QTL verification of grain protein content and its correlation with oil content by using connected RIL populations of high-oil maize

G.H. Yang, Y.B. Dong, Y.L. Li, Q.L. Wang, Q.L. Shi and Q. Zhou

Key Laboratory of Physiological Ecology and Genetic Improvement of Field
Crops in Henan Province, College of Agriculture,
Henan Agricultural University, Zhengzhou, Henan, China

Corresponding author: Y.L. Li
E-mail: yuling_li@126.com

ABSTRACT. Maize with high grain protein and oil contents offers great advantages for human food and animal feed. In this study, grain protein contents of 282 and 263 $F_{7:8}$ recombinant inbred lines (RILs) of 2 crosses were evaluated in 4 environments within and between populations. The RILs were developed from crosses between an inbred high-oil maize line and 2 normal dent inbred maize lines. A total of 16 single-population QTLs and 19 joint-population QTLs were identified for protein content, and 21 QTLs were detected for protein-oil in each of the 4 environments tested and in combination. Most of the QTLs for protein content were greatly influenced by variation among populations and environments. Seven QTLs showed generational consistency compared with QTLs detected in the 2 $F_{2:3}$ populations. However, 7 and 6 QTLs were detected in only the RIL and $F_{2:3}$ populations, respectively. Protein and protein-oil QTLs with the same parental effects were detected at bins 3.03-3.05, 5.04-5.06, 6.03-6.05, 8.03-8.04, and 8.04-8.06, demonstrating that tightly linked and/or pleiotropic QTLs are controlling both traits at these bins. Four single-population QTLs and

11 joint-population QTLs identified at bins 3.02-3.03, 3.05, 7.01, 8.02, 8.03, 8.04-8.05, 8.05, 9.03, and 9.05 with intervals <5 cM could be used in marker-assisted selection. Along with the previously detected QTLs qPRO1-8-1 and qPRO1-5-1 at bins 8.03-8.04 and 5.02-5.04, the QTLs detected herein could be used to develop near isogenic lines and chromosome segment substitution lines in future studies.

Key words: High-oil maize; Grain protein and oil content; RIL population; QTL verification; Joint QTL mapping; QTL consistency

INTRODUCTION

Maize is one of the most widely planted crops with multiple uses throughout the world. Maize grain with increased protein and oil contents has high nutritional value for human food, animal feed, and industrial processing (Han et al., 1987; Lambert, 2001). Since Hopkins (1899) first began to investigate selection of maize grain chemical compositions over 100 years ago, several germplasms have been developed, including Illinois high protein (IHP), Illinois low protein (ILP), Illinois high oil (IHO), and Illinois low oil (ILO) (Dudley et al., 1977; Dudley and Lambert, 1992, 2004) germplasms. To date, many researchers have demonstrated the genetic mechanisms contributing to grain protein and oil contents by using statistical and molecular models (Miller and Brimhall, 1951; Dudley et al., 1977; Dudley and Lambert, 1992, 2004; Goldman et al., 1993; Berke and Rocheford, 1995; Song et al., 2004; Wassom et al., 2008; Zhang et al., 2008; Li et al., 2009; Wang et al., 2010).

Using traditional quantitative genetic and statistical models, Dudley and Lambert (1992) revealed that the grain protein and oil content in maize was controlled by multiple genes with small individual and additive effects (Dudley et al., 1977). Miller and Brimhall (1951) reported that the total grain oil content was not related to the total grain protein content and was instead strongly positively related with grain protein content in the germ. However, more recently, several studies have demonstrated positive correlations between grain protein and oil contents (Dudley and Lambert, 2004; Wassom et al., 2008; Zhang et al., 2008). Further research has attempted to reveal the genetic mechanisms underlying kernel composition traits by mapping QTLs using different genetic germplasms such as IHP and ILP (Goldman et al., l993; Dudley et al., 2004), IHO and ILO (Alrefai et al., 1995; Berke and Rocheford, 1995; Laurie et al., 2004; Willmot et al., 2006; Wassom et al., 2008), Beinongda high-oil (BHO) (Song et al., 2004; Zhang et al., 2008), tropical high-oil (Mangolin et al., 2004), Alexho Single-Kernel (ASK) high-oil (Zheng et al., 2008; Li et al., 2009; Wang et al., 2010; Yang et al., 2012), and popcorn (Liu et al., 2008) maize populations. Based on data obtained from these studies and from the public genetic linkage map of maize (www.maizegdb.org), 79 QTLs for grain protein content and 52 QTLs for grain oil content have been detected. One QTL for grain oil content (qH06) has been cloned (Zheng et al., 2008).

Although QTLs for grain protein and oil content have been detected at the same or adjacent chromosome locations, effects of genetic background, environmental factors, and their interactions are also commonly observed (Song et al., 2004; Liu et al., 2008; Zhang et al., 2008; Li et al., 2009). Therefore, the genetic mechanism underlying grain protein and oil content can be best revealed by using many different populations. In our previous study, 2 connected $F_{2:3}$ popu-

lations were used to identify QTLs for grain protein and oil content, which were developed from 2 crosses between 1 inbred high-oil and 2 normal inbred maize lines (Li et al., 2009; Wang et al., 2010). Two connected recombinant inbred line (RIL) populations were previously derived and used to detect QTLs for ear kernel traits and grain oil content (Yang et al., 2012). In this study, these 2 RIL populations were used to map single-population and joint-population single-trait QTLs for grain protein content, as well as multiple-trait QTLs for grain protein and oil contents. Our first objective was to analyze the influence of the genetic backgrounds, environments, and generations on QTL detection for grain protein content. The second objective was to further determine the genetic mechanism controlling grain protein content and its correlation with grain oil content in high-oil maize. The third objective was to find consistent QTLs for further research on QTL cloning and marker assisted breeding for grain protein and oil content in high-oil maize.

MATERIAL AND METHODS

Population development

One individual of the inbred high-oil maize GY220 line was crossed with 2 normal individuals of the inbred maize 8984 and 8622 lines to obtain 2 connected RIL populations: 8984 x GY220 and 8622 x GY220. GY220 was derived from the cycle 27 Alexander high-oil maize population, a member of the Lancaster heterotic group, and was selected and provided by the China Agricultural University (Jiang et al., 2005; Li et al., 2009). The 2 normal dent maize inbreds (8984 and 8622), belonging to the Chinese Reid heterotic group, were developed from 2 different open-pollinated hybrid F_1 crosses in our laboratory. The 2 crosses were self-pollinated using the single-seed descent method to produce 282 and 263 $F_{7:8}$ RILs, which were designated as Pop.1 and Pop.2, respectively. These 2 populations were used for QTL detection of ear kernel traits and grain oil content in our previous study (Yang et al., 2012).

