Theor Appl Genet (2011) 122:1005–1016
DOI 10.1007/s00122-010-1505-4

ORIGINAL PAPER

Association mapping of grain color, phenolic content, flavonoid content and antioxidant capacity in dehulled rice

Yafang Shao · Liang Jin · Gan Zhang · Yan Lu · Yun Shen · Jinsong Bao

Received: 12 July 2010 / Accepted: 22 November 2010 / Published online: 15 December 2010

Abstract Phytochemicals such as phenolics and flavonoids in rice grain are antioxidants that are associated with reduced risk of developing chronic diseases including cardiovascular disease, type-2 diabetes and some cancers. Understanding the genetic basis of these traits is necessary for the improvement of nutritional quality by breeding. Association mapping based on linkage disequilibrium has emerged as a powerful strategy for identifying genes or quantitative trait loci (QTL) underlying complex traits in plants. In this study, genome-wide association mapping using models controlling both population structure (Q) and relative kinship (K) were performed to identify the marker loci/QTLs underlying the naturally occurring variations of grain color and nutritional quality traits in 416 rice germplasm accessions including red and black rice. A total of 41 marker loci were identified for all the traits, and it was confirmed that *Ra* (i.e., *Prp-b* for purple pericarp) and *Rc* (brown pericarp and seed coat) genes were main-effect loci for rice grain color and nutritional quality traits. RM228, RM339, *fgr* (fragrance gene) and RM316 were important markers associated with most of the traits. Association mapping for the traits of the 361 white or non-pigmented rice accessions (i.e., excluding the red and black rice) revealed a total of 11 markers for four color parameters, and one marker (RM346) for phenolic content. Among them, *Wx* gene locus was identified for the color parameters of lightness (L^*), redness (a^*) and hue angle (H^o). Our study suggested that the markers identified in this study can feasibly be used to improve nutritional quality or health benefit properties of rice by marker-assisted selection if the co-segregations of the marker–trait associations are validated in segregating populations.

Communicated by J. Yu.

Y. Shao and L. Jin have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-010-1505-4) contains supplementary material, which is available to authorized users.

Y. Shao · L. Jin · G. Zhang · Y. Lu · Y. Shen · J. Bao (✉)
Institute of Nuclear Agricultural Sciences,
Key Laboratory of Zhejiang Province and Chinese Ministry of Agriculture for Nuclear-Agricultural Sciences,
Zhejiang University, Hua Jiachi Campus,
Hangzhou 310029, People's Republic of China
e-mail: jsbao@zju.edu.cn

Introduction

Rice (*Oryza sativa*) is one of the most important cereal crops for human consumption in the world. The quality of rice affects consumers' acceptance and market value; therefore, it is one of main objectives of rice breeding activities. The quality traits encompass physical appearance, cooking and eating properties and, more recently, nutritional value (Fitzgerald et al. 2009).

The importance of nutritional quality can be viewed in two ways. On the one hand, micronutrient deficiency has been recognized in developing countries where rice is the main food, and fortification of nutrients by processing or biofortification by transgenic engineering to address particular deficiencies has emerged (Bouis et al. 2003; Welch and Graham 2004; Zimmermann and Hurrell 2002; Ye et al. 2000). On the other hand, the frequency of lifestyle-related diseases such as diabetes, hypertension and obesity has increased over the last few decades in developed countries (Takaiwa et al. 2008). Many epidemiological studies have provided evidence that reduced risk of these diseases and some cancers is associated with the intake of

whole grain including rice (Seal 2006; Vitaglione et al. 2008). Whole grains have become popular in western countries, but more gradually accepted in developing countries with the improvement of living standards.

Rice grain has a bark-like, protective hull, beneath which are the endosperm, bran and germ. After the hull is removed, it is called brown rice or dehulled rice. Further polishing of the brown rice produces milled white rice, the form that is generally consumed. However, polishing of the brown rice to obtain milled rice leads to loss of most of the nutritional components of the rice grain that are mostly deposited in the bran and germ. With only the inedible hull removed, brown rice is a whole grain, and it is an important source of dietary fiber, vitamins, phenolics and other components, which have been demonstrated to be beneficial for human health.

Most rice varieties that are planted and consumed throughout the world have white pericarp, but rice can also produce grains with red, and black or purple pericarp. The color is visible when the grains are dehulled, but it can be removed by polishing to reveal the white endosperm. Farming and consumption of colored varieties is limited in Western countries, but in some growing areas of Asia, traditional varieties with colored pericarp are particularly valued in local markets (Finocchiaro et al. 2007). Numerous studies have shown that consumption of colored rice causes decrease of oxidative stress and simultaneous increases of antioxidant capacity in the tested models, which are associated with reduced risk of developing chronic diseases, i.e., cardiovascular disease, type 2 diabetes and some cancers (Hu et al. 2003; Ling et al. 2001; Liu 2007; Toyokuni et al. 2002; Xia et al. 2003; Yawadio et al. 2007). Such health properties have been related to three well-known classes of antioxidant compounds present in rice: γ-oryzanols, tocols and polyphenols. Rice polyphenols include phenolic acids that are present in kernels of non-pigmented rice (Tian et al. 2004) and flavonoids that characterize pigmented rices (Reddy et al. 1995). The red pigment in red rice grains is proanthocyanidin, also called condensed tannins (Oki et al. 2002; Sweeney et al. 2006). Two loci have been identified using classical genetic analysis, *Rc* (brown pericarp and seed coat) and *Rd* (red pericarp and seed coat). When present together, these loci produce red seed color. *Rc* in the absence of *Rd* produces brown seeds, whereas *Rd* alone has no phenotype (Sweeney et al. 2006). Two groups independently cloned the *Rc* gene (Furukawa et al. 2007; Sweeney et al. 2006) and sequenced the wild-type allele (*Rc*), the domestication allele (*rc*) and a mutant allele (*Rc-s*; Sweeney et al. 2006). Both groups demonstrated that *rc* was a null-allele, produced by a 14-bp deletion, which caused a frame shift mutation and a premature stop codon (Brooks et al. 2008). The *Rd* has also been cloned, encoding a dihydroflavonol-4-reductase (Furukawa et al. 2007). The color of purple and black grain has been associated with the presence of anthocyanins (Reddy et al. 1995). Cyanidin-3-glucoside and peonidin-3-glucoside are the two main pigments deposited in grain pericarp of black rice (Abdel-Aal et al. 2006; Zhu et al. 2010). Classical genetic analysis indicated that two loci, *Pb* (*Prp-b*) and *Pp* (*Prp-a*), located on chromosome 4 and 1, respectively (Yoshimura et al. 1997), are required for the pericarp pigmentation with anthocyanins of black rice. Wang and Shu (2007) mapped *Pb* on rice chromosome 4 and suggested that the *Ra* gene may be the *Pb* gene, and the purple pericarp characteristic of rice is caused by a two bases deletion (GT) within exon 7 of the *Ra*.

