Theor Appl Genet (2014) 127:85–96
DOI 10.1007/s00122-013-2203-9

ORIGINAL PAPER

Association mapping of six yield-related traits in rapeseed (*Brassica napus* L.)

Dongfang Cai · Yingjie Xiao · Wei Yang · Wei Ye · Bo Wang · Muhammad Younas · Jiangsheng Wu · Kede Liu

Received: 2 April 2013 / Accepted: 22 September 2013 / Published online: 12 October 2013

Abstract Yield is one of the most important traits for rapeseed (*Brassica napus* L.) breeding, but its genetic basis remains largely ambiguous. Association mapping has provided a robust approach to understand the genetic basis of complex agronomic traits in crops. In this study, a panel of 192 inbred lines of *B. napus* from all over the world was genotyped using 451 single-locus microsatellite markers and 740 amplified fragment length polymorphism markers. Six yield-related traits of these inbred lines were investigated in three consecutive years with three replications, and genome-wide association studies were conducted for these six traits. Using the model controlling both population structure and relative kinship (Q + K), a total of 43 associations ($P < 0.001$) were detected using the means of the six yield-related traits across 3 years, with two to fourteen markers associated with individual traits. Among these, 18 markers were repeatedly detected in at least 2 years, and 12 markers were located within or close to QTLs identified in previous studies. Six markers commonly associated with correlated traits. Conditional association analysis indicated that five of the associations between markers and correlated traits are caused by one QTL with pleiotropic effects, and the remaining association is caused by linked but independent QTLs. The combination of favorable alleles of multiple associated markers significantly enhances trait performance, illustrating a great potential of utilization of the associations in rapeseed breeding programs.

Communicated by R. G. F. Visser.

D. Cai and Y. Xiao contribute equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2203-9) contains supplementary material, which is available to authorized users.

D. Cai · Y. Xiao · W. Yang · W. Ye · B. Wang · M. Younas · J. Wu · K. Liu (✉)
National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China
e-mail: kdliu@mail.hzau.edu.cn

Introduction

Rapeseed (*Brassica napus* L.) is a widely cultivated oil crop in the world. Rapeseed oil is not only used as edible oil, but also as industrial materials for lubricants and biodiesel. Yield has long been one of the most important traits in rapeseed breeding. Yield of a rapeseed cultivar is directly determined by three component traits: siliques per plant, seeds per silique and seed weight (Özer et al. 1999). It is also indirectly influenced by many yield-related traits such as plant height and first branch height (Quijada et al. 2006; Shi et al. 2009). Plant height has been proved to be an important yield-related trait in several crops (Salas Fernandez et al. 2009). In rice and wheat, the semi-dwarf varieties significantly increased yield through improving resistance to lodging and increasing harvest index (Pinstrup-Andersen and Hazell 1985). In rapeseed, many yield-related QTLs had been identified to be clustered with yield QTLs (Chen et al. 2007; Shi et al. 2009). Thus, understanding the genetic basis of yield-related traits will enable us to improve yield in *B. napus*.

Yield-related traits are usually controlled by multiple loci. In *B. napus*, QTL analyses were done to dissect the genetic bases of yield-related traits in segregating populations derived from biparental crosses. Hundreds of QTLs and epistatic interactions were detected across the whole genome of *B. napus* in different genetic backgrounds, and many of them had pleiotropic effects on different traits (Basunanda et al. 2010; Radoev et al. 2008; Shi et al. 2011; Udall et al. 2006; Zhang et al. 2011). These studies

demonstrated that genetic mapping of quantitative traits using genetic linkage map is an efficient approach to identify QTL underlying these agronomically important traits.

Association mapping with its high resolution is an alternative to the traditional QTL mapping to study the genetic architecture of complex traits. It is performed in natural populations with diverse genetic backgrounds, which can mitigate the inherent limitations such as time consuming and low resolution when using biparental populations for QTL identification, and can be applied directly to breeding programs (Jannink et al. 2010). In addition, association mapping has an ability to discover multiple alleles at the genetic loci in germplasm, thus enables the mining of elite alleles for breeding (Li et al. 2011b; Jia et al. 2012). Association studies had been successfully carried out in *B. napus*. Based on *Arabidopsis* orthologs of candidate genes for glucosinolate biosynthesis, 51 gene-linked SSR alleles have been detected to be associated with seed glucosinolate content in two sets of *B. napus* germplasm with 94 and 46 genotypes, respectively (Hasan et al. 2008). Genome-wide association analyses (GWAS) detected both previously identified QTLs and novel QTLs for oil content in the resynthesized and traditional *B. napus* populations of 103 and 69 lines, respectively (Zou et al. 2010), and for 14 agronomic traits in 84 European rapeseed genotypes of canola quality (Honsdorf et al. 2010). Raman et al. (2011) identified 150 associations with shatter resistance using 1513 unmapped DrAT, SSRs and two candidate genes in 192 genotypes of *B. napus*, *B. rapa*, *B. juncea* and *B. carinata*. Several additive QTLs for six agronomic traits and one epistatic QTL for flowering time were detected using two different biometrical approaches in multiple segregating populations (Wurschum et al. 2012). Associative transcriptomics identified genomic deletions that underlie two quantitative trait loci for glucosinolate content of seeds in tetraploid crop *B. napus*, which was especially applied to the crops with complex genomes (Harper et al. 2012). So far, the genetic basis of yield-related traits in rapeseed has not been systematically studied in the natural association population.

In this study, we genotyped a panel of 192 lines using 451 SSR and 740 AFLP markers and surveyed six yield-related traits in three consecutive years. The objectives of this study were to (1) analyze the influences of population structure on phenotypes; (2) reveal the genetic basis of yield-related traits in *B. napus*; and (3) evaluate the marker-based allelic effects for application in breeding programs.

Materials and methods

Plant materials

A panel of 192 *B. napus* cultivars and inbred lines from all over the world was used for association mapping in this study. The detail information including their origins was published previously as electronic supplementary materials (Xiao et al. 2012). These lines had been genotyped using 451 single-locus microsatellite markers and evaluated for population structure and linkage disequilibrium. Structure analysis and principal component analysis (PCA) classified the 192 rapeseed lines into two groups, P1 and P2, which are essentially consistent with the geographical origins of the rapeseed lines (Xiao et al. 2012). The relative kinship was weak in the panel, with more than 80 % pairwise relationship estimates between lines less than 0.05. The level of linkage disequilibrium in the rapeseed panel was low, with the distance of LD decay within 0.5–1 cm at the genome level (Xiao et al. 2012).

