Seed Science Research (2007) **17**, 243–251 DOI: 10.1017/S0960258507834957

Identification of quantitative trait loci controlling seed physical and nutrient traits in cotton

Xian-Liang Song[1,2] and Tian-Zhen Zhang[1]*

[1]National Key Laboratory of Crop Genetics and Germplasm Enhancement, Cotton Research Institute, Nanjing Agricultural University, Nanjing 210095, China; [2]College of Agronomy, State Key Laboratory of Crop Biology, Shandong Agricultural University, Taian 271018, China

Abstract

Cotton (*Gossypium* spp.) is an important source of edible oil and protein meals. Complex genetics and strong environmental effects hinder progress in seed quality trait breeding in this species. The use of molecular markers can improve an understanding of the genetic factors conditioning seed quality traits, and is expected to assist in selection of superior genotypes. This study was conducted to identify quantitative trail loci (QTL) associated with seed physical and nutrient traits in cotton. To achieve this objective, a population of 140 BC_1S_1 lines developed from a cross between 'TM-1' and 'Hai7124' was evaluated in 2003 and 2004. A linkage map consisting of 918 markers from this population was used to identify QTL using QTLNetwork-2.0 software. Eleven single QTL were identified for kernel percentage, kernel oil percentage, kernel protein percentage and seven amino acids (Asp, Ser, Gly, Ile, Leu, Phe and Arg). Phenotypic variation explained by each individual QTL ranged from 10.89 to 46.28%. Two epistatic QTL for Cys and Leu were detected, explaining 9.55 and 4.43% of the phenotypic variation. These QTL detected for seed quality traits in cotton are expected to be useful for further breeding programmes targeting development of cotton with improved nutrient quality.

Keywords: amino acids, cotton, *Gossypium*, marker-assisted selection, quantitative trait loci mapping, seed oil, seed protein

Introduction

Cotton (*Gossypium* spp.) is the leading fibre crop in the world, and secondarily an important source of edible oil and protein meals. The five largest producers (China, 27%; United States, 12%; former Soviet Union, 10%; India, 11%; Pakistan, 9%) of cottonseed oil from 1995 to 2003 accounted for 70% of global output. Cottonseed kernels contain 27.83–45.6% protein and 28.24–44.05% oil (Sun *et al.*, 1987). In addition to flavour stability, cottonseed oil also has superior nutritive qualities; it has a 3:1 ratio of unsaturated to saturated fatty acids, which meets the recommendations of many health professionals. Cottonseed meal is used principally as a protein concentrate in feed for livestock (http://www.cottonseed.com/publications).

Protein and oil concentration, kernel index and kernel percentage in cotton are controlled by multiple genes (Singh *et al.*, 1985; Dani and Kohel, 1989; Ye *et al.*, 2003) and are strongly influenced by the environment (Kohel and Cherry, 1983; Singh *et al.*, 1985; Ye *et al.* 2003). Seed traits may be simultaneously controlled by seed nuclear genes, cytoplasmic genes and maternal nuclear genes (Ye *et al.*, 2003). Previous studies indicated significant negative coefficients between oil content and protein content (Kohel and Cherry, 1983; Chen *et al.*, 1986; Sun *et al.*, 1987). Such factors may hinder progress in improvement of these traits in conventional cotton breeding programmes.

Genetic mapping provides an essential tool to understand the genetic architecture of quantitative traits at the molecular level. DNA markers linked to quantitative trait loci (QTL) controlling seed protein content have been identified in soybean (Chung *et al.*, 2003; Panthee *et al.*, 2005), rice (Tan *et al.*, 2001), barley (See *et al.*, 2002) and field pea (Tar'an *et al.*, 2004). DNA markers associated with loci controlling seed oil content or fatty acid composition have been identified in soybean (Kianian *et al.*, 1999), rapeseed (Zhao *et al.*, 2006), sunflower (Bert *et al.*, 2003; Pérez-Vich *et al.*, 2004), oilseed mustard (Gupta *et al.*, 2004) and canola (Hu *et al.*, 2006). However, no such results have been reported in cotton.

*Correspondence
Fax: 0086 25 84395307
Email: cotton@njau.edu.cn

A major function of proteins in nutrition is to supply adequate amounts of required amino acids, which can be divided into two groups, essential and non-essential. Essential amino acids cannot be synthesized in animals, but play a crucial role in metabolic processes; they are lysine (Lys), histidine (His), leucine (Leu), isoleucine (Ile), valine (Val), methionine (Met), threonine (Thr), tryptophan (Trp) and phenylalanine (Phe) (D'Mello, 2003). Lysine is an important amino acid for humans and animals; cotton kernels contain on average 2.37% Lys (dry weight of kernel powder basis), higher than rice (2.15%) and lower than wheat (2.7%) (Chen *et al.*, 1986).

Despite the central role of amino acids in animal nutrition, little information is available about the genetic control of amino acid content in cotton. Molecular analysis provides a good tool to understand the genetics of a trait in detail, to optimize its improvement through breeding. Such work has been conducted in corn (Wang and Larkins, 2001) and soybean (Panthee *et al.*, 2006). However, there are no such reports for cotton.

The objective of the present study was to identify and map genomic regions associated with kernel index, kernel percentage, hull percentage, kernel oil percentage, kernel protein percentage and amino acid composition, to facilitate the selection of these traits in cotton trait introgression and breeding.