It is evident that pigmented rice has higher amounts of phytochemicals than non-pigmented rice (Goffman and Bergman 2004; Shen et al. 2009). The color parameters of rice grain are positively correlated with the total phenolics, flavonoid contents and antioxidant capacity among a wide collection of rice germplasm including white, red and black rice, but the correlations among the white rice accessions are rather weak (Shen et al. 2009). For the improvement of the phytochemicals in rice grain by conventional breeding, it is necessary to understand the genetic bases for phenolic content, flavonoid content and antioxidant capacity and their relation to color parameters of brown rice. Linkage mapping of the quantitative trait loci (QTL) underlying the color parameter of milled rice flour was reported by Tan et al. (2001) using a recombinant inbred line population. They identified three, two and four QTLs for the color parameters of lightness (L^*), redness (a^*) and yellowness (b^*). Recently, Jin et al. (2009) reported the QTLs responsible for brown rice color, total phenolics and flavonoid contents and antioxidant capacity in rice grain using a doubled haploid population, the results indicating that each trait was controlled by several QTLs with minor effects.

In addition to the linkage mapping, association mapping, also known as linkage disequilibrium (LD) mapping, has recently emerged as an alternative approach to mapping QTLs by examining the marker–trait associations, and enables researchers to use modern genetic technologies to exploit natural diversity and locate valuable genes in the genome (Hall et al. 2010; Stich and Melchinger 2010; Zhu et al. 2008). In conventional linkage mapping studies, the LD required for mapping quantitative trait loci is generated as a result of the mating design, whereas association studies exploit the LD already present in the population of interest. The rate of decay of LD with physical or genetic distance is important to determine the marker density required for scanning previously unexploited regions of the genome, as well as the maximum resolution that can be achieved for association mapping (Stich and Melchinger 2010). The rate

of decay of LD varied in different plant taxa, such as maize, barley, *Arabidopsis thaliana*, rice, etc. (Stich and Melchinger 2010; Ersoz et al. 2009; Zhu et al. 2008). In rice, the studies of Garris et al. (2005), Olsen et al. (2006), Mather et al. (2007) and Rakshit et al. (2007) indicated that LD in rice decays at 1 cM or less using DNA sequences, whereas other studies indicated that LD decays at 20–30 cM or more using simple sequence repeat (SSR) markers (Agrama et al. 2007; Agrama and Eizenga 2008; Jin et al. 2010a). In large populations of autogamous species such as rice, the stretches of LD extending over several cM indicated that genome-wide LD mapping was possible in rice populations.

A concern about association mapping is that marker–trait associations may arise from confounding population structure, which may cause spurious correlations, leading to an elevated false-positive rate (Hall et al. 2010; Pritchard et al. 2000; Stich and Melchinger 2010; Zhao et al. 2007; Zhu et al. 2008). Population structure can be obtained with the software STRUCTURE version 2.1 (Pritchard et al. 2000). The assigned subpopulations could be integrated into the association mapping as covariates. Yu et al. (2006) introduced a mixed model, which takes genome-wide differences in relatedness into account via estimated pair-wise kinship coefficients (K) and population structure (Q). Population structure could also be controlled by other estimates, for example, using the most significant principal component analysis (PCA) axes as covariates (Price et al. 2006; Zhao et al. 2007). With these methods, the issue of false positives generated by population structure can be well dealt with accordingly (Price et al. 2006; Yu et al. 2006; Zhao et al. 2007).

Many successful examples using association mapping in plants and crops come from *Arabidopsis* (Zhao et al. 2007), maize (Thornsberry et al. 2001; Yang et al. 2010), wheat (Breseghello and Sorrells 2006), rice (Bao et al. 2006a, b; Agrama et al. 2007; Agrama and Eizenga 2008; Agrama and Yan 2009). In rice, Agrama et al. (2007) for the first time mapped the loci associated with the complex trait of yield, including the kernel width, kernel length, kernel width/length ratio and 1,000-kernel weight using marker–trait association mapping approach-based Q+K model. It was found that many of the associated markers were located in regions where QTL had previously been identified (Agrama et al. 2007). Since then, Agrama and Yan (2009) conducted association mapping of straighthead disorder induced by arsenic. Wen et al. (2009) detected the population structure and conducted association mapping on chromosome 7 for heading date, plant height and panicle length using a diverse panel of Chinese germplasm of rice. Ordonez et al. (2010) carried out association mapping to identify candidate molecular markers for grain quality and flowering traits in a collection of elite rice *japonica* inbred lines. Iwata et al. (2010) conducted genome-wide association study of grain shape variation among rice germplasm accessions. All the research suggested that association mapping in rice has become a viable alternative to linkage mapping based on crosses between different lines.

The objective of this study was to identify the marker loci/QTLs underlying the naturally occurring variations of grain color and nutritional quality traits by association mapping. As a result, we identified a total of 41 marker loci for all the traits of the total rice panel, and revealed a total of 12 loci for all the traits of the white non-pigmented rice panel.

Materials and methods

Materials

A total of 416 rice accessions including 361 white rice, 50 red rice and 6 black rice were used in this study. This set of rice had been genotyped with genome-wide 100 simple sequence repeat (SSR) with an average interval of 18.8 cM, and a total of 390 alleles were detected (Jin et al. 2010a). The population structure (Q) was estimated using STRUCTURE software version 2.2 (Pritchard et al. 2000), and seven subpopulations were revealed, denoted as POP1, POP2, POP3, POP4, POP5, POP6 and POP7. Linkage disequilibrium (LD) could extend to 25–50 cM in different subpopulations (Jin et al. 2010a).

Phenotypic traits

The traits, i.e., the color parameter of brown rice grain, phenolic content, flavonoid content and antioxidant activity of each accession that have been reported before (Shen et al. 2009) were used in this study. Color measurements were expressed as tristimulus parameters, L^*, a^* and b^*. L^* indicates lightness (100 = white and 0 = black). a^* indicates redness-greenness (positive = red) and b^* indicates yellowness-blueness (positive = yellow). In addition, the chroma (C) value indicates color intensity or saturation, calculated as $C = (a^{*2} + b^{*2})^{1/2}$, and hue angle was calculated as $H^{o} = \tan^{-1}(b^*/a^*)$. Total phenolic content was assayed by the Folin–Ciocalteu colorimetric method and was expressed as milligram of gallic acid equivalent (mg GAE) per 100 g of dry weight. Total flavonoid content was determined by a colorimetric method, which was calculated using the standard rutin curve, and expressed as mg rutin equivalent (mg RE) per 100 g of dry weight. Total antioxidant capacity of rice extracts was carried out using a spectrophotometer by the improved 2,2-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (ABTS)

radical cation method. Results were expressed in terms of Trolox equivalent antioxidant capacity (TEAC, mM Trolox equivalents per 100 g dry weight).