Field trials and trait measurement

The 192 lines were planted in the experimental field at Huazhong Agricultural University, Wuhan, China, in the winters of 2008, 2009 and 2010, and harvested in the springs of 2009, 2010 and 2011. The field experiments in the three consecutive years were treated as three independent environments. Field trials followed a completely randomized block design with three replicates each year. Each plot contained two rows, with 12 plants in each row and 20 cm between plants within each row and 30 cm between rows. Field management followed the standard agricultural practice. At mature stage, 10, 6 and 6 normally developed plants from the middle of the plots were harvested for trait investigation in the springs of 2009, 2010 and 2011, respectively.

Six yield-related traits, plant height (PH), first branch height (FBH), inflorescence length (IL), silique length (SL), seeds per silique (SPS) and seed weight (SW) were investigated. FBH was measured as the height from the base of the stem to the first primary branch. IL was measured as the length from the uppermost branch to the tip of the main inflorescence. PH was measured as the height from the base of the stem to the tip of the main inflorescence. SL was measured as the mean silique length of ten siliques sampled from the middle of the inflorescence. SPS was measured as the mean seed number of ten siliques sampled from the middle of the main inflorescence. SW was measured as the weight of 1,000 well-developed seeds.

Genotyping with microsatellite and AFLP markers

Genomic DNA was extracted from leaf tissues collected from a single plant of each inbred line. The 192 lines were previously genotyped using 451 single-locus microsatellite markers (Xiao et al. 2012), and were further genotyped using AFLP markers in this study. Genomic DNA of each line was digested with EcoRI and MseI, and the procedure

of AFLP analysis followed the method of Vos et al. (1995). The PCR products were visualized on 6 % polyacrylamide gels. Each polymorphic fragment was scored as '1' or '0' according to the presence or absence of amplified fragment. The polymorphic AFLP loci were named based on the primer combination followed by an Arabic number in ascending order according to the AFLP fragment size. Heterozygous genotypes and rare alleles with minor allele frequency (MAF) <0.05 were treated as missing data to increase the power of association analysis (Hirschhorn and Daly 2005). The markers with more than 15 % missing data were excluded, which resulted in a total of 369 SSR and 740 AFLP markers for further analyses.

Sequencing and chromosome assignment of associated markers

AFLP was performed according to Vos et al. (1995) and separated on 6 % denaturing polyacrylamide gels. The pieces of gel containing the target fragments corresponding to the associated markers were excised from the dried gel and incubated at 4 °C overnight in 50 μl ddH_2O, then incubated at 100 °C for 10 min before PCR amplification. After centrifugation, 6 μl was taken for PCR with the pre-amplification primers (10 μM) in a final volume of 20-μl reaction containing 2 × PCR buffer, 4 mM Mg^{2+}, 0.3 mM dNTP, 0.5 units of Taq DNA polymerase. PCR reaction was performed followed as the selective amplification procedure of AFLP fingerprinting. A 5 μl re-amplified product was separated on a 1.5 % agarose gel to confirm the PCR products. PCR products were directly sequenced with the BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems, Foster City, CA, USA).

All associated loci that have not been assigned to the BnaZNDH linkage maps (Cheng et al. 2009; Li et al. 2011a; Wang et al. 2011; Xu et al. 2010) were searched against the BRAD database (http://brassicadb.org/brad/) and the Bolbase database (http://www.ocri-genomics.org/bolbase/), and assigned to specific chromosomes with an *e* value <1e–10. If multiple positions were aligned, only the position with the best hit was selected for chromosome assignment.

Statistical analysis

Statistical analyses for all traits were carried out using the SAS software (The SAS Institute 1999). The effects of genotype, environment, and genotype by environment interaction (G × E) on phenotypic variation were evaluated using PROC GLM. Broad-sense heritability based on family mean was calculated as $H^2 = \delta_g^2 / \left(\delta_g^2 + \delta_{ge}^2/n + \delta_e^2/nr\right)$, where δ_g^2 is the genetic variance, δ_{ge}^2 is the variance due to G × E interaction, δ_e^2 is the residual error, n is the number of environments and r is the number of replicates within environment. The estimates of δ_g^2, δ_{ge}^2 and δ_e^2 were obtained from an analysis of variance (ANOVA) by considering environment as a random effect. The distributions of the six traits were tested for normality using the Shapiro–Wilk test at a significance level of $P < 0.05$. Phenotypic correlations (r_p) and genetic correlations (r_g) between all traits were calculated using PROC CORR and PROC MIXED, respectively. The effect of population structure on each trait was evaluated using PROC GLM.

Association analysis

The existence of population structure and relative kinship in natural populations always result in a high level of spurious positives in association mapping (Yu et al. 2006). The population structure (Q), principal component analysis (PCA) and relative kinship (K) in the panel of 192 rapeseed lines have been evaluated previously (Xiao et al. 2012), and their effects on associations were evaluated with six statistical models: (1) the GLM model; (2) the Q model; (3) the PCA model; (4) the K model; (5) the Q + K model considering both Q and K; (6) the PCA + K model considering both PCA and K. The GLM, Q and PCA models were performed using the GLM procedure, while the K, Q + K and PCA + K models were performed using the MLM procedure in TASSEL V2.1 (Bradbury et al. 2007). The quantile–quantile plots of estimated $-\log10(P)$ were displayed using the observed P values from marker–trait associations and the expected P values from the assumption that no association exists between markers and trait. The model having the observed P values to be closest to the expected P values was chosen as the optimal model to control the confounding of population structure.

Using the optimal statistical model, association analyses were carried out with the 369 SSR and 740 AFLP markers for all traits using the mean values across the 3 years and within each year. The significance of each marker was tested by setting the false discovery rate (FDR) at 0.2, as described previously in rapeseed (Honsdorf et al. 2010), to reduce the probability of false positives caused by multiple testing (Benjamini and Hochberg 1995). Given the distribution of empirical P values of 1,109 markers, the FDR of 0.2 corresponded to the P value of 0.001, which was employed as threshold of significance of tested markers in the association analysis. For markers associated with two traits, a conditional association analysis was conducted as previously described (Soranzo et al. 2009) to reveal if the associations were caused by a QTL with pleiotropic effects on the two traits or by two linked but independent QTLs. Linkage disequilibrium (LD) between associated markers was calculated to evaluate whether these markers inherited independently (Hasan et al. 2008).

