

Mapping quantitative trait loci for cottonseed oil, protein and gossypol content in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population

Jiwen Yu · Shuxun Yu · Shuli Fan ·
Meizhen Song · Honghong Zhai · Xingli Li ·
Jinfa Zhang

Received: 10 December 2011 / Accepted: 20 January 2012 / Published online: 11 February 2012
© Springer Science+Business Media B.V. 2012

Abstract Cotton is one of the most important oil-producing crops and the cottonseed meal provides important protein nutrients as animal feed. However, information on the genetic basis of cottonseed oil and protein contents is lacking. A backcross inbred line (BIL) population from a cross between *Gossypium hirsutum* as the recurrent parent and *G. barbadense* was used to identify quantitative trait loci (QTLs) for cottonseed oil, protein, and gossypol contents. The BIL population of 146 lines together with the two parental lines was tested in the same location for three years in China. Based on a genetic map of 392 SSR markers and a total genetic distance of 2,895.2 cM, 17 QTLs on 12 chromosomes for oil content, 22 QTLs on 12 chromosomes for protein content and three QTLs on two chromosomes for gossypol content were detected. Seed oil content was significantly and negatively correlated with seed protein content, which can be explained by eight QTLs for both oil and

protein contents co-localized in the same regions but with opposite additive effects. This research represents the first report using a permanent advanced backcross inbred population of an interspecific hybrid population to identify QTLs for seed quality traits in cotton in three environments.

Keywords Tetraploid cotton · Cottonseed · Oil · Protein · Gossypol · Linkage map · Backcross inbred line (BIL) · Quantitative trait locus (QTL)

Introduction

Cotton provides the most important natural fiber for the textile industry globally. Cotton is also a food and feed crop in that cottonseed as a byproduct produces edible oil for human consumption and protein meals for animal feed (Cherry et al. 1978a; Cherry 1983; Cherry and Leffler 1984). Cotton produces 150 kg of cottonseed for every 100 kg of lint fibers produced (O'Brien et al. 2005). As an oilseed crop, cottonseed production ranks third after soybean and rapeseed worldwide (USDA-FAS 2011). The use of its oil as biofuel has attracted increased attentions in recent years (Karaosmanoglu et al. 1999; Meneghetti et al. 2007). However, effort in genetics, breeding, and cultural practices to improve cottonseed quality including oil and protein contents is minimal and intermittent.

Cottonseed oil and protein contents are quantitative traits and both are usually negatively correlated with

J. Yu · S. Yu (✉) · S. Fan · M. Song · H. Zhai · X. Li
National Key Laboratory of Cotton Biology, Key
Laboratory of Cotton Genetic Improvement,
Cotton Research Institute, Chinese Academy of
Agricultural Sciences, Ministry of Agriculture,
Anyang 455112, Henan, China
e-mail: yujw666@hotmail.com

J. Zhang (✉)
Department of Plant and Environmental Sciences,
New Mexico State University, Las Cruces,
NM 88003, USA
e-mail: jinzhang@nmsu.edu

one another (Hanny et al. 1978; Wu et al. 2009). The two traits are affected by genotypic, developmental and environmental factors during cottonseed development (Cherry et al. 1981; Cherry and Leffler 1984). Cottonseed oil and protein contents vary in growing seasons, locations, and years (Cherry et al. 1978a, 1978b, 1981; Kohel and Cherry 1983; Turner et al. 1976a). Environmental factors include temperature (Gipson and Joham 1969) and fertilizers (Anderson and Worthington 1971; Leffler et al. 1977; Elmore et al. 1979) among others. Great genetic variations among cotton species and cultivars in cottonseed oil (17–27%) and protein (12–32%) contents also exist (Kohel 1980; Kohel et al. 1985; Wu et al. 2009; Dowd et al. 2010). However, the genetic basis controlling oil and protein contents has received little attention. Kohel (1980) estimated a moderate heritability based on a 20×5 NCII design and low heritability based on F_2/F_3 regression for cottonseed oil content. Other quantitative genetic designs including diallel crossing and generation mean analysis were also used to estimate genetic parameters for oil and protein contents. General and special combining abilities, both additive and non-additive including dominant efforts, and maternal effect were detected (Kohel 1980; Dani and Kohel 1989). Using a set of chromosomal substitution lines and an AD genetic model, Wu et al. (2009, 2010) confirmed low-moderate genetic variances for seed oil and protein content but moderate genetic variances for oil and protein index. Both dominant and additive variances and maternal effect existed for the four seed quality traits but additive variance for oil content was not detected (Wu et al. 2009, 2010). Song and Zhang (2007) reported 11 QTLs for kernel percentage, kernel oil percentage, kernel protein percentage, and seven amino acids using a BC_1S_1 population derived from a *G. hirsutum* \times *G. barbadense* cross.

However, the potential of cottonseed as food and feed has not fully utilized as it contains the toxic terpenoid gossypol. Gossypol can be removed from the cottonseed through genetic approaches or chemical techniques. Genetic variation in gossypol content exists among cotton species and genotypes. Gossypol-containing glands are distributed in all the cotton plant bodies including leaves, stems, flowers, bolls and seeds. Gossypol-free cotton plants and seeds can be bred by introduction of double recessive genes *gl₂gl₃* from Hopi cotton (McMichael 1960) or a dominant

Gl₂ gene from an Egyptian cotton (Kohel and Lee 1984). Gossypol-free glandless seed and glanded plants can be found in certain wild Australian cotton species but the introgression of the trait to Upland cotton has not been successful (Zhu et al. 2005; Benbouza et al. 2009). Gossypol content is usually higher in *G. barbadense* than in *G. hirsutum* and it also differs between genotypes in Upland cotton. However, there has been no report on QTLs controlling the quantitative variation of gossypol content in cotton.

