Unraveling the Complex Trait of Harvest Index with Association Mapping in Rice (*Oryza sativa* L.)

Xiaobai Li1,3, Wengui Yan2*, Hesham Agrama4, Limeng Jia1,2,4, Aaron Jackson2, Karen Moldenhauer4, Kathleen Yeater5, Anna McClung2, Dianxing Wu1*

1 State Key Lab of Rice Biology, International Atomic Energy Agency Collaborating Center, Zhejiang University, Hangzhou, People’s Republic of China, 2 Agricultural Research Service, United States Department of Agriculture, Dale Bumpers National Rice Research Center, Stuttgart, Arkansas, United States of America, 3 Institute of Horticulture, Zhejiang Academy of Agricultural Sciences, Hangzhou, People’s Republic of China, 4 University of Arkansas, Rice Research and Extension Center, Stuttgart, Arkansas, United States of America, 5 Agricultural Research Service, United States Department of Agriculture, Southern Plains Area, College Station, Texas, United States of America

Abstract

Harvest index is a measure of success in partitioning assimilated photosynthate. An improvement of harvest index means an increase in the economic portion of the plant. Our objective was to identify genetic markers associated with harvest index traits using 203 *O. sativa* accessions. The phenotyping for 14 traits was conducted in both temperate (Arkansas) and subtropical (Texas) climates and the genotyping used 154 SSRs and an indel marker. Heading, plant height and weight, and panicle length had negative correlations, while seed set and grain weight/panicle had positive correlations with harvest index across both locations. Subsequent genetic diversity and population structure analyses identified five groups in this collection, which corresponded to their geographic origins. Model comparisons revealed that different dimensions of principal components analysis (PCA) affected harvest index traits for mapping accuracy, and kinship did not help. In total, 36 markers in Arkansas and 28 markers in Texas were identified to be significantly associated with harvest index traits. Seven and two markers were consistently associated with two or more harvest index correlated traits in Arkansas and Texas, respectively. Additionally, four markers were constitutively identified at both locations, while 32 and 24 markers were identified specifically in Arkansas and Texas, respectively. Allelic analysis of four constitutive markers demonstrated that allele 253 bp of RM431 had significantly greater effect on decreasing plant height, and 390 bp of RM24011 had the greatest effect on decreasing panicle length across both locations. Many of these identified markers are located either nearby or flanking the regions where the QTLs for harvest index have been reported. Thus, the results from this association mapping study complement and enrich the information from linkage-based QTL studies and will be the basis for improving harvest index directly and indirectly in rice.

Introduction

In food production, optimizing grain yield, reducing production costs, and minimizing risks to the environment have been the primary objectives since the beginning of the twentieth century [1]. Food crops grow by developing a vegetative canopy that transpires water and carries out photosynthesis, and a root system that takes up water and nutrition, which leads to the production of biomass. Following the reproductive stage, a portion of the plant biomass is partitioned to various yield components and determines harvest index [2]. Harvest index is the ratio of grain yield to total biomass and is considered as a measure of biological success in partitioning assimilated photosynthate to the harvestable product [3,4,5]. In cereal crops, dramatic improvements in harvest index have made commercial cultivars greatly different from their wild ancestors [6]. Rice (*Oryza sativa* L.) is one of the most important staple foods [7]. It can be highly productive if high harvest index genotypes are grown with optimal management practices [2].

Harvest index of rice is the result of various integrated processes with an involvement of the number of panicles per unit area, the number of spikelets per panicle, the percentage of fully ripened grains, and the weight of 1,000 mature kernels [8]. Marri et al. [9] found that harvest index was negatively correlated with plant height, but positively correlated with grain number/panicle, grain number/plant, percentage spikelet fertility, test grain weight and yield/plant in rice. Sabouri et al. [10] verified the negative correlation of harvest index with plant height and positive correlation with spikelet number and grain weight per panicle, and reported the impact of some flag leaf characteristics on harvest index in rice. In maize, harvest index is negatively correlated with plant height, but positively correlated with grain yield/plant and grain number/plant, percentage spikelet fertility, test grain weight and yield/plant in rice. Sabouri et al. [10] verified the negative correlation of harvest index with plant height and positive correlation with spikelet number and grain weight per panicle, and reported the impact of some flag leaf characteristics on harvest index in rice. In sorghum, harvest index is negatively correlated with forage yield [12], but positively correlated with growth rate and grain filling rate [13]. Usually, the correlated traits are interrelated, so that increases in one component may lead to decreases or increases in others.
Therefore, scientists aim to identify genes/QTLs that directly improve a target trait without negatively affecting others, or improve the target trait indirectly through the improvement of its associated characteristics.