Table 1 Descriptive statistics, broad-sense heritability and percentage of phenotypic variance explained by population structure for six yield-related traits

Traits	Mean ± SD	Range	P_w^a	G	E	G × E	H^2 (%)[b] ± SE	R^2_{pop} (%)[c]
First branch height (cm)	43.71 ± 12.15	10.09–82.49	0.8643	**	**	**	88.69 ± 1.45	6.04
Inflorescence length (cm)	54.96 ± 7.62	24.94–86.33	0.0017	**	**	**	86.62 ± 1.72	1.71
Plant height (cm)	145.59 ± 14.89	99.72–180.08	0.0098	**	**	**	87.56 ± 1.60	14.13
Silique length (cm)	5.62 ± 0.78	3.84–9.39	0.0001	**	**	**	93.47 ± 0.84	7.49
Seeds per silique	20.15 ± 3.26	10.35–27.66	0.009	**	**	**	86.31 ± 1.75	4.85
Seed weight (g)	3.51 ± 0.59	2.20–5.49	0.0043	**	**	**	91.02 ± 1.14	3.51

** Significant at $P < 0.01$ for the effect of genotype (G), environment (E) and genotype by environment interaction (G × E) on phenotypic variance estimated by two-way ANOVA

[a] Stands for *P* values of the Shapiro–Wilk test

[b] Family mean-based broad-sense heritability

[c] Percentage of phenotypic variance explained by population structure based on means of the six traits across 3 years

Results

Phenotypic analysis of six yield-related traits

The 192 lines were planted in three consecutive years from 2009 to 2011 and six yield-related traits were investigated. Extensive phenotypic variations were observed for all these traits in the 192 rapeseed inbred lines (Table 1). First branch height (FBH) had 8.18 folds of variation, ranging from 10.09 to 82.49 cm with an average of 43.71 ± 12.15 cm. Plant height (PH) showed 1.81 folds of changes, ranging from 99.72 to 180.08 cm with an average of 145.59 ± 14.89 cm. Shapiro–Wilk normality test indicated that first branch height was normally distributed ($P = 0.86$), and the other five traits did not fit a normal distribution ($P < 0.01$) (Fig. 1; Table 1).

Analysis of variance indicated that genotype (G), environment (E) and genotype by environment interaction (G × E) have significant effects on all these traits ($P < 0.01$; Table 1). Broad-sense heritability (H^2) was calculated for the six traits (Table 1). All traits had an H^2 higher than 85 %, suggesting that all these traits are stably inherited. Phenotypic and genetic correlations were analyzed among the six traits. Most of these traits showed significant correlations with other traits ($P < 0.01$; Table 2). Highly significant positive correlations were observed between PH and the other five traits, with phenotypic correlation coefficients (r_p) ranging from 0.23 to 0.74 and genetic correlation coefficients (r_g) ranging from 0.24 to 0.86, suggesting that plant height is an important factor influencing the development of other yield-related traits. SPS showed a positive correlation with SL ($r_p = 0.47$ and $r_g = 0.39$), but showed a negative correlation with SW ($r_p = -0.29$ and $r_g = -0.30$), suggesting that longer silique will provide more space to develop more seeds, but more seeds per silique affect the development of seed size (Zhang et al. 2011).

Effects of population structure on phenotypes

The 192 rapeseed inbred lines were previously evaluated for population structure and classified into two separate groups, P1 and P2, which essentially corresponded to their geographic origins of inbred lines (Xiao et al. 2012). ANOVA results indicated that population structure explained 14.13, 6.04, 7.49, 4.85 and 3.51 % phenotypic variation of PH, FBH, SL, SPS and SW, respectively, but had basically no effect on IL (Table 1). Phenotypic variations of P1 and P2 were compared to further assess the effect of population structure on trait performance (Fig. 2). P1 contained more lines with higher phenotypic values than P2 for all traits, and P1 also had significantly higher mean values than P2 for all traits ($P < 0.01$).

Model comparison on controlling false associations

Association analyses for the six yield-related traits were performed to evaluate the effects of population structure (Q), principal component analysis (PCA) and familial relationship (K) for controlling false associations. For all six traits, the observed *P* values from the GLM model greatly deviated from the expected *P* values assuming that no association exists, followed by the Q and PCA models. The *P* values from the K, PCA + K and Q + K models were similar and close to the expected *P* values (Fig. 3). Whereas, the effects of the Q and PCA models on controlling false associations were similar, and varied from trait to trait (Fig. 3). The Q and PCA models which reduced the number of significant markers for PH by 14.34 and 14.25 % from the GLM model, respectively. But only reduced the number of markers by 0.99 and 0.81 % for IL, respectively ($P < 0.01$; Table S1). Using the Q + K model, only 1.17, 1.26, 1.44, 1.44, 1.26 and 1.80 % markers were significantly associated with FBH, IL, PH, SL, SPS and

