BMC Genomics

RESEARCH ARTICLE　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Identification of stable quantitative trait loci (QTLs) for fiber quality traits across multiple environments in *Gossypium hirsutum* recombinant inbred line population

Muhammad Jamshed[1†], Fei Jia[1†], Juwu Gong[1,2†], Koffi Kibalou Palanga[1†], Yuzhen Shi[1], Junwen Li[1], Haihong Shang[1], Aiying Liu[1], Tingting Chen[1], Zhen Zhang[1], Juan Cai[1], Qun Ge[1], Zhi Liu[3], Quanwei Lu[4], Xiaoying Deng[1], Yunna Tan[1], Harun or Rashid[1], Zareen Sarfraz[1], Murtaza Hassan[5], Wankui Gong[1*] and Youlu Yuan[1*]

Abstract

Background: The identification of quantitative trait loci (QTLs) that are stable and consistent across multiple environments and populations plays an essential role in marker-assisted selection (MAS). In the present study, we used 28,861 simple sequence repeat (SSR) markers, which included 12,560 *Gossypium raimondii* (D genome) sequence-based SSR markers to identify polymorphism between two upland cotton strains 0–153 and sGK9708. A total of 851 polymorphic primers were finally selected and used to genotype 196 recombinant inbred lines (RIL) derived from a cross between 0 and 153 and sGK9708 and used to construct a linkage map. The RIL population was evaluated for fiber quality traits in six locations in China for five years. Stable QTLs identified in this intraspecific cross could be used in future cotton breeding program and with fewer obstacles.

Results: The map covered a distance of 4,110 cM, which represents about 93.2 % of the upland cotton genome, and with an average distance of 5.2 cM between adjacent markers. We identified 165 QTLs for fiber quality traits, of which 47 QTLs were determined to be stable across multiple environments. Most of these QTLs aggregated into clusters with two or more traits. A total of 30 QTL clusters were identified which consisted of 103 QTLs. Sixteen clusters in the A_t sub-genome comprised 44 QTLs, whereas 14 clusters in the D_t sub-genome that included 59 QTLs for fiber quality were identified. Four chromosomes, including chromosome 4 (c4), c7, c14, and c25 were rich in clusters harboring 5, 4, 5, and 6 clusters respectively. A meta-analysis was performed using Biomercator V4.2 to integrate QTLs from 11 environmental datasets on the RIL populations of the above mentioned parents and previous QTL reports. Among the 165 identified QTLs, 90 were identified as common QTLs, whereas the remaining 75 QTLs were determined to be novel QTLs. The broad sense heritability estimates of fiber quality traits were high for fiber length (0.93), fiber strength (0.92), fiber micronaire (0.85), and fiber uniformity (0.80), but low for fiber elongation (0.27). Meta-clusters on c4, c7, c14 and c25 were identified as stable QTL clusters and were considered more valuable in MAS for the improvement of fiber quality of upland cotton.

(Continued on next page)

* Correspondence: wkgong@aliyun.com; yylcri@126.com
†Equal contributors
[1]State Key Laboratory of Cotton Biology, Key Laboratory of Biological and Genetic Breeding of Cotton, The Ministry of Agriculture, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455000, Henan, China
Full list of author information is available at the end of the article

Jamshed et al. BMC Genomics (2016) 17:197

Page 2 of 13

(Continued from previous page)

Conclusion: Multiple environmental evaluations of an intraspecific RIL population were conducted to identify stable QTLs. Meta-QTL analyses identified a common chromosomal region that plays an important role in fiber development. Therefore, QTLs identified in the present study are an ideal candidate for MAS in cotton breeding programs to improve fiber quality.

Keywords: Recombinant inbred line, Upland cotton, Multiple environments, SSR markers, Meta-QTL analyses, Stable QTLs

Background

Cotton (Genus *Gossypium*) is a well-known and highly important industrial crop that has been grown in more than 80 countries located in tropical and subtropical regions [1]. It is used as an important source of natural fiber, seed oil and proteins [2]. The genus *Gossypium* comprises approximately 45 diploid species and five tetraploid species. Two tetraploid species, *G. hirsutum* and *G. barbadense,* and two diploid species, *G. herbaceum* and *G. arboreum* have been extensively cultivated around the world, with *G. hirsutum* covering >90 % of the total world production and is generally referred to as upland cotton [3]. Upland cotton has a high yield potential, whereas *G. barbadense* has superior fiber quality attributes that subsequently gives it a 30–50 % price advantage over upland cotton [4], whereas the low yield and poor adaptation of *G. barbadense* restricts its production to specific regions around the world. To fulfill the global requirements of the growing human population and the recent advancement in spinning technology justify the need for increased cotton fiber yield and improved cotton fiber traits. Fiber quality traits and yield components are quantitative traits that are negatively correlated [5]. Therefore, it is very difficult to improve all these traits simultaneously by using conventional breeding procedures. Moreover, this would also be laborious and time consuming [6].

Marker-assisted selection (MAS) is prestigious blessing that breaks the linkage among these traits, as it directly selects genetic markers that are tightly linked to quantitative trait loci (QTLs) other than the conventional procedure of indirectly selecting strains with superior phenotypic performance for breeding. Recent developments in field of molecular markers have allowed plant breeders to identify and evaluate complex agronomical traits. The construction of a molecular genetic map is a foundation for the genetic dissection of important economical and agronomical traits, MAS, and map-based cloning [7]. The first molecular linkage map was constructed in 1994 [8]. Since then, several genetic maps have been constructed including interspecific [9–14] and intraspecific crosses [15–20], to explore the cotton genome and to identify QTLs. However, most fiber QTLs obtained from interspecific crosses have limited applications to upland cotton breeding programs [21, 22] as most of markers used in interspecific

cross do not show polymorphism in intraspecific crosses [23]. Saturated intraspecific upland cotton maps are useful but more challenging to construct because of the markedly low rate of polymorphisms of molecular markers within *G. hirsutum*. To overcome this obstacle scientists have employed different mapping populations or used whole-genome sequence-based markers. They used populations involving more than two parents, which have higher polymorphism rates in intraspecific crosses, namely, from 6.6 to 13.7 %, thereby ensuring a surge in genetic diversity and facilitating the identification of more QTLs [19, 23, 24].

Recently physical genome drafts of *G. raimondii* [25, 26] *G. arboreum* [27] and *G. hirsutum* [28, 29] have been completed which could be utilized in the construction of a high-density linkage map and investigate complex traits such as fiber quality. A previous study suggested that the tetraploid species originated from the hybridization of two diploid species, *G. arboreum* (A genome) and *G. raimondii* (D genome) about 1–2 million years ago [2]. Furthermore, more QTLs for fiber traits have been mapped to the D_t sub-genome of upland cotton compared to that in the A_t sub-genome, thus suggesting that it may play an important role in fiber developments [30–32]. A high-coverage genetic map constructed by Tang et al. [33] with SSR markers developed from *G. raimondii* BAC-end sequences has revealed that these D genome-based primers are widely distributed and suitable for whole-genome mapping. Therefore, because of the importance of the D_t sub-genome in determining fiber quality traits [23], we used D genome (*G. raimondii*) sequence-based SSR primers [26], together with SSR primers from Cotton Marker Database (http://www.cottonmarker.org/) to construct an intraspecific linkage map. Previously, Sun et al. [18] reported a linkage map based on an intraspecific cross of upland cotton cultivars sGK9708 and 0–153. They used 200 SSR markers to construct a genetic map and identified 50 QTLs for fiber quality in the F_2, $F_{2:3}$ and RIL populations in 4 environments. We added 603 primers to our published genetic map and identified QTLs for fiber quality in 11 environments, including four previously reported environments [18] (Table 1) to augment our previous results from the same intraspecific RIL ($F_{6:8}$) population of upland cotton. Furthermore we conducted a meta-analyses with

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 3 of 13

Table 1 Details of 11 environments used to evaluate 196 RIL along with their parents

Year	Environment	Abbreviation used	Replication	Layout
2007	Anyang[a]	Ay07	2	5× 0.8 m
2008	Anyang[a]	Ay08	2	5× 0.8 m
	Quzhou[a]	Qz08	2	5× 0.8 m
	Linqing[a]	Lq08	2	5× 0.8 m
2009	Anyang	Ay09	2	5× 0.8 m
	Quzhou	Qz09	2	5× 0.8 m
	Akesu	Ak09	2	2× 0.6 m
2010	Anyang	Ay010	2	5× 0.8 m
	Zhengzhou	Zz010	2	5× 0.8 m
	Gaoyi	Gy010	2	5× 0.6 m
2013	Anyang	Ay013	2	5× 0.8 m

[a]Data of these environments was reported in our previous report and used for QTL mapping but excluded in ANNOVA except Quzhou 2008

Biomercator V4.2 [34] using the fiber QTLs identified from the present study, those previously reported in F_2, $F_{2:3}$ and RIL population [18], and those generated from meta-analyses conducted by Said et al. [35, 36], along with three succeeding QTLs studies [33, 37, 38]. We identified some stable and consistent QTLs that aggregated into clusters in upland cotton. These QTL clusters can be made more valuable to MAS to improve the fiber quality of upland cotton.