Crop harvest index is also highly influenced by environmental factors [14], such as soil condition [15,16] and temperature [17,10]. However, genetic control of harvest index plays important role in crop production. Large variation was observed for harvest index in rice: about 0.25% among wild species, 0.30% among tall cultivars and more than 0.40% for semi-dwarf cultivars [19]. The intrinsic regulation of harvest index is controlled by many genes. A few reports in the literature have examined QTLs in rice associated with harvest index. Mao et al. [20] reported four main-effect QTLs for harvest index on chromosome (Chr) 1, 4, 5 and 11 and other epistatic interaction between two QTLs respectively on Chr 1 and Chr 5. Sabouri et al. [10] identified three QTLs mapped on Chr 2, 3 and 5, and two QTLs close to each other on Chr 4. Lanceras et al. [21] described harvest index QTLs on Chr 1 and 3. However, a recurring complication of the QTL data showed that different parental combinations and/or experiments conducted in different environments often result in partly or wholly non-overlapping sets of QTLs [22]. Therefore, it is necessary to explore constitutive QTLs across different environments and adaptive QTLs specifically for a given environment [23].

Classical QTL mapping reveals only a portion of the genetic control of a trait because there are only two alleles that can differ at any locus between the two parental lines. More comprehensive analyses of genetic architecture require consideration of a larger sample of the genetic variation in the species. One approach is association mapping, which maps the QTLs either among extant breeding lines with known pedigree relationships or in a diverse germplasm collection. Given pedigree and marker information, the probability for different lines in complex populations to share identity by descent QTLs can be defined, permitting estimation of the effects of each QTL [24]. Association mapping provides an alternate route into identifying the QTLs that have effects across a broader spectrum of germplasm, if false-positives caused by population structure can be minimized [25]. Whole-genome association scans are expected to be effective when linkage disequilibrium (LD) and marker density are sufficiently high, so that the random markers could have a greater chance of being in linkage disequilibrium with QTLs across diverse genetic materials [26]. Huang et al. [27] successfully performed genome-wide association study (GWAS) in a rice landrace collection of China for 14 agronomic traits and identified a substantial number of loci at close to gene resolution.

Results

Markers profile

The set of 154 SSRs and an indel with genome-wide distribution detected a total of 1993 alleles among 203 O. sativa accessions. The average number of alleles per locus was 12.86 ranging from 2 for RM338 to 57 for con673. Polymorphic Information Content (PIC) varied from 0.25 for AP5625-1 to 0.97 for con673 among the 155 markers with an average of 0.71. Nei’s (1983) genetic distances ranged from 0.0181 to 0.9667 with an average 0.7464 among each pair of 203 accessions in the URMC.

Population structure and geographic origin

Using STRUCTURE software with multi-loci genotype data, a five-group model was identified to sufficiently explain genetic structure among 203 accessions. Ancestry of each of these accessions was inferred for assignment into a genetic group (Figure 1A). A dendrogram tree created with PowerMarker had five main branches for the 203 accessions as well (Figure 1B). The principal components analysis (PCA) also displayed the pattern of genetic structure with five groups. The first three components of PCA for 45.07% of total variation were used to visualize the five groups derived from ancestry analyses (Figure 1C).

The resultant five groups of O. sativa categorized by the Q value (ancestry index) belong to indica (IND), temperate japonica (TEJ), tropical japonica (TRJ), aus (AUS) and aromatic (ARO) (Figure 1A), based on reference cultivars reported previously by Garris et al. [37], Agrama and Eizenga [38] and Agrama et al. [34]. Each accession with ancestry information was plotted on a world map using its latitude and longitude of geographic origin (Figure 2). TEJ accessions were mainly distributed between latitudes 30 and 50 degrees north and south of the equator (i.e. temperate zone) while the other four groups scattered between latitude N 50 and S 30 degrees (i.e. tropical and subtropical zone).