In the present study, a backcross inbred line (BIL) population was developed from a cross between Upland cotton and *G. barbadense* and used to identify QTLs for seed quality traits including oil, protein, and gossypol contents. Several consistent QTLs were identified in different tests and many QTLs for cottonseed oil and protein contents were co-localized.

Materials and methods

Generation of the BIL population and field tests

An interspecific BIL population of 146 lines was used in this study. The BIL population was developed from a cross between Upland cotton (*G. hirsutum*) SG 747 and *G. barbadense* Giza 75 through two generations of backcrossing using SG 747 as the recurrent parent followed by four generations of selfing. The 146 BILs and the two parents were planted in China Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang, Henan province in 2006, 2008, and 2009. This location represents one of the three major cotton production regions in China, i.e., Yellow River valley. The 148 entries were arranged in a randomized complete block design with two replications and single row plots. Seeds were sown in April and crop managements followed local recommendations. The plot length was 8 m with a row spacing of 0.8 m and seedlings were thinned to 32 plants per plot.

Determination of cottonseed oil, protein, and gossypol content

At plant maturity, 25 open bolls from each plot in each field test were hand harvested and ginned for evaluation of seed quality traits. Cottonseed harvested in 2006 was sent to Beijing Nutrient Research Institute, Beijing, China, to determine oil and protein content by

the Soxhlet (De Castro and Garcia-Ayuso 1998) and Kjeldahl (Feil et al. 2005) extraction methods, respectively. Using the same methods, cottonseed harvested in 2009 was also measured for oil and protein contents at Agricultural Product Testing Center at Zhenzhou, Henan, China. National standards GB/T 14772-2008 for oil and GB 5009.5-2010 for protein were followed. To reduce the cost in measurements, seed from the two replicates was combined based on genotypes in 2006 and 2009, respectively.

Seed samples harvested in 2008 and 2009 were also sent to Zhejinag University, Hangzhou, China, for determination of oil, protein, and gossypol contents using near infrared reflectance spectroscopy (FOSS NIR System 5000). Detailed information can be found in Wang et al. (2001) for oil, in Wang et al. (2010) for protein and in Birth and Ramey (1982) for gossypol.

Statistical Analysis

The results from 2008 were statistically analyzed using SAS. However, the results from Beijing 2006, Zhejiang 2009 and Zhenzhou 2009 were pooled to conduct the analysis of variance using SAS Proc MIXED with tests as replicates.

DNA extraction, maker analysis, and map construction

The genomic DNAs were extracted from young leaves of the 146 individual BIL lines and the two parents using a mini-prep method as described by Zhang and Stewart (2000). Simple sequence repeat markers (SSRs) were used to construct a genetic map for the BIL population using MAPMAKER ver. 3.0b program (Lander et al. 1987). The linkage map was published elsewhere (Yu et al. 2012). MAPMAKER has been used in constructing linkage maps for BIL populations of other crops including rice (e.g., Matsubara et al. 2008).

QTL mapping

QTLs were identified by composite interval mapping (Zeng 1994) using Windows QTL Cartographer 2.0 (Basten et al. 2001). This software has been used to detect QTLs in BIL populations of other crops such as rice (e.g., Yamagishi et al. 2004). The windows size was set at 5 cM and the walk speed at 1 cM. The

maximum ten background markers were used for genetic background control and LOD threshold values were estimated by 1,000 permutations to declare significant QTLs (Churchill and Doerge 1994). A LOD score of ≥ 2.5 was selected to detect significant QTLs. A location QTL confidence interval (95%) was set as a mapping distance interval corresponding to one LOD decline on either side of the peak. QTLs for the same trait across different years and environments were declared as a “common” QTL when their confidence intervals overlapped. The QTL nomenclature followed McCouch et al. (1997) in that a QTL designation begins with “q”, followed by an abbreviation of the trait name, year, location, chromosome name, and finally a serial number.

Results

Analysis of variance and performance of the BIL population

In the three years when the BILs were tested (Table 1), the Upland cotton (Gh hereafter) parent SG 747 had 5.0% lower seed oil content than the *Gossypium barbadense* Egyptian cotton (Gb hereafter) parent Giza 75 (33.08 vs. 34.81%), but 2–6% higher protein content (30.88 vs. 30.16% in 2008; 39.25 vs. 37.06% in 2006 and 2009). The mean oil and protein contents of BILs were closer to the recurrent parent, as expected for the BC₂-derived BIL population. However, the ranges of the BIL population in both seed oil

Table 1 Cottonseed oil, protein and gossypol contents of parents and their backcross inbred line (BIL) population, and variation and heritabilities in the BIL population

Trait	Oil%: 08	Pro%: 08	Goss%: 08	Oil%: 06 + 09	Pro%: 06 + 09
SG 747	33.02	30.88	1.02	33.13	39.25
Giza 75	34.89	30.16	1.00	34.72	37.06
BIL-Min	28.22	28.63	0.65	27.73	35.92
BIL-Max	36.76	34.27	1.39	36.28	44.84
BIL-Mean	33.48	30.66	0.96	32.06	39.69
F	4.18	4.90	2.74	2.34	22.66
LSD (0.05)	2.03	1.29	0.19	2.78	1.11
H (b)	0.81	0.83	0.73	0.70	0.96