Results
Assessment of phenotypic performance
The phenotypic performance of the five fiber traits was observed to continuously segregate, and transgressive segregation was observed. Very low absolute skewness and kurtosis values showed that these traits were normally distributed (Table 2). The results of correlation analyses of fiber quality traits in RILs are presented in Table 3. Positive correlations between any of the two traits, which included fiber elongation (FE), fiber length (FL), fiber strength (FS), and fiber uniformity (FU), were observed, with a significance level of 0.01. Fiber micronaire (FM) was negatively correlated with FL and FS. ANNOVA revealed that fiber quality traits presented significant environmental and genetic effects ($P < 0.01$, Table 4). A broad

sense heritability test was also performed for all fiber traits as defined elsewhere [39]. Fiber elongation had the lowest heritability (0.27), whereas that of other fiber traits was high, ranging from 0.80 (FU) to 0.93 (FL).

Construction of a genetic map
In the present study, we obtained 851 primer pairs that were clearly polymorphic between the two parents, 0–153 and sGK9708. These 851 primer pairs generated 997 loci, in which 132 pairs produced two loci, 13 pairs yielded three loci, and two pairs resulted in four loci. All 997 loci were used in the construction of a linkage map. A total of 793 loci were grouped into 76 linkage groups. Seventy three groups were assigned to 26 chromosomes of upland cotton (Additional file 1). Three groups could not be associated with any chromosome. We named these "UD" following the number. The total recombinant length of this map was 4,110 cM, which represented approximately 93.2 % [40] of the total length of the cotton genome, with an average distance of 5.2 cM between adjacent markers. The A_t sub-genome spanned 1,635 cM, consisted of 269 markers on 37 linkage groups, and with an average distance of 6.1 cM between adjacent markers. Thirty six groups were assigned to the D_t sub-genome and comprised 524 markers spanning 2,327.4 cM, with an average of 4.6 cM between adjacent loci (Table 5). Chromosomes c4, c5, c14, c16 and c25 had more markers compared to the other chromosomes. Among these, c25 had 113 loci that encompassed 204 cM, with an average distance of 1.9 cM between two adjacent markers. The smallest group, c11, had 8 markers, and a total length of 37.8 cM.

Segregation distortion of SSR markers
Segregation distortion is a common occurrence in plants [41], including cotton [7]. We observed severe segregation distortions at a rate of about 45 % (Table 5). Among the 361 distorted loci, 241 (67.1 %) favored sGK9708 alleles and 119 (32.9 %) involved 0–153 alleles. A total of 36 segregation distortion regions (SDRs) were detected on 20 chromosomes (Additional file 1). The A_t sub-genome contained 10 SDRs, whereas the D_t sub-genome comprised 26 SDRs. The largest SDR was on c25, which consisted of 26 distorted loci. The highest number of SDRs on one chromosome was 5, which was observed in

Table 2 The observed phenotypic performance of mean values of fiber quality traits of two parents and RILs in 11 environments

Trait[a]	0-153	sGK9708	RIL	Minimum	Maximum	Std. deviation	Kurtosis	Skewness
FL	30.25	27.40	29.20	23.81	33.69	1.288	−0.256	−0.059
FU	85.66	83.25	84.50	77.10	88.05	0.845	0.096	−0.318
FM	4.39	4.90	4.37	2.21	6.57	0.395	0.078	−0.080
FE	6.49	6.38	6.45	5.55	7.50	0.058	0.357	−0.181
FS	33.27	25.75	30.05	22.85	36.10	1.872	0.092	0.293

[a]*FL* fiber length, *FU* fiber uniformity, *FM* fiber mironaire, *FE* fiber elongation, *FS* fiber strength

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 4 of 13

Table 3 Correlation analyses among fiber quality traits based on eleven environments for RIL

Traits[a]	FL	FU	FM	FE
FU	0.616[b]			
FM	−0.371[b]	0.031		
FE	0.487[b]	0.619[b]	−0.023	
FS	0.740[b]	0.759[b]	−0.364[b]	0.592[b]

[a]For trait abbreviations see Table 2
[b]Indicates the correlation reaches the significant level at 0.01

c16 and c25. One chromosome (c21) contained 3 SDRs, 6 chromosomes (c4, c13, c14, c18, c20, and c26) comprised 2 SDRs, whereas the remaining 11 chromosomes (c2, c5–c9, c15, c17, c19, c23, and c24) harbored only 1 SDR.

Collinearity between the linkage and physical map

Loci collinearity between linkage map and the *G. hirsutum* physical map of various chromosomes is presented in Fig. 1. Some loci whose physical location was not confirmed were excluded from the analysis. The overall loci order on the genetic map was in agreement with the order of corresponding sequences on the A_t and D_t sub-genomes of *G. hirsutum*. In the A_t sub-genome (c1–c13), 1.76 GB corresponded to 1,635 cM, whereas in the D_t sub-genome (c14–c26) 774 Mb was equivalent to 2,327 cM.

QTL mapping of fiber quality traits

A total of 165 QTLs for five fiber traits were identified on 24 chromosomes using the composite interval mapping

Table 4 ANNOVA and a broad sense heritability of fiber quality traits in RIL population

Traits[a]	Source	DF	Mean square	*F* Value	Pr > F	H^2_B
FL	e	7	178.2	266.85	<.0001	0.93
	g	195	23.2	34.81	<.0001	
	g*e	1365	1.4	2.15	<.0001	
FU	e	7	364.9	298.18	<.0001	0.8
	g	195	11.6	9.47	<.0001	
	g*e	1365	2.0	1.63	<.0001	
FM	e	7	32.0	200.99	<.0001	0.85
	g	195	2.5	15.74	<.0001	
	g*e	1365	0.3	2.11	<.0001	
FE	e	7	217.9	4905.24	<.0001	0.27
	g	195	0.2	4.62	<.0001	
	g*e	1365	0.2	3.97	<.0001	
FS	e	7	1582.2	1056.22	<.0001	0.92
	g	195	44.9	30.01	<.0001	
	g*e	1365	3.7	2.49	<.0001	

[a]For trait abbreviations see Table 2
H^2_B is broad sense heritability, e is environment and g is genotype

method [42]., Forty seven QTLs identified in a minimum of 3 and a maximum of 10 environments were declared as stable QTLs, of which 12 QTLs were described as stable in our previous report [18], whereas 35 were novel. The physical map was also used to identify QTLs that confirmed 69 QTLs, including 43 stable ones. Two chromosomes, c14 and c25 had more QTLs. No QTL was detected on c1 and c8. Approximately 58 QTLs were identified on the A_t sub-genome chromosomes, whereas 107 QTLs were localized to the D_t sub-genome chromosomes. QTLs positions with their observed phenotypic variance (PV) and nearest loci are listed in Additional file 2 and graphically presented in Additional file 1.