Materials and methods

Plant material and trait measurements

The 140 BC_1 mapping population was derived from a backcross of [TM-1 (*G. hirsutum* L.) × Hai7124 (*G. barbadense* L.)] × TM-1], the latter a genetic standard accession of *G. hirsutum*, kindly provided by Drs R.J. Kohel and J. Yu from the Southern Plains Agricultural Research Center, College Station, Texas. All BC_1 plants were grown in the field during summer seasons and in a greenhouse during the winter at the Jiangpu Cotton Research Station of Nanjing Agricultural University (JCRSNAU), Nanjing, China, and were self-pollinated to produce BC_1S_1 seeds. Grafting was performed to maintain the BC_1 mapping population and provide enough BC_1S_1 seeds.

All BC_1S_1 plants and the two parents were grown at JCRSNAU in 2003 and 2004 under field conditions. Each BC_1S_1 plot was planted in two fully randomized replications, each having two rows in 2003, and one row in 2004, measuring 5.5 m. Eight rows of parents were randomly planted among BC_1S_1. The row spacing and plant spacing were 0.8 m and 0.4 m in the 2 years.

Normally opened bolls were harvested, and their seeds ginned and acid-delinted. Seeds were dried at 38°C in a forced-air oven to equalize moisture contents among samples. They were then separated into seed hull and kernel to determine kernel index (KID), hull index (HID) and kernel percentage (KP). Kernel index and hull index were the weight (g) of 100 kernels and hulls, respectively. Assays for kernel oil percentage (OP) and protein (N × 6.25) percentage (PP) (on a dry weight of kernel powder basis) were made in accordance with the standard methods described in Ye *et al.* (2003). Kernels were ground into power and dried to equilibrium at 38°C to equalize moisture contents for amino acid analysis. Amino acid content (% dry weight of kernel powder) was determined using a Biochrom30 amino acid analyser (Biochrom Ltd., Cambridge, UK), according to the manufacturer's instructions. In total, 17 amino acids, including threonine (Thr), valine (Val), methionine (Met), isoleucine (Ile), leucine (Leu), histidine (His), lysine (Lys), phenylalanine (Phe), tryptophan (Trp), aspartic acid (Asp), serine (Ser), glutamic acid (Glu), proline (Pro), glycine (Gly), alanine (Ala), cysteine (Cys) and arginine (Arg), were examined.

Construction of the SSR linkage map

QTL mapping was conducted based on the cotton genetic map (Han *et al.*, 2006). This linkage map was first constructed by Song *et al.* (2005) with the same TM-1/Hai7124//TM-1 BC_1 population and enhanced with expressed sequence tag–simple sequence repeat (EST–SSR) technology by Han *et al.* (2004, 2006). The present map consists of 918 marker loci assorted to 32 linkage groups, covering 5136.5 cM with an average distance of 5.60 cM between adjacent markers (Song *et al.*, 2005). A complete assignment of chromosomes of cotton has been constructed by Wang *et al.* (2006) using translocation and bacterial artificial chromosome–fluorescence *in situ* hybridization (BAC–FISH) technology. Only segments of linkage groups associated with QTL detected in the present experiment are shown.

QTL mapping

QTL analyses were carried out with QTLNetwork-2.0 software (http://ibi.zju.edu.cn/software/qtlnetwork), using the mixed-model based composite interval mapping (MCIM) with a 10 cM window size and a 1 cM walking speed. A 10 cM filtration window was used to distinguish two adjacent test statistic peaks whether they are two QTLs or not. One thousand permutation tests were performed on all traits in the combined data from two environments to calculate the critical *F* value at the 5% probability level. The Monte Carlo Markov Chain (MCMC) was used to estimate QTL effects.

https://doi.org/10.1017/S0960258507834957 Published online by Cambridge University Press

Results

Phenotypic variation and trait correlation

Significant differences in KID, KP, HID, OP and PP between TM-1 and Hai7124 were detected during 2003 and 2004. TM-1 had higher values of KID, KP and PP, and lower values of HID and OP (on average) than Hai7124 (Table 1). Of the seventeen amino acids examined, four essential and five non-essential amino acids differed significantly between the two parents. Except for Cys, in which the skewing exceeded 1.0, all other traits were normally distributed (skew <1.00), and were therefore suitable for QTL analysis (Table 1). All the traits expressed transgressive segregation in both directions in the BC_1S_1 population.

Correlation coefficients based on average data between phenotypic traits are given in Table 2. Significant positive correlations were detected between KID and HID and KP, KP and PP and OP. PP showed significant negative correlation to OP. No correlation was found between HID and other traits. Positive correlations were detected between PP and all the amino acids.

Single-locus analysis of quantitative trait loci for seed quality

Using joint analysis, 11 significant QTL for 10 seed quality traits were identified (Table 3; Fig. 1); none for KID, HID, Cys or Val were detected.

On chromosome D12, a QTL, *qKP-D12*-1, for kernel percentage (KP) was identified between a SSR marker BNL1227_180 and a SRAP marker DC1SA21_B, explaining 46.28% of the phenotypic variation (PV). The allele from Hai7124 increased KP. For kernel oil percentage (OP), a significant QTL *qOP-D8*-1 was detected in a region of 9.2 cM between two SSR markers BNL3860_190 and NAU1369_400, explaining 29.35% of PV. The genotype of TM-1 was in the direction of increasing OP. For kernel protein percentage (PP), a significant QTL *qPP-D9*-1 was mapped between BNL1672_140 and BNL3031_190, explaining 22.25% of PV. The genotype of the heterozygote was in the direction of increasing PP.

