



Article

QTL Mapping of Fiber Quality and Yield-Related Traits in an Intra-Specific Upland Cotton Using Genotype by Sequencing (GBS)

Latyr Diouf^{1,2,†}, Richard Odongo Magwanga^{1,3,†}, Wenfang Gong¹, Shoupu He¹, Zhaoe Pan¹, Yin Hua Jia¹, Joy Nyangasi Kirungu¹ and Xiongming Du^{1,*}

¹ State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455000, China; latyr@cricaas.com.cn (L.D.); richard@cricaas.com.cn (R.O.M.); gongwenfang@caas.cn (W.G.); heshoupu@caas.cn (S.H.); panzhaoe@caas.cn (Z.P.); jiayinhua@caas.cn (Y.H.J.); nyangasijoy@yahoo.com (J.N.K.)

² Senegalese River Valley Development Agency (SAED), Saint-Louis Bp74, Senegal

³ School of Physical and Biological Sciences (SPBS), Jaramogi Oginga Odinga University of Science and Technology (JOUST), Main Campus, P.O. Box 210-40601, Bondo, Kenya

* Correspondence: duxiongming@caas.cn; Tel.: +86-37-22562252

† These authors contributed equally to this work.

Received: 20 December 2017; Accepted: 28 January 2018; Published: 1 February 2018

Abstract: Fiber quality and yield improvement are crucial for cotton domestication and breeding. With the transformation in spinning techniques and multiplicity needs, the development of cotton fiber quality and yield is of great importance. A genetic map of 5178 Single Nucleotide Polymorphism (SNP) markers were generated using 277 F_{2:3} population, from an intra-specific cross between two upland cotton accessions, CCRI35 a high fiber quality as female and Nan Dan Ba Di Da Hua (NH), with good yield properties as male parent. The map spanned 4768.098 cM with an average distance of 0.92 cM. A total of 110 Quantitative Traits Loci (QTLs) were identified for 11 traits, but only 30 QTLs were consistent in at least two environments. The highest percentage of phenotypic variance explained by a single QTL was 15.45%. Two major cluster regions were found, cluster 1 (chromosome17-D03) and cluster 2 (chromosome26-D12). Five candidate genes were identified in the two QTL cluster regions. Based on GO functional annotation, all the genes were highly correlated with fiber development, with functions such as protein kinase and phosphorylation. The five genes were associated with various fiber traits as follows: *Gh_D03G0889* linked to qFM-D03_cb, *Gh_D12G0093*, *Gh_D12G0410*, *Gh_D12G0435* associated with qFS-D12_cb and *Gh_D12G0969* linked to qFY-D12_cb. Further structural annotation and fine mapping is needed to determine the specific role played by the five identified genes in fiber quality and yield related pathway.

Keywords: QTL mapping; fiber quality; upland cotton; intra-specific; yield related traits; gene ontology; genotype by sequencing

1. Introduction

Cotton is one of the most important natural fibers and oil crops in the world. Its annual global market value was estimated to be \$630.6 billion in 2011 [1]. Cotton fiber is the primary raw material in the textile industry [2]. The advancements in techniques and diversified methods of spinning have made cotton fiber quality and related yield traits of paramount significance in breeding and production of cotton [3]. Fiber quality is determined by a number of factors such as fiber strength, fiber length, fiber micronaire and fiber color, while yield is mainly determined by lint quantity [4]. However, lint yield and fiber quality have been found to be negatively correlated [5,6], which has long been a critical issue in cotton breeding [7]. Recently, Shang et al. [8] identified 20 QTLs for

fiber quality-related traits, however, four QTLs were validated. Moreover, five fiber quality traits were linked to 59 QTLs in an earlier report across five environments [9]. So far, few numbers of QTLs have been employed in marker-assisted selection (MAS) which is one of the enhanced breeding methods [10]. In all the identified and documented QTLs related to fiber and yield traits, most of them have been localized in a wide range of genomic regions and are often not stable across a wide genetic backgrounds [11]. Therefore, a dense interspecific map was generated, which included 2316 loci on the 26 cotton chromosomes in order to reduce and enhance accuracy in mapping [12]. However, these maps developed from interspecific hybridization have limited use in breeding due to limitation in controlling defective genes [2,5].

To overcome the inefficiency of maps developed from interspecific hybridization, it is therefore imperative to generate molecular maps based on an intraspecific population due to their ability to reduce the wide genome gap [2]. The employment of molecular marker techniques in cotton breeding through MAS and more advanced approaches such as genomic selection (GS) [13] would help break the bottleneck and, in turn, development of genetically advantaged genotypes. A small part of a DNA can be archived by reducing the complexity of the genome by restriction enzymes, such as genotyping-by-sequencing (GBS), the reduced-representation libraries (RRLs), restriction-site-associated DNA sequencing (RAD-seq) and next generation sequencing (NGS) [14].

The next-generation sequencing (NGS) of crop plant genomes have transformed the field of plant breeding. In the recent past, a lot of data generated has facilitated the discovery and use of large scale of single nucleotide polymorphisms (SNPs) in different genomes [15,16]. One of which was, genotype by sequencing (GBS), which holds the potential to narrow down the genotyping gap between references of large interest and mapping or breeding populations of local or specific interest [17]. GBS protocol techniques with their sample multiplicity have kept molecular research costs low while their output has diverse applications in many research areas, ranging from gene discovery to genomic-assisted breeding [18]. The ability of generating large amounts of unbiased markers in an inexpensive methods, has enabled GBS to become a more attractive approach to genotype and to construct high-resolution genetic maps, genomic selection and facilitated QTL mapping [19].

Mapping of QTLs has become an important technique to facilitate quantitative trait research and has been largely used in agricultural crops to map a number of beneficial agronomic traits including fiber quality and related yield traits.

In this investigation, a genetic map of 5178 SNP markers was generated using a 277 F_{2:3} intraspecific population developed from two tetraploid upland cotton accessions, mainly cultivated in China. CCRI35 with good fiber quality as female parent and Nan Dan Ba Di Da Hua (NH) known for high yield fiber as male parent. The map generated was employed to analyze QTLs related to fiber quality and yield related traits using QTL cartographer [20]. The aim of this study was to identify QTLs related to fiber quality, yield component traits, localize their position within the cotton genome and to identify the genes tightly linked to those QTLs. Findings of this research could provide valuable insights for breeders to develop cultivars with both traits, yield and quality fiber and enhance selection in cotton breeding.

2. Results

2.1. Phenotypic Variation between the Two Parents

In the determination of phenotypic variation of the 11 measured traits, Boll weight (BW), lint percentage (LP), fiber reflectance (FR), fiber yellowness (FY), spinning consistency index (SCI), and mature index (MI) were not used in analysis of the phenotypic variation between the parental lines due to the huge missing data throughout the phenotyping periods. The five traits used were fiber length (FL), fiber uniformity (FU), fiber strength (FS), fiber micronaire (FM) and fiber elongation (FE). FL, FU and FS showed significant differences between the parental lines. All traits were higher in CCRI35 than NH with exemption of FE which was higher in NH. In addition, no significant difference

was noted between the two parental lines for fiber micronaire (FM) and fiber elongation (FE), Figure 1. However, there was a wide range of phenotypic variation among the F_{2,3} population, with respect to all the measured traits; BW, LP, FL, FU, FM, FS, FE, FR, FY, SCI and MI. Across the three environments, 2014, 2015 and 2016 all the traits showed normal segregation with normalized distribution patterns (Figure 2).

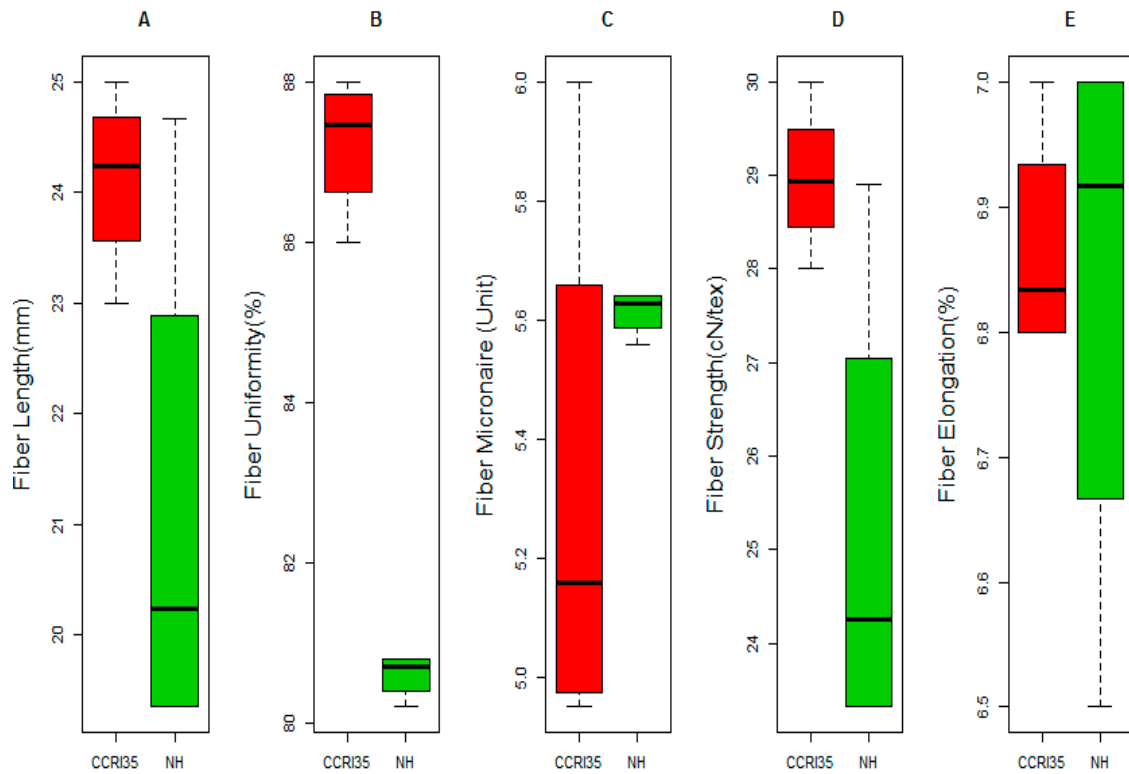


Figure 1. Phenotypic analysis of the two parents for fiber quality and yield-related traits; (A) Fiber Length (mm); (B) Fiber Uniformity (%); (C) Fiber Micronaire (Unit), (D) Fiber Strength (cN/tex); (E) Fiber Elongation (%). CCR135: female parent, NH: male parent.