Field experiment and trait evaluation

The 282 and 263 $F_{7:8}$ RIL populations, along with their respective parent lines, were planted according to an α-design in 2 adjacent trials with 1-row plots and 2 replications. The field trial was conducted under the same environmental conditions for each population in the summer sowings (June 12) of 2009 at 4 locations in Henan, Xuchuang, Xinxiang, and Zhengzhou and in spring sowing (April 15) at Yinchuan, China. Each row was 4 m long with 0.67 m between rows. Plots were planted by hand at a density of 60,000 plants/ha. Standard cultivation management practices were used at each location (Yang et al., 2012).

Three plants were sib-pollinated within each plot by hand to avoid the xenia effect. After maturity, the 3 plants were harvested and their ears were naturally dried (Yang et al., 2012). Grain protein and oil content was measured for grain samples mixed within each plot by using a MATRIX-1 NIR spectroscope (Bruker Corporation, Germany) according to methods described by Dudley and Lambert (1992).

Phenotypic data analysis

In order to verify the consistency of data across the 4 environments for grain pro-

tein content in both RIL populations, preliminary combined analysis of variance was used following standard procedures using a mixed model with a random genetic effect, fixed environment effect, and replicate effects, according to the Henderson III method (Henderson, 1953). Since the genotype x environment interactions were significant for grain protein content in both RIL populations, data were analyzed separately for each environment in further analyses. Correlation coefficients between protein content and oil content were calculated using the statistical software package SPSS 12.0. Broad-sense heritabilities (h^2_B) for the 2 connected RIL populations were calculated on an entry mean basis by dividing the genotypic variance by the phenotypic variance (Hallauer and Miranda, 1981). Confidence intervals on heritability estimates were determined according to the method reported by Knapp et al. (1985).

Single-population and joint-population QTL analyses

The integrated genetic map for the 2 populations has previously been described by Yang et al. (2012); this map includes 313 simple sequence repeat markers and was determined to be 2349.4 cM long with an average interval of 7.50 cM by using the BioMercator 2.1 software (Arcade et al., 2004).

QTL mapping and the estimation of each effect for each trait within populations was applied to data under each environmental condition and combined across the 4 environments in both connected RIL populations by composite interval mapping (CIM; Zeng, 1994). Thresholds for the logarithm of odds (LOD) score to identify QTLs were estimated by permutation tests with a minimum of 1000 replicates (Churchill and Doerge, 1994). The stepwise regression option of the CIM control was performed in Model 6 of the Zmapqtl procedure in QTL Cartographer, version 2.5, by using genetic background parameters and a window size of 10 cM on either side of the markers flanking the test site. QTL positions were assigned to the relevant regions at the point of the maximum LOD score. If 2 peaks for the same trait on the same chromosome were observed, and at least 2 markers separated the peaks at a minimum distance of 20 cM, they were considered to be different QTLs.

Based on the CIM results, interactions among QTLs were analyzed using the multiple interval method (MIM) in WinQTLCart (Kao et al., 1999; Wang et al., 2006). Joint-trait QTL analysis for grain protein and oil contents was analyzed according to the MIM method in WinQTLCart (Jiang and Zeng, 1995; Kao et al. 1999) with Cartographer, version 2.5 (Wang et al., 2006). A significance threshold was identified by the quick method for computing approximate thresholds for QTL detection (Piepho, 2001).

In order to combine data from the 2 populations, a joint-population analysis was performed using the joint inclusive composite interval mapping (JICIM) method (Li et al., 2011). QTL mapping for the nested association mapping (NAM) design in QTL IciMapping, version 3.2, was used to detect joint-population QTLs for grain protein content under each environment and in combination across the 4 environments. The threshold LOD score, determining significant additive QTLs, was specified at 2.5. The additive effects, positions, LOD values, and phenotypic variance explained (R^2) were obtained for each detected QTL.

Positive and negative additive effects indicated that the allele from the high oil

maize parent, GY220, and from the normal dent maize inbreds, 8984/8622, each increased the value of the trait. Single-trait QTLs were named according to "q" + "environment abbreviation" + "trait abbreviation (grain protein content, PRO)" + "population number" + "-" + "chromosome number" + "-" + "QTL number". For joint-trait QTLs, the "trait abbreviation (PRO)" was replaced by "two-trait abbreviation (protein and oil, PO)". For example, the first word "qc", "qx", "qz", "qy", and "q" signifies that the QTL was detected at Xuchang, Xinxiang, Zhengzhou, Yinchuan, and in the combined analysis, respectively. For joint-population QTLs, the "population number" term was omitted. Confidence intervals were calculated by subtracting 1 LOD unit on each side of the maximum LOD position (Lynch and Walsh, 1998).

RESULTS

Combined variance, heritability, and performance for grain protein content in the 2 connected RIL populations

The results of the combined analysis of variance showed that the genetic (σ_G^2), environmental (σ_E^2), and genotype x environment interaction (σ_{GE}^2) variances were all significant or highly significant for grain protein content in both populations, except for σ_E^2 in Pop.1 (Table 1). The broad sense heritability (h_B^2) estimate for grain protein content was slightly higher in Pop.1 (0.73) than in Pop.2 (0.67).

Table 1. Combined variance analysis for grain protein content and their heritabilities (h_B^2) in the two connected RIL populations.

Population	Variance component			Heritability	
	σ_G^2	σ_E^2	σ_{GE}^2	h_B^2	C.I. on h_B^2
Pop.1	4.36**	11.14	1.17*	0.73	0.66-0.79
Pop.2	4.00**	26.70*	1.32**	0.67	0.58-0.74

*Significant at P < 0.05, **Significant at P < 0.01.

The values of grain protein content for the 2 pairs of parents and the 2 connected RIL populations were variable across the 4 environments (Table 2). The value for grain protein content was higher in the high-oil maize inbred GY220 parental line than in the 2 normal maize inbreds, 8984 and 8622. There were minimal differences between the 2 normal inbred lines under all environments, although grain protein content was slightly higher in 8622 than in 8984. With respect to the connected RIL populations, grain protein content was higher in Pop.2 than in Pop.1 in all environments. Overall, the grain protein content data was normally distributed with a wide range of variation and transgressive segregations exceeding both parental values. The coefficients of variation for grain protein content were 6.69-9.17% in Pop.1 and 6.73-9.05% in Pop.2.

Highly significant positive phenotypic and genotypic correlations were observed between grain protein and oil contents in each of the 4 environments and in the combined analysis in both RIL populations. The correlation coefficients were 0.21-0.27 (P < 0.01) and 0.22-0.29 (P < 0.01) in Pop.1, and 0.21-0.28 (P < 0.01) and 0.23-0.30 (P < 0.01) in Pop.2, respectively (data not shown).

Table 2. Phenotypic performance of grain protein content for three parents and two RIL populations under each environment.