Fiber length

In total, 31 QTLs for FL were detected on 11 chromosomes, including c4, c6, c7, c14, c16, c18, c21, c22, c23, c24, and c25 (Additional file 3). The highest number of QTLs on one chromosome was 6 (c25). Four chromosomes, c6, c16, c22, and c24, harbored only one QTL. Twelve QTLs for FL were identified in only one environment and 5 QTLs were detected in two environments. Fourteen QTLs were identified in 3 or more environments and declared as stable QTLs. Nine stable QTLs for FL on c4, c7, c16, c23 and c25 have favorable alleles from parent 0–153, whereas 5 stable QTLs on c14, c18 and c21 showed favorable alleles from parent sGk9708. The QTL on c4, *qFL-C4-2*, was identified in three environments, explaining 5.8–8.1 % of the observed PV. Two QTLs on c7, *qFL-C7-1* and *qFL-C7-2* were also identified in 4 environments described in our previous report [18]. The QTL *qFL-C7-1* was stable and identified in 3 environments, explaining 5.8–12.1 % of the observed PV. Three QTLs on c14, *qFL-C14-1*, *qFL-C14-2* and *qFL-C14-3* were identified in 8, 6, and 3 environments, explaining 8.1–13.1 %, 7.1–11.5 %, and 6.3–8.1 % of the observed PVs, respectively. The QTL *qFL-C14-2* was also identified in our previous report [18] in 3 environments. The QTL on c16, *qFL-C16-1* was identified in three environments, explaining 5.7–7.5 % of the observed PV. The QTL on c18, *qFL-C18-3* was identified in a single environment in our previous report [18] and now in four environments, explaining 5.2–11.0 % of the detected PV. The QTL on c21, *qFL-21-1* was identified in seven environments, explaining 8.7–23.6 % of the observed PV. The QTL on c23, *qFL-C23-2* was identified in a single environment in our previous report [18] and now in three environments, explaining 9.0–14.9 % of the observed PV. Five QTLs on c25, *qFL-C25-2*, *qFL-C25-3*, *qFL-C25-4*, *qFL-C25-5*, and *qFL-C25-6* were respectively identified in 4, 6, 5, 5, and 3 environments, explaining 5.2–10 %, 6.8–9.4 %, 6.8–11.8 %, 6.5–10.5 % and 8.6–10.6 % of the observed PVs, respectively. Two QTLs, *qFL-C25-2* and *qFL-C25-3*, were

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 5 of 13

Table 5 Genomic distributions of SSR markers and identified QTLs

Chromosome	Linkage groups	Mapped markers	Distorted loci	Total distance covered	Aver. distance b/w markers	Max[a]. distance b/w markers	Min[a]. distance b/w markers
Chr1	3	17	6	158.2	9.0	18.4	2.09
Chr2	3	23	12	106.6	3.9	18.0	0.37
Chr3	2	15	5	108.4	11.3	28.7	0.68
Chr4	1	36	15	177.4	5.1	25.8	0.51
Chr5	2	31	9	170.6	6.3	27.9	0.50
Chr6	4	35	7	193.6	3.6	14.8	0.42
Chr7	3	25	14	103.8	4.3	12.7	0.40
Chr8	2	11	2	99.8	11.0	16.8	4.76
Chr9	5	14	3	132.6	13.0	15.9	11.44
Chr10	4	14	6	96.5	21.5	30.4	12.56
Chr11	3	8	3	37.9	9.5	17.5	1.60
Chr12	4	21	9	170.8	10.3	25.0	0.45
Chr13	1	19	9	79.2	4.4	16.4	0.35
Chr14	3	55	22	200.1	3.1	21.5	0.19
Chr15	2	41	13	161.2	3.3	14.9	0.36
Chr16	1	67	34	132.3	2.0	9.8	0.12
Chr17	4	18	9	149.9	7.6	20.2	1.73
Chr18	5	39	23	268.5	4.6	24.8	1.07
Chr19	3	26	11	170.1	4.2	14.5	0.17
Chr20	4	33	21	217.3	7.6	32.0	0.34
Chr21	2	28	14	213.3	6.0	16.6	1.73
Chr22	2	21	6	126.1	7.0	13.5	2.38
Chr23	3	39	12	224.4	5.7	20.7	0.78
Chr24	1	14	12	108.7	8.4	31.8	0.38
Chr25	1	113	62	204.8	1.9	13.8	0.01
Chr26	5	18	12	150.8	7.9	14.7	2.88
UD	3	12	10	147.3	10.8	21.4	0.08
Total (AD)	76	793	361	4110.0	5.2	32.0	0.01

[a] Max. Distance means maximum marker interval within linkage groups in that chromosome and Min distance means minimum marker interval between two markers in linkage groups of particular chromosome

also previously identified in four environments [18]. In total, 14 QTLs out of 31 were also identified during QTL analysis with the physical map including 11 stable QTLs.

Fiber strength

A total of 35 QTLs for FS were identified on 13 chromosomes including c4, c6, c7, c9, c11, c12, c13, c14, c18, c19, c21, c23, and c25 (Additional file 3). The highest number of QTLs on one chromosome was 7 (c25). Five chromosomes, c6, c9, c11, c12, and c19, harbor a single QTL. Twenty-one QTLs for FS were identified in only one environment and six QTLs were identified in two environments. Eight QTLs were detected in three or more environments and declared as stable QTLs. Six stable QTLs for FS on c7 and c25 have favorable alleles

from parent 0–153, whereas two stable QTLs on c14 showed favorable alleles from parent sGk9708. The QTL on c7, *qFS-C7-1,* was identified in 10 environments, explaining 12.2–26.7 % of the observed PV. The QTL *qFS-C7-2* was identified in seven environments, explaining 7.9–11.2 % of the observed PV. Both stable QTLs were also previously identified in four environments [18]. The QTL on c14, *qFS-C14-3* was identified in eight environments, explaining 4.9–13.7 % of the observed PV. The QTL *qFS-C14-4* was identified in four environments explaining 5.4–8.5 % of the detected PV. Four QTLs on c25, *qFS-C25-3, qFS-C25-4, qFS-C25-5,* and *qFS-C25-6,* were respectively identified in 3, 5, 6 and 7 environments, explaining 7.9–17.0 %, 8.4–15.0 %, 5.4–15.0 %, and 6.4–15.8 % of the observed PVs. Two QTLs,

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 6 of 13
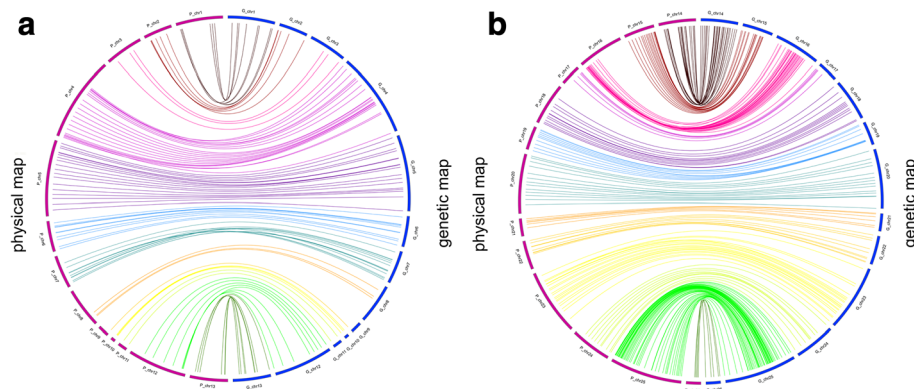
Fig. 1 Collinearity analyses between genetic map 0–153 and physical map of *G. hirsutum*. **a** Collinearity analyses between genetic map of 0–153 from C1-C13 (total distance 1635 cM) with corresponding sequence on At sub-genome (1.16GB) of *G. hirsutum*. **b** Collinearity analyses between C14-C26 (total distance 2327 cM) of genetic map with corresponding sequence of D$_t$ sub-genome (776 Mb) of *G. hirsutum*

qFS-C25-3 and *qFS-C25-4* were also earlier identified in four environments [18]. All eight stable QTLs were also detected and confirmed through physical map analysis.