Morphological analysis

Statistical analysis using a mixed model demonstrated that the differences due to genotypes and genotype x location interactions were highly significant at the 0.001 level of probability for all of the 14 traits (Table 1). The differences due to location were also significant for 12 traits except for panicle branches and seed set. Heritability was very high for all of these 14 traits. Heading had the highest heritability which was close to 100%. Although seed set had the lowest heritability, it was still above 70%. Heritability ranged from 77 to 97% among the other 12 traits. Harvest index had a heritability of 83% at Stuttgart and 90% at Beaumont. Correlation coefficients for each pair of the 14 traits were calculated using Spearman rank for each location and presented in Table S1A and S1B, respectively. To visualize the complex
relationship among the 14 traits, PCA was used to construct plots with the first two axes accounting for more than 50% phenotypic variation (Figure 3A, B). At Stuttgart, 47 out of 91 correlations among the 14 traits were significant (<0.0001) (Table S1A, Figure 3A), and 40 correlations were significant at Beaumont (Table S1B, Figure 3B). Thirty four correlations were uniformly significant across two locations and their correlation directions (positive or negative) were also same across two locations (Table S1A, S1B).

Six traits were significantly correlated with harvest index and these correlation directions were the same across the two locations. The correlations with harvest index were negative for heading (−0.46 at Stuttgart and −0.61 at Beaumont), plant height (−0.50 and −0.50), plant weight (−0.36 and −0.30), panicle length (−0.45 and −0.32), while positive for seed set (0.52 and 0.61) and grain weight/panicle (0.32 and 0.40) (Figure 3A, B). In the PCA based on phenotypic traits of 203 mini-core accessions, four traits negatively correlated with harvest index were plotted on opposing
### Table 1. Statistical analysis of 14 traits generated at Stuttgart, Arkansas and Beaumont, Texas in 2009 in the USDA rice mini-core collection.

| Trait                  | Location  | Mean ± SD | Range   | Heritability (%) | Genotype Location | Location | Genotype*Location |
|------------------------|-----------|-----------|---------|------------------|-------------------|----------|-------------------|
|                        |           |           |         |                  | F value  Pr>F     | F value  Pr>F | F value  Pr>F     |
| Heading (days)         | Stuttgart | 99.33 ±21.31 | 34.00–181.67 | 98.08  | 341.53 | 0.000000   | 2634.77  | 0.000001   | 12.45  | 0.000000   |
|                        | Beaumont  | 87.55 ±22.63 | 38.00–182.00 | 98.64  |          |           |          |           |       |           |
| Plant height (cm)      | Stuttgart | 109.73 ±20.20 | 61.08–153.92 | 97.11  | 127.48 | 0.000000   | 1676.50  | 0.000002   | 45.63  | 0.000000   |
|                        | Beaumont  | 124.74 ±22.45 | 67.00–178.78 | 95.73  |          |           |          |           |       |           |
| Plant weight (g)       | Stuttgart | 168.71 ±79.88 | 27.83–548.42 | 86.33  | 30.87  | 0.000000   | 122.48   | 0.000376   | 9.94   | 0.000000   |
|                        | Beaumont  | 219.02 ±87.70 | 35.93–558.02 | 86.83  |          |           |          |           |       |           |
| Tillers                | Stuttgart | 23.95 ±11.20  | 6.42–67.75   | 86.53  | 35.27  | 0.000000   | 818.76   | 0.000009   | 7.10   | 0.000000   |
|                        | Beaumont  | 41.13 ±15.83  | 13.00–85.89  | 87.16  |          |           |          |           |       |           |
| Grain yield (g)        | Stuttgart | 60.02 ±25.51  | 8.54–127.27  | 87.05  | 29.06  | 0.000000   | 98.37    | 0.000568   | 8.33   | 0.000000   |
|                        | Beaumont  | 76.67 ±30.05  | 5.64–165.97  | 84.03  |          |           |          |           |       |           |
| Harvest index (%)      | Stuttgart | 30.44 ±7.02   | 3.40–45.06   | 82.75  | 35.79  | 0.000000   | 2174.76  | 0.000000   | 6.10   | 0.000000   |
|                        | Beaumont  | 38.98 ±10.51  | 6.25–60.02   | 89.98  |          |           |          |           |       |           |
| Panicle length (cm)    | Stuttgart | 26.66 ±3.81   | 14.21–37.19  | 89.86  | 46.56  | 0.000000   | 293.26   | 0.000000   | 3.68   | 0.000000   |
|                        | Beaumont  | 24.75 ±3.44   | 16.84–38.40  | 88.34  |          |           |          |           |       |           |
| Panicle branches       | Stuttgart | 10.97 ±2.15   | 5.44–17.78   | 85.65  | 29.97  | 0.000000   | 31.18    | 0.004559   | 2.40   | 0.000000   |
|                        | Beaumont  | 10.64 ±2.06   | 5.63–16.33   | 81.68  |          |           |          |           |       |           |
| Kernels/panicle        | Stuttgart | 194.97 ±57.49 | 68.56–399.00 | 86.48  | 29.94  | 0.000000   | 367.90   | 0.000459   | 4.45   | 0.000000   |
|                        | Beaumont  | 155.77 ±45.46 | 50.00–318.33 | 86.92  |          |           |          |           |       |           |
| Seed set (%)           | Stuttgart | 78.15 ±15.23  | 25.48–96.97  | 78.39  | 15.39  | 0.000000   | 14.26    | 0.019138   | 4.39   | 0.000000   |
|                        | Beaumont  | 73.55 ±12.65  | 35.07–95.29  | 72.66  |          |           |          |           |       |           |
| 1000 Seed weight (g)   | Stuttgart | 25.77 ±5.07   | 11.17–44.74  | 91.79  | 69.00  | 0.000000   | 75.18    | 0.000477   | 3.94   | 0.000000   |
|                        | Beaumont  | 24.41 ±4.66   | 12.32–43.86  | 95.52  |          |           |          |           |       |           |
| Kernels/cm panicle     | Stuttgart | 7.30 ±1.80    | 3.25–14.61   | 84.71  | 28.72  | 0.000000   | 218.17   | 0.000104   | 3.60   | 0.000000   |
|                        | Beaumont  | 6.31 ±1.63    | 2.80–12.27   | 87.02  |          |           |          |           |       |           |
| Kernels/branch panicle | Stuttgart | 17.88 ±4.24   | 11.56–37.10  | 82.66  | 19.90  | 0.000000   | 353.27   | 0.000058   | 4.31   | 0.000000   |
|                        | Beaumont  | 14.67 ±2.98   | 9.61–23.23   | 77.42  |          |           |          |           |       |           |
| Grain weight/panicle   | Stuttgart | 3.79 ±1.18    | 0.68–8.62    | 82.29  | 21.86  | 0.000000   | 241.69   | 0.000075   | 3.94   | 0.000000   |
|                        | Beaumont  | 2.75 ±0.95    | 0.63–6.27    | 80.72  |          |           |          |           |       |           |