Population	Environment	Parent		RIL population				
		GY220	8984/8622[a]	Range	Means ± SD	CV%	Skewness	Kurtosis
Pop.1	Xuchang	14.34	12.29	9.42-16.31	13.26 ± 1.08	8.14	0.07	0.11
	Xinxiang	13.36	11.93	9.90-15.92	12.85 ± 0.97	7.55	0.33	0.34
	Zhengzhou	13.09	12.51	9.73-15.83	12.69 ± 1.02	8.04	0.26	0.21
	Yinchuan	13.78	11.28	9.82-15.34	12.45 ± 0.91	7.31	0.18	0.35
	Combined	13.64	12.00	9.72-15.79	12.82 ± 0.8	6.24	0.22	0.86
Pop.2	Xuchang	14.34	12.90	10.46-16.19	13.14 ± 0.93	7.08	0.26	0.26
	Xinxiang	13.36	12.91	9.61-14.81	12.67 ± 0.91	7.18	-0.04	0.04
	Zhengzhou	13.09	10.75	10.54-7.03	12.75 ± 0.97	7.61	0.18	0.67
	Yinchuan	13.78	12.07	9.52-15.90	12.89 ± 0.99	7.68	-0.08	0.59
	Combined	13.64	12.16	10.08-15.41	12.87 ± 0.72	5.61	-0.03	0.29

[a] 8984 was the parent in Pop.1, and 8622 was the parent in Pop.2.

QTL analysis for grain protein content in the 2 RIL populations

Since the analysis of variance indicated that the σ_{GE}^2 was significant for grain protein content in both populations, QTL mapping for grain protein content was conducted for each environment separately. For comparison, a combined analysis using means across the 4 environments was also conducted.

Sixteen QTLs were detected for protein content in the 4 environments and in the combined analysis in both RIL populations (Table 3). These were located on chromosomes 3 (4 QTLs), 5 (3 QTLs), 6 (1 QTL), 7 (1 QTL), 8 (6 QTLs), and 9 (1 QTL), with contributions to phenotypic variation for a single QTL ranging from 4.4 to 13.4%. Each QTL was detected in 1-2 cases in each population. The QTLs on chromosomes 3, 5, and 8 were all detected in both populations. However, the QTLs on chromosomes 3 and 5 were located at different bin loci in the 2 populations. The QTLs on chromosome 8 were located at the same or adjacent bin loci in 6 cases. Except for 2 QTLs at bins 3.05 and 7.01 in Pop.1, the positive alleles of the other 14 QTLs were contributed by the high oil maize parent GY220 in both populations.

In Pop.1, 10 QTLs were located on chromosomes 3 (1 QTL), 5 (2 QTLs), 6 (1 QTL), 7 (1 QTL), and 8 (5 QTLs), with contributions to phenotypic variation for a single QTL ranging from 4.4 to 13.4%. One, 2, 3, 3, and 1 QTLs were detected in the case of Xuchang, Xinxiang, Zhengzhou, Yinchuan, and in the combined analysis, respectively. Six QTLs detected in Pop.2 were located on chromosomes 3 (3 QTLs), 5 (1 QTL), 8 (1 QTL), and 9 (1 QTL). The contribution to phenotypic variation for an individual QTL varied between 5.9 and 9.7%. One, 1, 3, 0, and 1 QTLs were detected in the case of Xuchang, Xinxiang, Zhengzhou, Yinchuan, and in the combined analysis, respectively.

The intervals of QTL positions were 2.4-21.2 cM, with an average of 9.9 cM. The number of QTLs with intervals <5, 5-10, 10-15, and >15 cM were 4, 5, 4, and 3, accounting for 25.0, 31.25, 25.0, and 18.75%, respectively, of the total number of QTLs. Four QTLs (qxPRO1-8-1, qzPRO1-3-1, qyPRO1-7-1, and qyPRO1-8-1) at bins 8.04-8.05, 3.05, 7.01, and 8.03 with intervals of 3.9, 2.4, 2.5, and 4.7 cM, respectively, could be considered to be fine-mapped.

Table 3. QTL detected for grain protein content and joint QTL analysis with grain oil content in the two RIL populations under each environment and in combined analysis.