06: 2006; 08: 2008; 09: 2009; *Goss* gossypol; *H(b)* broad-sense heritability

and protein contents were beyond the values of the two parents, indicating transgressive segregations. For example, based on the results in 2008 (Table 1), the BIL with the lowest oil content was significantly lower (28.22%) than the lower recurrent parent (RP), while the BIL with the highest oil content (36.76%) was significantly higher than the RP, but the increase in oil content than the donor parent was insignificant. The same trend was noted when data from 2006 and 2009 were analyzed (Table 1). This indicated that both QTLs for reducing and increasing seed oil content have been transferred into the Upland cotton parent SG 747 through two generations of backcrossing.

For protein content, similar results were obtained. The BIL with the lowest protein content (28.63%) was significantly lower than the lower Gb parent, while the BIL with the highest protein content (34.27%) was significantly higher than the higher Gh parent. The same is true for the combined results from 2006 and 2009. This also indicated that both QTLs for reducing and increasing protein content have been transferred into the Upland cotton parent SG 747 through two generations of backcrossing.

For seed gossypol content, the two parents were similar (ca. 1%). However, transgressive segregation was also observed in that BILs with significantly lower (0.65%) or higher (1.39%) gossypol content were identified (Table 1).

Heritabilities for the three traits were moderately high to high (0.70–0.96; Table 1), indicating that the majority of the phenotypic variation was due to genotype. Therefore, selection for increasing seed oil or protein content is expected to make progress.

To evaluate the consistency between testing methods on oil and protein determination, correlation analysis was performed between tests (Table 2). It appeared that the results from the oil and protein analyses in Zhejiang 2008 and 2009 and Zhenzhou 2009 were highly correlated, but the oil contents were not correlated with the oil analysis performed in Beijing 2006. This indicated that the methods that were used in Zhejiang and Zhenzhou were overall congruent in determining seed oil and protein contents but different from that used in Beijing when determining cottonseed oil content. The results between 2008 and 2009 were also consistent. The lack of

Table 2 Coefficients of correlation between tests in cottonseed oil, protein, and gossypol contents and between the three cottonseed quality traits based on a backcross inbred line population

Trait or trait pair	Coefficient of correlation			
Oil%				
	Beijing 2006	Zhejiang 2008	Zhejiang 2009	Zhenzhou 2009
Beijing 2006	1	0.111	0.05	0.061
Zhejiang 2008		1	0.532	0.455
Zhejiang 2009			1	0.862
Zhenzhou 2009				1
Protein%				
	Zhejiang 2008	Zhejiang 2009	Zhenzhou 2009	
Zhejiang 2008	1	0.629	0.527	
Zhejiang 2009		1	0.833	
Zhenzhou 2009			1	
Oil–Protein%				
Zhejiang 2008	–0.905			
Zhejiang 2009	–0.818			
Zhenzhou 2009	–0.809			
Gossypol–Oil%				
Zhejiang 2008	0.195			
Zhejiang 2009	0.287			
Gossypol–Protein%				
Zhejiang 2008	–0.347			
Zhejiang 2009	–0.538			
Zhenzhou 2009	–0.538			

$$r_{0.05, df=146} = 0.157; r_{0.01, df=146} = 0.204$$

correlation between results from 2006 and these from 2008/2009 may be in large part due to genotype \times environment interactions.

The coefficients of correlation (-0.81 to -0.91) between seed oil and protein contents were highly significantly negative in three tests performed in Zhejiang 2008 and 2009 and Zhenzhou 2009. This indicated that cotton genotypes with higher oil content usually have lower protein content, or vice versa. Therefore, simultaneously improving seed oil and protein contents would be difficult.

Gossypol content is significantly and positively correlated with oil content, but significantly and negatively correlated with protein content. But the correlations were not very close. This indicated that reducing seed gossypol content may increase oil content but reduce protein content in cottonseed.

Brief description of the linkage map

A total of 392 SSR markers were developed for the BIL population to construct a linkage map of 29 linkage groups with a total genetic distance of 2,895.2 cM and an average genetic distance of 7.4 cM per marker. The results were published elsewhere (Yu et al. 2012). The linkage map was used to identify QTLs for seed oil, protein, and gossypol contents, as described in the following.

QTLs for seed oil content

Seventeen QTLs distributed on 12 chromosomes were detected from four tests (Table 3; Fig. 1) including, 4 from Beijing 2006, 8 from Zhejiang 2008, and only 3 and 2 QTLs from Zhejiang and Zhenzhou 2009, respectively. Interestingly, 2, 3 and 3 QTLs were detected on c12, c19, and c21, respectively. The three QTLs detected on c21 may be one common QTL since they were located in a similar region (peaked at 54.2–74.9 cM). The QTL on c12 at the peak of 127 cM and the two QTLs on c21 at the peaks of 54.2–60.5 contributed to the phenotypic variation (PV) by 22–26%. Therefore, they may be considered as major QTLs for seed oil content.