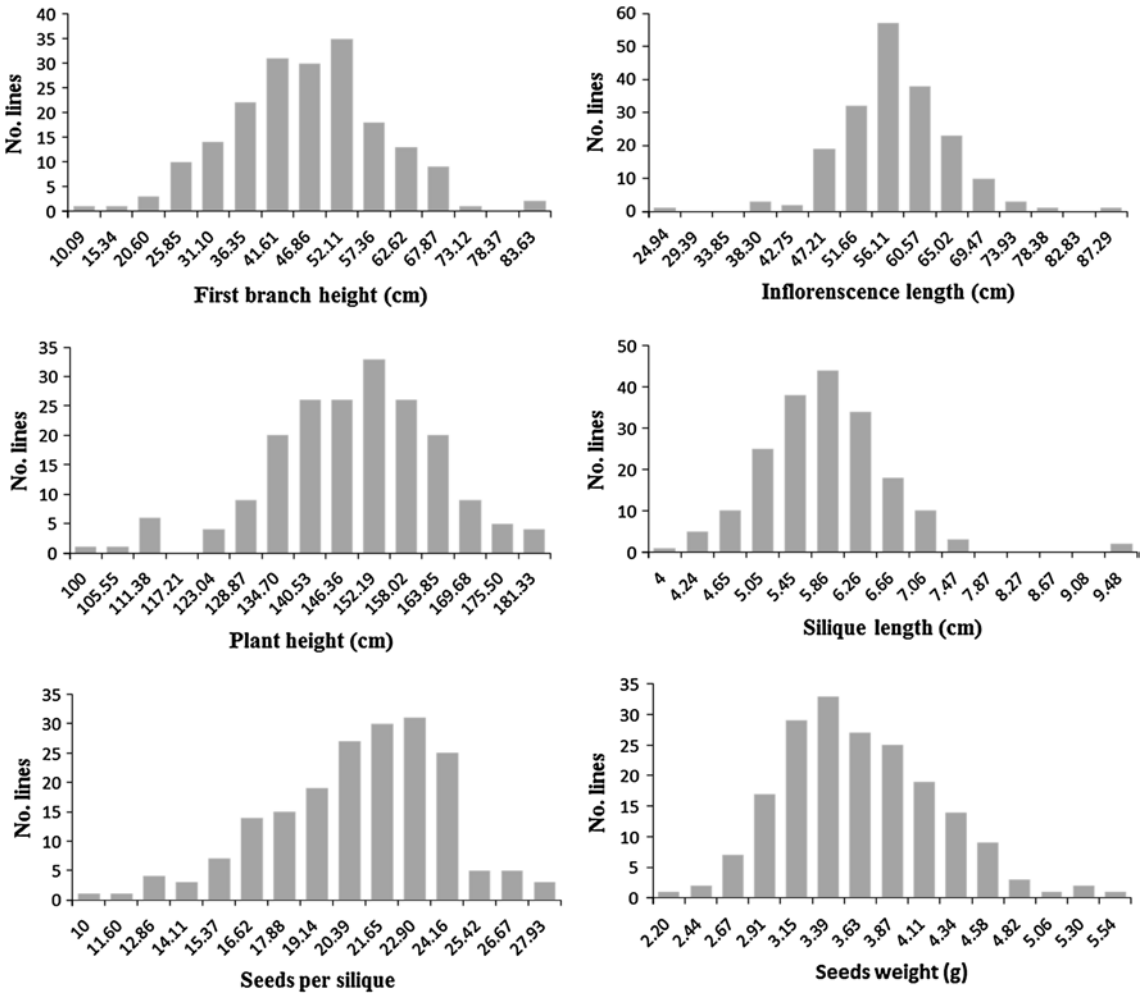

Fig. 1 Distributions of six yield-related traits with the means across 3 years

Table 2 Phenotypic and genetic correlations among six yield-related traits

	Plant height	First branch height	Inflorescence length	Silique length	Seeds per silique	Seed weight
Plant height		0.86	0.33	0.44	0.24	0.32
First branch height	0.74**		0.07	0.44	0.39	0.05
Inflorescence length	0.37**	−0.06		0.01	−0.16	0.54
Silique length	0.47**	0.34**	0.16		0.39	0.4
Seeds per silique	0.33**	0.28**	0.01	0.47**		−0.3
Seed weight	0.23**	0.09	0.35**	0.3**	−0.29**	

The values below and above the diagonal represent phenotypic correlation coefficients based on the means across 3 years, and genetic correlation coefficients based on the trait values in 3 years, respectively

** Significant at $P < 0.01$

SW, respectively, and the PCA + K model only detected 1.08, 0.54, 1.26, 1.26, 1.17 and 1.35 % markers significantly associated with FBH, IL, PH, SL, SPS and SW, respectively ($P < 0.01$; Table S1). Although the PCA + K model detected less associations than the Q + K model, the observed P values of the Q + K model were more closer to the expected P values than the PCA + K model, indicating that the Q + K model could effectively control false positive associations and avoid false negative associations. Thus, the Q + K model was chosen for association

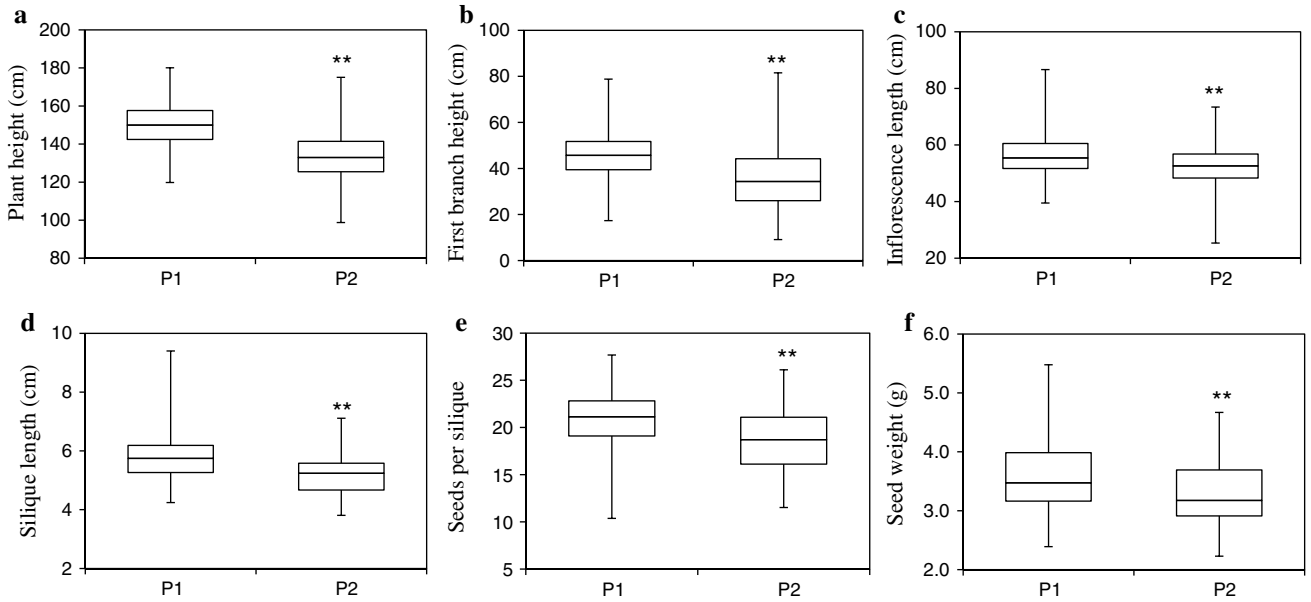

Fig. 2 Box plots of six yield-related traits in inbred lines grouped by population structure. The 192 rapeseed inbred lines were classified into two groups, P1 and P2, in previous study. The significance of difference between group means of each yield-related trait was estimated by *t* test ($P < 0.01$). **a** Plant height. **b** First branch height. **c** Inflorescence length. **d** Silique length. **e** Seeds per silique. **f** Seed weight