Population	Environment	QTL	Flanking marker	Interval (cM)[a]	Bin locus[b]	Position[c]	LOD	A[d]	R²%[e]
Pop.1	Xuchang	qcPO1-5-1	umc2026-umc1019	19.8 (16.0/3.8)	5.05-5.06	127.3	6.06		
		qcPO1-5-2	umc2305-bnlg1306	30 (14.0/16.0)	5.06-5.07	149.5	5.89		
		qcPO1-8-1	umc2075-bnlg2046	12.4 (4.0/8.4)	8.03-8.04	102.0	4.61		
		qcPRO1-8-1	umc2075-bnlg2046	12.4 (8.0/4.4)	8.03-8.04	106.0	4.24	0.33	7.8
	Xinxiang	qxPRO1-6-1	umc1462-umc1424	21.2 (6.0/15.2)	6.05-6.06	186.3	3.20	0.23	6.4
		qxPRO1-8-1	bnlg2046-umc1562	3.9 (2.0/1.9)	8.04-8.05	112.4	4.29	0.26	6.1
		qxPO1-8-1	bnlg2046-umc1562	3.9 (2.0/1.9)	8.04-8.05	112.4	7.48		
	Zhengzhou	qzPRO1-3-1	phi053-umc1174	2.4 (2.0/0.4)	3.05	189.5	3.28	-0.35	4.8
		qzPRO1-5-1	bnlg1879-umc1162	17.8 (10.0/7.8)	5.02-5.04	67.5	5.09	0.37	11.3
		qzPO1-5-1	bnlg1879-umc1162	17.8 (10.0/7.8)	5.02-5.04	67.5	5.04		
		qzPRO1-8-1	umc1149-umc1960	8.1 (8.0/0.1)	8.06	161.8	3.42	0.24	6.2
	Yinchuan	qyPO1-3-1	phi243966-umc1012	3.1 (0/3.1)	3.04	144.2	3.69		
		qyPRO1-5-1	bnlg2305-bnlg118	14.4 (0/14.4)	5.07	168.6	2.59	0.23	4.9
		qyPRO1-7-1	umc1066-umc1632	2.5 (0/2.5)	7.01	60.3	2.61	-0.19	4.4
		qyPO1-8-1	umc1360-bnlg1863	4.7 (2.0/2.7)	8.03	92.3	3.64		
		qyPRO1-8-1	umc1360-bnlg1863	4.7 (4.0/0.7)	8.03	94.3	4.04	0.24	6.8
	Combined	qPO1-1-1	umc1044-phi109275	14.9 (0/14.9)	1.03	51.1	4.75		
		qPO1-3-1	phi243966-umc1012	3.1 (2.0/1.1)	3.04	146.2	4.15		
		qPO1-4-1	umc1548-umc1329	18.7 (12.0/6.7)	4.05-4.06	85.7	4.78		
		qPO1-5-1	umc2026-umc1019	19.8 (14.0/5.8)	5.05-5.06	125.3	5.52		
		qPO1-5-2	umc2305-bnlg1306	30.0 (10.0/20.0)	5.06-5.07	145.5	5.46		
		qPO1-8-1	umc2075-bnlg2046	12.4 (8.0/4.4)	8.03-8.04	106.0	5.98		
		qPRO1-8-1	umc2075-bnlg2046	12.4 (10.0/2.4)	8.03-8.04	108.0	7.44	0.29	13.4
Pop.2	Xuchang	qcPRO2-3-1	umc2259-bnlg1447	13.7 (12.0/1.7)	3.02-3.03	69.2	3.11	0.29	6.5
	Xinxiang	qxPO2-6-1	umc2316-umc1979	8.3 (4.0/4.3)	6.03-6.04	70.8	7.44		
		qxPRO2-9-1	umc1657-umc1494	5.7 (4.0/1.7)	9.05	125.8	3.76	0.28	7.0
	Zhengzhou	qzPRO2-3-1	bnlg1325-bnlg1523	8.0 (0/8.0)	3.03	38.5	3.46	0.28	7.6
		qzPO2-3-1	bnlg1325-bnlg1523	8.0 (0/8.0)	3.03	38.5	4.09		
		qzPRO2-5-1	umc2305-phi085	17.2 (17.2/0)	5.06	164.8	3.19	0.26	5.9
		qzPO2-6-1	umc2316-umc1979	8.3 (4.0/4.3)	6.03-6.04	70.8	5.03		
		qzPRO2-8-1	phi115-phi100175	6.2 (2.0/4.2)	8.03	67.6	3.77	0.2	6.6
		qzPO2-8-1	phi115-phi100175	6.2 (6.0/0.2)	8.03	71.6	5.27		
	Combined	qPO2-3-1	bnlg1325-bnlg1523	8.0 (0/8.0)	3.03	38.5	5.09		
		qPRO2-3-1	bnlg1325-bnlg1523	8.0 (2.0/6.0)	3.03	40.5	4.15	0.26	9.7
		qPO2-6-1	umc2316-umc1979	8.3 (4.0/4.3)	6.03-6.04	70.8	8.11		
		qPO2-6-2	phi031-bnlg1617	12.8 (4.0/8.8)	6.04-6.05	91.5	5.85		
		qPO2-8-1	phi100175-bnlg1863	2.9 (2.0/0.9)	8.03	73.8	4.51		

[a]Values in the brackets were the intervals between QTL and its flanking markers. [b]Bin locations of the flanking markers from Maize GDB (http://www.maizegdb.org). [c]Genetic map position by cM. [d]A means additive effects estimated with QTL Cartographer. [e]R^2 means percent of phenotypic variance explained by each QTL.

Digenic epistasis among QTLs detected for grain protein content in 2 connected RIL populations

In the 2 RIL populations, 9 pairs of digenic interactions were identified among the detected QTLs for grain protein content in each environment and in the combined analysis (Table 4), including 6 QTLs x genetic background (marker intervals with no detected QTLs) and 3 genetic background x genetic background interactions. These interactions were linked to 7 chromosomes with the detected QTL distributions. In addition, 6 marker intervals on chromosomes 1, 2, 5, 8, and 9 were also included. Bin 3.03 appeared to be the most important chromosome region with epistasis, as it was related to 5 pairs of interactions. The values of the interaction effects were all very low, ranging from 0.4 from 4.8%. These results suggested that the overall contributions of digenic interactions to the performance related to grain protein content were minimal in both populations.

Table 4. Digenic epistatic interactions among detected QTL for grain protein content under each environment and in combined analysis for the two RIL populations.

Population	Environment	QTL/Marker interval 1	QTL/Marker interval 2	LOD	Effect	R² %
Pop.1	Xuchang	nc130-umc1679 (5.00-5.01)	qcPRO1-8-1 (8.03-8.04)	0.25	0.09	0.7
	Xinxiang	bnlg1879-umc1162 (5.02-5.04)	qxPRO1-6-1 (6.05-6.06)	1.36	0.19	4.8
	Zhengzhou	bnlg1325-umc1814 (3.03-3.02)	qzPRO1-8-1 (8.06)	0.43	-0.09	0.4
	Yinchuan	umc1448-umc2032 (2.04)	qxPRO1-7-1 (7.01)	0.41	-0.09	1.2
	Combined	umc1044-phi109275 (1.03)	umc2006-phi389203 (6.04-6.03)	0.27	0.05	0.5
Pop.2	Xuchang	umc1746-bnlg1325 (3.01-3.03)	phi080-phi015 (8.08)	0.64	0.17	2.0
	Xinxiang	bnlg1940-bnlg1329 (2.08)	bnlg1325-bnlg1523 (3.03)	1.30	0.21	3.9
	Zhengzhou	mmc0271-bnlg2144 (2.07-2.08)	qzPRO2-3-1 (3.03)	1.03	-0.17	2.5
	Combined	qPRO2-3-1 (3.03)	bnlg244-umc1033 (9.02)	0.52	0.10	1.6

Joint-population QTL analysis for grain protein content across the 2 connected RIL populations

Joint-population analysis across the 2 connected RIL populations by JICIM revealed 19 joint-population QTLs for grain protein content under each environment and in the combined analysis: 3 at Xuchang, 5 at Xinxiang, 5 at Zhengzhou, 2 at Yinchuan, and 4 in the combined analysis (Table 5). The overall contribution to phenotypic variation for an individual QTL varied between 8.7 and 29.0%: 0.9-47.5% in Pop.1 and 0.1-25.9% in Pop.2. The interval of the QTL positions was 0.8-36.4 cM, with an average of 9.4 cM. The number of QTLs with intervals <5 cM, 5-10 cM, 10-15 cM, and >15 cM were 11, 2, 3, and 3, accounting for 57.9, 11.5, 15.8, and 15.8% of the total, respectively.

Table 5. QTL for protein content detected by joint analysis for the two population using JICIM.