For aspartic acid (Asp), a significant QTL *qAsp-A11*-1 was detected in a region of 7.6 cM on chromosome A11, explaining 22.12% of PV. The heterozygote genotype increased Asp percentage in the kernel. For serine (Ser), a significant QTL *qSer-A8*-1 was identified between NAU1531_170 and NAU537_220, explaining 23.66% of PV. The allele from the *Gh* parent TM-1 increased kernel Ser percentage. QTL *qGly-A11*-1 and *qGly-A8*-1 were significant and detected for glycine (Gly). The former explained 16.97% of PV, while the latter explained 10.89% of PV. The positive alleles at *qGly-A11*-1 came from Hai7124 and the positive alleles at *qGly-A8*-1 from TM-1.

For isoleucine (Ile), a significant QTL *qIle-D3*-1 was mapped in the region of 9.4 cM between NAU1028_225 and BNL359_180, explaining 25.88% of PV. The heterozygote genotype was in the direction of increasing kernel Ile percentage. For leucine (Leu), a significant QTL was detected between SSR markers BNL3145_290 and NAU652_105 on chromosome D2, explaining 20.32% of PV. The genotype of the heterozygote was in the direction of increasing the kernel Leu percentage. For phenylalanine (Phe), a

Table 1. Cotton seed quality descriptive statistics of the BC_1S_1 population and its parents. Kernel percentage (KP) is based on dry weight of acid-delinted seeds; kernel oil (OP), protein (PP) and amino acid percentages are based on kernel powder dry weight. Data for all the traits are the mean of two environments

	BC_1S_1			Parents		
Trait	Range	Mean ± SD	Skew	TM-1	7124	*t*-value
Kernel index (KID, g)	5.091–8.78	7.081 ± 0.677	− 0.048	8.415	8.079	4.514*
Hull index (HID, g)	3.182–4.587	3.954 ± 0.021	− 0.382	4.144	4.281	− 3.012*
Kernel (KP, %)	59.201–67.923	63.900 ± 0.106	− 0.406	67.003	65.147	3.198*
Kernel oil (OP, %)	28.970–40.099	36.177 ± 0.152	− 0.656	35.728	37.441	− 6.157**
Kernel protein (PP, %)	32.612–47.125	39.301 ± 0.205	0.340	43.723	41.063	3.940*
Aspartic acid (%)	2.34–4.81	3.781 ± 0.056	0.179	3.204	2.738	3.132*
Serine (%)	0.995–2.09	1.656 ± 0.02	− 0.386	1.540	1.300	5.353**
Glycine (%)	1.460–2.72	1.817 ± 0.022	0.043	1.668	2.022	− 5.961**
Cysteine (%)	0.697–1.537	1.153 ± 0.024	1.101	0.824	1.256	− 5.812**
Valine (%)	1.855–3.61	2.250 ± 0.041	0.535	2.122	1.716	5.195**
Isoleucine (%)	0.985–2.21	1.749 ± 0.045	0.681	1.062	1.328	− 4.925**
Leucine (%)	1.680–4.02	2.761 ± 0.049	0.514	2.056	2.510	− 3.086*
Phenylalanine (%)	1.447–2.91	2.056 ± 0.028	− 0.315	1.902	1.430	5.121**
Arginine (%)	3.055–6.31	4.298 ± 0.067	− 0.270	3.568	3.272	3.174*

t-values: *difference significant at $\alpha = 0.05$; **difference significant at $\alpha = 0.01$.

https://doi.org/10.1017/S0960258507834957 Published online by Cambridge University Press

Table 2. Sample phenotypic correlation coefficients between cotton seed quality traits in an interspecific BC_1F_1 population of TM-1 × Hai7124. See Table 1 for basis of percentages