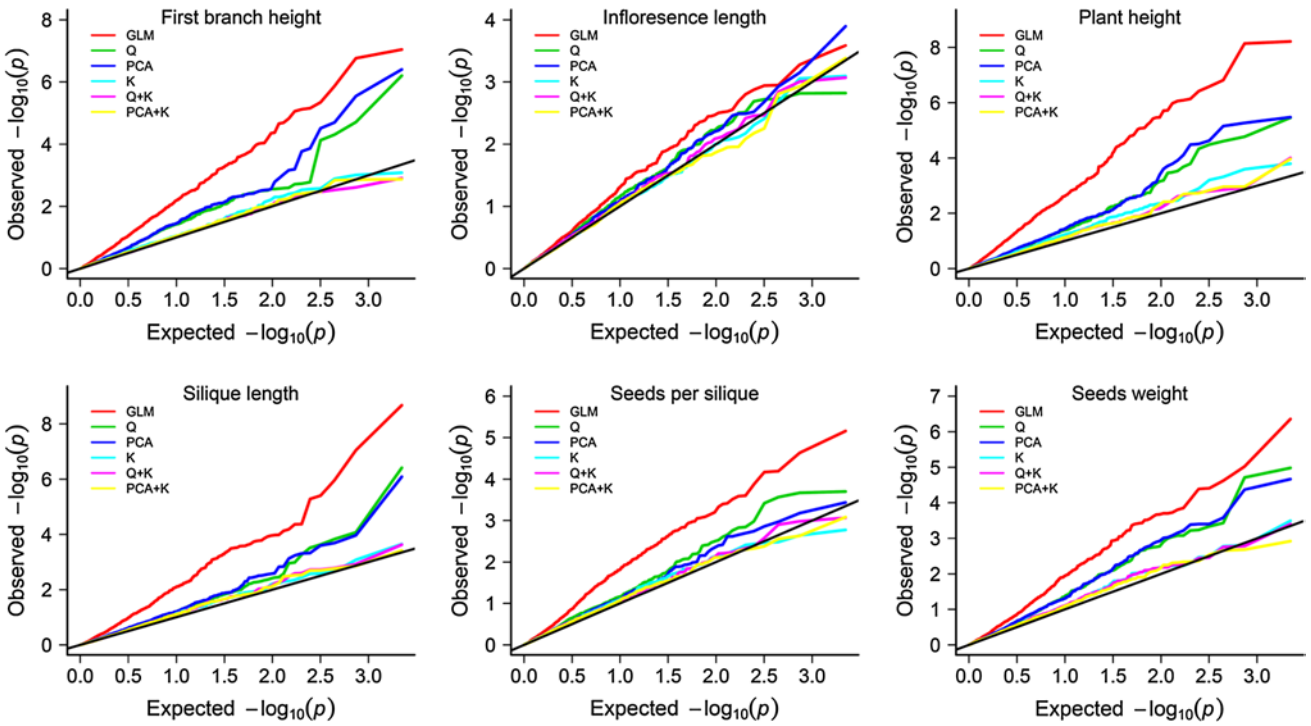

Fig. 3 Quantile–quantile plots of estimated −log10(*P*) from association analysis of six yield-related traits. The *black line* represents the expected *P* values with no association existed. The *red line* represents observed *P* values using the GLM model. The *green line* represents observed *P* values using the Q model. The *blue line* represents observed *P* values using the PCA model. The *cyan line* represents observed *P* values using the K model. The *pink line* represents observed *P* values using the Q + K model. The *yellow line* represents observed *P* values using the PCA + K model (color figure online)

analyses between the SSR and AFLP markers and the six yield-related traits.

Association mapping of six yield-related traits

Trait-marker associations were performed using the Q + K model. A total of 43 and 71 associations ($P < 0.001$) were detected for the six yield-related traits using the means across 3 years (Table 3) and within individual years (Table S2), respectively. Five markers were detected to be significantly associated with the mean values of FBH across 3 years. Of which, EA02MC04_7 and EA14MC08_1 were repeatedly detected in all 3 years, and EA14MC02_7 and EA09MC11_5 were repeatedly detected in 2 years (Table 3). EA14MC08_1 had the largest effect on FBH ($R^2 = 7.94$ %), with a phenotypic difference of 13.89 cm between alleles. Among these, EA02MC04_7 fell in the confidence interval of a FBH QTL *hpb13b* identified previously (Chen et al. 2007). Fourteen markers were significantly associated with the mean values of PH across 3 years. Among which, two markers (EA02MC04_7 and EA13MC08_2) were repeatedly detected in all 3 years, and seven markers (EA14MC08_1, EA09MC11_5, BnEMS56, BoGMS2306, EA03MC04_10, EA04MC14_7 and EA08MC14_9) were repeatedly detected in 2 years (Table 3). BrGMS2998 had the largest effect ($R^2 = 7.33$ %) with a plant height difference of 23.68 cm between alleles. Among the 14 associated markers, four (EA13MC08_2, BnEMS56, BoGMS2306 and EA03MC04_10) were co-localized with QTLs detected in previous studies (Chen et al. 2007; Shi et al. 2009). For SL, six significant markers were detected with the trait means across 3 years. Two markers, EA05MC08_1 and BrGMS1039, which were coincided with a major QTL (*cqSLA9*) for SL identified using recombinant inbred lines (Yang et al. 2012), were repeatedly detected in 3 years. Another two markers, BnEMS31 and BrGMS1040, were repeatedly detected in 2 years (Table 3). For SW, nine associated markers were detected with the mean values across 3 years. Among which, two markers (BoGMS2772 and BrGMS2199) were repeatedly detected in 3 years, and three markers (EA05MC08_1, BrGMS1409 and EA03MC04_12) were repeatedly detected in 2 years (Table 3). BrGMS2199 had the largest effect ($R^2 = 8.03$ %) with seed weight difference of 0.87 g between alleles. Four associations with SW are consistent with previous QTLs (Basunanda et al. 2010; Fan et al. 2010; Radoev et al. 2008; Shi et al. 2009). Two and seven markers were significantly associated with the mean values of IL and SPS across 3 years, respectively. But none was repeatedly detected for IL and SPS in two or three years (Table 3). Together, 18 markers were repeatedly detected with one of the six yield-related traits in at least 2 years. LD analyses showed that r^2 values of most pairs of the 18 markers were <0.15, except for the pairs of BrGMS1039, BrGMS1040 and EA05MC08_1 (Figure S1), suggesting that the majority of associated loci were not in high linkage disequilibrium with each other (Hasan et al. 2008), which make it feasible to combine the elite alleles from different loci into a single variety.

Five common markers (EA02MC04_7, EA14MC08_1, EA09MC11_5, EA14MC02_7 and EA03MC04_10) were associated with both FBH and PH, and one marker, EA05MC08_1, was associated with both SL and SW in 2 or 3 years (Table 3). To examine the independency of associations between associated loci and correlated traits, we conducted a conditional association analysis. Because FBH was correlated with PH, and SL was correlated with SW, conditional association analyses were performed using the Q + K model by setting one of the correlated traits as a covariate. The results of conditional association analyses indicated that the associations between EA02MC04_7, EA09MC11_5, EA14MC02_7 or EA03MC04_10 and both FBH and PH, and between EA05MC08_1 and SL and SW were probably caused by one QTL with pleiotropic effects, while the associations between EA14MC08_1 and PH and FBH might be caused by linked but independent QTLs (Table 4).