Environment	QTL	Position[a]	LOD	R²(%)[b]	Left marker (Bin locus)[c]	Right marker (Bin locus)	Interval (cM)[d]	A[e] in each population		R²(%) in each population	
								Pop.1	Pop.2	Pop.1	Pop.2
Xuchang	qcPRO3-1	70.0	2.75	15.4	umc2259 (3.03)	bnlg1447 (3.03)	13.7 (12.7/1.0)	-0.08	0.51*	0.9	17.3
	qcPRO8-1	87.5	3.64	16.0	umc2075 (8.03)	bnlg2046 (8.04)	3.6 (0.7/2.9)	-0.58*	-0.10	47.5	0.6
	qcPRO9-1	37.5	2.72	21.6	bnlg2122 (9.01)	bnlg244 (9.02)	36.4 (32.5/3.9)	0.38	-0.41	20.0	11.2
Xinxiang	qxPRO5-1	75.0	3.60	13.6	phi008 (5.03)	umc1162 (5.04)	13.9 (0.3/13.6)	-0.45	-0.26	27.9	6.2
	qxPRO6-1	130.0	3.31	10.9	nc012 (6.05)	umc1462 (6.05)	5.0 (1.6/3.4)	-0.41*	-0.19	23.5	3.2
	qxPRO8-1	25.0	2.5	11.1	umc1817 (8.02)	umc1304 (8.03)	3.6 (1.6/2.0)	-0.38	0.04	19.9	0.2
	qxPRO8-2	107.5	6.17	25.0	bnlg162 (8.05)	umc1562 (8.05)	3.5 (2.3/1.2)	-0.58*	0.03	47.3	0.1
	qxPRO9-1	127.5	3.67	12.9	umc1231 (9.05)	umc1494 (9.05)	3.2 (3.2/0)	-0.23	-0.44*	7.3	17.4
Zhengzhou	qzPRO3-1	40.0	4.70	29.0	bnlg1325 (3.03)	phi193225 (3.02)	3.0 (1.4/1.6)	0.23	-0.58*	7.3	25.9
	qzPRO3-2	135.0	2.92	8.7	phi053 (3.05)	umc1174 (3.05)	2.3 (1.2/1.1)	0.34*	0.31	16.6	7.4
	qzPRO5-1	100.0	4.76	18.3	umc2407 (5.04)	umc1389 (5.03)	0.8 (0.3/0.5)	-0.58*	-0.23	46.9	4.3
	qzPRO8-1	132.5	3.03	13.6	mmc0181 (8.06)	umc1960 (8.06)	8.1 (7.3/0.8)	-0.46*	0.04	29.1	0.2
	qzPRO9-1	92.5	2.70	9.3	umc1267 (9.03)	umc1688 (9.03)	4.2 (1.2/3.0)	-0.40	-0.22	22.8	3.9
Yinchuan	qyPRO1-1	152.5	2.70	12.4	umc1968 (3.04)	umc1184 (3.05)	17.2 (13.8/3.4)	-0.20	-0.46	5.4	16.0
	qyPRO8-1	85.0	2.77	9.4	umc1360 (8.03)	umc2075 (8.03)	2.2 (0.4/1.8)	-0.40*	-0.14	22.4	1.4
Combined	qaPRO3-1	40.0	2.63	17.5	bnlg1325 (3.03)	phi193225 (3.02)	3.0 (1.4/1.6)	0.17	-0.32*	4.1	13.5
	qaPRO5-1	80.0	2.76	11.5	phi008 (5.03)	umc1162 (5.04)	13.9 (5.3/8.6)	-0.33	-0.22	15.5	6.4
	qaPRO8-1	87.5	6.42	26.0	umc2075 (8.03)	bnlg2046 (8.04)	3.6 (0.7/2.9)	-0.53*	-0.20	39.1	5.3
	qaPRO9-1	40.0	3.2	23.4	bnlg2122 (9.01)	bnlg244 (9.02)	36.4 (35.0/1.4)	0.27	-0.31	10.4	12.5

[a]Genetic map position by cM. [b]R² means percent of phenotypic variance explained by each QTL. [c]Values in the brackets were the bin locations of the flanking markers from Maize GDB (http://www.maizegdb.org). [d]Values in the brackets were the intervals between QTL and its flanking markers. [e]A means additive effects estimated with JICIM.

In comparison, 12 of the 19 QTLs were detected in the individual population analysis at the same marker intervals or related with the same markers (Table 3). All the additive effects of these QTLs were significant in only 1 population, which could be considered as rare QTLs. This result demonstrates the great influence of population differences on QTL mapping. Although the QTLs at bins 5.06, 5.07, and 7.01 in the single-population analysis could not be detected, 7 QTLs, at bins 1.09, 5.03-5.04, 8.02, 9.01-9.02, and 9.03, detected in the joint-population analysis could be considered as additional QTLs. Furthermore, in the joint analysis, the contributions to phenotypic variance for most QTLs were much higher than those from individual population mapping, while their interval distances were much smaller.

Multiple-trait QTL analysis for grain protein and oil contents in the 2 connected RIL populations

To further analyze the genetic correlations between grain protein and oil content, multiple-trait analyses for protein and oil contents were conducted in both populations (Table 3). In total, 21 protein-oil QTLs were detected in the 4 environments and in the combined analysis in both RIL populations: 13 in Pop.1 and 8 in Pop.2. Compared with the QTLs for protein and oil contents detected by single-trait mapping, 7 QTLs in the joint analysis were located at different marker intervals. Three QTLs were only detected in the joint analysis, 2 of which were detected at bin 3.04 (phi243966-umc1012) in Pop.1 and the other was detected at bin 3.05 (unc2127-umc2166) in Pop.2. This result showed that the joint-trait analysis method had higher statistical power than did the single-trait analysis in QTL detection.