doi:10.1371/journal.pone.0029350.t001

**Figure 3. Relationship map constructed by PCA for 14 traits at A: Stuttgart, AR and B: Beaumont, TX.** The distance between traits is inversely proportional to the size of the correlation coefficients. Solid and dashed lines indicate positive and negative correlations, respectively. Trait names are T1: Heading; T2: Plant height; T3: Plant weight; T4: Tillers; T5: Grain yield; T6: Harvest index; T7: Panicle length; T8: Panicle branches; T9: Kernels/panicle; T10: Seed set; T11: 1000 Seed weight; T12: Kernels/cm panicle; T13: Kernels/branch panicle; T14: Weight/panicle. The variation explained by the principal components is showed in the brackets.

doi:10.1371/journal.pone.0029350.g003
axis from harvest index (Figure 3A,B). Conversely, two traits positively correlated with harvest index were plotted in the same axis relatively close to harvest index.

Model comparison and marker-trait associations

Dimension determination for PCA indicated that different dimensions should be included for testing associations for these traits. Further, relative performance of the association mapping models was also evaluated based on the criterion BIC (Table S2). The smaller BIC indicated the better model fit [25]. Among all possible models (naive, kinship, PCA, Q, PCA+kinship and Q+kinship), naive and kinship models showed the highest BIC value. The four other models (PCA, Q, P+kinship and Q+kinship) had a better performance, indicated by smaller BIC values. The model installed with kinship had a slightly higher BIC than the one without kinship. The PCA models containing different dimensions for different traits had the lowest BIC value. Thus, the PCA model was selected to conduct association mapping for harvest index traits.