	KID	HID	KP	PP	OP	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Cys	Val	Met	Ile	Leu	Tyr	Phe	His	Lys
HID	0.70**																				
KP	0.72**	0.12																			
PP	0.07	0.07	0.51**																		
OP	0.14	0.06	0.64**	−0.58**																	
Asp	−0.05	−0.16	0.29*	0.61**	−0.22																
Thr	−0.36*	−0.15	0.17	0.718*	−0.18	−0.26*															
Ser	−0.07	−0.17	0.37*	0.53**	−0.21	−0.15	0.95**														
Glu	−0.08	−0.18	0.27*	0.37*	−0.13	0.77**	−0.46**	−0.37*													
Pro	0.16	0.06	0.03	0.82**	−0.08	0.45**	0.6**	0.52**	0.11												
Gly	−0.04	−0.14	0.18	0.64**	−0.24*	0.08	0.48**	0.66**	−0.46**	0.37*											
Ala	−0.21	−0.11	0.17	0.34*	0.15	0.49**	0.37*	0.33*	0.04	0.75**	0.87**										
Cys	−0.02	−0.11	0.08	0.63**	0.13	−0.25	0.54**	0.69**	0.29*	−0.11	0.11	−0.48**									
Val	0.02	−0.1	0.13	0.71**	0.08	0.34*	0.51**	0.56**	0.18	0.67**	0.82**	0.92**	0.38*								
Met	0.31*	−0.07	0.08	0.52**	−0.26*	0.13	0.31*	0.46**	0.06	0.38*	0.01	0.18	0.48*	0.27							
Ile	−0.01	−0.13	0.12	0.55**	−0.18	0.88**	−0.24	−0.37*	0.8**	0.34	0.15	0.60**	−0.81**	0.46**	−0.17						
Leu	−0.12	−0.14	0.11	0.28*	−0.09	0.45**	−0.42**	−0.58**	0.89**	−0.25	0.44*	0.02	0.51**	0.08	−0.61**	0.67*					
Tyr	−0.24*	−0.15	0.09	0.45**	−0.34*	−0.26*	0.65**	0.64**	−0.41*	0.34*	0.68**	0.55**	0.05	0.62**	−0.36*	−0.16	−0.14				
Phe	0.12	0.08	0.1	0.52**	−0.35*	0.08	0.53**	0.52**	−0.13	0.36*	0.01	0.26	0.63**	0.01	0.33*	−0.3	0.18	0.31*			
His	−0.03	0.12	0.07	0.65**	−0.25	−0.26*	0.75**	0.76*	−0.26	0.35*	0.28	0.03	0.76**	0.19	0.54**	−0.58**	−0.68**	0.36*	0.82**		
Lys	0.14	−0.16	0.1	0.38*	0.09	0.21	−0.42*	−0.32*	0.41*	−0.22	−0.42*	−0.47**	0.3	−0.43*	0.14	0.02	0.33*	−0.6	0.22	−0.11	
Arg	−0.08	−0.1	−0.02	0.49**	−0.1	−0.17	0.41*	0.54**	−0.21	−0.11	0.39*	0.05	0.52**	0.39*	−0.26	−0.17	−0.17	0.24	0.38*	0.54**	0.07

*,**: Significant at 0.05, 0.01 level, respectively.

Table 3. QTL mapping for cotton seed quality traits in 2 years (environments), detected by QTLNetwork-2.0 (joint analysis). See Table 1 for basis of percentages

Trait	QTL	Interval	Position	Range[a]	A[b]	SE	*P* value	H^2 (a)[c]
Kernel (%)	*qKP-D12*-1	BNL1227_180–DC1SA21_B	32.8	32.4–33.8	−2.6200	0.0017	0.000000	0.4628
Kernel oil (%)	*qOP-D8*-1	BNL3860_190– NAU1369_400	53.7	48.7–57.9	1.5971	0.1549	0.000000	0.2935
Kernel protein (%)	*qPP-D9*-1	BNL1672_140– BNL3031_190	68.8	62.9–73.5	−2.46	0.0029	0.000000	0.2225
Aspartic acid (%)	*qAsp-A11*-1	NAU1162_300–BNL1681_110	71.1	65.5–73.1	−0.5825	0.0653	0.000000	0.2212
Serine (%)	*qSer-A8*-1	NAU1531_170– NAU537_220	35.1	33–42.6	0.2421	0.0261	0.000000	0.2366
Glycine (%)	*qGly-A11*-1	NAU1162_300– BNL1681_110	71.1	69.4–73.1	−0.4724	0.0443	0.000000	0.1697
	qGly-A8-1	BNL3627_180– BNL3792_225	22.2	14.6–24.1	0.3875	0.0433	0.000000	0.1089
Isoleucine (%)	*qIle-D3*-1	NAU1028_225– BNL359_180	7.2	3–12.4	−0.4911	0.0573	0.000000	0.2588
Leucine (%)	*qLeu-D2*-1	BNL3145_290– NAU652_105	15.0	14.2–16.7	−0.4676	0.0550	0.000000	0.2032
Phenylalanine (%)	*qPhe-A8*-1	BNL3792_225– BNL3474_193	24.1	23.2–26.1	0.3091	0.0260	0.000000	0.3113
Arginine (%)	*qArg-A5*-1	NAU1190_205– NAU797_170	82.9	80.7–90.8	−0.5832	0.0677	0.000000	0.1821

[a] Range means the support interval of QTL position.
[b] A is the additive genetic effects estimated at the testing points. Positive values mean that the TM-1 genotype has a positive effect on the trait. Negative values indicate that the heterozygote genotype has a positive effect on the trait.
[c] H^2(a) represents the phenotypic variations explained by the A value.

significant QTL *qPhe-A8*-1 was detected between SSR markers BNL3792_225 and BNL3474_193 on chromosome A8, explaining 31.13% of PV. The positive allele came from TM-1. For arginine (Arg), a significant QTL was identified in a region of 10.1 cM between two SSR markers NAU1190_205 and NAU797_170 on chromosome A5, explaining 18.21% of PV. The genotype of the heterozygote was in the direction of increasing the kernel Arg percentage.

Epistatic quantitative trait loci for amino acid composition

Two interactions for Cys and Leu were detected between the same two marker intervals, CIR099_80–CIR099_90 on D13 and Y1278B–CIR156_150 on D11, respectively. The former explained 9.55% of PV, while the latter explained 4.43% of PV (Table 4, Fig. 1). This indicates that the two regions on D11 and D13 were associated with Cys and Leu simultaneously, which agrees with the significant positive correlation between Cys and Leu. Such effects may have been due to linkage of multiple QTL, or to pleiotropic effects of a single gene on multiple traits (Jiang *et al.*, 2000).