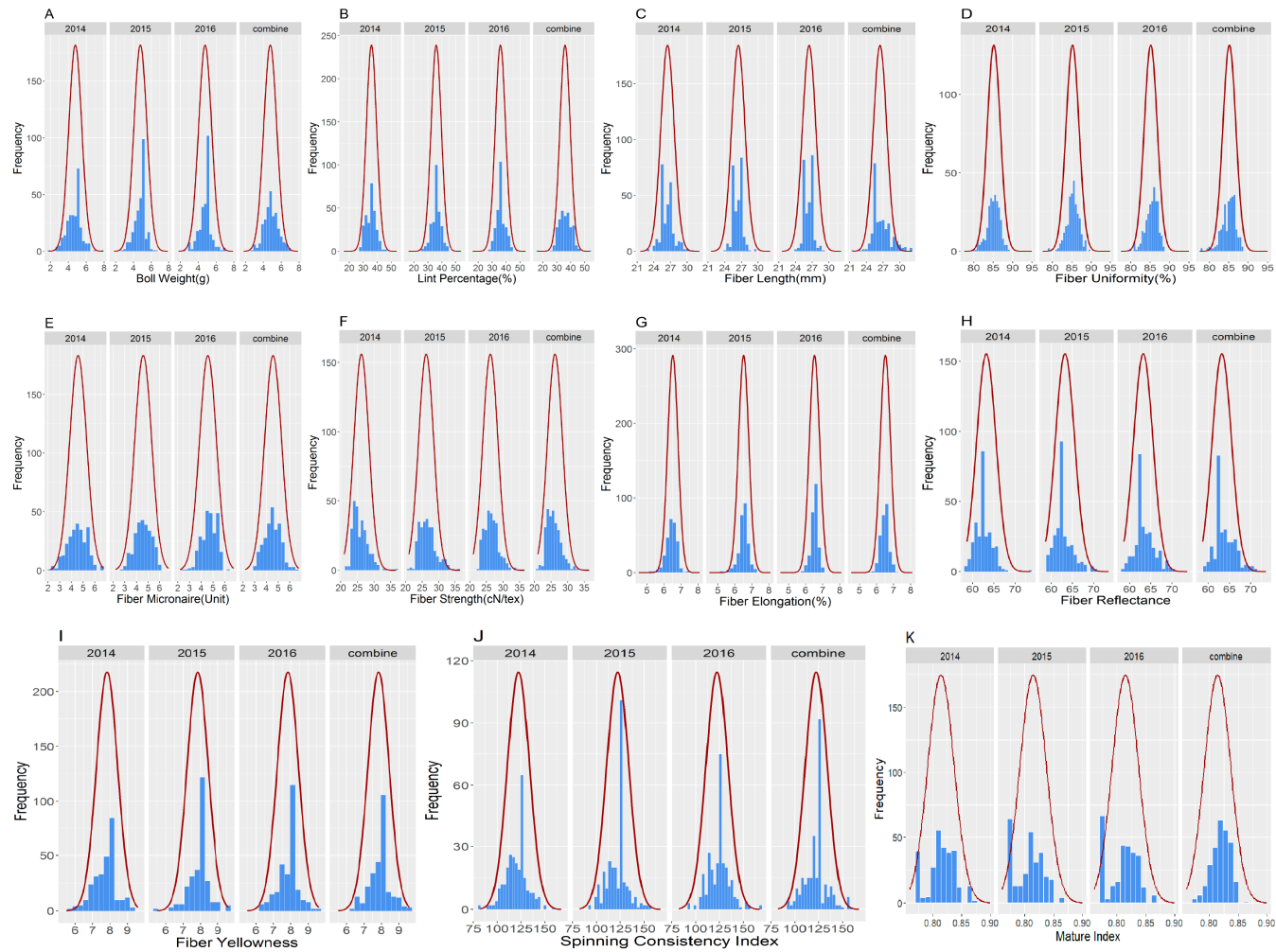


Figure 2. Frequency distribution of the 11 traits for fiber quality traits in F_{2,3}, (A) Boll Weight (g); (B) Lint Percentage (%); (C) Fiber Length (mm); (D) Fiber Uniformity (%); (E) Fiber Micronaire (Unit); (F) Fiber Strength (cN/tex); (G) Fiber Elongation (%); (H) Fiber Reflectance; (I) Fiber Yellowness; (J) Spinning Consistency Index; (K) Mature Index.

2.2. Correlation Analysis

To determine the correlations among different traits, a Pearson's correlation coefficient on yield-related and fiber quality trait was done using "Performance Analytics" package with Chart correlation function in R software version 3.4.2 [21]. Significant and positive correlations were noted between: BW with FL, FU, FM, FS, FE, FR, and MI; LP with FM and MI; FL with FU, FS, FE, FR, and SCI; FU with FS, FE, and SCI; FM with MI; FS with FE and SCI; FE with SCI and finally FR with SCI. Negative correlations were observed between: LP with FR and SCI; FL with FM; FM with FR and SCI; FY with SCI and finally SCI with MI (Figure 3). However, no significant correlation was noted between the other traits.

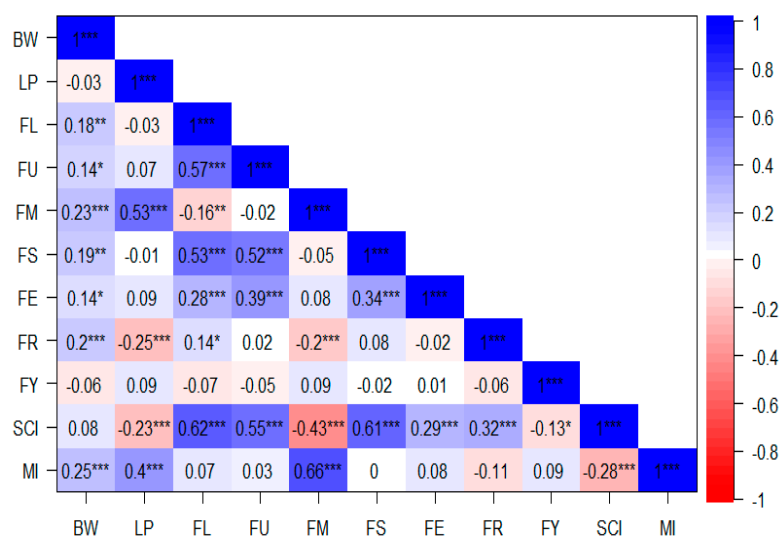


Figure 3. Pearson's correlation of the 11 traits for the $F_{2.3}$ in three environments. *, **, ***: significant levels of 0.5, 0.01 and 0.001 respectively. BW: Boll weight; LP: lint percentage; FL: fiber length; FU: fiber uniformity; FM: fiber micronaire; FS: fiber strength; FE: fiber elongation; FR: fiber reflectance; FY: fiber yellowness; SCI: spinning consistency index; MI: mature index. For the units, see in Figures 1 and 2.

2.3. ANOVA, Broad Sense Heritability and Phenotypic Analysis of Fiber Quality for the Two Parents and the $F_{2.3}$ Population

The ANOVA result revealed significant differences between the genotypes, environment and their interactions for all the traits (Table 1).

The broad sense heritability was much higher for the fiber quality traits as opposed to yield-related traits. The highest broad sense heritability was observed with fiber micronaire (FM), with 92.4% while the lowest broad sense heritability was observed in fiber elongation (FE) with 61.8%.

Table 1. ANOVA, broad sense heritability and phenotypic analysis of fiber quality and yield related traits for the two parents and the F_{2,3} population.

Trait	Source	DF	SS	MS	F	Pr > F	H _b (%)	P1	P2	P1 – P2	F _{2,3}					
											Mean	SD	Max	Min	Skew	Kurt
BW	e	3	327.4	109.1	1.9×10^{10}	<0.0001	67.5	-	-		4.77	0.73	7.9	1.8	-0.16	0.94
	g	276	728.6	2.6	4.59×10^{08}	<0.0001										
	g ² e	828	710.5	0.9	1.49×10^{08}	<0.0001										
LP	e	3	1856.7	619	9.5×10^{10}	<0.0001	82.1	-	-		35.97	3.71	54	16.08	0.09	1.2
	g	276	28,244	102.3	1.57×10^{10}	<0.0001										
	g ² e	828	15,200.7	18.4	2.82×10^{09}	<0.0001										
FL	e	3	1561.4	521	9.23×10^{24}	<0.0001	63.6	24.12	21.12	3	26.46	1.2	32.2	21.55	0.94	0.55
	g	276	1530.1	5.5	9.83×10^{22}	<0.0001										
	g ² e	828	1670.7	2	3.58×10^{22}	<0.0001										
FU	e	3	1200.6	400	9×10^{23}	<0.0001	77.1	87.23	80.6	6.63	85.14	1.69	89.3	77.8	-0.75	0.12
	g	276	4844.2	17.6	3.95×10^{22}	<0.0001										
	g ² e	828	3321.3	4	9.02×10^{21}	<0.0001										
FM	e	3	12.700	4	2.19×10^{25}	<0.0001	92.4	5.61	5.32	0.29	4.57	0.72	6.75	2.2	-0.19	-0.15
	g	276	1406.5	5.1	2.63×10^{25}	<0.0001										
	g ² e	828	320.3	0.4	1.99×10^{24}	<0.0001										
FS	e	3	2038.7	680	3.37×10^{25}	<0.0001	76.8	28.97	25.18	3.79	26.19	2.27	36.5	21	0.76	0.75
	g	276	8851.8	32.1	1.59×10^{24}	<0.0001										
	g ² e	828	6155.4	7.4	3.69×10^{23}	<0.0001										
FE	e	3	10.800	4	2.15×10^{24}	<0.0001	61.8	6.83	6.87	-0.04	6.51	0.3	8.1	4.5	-1.02	0.02
	g	276	137.1	0.5	2.96×10^{23}	<0.0001										
	g ² e	828	157.5	0.2	1.13×10^{23}	<0.0001										
FR	e	3	335.500	112	5.58×10^{22}	<0.0001	84.8	-	-		63.23	2.27	73.6	58.7	0.9	1.22
	g	276	11,577.9	41.9	2.09×10^{22}	<0.0001										
	g ² e	828	5282.3	6.4	3.18×10^{21}	<0.0001										
FY	e	3	38.2	12.7	8.03×10^{23}	<0.0001	86.3	-	-		7.83	0.61	9.6	5.5	-0.32	1.19
	g	276	845.6	3.1	1.93×10^{23}	<0.0001										
	g ² e	828	348	0.4	2.65×10^{22}	<0.0001										
SCI	e	3	24,493.6	8164.5	1.39×10^{25}	<0.0001	86.6	-	-		122.26	11.57	167	82	0.09	0.97
	g	276	299,967	1086.8	1.85×10^{24}	<0.0001										
	g ² e	828	120,302.4	145.3	2.48×10^{23}	<0.0001										
MI	e	3	0.5	0.2	6.73×10^{25}	<0.0001	66.7	-	-		0.81	0.02	0.87	0.75	-0.19	-0.79
	g	276	0.7	0	1.18×10^{24}	<0.0001										
	g ² e	828	0.5	0	2.77×10^{23}	<0.0001										

P1 = CCRI35: parental female with good fiber quality traits; P2 = NH: good yield fiber; DF: degree of freedom; SS: sum square; MS: mean square; F: F value; H_b (%): Broad sense heritability percentage; SD: Standard deviation; Min: Minimum; Max: Maximum; Skew: Skewness; Kurt: Kurtosis, (<0.0001): means significant at level $p < 0.001$.

2.4. GBS Genotyping, SNP Detection and Annotation

The genotypic data for the entire population was developed by use of the genotyping by sequencing (GBS) technique. Fifteen (15) individuals of each of the parents were sequenced and mapped on to the reference genome, which we obtained from the cotton research institute (available online: <http://mascotton.njau.edu.cn>). We obtained a total of 20,542,731 and 20,244,825 reads for CCRI35 and NH, respectively. An average of 80,372 and 112,128 SNPs were eventually identified for the female parent (CCRI35) and the male parent (NH), respectively, with an enzyme digestion efficiency of 99%. In genotyping the $F_{2.3}$ population, the enzyme efficiency was slightly lower compared to its efficiency in the parents, with efficiency of 98.9%. The overall mapped reads for the population and the two parents were 1,507,193,217, with an average of 4,909,424.16 mapped reads per individual which correspond to nearly 180.889 Gb of clean bases. The clean reads obtained were equivalent to 80.42-fold haploid genome coverage of raw paired-end Illumina reads by sequencing whole genome shotgun (WGS) libraries of homozygous cv. "TM-1" compared to Li et al. [22] in their study which generated a total of 445.7 Gb of clean reads translating to about 181-fold haploid genome coverage of raw paired-end Illumina reads by sequencing whole genome shotgun (WGS) libraries of homozygous cv. "TM-1" with fragment lengths ranging from 250 bp to 40 kb. The average GC content of the sequences was 38.25%, with a Q20 score of 94.66%. The parental lines were genotypes such as AC and AA, in which the female parent CCRI35 was heterozygous while the male parent (NH) was homozygous. The total resulting SNPs markers were 103,381 markers which were used to carry out further analysis. We assessed the distribution of the alleles across the $F_{2.3}$ population, and those markers which had a coverage threshold of 75% were filtered out, eventually, 34,090 markers were used. Markers with significant distortion ($p < 0.001$) were filtered and 6405 markers were retained with the purpose of determining bin markers.

2.5. Construction of the Linkage Maps

In the construction of the linkage groups, we used 6405 markers (Table S1) and phenotypic data of the $F_{2.3}$ population developed from an intra-specific cross of two tetraploid upland cottons were utilized for developing the intra-specific linkage map. A total of 5178 GBS markers were used for mapping the $F_{2.3}$ population, all the distorted markers were filtered out, the linkage groups were generated by the use of Join Map 4.0 [23]. Twenty six (26) LGs were generated from 5178 markers (Figure 4A and Figure S1, Table 2 and Table S2). Markers in linkage groups were ordered, rippled, and re-ordered according to pairwise recombination fractions, LOD scores (Logarithm of Odds) and linkage group length (Figure 4B). The 26 LGs were designated as A01 to A13 for A_t sub-genome and D01 to D13 for D_t sub-genome. The map generated had a map distance of 4768.098 cM, higher than the most current upland cotton linkage map with a map distance of 4450 cM [24]. The average distance between adjacent markers was 0.92 cM, the marker distances were narrowed in the map generated compared to earlier maps with 1.7 cM between adjacent markers [24]. The A_t sub-genome spanned 2611.43 cM, with a total of 3313 markers in the 13 linkage groups, with an average distance of 0.79 cM, while in D_t sub-genome, thirteen linkage groups comprised of 1865 markers spanning a distance of 2156.67 cM, with an average of 1.156 cM. The maximum gap between adjacent loci was 26.598 cM and 30.082 cM in A_t and D_t respectively, affirming the genome lengths between A_t and D_t [24] (Table 2). Chromosomes; A02, D02, A01, A05, A03, D01 and A10 exhibited higher marker loci with higher recombination frequency compared to the rest of the chromosomes such as D06 and D13 (Figure 4A,B). The chromosome with the highest marker loci was chromosome A02, 705 loci with map distance of 346.314 cM and an average distance of 0.49 cM, while the lowest marker loci was detected in chromosome D06 with only 16 markers, and a total length of 79.084 cM (Figure 4B).

