlated with the number of QTL on the chromosome. The coefficients of correlation (r) were 0.469 and 0.807 in the *Gh* population, respectively, and 0.636 and 0.712 in the *Gh* × *Gb* population, respectively ($r = 0.388$ at $P < 0.05$ and $r = 0.466$ at $P < 0.01$). However, in terms of number of QTL comprising

each cluster, the *Gh* population tended to carry greater numbers of QTL within large clusters when compared to the interspecific population. For example, c24 on the first cluster carried 56 QTL and contained five hotspots within the region in the *Gh* population, while the cluster in the *Gh* × *Gb* population with the most QTL was the first cluster on c6 with 24 QTL containing two hotspots within the region. In general, most clusters which carried larger numbers of QTL than others also had more hotspots except for a few chromosomes such as c17 whose first cluster carried 32 QTL but with only one hotspot.

QTL, cluster and hotspot analysis by traits

While analysis can be done for each trait, here we will focus on the most important economic traits (e.g., fiber quality and yield) with most QTL identified.

Fiber quality QTL, cluster, and hotspot analysis

Supplementary Tables 1 and 2 contain a complete summary of the number of each QTL trait for each chromosome in both populations. In the *Gh* population, 125, 124, 75, and 75 QTL were identified for FL, FL, FE, and micronaire, respectively. Of the four traits some chromosomes carried more QTL than others and are worth describing further. For FL, c7, c8, c14, and c21 carried the most QTL with 9, 14, 9, and 10 QTL, respectively. For FS, c7, c17, c23 and c24 carried the most QTL with 9, 9, 12, and 33 QTL, respectively. For FE, c5, c14, and c24 carried the most QTL with 7, 8, and 11 QTL, respectively. For micronaire, c5, c24, and c26 carried the most QTL with 7, 9, and 6 QTL, respectively. As shown above, c24 carried more QTLs for FS, FE, and micronaire.

In the *Gh* × *Gb* population, 156, 70, 80, and 189 were identified for the four traits, respectively. For FL, c1, c3, c19, and c26 carried the most QTL with 13, 15, 17, and 16 QTL, respectively. For FS, c21 and c23 carried the most QTL with 7 and 9 QTL, respectively. For FE, c12, c15, and c19 carried the most QTL with 7, 6, and 11 QTL, respectively. For micronaire, c1, c12, c15, and c21 carried 10 QTL, while c5, c6, c19, and c25 carried 13, 12, 13, and 16 QTL, respectively. As shown above, c1 and c19 carried more QTL for both FL and micronaire; c21 carried more QTL for FS and micronaire; c12 and c15 carried more QTL for FE and micronaire; and c12 and c15 carried more QTL for FE and micronaire. Chromosomes with the largest number of QTL for the four traits differed between populations.

In the intraspecific population, c1, c7, c8, c10, c14, c18, c21, and c24 carried FL QTL hotspots; c7, c8, c9, c16, c17, and c24 carried FS QTL hotspots and most notable is c24 carrying an FS hotspot with 27 QTL which may represent a major FS locus; c5, c14, c24, and c25 carried FE QTL hotspots; and c5, c14, c24, and c26 carried micronaire QTL hotspots. Chromosomes c7, c8, and c24 carried hotspots for FL and FS, while c14 and c24 for FL, FE and micronaire. In the interspecific population, c1, c3, c4, c12, c19, c24, and c26 carried FL QTL hotspots; c23 carried an FS QTL hotspot; and two FE hotspots were identified on c19 which also carried a hotspot for FL and micronaire, and one FE hotspot on c15; and c6, c12, c15, c16, c17, c18, c19, c21, c24, c25, and c26 carried micronaire QTL hotspots. Chromosomes c12, c24 and c26 each carried an FL hotspot and a micronaire hotspot.

The two populations did not share any chromosomes in common with the same hotspots for the same traits with the excepts of c1, c24, and c26 for which c24 carried FL and micronaire QTL hotspots, c1 carried FL QTL hotspots, and c26 carried micronaire QTL hotspots in both populations. QTL hotspot types mostly differed between the two populations in that c19, c21, c24 and c26 carried QTL hotspots for different fiber quality traits; however, the placements of hotspots in general often coincided and had common anchoring markers. For example, c17 in the *Gh* population carried an FS hotspot, and in the interspecific population a micronaire hotspot between markers E7M2_181 and E32M49-M185.2 from 0 to 32 cM; c24 in the *Gh* population carried an FS hotspot and in the interspecific population a micronaire hotspot between markers CIR026 and BNL3860 from 0 to 27 cM; c25 in the *Gh* population carried an FE hotspot and in the interspecific population a micronaire hotspot between markers E4M3_268 and BNL1047a from 0 to 33 cM; and c26 in the *Gh* population carried a micronaire hotspot and in the interspecific population an FL hotspot between markers pAR807 and BNL1045a from 0 to 22 cM.

While the declaration of a QTL hotspot in this study was four or more QTL, it is worth noting regions of the genome which carried between two and three fiber strength, length, and micronaire QTL. In the *Gh* population chromosomes c1, c3, c7, c9, c10, c15 and c20 carried 2 micronaire QTL each from approximately 0–14, 34–52, 18–33, 0–13, 0–20, 53–69, and 0–19 cM, respectively. Chromosomes c3, c5, c9, c11, c12, c17, c18, c20, and c26 carried 2 FL QTL from approximately 0–20, 0–20, 0–16, 0–20, 79–91, 0–23, 72–75, 0–27, and 0–28 cM, respectively, while c15, c18, c19, and c23 carried 3 QTL all from approximately 0–20 cM. On c3, c5, c10, and c19 2 FS QTL were identified from 34–66, 35–44, 0–3, and 0–25 cM, respectively, while c11 carried 3 FS QTL from approximately 0–20 cM. In the interspecific population c2, c4, c5, c13, c14, and c24 each carried 2 micronaire from 84–93, 34–49, 83–87, 80–93, 104–109, and 0–20 cM, respectively, while c10 carried 3 micronaire QTL from approximately 0–20 cM. Chromosomes c6, c7, c8, c9, c12, c13, c20, c21, and c25 all with 2 FL QTL were identified from approximately 0–17, 87–106,

20–25, 13–26, 23–25, 0–17, 30–37, 54–64, and 0–20 cM, respectively. Chromosomes c11 and c23 each carried 3 FL QTL from approximately 0–20 and 26–35 cM, respectively. Also 2 FS QTL were identified on c11 from approximately 0–20 cM.