Discussion

The complex genetic base for cottonseed quality traits has limited the progress of conventional breeding methods. Additionally, little attention has been paid to breeding for high nutrition quality lines. Molecular markers linked to QTL controlling seed oil content (Diers *et al.*, 1992; Mansur *et al.*, 1993; Goldman *et al.*, 1994; Alrefai *et al.*, 1995; Tanhuanpaa *et al.*, 1995a, b; Brummer *et al.*, 1997; Kianian *et al.*, 1999), seed protein content (Chung *et al.*, 2003; Panthee *et al.*, 2005) and amino acids (Wang and Larkins, 2001; Panthee *et al.*, 2006) have been identified in soybean, rape seed, maize and oat. Here is the first report on QTL mapping of cotton-seed nutritional quality traits based on a new SSR genetic map using a *G. hirsutum* × *G. barbadense* population. Eleven QTL and two epistatic interaction effects were identified for 11 seed quality traits based on 2 years of phenotypic data. Both *G. hirsutum* and *G. barbadense* contributed some alleles that increase, and others that decreased KP, PP, OP and amino acids, suggesting that there exists substantial opportunity to breed cotton varieties that transgress the current range of seed phenotypes. Molecular markers identified herein provide a useful base for genetic manipulation of these seed quality traits in breeding programmes.

Cotton-seed kernels contain a high percentage of protein, which is mainly water soluble (Zhang *et al.*, 1998). Zhu *et al.* (1995) showed that cotton-seed protein is mainly composed of subunits of 46, 41, 30, 27, 23, 16 and 14 kDa; and protein subunits of 34 and 18 kDa are specific for varieties with the A and AD genomes, respectively. Significant variation in protein components and relative content of the protein subunits were also investigated among varieties. Electrophoresis of cotton-seed ethanol-soluble proteins detected three special bands distinctive to all eight *G. hirsutum* varieties studied and three special bands distinctive to all four *G. barbadense* varieties (Gao *et al.*, 2003). These studies indicate a high polymorphism in cotton-seed protein components. In this paper, only one significant QTL (*qPP-D9*-1) for total protein percentage was identified (Table 3). This may be because total protein content did not reflect large variations in protein components between the

https://doi.org/10.1017/S0960258507834957 Published online by Cambridge University Press

two parents. The same reason may explain partly why only one QTL for kernel oil content was detected. The original purpose of breeding a *G. hirsutum* × *G. barbadense* population was to construct a saturated molecular genetic map of tetraploid cotton and to integrate the high yield of *G. hirsutum* and good fibre quality of *G. barbadense*; thus, little attention was paid to seed quality variation in selecting parental varieties, which may have resulted in low polymorphism at loci controlling seed quality traits.

The nutritional quality of protein is determined by its amino acid composition. Panthee *et al.* (2006) detected genomic regions associated with amino acid composition in soybean, and found that most of the amino acid QTL had been reported previously as protein QTL in one or more populations. In the present study, a single QTL and/or epistatic QTL was associated with three essential amino acids Ile, Leu and Phe, as well as five non-essential amino acids Ser, Asp, Gly, Arg and Cys in cotton seeds (Tables 3 and 4). These QTL may be useful to obtain balanced amino acid protein contents by breeding.

Only QTL with associations in two environments (two years) are reported here, but most of them could explain more than 10% of PV, indicating that they are major QTL (Falconer and Mackay, 1996). However, it is recognized that it is necessary to validate these putative QTL in more environments and populations before using them in marker-assisted selection.