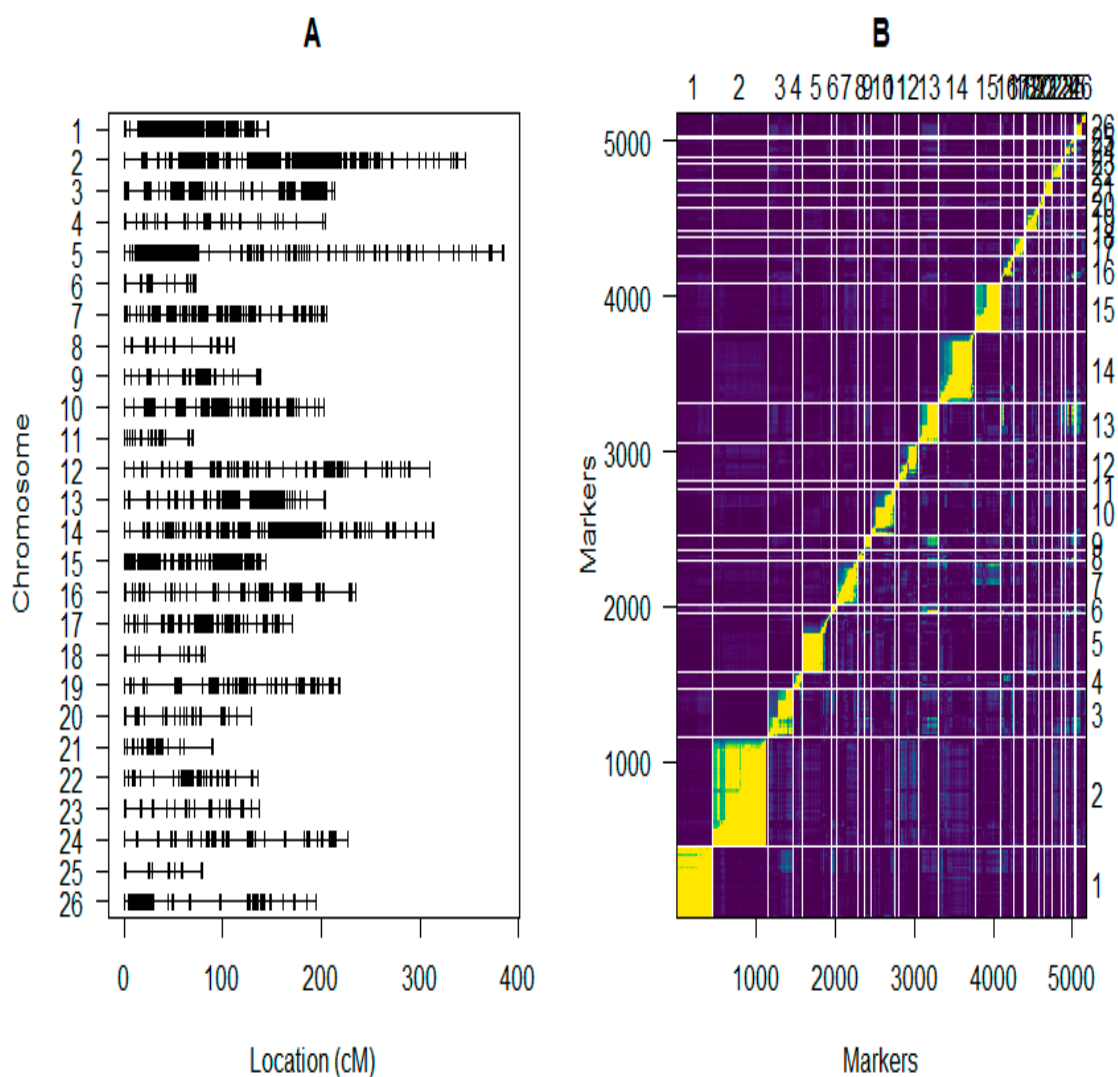


Figure 4. (A) Genetic linkage map constructed using the F_{2,3} Population; (B) Plot of estimated recombination fractions of all markers used in the F_{2,3} population. X and Y axis are the markers and Z is the linkage groups (LGs).

Table 2. Genomic distributions of SNPs markers.

Group	Marker Number	Map Length (cM)	Av Distance (cM)	Max Gap (cM)	<10 cM	>10 cM	Ratio
A01(c1)	448	146.704	0.33	8.505	447	0	1
A02(c2)	705	346.314	0.49	17.848	699	5	0.99
A03(c3)	323	213.937	0.66	17.145	319	3	0.99
A04(c4)	106	203.891	1.92	26.598	99	6	0.93
A05(c5)	378	385.092	1.02	21.198	365	12	0.97
A06(c6)	58	73.063	1.26	15.032	54	3	0.93
A07(c7)	279	205.892	0.74	11.622	276	2	0.99
A08(c8)	69	112.137	1.63	18.894	65	3	0.94
A09(c9)	98	138.501	1.41	19.234	95	2	0.97
A10(c10)	292	202.134	0.69	10.551	287	4	0.98
A11(c11)	51	70.548	1.38	23.241	49	1	0.96
A12(c12)	244	309.608	1.27	19.593	236	7	0.97
A13(c13)	262	203.61	0.78	17.425	256	5	0.98
Subtotal A _t	3313	2611.43	0.79	26.598	3247	53	0.98

Table 2. Cont.

Group	Marker Number	Map Length (cM)	Av Distance (cM)	Max Gap (cM)	<10 cM	>10 cM	Ratio
D01(c15)	319	144.092	0.45	6.351	318	0	1
D02(c14)	454	313.268	0.69	14.541	450	3	0.99
D03(c17)	133	170.555	1.28	14.993	131	1	0.98
D04(c22)	114	136.228	1.19	20.275	110	3	0.96
D05(c19)	153	218.788	1.43	27.062	148	4	0.97
D06(c25)	16	79.084	4.94	22.389	12	3	0.75
D07(c16)	169	235.366	1.39	26.041	161	7	0.95
D08(c24)	118	226.688	1.92	20.878	109	8	0.92
D09(c23)	40	136.744	3.42	14.48	33	6	0.83
D10(c20)	80	129.051	1.61	20.539	76	3	0.95
D11(c21)	98	89.782	0.92	27.564	95	2	0.97
D12(c26)	143	194.735	1.36	30.082	135	7	0.94
D13(c18)	28	82.286	2.94	20.917	25	2	0.89
Subtotal D _t	1865	2156.67	1.156	30.082	1803	49	0.97
TOTAL (A _t + D _t)	5178	4768.1	0.92	30.082	5050	102	0.98

Ratio: number of markers less than (<) 10 cM divided by total number of markers within chromosome. Av: Average; Max: Maximum.

2.6. Identification of Consistent and Clustering QTLs for Yield Related and Fiber Quality Traits

Thirty (30) QTLs were consistent among all the 110 QTLs identified for 11 traits in at least two environments (Table 3 and Figure S1). The 30 consistent QTLs were located on 16 chromosomes; A02 (2), A03 (1), A05 (2), A09 (3), A10 (2), A12 (1), D01 (1), D02 (1), D03 (4), D04 (1), D05 (2), D08 (2), D10 (2), D11 (1), D12 (4), and D13 (1). The distribution of the QTLs within the identified chromosomes, exhibited multiple position as illustrated in Figure S1 and Table S2 and Table 3. Of the 30 detected QTLs, 11 were localized on A_t sub-genome while the remaining 19 were mapped on the D_t sub-genome. The contributions of the parents toward the QTLs: 19 QTLs were linked to the good fiber quality parent (CCRI35) while only 11 QTLs were contributed by the high yield fiber parent (NH). Only 16 chromosomes out of 26 were found to harbor consistent QTLs for ten traits except MI (Mature Index) for yield-related and fiber quality (Table S2 and Table 3).

Four types of gene actions were revealed by the genetic effects of which one gene exhibited dominant effects (De), four partial dominances (PD), 20 over dominances (OD) and five additive effects (Ae). OD was observed for most of the traits in response to yield-related and fiber quality traits.

The highest percentage of phenotypic variance explained by a single QTL was 15.45%. The highest percentage of phenotypic variance was noted in lint percentage (LP), with a range of 10.03–15.46%. The distribution of the QTLs within the identified chromosomes, exhibited multiple positions in some chromosomes; A02, A03, A09, A10, D01, D03, D05, D08, D12, and D13 as illustrated in Table S2 and Table 3 and Figure S1. Moreover, a total of two important clusters with more than three traits per region, with high broad sense heritability and high percentage of phenotypic variation were identified as D03 (c17) and D12 (c26), which we designated as cluster 1 and cluster 2, respectively (Table 3, Figures 5 and 6).

Table 3. Cont.