Molecular markers and candidate gene markers

DNA was extracted following a CTAB procedure (Doyle 1991). A total of 108 molecular markers including 100 SSR markers (Jin et al. 2010a, 2010b), seven starch biosynthesizing gene markers (Bao et al. 2006a, b) and one fragrance gene (*fgr*) marker (Jin et al. 2010b) have been genotyped. *Rc* (Sweeney et al. 2006) and *Ra* (Wang and Shu 2007) gene markers were genotyped in this study (see below). All the markers with an allele frequency of 1% or greater were included in the association analysis.

For genotyping *Rc* and *Ra*, we designed primers using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome) according to their sequences reported by Sweeney et al. (2006) and Wang and Shu (2007), respectively. The primers for amplification of the 14-bp deletion in the *Rc* gene were designed according to the gene accession no. DQ204737 (Sweeney et al. 2006), with forward primer 5′-ATCAGTCCAGGCACCACA-3′, and reverse primer 5′-CCAAAGATCGCAGAATTATGA-3′. The length of the PCR product was 229 bp.

The CAPS (cleaved amplified polymorphic sequence) marker was designed to genotype the *Ra* gene (LOC_Os04g47080). The PCR product of 853 bp amplified by the forward primer 5′-ATTTCTTTGGCCACAGG CGA-3′ and reverse primer 5′-CCCAGATTCGGAACAA GAAC-3′ was digested by the restriction endonuclease *Bam*HI. The 2-bp (GT) deletion within the exon 7 of the *Ra* gene in black rice generated a *Bam*HI restriction site (Wang and Shu 2007), resulting in two digested products of 533 and 318 bp.

PCR amplification was carried out in a total volume of 20 μL containing 10 mM Tris–HCl (pH 9.0), 50 mM KCl, 0.1% Triton X-100, 2 mM $MgCl_2$, 0.1 mM dNTPs, 200 nM primers, 1 unit of *Taq* polymerase and 30 ng of genomic DNA. All amplifications were performed on an MG96G thermal cycler (Hangzhou LongGene Scientific Instruments Co. Ltd., Hangzhou, China) under the following conditions: (1) pre-denature at 95°C for 5 min; (2) 35 cycles of run, each followed by denature at 95°C for 1 min, anneal at 55°C for 45 s and extension at 72°C for 1 min; (3) final extension at 72°C for 10 min. The digested PCR products of *Ra* marker were separated by electrophoresis in 1.2% agarose in 0.5 × tris–borate EDTA (TBE) buffer, and those of *Rc* in 8% denaturing polyacrylamide gel (PAGE) with 3.4% cross-linker (ratio of bis-acrylamide to acrylamide) in 1.0 × TBE buffer and marker bands were visualized using silver staining.

Statistical analyses and association mapping

Analysis of variance (ANOVA) using general linear model (GLM) was performed with the SAS System for Windows version 8 (SAS Institute Inc., Cary, NC, USA).

To avoid possible spurious associations, both the Q (population structure) model and the Q+K (kinship) model (Yu et al. 2006) were used to account for population structure and relatedness of individuals among 416 rice accessions. The genetic structure (Q) among the 416 entries was previously estimated by 100 SSRs (Jin et al. 2010a) with the program STRUCTURE version 2.2 (Pritchard et al. 2000). The relative kinship (K) matrix was calculated on the basis of 100 SSR loci using the method proposed by Ritland (1996), which is built into the program SPAGeDi (Hardy and Vekemans 2002). The Q model was performed using GLM in TASSEL V2.1 using 50,000 time permutations for the correction of multiple testing. Markers with the adjusted P value <0.05 were regarded as significant. The K + Q model was performed using mixed linear model (MLM) in TASSEL V2.1 (Yu et al. 2006; Bradbury et al. 2007). The default run parameters of the convergence criterion set at 1.0×10^{-4} and the maximum number of iterations set at 200 were used. Correction for multiple testing was performed using the *q*FDR value, which is an extension of the false discovery rate (FDR) method (Benjamini and Hochberg 1995). The q values were calculated with the QVALUE R package using the smoother method (Storey and Tibshirani 2003). Markers with DFR $q < 0.05$ were regarded as significant.

Results

Analysis of *Rc* and *Ra* gene markers

A total of 416 rice accessions were checked for their genotype of *Rc* and *Ra* using gene markers (Fig. 1). For the *Rc* gene, among 50 red rice accessions, 43 accessions had the homozygous *RcRc* genotype, 5 accessions had the heterozygous *Rc/rc* genotype, and the other 2 had the homozygous *rcrc* genotype (Fig. 1a). White and black rice accessions all had the homozygous *rcrc* genotype.

The PCR products of six rice accessions with black pericarp could be digested by *Bam*HI that produced two fragments of 533 and 318 bp, indicating that they harbored the GT deletion. Other rice with white and red pericarp could not be cleaved by *Bam*HI (Fig. 1b).

Diversity of phenotypic traits

Substantial phenotypic diversity existed in grain color and nutritional quality measured in this diverse germplasm

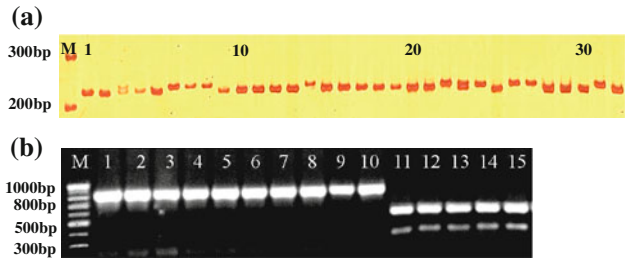

Fig. 1 Genotyping the *Rc* and *Ra* genes with gene-specific markers. **a** Differences of one *Rc* fragment containing a 14-bp deletion in the association population, 1–32: association population accessions; *M* marker. **b** Analysis of the PCR products of *Ra* after being digested by *BamHI*, 1–5: white rice accessions; 6–10: red rice accessions; 11–15: black rice accessions; *M* marker

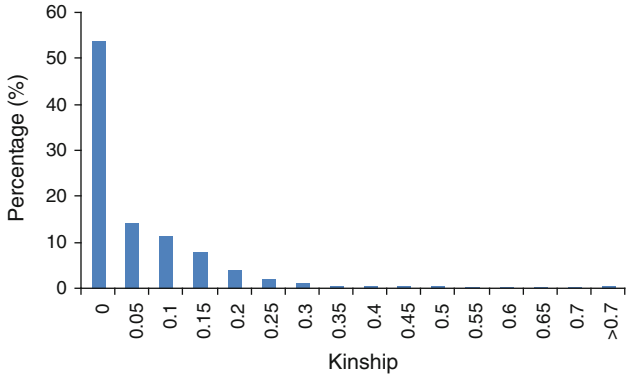

Fig. 2 Distribution of the pair-wise relative kinship estimates between 416 rice accessions. Values are from SPAGeDi estimates using 100 SSRs

panel (Shen et al. 2009). $H°$ was the most striking example of phenotypic variation, ranging from 193.3 to 101.4. The phenolic content ranged from 108.1 to 1244.9 mg of gallic acid equivalent (mg GAE) per 100 g of dry weight (Shen et al. 2009; Jin et al. 2010a). There were significant differences in these traits among different subpopulations (Table 1). POP2 and POP4 had larger mean lightness (L^*), yellowness (b^*) and chroma (C) values of color parameters than other subpopulations, while the POP6 had larger H^o value and higher phenolic content, flavonoid content and antioxidant capacity than other subpopulations (Table 1). However, the population structure was not the dominant factor for the phenotypic variation, accounting for less than 10% of the phenotypic variation across all traits in this study except the a^* of grain color (Jin et al. 2010a; also see Supplementary Table 1).