Fiber elongation

For the FE trait, 32 QTLs were identified and located on 13 chromosomes including c3, c4, c7, c10, c13, c14, c15, c19, c21, c22, c23, c25, and c26, explaining 3.15–17.9 % of the observed PV (Additional file 3). The highest number of QTLs on one chromosome was 8 (c25). Six chromosomes, c10, c13, c21, c22, c23, and c26, harbored a single QTL. Eighteen QTLs were identified in one environment, whereas four QTLs were identified in two environments. Ten QTLs for FE were detected and described as stable QTLs. Six stable QTLs for FE on c4, c22, and c25 have favorable alleles from parent 0–153, whereas four stable QTLs on c14 showed favorable alleles from parent sGk9708. Two QTLs on c4, *qFE-C4-2* and *qFE-C4-3* were respectively identified in three and five environments, explaining 4.6–8.5 % and 5.5–12.4 % of the observed PVs, respectively. The QTL, *qFE-C4-2* was also previously identified in four environments [18]. Four QTLs on c14 *qFE-C14-1, qFE-C14-2, qFE-C14-3,* and *qFE-C14-4* were respectively identified in 4, 3, 4 and 3 environments, explaining 8.8–17.9 %, 7.4–14.3 %, 8–15 % and 9.8–11.8 % of the observed PVs. The QTL on c22, *qFE-C22-1,* was identified in three environments, explaining 7.2–13.8 % of the observed PV. Three stable QTLs on c25, *qFE-C25-4, qFE-C25-5,* and *qFE-C25-6* were identified in 3, 4, and 3 environments, explaining 5.6–9.4 %, 5.6–10.1 % and 6.8–10.4 % of the observed PVs, respectively. The QTL *qFE-C25-4* was also earlier identified in three environments [18]. In total, 13 QTLs, including 10 stable ones were also identified and confirmed through physical map-based QTL analysis.

Fiber uniformity

For FU, 32 QTLs were identified and located on 14 chromosomes including c2, c4, c5, c6, c7, c10, c12, c13, c14, c16, c18, c19, c23, and c25, explaining 1.8–18.2 % of the observed PV (Additional file 3). The highest number of QTLs on one Chromosome was 7 (c25). Seven chromosomes, c5, c6, c12, c13, c18, c19, and c23 harbored a single QTL. Twenty QTLs were identified in one environment, whereas seven 7 QTLs were detected in two environments. Five QTLs for FU were identified as stable QTLs. Three stable QTLs for FU on c7, c13, and c25 have favorable alleles from parent 0–153, whereas two stable QTLs on c14 showed favorable alleles from parent sGk9708. The QTL on c7, *qFU-C7-1* was identified in six environments, explaining 7.0–18.2 % of the observed PV. This was also previously identified in the F$_{2:3}$ and RIL populations in two environments [18]. The QTL on c13, *qFU-C13-1* was identified in three environments, explaining 4.4–6.5 % of the observed PV. It was same QTL that we earlier identified in two environments [18]. Two stable QTLs on c14, *qFU-C14-2* and *qFU-C14-3,* were respectively identified in five and four environments, explaining 6.7–14.2 % and 7.6–10.1 % of the observed PVs. The QTL on c25, *qFU-C25-5* was identified in four environments, explaining 6.4–8.0 % of the observed PV. This was also earlier identified in four environments [18]. In total, 13 QTLs including five stable ones were also confirmed through QTL analysis using a physical map.

Fiber micronaire

A total of 35 QTLs were identified for FM on 16 chromosomes including c3, c4, c5, c6, c7, c10, c13, c14, c15, c16, c17, c20, c21,c23, c24, and c25 (Additional file 3). The highest number of QTLs on one chromosome was 6 (c25). Eight chromosomes, c3, c7, c10, c13, c17, c21, c23,

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 7 of 13

and c24 harbored a single QTL. Eighteen QTLs were identified in one environment, whereas seven QTLs were identified in two environments. Ten QTLs were identified as stable QTLs. Three stable QTLs on c3, c4 and c16 have favorable alleles from parent 0–153, whereas seven stable QTLs on c7, c14 and c25 comprised favorable alleles from parent sGk9708. The QTL on c3, *qFM-C3-1* was identified in three environments, explaining 5.3–5.6 % of the observed PV. It was also identified in our previous report in one environment [18]. The QTL on c4, *qFM-C4-2* was identified in four environments, explaining 7.7–8.7 % of the observed PV. The QTL on c7, *qFM-C7-1,* was identified in five environments, explaining 9.6–16.7 % of the observed PV. The QTL on c16, *qFM-C16-3,* was identified in three environments, explaining 5.2–7.9 % of the observed PV. It was also identified in our previous report [18] in one environment. Two QTLs on c14, *qFM-C14-2* and *qFM-C14-3,* were identified in four environments, explaining 6.1–9.1 % and 6.5–8.6 % of the observed PVs, respectively. The QTL, *qFM-C14-3* was also identified in our previous report [18] in one environment. Four QTLs on c25, *qFM-C25-1, qFM-C25-2, qFM-C25-4,* and *qFM-C25-5,* were respectively identified in 4, 5, 3, and 4 environments explaining 7.3–10.3 %, 5.2–9.9 %, 6.3–8.5 % and 6.2–10.5 % of the observed PVs. The QTL, *qFM-C25-4* was also previously identified [18] in four environments. In total, 15 QTLs, including eight stable ones, were also identified and confirmed through physical map analysis.

QTL clusters and meta-analysis

QTL clustering is a common phenomenon in plants and also observed in cotton [32, 43, 44]. We identified 30 clusters on 11 chromosomes including c3, c4, c6, c7, c10, c12, c13, c14, c16, c21 and c25. Most of stable QTLs fall in these cluster regions. Six clusters having QTLs for all five fiber traits were identified on c7, c14 and c25, among which, the cluster on c7 *c7-cluster 1* contained five QTLs that were tightly linked to markers PGML00802 and NAU2627 explaining 5.9–26.7 % of the observed PV. Two QTL clusters on c25, *c25-cluster 2* and *c25-cluster 4* contained nine and five QTLs that were tightly linked to markers TMK19, BNL3806b, PGML00463b, SWU19198, and NBRI1529 explaining 5.5–17.0 % and 6.2–14.5 % of the observed PVs, respectively. Three clusters on c14, *c14-cluster 2, c14-cluster 3,* and *c14-cluster 4* each contained five QTLs that were tightly linked to marker SWU14535, PGML00989, NAU3393, SWU 14507, CSHES150, BNL3099, and COT99 and explaining 5.7–15 %, 5.4–11.8 % and 6.3–10.0 % of the observed PVs, respectively. The details of each cluster are summarized in Additional file 4.
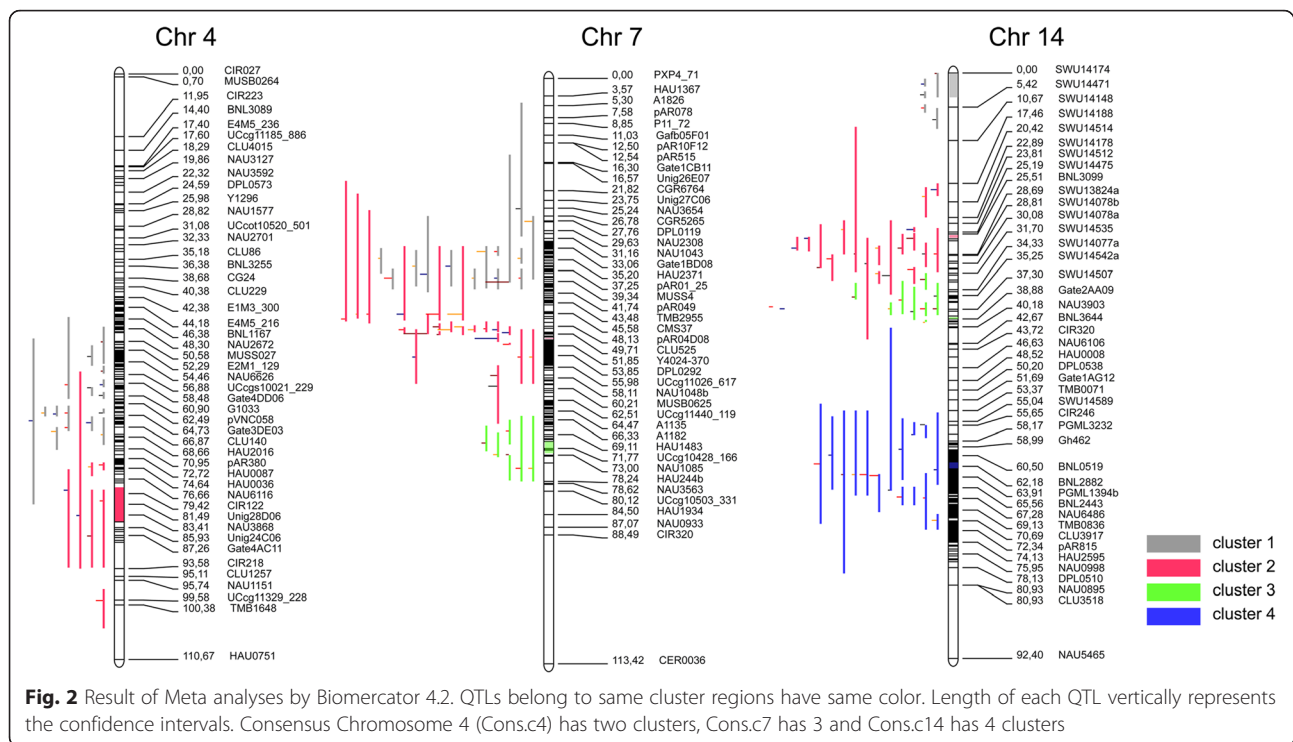
In the meta-analysis, a total of 38 meta-cluster regions on 11 chromosomes were identified, which included c4, c5, c7, c12, c13, c14, c15, c16, c20, c23, and c25 (Additional file 5). The results showed that some clusters in the 0-153xsGK9708 genetic map, (which were very close) were grouped into the same 20-cM meta-cluster region on the consensus map and part of same meta-cluster (Additional files 6 and 7). Twenty-nine QTLs were projected on consensus chromosome 4 (Cons.c4), which resulted into 2 QTL meta-clusters. *C4-m-cluster-1* has 14 QTLs, while *C4-m-cluster-2* has seven QTLs (Fig. 2). Fifty-three QTLs were projected on Cons.c7 which yielded three QTL clusters. *C7-m-cluster-1*, *C7-m-cluster 2,* and *C7-m-cluster-3* contained 12, 21 and 6 QTLs respectively (Fig. 2). Seventy-six QTLs were projected on Cons.c14 which resulted in four meta-clusters. *C14-m-cluster-1, C14-m-cluster-2, C14-m-cluster-3,* and *C14-m-cluster-4* contained 5, 16, 8, and 15 QTLs respectively (Fig. 2). Sixty-eight QTLs were projected on Cons.c25 which resulted in four QTL clusters. *C25-m-cluster-1, C25-m-cluster-2, C25-m-cluster-3,* and *C25-m-cluster-4* contained 18, 15, 21, and 6 QTLs, respectively. The details of the remaining QTLs are summarized in Additional file 5. The cluster on Cons.c4, *C4-m-cluster-1* from the 45–65 cM interval was situated between markers DPL0196 and NAU3093 (40,406,319–59,166,290 bp). Cluster, *C4-m-cluster-2* from 73 to 93 cM interval was located between markers DPL0451 and CIR218 (60,349,199–62,668,683 bp). The cluster on Cons.c7, *C7-m-cluster-1,* from the 20–36 cM interval was localized between markers NAU5303 and NAU3918 (3,545,485–8,213,231 bp). Cluster, *C7-m-cluster-2* from the 40–58 cM interval was located between markers BNL1597 and NAU2186 (9,280,354–15,534,308 bp). Cluster, *C7-m-cluster-3* from 60 to 72 cM interval was situated between markers NAU1085 and CIR238 (16,350,941–2,178,086 bp). The cluster on Cons.c14, *C14-m-cluster-1,* from 0 to 15 cM interval was localized between markers SWU14174 and SWU14188 (17,025,534–21,454,750 bp). Cluster, *C14-m-cluster-2* from 20 to 36 cM interval was localized between markers BNL3099 and COT099 (49,640,545–50,515,032 bp). *C14-m-cluster-3* from 38 to 54 cM interval was between markers NAU3393 and PGML0989 (11,844,310–17,029,745 bp) and *C14-m-cluster-4* from the 58–78 cM interval was between markers DPL0354 and BNL3033 (62,556,024–70,746,352 bp).