Yield and yield-related trait QTL

In the *Gh* population 26, 18, 51, and 97 QTL were identified for seed cotton yield (SCY), lint yield (LY), BW, and LP, respectively. While QTL were not highly localized on any given chromosome for SCY or LY, c3, c7, c13, and c16 carried 9, 9, 13, and 12 LP QTL, respectively. Chromosomes c7, c16, c20, and c22 carried 2 SCY QTL from approximately 4–18, 41–64, 0–27, and 0–16, respectively. Chromosomes c14 and c16 carried 2 LY QTL each from approximately 0–20 cM. Chromosomes c21, c22, and c25 carried BW QTL hotspots, most notably c25 carried 12 BW QTL within the hotspot. Chromosomes c14, c17, and c26 each carried 2 BW QTL from approximately 0–20, 0–32, and 0–28 cM, respectively, while c6 and c8 each carried 3 BW QTL from approximately 0–20 cM. Chromosome c3 carried 2 LP QTL hotspots, while c7, c11, c13, c16, and c24 each carried an LP QTL hotspot. Chromosomes c2, c4, c10, c19, and c22 carried 2 QTL each from approximately 0–30, 0–20, 0–9, and 0–25, 0–12 cM, respectively. Chromosomes c9, c12, c17, and c22 carried 3 QTL from approximately 6–26, 48–57, 0–20, and 22–44 cM, respectively.

In the interspecific population 21, 11, 18, and 10 QTL were identified for SCY, LY, BW, and LP, respectively, which were not highly localized, and no hotspots were identified. Chromosomes c9 and c18 both carried 3 SCY QTL, c1, c7, and c18 all carried 2 LY QTL, c18 carried 3 BW QTL, and c7, c16, and c23 all carried 2 LP QTL. No chromosome carried more than 3 QTL for any of the four traits in the *Gh* × *Gb* population. There was a considerable difference between the number and localization of QTL for these four traits between the two populations.

QTL placement analysis

When the chromosomes are divided into 20 cM regions with region 1 spanning from 0 to 20 cM, clear similarities in QTL distribution among chromosomes were observed. All chromosomes with the exception of c20 in the *Gh* × *Gb* population carried a cluster within region 1 of the chromosome. The most distal region of each chromosome (the last 20 cM) was devoid of QTL clusters or hotspots with the exception of c5 in the *Gh* population and c4, c7, c14, c17, c18, and c23 in the interspecific population. The distribution of QTL on each chromosome appears to be biased towards the more proximal end of the chromosome as relatively few clusters or hotspots were observed in the distal region. As this was consistent in both populations, it may be an evolutionary signature that QTL aggregate more on the proximal regions of chromosomes in *Gossypium*.

The distribution of QTL along the genome in both populations was examined using a Chi-square test. In the intraspecific *Gh* and interspecific populations, chromosomes were not evenly distributed over the genome ($\chi^2 = 212.092 > 37.652$ and $\chi^2 = 118.94 > 37.652$ with 25 *df* and a *P* value = 0.05, respectively). The *Gh* population showed a significant difference in QTL distribution between the A and D subgenomes with the D subgenome having significantly more QTL ($\chi^2 = 23.517 > 3.841$ with 1 *df* and a *P* value = 0.05). In the interspecific population, however, no significant difference was observed between the A and D subgenomes ($\chi^2 = 0.159 > 3.841$ with 1 *df* and a *P* value = 0.05).

In the *Gh* population, homeologous chromosomes of the A and D subgenomes did not share many similarities in terms of QTL cluster or hotspot placement or type. The exceptions to this were homeologous chromosomes c7 and c16 which both carried an FS and LP QTL hotspot within 0–20 cM and homeologous chromosomes c8 and c24 which both carried an FL and FS QTL hotspot between 0 and 30 cM. In the interspecific population, homeologous c2 and c14 each contained 3 QTL clusters which spanned a region approximately 0–77 cM, although the general size of the clusters between the two populations differed. A micronaire QTL hotspot was found on homeologous c6 and c25 in the same region and both also carried fiber color QTL hotspots although the placement differed. Homeologous c10 and c20 both carried two QTL clusters although their placements did not coincide within the same regions. Homeologous c12 and c26 both carried fiber length and micronaire QTL hotspots; however, their placements were not in the same region. Overall, homeologous chromosomes between the A and D subgenomes were not comparable in either population in terms of QTL clusters and hotspot placements.