Estimation of relative kinship

Relative kinship estimates based on the SSR data showed that 54.4% of the pair-wise kinship estimates were equal to 0, suggesting that more than half of the total pairs of accessions had no relationship. As much as 40% of the estimates were less than 0.25, which indicated those accessions had weak relationship with the others in this rice panel (Fig. 2).

Association mapping among the whole rice panel

Association mapping using the *Rc*, *Ra*, *fragrance* (*fgr*), starch gene markers (Jin et al. 2010b) and 100 SSRs (Jin et al. 2010a) based on both Q model and Q+K model were performed, respectively. Without correction for multiple test, the number of markers identified in association with the traits were decreased with the increase of significance level, with 50% decrease from $P < 0.05$ to 0.01, and another ~50% decrease from $P < 0.01$ to $P < 0.001$. Compared to the average number of significant markers identified with the Q model, the average number identified with Q+K model was reduced by 23, 21 and 14% at $P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively (Table 2). It was found that the loci identified using the Q+K model at $P < 0.001$ (Table 2) were similar with those using Q model at $P < 0.001$ level, except that four and two more loci identified by the Q model for the color parameter C and a*, respectively. After 50,000 times of permutation test, the markers identified with the Q model at adjusted $P < 0.05$ was similar to those at $P < 0.001$ without permutation test except for one more marker for L^*, b^*, phenolic content and antioxidant capacity, respectively (Table 2; Supplementary Table 2). However, four more

Table 1 Mean values of the phenotypic traits among seven subpopulations

Subpopulation	L^*	a^*	b^*	C	H^o	Phenolic content	Flavonoid content	Antioxidant capacity
POP1	51.9 bc	7.01 a	14.87 bc	15.69 ab	63.6 cd	170.6 bc	135.7 abc	226.1 c
POP2	57.4 a	4.54 b	14.49 b	15.22 bc	72.5 abc	138.3 c	126.8 c	202.3 c
POP3	53.9 ab	2.10 c	14.49 b	15.00 c	78.1 ab	229.8 ab	131.6 bc	637.9 ab
POP4	54.6 ab	4.07 b	15.43 a	16.14 a	75.4 abc	163.7 bc	133.3 abc	229.5 c
POP5	48.8 c	4.03 b	14.19 b	15.28 bc	65.7 bcd	177.6 bc	131.6 bc	324.6 bc
POP6	48.9 c	1.87 c	14.30 b	14.62 c	82.9 a	277.6 a	139.9 ab	812.1 a
POP7	51.2 bc	4.14 b	13.07 c	14.63 c	57.5 d	210.9 ab	142.0 a	410.5 bc

Different letter in the same column was significant at $P < 0.05$

Table 2 The number of significant markers associated with eight traits using two statistic models, Q and Q+K, at different significant levels

	All rice						White (non-pigmented) rice					
	Q			Q+K			Q			Q+K		
	$P < 0.05$	$P < 0.01$	$P < 0.001$	Padj < 0.05[a]	$P < 0.05$	$P < 0.01$	$P < 0.001$	$q < 0.05$[b]	$P < 0.05$	$P < 0.01$	$P < 0.001$	Padj < 0.05[a]	$P < 0.05$	$P < 0.01$	$P < 0.001$	$q < 0.05$[b]
L^*	13	6	3	4	14	6	3	3	25	9	6	7	11	5	2	2
a^*	30	15	8	8	19	10	6	8	30	15	8	8	13	6	2	2
b^*	23	11	6	7	19	9	6	6	22	13	5	5	7	4	2	2
C	25	10	6	6	9	5	2	3	18	10	6	7	10	5	0	0
H^o	20	11	5	5	20	11	5	5	33	17	10	10	18	7	5	5
Phenolic content	16	8	4	5	15	8	4	4	20	10	5	5	14	6	1	1
Flavonoid content	23	11	5	5	18	7	5	6	21	4	0	1	14	3	1	0
Antioxidant capacity	18	9	6	7	15	8	6	6	20	8	1	2	4	0	0	0
Average	21	10.1	5.4	5.9	16.1	8.0	4.6	5.1	23.6	10.8	5.1	5.6	11.4	4.5	1.6	1.5

[a] The adjusted P values were obtained after 50,000 permutation test, and these markers are shown in Supplementary Tables 2 and 3, respectively

[b] The false discovery rate (DFR) or q values were obtained from QVALUE R package, and these markers are shown in Tables 3 and 4, respectively

markers for a^*, C and flavonoid content were identified with the Q+K model at qFDR<0.05 than at $P < 0.001$ without correction (Table 2). Table 3 summarizes the significant markers identified by the Q+K model.

For the five grain color traits, a total of 25 marker–trait associations were identified with 11 different markers (Table 3). Three genes were identified to be associated with grain color. *Ra* was significant for all the five traits and could explain the percentage of phenotypic variance ranging from 3.99% for parameter C and 86.7% for parameter H^o. *Rc* was significant for L^* and a*, and could explain 31.2% and 4.42% of the total variation for L^* and a^*, respectively. The fragrance gene (*fgr*) on chromosome 8 was identified for b^*, C and H^o.

Among the 8 SSRs, RM339 on chromosome 8 was detected for a^*, b^* and H^o, explaining relatively small variations (3.1–9.1%) for each trait. Because RM339 is closely linked to the *fgr* locus (Lorieux et al. 1996; also see http://www.gramene.org), its association could be derived from its strong linkage disequilibrium with *fgr*. RM316 on chromosome 9 was significant for a^*, b^*, and H^o and explained 6.5, 11 and 21.3% of the total phenotypic variance for each trait, which indicated that it might be a major QTL for H^o. RM228 on chromosome 10 was identified significant for L^*, b^* and C. In addition, RM297, RM337, RM224 and RM309 were also significant for a^*.

For three nutritional quality traits, a total of 16 marker–trait associations were identified with six different markers (Table 3). Four of these, *Ra*, *Rc*, RM339 and RM316, were common loci for phenolic content, flavonoids content and antioxidant capacity. For major genes, *Ra* had a larger effect than *Rc* for all the three traits. For minor loci, RM316 contributed more than RM339 to all the traits. In addition to these four loci, *fgr* and RM228 were significant for both flavonoid content and antioxidant capacity. The result confirmed that *Ra* and *Rc* were main-effect loci for rice grain color and nutritional quality traits, whereas RM316 might be another important loci with unknown function.