Discussion

Genetic map

The identification of stable QTLs for superior agronomically significant traits and the construction of a high-resolution map are essential for MAS. Several intraspecific genetic maps have been reported; however, these contain some gaps that limit its applicability in generating a high-density genetic map. Major obstacles in the construction of a high-resolution map in intraspecific crosses include a low rate of polymorphism within *G. hirsutum* and the presence of fixed homozygous genetic

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 8 of 13

Fig. 2 Result of Meta analyses by Biomercator 4.2. QTLs belong to same cluster regions have same color. Length of each QTL vertically represents the confidence intervals. Consensus Chromosome 4 (Cons.c4) has two clusters, Cons.c7 has 3 and Cons.c14 has 4 clusters

blocks [23]. Therefore, there is a need to identify additional markers that covers these gaps in the genetic map. In the present study, an updated genetic map based on our previous report showing 190 markers [18] is described. We have added 586 markers including 386 (41 % of the total number of markers) novel SWU primers. Among these 793 markers, 524 were mapped to the D_t sub-genome and 269 were mapped to the A_t sub-genome. In our previous report, chromosomes c4, c7, c13, c14, c18, and c25 were identified as important and rich in QTLs for fiber quality traits [18]. Most of the new markers that we have successfully added to the map have been localized to these chromosomes, thereby enabling us to dissect these QTLs into clusters at a higher resolution, as well as identify some important stable QTLs for specific superior features. In the current map, 20 chromosomes harbored more than one linkage group, which indicates a relatively low rate of polymorphism in intra specific crosses which was observed at a rate of 2.9 % in the present study. The observed relatively low rate of polymorphism suggests that the genetic distance between the two parents was very narrow, thereby indicating the need for a saturated intra-specific map. Therefore, our next goal is to develop new SSR and SNP primers that would facilitate in the construction of a saturated genetic map.

Segregation distortion
Among the 793 mapped primers, 361 showed distortion from the normal Mendelian ratio, which is 1:1 in the case of RILs. This severe distortion was also reported by Sun et al. [18] and commonly occurs in RIL populations that were developed from an introgressed line parent. This high ratio of segregation distortion in our population may be attributed to parent 0–153, which is an introgressed line. Tang et al. [33] also reported similar results (41.8 %) in their RIL population with introgressed parental line, 7235. Segregation distortion could be influenced by various factors including genetic factors such as genetic drift [45] and the environment. However, it does not significantly impact the estimation of QTL position and effect [46]. The broad sense heritability estimates of fiber quality traits were high for FL, FM, FS and FU, indicating that the QTLs identified in this population are more reliable and useful in MAS for cotton breeding.

Distribution of QTLs among A_t and D_t sub-genomes
The distribution of QTLs was not uniform in the A_t and D_t sub-genomes. Among the 165 QTLs identified, 58 QTLs (35 % of the total) were identified in the A_t sub-genome, whereas 107(65 % of the total) were identified in the D_t sub-genome. Previous comparative meta-analyses conducted by Rong et al. [32], Lacape et al. [43] and Said et al. [36] have indicated that in cotton a higher number of QTLs for fiber traits resided within D_t sub-genome chromosomes, and gene expression among homologous pairs were not uniform [44, 47]. Yu et al. [48] also observed 35 % more QTLs in the D_t sub-genome in an inter specific backcross inbred line population. In the present

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 9 of 13

study a higher number of loci were mapped to the D_t sub-genome. This observation might be due to the presence of more SSR markers that were developed from the D genome sequence [26], although this phenomenon has also been previously described by Yu et al. [49] in their BC_1 population. However we also observed that some A_t sub-genome chromosomes also have more loci than its homologous counterparts in D_t sub-genome chromosomes. This unequal distribution of loci indicates the presence of active regions with more recombination frequencies in the upland cotton genome [4]. Similarly, QTLs on both pairs were also not homogeneous. Most importantly, homology was observed between homologous pair c6-c25 and c7-c16, which harbored QTL clusters and were in agreement with the findings of previous reports [23, 43].

Comparison of the tetraploid cotton genome with its ancestors shows that only the A genome (*G. arboreum*) produces spinnable fibers, whereas the D genome (*G. raimondii*) lacks this characteristic. After polyploidization, transposable elements tend to be more active, especially in the D_t sub-genome, compared to that in the A_t sub-genome. Furthermore, the D_t sub-genome also has a higher mutation rate than the A_t sub-genome [28]. These findings might also contribute to our observation that the D_t sub-genome harbored more QTLs than the A_t sub-genome. However, the additional of novel markers for the A_t Sub-genome may improve the assessment of the contribution of each sub-genome in fiber quality traits.