Candidate genes in QTL hotspots

Putative candidate genes identified in this meta-analysis will require additional studies to confirm their presence or absence and roles in QTL hotspot regions. While the AD1 and AD2 genomes have not been fully sequenced, the A2 and D5 genome sequences are available. The identification of putative genes in all of the hotspots identified in this study would present an immense undertaking considering the wide confidence intervals and possible lack of exact collinearity between the tetraploid and their ancestral diploid genomes. However, a small number of hotspot regions were selected to validate the presence of putative genes in hotspot regions. In cotton a centimorgan constitutes

approximately 500,000 bps meaning that 1 cM can potentially contain multiple ORFs. On the A subgenome in the interspecific population two QTL hotspots within relatively narrow confidence intervals were selected. The pollen color QTL hotspot (which in fact resides a major Mendelian gene P1) from 89 to 103 cM which constitutes 7 Mbps of the genome and the fiber uniformity QTL hotspot on chromosome c12 from 23 to 38 cM constituting 7.5 Mbps were selected. In the pollen color hotspot region from 89 to 103 cM (44,500,000–51,500,000 bps), 172 putative gene sequences were identified. Supplementary Figure 7 shows these genes identified within the pollen color hotspot region on c5. In the fiber uniformity QTL hotspot region from 23 to 38 cM (11,500,000–19,000,000 bps) 152 putative genes were identified. Supplementary Figure 8 shows the genes identified within the fiber uniformity hotspot on c12.

On the D subgenome two QTL hotspots identified with narrow confidence intervals in the *G. hirsutum* population were a fiber length hotspot on c14 from 10 to 20 cM (5,000,000–10,000,000 bps) and an amino acid hotspot on c22 from 0 to 3 cM (0–1,500,000 bps). The putative fiber length QTL(s) on c14 are within 5 million nucleotides of the identified region, and the amino acid QTL(s) on c22 are within 1.5 million nucleotides. This does not mean that all putative genes related to those traits as multiple QTLs of other traits were also found within the regions and no doubt major genes are also contained in the regions. As the majority of QTL in both populations are found within the proximal portion of the chromosomes multiple ORFs were identified in both putative hotspot regions which differs from the two hotspot regions investigated in the A subgenome which were positioned further down on the chromosome. On chromosome c14 from 10 to 20 cM (5 Mbp) 488 putative genes sequences were identified. Supplementary Figure 9 shows the portion of chromosome c14 from 10 to 20 cM with the genes identified. On chromosome c22 from 0 to 3 cM (1.5 Mbp) 382 putative genes were identified. Supplementary Figure 10 shows the portion of chromosome c22 from 0 to 3 cM with the genes identified. These putative genes identified represent candidate genes for the QTL hotspots identified in this study.

Discussion

The ratio of QTL between the *G. hirsutum* (1,075 QTL) and interspecific population (1,059 QTL) was proportional, but our results illustrate clear differences and similarities in QTL cluster and hotspot placement. Therefore, it is interesting to note that some chromosomes were more densely populated with QTL, clusters, hotspots, and/or hotspot clusters in one population than in the other. In terms of hotspot trait type there were largely no similarities between homologous chromosomes between the two populations or between the homeologous chromosomes from the A and D subgenomes. While the QTL hotspot type usually differed between populations, the regional placement of hotspots on the chromosome was often the same. QTL hotspots and important mQTL regions identified by the software provide valuable maps and insight into commercially interesting regions of the genome. Marker-assisted breeding programs can utilize these differences to select for markers which flank the most desirable trait QTL between populations. Hotspot clusters identified in both populations on different chromosomes will also aid marker-assisted breeding programs in the selection of desirable QTL aggregates which may be absent in one population or the other, therefore, increasing genetic gains. Hotspot clusters contained multiple mQTL and should be considered important for MAS programs. The [“Guazuncho2” (*G. hirsutum*) × “VH8-4602” (*G. barbadense*)] map with common anchoring markers which flank the largest number of desirable QTL can be used to improve marker-assisted breeding efforts. The comparison map between the two populations can also be used to identify possible genetic differences which exist between two populations. Flanking markers which surround such regions can be the focus of breeding programs attempting to increase gains between interspecific crosses. While not declared as hotspots many regions carried two or three fiber strength, length, micronaire or yield QTL which can be also used in MAS to improve commercially valuable traits. In the meta-analysis by Rong et al. (2007), three QTL were used as the basis for declaring a cluster (with a false-positive rate of 12.5 %); however, in this study the number was increased to at least four to reduce the false-positive rate at or below 6.25 %. Since fiber quality-related QTL composed the majority of QTL in the study and almost saturated the genome, it is unclear if small aggregates of fiber quality QTL found between populations on homologous chromosomes are a result of conservation or not.

Genetically it is interesting that the majority of QTL were localized around the first 20 cM of every chromosome with a few exceptions. The fact that this was observed in both populations may be an evolutionary signature remaining from the common ancestor of *G. hirsutum* and *G. barbadense*. Tightly clustered QTL within a narrow range may indicate pleiotropic-effect QTL responsible for multiple traits. While the placement of QTL in the first region seems to have been evolutionarily conserved, QTL trait type distribution was not. Of the hotspots identified except for the nematode resistance hotspot on c11, and the FL and micronaire QTL hotspots shared on c24, no other hotspots were shared by both populations. Since the distribution of QTL was fairly even (1,075 vs. 1,059) between the two populations, it is interesting to note that there was no significant

difference between the QTL distribution between the A and D subgenome in the interspecific population. In the *Gh* population, however, not only were there significantly more QTL detected on the D subgenome, but resistance QTL for nematode and Verticillium wilt were highly localized on two chromosomes each. This shows a level of QTL organization and localization which is not observed in the interspecific population. Future directions of study may investigate the intraspecific *G. barbadense* population and compare QTL localization with the intraspecific *G. hirsutum* population in this study. It is possible that in the interspecific cross much of the QTL localization observed within *G. hirsutum* or perhaps within *G. barbadense* is disrupted due to cryptic structural differences in chromosomes between the two species (Stephens 1949).