Association mapping in white (non-pigmented) rice accessions

The presence of pigmented rice accessions (red and black rice) might mask the effects of loci with minor contributions. Thus, it was necessary to conduct association mapping between the marker loci and grain traits among the 361 white rice accessions. Again, both of the Q model and Q+K model were tried. Without correction for multiple tests, the number of markers identified in association with the traits decreased with the increase of significance level, as reported for the whole rice panel. Unlike the whole rice panel, compared to the average number of significant

Table 3 The marker loci associated with grain color parameters, phenolic content, flavonoid content and antioxidant capacity in rice grain detected with Q+K model among the whole rice panel

Trait	Locus[a]	Chro. no.	Position (cM)	*P*	*qFDR*	R^{2b}
*L**	*Ra*	4	113.2	1.74×10^{-20}	9.57×10^{-19}	0.1665
	Rc	7	43.8–44.4	5.63×10^{-42}	6.19×10^{-40}	0.3123
	RM228	10	96.3	6.89×10^{-04}	0.0252	0.0496
*a**	RM297	1	155.9	0.0066	0.0415	0.0151
	Ra	4	113.2	4.36×10^{-48}	2.74×10^{-46}	0.1965
	Rc	7	43.8–44.4	7.53×10^{-10}	1.58×10^{-8}	0.0442
	RM337	8	1.1	0.0047	0.0369	0.017
	RM339	8	72.2	5.78×10^{-4}	0.0061	0.031
	RM316	9	1.8	9.43×10^{-12}	2.97×10^{-10}	0.0646
	RM224	11	120.1	2.09×10^{-4}	0.0033	0.0264
	RM309	12	74.5	3.87×10^{-4}	0.0049	0.0195
*b**	*Ra*	4	113.2	1.99×10^{-75}	1.39×10^{-73}	0.4982
	RM30	6	125.4	4.50×10^{-6}	0.0002	0.0614
	fgr	8	71.7–72.2	8.90×10^{-6}	0.0002	0.0438
	RM339	8	72.2	8.40×10^{-5}	0.0018	0.0666
	RM316	9	1.8	1.19×10^{-10}	6.55×10^{-9}	0.1099
	RM228	10	96.3	1.33×10^{-4}	0.0024	0.0608
C	*Ra*	4	113.2	3.01×10^{-8}	1.77×10^{-6}	0.0399
	fgr	8	71.7–72.2	0.0013	0.0255	0.0139
	RM228	10	96.3	1.09×10^{-4}	0.0032	0.0367
H°	*Ra*	4	113.2	1.11×10^{-212}	1.22×10^{-210}	0.8668
	RM30	6	125.4	1.38×10^{-5}	0.0003	0.0642
	fgr	8	71.7–72.2	4.06×10^{-6}	0.0001	0.054
	RM339	8	72.2	7.63×10^{-6}	0.0002	0.0905
	RM316	9	1.8	4.07×10^{-19}	2.24×10^{-17}	0.2132
Phenolic content	*Ra*	4	113.2	5.49×10^{-55}	6.05×10^{-53}	0.3967
	Rc	7	43.8–44.4	1.38×10^{-28}	7.59×10^{-27}	0.2373
	RM339	8	72.2	3.29×10^{-4}	0.0090	0.0587
	RM316	9	1.8	6.51×10^{-10}	2.39×10^{-8}	0.1022
Flavonoid content	*Ra*	4	113.2	3.20×10^{-51}	2.82×10^{-49}	0.3535
	Rc	7	43.8–44.4	3.92×10^{-6}	0.0001	0.0442
	fgr	8	71.7–72.2	3.59×10^{-4}	0.0079	0.0264
	RM339	8	72.2	5.38×10^{-4}	0.0094	0.0522
	RM316	9	1.8	6.23×10^{-9}	2.74×10^{-7}	0.0855
	RM228	10	96.3	0.0026	0.0382	0.0417
Antioxidant capacity	*Ra*	4	113.2	1.38×10^{-58}	1.39×10^{-56}	0.4254
	Rc	7	43.8–44.4	7.89×10^{-24}	3.96×10^{-22}	0.2067
	fgr	8	71.7–72.2	7.24×10^{-4}	0.0121	0.0259
	RM339	8	72.2	6.69×10^{-5}	0.0017	0.0687
	RM316	9	1.8	1.30×10^{-8}	4.35×10^{-7}	0.0915
	RM228	10	96.3	5.21×10^{-4}	0.0105	0.0542

[a] Markers with a significant marker–trait association are reported at *q*FDR value <0.05

[b] R^2 indicates the percentage of the total variation explained by each locus

markers identified with Q model, the average number identified with Q+K model was reduced by 51, 58 and 68% at $P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively (Table 2). After 50,000 times of permutation test, the markers identified with the Q model at adjusted $P < 0.05$ was similar to those at $P < 0.001$ without permutation test,

Table 4 The marker loci associated with grain color parameters, phenolic content, flavonoid content and antioxidant capacity in rice grain detected with Q+K model among the white (non-pigmented) rice panel

Trait	Locus[a]	Chro. no.	Position (cM)	*P* value	*q*-FDR	R^{2b}
*L**	*Wx* SNP	6	7.4	1.44×10^{-7}	1.43×10^{-5}	0.0194
	Wx SSR	6	7.4	2.36×10^{-5}	0.0012	0.0249
*a**	*Wx* SSR	6	7.4	3.24×10^{-4}	0.0122	0.0040
	RM454	6	99.3	3.06×10^{-4}	0.0122	0.0040
*b**	RM201	9	81.2	9.01×10^{-4}	0.0464	0.0237
	RM171	10	73.0	0.0023	0.0464	0.0148
H^o	*Wx* SSR	6	7.4	8.60×10^{-5}	0.0050	0.0115
	RM454	6	99.3	1.62×10^{-4}	0.0058	0.0051
	RM219	9	11.7	0.0023	0.0497	0.006
	RM286	11	0.0	9.35×10^{-5}	0.0050	0.0065
	RM224	11	120.1	6.50×10^{-4}	0.0176	0.007
Phenolic content	RM346	7	47.0	2.33×10^{-4}	0.0219	0.0247

[a] Markers with a significant marker–trait association are reported at *q*FDR value <0.05

[b] R^2 indicates the percentage of the total variation explained by each locus

except one more marker for the *L**, *C*, flavonoid content and antioxidant capacity, respectively (Table 2; Supplementary Table 3). The number of markers identified with Q+K model at qFDR<0.05 was similar to that at $P < 0.001$ without correction (Table 2). Table 4 summarizes the significant markers identified by the Q+K model.