Except for the QTL on c1 whose allele from the Gb parent had positive additive effect, all the oil QTL alleles from the Gh parent, the lower parent in oil content, had positive additive effects. This may explain the negative transgressive segregation in oil

content in the BIL population, because the Gb alleles for most of the oil QTLs detected had negative additive effects. There may be more positive alleles from the Gb parent contributing to the positive transgressive segregation in cottonseed oil content, but this experiment failed to detect most of them.

QTLs for seed protein content

A total of 22 QTLs distributed on 12 chromosomes for cottonseed protein content were detected including, 16 from Zhejiang 2008, 4 from Zhejiang 2009 and 2 from Zhenzhou 2009 (Table 3; Fig. 1). Chromosome c3 and c19 each carried 3 QTLs, while c5, c12, c20, c21, and c24 each carried 2 QTLs. The two QTLs on c3, c21, and c24 may be common QTLs because of their close proximity within a 25 cM region.

Although the Gh parent had higher cottonseed protein content than the Gb parent, the difference was insignificant. Unexpectedly, all the QTL alleles from the Gh parent displayed negative additive effects on protein content. This may explain the positive transgressive segregation in protein content of the BIL population because the alleles from the Gb parent had positive contributions to the PV of protein content in cottonseed. In the current study, positive QTL alleles from the Gh parent were not detected. Also, similar to oil content, 9 QTLs for protein content contributed to the protein PV by more than 20% based on simple regression analysis.

QTLs for seed gossypol content

Three QTLs on two chromosomes (c13 and c19) were detected (Table 3; Fig. 1). The two QTLs on c19 were separated by >90 cM and also possessed opposite gene effects. Therefore, they are different QTLs. Each of the QTLs contributed to PV by ca. 6% and therefore they were minor QTLs.

Co-localization of QTLs

A number of QTLs for cottonseed oil and protein contents were found to be co-localized on chromosomes. For example, one oil QTL and two protein QTLs on c3, one oil and one protein QTL on c5, one oil and one protein QTL on c12, one oil and one protein QTL on c15, one oil and one protein QTL on c16, one oil and one protein QTL on c19, three oil and

Table 3 QTLs, chromosome locations and effects for cottonseed oil, protein, and gossypol content in a backcross inbred line population

Trait	Environment	QTLname	Chr	Position	LOD	Marker Interval	A	R ²
Oil	Zhejiang2008	qOil2-c1-1	1	36.7	2.02	NAU3135-CIR004	-1.12	7.45
Protein	Zhejiang2008	qPro1-c2-1	2	36.2	4.84	NAU3684-BNL3971	-1.94	33.39
Oil	Zhejiang2009	qOil3-c3-1	3	48.9	2.53	BNL3259-NAU3541	2.47	8.94
Protein	Zhejiang2008	qPro1-c3-1	3	40.9	3.17	BNL3259-NAU3541	-1.79	32.11
Protein	Zhejiang2009	qPro2-c3-1	3	16	2.44	NAU5289-CIR068	-1.32	5.77
Protein	Zhejiang2009	qPro2-c3-2	3	69.7	2.99	NAU3671-CIR228	-2.38	15.15
Protein	Zhejiang2008	qPro1-c4-1	4	96.8	2.14	NAU3093-NAU5236	-1.68	34.98
Oil	Zhejiang2008	qOil2-c5-1	5	1.6	3.37	NAU3607-NAU3405	1.32	8.24
Protein	Zhejiang2008	qPro1-c5-1	5	1.6	3.65	NAU3607-NAU3405	-0.99	9.85
Protein	Zhenzhou2009	qPro3-c5-1	5	94.9	2.23	BNL3992-BNL1038	-1.43	5.61
Protein	Zhenzhou2009	qPro3-c6-1	6	0	2.60	CIR203-BNL2569	-1.79	6.55
Oil	Beijing2006	qOil1-c11-1	11	85.2	2.27	BNL3442-NAU3341	1.73	5.45
Oil	Beijing2006	qOil1-c12-1	12	27.1	2.64	NAU3041-CIR362	1.68	6.47
Oil	Zhejiang2008	qOil2-c12-1	12	127	2.26	BNL4059-BNL2717	3.26	23.64
Protein	Zhejiang2008	qPro1-c12-1	12	51.4	2.17	NAU0943-NAU3109	-1.41	13.91
Protein	Zhejiang2008	qPro1-c12-2	12	129	3.42	BNL4059-BNL2717	-2.23	27.15
Protein	Zhejiang2008	qPro1-c14(2)-1	14	23.5	2.25	BNL0891-CIR097	-1.77	34.04
Oil	Zhenzhou2009	qOil4-c15-1	15	51.5	3.06	BNL1667-NAU3680	2.78	15.98
Protein	Zhejiang2008	qPro1-c15-1	15	75.2	2.63	BNL2646-NAU3922	-1.94	31.69
Oil	Zhejiang2008	qOil2-c16-1	16	65.5	3.06	BNL2734-NAU5024	1.53	10.65
Protein	Zhejiang2008	qPro1-c16-1	16	65.5	5.10	BNL2734-NAU5024	-1.43	18.09
Oil	Beijing2006	qOil1-c19-1	19	54.1	2.92	BNL1690-BNL1611	2.38	7.24
Oil	Zhejiang2008	qOil2-c19-1	19	14.2	2.59	NAU3405-BNL1706	1.34	11.48
Oil	Zhejiang2009	qOil3-c19-1	19	157.2	2.01	BNL1671-NAU3946	1.57	13.18
Protein	Zhejiang2008	qPro1-c19-1	19	85.1	2.15	NAU3416-NAU3664	-1.47	14.11
Protein	Zhejiang2008	qPro1-c19-2	19	132.4	5.28	CIR179-BNL1671	-1.53	30.52
Protein	Zhejiang2009	qPro2-c19-1	19	16.2	3.74	NAU3405-BNL1706	-1.57	10.24
Oil	Zhejiang2008	qOil2-c20-1	20	25.7	2.15	NAU3434-NAU3531	1.00	5.78
Protein	Zhejiang2008	qPro1-c20-1	20	25.5	2.13	NAU3407-NAU3531	-0.69	5.99
Protein	Zhejiang2008	qPro1-c20-2	20	57.7	3.25	BNL0119-NAU3368	-0.14	19.79
Oil	Zhejiang2008	qOil2-c21-1	21	60.5	2.45	BNL1404-BNL3449	1.98	25.60
Oil	Zhejiang2009	qOil3-c21-1	21	74.9	2.79	NAU3381-NAU3731	1.94	7.09
Oil	Zhenzhou2009	qOil4-c21-1	21	54.2	2.58	BNL1404-BNL3449	2.47	22.20
Protein	Zhejiang2008	qPro1-c21-1	21	54.2	3.43	NAU5212-BNL1404	-1.70	32.50
Protein	Zhejiang2009	qPro2-c21-1	21	60.5	2.08	BNL1404-BNL3449	-1.71	16.96
Oil	Zhejiang2008	qOil2-c24-1	24	62.4	3.22	DPL0068-BNL3638	0.74	18.40
Protein	Zhejiang2008	qPro1-c24-1	24	42.4	4.10	NAU3904-NAU3158	-2.12	28.17
Protein	Zhejiang2008	qPro1-c24-2	24	60.4	2.87	DPL0068-BNL3638	-0.59	33.10
Oil	Beijing2006	qOil1-c25-1	25	64.4	2.56	NAU3306-BNL3103	-2.34	8.12
Gossypol	Zhejiang2008	qGos1-c13-1	13	35.9	2.22	BNL1438-BNL2652	0.10	5.79
Gossypol	Zhejiang2008	qGos1-c19-1	19	92.3	2.45	NAU5347-BNL1878	0.09	6.47
Gossypol	Zhejiang2009	qGos2-c19-1	19	0	2.35	NAU3656-DPL0247	-0.24	5.90