The identification of putative genes in regions of the A2 and D5 genomes where hotspots were identified in the intraspecific *Gh* and interspecific *Gh* × *Gb* populations was revealing. It is difficult to speculate at this time which of the putative genes in the regions are related directly to the trait QTL hotspots identified; however, it is interesting that more putative genes were identified in the first 20 cM regions of the chromosomes than beyond. This further lends confidence to the notion that in *Gossypium* the majority of QTLs and other coding regions are more concentrated in the proximal regions of chromosomes than other regions. On chromosome c22 in the narrow span of 3 cM 382 putative genes were identified, whereas beyond the first 20 cM of the chromosomes within the span of 89–103 on c5 and 23–38 cM on c12, only 172 and 152 putative genes were identified, respectively. Chromosome c14 also lends confidence to this observation having 488 putative genes identified from 10 to 20 cM. Once the tetraploid cotton genomes, from which mapping and breeding populations are made, sequenced, candidate genes within the QTL clusters and hotspots can be more reliably identified for further studies including high resolution mapping of QTL, candidate gene cloning and functional analysis.

If QTL type and placement of *G. barbadense* were similar to that of *G. hirsutum*, the interspecific cross should produce a hybrid with similar QTL placement as homologous chromosomes during recombination would share the same or similar QTL. However, in this study we see a complete lack of homology in QTL type and placement which shows a rapid QTL divergence between the *G. hirsutum* and *G. barbadense* populations from their common ancestor. This lack of conservation between the two types of genetic populations is revealing as it would appear that QTL have moved positions extensively throughout the genome since their origin from a common ancestor within the last 1–2 million years (Wendel and Cronn 2002). This indicates that the position of some QTL is not under strict purifying selection. It will be revealing in future studies which have the sequenced genomes of both species to examine how much the two populations have diverged.

In the *Gh* population RN-nematode (c11 and c14) and Verticillium wilt (c16 and c23) resistance QTL were highly localized. However, in the interspecific population they appear to be relatively distributed throughout the genome and largely delocalized. This indicates that some features such as resistance-related QTL were under selection in terms of localization in one population, but not the other. In the *Gh* population yield and fiber quality QTL were often co-localized in the same region and overlapping hotspots between the two trait types were observed. For example, c7 carried an LP, FL, and FS QTL hotspot which overlapped, while c16 carried an LP and FS hotspot which also overlapped. Chromosomes c21 carried a BW and FL hotspot which overlapped as did c25 with a BW and FE QTL hotspot. Most notable was c24 which carried an LP, FL, FS, FE, and micronaire QTL hotspot all within the same region. These yield and fiber quality QTL hotspot clusters may help to explain the close correlations seen between yield and fiber quality as they are linked at the same loci. Due to hybrid breakdown in early segregating populations in the interspecific *Gh* x *Gb* population, yield traits have not been extensively studied. The relatively few and delocalized yield trait QTL makes it difficult to speculate if linkages exist between possible aggregates of yield and fiber quality QTL.

In this study, regions of the genome containing more QTL hotspot clusters pertaining to yield, fiber quality, and disease resistance were identified. These regions can be used in MAS to increase genetic gains when working with intraspecific or interspecific populations. In the *Gh* population, c24 had the most prominent hotspot cluster carrying yield (LP) and fiber quality (FL, FE, and FS) QTL hotspots between markers CIR026 (0 cM) and NAU2407b (19.57 cM). Chromosome c7 carried a yield (SI and LP) and fiber quality (FL and FS) QTL hotspot cluster between the flanking markers E1M7_80 (0 cM) and CG05a (22.31 cM). A hotspot cluster was identified on c14 carrying fiber quality (FE, FL, and micronaire) and morphological (PH) QTL hotspots between markers CIR246 (0 cM) and G1012 (20.20 cM). On chromosome c16 the region between E5M4_480 (0 cM) and pAR544 (23.97 cM) represents a hotspot cluster carrying yield (LP), fiber quality (FS), and resistance (VW) QTL hotspots. In the *Gh* × *Gb* population, QTL were much less localized and QTL hotspots were less frequent and did not contain yield, fiber quality, and resistance hotspots together within a region that could be used in MAS breeding programs. The largest QTL hotspot in this population was on c15 between markers E7M5_121 (20.56 cM) and E1M1_174 (44.36) which carried FE, LM, and micronaire hotspots. The fairly even distribution of QTL in the *Gh* × *Gb* population

makes it less ideal for MAS programs when compared to *Gh* as the population has large aggregates of various hotspot trait types to select from. The hotspot cluster regions with anchoring markers mentioned from the *Gh* population above are ideal for introgression into populations which lack those QTL.

This study provides a useful resource for breeders and geneticists to select chromosomal regions of their interest for further studies. For example, a breeder can choose some of the chromosomal regions with QTL hotspots for fiber length and select markers in the regions for a MAS scheme. A geneticist can further define these regions and identify putative genes for a further analysis. To facilitate the selection process and to allow addition of more QTL studies to the current meta-analysis, an open database from this study will be established and shared with the cotton research community.

Author contribution JFZ conceived the research; JIS conducted data analysis and draft the manuscript; HTW, ZXL and XLZ provided results published from China; JFZ, MS and DDF contributed to the result synthesis and edited the manuscript. All authors read the manuscript.

Conflict of interest The authors declare that they have no conflict of interest.