Based on Q model, a total of 37 loci were identified as significant for grain color at adjusted $P < 0.05$, while 5 loci were identified for phenolic content, one locus for flavonoid content and two loci for antioxidant capacity (Table 2; Supplementary Table 3). Only 11 loci in total were identified based on Q+K model for color parameters, and only one locus (RM346) was significant for phenolic content (Table 4). No locus was significant for color parameter *C*, flavonoid content and antioxidant capacity based on Q+K model (qDFR < 0.05).

The *Wx* gene on chromosome 6 was highly associated with color parameters *L**, *a** and H^o. Another marker, RM454 also on chromosome 6, was associated with color parameter *a** and H^o. In addition, three markers, RM219, RM286 and RM224, were detected for H^o, and two markers, RM201 and RM171, were detected for *b**. All markers only explained a small part of phenotypic variations (Table 4).

Allelic effect

The model-based simulation of population structure using 100 SSR markers showed that there were seven subpopulations, denoted as POP1, POP2, POP3, POP4, POP5, POP6 and POP7 (Jin et al. 2010a). The allelic effect for each loci associated with the traits could be estimated in comparison to the mean phenotypic data of each trait for each locus (Fig. 3). *Wx* gene was associated with *L**, *a** and H^o (Table 4). Especially, the single nucleotide polymorphism (SNP) of *Wx* was highly associated with *L** with qDFR value of 1.43×10^{-5}. We compared the *L** value of different SNP alleles of *Wx* among the whole non-pigmented rice and each subpopulations. Among all the white rice, rice carrying allele 2 had larger *L** value than allele 1. Among each subpopulation, rice carrying allele 2 still showed larger *L** value than allele 1 (Fig. 3a).

Only RM346 was identified for phenolic content. After removal of the rare alleles, four alleles were used to

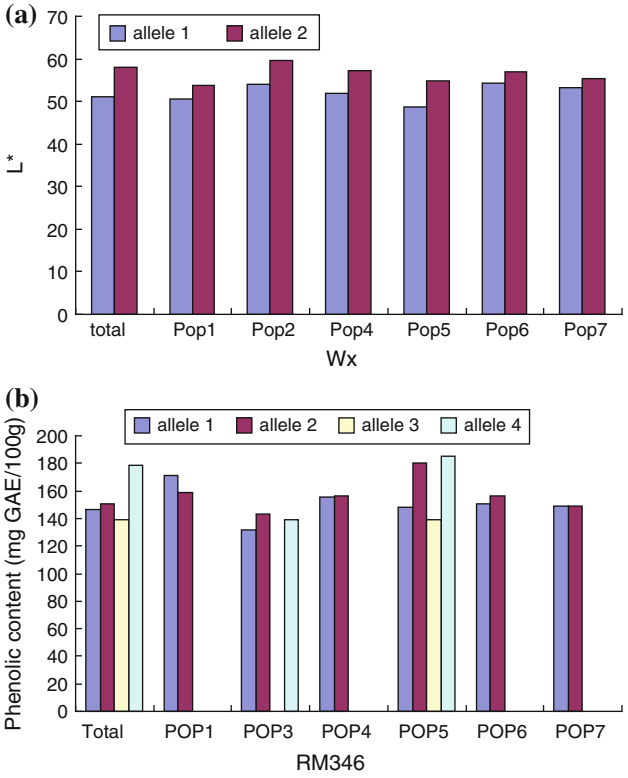

Fig. 3 Comparisons of the mean values of *L** and phenolic content in white rice grain among alleles of *Wx* SNP (**a**) and RM346 (**b**), respectively, identified by association mapping

compare the phenolic contents. Among all the non-pigmented rice, rice carrying the allele 1 had lower phenolic content than those with allele 2 and 4. Similarly, among the subpopulation POP3, POP5 and POP6, the phenolic content of rice with allele 1 was lower than that of allele 2 and 4 (Fig. 3b). Allele 3 was only present in the POP5, but rice with this allele had lowest phenolic content (Fig. 3b).

Discussion

There is a growing interest in applying association mapping to a wide range of plants to identify genes/QTLs responsible for quantitative variation of complex traits with agricultural and evolutionary importance (Ersoz et al. 2009; Hall et al. 2010; Stich and Melchinger 2010; Zhu et al. 2008). Understanding of the genetic bases underlying the naturally occurring genetic diversity in the nutraceutical properties and health benefit of whole grain rice among diverse rice accessions could help accelerate the breeding process. The present study is apparently the first to dissect the QTLs underlying the grain appearance traits and some health benefit traits by association mapping. Among other crops, only identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping was reported recently (Rezaeizad et al. 2010).

It is widely accepted that differential relatedness among subgroups, or population structure, can inflate the number of spurious marker–trait associations identified through association mapping (Pritchard et al. 2000; Stich et al. 2008; Yu et al. 2006; Zhao et al. 2007). When using the Q+K model to control the population structure and genetic relatedness, better performance than the model only controlling population structure (Q) has been well demonstrated (Yang et al. 2010; Yu et al. 2006; Zhao et al. 2007). However, the degree of the effects of Q and K for controlling false associations was different for different rice panels. In the whole rice panel, Q+K model performed similarly to the Q model, with only slightly fewer significant markers identified than Q model (Table 2). This may be derived from the weak relationship among this rice panels (Fig. 2). Whereas in the white rice panel, the Q+K model performed well for traits with 60% of significant markers reduced at different *P* value (Table 2). Yang et al. (2010) indicated that the effects of Q and K for controlling false associations also varied for different traits. In this study, it was found that Q+K model had similar effect on different traits. It is possible that all the traits in this study were correlated with one another to some extent, whereas in the study of Yang et al. (2010) more diverse traits were used. Multiple testing corrections such as permutation test for the Q model and qDFR test for the Q+K model sharply reduced the number of markers associated with the traits (Table 2), thus reducing the false-positive markers effectively. Comparing the number of markers identified after multiple testing corrections, the number of significant markers was similar to those at $P < 0.001$ in both rice panels without conducting the multiple testing correction. It is implied that more strict *P* values should be used for association mapping with this rice panel even though the Q+K model has been used to control the population structure and genetic relatedness.

The gene markers for *Rc* (brown pericarp and seed coat) for red rice pigmentation and *Ra* for black rice pigmentation were genotyped. Among the 50 red rice accessions, we found two rice were homozygous *rc* produced by a 14-bp deletion. Brooks et al. (2008) found that an SNP mutation within the *rc* restored the reading frame of the gene and reverted the bran layer pigmentation to red (wild type). Whether the two red rice accessions have the same point mutation as observed by Brooks et al. (2008) needs to be verified in future research. The association mapping showed that *Rc* and *Ra* gene markers were strongly associated with grain color traits (Table 3). The results are as expected since pigmented rice had pericarp color distinct from white rice, and they also had higher antioxidant activities than brown rice due to having more phenolics and flavonoids in the grain (Goffman and Bergman 2004; Shen et al. 2009). In addition to the major gene loci, other markers with minor effects have also been identified. Among them, four loci, i.e., RM228, RM339, *fgr* and RM316 are of great importance, because they are responsible for at least three color parameters simultaneously. Most of them explained a small part of the phenotypic variation, but RM316 could explain 21% of variation of H^o, indicating it might be a putative major QTL. The six loci, *Rc*, *Ra*, RM228, RM339, *fgr* and RM316, are also responsible for the genetic variation of phenolic content, flavonoid content and antioxidant capacity, which could be deduced as these three traits are correlated with the color parameters (Shen et al. 2009).