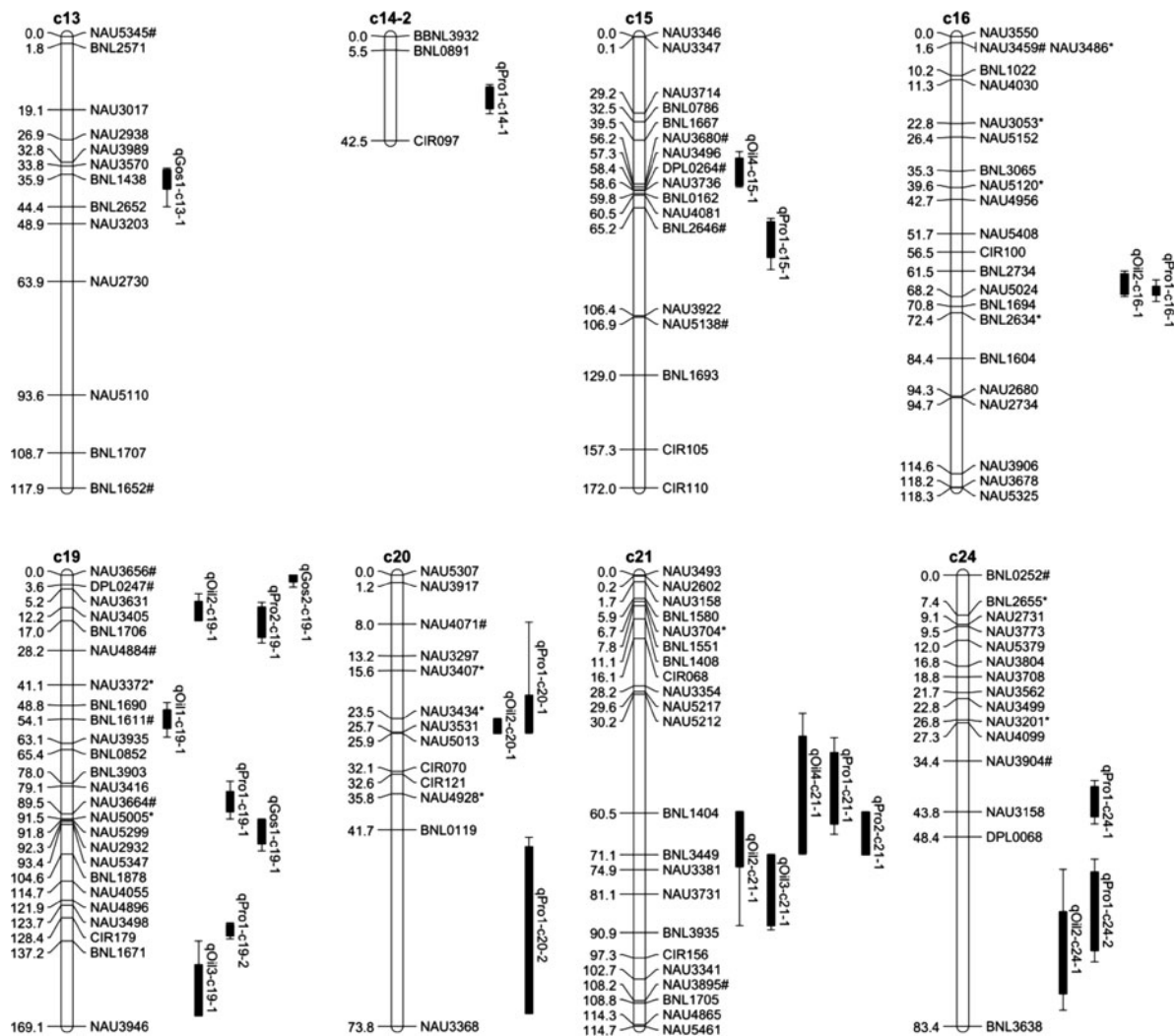


Fig. 1 continued

within a 14 cM region in two of the three tests. Another QTL for protein was detected in an 18 cM region on c24 in Zhejiang 2008, where an oil QTL was also detected.

Discussion

Heritability and transgressive segregation in seed quality traits

There have been a few reports on heritabilities and genetic effects concerning cottonseed oil and protein contents using classical quantitative genetic techniques such as diallel crossing and generation mean

analysis in a single environment (Kohel 1980; Dani and Kohel 1989; Ye et al. 2003). The results reported in the current study represent the first study on the genetic basis of cottonseed oil and protein contents using a permanent genetic population tested in replicated field trials in different environments. Compared with fiber yield and quality traits, moderately high to high broad-sense heritabilities were estimated for cottonseed oil, protein, and gossypol content. This is not unexpected since the measurements were on an average genotype basis rather than on single plant basis with no replicates. Therefore, selection for high cottonseed oil or protein content can be efficiently accomplished using replicated genotype means as the selection unit.

Since the BILs were developed by two generations of backcrossing followed by several times of selfing, the BIL population carried an overall 11% of the Gb genetic materials. However, the actual percentage of Gb genetic materials differed among individual BILs, following a binomial distribution. For the three seed quality traits, both negative and positive transgressive segregations were observed, indicating the successful transfer of both negative and positive QTL alleles from the donor Gb parent to the recipient Gh parent. However, from a cotton breeding's perspective, only the positive transgression in oil and protein content and negative transgressive segregation in gossypol content will be beneficial. The best BIL increased cottonseed oil by ca. 10% than the current parent, and even by 5% than the donor parent (Table 1). For cottonseed protein content, the introduction of Gb alleles resulted in its increase by ca. 12% in the best BIL than the recurrent parent (Table 1). The BIL with the lowest cottonseed gossypol (0.65%) content had gossypol content 37% lower than the recurrent Gh parent (1.0%). The results demonstrate that backcrossing followed by selfing is an efficient method to enhance cottonseed oil or protein content and to reduce gossypol content in interspecific breeding between Upland cotton and *G. barbadense*.