The 192 inbred lines were grouped according to combined genotypes at two associated loci to analyze their combined effects on yield-related traits (Table S3). EA02MC04_7 (A locus) and EA14MC08_1 (B locus) were associated with FBH and each detected two alleles, A_1 and A_0 for EA02MC04_7, and B_1 and B_0 for EA14MC08_1, respectively. The alleles at two loci formed four combined genotypes, A_1B_1, A_0B_1, A_1B_0 and A_0B_0. A_1B_1 had the highest FBH (62.86 ± 13.07 cm, $n = 8$), followed by A_0B_1 (51.33 ± 13.61 cm, $n = 7$), A_1B_0 (46.52 ± 10.72 cm, $n = 23$) and A_0B_0 (41.53 ± 11.14 cm, $n = 141$). The alleles at BoGMS2772 (A locus) and BrGMS2199 (B locus) formed six combined genotypes for SW, A_1B_1, A_1B_2, A_1B_3, A_2B_1, A_2B_2 and A_2B_3. A_1B_3 had the highest SW (4.87 ± 0.58 g, $n = 5$), followed by A_2B_3 (4.00 ± 0.73 g, $n = 9$), A_1B_1 (3.66 ± 0.52 g, $n = 17$), A_1B_2 (3.61 ± 0.29 g, $n = 3$), A_2B_2 (3.41 ± 0.53 g, $n = 7$) and A_2B_1 (3.37 ± 0.54 g, $n = 140$) (Table S3).

Discussion

The 192 lines have been divided into two groups previously, P1 and P2, which are consistent with their geographic origins of the inbred lines. Most lines in P1 were semi-winter type from China, and most lines in P2 were spring and winter types from Europe and Canada (Xiao et al. 2012). In this study, the P1 group contained more lines with high values of yield-related traits than the P2 group (Fig. 2), which

Table 3 Associations detected with the means of six yield-related traits

Traits	Loci[a]	MAF[b]	Genetic position		Physical position		−LogP[c]	R^2 (%)[d]	Effect[e]	Env[f]	Known loci[g]
			Chr.	Pos. (cM)	Chr.	Pos. (bp)					
First branch height	***EA02MC04_7***	0.17			C03	4,828,759	4.56	5.85	8.63	09, 10, 11	*hpb13b* (Chen et al. 2007)
	EA14MC08_1	0.08					6.09	7.94	13.89	09, 10, 11	
	EA09MC11_5	0.09					4.07	5	12.89	10, 11	
	EA14MC02_7	0.21			A08	21,595,819	4.43	5.65	11.17	09, 10	
	EA03MC04_10	0.41					3.11	3.82	2.24	mean	
Inflorescence length	BnGMS275	0.22			A05	1,300,572	3.22	4.58	3.99	10	
	EA04MG07_11	0.15					3.42	4.63	3.97	mean	
Plant height	***EA02MC04_7***	0.17			C03	4,828,759	5.51	6.67	10.88	09, 10, 11	
	EA13MC08_2	0.07			A03	20,262,420	4.80	5.7	19.09	09, 10, 11	*qPH.A3-2* (Shi et al. 2009)
	EA14MC08_1	0.08					4.75	5.67	13.47	09, 10	
	EA09MC11_5	0.09					4.37	4.96	16.46	10, 11	
	BnEMS56	0.14	A03	19.6	A03	26,833,930	3.32	5.58	11.41	09, 11	*qPH.A3-4* (Shi et al. 2009)
	BoGMS2306	0.33			C04	27,264,735	3.68	4.36	13.04	09, 11	*Hph14, ph14* (Chen et al. 2007)
	EA03MC04_10	0.41	A03	94.6			5.13	6.16	3.53	09, 10	*ph3* (Chen et al. 2007)
	EA04MC14_7	0.06			C06	6,122,537	4.09	4.77	21.09	09, 10	
	EA08MC14_9	0.29	A10	33.7			3.12	3.55	6.25	09, 11	
	EA14MC02_7	0.21					3.55	4.12	13.23	10	
	BrGMS2998	0.05	A03		A01	7,471,778	4.02	7.33	23.68	11	
	BrGMS3778	0.46			A03	31,765,688	3.05	3.67	3.99	mean	
	EA06MG11_8	0.19					3.08	3.51	7.65	mean	
	EA14MC12_5	0.08			C08	37,026,509	3.27	3.65	15.93	11	
Silique length	***EA05MC08_1***	0.23			A09	29,608,603	6.75	9.02	0.79	09, 10, 11	*cqSLA9* (Yang et al. 2012)
	BrGMS1039	0.27	A09		A09	29,998,693	4.22	5.41	0.68	09, 10, 11	*cqSLA9* (Yang et al. 2012)
	BnEMS31	0.13	A05	78.1	A05	21,375,632	3.76	5.4	0.50	10, 11	
	BrGMS1040	0.3	A09		A09	30,042,496	3.83	5.06	0.61	09, 11	*cqSLA9* (Yang et al. 2012)
	EA04MG07_10	0.45					3.29	3.96	0.17	10	
	EA01MC14_1	0.13			C04	5,550,490	3.16	3.78	0.75	09	
Seeds per silique	EA08MC10_5	0.32					3.45	4.44	1.10	09	
	BnGMS583	0.17			A06	15,358,871	3.63	4.78	0.73	09	
	EA06MC09_10	0.17					3.45	4.44	1.35	09	
	BrGMS4507	0.10	A08		A07	8,857,992	3.14	5.22	1.38	09	
	EA05MC08_8	0.35					3.07	3.96	0.78	mean	
	EA06MG08_9	0.46					3.02	3.85	2.11	10	
	EA06MG10_5	0.32					3.63	4.7	1.45	10	

Table 3 continued

Traits	Loci[a]	MAF[b]	Genetic position		Physical position		−LogP[c]	R^2 (%)[d]	Effect[e]	Env[f]	Known loci[g]
			Chr.	Pos. (cM)	Chr.	Pos. (bp)					
Seeds weight	BoGMS2772	0.15			C08	36,177,859	3.73	4.91	0.40	09, 10, 11	*qSW.C8-2* (Shi et al. 2009)
	BrGMS2199	0.05			A09	31,248,597	5.00	8.03	0.87	09, 10, 11	
	EA05MC08_1	0.23			A09	29,608,603	3.31	4.31	0.40	09, 10	*qSW.A9-6* (Shi et al. 2009)
	BrGMS1409	0.22	A02		A02	6,060,396	4.52	7.46	0.57	09, 10	
	EA03MC04_12	0.36			C09	20,712,124	3.00	3.78	0.30	10, 11	*TkwN19* (Radoev et al. 2008)
	BrGMS1389	0.27	A02		A02	8,160,663	3.31	6.02	0.46	11	
	BrGMS3983	0.07			A07	829,518	3.15	5.63	0.54	09	*qSW.A7-2* (Shi et al. 2009); *TSWA7b* (Fan et al. 2010); *TkwN7* (Radoev et al. 2008); DH-tsm06, DH-tsm07, MPH-tsm06, TH-tsm07 (Basunanda et al. 2010)
	EA06MG11_4	0.20			A06	20,873,888	3.25	4.17	0.39	10	
	EA09MC07_12	0.38			A10	16,407,166	3.24	4.31	0.38	09	