At Stuttgart, a total of 36 markers were identified to be significantly associated with harvest index traits at the 6.45 \times 10^{-3} level of probability (the Bonferroni corrected significance level) (Table S3). Among 36 markers, seven were associated with harvest index per se, five with heading, three with plant height, six with plant weight, five with panicle length, nine with seed set and one with grain weight/panicle. Eight of these trait-marker associations have been reported previously (Table S3). Additionally, seven markers were consistently associated with two or more harvest index traits [39]. Out of the seven consistent markers, RM600, RM5 and RM302 were co-associated with harvest index and seed set, RM431 with heading and seed set, RM341 with plant height and panicle length, RM471 with heading and plant weight, and RM510 with three traits, plant height, harvest index and seed set.

At Beaumont, we identified 28 markers significantly associated with harvest index traits (Table S3). Among these, two were associated with harvest index, three with heading, nine with plant height, six with plant weight, four with panicle length, three with seed set and one with grain weight/panicle. At Stuttgart, eight of the trait-marker associations have been identified in previous QTL studies. Two consistent markers were RM208 co-associated with harvest index and seed set, and RM55 co-associated with plant height and plant weight.

Across two locations, the associations of RM431 with plant height, Rid12 and RM471 with plant weight, and RM24011 with panicle length were consistently true. The four markers that associated with the same trait across both locations are called “constitutive QTL” markers, while others that associated with a certain trait only at one location are called “adaptive QTL” markers [23].

Allelic effects

The allelic effects of the constitutive markers associated with their traits were estimated using the least square mean (LSMEAN) of phenotypic values and are presented in Figure 4 and Table S4. For RM431, allele 253 bp had a significantly larger effect than all other 6 alleles at Beaumont and than 4 others at Stuttgart to reduce plant height. For RM24011, allele 390 bp had the greatest effect on decreasing panicle length while allele 411 bp had the largest effect on increasing panicle length at both locations. However, for Rid12, the allelic effects were opposite between two locations.

Figure 4. Comparisons of allelic effects of four constitutive marker loci. A: RM431 associated with plant height, B: RM471 and C: Rid12 associated with plant weight, D: RM24011 associated with panicle length constitutively at both Stuttgart, Arkansas and Beaumont, Texas.

doi:10.1371/journal.pone.0029350.g004
locations. Allele 151 bp of Rid12 had a decreasing effect on plant weight at Stuttgart, but an increasing effect at Beaumont instead. The 165 allele of Rid12 had an opposite effect to 151 bp on plant weight. For RM471, the allelic effects on plant weight were not consistent from one location to another. The 109 bp allele was associated with one of the lowest means for plant weight at Stuttgart, but one of the largest means for plant weight at Beaumont.

Discussion

Genetic diversity and genetic structure

The average number of alleles per locus was 12.86 among 203 accessions in the URMC genotyped with 155 markers. The allele number per locus is the highest among the rice collections that have been reported to date [37,40], including an Indian germplasm collection [41], an Indonesian landrace collection [42] and a Brazilian rice core collection [43], with an exception of an Indonesian traditional and improved rice collection with 13 alleles per locus reported by Thomson et al. [44]. The average polymorphic information content (PIC) value in this study was 0.71, which is also the highest among previous studies for rice populations [37,40,41,42,44,45] with an exception of 0.75 PIC value in a study reported by Borba et al. [43]. The wide range of genetic diversity along with the manageable number of accessions in the URMC makes it one of the best collections for mining valuable genes in rice.

Population structure is an important component in association mapping analyses because it can be a source of Type I error in an autogamous species such as barley and rice [46,47,48]. In this study, the 203 O. sativa accessions in the URMC were divided into five model-based groups from ancestry analysis (Figure 1A). Both the dendrogram tree (Figure 1B) and the PCA analysis (Figure 1C) reached similar conclusions regarding population structure in this collection. The results obtained from these three separate analyses supported each other. The classification agreed with the previous study [33] except for the group of wild relatives of rice having a high rate of rare alleles. The high rate of rare alleles was suggested by its high percentage of private alleles and the small size of the group [33]. The wild rice accessions were not integrated into association mapping since low frequency alleles are known to inflate variance estimates of linkage disequilibrium and produce a greater chance of Type I error [46,47,49]. In addition, the population structure was observed to be tied with geographic origins, e.g. TEJ mainly distributed in the temperate zone (Figure 2) and wild rice relatives were from a relatively isolated area (data not shown). The distinctive geographic origins corresponding to the difference of ecological environments could be partially responsible for the genetic differentiation, which in turn contributes to the different responses to environmental factors and rare alleles in the germplasm accessions of wild relative species.