Consistency with previously reported fiber QTLs

It is very difficult to compare different QTLs that have been reported in various populations, although this is necessary to fully understand the behavior of complex traits, particularly in a changing environment. In present study, 325 markers were designated as novel SSRs (Additional file 8). However, some regions did not have common markers at QTLs and thus we were unable to compare these with the findings of previous reports. However some stable QTLs with common markers have been identified and were used in our meta-analyses. We identified 38 cluster regions. When a meta-cluster contained stable QTLs from our RIL population and QTLs were identified by recent meta-analyses report [35], this was considered as the same cluster. We also confirmed the previous meta-analyses report [35], which in turn allowed us to declare a true stable QTL in this consensus genomic region. For example Lacape et al. [12], Shen et al. [5, 6] and Sun et al. [18] reported QTLs for fiber strength and length that were linked to primers BNL3806, TMK19, and BNL1440 on c25. We have identified two clusters that were tightly linked to these primers. Four QTLs for fiber quality traits FE, FL, FM and FS were closely linked to primer BNL3806 and TMK19. Four QTLs for the fiber quality

traits, FE, FL, FS, and FU were tightly linked to BNL1440. These QTLs were in two meta-cluster regions *C25-cluster-1:0–20 cM* and *C25-cluster-2-25-45 cM*. Our results confirm the findings of Said et al. [36] as well as declare that these QTLs are indeed stable. We also verified its physical position in the genome sequence of *G. hirsutum*. QTL analysis on the basis of the physical map also confirmed that these loci were closely linked to these fiber quality traits. However, additional studies confirming the presence of putative genes in this region are warranted. Meta-clusters that harbor QTLs from our RIL population and the latest QTL studies except for those identified by Said et al. [36] were regarded as new meta-clusters in the present study. Of the 38 meta-clusters, 31 clusters with 314 QTLs were considered similar to that of a previous report [36]. In Addition, we identified seven novel cluster regions with 55 QTLs for fiber quality traits in the present study. The cluster on Cons.c4, C4-m-cluster-1, which contained 14 QTLs including five fiber quality traits FE, FS, FL, FU, and FM was considered as novel. Three stable QTLs identified in our RIL population *qFE-C4-3, qFM-C4-2,* and *qFL-C4-2* and one stable QTL identified by Tang et al. [33], *qFS04.1* were also detected in this cluster region. The cluster on c7, *C7-m-cluster-3* which contained six QTLs for three fiber traits FL, FS, and FU was considered as a novel cluster. One stable QTL, *qFS-C7-2*, which was identified in our RIL population and one QTL, *qFU07.1* identified by Tang et al. [33], were also confined in this cluster region. On Cons.c14, *C14-m-cluster-2* and *C14-m-cluster-3* were respectively identified as novel clusters. The *C14-m-cluster-2*, contained 16 QTLs including six stable QTLs for five fiber quality traits, were identified in our RIL population. *C14-m-cluster-3* contained three stable QTLs that were identified in our RIL population and one stable QTL *qFS14.1,* that was earlier identified by Tang et al. [33]. On c15 and c20, *C15-m-cluster-4 and C20-m-cluster-3* were considered as novel clusters, respectively (Additional file 5). On c25, *C25-m-cluster-4* which contained six QTLs for fiber quality trait was considered as a novel cluster. Fine mapping of c25 was also performed and discussed separately [50].

Conclusion

QTLs detected in different environments are stable QTLs [51], that may be utilized in MAS and RIL population are useful in the detection of stable QTLs in multiple environments [52]. We have identified 165 QTLs, of which 30 QTL clusters were identified in an intraspecific RIL population in 11 environments. Meta analyses results have revealed that 90 fiber QTLs in the RIL population were in agreement with the findings of previous reports. We have identified seven novel cluster regions that contained 55

Jamshed et al. BMC Genomics (2016) 17:197

Page 10 of 13

fiber QTLs, including 33 QTLs from the RIL population. QTL clusters on c4, c7, c14 and c25 were identified as stable across multiple environments and populations. Therefore, these clusters were considered important for cotton breeders and can be utilized in MAS to improve fiber quality.

Methods

Mapping population
A segregation population consisting of 196 $F_{6:8}$ RIL individuals were derived from a cross between two upland cotton strains, 0–153 and sGK9708. Strain sGK9708 is insect resistant with moderate fiber quality and high yield potential, whereas strain 0–153 has excellent fiber quality with low yield. The cross was made in 2001 and recombinant inbred lines were developed as detailed by Sun et al. [18]. From 2007 to 2013 multi-environmental evaluations were conducted in six different locations throughout China with two replications in each environment (Table 1). Sun et al. [18] reported four environments from the year 2007 to 2008. We added seven more environments with three additional locations to the total phenotypic data set (Table 1). These evaluation procedures were also earlier described by Zhang et al. [50].

Phenotyping
Fiber samples were collected from each line to investigate fiber quality traits. 30 normally opened bolls were collected from each plot. Fiber quality traits were measured using an HVI-100 instrument (user technologies, Switzerland) at the Cotton Fiber Quality Inspection and Testing Center of Ministry of Agriculture, Anyang, China. The fiber quality traits included FE, FL, FM, FS and FU. These observed phenotypic data were analyzed by using the software SPSS20.0 (SPSS, Chicago, IL, USA). For ANOVA, we used the SAS statistical software (version 8.1; SAS institute, Cary NC). To calculate broad sense heritability the following equation was used

$$H^2_B(\text{Broad sense heritability}) = \sigma^2 G/(\sigma G + \sigma^2 G * E/n_e + \sigma^2 E/n_e n_r)$$

Where $\sigma^2 G$ is genotypic variance, $\sigma^2 G*E$ is genotype * environment variance, and $\sigma^2 E$ is variance of error.

Genotyping
DNA extraction
Young leaves were collected from each line and stored at −80 °C. Genomic DNA from the parents and 196 RILs was extracted using a modified CTAB method as described by Paterson et al. [53]. PCR amplification was performed in a total reaction volume of 10 μL containing 6.15 μL ddH$_2$O, 1 μL 10× buffer (with 1.5 mL Mg$^+$), 0.5 μL dNTPs (10 mM), 0.5 μL each primer, 0.15 μL of Taq polymerase (500U) and 1.2 μL of genomic DNA (30 ng/μL). PCR amplification conditions comprised of an initial denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 94 °C for 1 min, annealing at 57 °C for 30s and an extension at 72 °C for 60s followed by a final elongation at 72 °C for 5 min, and then held at 4 °C until analysis. PCR products were electrophoresed on an 8 % non-denatured polyacrylamide gel and silver staining was used for visualization of bands.

SSR analyses
A total of 28,891 primers pairs, including 12,560 SWU primers (D genome sequence-based), were used to detect polymorphisms between the two parents. Approximately 851 polymorphic primers were selected and used in genotyping 196 recombinant inbred lines. All loci were named according to their respective primer names. In the case of multiple loci generated by single primer pair that showed a different segregation pattern from that of the main band, a suffix of a/b/c was used after the primer name to differentiate loci according to increasing molecular size. The details of the primers used in the present study are listed in Additional file 7. The SWU primers were synthesized by Beijing Genomics Institute (Beijing, China), whereas all other primers were synthesized by Invitrogen, Co. Ltd. (Shanghai, China) and Bio Asia, China (Beijing, China).

Construction of the genetic map and QTL analyses
A linkage map was constructed using JoinMap 4.0 [54] with a logarithm of odds (LOD) threshold of >7 and a maximal distance of 50 cM. Recombination frequencies were converted to map distance using the Kosambi map function [55]. For some groups that have mixed markers belonging to different chromosomes, a higher LOD score of >9 was used to separate these into small groups. Linkage groups were assigned to its respective chromosome based on previous reports [18, 19, 20, 33, 5657] and marker mapping information from the CottonGen database (http://www.cottongen.org/). Small groups that were mapped to the same chromosome were recalculated to combine these into one group. A minimum LOD score of 6 was used to combine these groups. In the case of c20 and c23, an LOD score of 5 was used to combine small linkage groups into one. The G. hirsutum fasta sequence was downloaded from http://www.cottongen.org/ and used to check co-linearity of loci between the linkage map and the G. hirsutum physical map.