Numbers of QTLs controlling cottonseed oil and protein content

Wu et al. (2009, 2010) studied cottonseed quality traits using 13 chromosome substitution (CS-B) lines each of which carried a Gb chromosome or an arm in one or five Upland cotton backgrounds. In the TM-1 genetic background (Wu et al. 2009), more than eight chromosomes or arms are involved in determining cottonseed oil and protein contents. For example, Gb chromosome 2, 6, 17, and 18 were associated with increase in protein content, while Gb chromosome 4, 7, 14sh, and 15sh were associated with reduction in protein content. For cottonseed oil content, Gb chromosome 4 and 18 had positive effects, while Gb chromosome 2, 6, 7, 17, 5sh, 14sh, 22sh, and 22Lo had negative effects. Both the short and long arm of chromosome 22 contributed to oil content negatively. In crossing with five Upland cotton cultivars, Wu et al. (2010) further confirmed that Gb chromosome 2, 4, 25, 5sh, 14sh, and 15sh had significant additive effects on protein content, while Gb chromosome 4, 17, 18, 15sh,

and 22Lo carried significant additive effects on oil content. The authors also detected significant homozygous dominant effects on oil content from seven Gb chromosomes or arms and heterozygous dominant effects on oil content from 12 Gb chromosomes or arms (except for chromosome 4) in one or more than one genetic background. This clearly indicated almost each Gb chromosome may carry genetic factors affecting cottonseed oil or protein content.

In the current study, 17 QTLs for cottonseed oil content were identified to be located on 12 chromosomes, six of which were on five chromosomes (c1, c3, c5, c11, and c12) of the Ah sub-genome and 11 were on seven chromosomes (c15, c16, c19, c20, c21, c24, and c25) of the Dh sub-genome. For cottonseed protein content, 22 QTLs were found on 12 chromosomes (c2, c3, c4, c5, c6, c12, c15, c16, c19, c20, c21, and c24), and six chromosomes in each of the two sub-genomes carried 10 or 12 QTLs. Therefore, almost half (12/26) of the chromosomes in the tetraploid cotton contain genetic factors affecting cottonseed oil or protein formation. The results obtained in this study support the notion that genetic factors controlling cottonseed oil and protein are scattered on many cotton chromosomes, as demonstrated by Wu et al. (2009, 2010) using chromosome substitution lines.