The graph of the LOD curve peaks of protein and oil contents changed simultaneously and in the same direction in 10 cases in Pop.1 (at marker intervals bnlg1879-umc1162 on chromosome 5 at Zhengzhou; umc2026-umc1019 and umc2305-bnlg1306 on chromosome 5 at Xuchang and in the combined analysis; bnlg1863-umc2075 on chromosome 8 at Yinchuan; umc2075-bnlg2046 at Xuchang, Xinxiang, and in the combined analysis; and bnlg2046-umc1562 at Xinxiang), and in 7 cases in Pop.2 (at marker intervals umc2127-umc2166 on chromosome 3 at Zhengzhou; umc2316-umc1979 on chromosome 6 at Xinxiang, Zhengzhou, and in the combined analysis; phi031-bnlg1617 on chromosome 6 in the combined analysis; phi115-phi100175 on chromosome 8 at Zhengzhou; and phi100175-bnlg1863 on chromosome 8 in the combined analysis) (Figure 1). These data suggested that pleiotropic QTLs were controlling protein and oil contents simultaneously. The peaks of the LOD curve for protein and oil contents changed in the same direction in 3 cases in Pop.1 (at marker intervals phi243966-umc1012 on chromosome 3 at Yinchuan and in the combined analysis and bnlg1879-umc1162 on chromosome 5 at Zhengzhou), and in 1 case in Pop.2 (at marker interval bnlg1325-bnlg1523 on chromosome 3 in the combined analysis), suggesting that the QTLs controlled protein and oil contents with a tight linkage in these marker intervals. However, 4 QTLs for protein content and 3 QTLs for oil content in Pop.1 and 3 QTLs for protein content and 6 QTLs for oil content in Pop.2 failed to show significant effects in the joint-trait analysis (Table 3). This suggests that these QTLs might affect protein and oil contents in opposite directions.
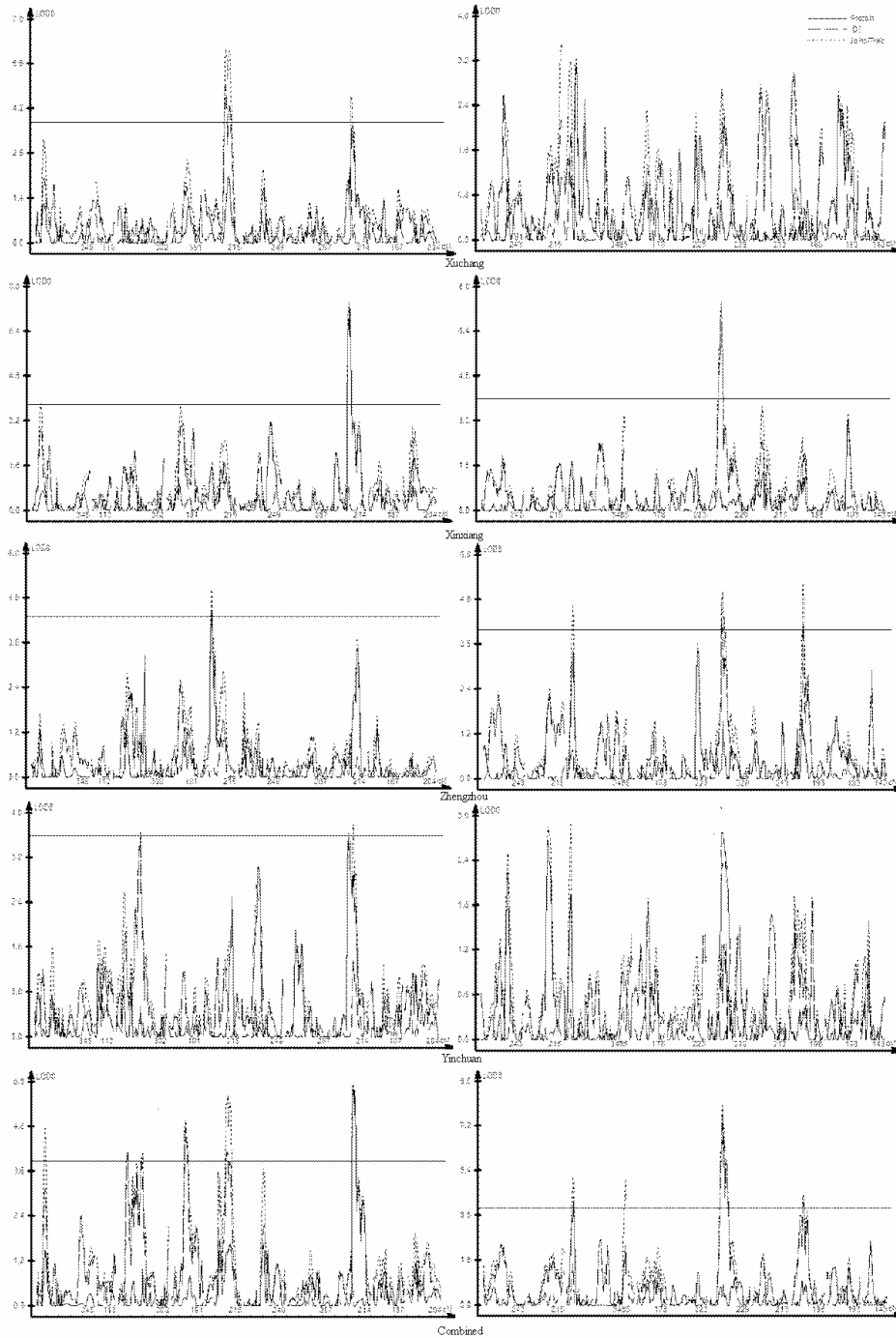
Figure 1. Joint QTL analysis of protein content with oil content in two RIL populations. Figures on the left line came from Pop.1 and the figures on the right came from Pop.2.

DISCUSSION

Comparison of QTLs detected for grain protein content between the 2 connected RIL populations and with previous studies

In this study, 16 single-population QTLs were detected for grain protein content in 2 connected RIL populations under 4 environments and in the combined analysis: 10 in Pop.1 and 6 in Pop.2. No QTL showed consistency across both RIL populations. Although all the QTLs were detected on chromosomes 3, 5, and 8, they were located at different marker intervals in the 2 populations. In the joint-population analysis, 12 of the 19 QTLs had significant genetic effects in only 1 population. This result demonstrated that all the detected QTLs showed population-specific effects and were greatly affected by different genetic backgrounds. Across the 4 environments and in the combined analysis, only 2 QTLs were detected under 2 environments simultaneously, which showed that most QTLs were greatly influenced by environmental differences. The same tendency was observed in our previous study using the 2 $F_{2:3}$ populations developed from the same crosses used in this study (Li et al., 2009).

Compared with the QTLs previously detected in the 2 $F_{2:3}$ populations (Li et al., 2009), 7 were detected in both generations and showed generational consistency, which were located at marker intervals bnlg1879-umc1162, umc2305-phi085, bnlg2305-bnlg118, umc1360-bnlg1863, umc2075-bnlg2046, bnlg2046-umc1562, and umc1657-umc1494. However, 7 QTLs, located at marker intervals umc2259-bnlg1447, bnlg1325-bnlg1523, phi053-umc1174, umc1462-umc1424, umc1066-umc1632, phi115-phi100175, and umc1149-umc1960, were only detected in the RIL populations, and 6 QTLs, located at marker intervals umc1320-bnlg1754, umc1653-umc1127, bnlg1067-bnlg2082, umc1231-umc1771, umc1657-umc1494, and phi96342-umc1938, were only detected in the $F_{2:3}$ populations. Furthermore, these QTLs were generation-specific (Table 6). Strong influences of generation and population are commonly reported in QTLs detected for grain and yield components (Stuber et al., 1992; Austin and Lee, 1996a; Austin et al., 2000; Li et al., 2009), flowering, and plant height traits (Austin and Lee, 1996b).

Table 6. Comparison of QTL for grain protein content detected in the two RIL populations and in previous studies.