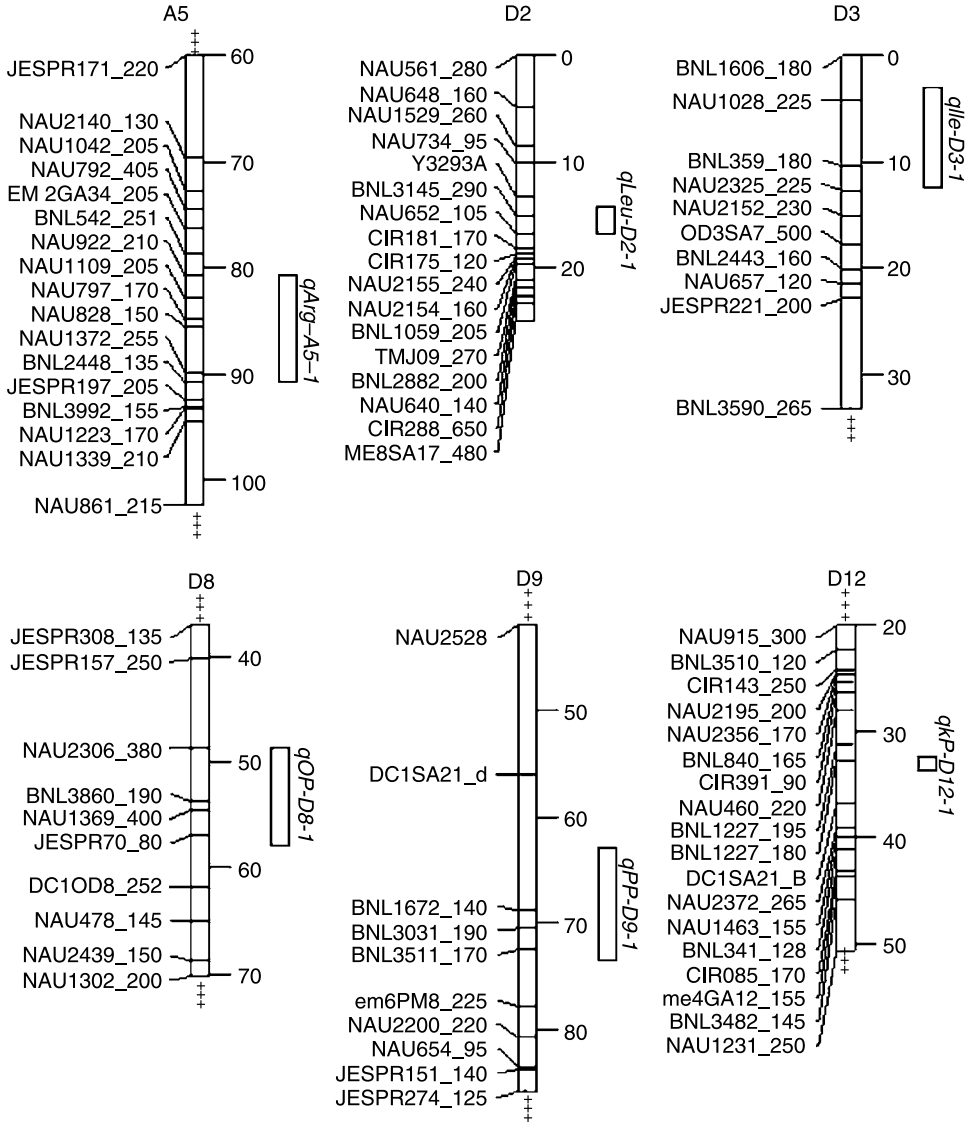

Figure 1. (*continued*)

https://doi.org/10.1017/S0960258507834957 Published online by Cambridge University Press

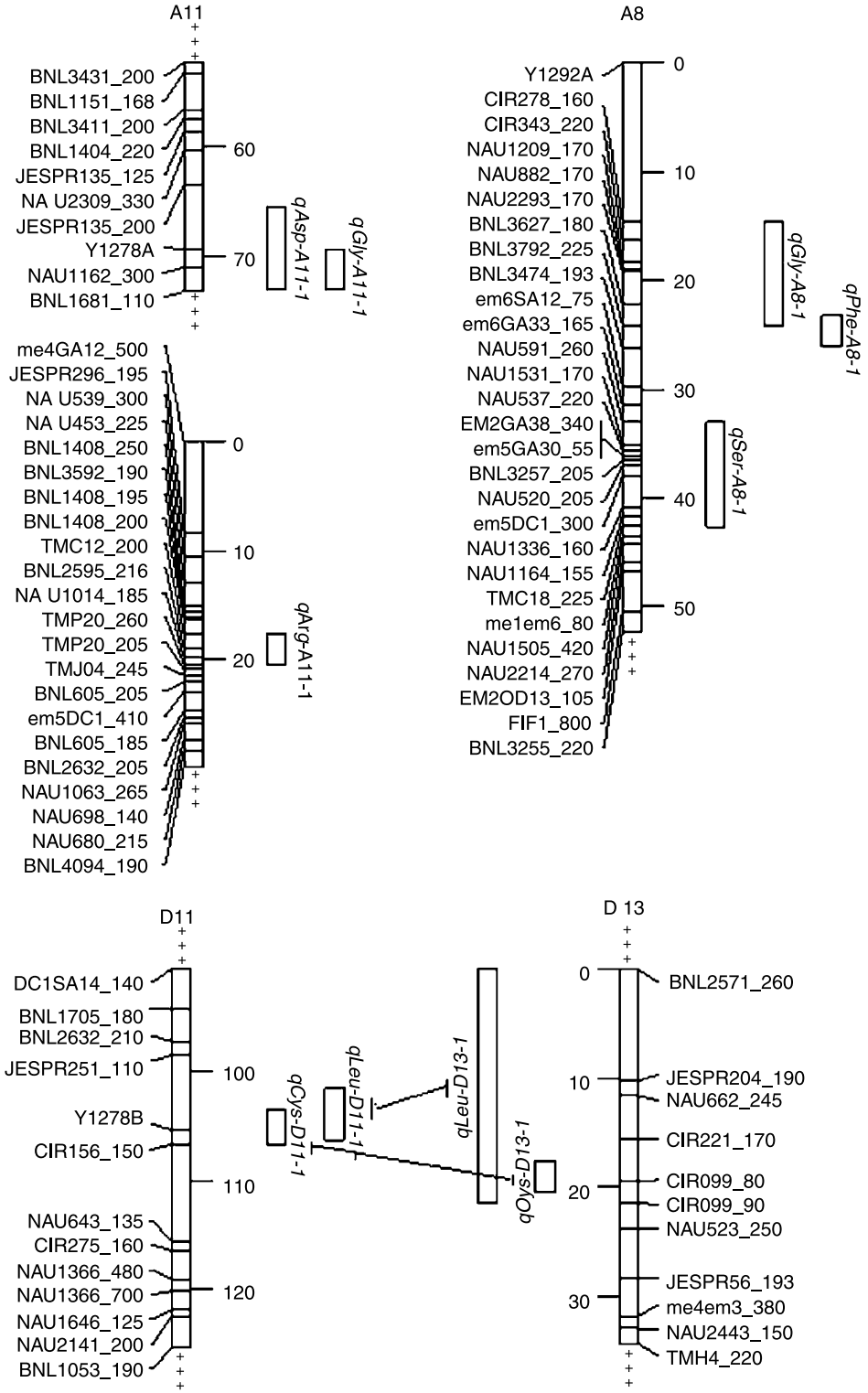

Figure 1. Significant quantitative trait loci (QTL) for cotton-seed quality traits detected by QTLNetwork-2.0 using joint analysis. A line between linkage groups indicates the epistatic QTL. Only segments of linkage groups related to QTL detected in this study are shown; for detailed information of the BC_1 map see Han *et al.* (2006).