Trait	QTL	Chr	Start Marker	End Marker	Start Marker (bp)	End Marker (bp)	Start Marker (cM)	End Marker (cM)	Position (cM)	LOD	Ae	De	d/a	GA	R ² (%)	DPE
FY	qFY-D03_15	D03	mk12142_D03	mk12159_D03	36,697,656	38,616,587	125.137	134.584	130.21	2.610206	0.1828	−0.0298	0.1630197	A	4.7288	CCRI35
	qFY-D03_15	D03	mk12154_D03	mk12155_D03	37,676,414	37,682,981	141.665	142.221	141.71	3.806732	0.184	−0.1844	1.0021739	D	6.9294	CCRI35
	qFY-D03_15	D03	mk12158_D03	mk12161_D03	37,938,158	39,407,242	145.337	150.198	148.31	4.210641	0.2038	−0.2276	1.1167812	D	8.0817	CCRI35
	qFY-D03_cb	D03	mk12109_D03	mk12111_D03	27,707,667	29,136,194	105.804	106.491	105.81	2.571118	0.0879	−0.0931	1.0591581	D	4.7115	CCRI35
LP	qLP-D03_15	D03	mk12142_D03	mk12159_D03	36,697,656	38,616,587	125.137	134.584	129.21	9.233442	1.7674	−0.2922	0.1653276	A	15.4584	CCRI35
	qLP-D03_15	D03	mk12152_D03	mk12160_D03	37,665,167	38,832,736	134.986	141.665	138.01	8.214984	1.6425	−0.399	0.2429224	PD	13.9321	CCRI35
	qLP-D03_16	D03	mk12152_D03	mk12160_D03	37,665,167	38,832,736	134.986	141.665	138.01	5.439739	1.4767	−0.9737	0.6593756	PD	9.8343	CCRI35
	qLP-D03_16	D03	mk12153_D03	mk12158_D03	37,668,262	37,938,158	143.823	145.337	144.81	4.621064	1.2806	−1.0766	0.8406997	D	7.9495	CCRI35
	qLP-D03_cb	D03	mk12152_D03	mk12160_D03	37,665,167	38,832,736	134.986	141.665	139.01	8.169381	1.2356	−0.1607	0.1300583	A	13.3867	CCRI35
	qLP-D03_cb	D03	mk12158_D03	mk12161_D03	37,938,158	39,407,242	145.337	150.198	147.31	7.090119	1.1317	−0.3152	0.278519	PD	12.0557	CCRI35
FR	qFR-D04_15	D04	MulMa448_D04	MulMa451_D04	49,867,798	50,187,192	4.736	8.337	5.01	2.660152	−0.1841	−1.5832	8.5996741	OD	0.0079	NH
	qFR-D04_cb	D04	MulMa448_D04	MulMa451_D04	49,867,798	50,187,192	4.736	8.337	5.01	2.781759	−0.227	−0.8557	3.7696035	OD	0.6158	NH
FM	qFM-D05_15	D05	mk12822_D05	mk12824_D05	30,214,244	30,216,527	152.95	153.434	153.01	5.061889	0.2063	0.008	0.0387785	A	6.8075	CCRI35
	qFM-D05_cb	D05	MulMa463-m_D05	mk12861_D05	30,373,354	31,354,896	144.668	148.809	146.71	4.021716	0.08	−0.1214	1.5175	OD	6.6096	CCRI35
	qFM-D05_cb	D05	mk12822_D05	mk12824_D05	30,214,244	30,216,527	152.95	153.434	153.41	4.627579	0.0886	−0.0879	0.9920993	D	7.3098	CCRI35
SCI	qSCI-D05_15	D05	MulMa463-m_D05	mk12861_D05	30,373,354	31,354,896	144.668	148.809	144.71	2.566775	−2.7495	3.5799	1.3020185	OD	4.5444	NH
	qSCI-D05_cb	D05	MulMa463-m_D05	mk12861_D05	30,373,354	31,354,896	144.668	148.809	144.71	2.579805	−1.778	1.3725	0.7719348	PD	4.7561	NH
FR	qFR-D08_15	D08	mk15992_D08	mk15995_D08	56,628,640	56,628,844	181.945	182.082	182.01	2.655809	0.587	−0.2932	0.4994889	PD	4.6002	CCRI35
	qFR-D08_cb	D08	mk15992_D08	mk15995_D08	56,628,640	56,628,844	181.945	182.082	182.01	3.441911	0.408	−0.2977	0.7296569	PD	6.2705	CCRI35
FL	qFL-D08_14	D08	MulMa514_D08	mk16004_D08	54,937,781	58,533,805	187.587	196.342	196.31	3.583062	0.3322	−0.0152	0.0457556	A	5.8804	CCRI35
	qFL-D08_cb	D08	MulMa514_D08	mk16004_D08	54,937,781	5,853,3805	187.587	196.342	196.31	3.639522	0.1799	0.0892	0.495831	PD	4.9405	CCRI35
	qFL-D08_cb	D08	mk16017_D08	mk16020_D08	59,691,087	59,698,388	208.553	208.76	208.61	3.255157	0.1595	0.1876	1.1761755	D	3.2928	CCRI35
FE	qFE-D10_14	D10	mk17141_D10	MulMa593-m_D10	56,817,432	56,887,380	102.161	106.55	106.21	2.523344	−0.0328	0.082	2.5	OD	3.7527	NH
	qFE-D10_15	D10	MulMa575-m_D10	MulMa581_D10	24,793,863	24,918,141	1.604	2.047	1.91	2.959826	0.0751	0.0306	0.4074567	PD	3.9966	CCRI35
FL	qFL-D10_14	D10	mk17141_D10	MulMa593-m_D10	56,817,432	56,887,380	102.161	106.55	105.21	3.411509	−0.1394	0.7384	5.2969871	OD	3.6287	NH
	qFL-D10_cb	D10	mk2492	MulMa366-m	519	33,037	78.155	98.694	83.21	2.614549	−0.045	0.464	10.311111	OD	2.0017	NH
FL	qFL-D11_16	D11	mk17462_D11	mk17463_D11	15,695,804	15,711,598	88.706	88.935	88.71	2.831705	0.0011	0.1765	160.45455	OD	0.8561	CCRI35
	qFL-D11_cb	D11	mk17464_D11	mk17514_D11	15,711,711	21,298,890	61.142	88.706	82.21	3.079262	−0.0461	0.5185	11.247289	OD	2.2603	NH
FS	qFS-D12_14	D12	mk18221_D12	MulMa605_D12	50,554,371	5,129,3378	148.99	160.984	153.01	3.270358	0.0077	0.8261	107.28571	OD	0.6079	CCRI35
	qFS-D12_15	D12	mk17994_D12	mk17997_D12	3,798,8313	3,8143,957	65.608	66.056	65.61	3.743757	0.3487	−0.7277	2.0868942	OD	5.7325	CCRI35
	qFS-D12_cb	D12	mk19853	mk17913_D12	101,319	13,479,261	16.392	16.946	16.41	4.049946	0.1807	−0.3037	1.6806862	OD	6.3446	CCRI35
FY	qFY-D12_15	D12	mk17995_D12	mk18057_D12	38,058,755	41,722,495	66.861	96.943	70.91	2.677524	0.047	−0.4721	10.044681	OD	2.0175	CCRI35
	qFY-D12_cb	D12	mk1009	mk17992_D12	18,989	37,732,030	66.348	66.861	66.41	3.252986	0.0211	−0.2739	12.981043	OD	1.9523	CCRI35
SCI	qSCI-D12_14	D12	mk18202_D12	mk18207_D12	48,411,387	48,718,111	132.168	133.627	133.61	3.14658	1.3176	−0.8128	0.6168792	PD	6.2241	CCRI35
	qSCI-D12_15	D12	mk19857	mk17916_D12	117,142	15,801,265	23.692	24.245	23.71	3.072747	2.636	4.6094	1.7486343	OD	1.4248	CCRI35
FL	qFL-D12_14	D12	mk18210_D12	mk18214_D12	48,923,084	49,133,419	135.788	138.113	135.81	2.851249	0.281	−0.216	0.7686833	PD	5.2886	CCRI35
	qFL-D12_cb	D12	mk18221_D12	MulMa605_D12	50,554,371	51,293,378	148.99	160.984	158.01	2.896851	0.1755	0.1739	0.9908832	D	3.2967	CCRI35
	qFL-D12_cb	D12	MulMa604-m_D12	mk18232_D12	51,286,859	52,905,207	161.79	172.395	165.81	2.773073	0.1653	0.1375	0.8318209	D	3.0451	CCRI35
FL	qFL-D13_14	D13	mk18377_D13	MulMa619_D13	4,171,037	34,392,983	36.052	56.969	46.11	3.237785	0.1336	−0.7846	5.8727545	OD	3.0376	CCRI35
	qFL-D13_15	D13	mk18378_D13	mk18379_D13	4,310,490	4,329,364	64.756	65.409	64.81	2.529859	0.2624	0.2724	1.0381098	D	2.1912	CCRI35
	qFL-D13_cb	D13	mk18516_D13	mk18533_D13	41,759,681	47,859,575	2.023	10.429	9.01	3.051031	0.17	−0.0572	0.3364706	PD	5.3187	CCRI35
	qFL-D13_cb	D13	mk20382	mk18378_D13	141	4,310,490	65.409	72.523	65.41	3.072747	0.1705	−0.0658	0.3859238	PD	5.4653	CCRI35

LOD: logarithm of odds; $0 < Ae$ (additive effect) < 0.20 ; $0.21 < PD$ (partial dominance) < 0.80 ; $0.81 < De$ (dominance effect) < 1.20 ; OD (over dominance) > 1.20 ; $|d/a| = De/Ae$; GA: gene action; DPE: direction of phenotypic explanation. For traits meaning see Figure 1 or Figure 2.

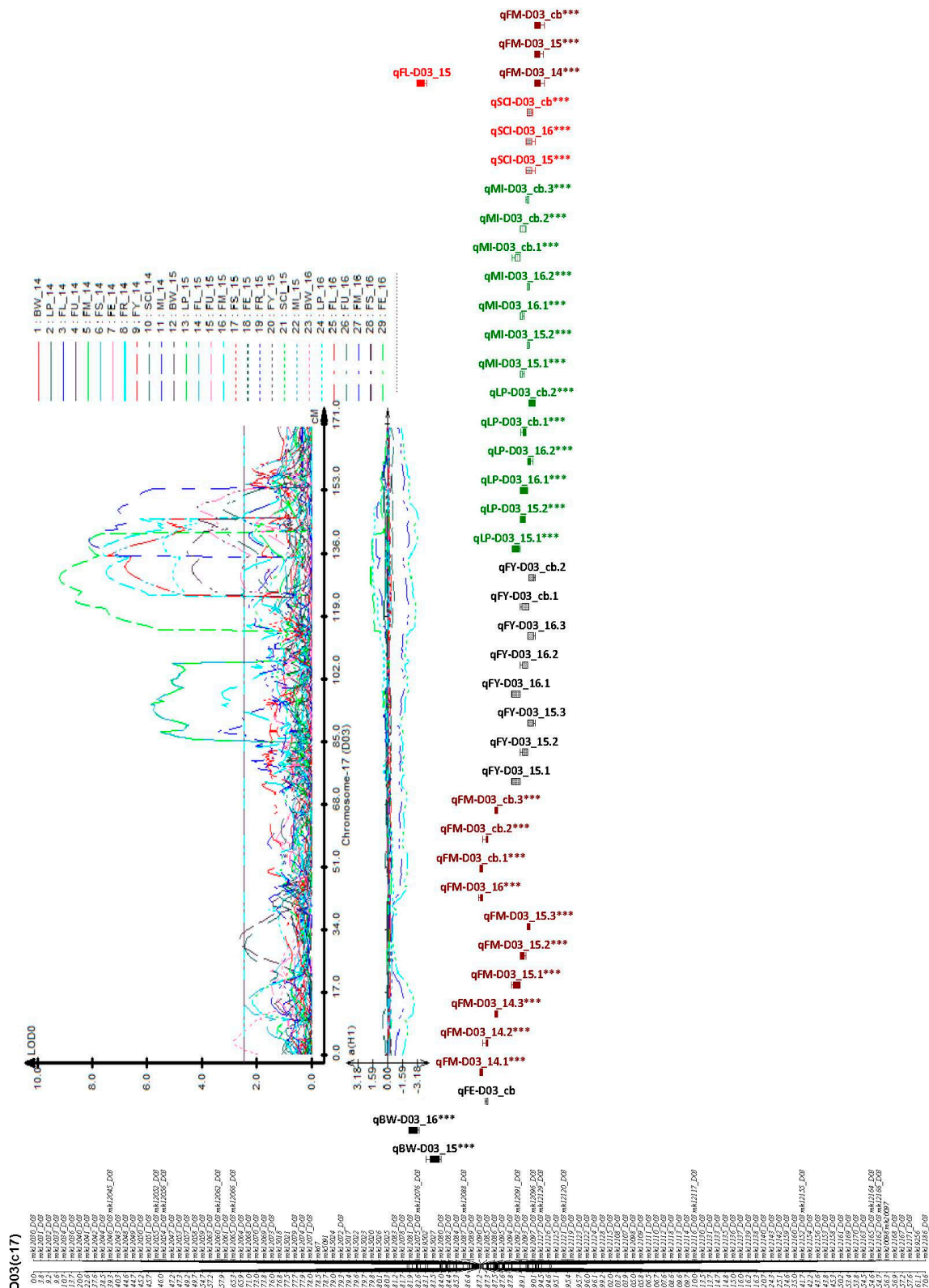


Figure 5. Clustered QTLs identified in D03 (c17) of yield-related and fiber quality traits. Bars and lines on the right-hand side of the linkage groups show the QTL likelihood intervals. Map distances in centiMorgan (cM) are indicated on the left-hand side of the linkage groups. For trait meanings, see Figure 1 or Figure 2, *** asterisk means the QTL is consistent.