Morphological environment-sensitivity and trait-trait correlation

All 14 traits were significantly affected by environment and environment X genotype interaction, which suggested genotypic sensitivities to differences in environmental conditions at the two locations (Table 1). The sensitivity of panicle heading to temperature change and the variation of harvest index in response to photoperiod were previously observed in rice [50]. Others have reported that rice accessions derived from different geographic regions react to environmental signals differently as well [51,52]. Information on germplasm and environmental interaction is helpful for parental selection for a specific or broad adaptation to environments.

The correlations among the 14 traits exhibited a complex relationship between pairs of traits. At both locations, the harvest index increased with an increase of seed set and grain weight/panicle, while decreased with an increase of heading, panicle length, plant height and plant weight. The negative and significant correlation between heading and harvest index was also reported in spring wheat [53], rice [54] and sorghum [11]. These studies concluded that harvest index could be easily influenced not only during the grain filling period [55,56], but also during the period from panicle initiation to heading [54] as affected by planting dates and temperature during the growing season [57]. Plant height is another important agronomic trait that is directly linked to harvest index [9,58]. Yoshida et al. [15] also reported a similar result to this study where harvest index was inversely correlated with plant height, which may be due to lodging in the tall varieties [54], or greater translocation of photosynthate from the vegetative tissues to grain in semi-dwarf varieties [59]. The positive correlation between harvest index and grain weight/panicle was also reported by Sabouri et al. [10]. However, panicle length was not found to be correlated with harvest index in Marri’s study [9]. Similarly, plant weight was not correlated with harvest index in Sabouri’s study [10]. These different results are understandable since different materials were used in those studies. In practice, highly correlated traits, such as heading, can be used to obtain indirect estimates of harvest index when direct estimates are difficult or impractical to obtain. Thus improvement of harvest index can be manipulated indirectly. In theory, the correlation of harvest index with its related traits determined in this study, indicates an interrelationship of physiological pathways controlling these traits.

Model comparison for association mapping of harvest index’s traits

For harvest index traits, the number of dimensions in PCA was tested for each trait, and the appropriate number of dimensions was determined on the basis of BIC. Our simulated experiments showed that the dimension of PCA can exhibit phenotypic specificity. As an example with heading, the PCA model required a higher dimension number to capture the true population structure effects. Traditionally, the number of dimensions has been generally determined on the basis of random marker information without considering phenotypic information. However, the effects of population structure on different complex traits vary dramatically [60,61] and it is logical to hypothesize that the numbers of dimensions required for cofactors in detecting marker-trait association are not necessarily the same [62].

Comparing with other five models (naive, kinship, PCA+Kinship, Q and Q+Kinship model), the PCA showed the best fit with the smallest BIC value for harvest index traits. Interestingly, correction of the kinship model was not observed to be better than the naive model. Similarly, the models with Q+kinship or PCA+kinship did not perform better than the ones with only Q or PCA, either. Shao et al. [63] also found that Q+kinship model performed similarly to the Q model alone in a rice panel. The result did not agree with some other studies on cross-pollinated plants and humans [25,62], where the relatedness among accessions in a population is quite complex because of the mating style. The low complex relatedness in the URMC rice collection could be attributable to the restricted gene flow among these self-pollinated accessions and the diverse global origination of these accessions. Moreover, the low complex relatedness may be a result of the M strategy based on 26 phenotypic traits and 70 molecular
Environmental sensitivity and marker-assisted selection

Genetic dissection of harvest index

Harvest index is an integrative trait including the net effect of all physiological processes during the crop cycle and its phenotypic expression is generally affected by genes responsible for non-target traits, such as heading [20,65], plant height [20] and panicle architecture [66]. The magnitude and direction of these gene functions on different phenotypes would bear heavily on the utility of such genes for improvement of these traits. In the current study, the traits like heading, plant height, plant weight and panicle length had a strong negative correlation with harvest index, while seed set and grain weight/panicle were positively correlated with harvest index. These phenotypic correlations were consistently reflected in the identification of molecular markers associated with harvest index and related traits. For example, four consistent markers at Stuttgart, RM600, RM302, RM25, and RM431, were associated with not only harvest index itself, but also for one or more traits consistently correlated with harvest index. Another consistent marker, Rid12, associated with both heading and plant weight, was close to a reported QTL “qHID7-1” responsible for harvest index [67] and the gene “Ghd7” having major effects on grains per panicle, plant height and heading in rice [68]. At Beaumont, the consistent marker RM55, associated with both plant height and plant weight, was adjacent to a QTL “qHID7-2” for control of harvest index [67]. RM431 co-associated with plant height and harvest index in this study has been reported to be closely linked to gene “sd1” [69,70]. The sd1 that is involved in gibberellic acid biosynthesis decreases plant height, thus increases harvest index. The decreased height reduces lodging susceptibility, is tolerant to heavy applications of nitrogen fertilizer, and can be planted at relatively high density, all contributing to improved grain yield that has resulted in the Green Revolution in cereal crops including rice [71].