QTL analyses and meta-analysis
Windows QTL Cartographer 2.5 [57] was used for QTL mapping. The composite interval mapping method [42] was used at a walking speed of 1 cM and using a

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 11 of 13

1000-permutation test. QTLs for the same trait across different environments were declared the same when its confidence interval overlapped. A QTL identified in at least three environments was declared as stable. We used a physical map in which loci were arranged according to their position on the *G. hirsutum* genome, and QTL analysis was performed using the composite interval mapping method as earlier described.

Meta-analysis was performed with Biomercator 4.2 [34] as described elsewhere [36]. A previous meta-QTL analyses established a QTL data-base [35] consisting of 2,274 QTLs, which included 437 highly consistent QTLs for fiber quality traits from 58 QTL reports on upland cotton [35]. We downloaded its QTL information, including names and CI from www.cottonqtldb.org. We used the high-density consensus map [58] as a reference to project our QTLs and performed chromosome-wise meta-analyses. A total of 850 fiber QTLs from six QTLs reports including 165 fiber QTLs from our RIL population, 50 fiber QTLs from the F_2,$F_{2:3}$, and RIL populations of same parents [18], and 635 fiber QTLs from previous reports literatures [33, 35, 37, 38] were thereby generated.

For meta-analyses, two separate input files were prepared, a map file and a QTL file. The map file contained distances between markers on each chromosome, and the QTL file contained 12 columns, where each row represented a single QTL in a given environment, i.e., QTL name, trait name, trait ontology, experiment place, year, chromosome name, linkage group name, LOD score, observed PV value (R2), most likely position of the QTL, CI start position and CI end position. First both files were loaded into the software and checked for map connectivity. Then QTLs were projected on a consensus map and meta-analyses were performed for each trait. Four models were thus generated, each with an Akaike information criterion (AIC) value. The model with lowest AIC value was selected and used for the identification of mQTL position, whereas QTL clusters were determined manually. The QTLs within the region of 20 cM on the consensus map were considered as part of same cluster as earlier defined by Said et al. [36].

Additional files

Additional file 1: Genetic linkage map of an intraspecific RIL population. (ZIP 7480 kb)

Additional file 2: Fiber QTLs identified in an intraspecific RIL population and their position. (XLSX 39 kb)

Additional file 3: Position and markers details of identified QTLs for fiber quality traits in an intraspecific RIL. (XLSX 37 kb)

Additional file 4: Details of QTL clusters on the 0–153 genetic map of a RIL. (XLSX 29 kb)

Additional file 5: Summary of meta-clusters on a consensus map. (XLSX 10 kb)

Additional file 6: Identified QTLs for fiber traits in a RIL and comparison with the findings of previous reports. (XLSX 46 kb)

Additional file 7: Meta-analysis results of the remaining chromosomes. (DOCX 672 kb)

Additional file 8: Details of the primers used in the study. (XLSX 10 kb)

Abbreviations
AIC: Akaike information criterion; c: chromosome; CI: confidence interval; FE: fiber elongation; FL: fiber length; FM: fiber micronaire; FS: fiber Strength; FU: fiber uniformity; H^2_B: broad sense heritability; LOD: logarithm of odds; MAS: marker-assisted selection; PV: phenotypic variance; QTL: quantitative trait loci; SDRs: segregation distortion regions; SSR: simple sequence repeats.

Authors' contributions
YLY and WKG conceived and designed the experiments. MJ performed D genome-based SSR analysis. JF, JG and KKP performed analysis using the remaining SSR primers. MJ drafted the manuscript. YLY and WKG contributed to the final editing of manuscript. YZS, JWG, HHS, AYL, TTC, ZZ, JC, QG, ZL, QWL, XYD, YNT, HR, ZS, and MH collected field data from six experimental areas during different years. All authors contributed in the interpretation of results and approved the final manuscript.

Author details
[1]State Key Laboratory of Cotton Biology, Key Laboratory of Biological and Genetic Breeding of Cotton, The Ministry of Agriculture, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455000, Henan, China. [2]College of Agronomy, Xinjiang Agricultural University, Key Laboratory of Agro-Biotechnology, Urumqi 830052, Xinjiang, China. [3]College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha 410128, Hunan, China. [4]Anyang College of Technology, Anyang 455000, Henan, China. [5]Department of Materials Science and Engineering, College of Engineering, Peking University, Beijing 100871, China.

References
1. Alkuddsi YA, Rao MG, Patil SS, Joshi M, Gowda TH. Heterosis Studies and per se Performance of Intra *Hirsutum* Hybrids (*G. hirsutum x G. hirsutum*) for Kapas Yield and its Components in Cotton. Cotton Genomics Genet. 2013;4(6):73–92. doi:10.5376/cgg.2013.04.0006.
2. Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD et al. Toward sequencing cotton (*Gossypium*) genomes. Plant physiol. 2007;145(4):1303–10. doi:10.1104/pp.107.107672.
3. Page JT, Huynh MD, Liechty ZS, Grupp K, Stelly D, Hulse AM, Ashrafi H, Van Deynze A, Wendel JF, Udall JA. Insights into the evolution of cotton diploids and polyploids from whole-genome re-sequencing. G3 (Bethesda). 2013;3:1809–18. doi:10.1534/g3.113.007229.
4. Ulloa M, Saha S, Jenkins JN, Meredith WR, McCarty JC, Stelly DM. Chromosomal assignment of RFLP linkage groups harboring important QTLs on an intraspecific cotton (*Gossypium hirsutum* L.) joinmap. J Hered. 2005;96(2):132–44.
5. Shen XL, Guo WZ, Lu QX, Zhu XF, Yuan YL, Zhang TZ. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica. 2007;155(3):371–80.

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 12 of 13

6. Shen XL, Guo WZ, Zhu XF, Yuan YL, Yu JZ, Kohel RJ, Zhang TZ. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. Mol Breeding. 2005;15(2):169–81.

7. Shi Y, Li W, Li A, Ge R, Zhang B, Li J, Liu G, Liu A, Shang H, Gong J. Constructing a high-density linkage map for *Gossypium hirsutum* x *Gossypium barbadense* and identifying QTLs for lint percentage. J integr plant biol. 2015;57(5):450–67. doi:10.1111/jipb.12288.

8. Reinisch AJ, Dong JM, Brubaker CL, Stelly DM, Wendel JF, Paterson AH. A detailed RFLP map of cotton, *Gossypium hirsutum* x *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics. 1994;138(3):829–47.

9. Jiang C, Wright RJ, El-Zik KM, Paterson AH. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). Proc Natl Acad Sci U S A. 1998;95(8):4419–24.

10. Kohel RJ, Yu J, Park YH, Lazo GR. Molecular mapping and characterization of traits controlling fiber quality in cotton. Euphytica. 2001;121(2):163–72.

11. Mei M, Syed NH, Gao W, Thaxton PM, Smith CW, Stelly DM, Chen ZJ. Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). Theor Appl Genet. 2004;108(2):280–91.

12. Lacape JM, Nguyen TB, Courtois B, Belot JL, Giband M, Gourlot JP, Gawryziak G, Roques S, Hau B. QTL analysis of cotton fiber quality using multiple *Gossypium hirsutum* x *Gossypium barbadense* backcross generations. Crop Sci. 2005;45(1):123–40.

13. Frelichowski JE, Palmer MB, Main D, Tomkins JP, Cantrell RG, Stelly DM, Yu J, Kohel RJ, Ulloa M. Cotton genome mapping with new microsatellites from Acala 'Maxxa' BAC-ends. Mol Genet Genomics. 2006;275(5):479–91.

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

15. Shappley ZW, Jenkins JN, Meredith WR, McCarty JC. An RFLP linkage map of Upland cotton, *Gossypium hirsutum* L. Theor Appl Genet. 1998;97(5–6):756–61.

16. Zhang ZS, Xiao YH, Luo M, Li XB, Luo XY, Hou L, Li DM, Pei Y. Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (*Gossypium hirsutum* L.). Euphytica. 2005;144(1–2):91–9.

17. Wang BH, Guo WZ, Zhu XF, Wu YT, Huang NT, Zhang TZ. QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica. 2006;152(3):367–78.

18. Sun FD, Zhang JH, Wang SF, Gong WK, Shi YZ, Liu AY, Li JW, Gong JW, Shang HH, Yuan YL. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Mol Breeding. 2012;30(1):569–82.