Trait association, QTL co-localization and implications in breeding

This study also detected significant negative correlation between oil and protein content in cottonseed, as Turner et al. (1976b), Leffler et al. (1977), Hanny et al. (1978), and Shaver and Dilday (1982) reported. However, the association between the two traits was weak in a set of chromosome substitution lines, their parents and five Upland cotton cultivars (Wu et al. 2009). Similar to Hanny et al. (1978) and Shaver and Dilday (1982), this study also detected a weak but positive correlation between cottonseed oil and gossypol content and negative correlation between protein and gossypol content.

In the current study, nine chromosomes (c3, c5, c12, c15, c16, c19, c20, c21, and c24) were found to carry QTLs for both cottonseed oil and protein contents, consistent with Wu et al. (2009, 2010). Interestingly, 8 of the 9 chromosomes carried QTLs for both oil and protein contents but with opposite additive effects, which were located in a close

proximity. The results for the first time provide evidence that the negative association between cottonseed oil and protein contents is likely due to pleiotropy or tightly linked QTLs for both oil and protein formation. It implies that breeding for increasing both oil and protein contents in cottonseed is difficult and will need to break the tight linkage between QTLs in a repulsion phase for the two traits. High resolution mapping of QTLs for both cottonseed oil and protein contents will shed more light on the relative importance of pleiotropy and linkage in determining the association of the two traits.

Another interesting finding of this study is that almost all the positive alleles for oil and protein content were from the Gh and Gb parents, respectively. This is unexpected based on the parental differences in that the Gh parent had lower oil but higher protein content than the Gb parent. According to Wu et al. (2010), of the five Gb chromosomes with significant additive effects on oil content, four had positive effects; of the six Gb chromosomes with significant additive effects on protein content, five had negative effects. The results from this study appeared to be contradictory to Wu et al. (2010). Further studies will be needed to clarify this important issue.

QTLs for cottonseed gossypol content

Glandedness in cotton plants are controlled by two major genes *Gl₂Gl₃* and different alleles and they are located on c12 and c26, respectively. The double recessive mutant *gl₂gl₃* resulted in glandless cotton plants and seed (Percy and Kohel 1999). Several other major glanded loci were also reported. In this study, three QTLs (1 on c13 and 2 on c19) were identified to contribute to quantitative variation in cottonseed gossypol content and none of them were located on c12 and c26, indicating that these QTLs are different genes from the two major glanded genes. Especially, the QTL allele for qGos2-c19-1 from the Gb parent had a positive effect contributing to increased gossypol content. This may partially explain the positive transgressive segregation of gossypol content in the BIL population.

Acknowledgments This research was supported by the National Key Basic Research and Development Program of China (973 Program) (No. 2010CB126006), and the New Mexico Agricultural Experiment Station.

References

- Anderson OE, Worthington RE (1971) Boron and manganese effects on protein, oil content, and fatty acid composition of cottonseed. *Agron J* 63:566–569
- Basten CJ, Weir BS, Zeng ZB (2001) QTL Cartographer, Version 1.15. Department of Statistics, North Carolina State University, Raleigh, NC
- Benbouza H, Lognay G, Scheffler J, Baudoin JP, Mergeai G (2009) Expression of the “glanded-plant and glandless-seed” trait of Australian diploid cottons in different genetic backgrounds. *Euphytica* 154:211–221
- Birth GS, Ramey HH (1982) Near-infrared reflectance for analysis of cottonseed for gossypol. *Cereal Chem* 59: 516–519
- Cherry JP (1983) Cottonseed oil. *J Am Oil Chem Soc* 60:360–367
- Cherry JP, Leffler HR (1984) Seed. In Kohel RJ and Lewis CF (eds) *Cotton*, Agronomy Monograph no. 24, Crop Science Society of America, Madison, pp 511–531
- Cherry JP, Simmons JG, Kohel RJ (1978a) Potential for improving cottonseed quality by genetics and agronomic practices. In: Friedman M (ed) *Nutritional improvement of food and feed proteins*. Plenum Press, New York, pp 343–364
- Cherry JP, Simmons JG, Kohel RJ (1978b) Cottonseed composition of national variety test cultivars grown at different Texas locations. In: Brown JM (ed) *Proceedings of the Beltwide cotton production research conference*, Dallas. National Cotton Council, Memphis, pp 47–50
- Cherry JP, Kohel RJ, Jones LA, Powell WH (1981) Cottonseed quality: factors affecting feed and food uses. In: Brown JM (ed) *Proc. Beltwide cotton production research conference*, New Orleans. National Cotton Council: Memphis, pp 266–282
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963–971
- Dani RG, Kohel RJ (1989) Maternal effects and generation mean analysis of seed-oil content in cotton (*Gossypium hirsutum* L.). *Theor Appl Genet* 77:569–575
- De Castro MDL, Garcia-Ayuso LE (1998) Soxhlet extraction of solid materials: an outdated technique with a promising innovative future. *Anal Chim Acta* 369:1–10
- Dowd MK, Boykin DL, Meredith WR, Campbell BT, Bourland FM, Gannaway JR, Glass KM, Zhang JF (2010) Fatty acid profiles of cottonseed genotypes from the national cotton variety trials. *J Cotton Sci* 14:64–73
- Elmore CD, Spurgeon WI, Thom WQ (1979) Nitrogen fertilization increases N and alters amino acid concentration of cottonseed. *Agron J* 71:713–716
- Feil B, Moser SB, Jampatong S, Stamp P (2005) Mineral composition of the grains of tropical maize varieties as affected by pre-anthesis drought and rate of nitrogen fertilization. *Crop Sci* 45:516–523
- Gipson JR, Joham HE (1969) Influence of night temperature on growth and development of cotton (*Gossypium hirsutum* L.). IV. Seed quality *Agron J* 61:365–367
- Hanny BW, Meredith WR, Bailey JC, Harvey AJ (1978) Genetic relationships among chemical constituents in seeds, flower buds, terminals, and mature leaves of cotton. *Crop Sci* 18:1071–1074