References

An CF, Jenkins JN, Wu JX, Guo YF, McCarty JC (2010) Use of fiber and fuzz mutants to detect QTL for yield components, seed, and fiber traits of upland cotton. Euphytica 172:21–34

Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J (2004) BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. Bioinformatics 20:2324–2326

Blenda A, Scheffler J, Sheffler B, Palmer M, Lacape JM, Yu JZ, Jesudurai C, Jung S, Muthukumar S, Yellambalase P, Ficklin S et al (2006) CMD: a cotton microsatellite database resource for *Gossypium* genomics. BMC Genom 7:132

Chee P, Draye X, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH (2005a) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: I Fiber elongation. Theor Appl Genet 111:757–763

Chee PW, Draye X, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH (2005b) Molecular dissection of phenotypic variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: III. Fiber length. Theor Appl Genet 111:772–781

Chen L, Zhang ZS, Hu MC, Wang W, Zhang J, Liu DJ, Zheng J, Zheng FM, Ma J (2008) Genetic linkage map construction and QTL mapping for yield and fiber quality in Upland cotton (*Gossypium hirsutum L.*). Acta Agron Sin 34:1199–1205

Chen H, Qian N, Guo WZ, Song QP, Li BC, Deng FJ, Dong CG, Zhang TZ (2009) Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro. D8 in Upland cotton. Theor Appl Genet 119:605–612

Draye X, Chee P, Jiang CX, Decanini L, Delmonte TA, Bredhauer R, Smith CW, Paterson AH (2005) Molecular dissection of interspecific variation between *Gossypium hirsutum* and *G barbadense* (cotton) by a backcross-self approach: II. Fiber fineness. Theor Appl Genet 111:764–771

Fang H, Zhou H, Sanogo S, Flynn R, Percy RG, Hughs SE, Ulloa M, Jones DC, Zhang JF (2013) Quantitative trait locus mapping for verticillium wilt resistance in a backcross inbred line population of cotton (*Gossypium hirsutum* × *Gossypium barbadense*) based on RGA-AFLP analysis. Euphytica 194:79–91

Fang H, Zhou H, Sanogo S, Lipka AE, Fang DD, Percy RG, Hughs SE, Jones DC, Gore MA, Zhang JF (2014) Quantitative trait locus analysis of Verticillium wilt resistance in an introgressed recombinant inbred population of Upland cotton. Mol Breed 33:709–720

Fryxell PA (1992) A revised taxonomic interpretation of *Gossypium* L. (*Malvaceae*). Rheedea 2:108–165

Ge HY, Wang YC, Guo WZ, Zhang TZ (2008) Inheritance and molecular tagging of resistance against *Verticillium* wilt in upland cotton. Cotton Sci 20:19–22

Gore MA, Fang DD, Poland JA, Zhang JF, Percy RG, Cantrell RG, Thyssen G, Lipka AE (2014) Linkage map construction and quantitative trait locus analysis of agronomic and fiber quality traits in cotton. Plant Genome 7:1–10. doi:10.3835/plantgenome2013.07.0023

Guo YF, McCarty JC, Jenkins JN, Saha S (2008) QTLs for node of first fruiting branch in a cross of an upland cotton, *Gossypium hirsutum* L., cultivar with primitive accession Texas 701. Euphytica 163:113–122

Gutierrez OA, Jenkins JN, McCarty JC, Wubben MJ, Hayes RW, Callahan FE (2010) SSR markers closely associated with genes for resistance to root-knot nematode on chromosomes 11 and 14 of Upland cotton. Theor Appl Genet 121:1323–1337

Gutierrez OA, Robinson AF, Jenkins JN, McCarty JC, Wubben MJ, Callahan FE, Nichols RL (2011) Identification of QTL regions and SSR markers associated with resistance to reniform nematode in *Gossypium barbadense* L. accession GB713. Theor Appl Genet 122:271–280

Hu WJ, Zhang XY, Zhang TZ, Guo WZ (2008) Molecular tagging and source analysis of QTL for elite fiber quality in Upland Cotton. Acta Agron Sin 34:578–586

Jia F, Sun FD, Li JW, Liu AY, Shi YZ, Gong JW, Shang HH, Gong WK, Wang T, Liu Z, Yuan YL (2011) Identification of QTL for boll weight and lint percentage of Upland cotton (*Gossypium hirsutum* L.) RIL population in multiple environments. Mol Plant Breed 9:318–326

Jiang CX, Wright RJ, El-Zik KM, Paterson AH (1998) Polyploid formation created unique avenues for response to selection in Gossypium (cotton). Proc Natl Acad Sci USA 95:4419–4424

Jiang C, Wright RJ, Woo SS, DelMonte TA, Paterson AH (2000) QTL analysis of leaf morphology in tetraploid Gossypium (cotton). Theor Appl Genet 100:409–418

Jiang F, Zhao J, Zhou L, Guo WZ, Zhang TZ (2009) Molecular mapping of Verticillium wilt resistance QTL clustered on chromosomes D7 and D9 in Upland cotton. Sci China C Life Sci 52:872–884

Kong FJ, Li JW, Gong JW, Shi YZ, Liu AY, Shang HH, Gong WK, Wang T, Yuan YL (2011) QTL mapping for lint percentage and seed index in Upland cotton (*Gossypium hirsutum* L.) of different genetic backgrounds. Chin Agric Sci Bull 27:104–109

Kumar P, Singh R, Lubbers EL, Shen X, Paterson AH, Campbell BT, Jones DC, Chee PW (2012) Mapping and validation of fiber strength quantitative trait loci on chromosome 24 in upland cotton. Crop Sci 52:1115–1122

Lacape JM, Nguyen TB, Courtois B, Belot JL, Giband M, Gourlot JP, Gawryziak G, Roques S, Hau B (2005) QTL analysis of cotton fiber quality using multiple *Gossypium hirsutum* × *Gossypium barbadense* backcross generations. Crop Sci 45:123–140

Lacape JM, Jacobs J, Arioli T, Derijcker R, Forestier-Chiron N, Llewellyn D, Jean J, Thomas E, Viot C (2009) A new interspecific, *Gossypium hirsutum* × *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. Theor Appl Genet 119:281–292