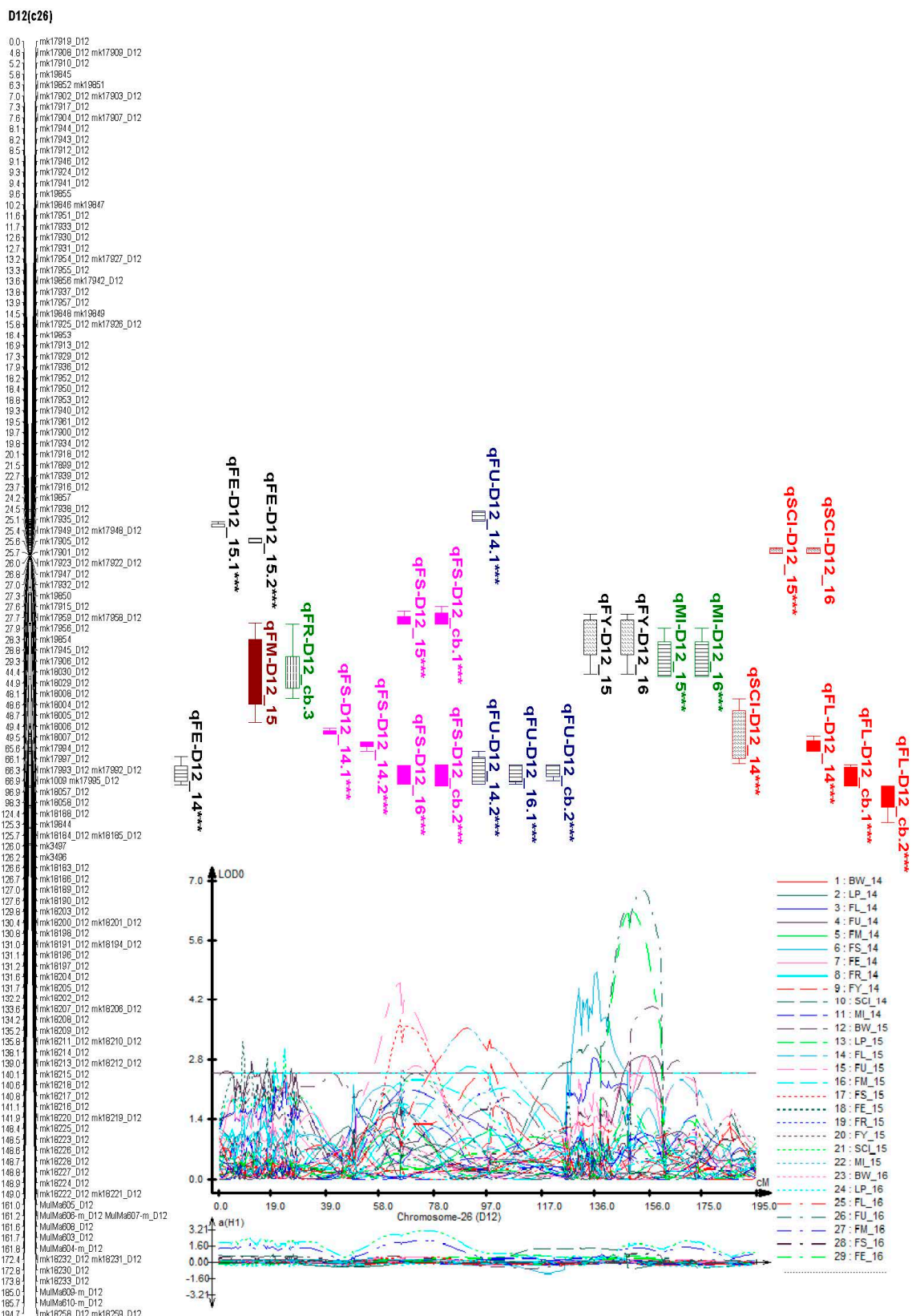


Figure 6. Clustered QTLs identified in D12 (c26) of yield-related and fiber quality traits. Bars and lines on the right-hand side of the linkage groups show the QTL likelihood intervals. Map distances in centiMorgan (cM) are indicated on the left-hand side of the linkage groups. For trait meanings, see Figure 1 or Figure 2, *** asterisk means the QTL is consistent.

2.7. The Gene Ontology Enrichment Analysis Based on QTL Clusters

Based on phenotype variation and QTL frequency, D_t-sub genome of the whole tetraploid chromosomes, harbored the highest number of stable QTLs with the highest level of phenotypic variation. In lieu of this, chromosome 17 (D03) and chromosome 26 (D12) had two clusters with four QTLs in each. Within the two cluster regions, we were able to mine the putative genes which could be having a role in fiber and yield-related traits. In cluster 1 (Chr17, D03), 136 genes were obtained, in which 14 were found to be highly expressed based on the RNA sequence while in cluster 2 (Chr26, D12), a total of 1280 genes were mined, out of which 153 were highly expressed at various stages of fiber development, 5 DPA, 10 DPA, 20 DPA and 25 DPA.

Moreover, in order to identify the set of the most robust candidate genes for yield-related traits and fiber quality; we mainly focused on the 153 highly expressed genes as obtained from “TM-1”_RNA-seq data (available online: <http://mascotton.njau.edu.cn>). Out of 153 highly expressed genes, five showed high level of expression across the various stages of fiber development, and therefore, the five genes could be the potential candidate genes with greater roles in the regulation of various fiber traits Table S3. Furthermore, all the five genes were localized in different positions of the genome: one gene (*Gh_D03G0889*) was located in cluster 1 (D03 (Chr 17)) within the marker mk12119_D03 (30,535,745 bp) to marker mk12123_D03 (30,566,883 bp), the trait localized in this region was fiber micronaire (FM); while the other three genes: *Gh_D12G0093*, *Gh_D12G0410*, and *Gh_D12G0435* were localized in cluster 2 (D12 (c26)) within the marker mk19853 (101,319 bp) to marker mk17913_D12 (13,479,261 bp), the trait localized in the genome region was fiber strength (FS). Finally, the fifth gene, *Gh_D12G0969* was also mapped in cluster 2 (D12 (Chr 26)), from marker mk1009 (18,989 bp) to marker mk17992_D12 (37,732,030 bp), the trait localized in that area was fiber yellowness (FY). Based on the expression profile and GO functional annotation, these five genes were therefore found to be the most robust and possibly the putative candidate genes for fiber quality and yield related traits (Table S3, Figures 7 and 8).

Based on GO enrichment analysis, the five highly up regulated genes were as follows: *Gh_D03G0889* was mainly involved in molecular function and biological processes, such as, up regulation of translational elongation (GO: 0003746), poly-A RNA binding (GO: 0003723), ribosome receptor activity (GO: 0043022), hypusine anabolism (GO: 0008612), translation elongation factor (GO: 0003746), regulation of translation elongation (GO: 0045901) and regulation of translation termination (GO: 0045905). The second gene, *Gh_D12G0093* was involved only in molecular function, protein amino acid binding (GO: 0005515). The third gene, *Gh_D12G0410* was involved in all the GO functional annotation, in biological process, it was mainly involved in translation elongation (GO: 0006414), molecular function, it was mainly involved in translation elongation factor activity (GO: 0003746) and protein binding (GO: 0005515) while in cellular component, it was found to be involved in eukaryotic translation elongation factor 1 complex (GO: 0005853). The fourth gene, *Gh_D12G0435*, had no functional annotation, however it was found to function in nucleoside diphosphate kinase activity and the last gene, *Gh_D12G0969*, functions both in biological process and molecular function, nucleoside diphosphate kinase activity (GO: 0004550), nucleoside diphosphate phosphorylation (GO: 0006165), GTP biosynthetic process (GO: 0006183), UTP biosynthetic process (GO: 0006228), CTP biosynthetic process (GO: 0006241) and ATP binding (GO: 0005524). In relation to gene action analysis, the five putative and robust genes with direct role in fiber development in cotton were all contributed by the female parent, CCRI35, known for its superior fiber quality (Table S3 and Figures 7 and 8). The five genes had similar sequences based on phylogenetic tree analysis; the same was affirmed by their expression profile and all from D_t-sub genome. High quality fiber attributes are highly linked to the D-genome of the diploid cotton such as *G. barbadense*, and being tetraploid cotton originated from the polyploidization of the A and D genomes of the diploid cotton.

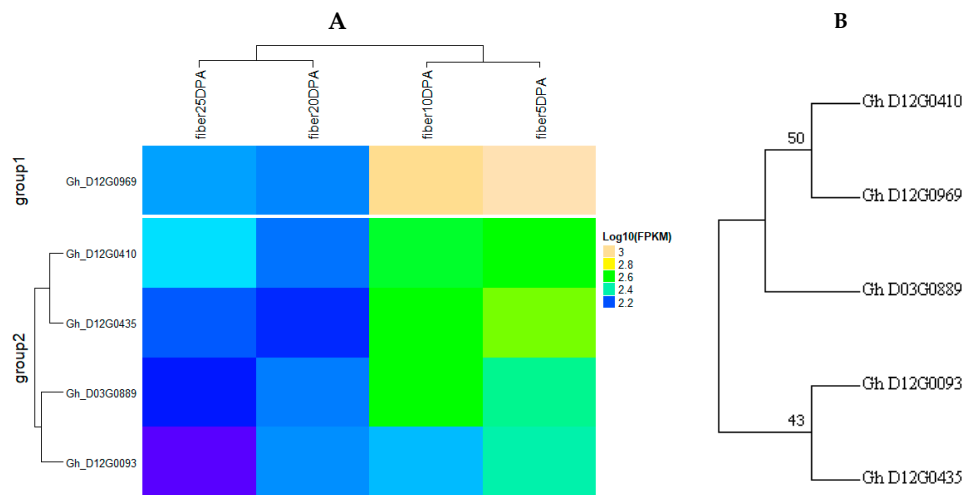


Figure 7. (A) Candidate genes involved in yield and/or fiber quality traits in this study; (B) phylogenetic tree analysis of the five involved genes in yield related and fiber quality traits of upland cotton.

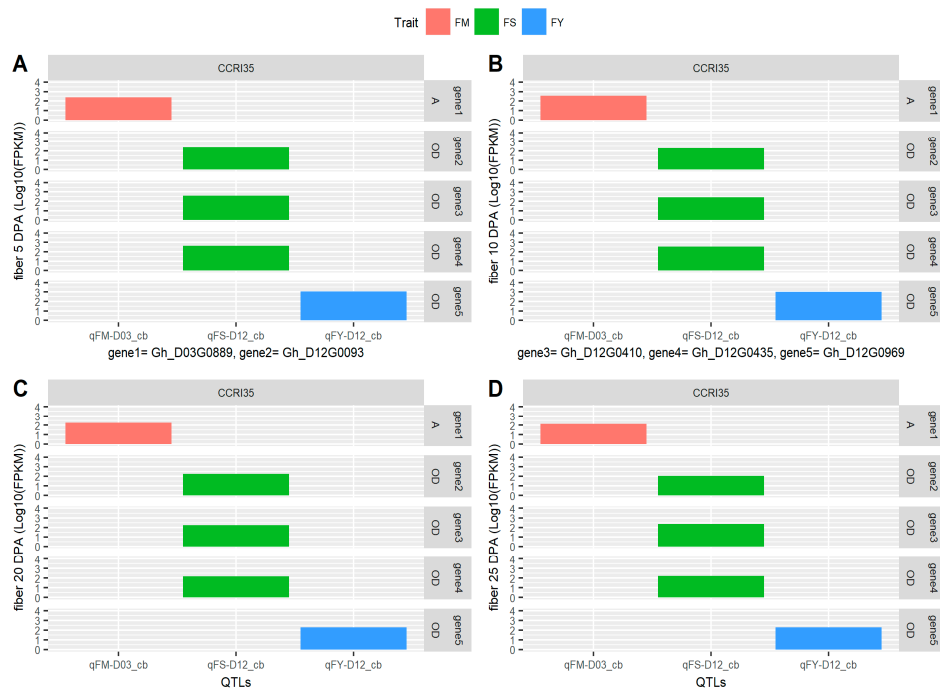


Figure 8. The five genes with highest expression in different stages of fiber development. *x* axis: QTLs for fiber Quality; CCR135: good fiber quality parental line; *y* axis: fiber DPA (Log10(FPKM)); FM: fiber micronaire (%); FS: fiber strength (cN/tex); FY: fiber yellowness; *Gh_D03* or *Gh_D12* are genes identified with high expression and involved in fiber development; cb: combine analysis; 0 < Ae (additive effect) < 0.20; OD (over dominance) > 1.20; (A–D) are respectively fiber expression at 5, 10, 20 and 25 DPA. For trait meanings, see Figure 1 or Figure 2.

3. Discussion

The determination of stable QTLs for superior agronomical traits and the construction of a high-resolution map are crucial for MAS. Several intra-specific genetic maps have been generated and used for QTL detection related to fiber and yield components [2]. Even though these maps have been used, they are limited in scope and accuracy due to huge marker intervals and narrow genome

coverage. The greatest impediment in the construction of a high-resolution map in intraspecific crosses is due to low rate of polymorphism within *G. hirsutum* and the presence of fixed homozygous genetic blocks [11,25]. Therefore, there is a need to find additional markers to fill in the gaps in the genetic map [11]. In this current research, a genetic map consisting of 5178 SNP markers obtained through the GBS technique was developed using a 277 F_{2:3} population derived from an intra-specific cross. In addition, the contrasting difference between the two parental lines used in this investigation could be explained based on inherent genetic characteristics. The male parent is known for superior agronomic traits such as early flowering and the ability to generate a high percentage of fruits with large size, while the female parent is known for superior fiber traits. Fiber length (FL), fiber uniformity (FU), fiber micronaire (FM), and fiber strength (FS) showed significant differences between the two parental lines. These traits were attributed to CCRI35, except FE which was linked to NH. There was no significant difference noted between the two parents for FE and FM. This result confirmed the good quality fiber trait of the female parent, CCRI35, compared to the male, NH.