Due to the effect of major genes, the results of the association mapping with the whole rice panel (Table 3) are very different from those with white non-pigmented rice panel (Table 4). Only RM224 was identified in both panels for different traits. However, on comparing the markers indentified with Q model in both rice panels, more markers were found in common, such as RM228, RM297, RM339, RM224, RM507 and RM171 (Supplementary Tables 2; 3). The results of association mapping in white rice accessions showed that these color traits were controlled by polygenes with minor effects. The color of white rice pericarp may be influenced by the starch content. For example, glutinous rice is white opaque, and the pericarps of *indica* and *japonica* rice apparently have different color. *Wx SSR* was identified to be significantly associated with

three color traits. From QTL mapping, Tan et al. (2001) have mapped the *Wx* locus for color parameters of milled rice flour with a population derived from the cross between two white *indica* rice. Jin et al. (2009) also mapped the color parameters near the *Wx* locus for color parameters of brown rice grain using a population derived from *indica* and *japonica* cross. Association mapping of nutritional quality in white rice accessions showed that only one marker (RM346) was significantly associated with phenolic content (Table 4), though more markers were identified as potentially associated with nutritional quality and antioxidant capacity by Q model (Supplementary Table 3). The results of the association mapping are quite different from those of linkage mapping (Jin et al. 2009), but they still had few possible common loci. For example, *Wx* locus on the short arm of chromosome 6 was identified by both studies for *L**, and the genomic region close to the *SBE1* locus on the long arm of chromosome 6 was identified by both studies for *b** (Supplementary Table 3).

A better understanding of the factors that contribute to the grain micronutrient quality of rice will lay the foundation for developing new breeding and selection strategies for high quality. The putative genomic regions associated with phenolic content, flavonoid content and antioxidant capacity identified in this study could be used for marker-assisted selection (MAS) in rice breeding programs, thus making conventional breeding faster and more efficient. Comparison of the allelic effect among different alleles at the same locus would find the superior or inferior alleles that could be used or not used in MAS. For example, the allele 1 of RM346 would be an inferior allele, because it seemed to lower the phenolic content as compared to other alleles (Fig. 3b). Validation or fine mapping for marker–trait association is necessary because the marker alleles are correlated with, but not entirely predictive of, the gene alleles. Co-segregation test of the marker–trait association could be conducted in a segregating population (i.e., F_2) derived from parents with different target traits and different marker alleles. Once a marker–trait association is confirmed, it can be used as a selection target for the trait indirectly. The indirect selection is important for the micronutrients because chemical analysis is time and resource consuming (Zhang et al. 2008). The results of this study also showed that different markers should be used in breeding of pigmented and non-pigmented rice, respectively.

In conclusion, we conducted association mapping for grain color traits, phenolic content, flavonoid content and antioxidant capacity traits for the purpose of aiding breeding to increase the nutrient density in the rice grain. In addition to many QTLs with minor effect, the candidate genes *Rc* and *Ra* were identified among 416 rice panel including red and black rice, and the *Wx* gene markers could be identified in the white rice panel. Results of the present study demonstrated that genome-wide association mapping in rice could complement and enhance the information from linkage-based QTL studies toward the implementation of marker-assisted selection.

Acknowledgments Financial support for this work was provided in part by the Zhejiang Provincial Natural Science Foundation of China (R3080016), the Science and Technology Department of Zhejiang Province (2010C32082), the Program for the New Century Excellent Talents in University (NCET-08-0484) and the ANTA project from the Chinese Ministry of Agriculture (200803034).

References

Abdel-Aal ESM, Young JC, Rabalski I (2006) Anthocyanin composition in black, blue, pink, purple, and red cereal grains. J Agric Food Chem 54:4696–4704

Agrama HA, Eizenga GC (2008) Molecular diversity and genome-wide linkage disequilibrium patterns in a worldwide collection of *Oryza sativa* and its wild relatives. Euphytica 160:339–355

Agrama HA, Yan W (2009) Association mapping of straighthead disorder induced by arsenic in *Oryza sativa*. Plant Breed 128:551–558

Agrama HA, Eizenga GC, Yan W (2007) Association mapping of yield and its components in rice cultivars. Mol Breed 19:341–356

Bao JS, Corke H, Sun M (2006a) Microsatellites single nucleotide polymorphisms and a sequence tagged site in starch-synthesizing genes in relation to starch physicochemical properties in nonwaxy rice (*Oryza sativa* L.). Theor Appl Genet 113:1185–1196

Bao JS, Corke H, Sun M (2006b) Nucleotide diversity in starch synthase IIa and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.). Theor Appl Genet 113:1171–1183

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

Bouis HE, Chassy BM, Ochanda JO (2003) Genetically modified food crops and their contribution to human nutrition and food quality. Trend Food Sci Tech 14:191–209

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

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177

Brooks SA, Yan W, Jackson AK, Deren CW (2008) A natural mutation in rc reverts white-rice-pericarp to red and results in a new, dominant, wild-type allele: Rc-g. Theor Appl Genet 117:575–580

Doyle JJ (1991) DNA protocols for plants-CTAB total DNA isolation. In: Hewitt GM (ed) Molecular techniques in taxonomy. Springer, Berlin pp 283–293

Ersoz ES, Yu J, Buckler ES (2009) Applications of linkage disequilibrium and association mapping in maize. In: Kriz A, Larkins B (eds) Molecular genetic approaches to maize improvement. Springer, Berlin, p173–195

Finocchiaro F, Ferrari B, Gianinetti A, Dall'asta C, Galaverna G, Scazzina F, Pellegrini N (2007) Characterization of antioxidant