19. Zhang K, Zhang J, Ma J, Tang SY, Liu DJ, Teng ZH, Liu DX, Zhang ZS. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (*Gossypium hirsutum* L.). Mol Breeding. 2012;29(2):335–48.

20. Liang QZ, Hu C, Hua H, Li ZH, Hua JP. Construction of a linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Chinese Sci Bull. 2013;58(26):3233–43.

21. Lin ZX, Zhang YX, Zhang XL, Guo XP. A high-density integrative linkage map for *Gossypium hirsutum*. Euphytica. 2009;166(1):35–45.

22. Shang L, Liang Q, Wang Y, Wang X, Wang K, Abduweli A, Ma L, Cai S, Hua J. Identification of stable QTLs controlling fiber traits properties in multi-environment using recombinant inbred lines in Upland cotton (*Gossypium hirsutum* L.). Euphytica. 2015;205(3):877–88.

23. Fang DD, Jenkins JN, Deng DD, McCarty JC, Li P, Wu J. Quantitative trait loci analysis of fiber quality traits using a random-mated recombinant inbred population in Upland cotton (*Gossypium hirsutum* L.). BMC Genomics. 2014;15:397. doi:10.1186/1471-2164-15-397.

24. Qin H, Guo W, Zhang YM, Zhang T. QTL mapping of yield and fiber traits based on a four-way cross population in *Gossypium hirsutum* L. Theor Appl Genet. 2008;117(6):883–94.

25. Wang KB, Wang ZW, Li FG, Ye WW, Wang JY, Song GL, Yue Z, Cong L, Shang HH, Zhu SL et al. The draft genome of a diploid cotton *Gossypium raimondii*. Nat Genet. 2012;44(10):1098–103. doi:10.1038/ng.2371.

26. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin DC, Llewellyn D, Showmaker KC, Shu SQ, Udall J et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature. 2012;492(7429):423–7. doi:10.1038/nature11798.

27. Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, Li Q, Ma Z, Lu C, Zou C et al. Genome sequence of the cultivated cotton *Gossypium arboreum*. Nat Genet. 2014;46:567–72. doi:10.1038/ng.2987.

28. Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, Ma Z, Shang H, Ma X, Wu J et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. Nat Biotech. 2015;33:524–30. doi:10.1038/nbt.3208.

29. Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat Biotech. 2015;33:531–7. doi:10.1038/nbt.3207.

30. Jiang C, Wright RJ, Woo SS, DelMonte TA, Paterson AH. QTL analysis of leaf morphology in tetraploid *Gossypium* (cotton). Theor Appl Genet. 2000;100(3–4):409–18.

31. Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ. QTL analysis of genotype x environment interactions affecting cotton fiber quality. Theor Appl Genet. 2003;106(3):384–96. doi:10.1007/s00122-002-1025-y.

32. Rong J, Feltus EA, Waghmare VN, Pierce GJ, Chee PW, Draye X, Saranga Y, Wright RJ, Wilkins TA, May OL et al. 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. 2007;176(4):2577–88.

33. Tang SY, Teng ZH, Zhai TF, Fang XM, Liu F, Liu DJ, Zhang J, Liu DX, Wang SF, Zhang K et al. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (*Gossypium hirsutum* L.). Euphytica. 2015;201(2):195–213.

34. Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J. BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. Bioinformatics. 2004;20(14):2324–6.

35. Said JI, Knapka JA, Song M, Zhang J. Cotton QTLdb: a cotton QTL database for QTL analysis, visualization, and comparison between *Gossypium hirsutum* and *G. hirsutum × G. barbadense* populations. Mol Genet Genomics. 2015;290(4):1615–25.

36. Said JI, Song M, Wang H, Lin Z, Zhang X, Fang DD, Zhang J. A comparative meta-analysis of QTL between intraspecific *Gossypium hirsutum* and interspecific *G. hirsutum × G. barbadense* populations. Mol Genet Genomics. 2014;290(3):1003–25.

37. Tan ZY, Fang XM, Tang SY, Zhang J, Liu DJ, Teng ZH, Li L, Ni HJ, Zheng FM, Liu DX et al. Genetic map and QTL controlling fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Euphytica. 2015;203(3):615–28. doi:10.1007/s10681-014-1288-9.

38. Shao QS, Zhang FJ, Tang SY, Liu Y, Fang XM, Liu DX, Liu DJ, Zhang J, Teng ZH, Paterson AH et al. Identifying QTL for fiber quality traits with three upland cotton (*Gossypium hirsutum* L.) populations. Euphytica. 2014;198:43–58.

39. Hallauer AR, Carena MJ, Miranda Filho JB. Quantitative Genetics in Maize Breeding. New York: Springer; 2010. p. 664.

40. Rong JK, Abbey C, Bowers JE, Brubaker CL, Chang C, Chee PW, Delmonte TA, Ding XL, Garza JJ, Marler BS et al. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). Genetics. 2004;166(1):389–417.

41. Xian-Liang S, Xue-Zhen S, Tian-Zhen Z. Segregation distortion and its effect on genetic mapping in plants. Chin J Agric Biotechnol. 2006;3(03):163–9.

42. Zeng ZB. Precision mapping of quantitative trait loci. Genetics. 1994;136(4):1457–68.

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

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

45. Zhang ZS, Hu MC, Zhang J, Liu DJ, Zheng J, Zhang K, Wang W, Wan Q. Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Mol Breeding. 2009;24(1):49–61.

46. Zhang L, Wang S, Li H, Deng Q, Zheng A, Li S, Li P, Li Z, Wang J. Effects of missing marker and segregation distortion on QTL mapping in F2 populations. Theor Appl Genet. 2010;121(6):1071–82.

47. Flagel L, Udall J, Nettleton D, Wendel J. Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. BMC Biol. 2008;6:16.

Jamshed *et al. BMC Genomics* (2016) 17:197

Page 13 of 13

48. Yu J, Zhang K, Li S, Yu S, Zhai H, Wu M, Li X, Fan S, Song M, Yang D et al. Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* x *Gossypium barbadense* backcross inbred line population. Theor Appl Genet. 2013;126(1):275–87.

49. Yu Y, Yuan D, Liang S, Li X, Wang X, Lin Z, Zhang X. Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC1 population between *Gossypium hirsutum* and *G. barbadense*. BMC Genomics. 2011;12(1):1–14.

50. Zhang Z, Li J, Muhammad J, Cai J, Jia F, Shi Y, Gong J, Shang H, Liu A, Chen T, Ge Q, Palanga KK, Lu Q, Deng X, Tan Y, Li W, Sun L, Gong W, Yuan Y. High Resolution Consensus Mapping of Quantitative Trait Loci for Fiber Strength, Length and Micronaire on Chromosome 25 of the Upland Cotton (*Gossypium hirsutum* L.). PloS ONE. 2015;10(8):e0135430.

51. Su CF, Lu WG, Zhao TJ, Gai JY. Verification and fine-mapping of QTLs conferring days to flowering in soybean using residual heterozygous lines. Chinese Sci Bull. 2010;55(6):499–508.

52. Ning ZY, Chen H, Mei HX, Zhang TZ. Molecular tagging of QTLs for fiber quality and yield in the upland cotton cultivar Acala-Prema. Euphytica. 2014;195(1):143–56.

53. Paterson AH, Brubaker CL, Wendel JF. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. Plant Mol Biol Rep. 1993;11(2):122–7.

54. van Ooijen JW. JoinMap 4.0: Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Wageningen: Kyazma B.V; 2006.

55. Kosambi DD. The estimation of map distance from recombination values. Ann Eugen. 1944;12:172–5.

56. Liu D, Liu F, Shan X, Zhang J, Tang S, Fang X, Liu X, Wang W, Tan Z, Teng Z et al. Construction of a high-density genetic map and lint percentage and cottonseed nutrient trait QTL identification in upland cotton (*Gossypium hirsutum* L.). Mol Genet Genomics. 2015;290(5):1683–700.

57. Wang S, Basten CJ, Zeng ZB. Windows QTL Cartographer 2.5. Raleigh: Department of Statistics, North Carolina State University; 2012. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.

58. Blenda A, Fang DD, Rami JF, Garsmeur O, Luo F, Lacape JM. A high density consensus genetic map of tetraploid cotton that integrates multiple component maps through molecular marker redundancy check. Plos ONE. 2012;7(9):e45739.