- Karaosmanoglu F, Tuter M, Gollu E, Yanmaz S, Altintig E (1999) Fuel properties of cottonseed oil. *Energy Sources* 21:821–828
- Kohel RJ (1980) Genetic studies of seed oil in cotton. *Crop Sci* 20:784–787
- Kohel RJ, Cherry JP (1983) Variation of cottonseed quality with stratified harvests. *Crop Sci* 23:1119–1124
- Kohel RJ, Lee JA (1984) Genetic analysis of Egyptian glandless cotton. *Crop Sci* 24:1119–1121
- Kohel RJ, Glueck J, Rooney LW (1985) Comparison of cotton germplasm collections for seed-protein content. *Crop Sci* 25:961–963
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Leffler HR, Elmore CD, Hesketh JD (1977) Seasonal and fertility-related changes in cottonseed protein quantity and quality. *Crop Sci* 17:953–956
- Matsubara K, Kono I, Hori K, Nonoue Y, Ono N, Shomura A, Mizubayashi T, Yamamoto S, Yamanouchi U, Shirasawa K, Nishio T, Yano M (2008) Novel QTLs for photoperiodic flowering revealed by using reciprocal backcross inbred lines from crosses between *japonica* rice cultivars. *Theor Appl Genet* 117:935–945
- McCouch SR, Chen XL, Panaud O, Temnykh S, Xu YB, Cho YG, Huang N, Ishii T, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol Biol* 35:89–99
- McMichael SC (1960) Combined effects of the glandless genes *gl₂*, and *gl₃*, on pigment glands in the cotton plant. *Agron J* 46:385–386
- Meneghetti SMP, Meneghetti MR, Serra TM, Barbosa DC, Wolf CR (2007) Biodiesel production from vegetable oil mixtures: cottonseed, soybean, and castor oils. *Energy Fuels* 21:3746–3747
- O'Brien RD, Jones LA, King CC, Wakelyn PJ, Wan PJ (2005) Cottonseed oil. In: Shahidi F (ed) *Bailey's industrial oil and fat products*, 6th edn. Wiley, New Jersey
- Percy RG, Kohel RJ (1999) Qualitative genetics. In: Smith CW, Cothren JT (eds) *Cotton: origin, history, technology, and production*. Wiley, New York, pp 319–360
- Shaver TN, Dilday (1982) Measurement of correlation among selected seed quality factor for 36 Texas race stocks of cotton. *Crop Sci* 22:779–781
- Song X, Zhang TZ (2007) Identification of quantitative trait loci controlling seed physical and nutrient traits in cotton. *Seed Sci Res* 27:243–251
- Turner JH, Ramey HH, Worley S (1976a) Influence of environment on seed quality of four cotton cultivars. *Crop Sci* 16:407–409
- Turner JH, Ramey HH, Worley S (1976b) Relationship of yield, seed quality, and fiber properties in Upland cotton. *Crop Sci* 16:578–580
- USDA-FAS (2011) Oilseeds: World markets and trade. Sept. 2011
- Wang XS, Lu Y, Wu JG (2001) Determination of oil component in cottonseeds with near infrared reflectance spectroscopy. *Acta Agriculturae Zhejiangensis* 13:218–222
- Wang L, Meng QX, Ren LP, Yang JS (2010) Near infrared reflectance spectroscopy and its application in the determination for the quality of animal feed and products. *Spectroscopy Spectral Ana* 30:1482–1487
- Wu J, Jenkins JN, McCarty JC, Thaxton P (2009) Seed trait evaluation of *Gossypium barbadense* L. chromosomes/arms in a *G. hirsutum* L. background. *Euphytica* 167:371–380
- Wu J, McCarty JC, Jenkins JN (2010) Cotton chromosome substitution lines crossed with cultivars: Genetic model evaluation and seed trait analyses. *Theor Appl Genet* 120:1473–1483
- Yamagishi J, Miyamoto N, Hirotsu S, Laza RC, Nemoto K (2004) QTLs for branching, floret formation, and pre-flowering floret abortion of rice panicle in a temperate *japonica* x tropical *japonica* cross. *Theor Appl Genet* 109:1555–1561
- Ye ZH, Lu ZZ, Zhu J (2003) Genetic analysis for developmental behavior of some seed quality traits in Upland cotton (*Gossypium hirsutum* L.). *Euphytica* 129:183–191
- Yu JW, Zhang K, Yu SX, Fan SL, Song MZ, Zhai HH, Li XL, Zhang JF (2012) Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* x *Gossypium barbadense* backcross inbred line population. (Submitted)
- Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457–1468
- Zhang JF, Stewart JM (2000) Economical and rapid method for extracting cotton genomic DNA. *J Cotton Sci* 4:193–201
- Zhu SJ, Reddy N, Jiang YR (2005) Introgression of a gene for delayed pigment gland morphogenesis from *Gossypium bickii* into upland cotton. *Plant Breed* 124:590–594