In addition, there was a wide range of phenotypic variation among the F_{2:3} population, with respects to the following measured traits: BW, LP, FL, FU, FM, FS, FE, FR, FY, SCI, and MI. In the three environments, all traits exhibited normal segregation patterns, with equal distribution. The low absolute values for skewness and kurtosis showed that these traits had normal distribution. In addition, in the F_{2:3} population, the maximum phenotypic data values in all the variables were much higher than in CCRI35, the parent known for superior fiber traits, fiber length (FL), fiber uniformity (FU), fiber micronaire (FM), fiber strength (FS), and fiber elongation (FE). This finding showed that all traits were transgressively segregated in the F_{2:3} population. Previous research reported that transgression was the difference observed between the mapping parents of upland cotton [11,26–28].

Furthermore, positive correlations were noted between the following traits: boll weight (BW) with fiber length (FL), fiber uniformity (FU), fiber micronaire (FM), fiber strength (FS), fiber elongation (FE), fiber reflectance (FR), and mature index (MI); lint percentage (LP) with FM, FL with FU, FS, FE, and spinning consistency index (SCI); FU with FS, FE, and SCI; FM with MI; FS with FE, and SCI; FE with SCI and finally FR with SCI. However negative correlations were observed in the following traits: LP with FR, and SCI; FL with FM; FM with FR, and SCI; finally, SCI with MI. This result is consistent with those from Jamshed et al. [11] which showed that positive correlations were observed between: fiber elongation (FE), fiber length (FL), fiber strength (FS), and fiber uniformity (FU), with a significance level of 0.01. Moreover, FL and FS were both negatively correlated with fiber micronaire (FM). In this study, the correlations between FM with and FS were found to be negative but were not significant, which does not agree with previous findings. This deviation could be attributed to the population background used in this study.

It is known that broad sense heritability with high percentage is more useful and very easy to manipulate in MAS. Therefore, the extent of transmission of traits from the parents to the descendants or offspring was determined by level of heritability, hence traits with high broad sense heritability could be easier to manipulate [29]. The broad sense heritability was high for LP (82.65%), FM (91.68%), FR ((86.08%), FY (88.89%), SCI (87.47%), and moderate for BW (68.92%), FL (61.66%), FU (76.42%), FS (76.54%), FE (60.58%), and MI (76.61%). The lowest broad sense heritability was noted for fiber elongation (FE), 60.58%. Similar findings were observed with Jamshed et al. (2016) who found that fiber elongation had the lowest broad sense heritability (27%), whereas other fiber traits were higher, ranging from 80% (FU) to 93% (FL) [11,28].

A total map distance of 4768.098 cM was generated, higher than the most current linkage map with a map distance of the 4450 cM of cotton genome [24]. This is the densest intra-specific map developed in upland cotton. This map could be helpful for further studies in MAS, especially in fine mapping. The average distance of the adjacent markers was 0.92 cM. A_t sub-genome spanned 2611.43 cM, and consisted of 3313 markers with 13 LGs. The average marker distance in A_t sub-genome was 0.79 cM with a maximum gap of 26.598 cM of the adjacent markers. In D_t sub-genome, 13 LGs were assigned which comprised of 1865 markers spanning 2156.67 cM, with an average of 1.156 cM. The maximum

gap was 30.082 cM between adjacent loci. Due to the nature of upland cotton genome, mapping QTLs not only for fiber as in this research but for other agronomic traits has been difficult. This is because of the narrow genetic background, which resulted in low diversity of alleles with a significant role in fiber quality traits between two given varieties [30]. Therefore, only few QTLs could be mapped based on two parent crossing populations, which has been verified by previous reports [3,6,31–39]. In this current study, a total of 110 QTLs were identified for 11 traits, but only 30 QTLs were consistent in at least two environments. The 30 consistent QTLs were located on 16 chromosomes; A02, A03, A05, A09, A10, A12, D01, D02, D03, D04, D05, D08, D10, D11, D12, and D13 with 2, 1, 2, 3, 2, 1, 1, 1, 4, 1, 2, 2, 2, 1, 4, and 1 QTL respectively. Of the 30 detected QTLs, 11 were located on A_t sub-genome while the remaining 19 were located on the D_t sub-genome. This finding is consistent with previous reports in which 58 QTLs were found on the A_t sub-genome, whereas 107 QTLs were localized on the D_t sub-genome [11]. Fifty-eight QTLs were located on the A_t sub-genome (Chr01–Chr13), and 73 QTLs on the D_t sub-genome [25]. These QTLs explained from 2.03 to 16.85% of phenotypic variation, with an average of 6.26% explained in all five fiber quality traits [40].

Most of the QTLs distributed in the cotton genome revealed the complexity of the cotton genome and arduousness of QTL mapping in cotton. Therefore, comparing our QTLs with other QTLs mapped from previous studies could be of great help in determining the reliability of the QTLs detected [41]. Up to now, 4268 QTLs from 140 publications of cotton have been documented in the collected Cotton QTL Database (available online: <http://www2.cottonqtl.org:8081/index>). In this study, the GBS-SNP markers are unique and thus lack common identity with the SSR-based markers. However, five QTL clusters in this investigation were found to have a common bearing to those documented by Said et al. [42], which have been known as one of the strongest references in QTL mapping in recent years. The five QTL clusters were: cluster A07 was identical to c7-cluster-Gh × Gb-4:55–79 cM; cluster A08 had an approximate position of 4.81–110.81 cM, which was similar to c8-cluster-Gh-2:21–31 cM; cluster D01, had an approximate position of 2.21–139.31 cM, similar to c15-cluster-Gh-3:49–68 cM; cluster D02, had an approximate position of 0.01–206.11 cM, similar to c14-cluster-Gh-2:76–91 cM and lastly cluster D08, had an approximate position of 100.71–208.61 cM, nearby to c24-cluster-Gh-2:41–62 cM. The high correlation of the QTLs detected in this study to the previous finding, provides the opportunity for the utilization of these QTLs in MAS to improve the fiber quality of Upland cotton.

Gene Ontology enrichment analysis revealed five genes with very high expression and were linked to three fiber quality traits, FM, FS, and FY. Interestingly, the five genes took their alleles from the parental line known for superior fiber quality CCRI35. This result supported our study. Our findings provide an opportunity in the improvement for fiber quality especially fiber color (FY: fiber yellowness). Cotton fiber development occurs through various stages, namely fiber initiation, elongation, secondary cell wall formation and maturation [43]. Cotton fiber development is controlled by a multi-complex of genes interactions rather than a single gene effect [44,45]. *GhD12G0969* was mainly found to have a functional role in phosphorylation; phosphorylation is a process mediated by protein kinases to activate critical cellular pathways such as metabolism, cell division and cell differentiation during initiation stages in cotton fiber development [46]. In addition, *Gh_D12G0435* was found to be involved in kinase activity; protein kinase activity plays an important role in signal transduction through the phosphorylation process during cotton fiber development [47]. Therefore, the five highly up regulated genes could possibly be the key genes with major functional roles in fiber development and in turn superior quality as evident in the CCRI35, female parent.

4. Materials and Methods

4.1. Plant Materials, Growth Conditions and Trait Data Collection

The accessions used in this research were, Nan Dan Ba Di Da Hua in Chinese annotation, but for simplicity, we abbreviated the name as (NH), the male parent; it has moderate fiber quality traits but

high yielding in fiber [48,49]. The female parent was Zhong35, also with the Chinese name, was then abbreviated as CCRI35; it is known for high fiber quality traits but with moderate yield [9]. The parental lines and 277 F_{2.3} population were evaluated for fiber quality traits and yield components in Anyang research station (36°100' N, 114°350' E), Henan province, Yellow River. The field experiment was carried out during summer periods in three consecutive years, 2014, 2015 and 2016. The experimental layout adopted, was complete randomized block design (CRBD) with three replicates. The plot sizes were 5 m long with row spacing of 0.75 m. Fiber quality and yield component traits were collected following the laid down scheme as described by [41]. Fully opened bolls in each sampled plant were collected within the middle region of the plant, 25 bolls were collected from each line for fiber quality and yielded component determination. The balls were ginned for the determination of lint percentage (LP), fiber length (FL), fiber uniformity (FU), fiber micronaire (FM), fiber elongation (FE), fiber strength (FS), fiber reflectance (FR), fiber yellowness [50], spanning consistency index (SCI) and mature index (MI) by the HVI 900 fiber testing system, which was done in our cotton fiber quality testing unit, cotton research institute, Anyang, China. The test conditions were set with temperature at 20 °C and relative humidity of 65%.

4.2. Sample Collection, Library Preparation, Sequencing and SNP Genotyping

4.2.1. DNA Extraction, Quantification and Quality Determination

Fresh leaf samples were obtained from each line, together with the two parents and immediately frozen in liquid nitrogen then stored under −80 °C before DNA extraction. DNA of the F_{2.3} populations of 277 individuals and 10 samples for individual parents was extracted by the CTAB method as described by Paterson et al. [51]. Each sample was then crushed separately in liquid nitrogen to fine powder, then immediately added to CTAB solution. In every 100 mg ground tissues, we added 500 µL of CTAB Buffer. The samples were then shaken for 15 min then centrifuged. The centrifuged mixture was then put in a water bath at 60 °C for 30 min. Then, samples were centrifuged for 5 min at 12,000 revolutions per minute (rpm) for 5 min. After centrifuging, the supernatant transferred to a new tube. Then, 5 µL of RNase solution was added to digest RNA and then incubated for 20 min at 32 °C. Equal amount in volume of chloroform/isoamyl alcohol (24:1) was added then shaken for 5 s before centrifuging the samples for 1 min to separate the phases. We pipetted the aqueous upper phase to a new tube; the method was then redone until the upper phase was clear. The upper clear phase was then pipetted into a new tube. DNA samples were later precipitated by adding 70% by volume of ice-cold isopropanol and incubated for 15 min at −20 °C. The condensed DNA samples were then centrifuged at 12,000 rpm for 10 min. The supernatant was then decanted and subsequently washed with 500 µL ice cold 70% ethanol twice then absolute alcohol. DNAs were later dissolved in 20 µL TE buffer (10 mM Tris, pH 8, 1 mM EDTA) [52]. The degradation and contamination of DNA was checked through 1% agarose gels. The purity of DNA was determined by using a Nano Photometer[®] spectrophotometer (IMPLEN, Westlake Village, CA, USA). The ratio of absorbance at 260 and 280 nm was used to assess the purity of DNA. The DNA samples with the ratio of ~1.8 were then qualified as pure [53]. The concentration of DNA was done by using Qubit[®] DNA Assay Kit in Qubit[®] 2.0 Fluorimeter (Life Technologies, Carlsbad, CA, USA). The Qubit[®] dsDNA HS (High Sensitivity) Assay Kits make DNA quantitation easy and accurate. The kits contain concentrated assay reagent, dilution buffer, and prediluted DNA standards. The reagents were mixed with the buffer solution, and then added 1–20 µL of each DNA samples.

The concentrations were read using the Qubit[®] Fluorometer (Life Technologies, Carlsbad, CA, USA); only the DNA samples with concentration range of 10 pg/µL to 100 ng/µL were finally used (available online: https://tools.thermofisher.com/content/sfs/manuals/Qubit_dsDNA_HS_Assay_UG.pdf).

4.2.2. GBS Library Preparation, Sequencing and SNP Genotyping

GBS is a low cost and an efficient method of large-scale genotyping, which is based on high-throughput sequencing but with a reduced-representation library (RRL). The following were step by step processes in GBS technique; firstly, we carried out a GBS pre-design experiment to test the accuracy of the GBS protocol and quality of the output data. The enzymes and sizes of restriction fragments were examined by using training data. Three basic criteria were followed: (a) the suitability of the number of tags to the project needs; (b) the homogenous distribution of the enzymatic tags throughout the examined sequences; (c) elimination of redundant tags (repeated tags must be avoided). This was to ensure the effectiveness and accuracy of data obtained from GBS reads; 50 bp was the selection criterion to ensure sequence depth uniformity.

Secondly, we constructed the GBS library using the pre-designed scheme. The genomic DNA of the F_{2:3} population were incubated at 37 °C with MseI Restriction Enzyme obtained from New England BioLabs (Ipswich, MA, USA), NEB, T4 DNA ligase and ATP. MseI Y adapter N containing barcode. Restriction-ligation reactions were activated at 65 °C, followed by digestion for additional restriction enzyme NlaIII at a temperature of 37 °C. The samples were then purified by using Agencourt AMPure XP (Beckman, Brea, CA, USA). Then carried out polymerization chain reaction (PCR) using the purified samples, Phusion Master Mix universal primer and index primer were used to add index, complete i5 and i7 sequence. The Agencourt AMPure XP (Beckman) was used to purify the PCR products, which were pooled then ran through 2% agarose gels. Fragments with 375–400 bp (with indexes and adaptors) in size were obtained by using a Gel Extraction Kit (Qiagen, Hilden, Germany). The isolated fragment products were then purified using Agencourt AMPure XP (Beckman), and finally diluted for sequencing.