Bin locus	Environment[a]	Reference	Flanking marker
3.02-3.03	c	Goldman et al., 1993; Wang, 2007	umc2259-bnlg1447
3.03	z, a		bnlg1325-bnlg1523
3.05	Z	Goldman et al., 1993	phi053-umc1174
5.02-5.04	Z	Goldman et al., 1993; Wang, 2007; Wassom et al., 2008	bnlg1879-umc1162
5.06	z	Goldman et al., 1993	umc2305-phi085
5.07	Y		bnlg2305-bnlg118
6.05-6.06	X	Goldman et al., 1993; Melchinger et al., 1998; Willmot et al., 2006; Zhang et al., 2008; Wang, 2007; Liu et al., 2008; Wassom et al., 2008; Li et al., 2009	umc1462-umc1424
7.01	Y	Liu et al., 2008	umc1066-umc1632
8.03	z	Berke and Rocheford, 1995; Melchinger et al., 1998; Dudley et al., 2004; Willmot et al., 2006; Zhang et al., 2008; Liu et al., 2008; Wassom et al., 2008; Li et al., 2009	phi115-phi100175
8.03	Y		umc1360-bnlg1863
8.03-8.04	C, A		umc2075-bnlg2046
8.04-8.05	X		bnlg2046-umc1562
8.06	Z		umc1149-umc1960
9.05	x	Goldman et al., 1993; Zhang et al., 2008; Wang, 2007	umc1657-umc1494

[a]C, X, Z, Y, and A present QTL detected at Xuchang, Xinxiang, Zhengzhou, Yinchuan and in combined analysis in Pop.1, respectively; c, x, z, y and a present QTL detected at Xuchang, Xinxiang, Zhengzhou, Yinchuan, and in combined analysis in Pop.2, respectively.

Previous studies on QTL detection for grain protein content have used F_3 populations derived from the IHP x ILP cross (Goldman et al., l993); S_1 families derived from the IHO x ILO cross (Berke and Rocheford, 1995); IHO x B73 backcross-derived lines (Wassom et al., 2008); $F_{2:3}$, $F_{2:4}$, and $F_{7:8}$ populations from normal corn and BHO inbreds (Zhang et al., 2008); $F_{2:3}$ derived from ASK high-oil (Li et al., 2009); and $F_{2:3}$ and BC_2F_2 populations derived from normal corn x popcorn inbreds (Liu et al., 2008). In addition, qzPRO1-5-1, which was detected at bin 5.02-5.04 with a 11.3% contribution to phenotypic variation in the present study, was also identified by Goldman et al. (1993), Li et al. (2009), and Wassom et al. (2008); qx-PRO1-6-1, which was detected at bin 6.05-6.06 with a 6.4% contribution to phenotypic variation, has also been identified by Goldman et al. (1993), Melchinger et al. (1998), Willmot et al. (2006), Zhang et al. (2008), Li et al. (2009), Liu et al. (2008), Wassom et al. (2008), and Li et al. (2009); and qPRO1-8-1, which was detected at bin 8.03-8.04 with a 13.4% contribution to phenotypic variation in the present study, has also been identified by Berke et al. (1995), Melchinger et al. (1998), Dudley et al. (2004), Willmot et al. (2006), Zhang et al. (2008), Liu et al. (2008), Wassom et al. (2008), and Li et al. (2009). The consistency of detection of these QTLs across studies suggests that they are less influenced by their environmental and genetic backgrounds. According to the public genetic linkage map (www.maizegdb.org), several functional genes were located at these bin loci, such as proline responding 1 (*pro1*) at bin 8.04, sucrose phosphate synthase 1 (*sps1*) at bin 8.06, opaque 2 heterodimerizing protein 2 (*ohp2*) at bin 5.01, brittle endosperm 1 (*bt1*) at bin 5.04, and glutamate-oxaloacetic transaminase 3 (*got3*) and beta amylase 2 (*amyb2*) at bin 5.03. Considering the simultaneous contribution of individual QTLs to the overall phenotypic variation, qPRO1-8-1 and qPRO1-5-1 detected herein and their chromosome intervals at bins 8.03-8.04 and 5.02-5.04 could be used as targets for the development of near isogenic lines (NILs) and chromosome segment substitution lines by using marker-assisted selection in future studies.

Comparison of QTLs detected for grain protein and oil contents and their correlations

Approximately 75-80% of the grain protein content is stored in the endosperm and 85% of oil concentration is located in the embryo (Orthoefer and Sinram, 1987; Watson, 1987). A positive relationship between grain protein and oil content has been reported in studies using Illinois maize strains (Goldman et al., 1993; Lambert, 2001; Wassom et al., 2008), BHO (Zhang et al., 2008), ASK high-oil (Li et al., 2009; Wang et al., 2010), and popcorn germplasms (Liu et al., 2008). Curtis et al. (1968) showed that the total germ weight and percentage increased with selection of oil and fatty acid composition in IHO after 65 generations, and the protein content also increased. This finding is likely a result of the association between germ weight and percentage with the protein and oil content in the kernel.

In this study, significantly positive phenotypic and genotypic correlations were found between grain protein and oil contents in the 2 RIL populations under each environment and in the combined analysis. Comparing the single-trait QTLs detected for grain protein and oil contents, 1 QTL in the same marker interval (phi115-phi100175) on chromosome 8 in Pop.2 and 7 QTLs at the same or adjacent bin loci (such as bins 3.03-3.05, 5.04-5.06, 6.03-6.05, 8.03-8.04, and 8.04-8.06) were detected for both traits. All these QTLs had the same parental effects, in which the favorable alleles were contributed by the same high-oil parent, GY220. In the joint-trait analysis, the protein-oil QTLs included most of the single-trait QTLs with

the same parental effects for grain protein and oil contents. Both results of the single-trait and joint-trait analyses suggested the presence of tightly linked and/or pleiotropic QTLs with the same effects controlling both traits in these bin loci. Similar results have been reported by Li et al. (2009), Zhang et al. (2008), and Wassom et al. (2008). Therefore, since increases in grain oil content might be associated with increases in grain protein content, both traits could be improved simultaneously.

ACKNOWLEDGMENTS

REFERENCES

Alrefai R, Berke TG and Rocheford TR (1995). Quantitative trait locus analysis of fatty acid concentrations in maize. *Genome* 38: 894-901.

Arcade A, Labourdette A, Falque M, Mangin B, et al. (2004). BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* 20: 2324-2326.

Austin DF and Lee M (1996a). Comparative mapping in $F_{2:3}$ and $F_{6:7}$ generations of quantitative trait locus for grain yield and yield composition traits in maize. *Theor. Appl. Genet.* 92: 817-826.

Austin DF and Lee M (1996b). Genetic resolution and verification of quantitative trait loci for flowering and plant height with recombinant inbred lines of maize. *Genome* 39: 957-968.

Austin DF, Lee M, Veldboom LR and Hallauer AR (2000). Genetic mapping in maize with hybrid progeny across testers and generations: grain yield and grain moisture. *Crop Sci.* 40: 30-39.