Lacape JM, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, Al-Ghazi Y, Liu SM, Palai O, Georges S et al (2010) Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* × *G. barbadense* RIL population. BMC Plant Biol 10:132

Lacape JM, Gawrysiak G, Cao TV, Viot C, Liewellyn D, Liu S, Jacobs J, Becker D, Barroso PAV et al (2013) Mapping QTLs for traits related to phenology, morphology and yield components in an inter-specific *Gossypium hirsutum* × *G. barbadense* cotton RIL population. Field Crops Res 144:256–267

Li PB, Cao ML, Liu HM, Yang LL, Chen G (2006) Genetic map construction and QTL mapping of fibre quality in Upland cotton (*Gossypium hirsutum* L.). Acta Bot Boreall Occidentalia Sin 26:1098–1104

Li CQ, Guo WZ, Ma XL, Zhang TZ (2008) Tagging and mapping of QTL for yield and its components in upland cotton (*Gossypium hirsutum* L.) population with varied lint percentage. Cotton Sci 20:163–169

Li K, Yang DG, Ma XF, Wang HF, Zhou XJ, Meng QQ, Pei XY (2010) QTL mapping of yield traits in high advantage combination of cotton. Mol Plant Breed 8:673–679

Li CQ, Wang CB, Dong N, Wang XY, Zhao HH, Converse R, Xia Z, Wang R, Wang QL (2012) QTL detection for node of first fruiting branch and its height in Upland cotton (*Gossypium hirsutum* L.). Euphytica 188:441–451

Liang Q, Hu C, Hua H, Li Z, Hua J (2013) Construction of a linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Chin Sci Bull 58:3233–3243

Liu DJ, Zhang J, Zhang K, Wang W, Zhang ZS (2010) QTL mapping of seed physical traits in upland cotton (*Gossypium hirsutum* L.). Acta Agron Sin 36:53–60

Liu HY, Quampah A, Chen JH, Li JR, Huang ZR, He QL, Shi CH, Zhu SJ (2012a) QTL analysis for gossypol and protein contents in upland cottonseeds with two different genetic systems across environments. Euphytica 188:453–463

Liu RZ, Wang BH, Guo WZ, Qin YS, Wang LG, Zhang YM, Zhang TZ (2012b) Quantitative trait loci mapping for yield and its components by using two immortalized populations of a heterotic hybrid in *Gossypium hirsutum* L. Mol Breed 29:297–311

Liu H, Quampah A, Chen J, Li J, Huang Z, He Q, Zhu S, Shi C (2013a) QTL mapping based on different genetic systems for essential amino acid contents in cottonseeds in different environments. PLoS ONE 8:e57531

Liu R, Ai N, Zhu X, Liu F, Guo W, Zhang T (2013b) Genetic analysis of plant height using two immortalized populations of "CRI12 × J8891" in *Gossypium hirsutum* L. Euphytica 196:51–60

Liu XF, Li JW, Yu XK, Shi YZ, Jia F, Sun DF, Liu AY, Gong JW, Shang HH, Gong WK, Wang T, Deng HB, Yuan YL (2013c) Identification of QTL for cottonseed oil and protein content in upland cotton (*Gossypium hirsutum* L.) based on a RIL population. Mol Plant Breed 11:520–528

Lopez-Lavalle LAB, Gillespie VJ, Tate WA, Ellis MH, Stiller WN, Liewellyn DL, Wilson IW (2012) Molecular mapping of a new source of Fusarium wilt resistance in tetraploid cotton (*Gossypium hirsutum* L.). Mol Breed 30:1181–1191

Mauricio R (2001) Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. Nat Rev Genet 2:370–381

Mei M, Syed NH, Gao W, Thaxton PM, Smith CW, Stelly DM, Chen ZJ (2004) Genetic mapping and QTL analysis of fiber-related traits in cotton (Gossypium). Theor Appl Genet 108:280–291

Mei H, Ai N, Zhang X, Ning Z, Zhang T (2014) QTLs conferring FOV 7 resistance detected by linkage and association mapping in upland cotton. Euphytica 197:237–249

Miles C, Wayne M (2008) Quantitative trait locus (QTL) analysis. Nat Educ 1:1

Ning Z, Chen H, Mei H, Zhang T (2014) Molecular tagging of QTLs for fiber quality and yield in the upland cotton cultivar Acala-Prema. Euphytica 195:143–156

Nusurat O, Yu SX, Fan SL, Pang CY (2012) Correlation of agronomic characters and QTL mapping in mechanical harvest cotton (*Gossypum hirsutum* L.). Xinjiang Agric Sci 9:791–795

Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ (2003) QTL analysis of genotype × environment interactions affecting cotton fiber quality. Theor Appl Genet 106:384–396

Percival AE, Wendel JE, Stewart JM (1999) Taxonomy and germplasm resources. In: Smith CW, Cothren JT (eds) Cotton: origin, history, technology, and production. Wiley, New York, pp 33–63

Qin HD, Guo WZ, Zhang YM, Zhang TZ (2008) QTL mapping of yield and fiber traits based on a four-way cross population in *Gossypium hirsutum* L. Theor Appl Genet 117:883–894

Qin YS, Liu RZ, Mei HX, Zhang TZ, GUO WZ (2009) QTL mapping for yield traits in Upland Cotton (*Gossypium hirsutum* L.). Acta Agron Sin 5:1812–1821

Romano GB, Sacks EJ, Stetina SR, Robinson AF, Fang DD, Gutierrez OA, Scheffler JA (2009) Identification and genomic location of a reinform nematode (*Rotylenchulus reniformis*) resistance locus (Ren^{ari}) introgressed from *Gossypium aridum* into upland cotton (*G. hirsutum*). Theor Appl Genet 120:139–150