GBS analysis was strictly carried out as outlined by Elshire et al. (2011) [54]; integrating 3 of 96-well plates across 288 barcodes for library preparation and sequencing. For SNP calling, the raw sequence data for the 277 F_{2:3} population together with the F₁ generation was processed through the TASSEL 3.0 Genotype By Sequencing (GBS) pipeline [55] using the *Gossypium_hirsutum_v1.1.fa* as the reference genome [56] which was obtained from Cotton research institute (available online: <http://mascotton.njau.edu.cn/info/1054/1118.htm>), for alignment and the Burrows–Wheeler Aligner (BWA) mem [57] with default parameters. The output consisted of variant call format (VCF) file version 4.1 [58] including single nucleotide polymorphisms (SNPs) present in at least 40% of the progeny and with a minor allele frequency (MAF) 0.1. Subsequently, the data in variant call format (VCF) was filtered using VCF tools version.1.12a [58] and TASSEL [59] versions 3.0 and 4.0. A total of 93,384 single nucleotide polymorphisms (SNPs) were identified in 277 F_{2:3} population by TASSEL 3.0, then a custom filtering process was applied for alignment. The filtering was based on maintaining sites with a minimum read depth of 6% and 75% completeness by site across progeny and by progeny across sites. Results were obtained as a TASSEL hapmap file.

Finally, using a custom perl script marker heterozygous in the F₁ generations and with a co-dominant segregation ratio of 1:2:1 among the F_{2:3} population were identified using a chi-squared (χ^2) goodness-of-fit test at $\alpha \leq 0.01$. These were reconverted and imported in JoinMap[®] 4.1 for linkage group generation. A total of 26 LGs were obtained, each linkage group was assigned to its corresponding chromosome by using BLASTN-search (available online: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>), for the marker sequence.

4.3. Data Analysis and Linkage Map Construction

Analysis of variance (ANOVA) was performed by using field phenotype data of the three consecutive seasons 2014, 2015, 2016, and the combine analysis (cb). A mixed procedure was used; the genotypes and the environments were fixed as factors in order to detect the heritability [60]. Post hoc test (Turkey's) to compare means was done [60]. The broad-sense heritability percentage, H_b (%), was calculated for each trait using the formula described by [61].

$$H = \sigma^2G / \sigma^2G + (\sigma^2e/r)$$

With σ^2G is the genotypic variance; σ^2e : phenotypic variance and r : replication.

Most of the data were analyzed using R software version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria) [21]. Markers were ordered, rippled, and re-ordered according to pairwise recombination fractions, LOD scores and linkage group length [62]. Linkage group analyses were conducted using Join Map 4.0 [23] with a recombination frequency of 0.40 and a logarithm of odds (LOD) score of 2.5 for the $F_{2:3}$ population. The Kosambi mapping function was employed in the conversion of the recombination frequencies to map distances. Each data point represented the mean of three replications. Fiber quality and yield-related traits such as boll weight (BW), lint percentage (LP), fiber length (FL), fiber uniformity (FU), fiber micronaire (FM), fiber strength (FS), fiber elongation (FE), fiber reflectance (FR), fiber yellowness (FY), spinning consistency index (SCI) and mature index (MI) were used to conduct QTL analysis. The quantitative trait loci (QTLs) were detected using composite interval mapping (CIM) [63] by WinQTL Cartographer version 2.5 [20]. In the CIM mapping method, version 6, forward–reverse regression method with 1 cM walking speed, a probability into and out of the model of $p = 0.01$ and window size set at 10 cM. The LOD [64] threshold value was determined by 1000 permutation tests for all traits and was used to declare the significant QTLs with a significance level of $p = 0.05$. In addition, QTLs with LOD threshold of 2.5 in more than one environment were considered as common QTLs based on the explanation by Lander and Kruglyak [65].

QTL nomenclature was done based on the description by Liang et al. [2]. The proportion of observed phenotypic variance explained by each QTL was estimated by the coefficient of determination R^2 (%) as a percentage. The additive and dominance effects from QTL cartographer results were used to calculate genetic effects ($|d/a|$). The results were used to classify the QTL as additive effect (Ae) (0–0.20), partial dominant (PD) (0.21–0.80), dominant effect (De) (0.81–1.20) and over dominant (OD) >1.20 according to Stuber et al. (1987) [66]. The graphic presentation of the linkage group and QTLs marked were created by R software version 3.4.2 [21] and Map Chart 2.2 [67], respectively

4.4. Gene Mining and Expression Analysis

In this study, only segments of linkage groups associated with significantly detected QTLs were presented. The detected consistent QTLs were used to identify the crucial candidate genes for fiber yield and fiber quality-related traits. The genes identified were searched through the available resources [68] (available online: <https://cottonfgd.org>). The physical position of the GBS-SNP markers flanking major QTLs for fiber quality and yield-related traits were used to find the gene located in each QTL region. The function of the identified genes was determined through gene annotation. Furthermore, the expression profile of the candidate genes was analyzed by mapping it in the “TM-1”_RNA-seq transcriptome data of cotton (available online: <https://cottonfgd.org>). The expression values for each gene mined were used to generate the heat map using R-software script [21].

4.5. The Gene Ontology Enrichment Analysis BaseD on QTL Clusters

In order to determine the functions of the identified genes, we carried out gene ontology enrichment analysis through online software, Blast2GO (available online: <https://www.blast2go.com/>). Gene ontology describes the genes in three functional annotations, namely cellular component (CC), biological process (BP) and molecular functions (MF); three functions provide information on the possible roles played by the genes in the plant; of interest were the genes responsible or fiber qualities and yield-related traits. The choice of genes used for GO analysis was based on the genes mined from the two clusters, cluster 1 (D03) and 2 (D12), which had high percentage of phenotypic variation (PV) and heritability (Hb).

5. Conclusions

A genetic linkage map comprising of 5178 SNP markers, obtained by the GBS genotyping method, was generated using a 277 F_{2:3} population derived from an intra-specific cross of two tetraploid upland cotton. The map constructed in this study is the highest dense genetic map ever developed from an intra-specific population of the tetraploid upland cotton. The average distance of 0.92 cM was observed between adjacent markers. A total of 110 QTLs were obtained for 11 traits, however, only 30 QTLs were consistent in more than one environment. In addition, we identified 1709 genes that were found in the two main hot spot regions, named as cluster 1 and 2, with four QTLs in each. Out of the 1709 genes, 153 genes exhibited higher expression levels while the rest showed lower expression levels in all stages of fiber development. We further identified five key genes: *Gh_D03G0889*, *Gh_D12G0093*, *Gh_D12G0969*, *Gh_D12G0410*, and *Gh_D12G0435* to be the candidate genes involved in fiber development. This research provides the very first foundation in which future molecular work can be done, such as cloning of the identified genes and/or saturation of the genes to boost the current elite cultivated cotton cultivars.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/1422-0067/19/2/441/s1>.

Acknowledgments: We thank Xiongming Du (Cotton Research Institute/Chinese Academy of Agricultural Science, Anyang, China), Wenfang Gong (Cotton Research Institute/Chinese Academy of Agricultural Science, Anyang, China), Shoupu He (Cotton Research Institute/Chinese Academy of Agricultural Science, Anyang, China), all our lab mates and friends for their helpful suggestion and useful comments on the manuscript.

Author Contributions: Conceived and designed the experiments: Xiongming Du, Wenfang Gong, Shoupu He, Zhaoe Pan, Latyr Diouf, Yin Hua Jia, Richard Odongo Magwanga and Joy Nyangasi Kirungu. Performed the experiments: Latyr Diouf. Analyzed the data: Latyr Diouf. Contributed reagents/materials/analysis tools: Xiongming Du and Zhaoe Pan. Wrote the paper: Latyr Diouf and Richard Odongo Magwanga.

Conflicts of Interest: The authors declare no conflict of interest.

Funding: This program was financially sponsored by the National Science and Technology Support Program of China (No. 68612014028) and the National Key R&D Program of China (2017YFD0101600). We are very grateful to the China Scholarship Council (CSC).

References

1. Wang, K.; Wang, Z.; Li, F.; Ye, W.; Wang, J.; Song, G.; Yue, Z.; Cong, L.; Shang, H.; Zhu, S.; et al. The draft genome of a diploid cotton *Gossypium raimondii*. *Nat. Genet.* **2012**, *44*, 1098–1103. [[CrossRef](#)] [[PubMed](#)]
2. Liang, Q.; Hu, C.; Hua, H.; Li, Z.; Hua, J. Construction of a linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Chin. Sci. Bull.* **2013**, *58*, 3233–3243. [[CrossRef](#)]
3. Kohel, R.J.; Yu, J.; Park, Y.H. Molecular mapping and characterization of traits controlling fiber quality in cotton. *Euphytica* **2001**, *121*, 163–172. [[CrossRef](#)]
4. Lokhande, S.; Reddy, K.R. Quantifying Temperature Effects on Cotton Reproductive Efficiency and Fiber Quality. *Agron. J.* **2014**, *106*, 1275–1282. [[CrossRef](#)]
5. Zhang, Z.S.; Li, X.B.; Xiao, Y.H. Combining ability and heterosis between high strength lines and transgenic Bt (*Bacillus thuringiensis*) bollworm-resistant lines in upland cotton (*Gossypium hirsutum* L.). *Agric. Sci. China* **2003**, *2*, 13–18.
6. Shen, X.; Guo, W.; Lu, Q. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in upland cotton. *Euphytica* **2007**, *155*, 371–380. [[CrossRef](#)]
7. 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* × *Gossypium barbadense* backcross inbred line population. *Theor. Appl. Genet.* **2013**, *126*, 275–287. [[CrossRef](#)] [[PubMed](#)]
8. 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*, 877–888. [[CrossRef](#)]

9. Tan, Z.Y.; Fang, X.M.; Tang, S.Y.; Zhang, J.; Liu, D.J.; Teng, Z.H.; Li, L.; Ni, H.; Zheng, F.; Liu, D.; et al. Genetic map and QTL controlling fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Euphytica* **2015**, *203*, 615–628. [[CrossRef](#)]
10. Collard, B.C.Y.; Mackill, D.J. Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. B Biol. Sci.* **2008**, *363*, 557–572. [[CrossRef](#)] [[PubMed](#)]
11. Jamshed, M.; Jia, F.; Gong, J.; Palanga, K.K.; Shi, Y.; Li, J.; Shang, H.; Liu, A.; Chen, T.; Zhang, Z.; et al. Identification of stable quantitative trait loci (QTLs) for fiber quality traits across multiple environments in *Gossypium hirsutum* recombinant inbred line population. *BMC Genom.* **2016**, *17*, 197. [[CrossRef](#)] [[PubMed](#)]
12. Yu, Y.; Yuan, D.; Liang, S. Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC1 population between *Gossypium hirsutum* and *G. barbadense*. *BMC Genom.* **2011**, *12*, 15. [[CrossRef](#)] [[PubMed](#)]
13. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **2001**, *157*, 1819–1829. [[PubMed](#)]
14. Davey, J.W.; Hohenlohe, P.A.; Etter, P.D.; Boone, J.Q.; Catchen, J.M.; Blaxter, M.L. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **2011**, *12*, 499–510. [[CrossRef](#)] [[PubMed](#)]
15. Huang, X.; Kurata, N.; Wei, X.; Wang, Z.-X.; Wang, A.; Zhao, Q.; Zhao, Y.; Liu, K.; Lu, H.; Li, W.; et al. A map of rice genome variation reveals the origin of cultivated rice. *Nature* **2012**, *490*, 497–501. [[CrossRef](#)] [[PubMed](#)]
16. Xu, X.; Liu, X.; Ge, S.; Jensen, J.D.; Hu, F.; Li, X.; Dong, Y.; Gutenkunst, R.N.; Fang, L.; Huang, L.; et al. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* **2012**, *30*, 105–111. [[CrossRef](#)] [[PubMed](#)]
17. Spindel, J.; Wright, M.; Chen, C.; Cobb, J.; Gage, J.; Harrington, S.; Lorieux, M.; Ahmadi, N.; McCouch, S. Bridging the genotyping gap: Using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor. Appl. Genet.* **2013**, *126*, 2699–2716. [[CrossRef](#)] [[PubMed](#)]
18. Thomson, M.J.; Zhao, K.; Wright, M.; McNally, K.; Rey, J.; Tung, C.W.; Reynolds, A.; Scheffler, B.; Eizenga, G.; McClung, A.K.H.; et al. High-throughput single nucleotide polymorphism genotyping for breeding applications in rice using the BeadXpress platform. *Mol. Breed.* **2012**, *29*, 875–886. [[CrossRef](#)]
19. Poland, J.A.; Rife, T.W. Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* **2012**, *5*, 92–102. [[CrossRef](#)]
20. Wang, S.; Zeng, Z.B. *Department of Statistics*; North Carolina State University: Raleigh, NC, USA, 2007.
21. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008.
22. Li, F.; Fan, G.; Lu, C.; Xiao, G.; Zou, C.; Kohel, R.J.; 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. Biotechnol.* **2015**, *33*, 524–530. [[CrossRef](#)] [[PubMed](#)]
23. Stam, P.; van Ooijen, J.W. *JoinMap™ Version 2.0: Software for the Calculation of Genetic Linkage Maps*; Plant Research International B.V. and Kyazma B.V.: Wageningen, The Netherlands, 1995; CPRO-DLO: Wageningen, The Netherlands, 1995.
24. Rong, J.; Abbey, C.; Bowers, J.E. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). *Genet. Mol. Biol.* **2004**, *166*, 389–417. [[CrossRef](#)]
25. Fang, D.D.; Jenkins, J.N.; Deng, D.D.; McCarty, J.C.; 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 Genom.* **2014**, *15*, 397. [[CrossRef](#)] [[PubMed](#)]
26. Wang, K.; Wang, H.; Huang, C.; Guo, H.; Li, X.; Zhao, W.; Dai, B.; Yan, Z.; Lin, Z. QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. *PLoS ONE* **2015**, *10*, e0130742. [[CrossRef](#)] [[PubMed](#)]
27. Oluoch, G.; Zheng, J.; Wang, X.; Khan, M.K.R.; Zhou, Z.; Cai, X.; Wang, C.; Wang, Y.; Li, X.; Wang, H.; et al. QTL mapping for salt tolerance at seedling stage in the interspecific cross of *Gossypium tomentosum* with *Gossypium hirsutum*. *Euphytica* **2016**, *209*, 223–235. [[CrossRef](#)]