Berke TG and Rocheford TR (1995). Quantitative trait locus for flowering, plant, and ear height, and kernel traits in maize. *Crop Sci.* 35: 1542-1549.

Churchill GA and Doerge RW (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963-971.

Curtis PE, Leng ER and Hageman RH (1968). Development changes in oil and fatty acid content of maize strains varying in oil content. *Crop Sci.* 8: 689-693.

Dudley JW and Lambert RJ (1992). Ninety generations of selection for oil and protein in maize. *Maydica* 37: 81-87.

Dudley JW and Lambert RJ (2004). 100 generations of selection for oil and protein in corn. *Plant Breed. Rev.* 24: 79-110.

Dudley JW, Lambert RJ and de la Roche IA (1977). Genetic analysis of crosses among corn strains divergently selected for percent oil and protein. *Corp Sci.* 17: 111-117.

Dudley JW, Dijkhuizen A, Paul C, Coates ST, et al. (2004). Effects of random mating on marker-QTL associations in the cross of the Illinois High Protein-Illinois Low Protein Maize strains. *Crop Sci.* 44: 1419-1428.

Goldman IL, Rocheford TR and Dudley JW (1993). Quantitative trait loci influencing protein and starch concentration in the Illinois Long Term selection maize strains. *Theor. Appl. Genet.* 87: 217-224.

Hallauer AR and Miranda JB (1981). Quantitative Genetics in Maize Breeding. Iowa State University Press.

Han Y, Parsons CM and Alexander DE (1987). Nutritive value of high oil corn for poultry. *Poult. Sci.* 66: 103-111.

Henderson CR (1953). Estimation of variance and covariance components. *Biometrics* 9: 226-252.

Hopkins CG (1899). Improvement in the chemical composition of the corn kernel. *Ill. Agr. Exp. Sta. Bull.* 55: 205-240.

Jiang HY, Chen SJ, Gao LF, Xing JM, et al. (2005). Analysis on heterotic groups and heterotic patterns of high oil corn inbred lines. *Acta Agron. Sin.* 31: 361-367.

Jiang C and Zeng ZB (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140: 1111-1127.

Kao CH, Zeng ZB and Teasdale RD (1999). Multiple interval mapping for quantitative trait loci. *Genetics* 152: 1203-1216.

Knapp SJ, Stroup WW and Ross WM (1985). Exact confidence intervals for heritability on progeny mean basis. *Crop Sci.* 25: 192-194.

Lambert RJ (2001). High-Oil Corn Hybrids. In: Specialty Corns (Hallauer AR, ed.). CRC, Boca Raton, 131-154.

Laurie CC, Chasalow SD, LeDeaux JR, McCarroll R, et al. (2004). The genetic architecture of oil concentration in the maize kernel after 70 generations of divergent selection. *Genetics* 168: 2141-2155.

Li HH, Bradbury P, Ersoz E, Buckler ES, et al. (2011). Joint QTL linkage mapping for multiple-cross mating design sharing one common parent. *PloS One* 6: e17573.

Li Y, Wang Y, Wei M, Li X, et al. (2009). QTL identification of grain protein concentration and its genetic correlation with starch concentration and grain weight using two populations in maize (*Zea mays* L.). *J. Genet.* 88: 61-67.

Liu YY, Dong YB, Niu SZ, Cui DQ, et al. (2008). QTL identification of kernel composition traits with popcorn using both $F_{2:3}$ and BC_2F_2 populations developed from the same cross. *J. Cereal Sci.* 48: 625-631.

Lynch M and Walsh B (1998). Genetics and Analysis of Quantitative Traits. Sinauer Associates Inc., Sunderland Massachusetts.

Mangolin CA, de Souza Jr CJ, Garcia AAF, Garcia AF, et al. (2004). Mapping QTLs for kernel oil content in a tropical maize population. *Euphytica* 137: 251-259.

Melchinger AE, Utz HF and Schon CC (1998). Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* 149: 383-403.

Miller PA and Brimhall B (1951). Factors influencing the oil and protein content of corn grain. *Agron. J.* 43: 305-311.

Orthoefer FT and Sinram RD (1987). Corn Oil: Composition, Processing, and Utilization. In: Corn: Chemistry and Technology (Watson SA and Ranstad PE, eds.). American Association of Cereal Chemists, Inc., St. Paul, Minnesota, 535-551.

Piepho HP (2001). A quick method for computing approximate thresholds for quantitative trait loci detection. *Genetics* 157: 425-432.

Song XF, Song TM, Dai JR, Rocheford T, et al. (2004). QTL mapping of kernel oil concentration with high-oil maize by SSR markers. *Maydica* 49: 41-48.

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, et al. (1992). Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* 132: 823-839.

Wang YZ (2007). QTL Mapping for Kernel Nutritional Quality Characters and their Genetic Relationship in Maize. Ph.D thesis. Henan Agricultural University, Zhengzhou.

Wang S, Basten CJ and Zeng ZB (2006). Windows QTL Cartographer 2.5. Department of Statistics. North Carolina State University, Raleigh, NC. Available at [http://statgen.ncsu.edu/qtlcart/WQTLCart.htm]. Accessed March 12, 2011.

Wang YZ, Li JZ, Li YL, Wei MG, et al. (2010). QTL detection for grain oil and starch content and their associations in two connected $F_{2:3}$ populations in high-oil maize. *Euphytica* 174: 239-252.

Wassom JJ, Wong JC, Martinez E, King JJ, et al. (2008). QTL associated with maize kernel oil, protein, and starch concentrations; kernel mass; and grain yield in Illinois High OilxB73 backcross-derived lines. *Crop Sci.* 48: 243-252.

Watson SA(1987). Structure and Composition. In Corn: Chemistry and Technology (eds Watson SA and Ramstad PT), American Association Cereal Chemistry, St. Paul, 53-82.

Willmot DB, Dudley JW, Rocheford TR and Bari A (2006). Effect of random mating on marker-QTL associations for grain quality traits in the cross of Illinois High Oil x Illinois Low Oil. *Maydica* 51: 187-199.

Yang GH, Li YL, Wang QL, Zhou YG, et al. (2012). Detection and integration of quantitative trait loci for grain yield components and oil content in two connected recombinant inbred line populations of high-oil maize. *Mol. Breed.* 29: 313-333.

Zeng ZB (1994). Precision mapping of quantitative trait loci. *Genetics* 136: 1457-1468.

Zhang J, Lu XQ, Song XF, Yan JB, et al. (2008). Mapping quantitative trait locus for oil, starch, and protein concentrations in grain with high-oil maize by SSR markers. *Euphytica* 162: 335-344.

Zheng P, Allen WB, Roesler K, Williams ME, et al. (2008). A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat. Genet.* 40: 367-372.