https://doi.org/10.1017/S0960258507834957 Published online by Cambridge University Press

Table 4. Epistatic QTL detected by QTLNetwork-2.0 for cysteine and leucine percentage in kernels of cotton

Trait	Chr[a]	Interval	Position	Range[b]	AA	SE	*P* value	H^2(aa)[c]
Cysteine (%)	D13	CIR099_80–CIR099_90	19.5	17.6–20.5	0.1832	0.0337	0.000000	0.0955
	D11	Y1278B–CIR156_150	105.3	103.5–106.7				
Leucine (%)	D13	CIR099_80–CIR099_90	19.5	0.0–21.5	0.3089	0.0762	0.000050	0.0443
	D11	Y1278B–CIR156_150	105.3	101.5–106.3				

[a] Chr is the chromosome number of the points tested.
[b] Ranges are the position support intervals of the two QTL.
[c] H^2 (aa) are the phenotypic variations explained by AA.

Acknowledgements

This program was supported financially by the National High-tech Program (2004AA211172), National Science Foundation in China (30070483, 30270806), the Key Project of Chinese Ministry of Education (10 418), Postdoctoral Science Foundation of China (2005037736) and Jiangsu High-tech Project (BG2002306) and IAEA (12 846).

References

Alrefai, R., Berke, T.G. and Rocheford, T.R. (1995) Quantitative trait locus analysis of fatty acid concentrations in maize. *Genome* **38**, 894–901.

Bert, P.-F., Jouan, I., Tourvieille de Labrouhe, D., Serre, F., Philippon, J., Nicolas, P. and Vear, F. (2003) Comparative genetic analysis of quantitative traits in sunflower (*Helianthus annuus* L.). 2. Characterisation of QTL involved in developmental and agronomic traits. *Theoretical and Applied Genetics* **107**, 181–189.

Brummer, E.C., Graef, G.L., Orf, J., Wilcox, J.R. and Shoemaker, R.C. (1997) Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Science* **37**, 370–378.

Chen, Z.F., Zhang, Z.W. and Cheng, H.L. (1986) The analysis of Upland cotton quality. *Acta Agronomica Sinica* **12**, 195–200.

Chung, J., Babka, H.L., Graef, G.L., Staswick, P.E., Lee, D.J., Cregan, P.B., Shoemaker, R.C. and Specht, J.E. (2003) The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Science* **43**, 1053–1067.

Dani, R.G. and Kohel, R.J. (1989) Maternal effects and generation mean analysis of seed-oil content in cotton (*Gossypium hirsutum* L.). *Theoretical and Applied Genetics* **77**, 569–575.

Diers, B.W., Keim, P., Fehr, W.R. and Shoemaker, R.C. (1992) RFLP analysis of soybean seed protein and oil content. *Theoretical and Applied Genetics* **83**, 608–612.

D'Mello, J.P.F. (2003) Amino acids as multifunctional molecules. pp. 1–4 *in* D'Mello, J.P.F. (Ed.) *Amino acids in animal nutrition* (2nd edition). Wallingford, UK, CABI Publishing.

Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to quantitative genetics* (4th edition). Harlow, UK, Longmans Green.

Gao, G.Q., Lu, T.X., Su, X.H., Liu, Y.X., Wu, D.F. and Zhu, B.D. (2003) Seed protein electrophoresis for identification of fine fiber cotton line. *Acta Agriculturae Nucleatae Sinica* **5**, 340–342.

Goldman, I.L., Rocheford, T.R. and Dudley, J.W. (1994) Molecular markers associated with maize kernel oil concentration in an Illinois high protein × Illinois low protein cross. *Crop Science* **34**, 908–915.

Gupta, V., Mukhopadhyay, A., Arumugam, N., Sodhi, Y.S., Pental, D. and Pradhan, A.K. (2004) Molecular tagging of erucic acid trait in oilseed mustard (*Brassica juncea*) by QTL mapping and single nucleotide polymorphisms in *FAE1* gene. *Theoretical and Applied Genetics* **108**, 743–749.

Han, Z.G., Guo, W.Z., Song, X.L. and Zhang, T.Z. (2004) Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. *Molecular Genetics and Genomics* **272**, 308–327.

Han, Z.G., Wang, C.B., Song, X.L., Guo, W.Z., Gou, J.Y., Li, C.H., Chen, X.Y. and Zhang, T.Z. (2006) Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. *Theoretical and Applied Genetics* **112**, 430–439.