Rong J, Feltus EA, Waghmare VN, Pierce GJ, Chee PW, Draye X, Saranga Y, Wright RJ, Wilkins TA, May OL et al (2007) Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics 176:2577–2588

Said JI, Lin Z, Zhang X, Song M, Zhang J (2013) A comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genom 14:776

Saranga Y, Jiang CX, Wright RJ, Yakir D, Paterson AH (2004) Genetic dissection of cotton physiological responses to arid conditions and their inter-relationships with productivity. Plant Cell Environ 27:263–277

Shen XL, Guo WZ, Zhu XF, Yuan YL, Yu JZ, Kohel RJ, Zhang TZ (2005) Molecular mapping of QTLs for fiber qualities in three diverse lines in upland cotton using SSR markers. Mol Breed 15:169–181

Shen X, Van Becelaere G, Kumar P, Davis RF, May OL, Chee P (2006a) QTL mapping for resistance to root-knot nematodes in the M-120 RNR upland cotton line (*Gossypium hirsutum* L.) of the Auburn 623 RNR source. Theor Appl Genet 11:1539–1549

Shen XL, Zhang TZ, Guo WZ, Zhu XF, Zhang XY (2006b) Mapping fiber and yield QTLs with main, epistatic, and QTL X environment interaction effects in recombinant inbred lines of upland cotton. Crop Sci 46:61–66

Shen X, He Y, Lubbers EL, Davis RF, Nichols RL, Chee PW (2010) Fine mapping QMi-C11 a major QTL controlling root-knot nematodes resistance in upland cotton. Theor Appl Genet 121:1623–1631

SnapGene® software (from GSL Biotech; www.snapgene.com)

Stephens SG (1949) The cytogenetics of speciation in Gossypium. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 32:627–637

Su C, Wang W, Qui X, Yang L, Li S, Wang M, Pan Q (2013) Fine-mapping a fiber strength QTL QFS-D11-1 on cotton chromosome 21 using introgressed lines. Plant Breed 132:725–730

Sun FD, Zhang JH, Wang SF, Gong WK, Shi YZ, Liu AY, Li JW, Gong JW, Shang HH, Yuan YL (2012) QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Mol Breed 30:569–582

Ulloa M, Wang C, Roberts PA (2009) Gene action analysis by inheritance and quantitative trait loci mapping of resistance to root-knot nematodes in cotton. Plant Breed 129:541–550

Ulloa M, Wang C, Hutmacher RB, Wright SD, Davis RM, Saski CA, Robers PA (2011) Mapping Fusarium wilt race 1 resistance genes in cotton by inheritance, QTL and sequencing composition. Mol Genet Genomics 286:21–36

Ulloa M, Hutmacher RB, Robers PA, Wright SD, Nichols RL, Davis RM (2013) Inheritance and QTL mapping of Fusarium wilt race 4 resistance in cotton. Theor Appl Genet 126:1405–1418

Wang BH, Guo WZ, Zhu XF, Wu YT, Huang NT, Zhang TZ (2006) QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica 152:367–378

Wang FR, Liu RZ, Wang LM, Zhang CY, Liu GD, Liu QH, Ma XB, Zhang J (2007a) Molecular markers of Verticillium wilt resistance in Upland Cotton(*Gossypium hirsutum* L.) cultivar and their effects on assisted phenotypic selection. Cotton Sci 19:424–430

Wang J, Guo WZ, Zhang TZ (2007b) QTL mapping for fiber quality properties in cotton cultivar Yumian 1. Acta Agron Sin 33:1915–1921

Wang PZ, Qin L, Su L, Hu BM, Zhang TZ (2007c) QTL Mapping of some plant morphological traits for main Upland Cotton cultivars planted in Xinjiang. Mol Plant Breed 5:811–818

Wang HM, Lin ZX, Zhang XL, Chen W, Guo XP, Nie YC, Li YH (2008) Mapping and quantitative trait loci analysis of Verticillium wilt resistance genes in cotton. J Integr Plant Biol 50:174–182

Wang PZ, Su L, Qin L, Hu BM, Guo WZ, Zhang TZ (2009) Identification and molecular mapping of a Fusarium wilt resistant gene in upland cotton. Theor Appl Genet 119:733–739

Wang YQ, Li JW, Shi YZ, Liu AY, Shang HH, Gong JW, Wang T, Gong WK, Yuan YL (2010) Molecular marker of QTL for fiber quality traits in upland cotton with elite fiber quality. Cotton Sci 22:533–538

Wang XK, Pan ZE, Sun JL, He SP, Tang CM, Du XM (2011) Correlation analysis and QTL mapping for plant height and fiber quality of dwarf mutant in Upland cotton (*Gossypium hirsutum* L.). J Nuclear Agric Sci 25:448–455

Wang C, Ulloa M, Mullens TR, Yu JZ, Robers PA (2012a) QTL analysis for transgressive resistance to root-knot nematode in interspecific cotton (Gossypium spp.) progeny derived from susceptible parents. PLoS ONE 7:e34874

Wang L, Liu F, Li SH, Wang CY, Zhang XD, Wang YH, Hua JP, Wang KB (2012b) QTL mapping for fiber quality properties in Lumianyan 15. Cotton Sci 24:97–105

Wang P, Zhu YJ, Song XL, Cao ZB, Ding YZ, Liu BL, Zhu XF, Wang S, Guo WZ, Zhang TZ (2012c) Inheritance of long staple fiber quality traits of *Gossypium barbadense* in *G. hirsutum* background using CSILs. Theor Appl Genet 124:1415–1428