[a] Bold italics are common markers detected in two traits

[b] Minor allele frequency for each associated marker

[c] Negative logarithm value of *P* value of each associated marker

[d] Percentage of phenotypic variance explained by each associated marker

[e] Phenotypic differences between different genotypes classified on alleles of associated markers

[f] 09, 10 and 11 represented the environments of years 2009, 2010, and 2011, respectively; 'mean' represented associations only detected with the mean values across 3 years

[g] Comparision of trait-marker associations identified in this study with QTLs identified in previous studies

Table 4 Conditional association analyses for first branch height, plant height, silique length and seeds weight

Traits	Covariate	Loci	−Log10(*P*)	
			Unconditioned	Conditioned
First branch height	Plant height	EA02MC04_7	5.49	0.70
		EA14MC08_1	4.79	2.11
		EA09MC11_5	4.36	0.89
		EA14MC02_7	3.63	1.51
		EA03MC04_10	5.26	0.10
Seed weight	Silique length	EA05MC08_1	3.36	1.92

is accordance with the ANOVA results (Table 1), indicating that population structure had significant impacts on phenotypic variation for all the traits, which is similar to the findings observed in *B. rapa* (Del Carpio et al. 2011) and maize (Yang et al. 2010). Associations between the six traits and population structure may be caused not only by the genetic differentiation between P1 and P2 with regard to genetic diversity, *Fst* and LD level (Xiao et al. 2012), but also by the adaptation of inbred lines to distinct environments. Winter-type oilseed rape (OSR) is more adaptable to winter hardiness and requires strong vernalization to flower. Winter-type OSR generally had a prostrate growth at the vegetative stage and late transition from vegetative to reproductive development in the semi-winter growing environment of Wuhan, which results in a lower FBH and PH compared to semi-winter and spring-type OSR.

The genetic bases of yield-related traits have been thoroughly dissected in many crops by QTL mapping (Maccaferri et al. 2008; Xing and Zhang. 2010), and several QTLs with major effects had been identified using the strategy of map-based cloning (Jiao et al. 2010; Li et al. 2011c). In this study, the genetic bases of six yield-related traits were analyzed using association mapping and a total of 43 associations were detected using the means of six yield-related traits across 3 years. Of these, 18 markers were repeatedly detected in 2 or 3 years ($P < 0.001$; Table 3), suggesting that the QTLs associated with these markers were insensitive to the growing environments (Wen et al. 2009; Zorić et al. 2012). While the other associations were environment-specific, indicating that phenotypic plasticity plays an important role in plant agronomic diversity (Ungerer et al. 2003). Based on alignment of the associated markers and common markers to genomic sequences, 12 associations were identified to be co-localized with QTLs identified in previously studies (Table 3). However, due to lack of common markers between the genetic maps, the other marker-trait associations could not align to previous QTLs. Thus, the loci that were repeatedly detected in 2 or 3 years or consistent with QTLs identified in previous studies would be very useful for marker-assisted selection of the yield-related traits (Table 3). Fine mapping of such chromosomal regions would help to determine the candidate genes responsible for natural variation of these yield-related traits.

Yield and yield-related traits are mainly inherited in an additive manner (Xing and Zhang 2010; He et al. 2010), which makes breeding for these traits more straightforward. Although the associated markers individually explained a small portion of phenotypic variance in this study (<10 %; Table 3), the combination of the favorable alleles from a few significant loci had the potential of exhibiting much larger effects (Jia et al. 2012; Li et al. 2013). In our study, EA02MC04_7 and EA13MC08_2 explained 6.67 and 5.70 % of phenotypic variance of PH, respectively, while they jointly explained 20.92 % of phenotypic variance. The plants with A_0B_1 genotype ($n = 9$, 120.90 ± 16.30 cm) were much shorter than those with the other three genotypes (145.15 ± 13.46, 150.65 ± 14.44 and 154.89 ± 12.98 cm, respectively) (Table S3). Additionally, two inbred lines [Bugle (TT) and Karod (TT)] contained nine favorable alleles at the nine associated markers had the lowest stature (99.72 and 102.92 cm, respectively), while the line (H-5) contained only two favorable alleles had the highest stature (180.08 cm). Three lines (Guizhouqianxuan6, 324, Guizhouqianxuan) having five favorable SW alleles at the five associated markers had the highest seed weight (5.09, 5.26 and 5.49 g, respectively), while the lines (Bvonowshi-DH and Bolko) containing no favorable alleles had the lowest seed weight (2.20 and 2.24 g, respectively). These results suggested that the pyramiding of favorable alleles with minor effects is an effective way to enhance trait performance of rapeseed variety for yield-related traits, which depict us the perspective of application in molecular breeding.

Acknowledgments The research was supported by the National Natural Science Foundation of China (No. 31071452) and the Doctoral Fund of Ministry of Education of China (No. 20100146110019).

References

Basunanda P, Radoev M, Ecke W, Friedt W, Becker HC, Snowdon RJ (2010) Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). Theor Appl Genet 120:271–281

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Chen W, Zhang Y, Liu X, Chen B, Tu J, Fu T (2007) Detection of QTL for six yield-related traits in oilseed rape (*Brassica napus*) using DH and immortalized F2 populations. Theor Appl Genet 115:849–858

Cheng X, Xu J, Xia S, Gu J, Yang Y, Fu J, Qian X, Zhang S, Wu J, Liu K (2009) Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. Theor Appl Genet 118:1121–1131

Del Carpio DP, Basnet RK, De Vos RCH, Maliepaard C, Paulo MJ, Bonnema G (2011) Comparative methods for association studies: a case study on metabolite variation in a *Brassica rapa* core collection. PLoS ONE 6:e19624

Fan C, Cai G, Qin J, Li Q, Yang M, Wu J, Fu T, Liu K, Zhou Y (2010) Mapping of quantitative trait loci and development of allele-specific markers for seed weight in *Brassica napus*. Theor Appl Genet 121:1289–1301

Harper AL, Trick M, Higgins J, Fraser F, Clissold L, Wells R, Hattori C, Werner P, Bancroft I (2012) Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. Nat Biotechnol 30:798–802

Hasan M, Friedt W, Pons-Kühnemann J, Freitag NM, Link K, Snowdon RJ (2008) Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*). Theor Appl Genet 116:1035–1049

He Q, Zhang K, Xu C, Xing Y (2010) Additive and additive × additive interaction makes important contributions to spikelets per panicle in rice near isogenic (*Oryza sativa* L.) lines. J Genet Genomics 37:795–803