28. Lacape, J.-M.; Llewellyn, D.; Jacobs, J.; Arioli, T.; Becker, D.; Calhoun, S.; Al-Ghazi, Y.; Liu, S.; Palaï, O.; Georges, S.; et al. Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* × *G. barbadense* RIL population. *BMC Plant Biol.* **2010**, *10*, 132. [[CrossRef](#)] [[PubMed](#)]
29. Saba, J.; Moghaddam, M.; Ghassemi, K.; Nishabouri, M.R. Genetic properties of drought resistance indices. *J. Agric. Sci. Technol.* **2001**, *3*, 43–49.
30. Shao, Q.; Zhang, F.; Tang, S.; Liu, Y.; Fang, X.; Liu, D.; Liu, D.; Zhang, J.; Teng, Z.; Paterson, A.H.; et al. Identifying QTL for fiber quality traits with three upland cotton (*Gossypium hirsutum* L.) populations. *Euphytica* **2014**, *198*, 43–58. [[CrossRef](#)]
31. Lin, Z.; He, D.; Zhang, X.; Nie, Y.; Guo, X.; Feng, C.; Stewart, J.M. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. *Plant Breed.* **2005**, *124*, 180–187. [[CrossRef](#)]
32. Shen, X.L.; Guo, W.Z.; Zhu, X.F.; Yuan, Y.L.; Yu, J.Z.; Kohel, R.J.; Zhang, T.Z. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. *Mol. Breed.* **2005**, *15*, 169–181. [[CrossRef](#)]
33. Zhang, Z.-S.; Xiao, Y.-H.; Luo, M.; Li, X.-B.; Luo, X.-Y.; Hou, L.; Li, D.-M.; Pei, Y. Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (*Gossypium hirsutum* L.). *Euphytica* **2005**, *144*, 91–99. [[CrossRef](#)]
34. Zhang, Z.-S.; Hu, M.-C.; Zhang, J.; Liu, D.-J.; 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. Breed.* **2009**, *24*, 49–61. [[CrossRef](#)]
35. Wang, B.H.; Guo, W.Z.; Zhu, X.F.; Wu, Y.T.; Huang, N.T.; Zhang, T. QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. *Euphytica* **2006**, *152*, 367–378. [[CrossRef](#)]
36. He, D.H.; Lin, Z.X.; Zhang, X.L.; Nie, Y.C.; Guo, X.P.; Zhang, Y.X.; Li, W. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* × *Gossypium barbadense*. *Euphytica* **2007**, *153*, 181–197. [[CrossRef](#)]
37. Rong, J.; Feltus, F.A.; Waghmare, V.N.; Pierce, G.J.; Chee, P.W.; Draye, X.; Saranga, Y.; Wright, R.J.; Wilkins, T.A.; May, O.L.; et al. Meta-analysis of polyploidy cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. *Genetics* **2007**, *176*, 2577–2588. [[CrossRef](#)] [[PubMed](#)]
38. Ma, X.; Ding, Y.; Zhou, B.; Guo, W.; Lv, Y.; Zhu, X.; Zhang, T. QTL mapping in A-genome diploid Asiatic cotton and their congruence analysis with AD-genome tetraploid cotton in genus *Gossypium*. *J. Genet. Genom.* **2008**, *35*, 751–762. [[CrossRef](#)]
39. Sun, F.-D.; Zhang, J.-H.; Wang, S.-F.; Gong, W.-K.; Shi, Y.-Z.; Liu, A.-Y.; Li, J.-W.; Gong, J.-W.; Shang, H.-H.; Yuan, Y.-L. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. *Mol. Breed.* **2012**, *30*, 569–582. [[CrossRef](#)]
40. Zhang, S.W.; Zhu, X.F.; Feng, L.C.; Gao, X.; Yang, B.; Zhang, T.Z.; Zhou, B.L. Mapping of fiber quality QTLs reveals useful variation and footprints of cotton domestication using introgression lines. *Sci. Rep.* **2016**, *6*, 31954. [[CrossRef](#)] [[PubMed](#)]
41. Wang, H.; Huang, C.; Zhao, W.; Dai, B.; Shen, C.; Zhang, B.; Li, D.; Lin, Z. Identification of QTL for Fiber Quality and Yield Traits Using Two Immortalized Backcross Populations in Upland Cotton. *PLoS ONE* **2016**, *11*, e0166970. [[CrossRef](#)] [[PubMed](#)]
42. Said, J.I.; 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 Genom.* **2013**, *14*, 776. [[CrossRef](#)] [[PubMed](#)]
43. Tiwari, S.C.; Wilkins, T.A. Cotton (*Gossypium hirsutum*) Seed Trichomes Expand via Diffuse Growing Mechanism. *Can. J. Bot. Rev. Can. Bot.* **1995**, *73*, 746–757. [[CrossRef](#)]
44. Guan, X.; Lee, J.J.; Pang, M.; Shi, X.; Stelly, D.M.; Chen, Z.J. Activation of Arabidopsis Seed Hair Development by Cotton Fiber-Related Genes. *PLoS ONE* **2011**, *6*, e21301. [[CrossRef](#)] [[PubMed](#)]
45. Pu, L.; Li, Q.; Fan, X.; Yang, W.; Xue, Y. The R2R3 MYB Transcription Factor GhMYB109 Is Required for Cotton Fiber Development. *Genetics* **2008**, *180*, 811–820. [[CrossRef](#)] [[PubMed](#)]
46. Ma, Q.; Wu, M.; Pei, W.; Li, H.; Li, X.; Zhang, J.; Yu, J.; Yu, S. Quantitative phosphoproteomic profiling of fiber differentiation and initiation in a fiberless mutant of cotton. *BMC Genom.* **2014**, *15*, 466. [[CrossRef](#)] [[PubMed](#)]

47. Chaudhary, B.; Hovav, R.; Rapp, R.; Verma, N.; Udall, J.A.; Wendel, J.F. Global Analysis of Gene Expression in Cotton Fibers from Wild and Domesticated *Gossypium barbadense*. *Evol. Dev.* **2008**, *10*, 567–582. [[CrossRef](#)] [[PubMed](#)]
48. Peng, Z.; He, S.; Gong, W.; Sun, J.; Pan, Z.; Xu, F.; Lu, Y.; Du, X. Comprehensive analysis of differentially expressed genes and transcriptional regulation induced by salt stress in two contrasting cotton genotypes. *BMC Genom.* **2014**, *15*, 760. [[CrossRef](#)] [[PubMed](#)]
49. Gong, W.; Xu, F.; Sun, J.; Peng, Z.; He, S.; Pan, Z.; Du, X. iTRAQ-Based Comparative Proteomic Analysis of Seedling Leaves of Two Upland Cotton Genotypes Differing in Salt Tolerance. *Front. Plant Sci.* **2017**, *8*, 2113. [[CrossRef](#)] [[PubMed](#)]
50. Socquet-Juglard, D.; Christen, D.; Devenes, G.; Gessler, C.; Duffy, B.; Patocchi, A. Mapping Architectural, Phenological, and Fruit Quality QTLs in Apricot. *Plant Mol. Biol. Rep.* **2013**, *31*, 387–397. [[CrossRef](#)]
51. Paterson, A.H.; Brubaker, C.; Wendel, J.F. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep.* **1993**, *11*, 122–127. [[CrossRef](#)]
52. Krizman, M.; Jakse, J.; Baricevic, D.; Javornik, B.; Prosek, M. Robust CTAB-activated charcoal protocol for plant DNA extraction. *Acta Agric. Slov.* **2006**, *87*, 427–433.
53. Wilfinger, W.W.; Chomczynski, M.K. P 260/280 and 260/230 Ratios NanoDrop ©ND-1000 and ND-8000 8-Sample Spectrophotometers. *BioTechniques* **1997**, *22*, 474–481. [[PubMed](#)]
54. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [[CrossRef](#)] [[PubMed](#)]
55. Glaubitz, J.C.; Casstevens, T.M.; Lu, F.; Harriman, J.; Elshire, R.J.; Sun, Q.; Buckler, E.S. TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLoS ONE* **2014**, *9*, e90346. [[CrossRef](#)] [[PubMed](#)]
56. Paten, B.; Novak, A.; Haussler, D. Mapping to a Reference Genome Structure. *arXiv*, 2014.
57. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
58. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)] [[PubMed](#)]
59. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**, *23*, 2633–2635. [[CrossRef](#)] [[PubMed](#)]
60. SPSS Inc. *SPSS Base 10.0 for Windows User's Guide*; SPSS Inc.: Chicago, IL, USA, 1999.
61. Singh, U.M.; Yadav, S.; Dixit, S.; Ramayya, P.J.; Devi, M.N.; Raman, K.A.; Kumar, A. QTL Hotspots for Early Vigor and Related Traits under Dry Direct-Seeded System in Rice (*Oryza sativa* L.). *Front. Plant Sci.* **2017**, *8*, 286. [[CrossRef](#)] [[PubMed](#)]
62. Broman, K.W.; Wu, H.; Sen, S.; Churchill, G.A. R/QTL: QTL mapping in experimental crosses. *Bioinformatics* **2003**, *19*, 889–890. [[CrossRef](#)] [[PubMed](#)]
63. Da Silva, M.J.R.; Ferreira, T.E.; Domiciano, S.; Paula, A.; Paiva, M.; Tecchio, M.A.; Leonel, S. Phenology, yield and fruit quality of four persimmon (*Diospyros kaki* L.) cultivars in So Paulos Midwest countryside, Brazil. *Afr. J. Agric. Res.* **2016**, *11*, 5171–5177.
64. Podliesna, S.; Bezzina, C.R.; Lodder, E.M. Complex Genetics of Cardiovascular Traits in Mice: F2-Mapping of QTLs and Their Underlying Genes. *Methods Mol. Biol.* **2017**, *1488*, 431–454. [[PubMed](#)]
65. Lander, E.; Kruglyak, L. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat. Genet.* **1995**, *1*, 1241–1247. [[CrossRef](#)] [[PubMed](#)]
66. Stuber, C.; Edwards, M.D.; Wendel, J.F. Molecular marker facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. *Crop Sci.* **1987**, *27*, 639–648. [[CrossRef](#)]
67. Voorrips, R.E. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [[CrossRef](#)] [[PubMed](#)]
68. Zhu, T.; Liang, C.; Meng, Z.; Sun, G.; Meng, Z.; Guo, S.; Zhang, R. CottonFGD: An integrated functional genomics database for cotton. *BMC Plant Biol.* **2017**, *17*, 101. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).