compounds of red and white rice and changes in total antioxidant capacity during processing. Mol Nutr Food Res 51:1006–1019
Fitzgerald MA, McCouch SR, Hall RD (2009) Not just a grain of rice: the quest for quality. Trend Plant Sci 14:133–139
Furukawa T, Maekawa M, Oki T, Suda I, Iida S, Shimada H, Takamure I, Kadowaki KI (2007) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. Plant J 49:91–102
Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch SR (2005) Genetic structure and diversity in *Oryza sativa* L. Genetics 169:1631–1638
Goffman FD, Bergman CJ (2004) Rice kernel phenolic content and its relationship with antiradical efficiency. J Sci Food Agric 10:1002–1007
Hall D, Tegstrom C, Ingvarsson PK (2010) Using association mapping to dissect the genetic basis of complex traits in plants. Brief Funct Genom 9:157–165
Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620
Hu C, Zawistowski J, Ling W, Kitts DD (2003) Black rice (*Oryza sativa* L. indica) pigmented fraction suppresses both reactive oxygen species and nitric oxide in chemical and biological model systems. J Agric Food Chem 51:5271–5277
Iwata H, Ebana K, Uga Y, Hayashi T, Jannink J-L (2010) Genome-wide association study of grain shape variation among *Oryza sativa* L. germplasms based on elliptic Fourier analysis. Mol Breed 25:203–215
Jin L, Xiao P, Lu Y, Shao YF, Shen Y, Bao JS (2009) Quantitative trait loci for brown rice color, total phenolics and, flavonoid contents and antioxidant capacity in rice grain. Cereal Chem 86:609–615
Jin L, Lu Y, Xiao P, Sun M, Corke H, Bao JS (2010a) Genetic diversity and population structure of a diverse set of rice germplasm for association mapping. Theor Appl Genet 121:475–487
Jin L, Lu Y, Shao YF, Zhang G, Xiao P, Shen SQ, Corke H, Bao JS (2010b) Molecular marker assisted selection for improvement of the eating, cooking and sensory quality of rice (*Oryza sativa* L.). J Cereal Sci 51:159–164
Ling WH, Cheng QX, Ma J, Wang T (2001) Red and black rice decrease atherosclerotic plaque formation and increase antioxidant status in rabbits. J Nutr 131:1421–1426
Liu RH (2007) Whole grain phytochemicals and health. J Cereal Sci 46:207–219
Lorieux M, Petrov M, Huang N, Guiderdoni E, Ghesquiere A (1996) Aroma in rice: genetic analysis of a quantitative trait. Theor Appl Genet 93:1145–1151
Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD (2007) The extent of linkage disequilibrium in rice (*Oryza sativa* L.). Genetics 177:2223–2232
Oki T, Masuda M, Kobayashi M, Nishiba Y, Furuta S, Suda I, Sato T (2002) Polymeric procyanidins as radical-scavenging components in red-hulled rice. J Agric Food Chem 50:7524–7529
Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan D (2006) Selection under domestication: evidence for a sweep in the rice *Waxy* genomic region. Genetics 173:975–983
Ordonez SA Jr, Silva J, Oard JH (2010) Association mapping of grain quality and flowering time in elite japonica rice germplasm. J Cereal Sci 51:337–343
Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y, Ito A, Ishii T, Miyashita NT, Terauchi R (2007) Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. Theor Appl Genet 114:731–743
Reddy VS, Dash S, Reddy AR (1995) Anthocyanin pathway in rice (*Oryza sativa* L.): identification of a mutant showing dominant inhibition of anthocyanins in leaf and accumulation of proanthocyanidins in pericarp. Theor Appl Genet 91:301–312
Rezaeizad A, Wittkop B, Snowdon R, Hasan M, Mohammadi V, Zali A, Friedt W (2010) Identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping using SSR markers. Euphytica. doi:10.1007/s10681-010-0231-y
Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. Genet Res 67:175–185
Seal CJ (2006) Whole grains and CVD risk. Proc Nutr Soc 65:24–34
Shen Y, Jin L, Xiao P, Lu Y, Bao JS (2009) Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. J Cereal Sci 49:106–111
Stich B, Melchinger AE (2010) An introduction to association mapping in plants. CAB Rev 5:039
Stich B, Mohring J, Piepho H, Heckenberger M, Buckler ES, Melchinger AE (2008) Comparison of mixed-model approaches for association mapping. Genetics 178:1745–1754
Storey JD, Tibshirani R (2003) Statistical significance for genome-wide studies. Proc Natl Acad Sci USA 100:9440–9445
Sweeney MT, Thomson MJ, Pfeil BE, McCouch SR (2006) Caught red-handed: Rc encodes a basic helix–loop–helix protein conditioning red pericarp in rice. Plant Cell 18:283–294
Takaiwa F, Yang L, Yasuda H (2008) Health-promoting transgenic rice: application of rice seeds as a direct delivery system for bioactive peptides in human health. In: Hirano HY, Sano Y, Hirai A, Sasaki T (eds) Rice biology in the genomics era. Springer, Berlin pp 357–373
Tan YF, Sun M, Xing YZ, Hua JP, Xun XL, Zhang QF, 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. Theor Appl Genet 103:1037–1045
Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, ES Buckler IV (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat Genet 28:286–289
Tian S, Nakamura K, Kayara H (2004) Analysis of phenolic compounds in white rice, brown rice and germinated brown rice. J Agric Food Chem 52:4808–4813
Toyokuni S, Itani T, Morimitsu Y, Okada K, Ozeki M, Kondo S, Uchida K, Osawa T, Hiai H, Tashiro T (2002) Protective effect of coloured rice over white rice on Fenton reaction-based renal lipid peroxidation in rats. Free Radic Res35:583–592
Vitaglione P, Napolitano A, Fogliano V (2008) Cereal dietary fibre: a natural functional ingredient to deliver phenolic compounds into the gut. Trend Food Sci Tech 19:451–463
Wang CX, Shu QY (2007) Fine mapping and candidate gene analysis of purple pericarp gene Pb in rice (*Oryza sativa* L.). Chin Sci Bull 52:3097–3104
Welch RM, Graham R (2004) Breeding for micronutrients in staple food crops from a human nutrition perspective. J Exp Bot 55:353–364
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
Xia M, Ling WH, Ma J, Kitts DD, Zawistowski J (2003) Supplementation of diets with the black rice pigment fraction attenuates atherosclerotic plaque formation in apolipoprotein E deficient mice. J Nutr 133:744–751

Yang XH, Yan JB, Shah T, Warburton ML, Li Q, Li L, Gao YF, Chai YC, Fu ZY, Zhou Y, Xu ST, Bai GH, Meng YJ, Zheng YP, Li JS (2010) Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. Theor Appl Genet 121:417–431

Yawadio R, Tanimori S, Morita N (2007) Identification of phenolic compounds isolated from pigmented rices and their aldose reductase inhibitory activities. Food Chem 101:1644–1653

Ye X, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrokus I (2000) Engineering the pro-vitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287:303–305

Yoshimura A, Ideta O, Iwata N (1997) Linkage map of phenotype and RFLP markers in rice. Plant Mol Biol 35:49–60

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

Zhang CY, Shen Y, Chen J, Xiao P, Bao JS (2008) Nondestructive prediction of total phenolics and flavonoid contents, and antioxidant capacity of rice grain using near-infrared spectroscopy. J Agric Food Chem 56:8268–8272

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M (2007) An Arabidopsis example of association mapping in structured samples. PLoS Genet 3:e4

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20

Zhu F, Cai YZ, Bao JS, Corke H (2010) Effect of γ-irradiation on phenolic compounds in rice grain. Food Chem 120:74–77

Zimmermann MB, Hurrell RF (2002) Improving iron, zinc and vitamin A nutrition through plant biotechnology. Curr Opin Biotech 13:142–145