Hu, X., Sullivan-Gilbert, M., Gupta, M. and Thompson, S.A. (2006) Mapping of the loci controlling oleic and linolenic acid contents and development of *fad2* and *fad3* allele-specific markers in canola (*Brassica napus* L.). *Theoretical and Applied Genetics* **113**, 497–507.

Jiang, C., Wright, R.J., Woo, S.S., Delmonte, T.A. and Paterson, A.H. (2000) QTL analysis of leaf morphology in tetraploid *Gossypium* (cotton). *Theoretical and Applied Genetics* **100**, 409–418.

Kianian, S.F., Egli, M.A., Phillips, R.L., Rines, H.W., Somers, D.A., Gengenbach, B.G., Webster, F.H., Livingston, S.M., Groh, S., O'Donoughue, L.S., Sorrells, M.E., Wesenberg, D.M., Stuthman, D.D. and Fulcher, R.G. (1999) Association of a major groat oil content QTL and an acetyl-CoA carboxylase gene in oat. *Theoretical and Applied Genetics* **98**, 884–894.

Kohel, R.J. and Cherry, J.P. (1983) Variation of cottonseed quality with stratified harvests. *Crop Science* **23**, 1119–1124.

Mansur, L.M., Lark, K.G., Kross, H. and Oliveira, A. (1993) Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). *Theoretical and Applied Genetics* **86**, 907–913.

Panthee, D.R., Pantalone, V.R., West, D.R., Saxton, A.M. and Sams, C.E. (2005) Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. *Crop Science* **45**, 2015–2022.

Panthee, D.R., Pantalone, V.R., Saxton, A.M., West, D.R. and Sams, C.E. (2006) Genomic regions associated with amino acid composition in soybean. *Molecular Breeding* **17**, 79–89.

Pérez-Vich, B., Knapp, S.J., Leon, A.J., Fernández-Martínez, J.M. and Berry, S.T. (2004) Mapping minor QTL for increased stearic acid content in sunflower seed oil. *Molecular Breeding* **13**, 313–322.

See, D., Kanazin, V., Kephart, K. and Blake, T. (2002) Mapping genes controlling variation in barley grain protein concentration. *Crop Science* **42**, 680–685.

Singh, M., Singh, T.H. and Chahal, G.S. (1985) Genetic analysis of some seed quality characters in upland cotton (*Gossypium hirsutum* L.). *Theoretical and Applied Genetics* **71**, 126–128.

Song, X.L., Wang, K., Guo, W.Z., Zhang, J. and Zhang, T.Z. (2005) A comparison of genetic maps constructed from haploid and BC_1 mapping populations from the same crossing between *Gossypium hirsutum* L. and *Gossypium barbadense* L. *Genome* **48**, 378–390.

Sun, S.K., Chen, J.H., Xian, S.K. and Wei, S.J. (1987) Study on the nutritional quality of cotton seeds. *Scientia Agricultura Sinica* **5**, 12–16.

Tan, Y.F., Sun, M., Xing, Y.Z., Hua, J.P., Sun, X.L., Zhang, Q.F. and Corke, H. (2001) Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. *Theoretical and Applied Genetics* **103**, 1037–1045.

Tanhuanpaa, P.K., Vilkki, J.P. and Vilkki, H.J. (1995a) Identification of a RAPD marker for palmitic acid concentration in the seed of spring turnip rape (*Brassica rapa* ssp *oleifera*). *Theoretical and Applied Genetics* **91**, 477–480.

Tanhuanpaa, P.K., Vilkki, J.P. and Vilkki, H.J. (1995b) Association of a RAPD marker with linolenic acid concentration in the seed oil of rapeseed (*Brassica napus* L.). *Genome* **38**, 414–416.

Tar'an, B., Warkentin, T., Somers, D.J., Miranda, D., Vandenberg, A., Blade, S. and Bing, D. (2004) Identification of quantitative trait loci for grain yield, seed protein concentration and maturity in field pea (*Pisum sativum* L.). *Euphytica* **136**, 297–306.

Wang, K., Song, X.L., Han, Z.G., Guo, W.Z., Yu, J.Z., Sun, J., Pan, J.J., Kohel, R.J. and Zhang, T.Z. (2006) Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence *in situ* hybridization mapping. *Theoretical and Applied Genetics* **113**, 73–80.

Wang, X.L. and Larkins, B.A. (2001) Genetic analysis of amino acid accumulation in *opaque-2* maize endosperm. *Plant Physiology* **125**, 1766–1777.

Ye, Z.H., Lu, Z.Z. and Zhu, J. (2003) Genetic analysis for developmental behavior of some seed quality traits in Upland cotton (*Gossypum hirsutum* L.). *Euphytica* **129**, 183–191.

Zhang, C.Q., Yin, Y.P., Gao, R.Q. and Jia, J.Z. (1998) Polymorphism of seed protein and identification of cotton cultivar. *Scientia Agricultura Sinica* **31**, 16–19.

Zhao, J.Y., Becker, H.C., Zhang, D.Q., Zhang, Y.F. and Ecke, W. (2006) Conditional QTL mapping of oil content in rapeseed with respect to protein content and traits related to plant development and grain yield. *Theoretical and Applied Genetics* **113**, 33–38.

Zhu, X.C., Zhao, W.M. and Liao, X.R. (1995) Varieties of *Gossypium* seed protein components. *Acta Gossypium Sinica* **4**, 223–225.

Received 14 November 2006
accepted after revision 8 August 2007