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108

Honsdorf N, Becker HC, Ecke W (2010) Association mapping for phenological, morphological, and quality traits in canola quality winter rapeseed (*Brassica napus* L.). Genome 53:899–907

Jannink JL, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics 9:166–177

Jia L, Yan W, Zhu C, Agrama HA, Jackson A, Yeater K, Li X, Huang B, Hu B, McClung A, Wu D (2012) Allelic analysis of sheath blight resistance with association mapping in rice. PLoS ONE 7:e32703

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X (2010) Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. Nat Genet 42:541–544

Li H, Chen X, Yang Y, Xu J, Gu J, Fu J, Qian X, Zhang S, Wu J, Liu K (2011a) Development and genetic mapping of microsatellite markers from whole genome shotgun sequences in *Brassica oleracea*. Mol Breed 28:585–596

Li X, Yan W, Agrama H, Jia L, Shen X, Jackson A, Moldenhauer K, Yeater K, McClung A, Wu D (2011b) Mapping QTLs for improving grain yield using the USDA rice mini-core collection. Planta 234:347–361

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J, He Y, Zhang Q (2011c) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. Nat Genet 43:1266–1269

Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N, Liu J, Warburton ML, Cheng Y, Hao X, Zhang P, Zhao J, Liu Y, Wang G, Li J, Yan J (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet 45:43–50

Maccaferri M, Sanguineti MC, Corneti S, Ortega JL, Salem MB, Bort J, DeAmbrogio E, del Moral LF, Demontis A, El-Ahmed A, Maalouf F, Machlab H, Martos V, Moragues M, Motawaj J, Nachit M, Nserallah N, Ouabbou H, Royo C, Slama A, Tuberosa R (2008) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. Genetics 178:489–511

Özer H, Oral E, Dogru U (1999) Relationships between yield and yield components on currently improved spring rapeseed cultivars. Tr J Agric For 23:603–607

Pinstrup-Andersen P, Hazell PBR (1985) The impact of the green revolution and prospects for the future. Food Rev Int 1:1–25

Quijada P, Udall J, Lambert B, Osborn T (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 1. Identification of genomic regions from winter germplasm. Theor Appl Genet 113:549–561

Radoev M, Becker HC, Ecke W (2008) Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by QTL mapping. Genetics 179:1547–1558

Raman R, Raman H, Kadkol G, Coombes N, Taylor B, Luckett D (2011) Genome-wide association analyses of loci for shatter resistance in *Brassicas*. In: Australian Research Assembly on *Brassicas*: Wagga Wagga NSW Australia, pp 36–41

Salas Fernandez MG, Becraft PW, Yin Y, Lübberstedt T (2009) From dwarves to giants? Plant height manipulation for biomass yield. Trends Plant Sci 14:454–461

Shi J, Li R, Qiu D, Jiang C, Long Y, Morgan C, Bancroft I, Zhao J, Meng J (2009) Unraveling the complex trait of crop yield with quantitative trait loci mapping in *Brassica napus*. Genetics 182:851–861

Shi J, Li R, Zou J, Long Y, Meng J (2011) A dynamic and complex network regulates the heterosis of yield-correlated traits in rapeseed (*Brassica napus* L.). PLoS ONE 6:e21645

Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, Stolk L, Nica A, Inouye M, Hofman A, Stephens J, Wheeler E, Arp P, Gwilliam R, Jhamai PM, Potter S, Chaney A, Ghori MJR, Ravindrarajah R, Ermakov S, Estrada K, Pols HAP, Williams FM, McArdle WL, van Meurs JB, Loos RJF, Dermitzakis ET, Ahmadi KR, Hart DJ, Ouwehand WH, Wareham NJ, Barroso I, Sandhu MS, Strachan DP, Livshits G, Spector TD, Uitterlinden AG, Deloukas P (2009) Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. PLoS Genet 5:e1000445

The SAS Institute (1999) SAS/STAT user's guide, version 8. SAS Institute, Cary

Udall J, Quijada P, Lambert B, Osborn T (2006) Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 2. Identification of alleles from unadapted germplasm. Theor Appl Genet 113:597–609

Ungerer MC, Halldorsdottir SS, Purugganan MA, Mackay TFC (2003) Genotype-environment interactions at quantitative trait loci affecting inflorescence development in *Arabidopsis thaliana*. Genetics 165:353–365

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wang F, Wang X, Chen X, Xiao Y, Li H, Zhang S, Xu J, Fu J, Huang L, Wu J, Liu K (2011) Abundance, marker development and genetic mapping of microsatellites from unigenes in *Brassica napus*. Mol Breed 30:731–744

Wen W, Mei H, Feng F, Yu S, Huang Z, Wu J, Chen L, Xu X, Luo L (2009) Population structure and association mapping on chromosome 7 using a diverse panel of Chinese germplasm of rice (*Oryza sativa* L.). Theor Appl Genet 119:459–470

Wurschum T, Liu W, Maurer HP, Abel S, Reif JC (2012) Dissecting the genetic architecture of agronomic traits in multiple segregating populations in rapeseed (*Brassica napus* L.). Theor Appl Genet 124:153–161

Xiao Y, Cai D, Yang W, Ye W, Younas M, Wu J, Liu K (2012) Genetic structure and linkage disequilibrium pattern of a rapeseed (*Brassica napus* L.) association mapping panel revealed by microsatellites. Theor Appl Genet 125:437–447

Xing Y, Zhang Q (2010) Genetic and molecular basis of rice yield. Annu Rev Plant Biol 61:421–442

Xu J, Qian X, Wang X, Li R, Cheng X, Yang Y, Fu J, Zhang S, King G, Wu J, Liu K (2010) Construction of an integrated genetic linkage map for the A genome of *Brassica napus* using SSR markers derived from sequenced BACs in *B. rapa*. BMC Genomics 11:594

Yang X, Gao S, Xu S, Zhang Z, Prasanna BM, Li L, Li J, Yan J (2010) Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. Mol Breeding 28:511–526

Yang P, Shu C, Chen L, Xu J, Wu J, Liu K (2012) Identification of a major QTL for silique length and seed weight in oilseed rape (*Brassica napus* L.). Theor Appl Genet 125:285–296

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zhang L, Yang G, Liu P, Hong D, Li S, He Q (2011) Genetic and correlation analysis of silique-traits in *Brassica napus* L. by quantitative trait locus mapping. Theor Appl Genet 122:21–31

Zorić M, Dodig D, Kobiljski B, Quarrie S, Barnes J (2012) Population structure in a wheat core collection and genomic loci associated with yield under contrasting environments. Genetica 140:259–275

Zou J, Jiang C, Cao Z, Li R, Long Y, Chen S, Meng J (2010) Association mapping of seed oil content in *Brassica napus* and comparison with quantitative trait loci identified from linkage mapping. Genome 53:908–916