Wang F, Zhenzhen X, Sun R, Gong Y, Liu G, Zhang J, Wang L, Zhang C, Fan S, Zhang J (2013) Genetic dissection of the introgressive genomic components from *Gossypium barbadense* L. that contribute to improved fiber quality in *Gossypium hirsutum* L. Mol Breed 32:1–16

Wendel JF, Cronn RC (2002) Polyploidy and the evolutionary history of cotton. Adv Agron 78:139–186

Wright RJ, Thaxton PM, El-Zik KM, Paterson AH (1998) D-subgenome bias of Xcm resistance genes in tetraploid Gossypium (Cotton) suggests that polyploid formation has created novel avenues for evolution. Genetics 149:1987–1996

Wright RJ, Thaxton PM, El-Zik KH, Paterson AH (1999) Molecular mapping of genes affecting pubescence of cotton. J Hered 90:215–219

Wu JX, Gutierrez OA, Jenkins JN, McCarty JC, Zhu J (2009) Quantitative analysis and QTL mapping for agronomic and fiber traits in an RI population of upland cotton. Euphytica 165:231–245

Xu P, Guo TT, Zhang XG, Ni WC, Shen XL (2010) QTL mapping of seed oil content in Upland cotton (*Gossypium hirsutum* L.). China Cotton Society 2010 annual meeting proceedings

Yang C, Guo WZ, Zhang TZ (2007) QTL mapping for resistance to verticillium wilt, fiber quality and yield traits in upland cotton (Gossypium hirsutum L.). Mol Plant Breed 5:797–805

Yang C, Guo WZ, Li GY, Gao F, Lin SS, Zhang TZ (2008) QTLs mapping for verticillium wilt resistance at seedling and maturity stages in *Gossypium barbadense* L. Plant Sci 174:290–298

Yang XL, Wang ZW, Zhang GY, Pan YX, Wu LQ, Li ZK, Wang XF, Ma ZY (2009) Construction of molecular genetic map and QTL analysis of fiber quality in cotton. Acta Agron Sin 35:2159–2166

Yao JB, Zhang YS, Chen W, Luan MB, Feng ZD, Guo XM (2010) Tagging QTLs of yield-related traits in chromosome 22sh of allotetraploid cotton using substitution line. Cotton Sci 22:521–526

Yin JM, Wu YT, Zhang J, Zhang TZ, Guo WZ, Zhu XF (2002) Tagging and mapping of QTLs controlling lint yield and yield components in upland cotton (*Gossypium hirsutum* L.) using SSR and RAPD markers. Chin J Biotech 18:162–166

Yu JW, Yu SX, Fan SL, Song MZ, Zhai HH, Li XL, Zhang JF (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, Jung S, Cheng CH, Ficklin SP, Lee T, Zheng P, Jones D, Percy R, Main D (2013a) CottonGen: a genomics, genetics and breeding database for cotton research. Nucleic Acids Res 42:1229–1236

Yu JW, Yu S, Gore M, Wu M, Zhai H, Li X, Fan S, Song M, Zhang JF (2013b) Identification of quantitative trait loci across interspecific F2, F2:3 and testcross populations for agronomic and fiber traits in tetraploid cotton. Euphytica 191:375–389

Yu JW, Zhang K, Li S, Yu S, Zhai H, Wu M, Li X, Fan S, Song M, Yang D, Li Y, Zhang JF (2013c) 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 PT, Zhu XF, Guo WZ, Zhang TZ (2005) Genetic analysis and OTLs tagging of lint percentage and its closely related yield components in upland Cotton. Jiangsu J Agric Sci 21:264–271

Zhang PT, Guo WZ, Zhu XF, Yu JZ, Zhang TZ (2006) Molecular tagging of QTLs and its components of *G. hirsutum*. cv. Simian3. Acta Agron Sin 32:1197–1203

Zhang ZS, Hu MC, Zhang J, Liu DJ, Zheng J, Zhang K, Wang W, Wan Q (2009) Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Mol Breed 24:49–61

Zhang J, Chen X, Zhang K, Liu DJ, Wei XQ, Zhang ZS (2010) QTL mapping of yield traits with composite cross population in Upland cotton (*Gossypium hirsutum* L.). J Agric Biotech 18:476–481

Zhang J, Liu DJ, Lin G, Zhang ZS (2011a) QTL mapping for chlorophyll content in upland cotton (*Gossypium hirsutum* L.). J Southwest Univ 33(4):1–4

Zhang J, Ma J, Chen X, Liu DJ, Zhang ZS (2011b) QTL mapping of fiber quality traits with a composite cross population in Upland cotton (*Gossypium hirsutum* L.). J Agric Biotech 19:230–235

Zhang W, Liu F, Li SH, Wang W, Wang CY, Zhang XD, Wang YH, Song GL, Wang KB (2011c) QTL analysis on yield and its components in upland cotton RIL. Acta Agron Sin 37:433–442

Zhang K, Zhang J, Ma J, Tang SY, Liu DJ, Teng ZH, Liu DX, Zhang ZS (2012) Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (*Gossypium hirsutum* L.). Mol Breed 29:335–348

Zhang T, Qian N, Zhu X, Chen H, Wang S, Mei H, Zhang Y (2013) Variations and transmission of QTL alleles for yield and fiber qualities in Upland cotton cultivars developed in China. PLoS ONE 8:e57220

Zhang J, Percy RG, McCarty JC Jr (2014) Introgression genetics and breeding between Upland and Pima cotton: a review. Euphytica 198:1–12

Zhao Y, Wang H, Chen W, Li Y (2014) Genetic structure, linkage disequilibrium and association mapping of Verticillium wilt resistance in elite cotton (*Gossypium hirsutum* L.) germplasm population. PLoS ONE 9:e86308