Other markers were associated with the traits correlated with harvest index, but not with harvest index directly in this study. These markers have been reported either nearby or flanking the QTLs for harvest index. RM5, which was associated with plant height in the Stuttgart study, was close to a reported QTL for harvest index on Chr 1 [9]. RM471 associated with plant weight was close to the reported qHID4-1 and qHID4-2 for harvest index [67]. Furthermore, RM257 and RM22559 associated with seed set were co-localized with a known QTL on Chr 9 [9], and with qHID8-1 [67] for harvest index, respectively. Similarly, at Beaumont, RM44 associated with plant height was close to qHID8-1 [67], and RM263 associated with heading was adjacent to hid2-1 [9]. The chromosomal regions where numerous correlated traits are mapped indicate either pleiotropy of a single gene or tight linkage of multiple genes. Fine-mapping of such chromosomal regions would help discern the actual genetic control of these congruent traits. Development of markers for such traits in specific regions could lead to a highly effective strategy of marker-assisted selection for improving harvest index.

Environmental sensitivity and marker-assisted selection

Quantitative traits show a range of sensitivities to environmental changes [67]. In this study, 32 marker-trait associations were identified specifically adaptive to Stuttgart, whereas 24 marker-trait associations were adaptive to Beaumont. More importantly, we identified four constitutive markers associated with harvest index traits in both environments.

Environment-specific QTLs can be used for marker-assisted selection (MAS) at specific environments. For example, RM431 could be used to improve harvest index directly and indirectly through decreasing plant height and increasing seed set in Arkansas because it was co-associated with harvest index, plant height, and seed set. The constitutive marker-trait associations over multiple environments can be applied to MAS programs in a wide area. For example, results suggest that the constitutive markers Rid12 and RM471 could be used to improve harvest index indirectly through decreasing plant weight in the southern states of the USA.

Comparison of allelic effects of these constitutive markers can classify the alleles within a marker locus into superior or inferior ones, which helps decide which to use for MAS in the southern states. For example, allele 253 bp of RM431 and allele 390 bp of RM24011 had the largest effect on decreasing two traits, plant height and panicle length, negatively associated with harvest index. Thus, these superior alleles can be introduced for improvement of harvest index indirectly through decreasing the negative traits at both locations. Conversely, the allele 411 bp of RM24011 had the largest effect on increasing the panicle length and thus would not be useful for improving harvest index using MAS at either location. Interestingly, the two alleles of Rid12 associated with plant weight had opposite effects at the two locations. Allelic choice for this marker should be dependent on the particular environment targeted for breeding.

Results of the present study demonstrated that genome-wide association mapping in the URMRC could complement and enrich the information derived from linkage-based QTL studies. After validation or fine mapping of these putative genomic regions, the information will help secure food production through either direct improvement of harvest index or indirect improvement via changes in seed set, grain weight per panicle, heading, plant height and weight, and panicle length using the MAS.

Materials and Methods

Rice association panel

Of 217 accessions in the URMRC, 203 belong to O. sativa whereas the remaining belongs to other species in O. glaberrima. Pure seed of these accessions were provided by the Genetic Stock Oryza Collection (GSOR) (www.ars.usda.gov/spa/dbnrrc/gsor) with cultivar name or designation, accession number, registration year, place of origin, longitude and latitude of origin, pedigree or genetic background (if available), morphological characteristics and references. The GSOR supplies seeds for research purposes to national and international users upon to request. In this study, only 203 O. sativa accessions were used for the following analysis because the wild relatives, O. glaberrima, nivara, rufipogon, glumaepatula and latifolia, contain many rare alleles. Rare alleles are one of the factors that increase the risk of Type I errors or spurious associations [47].

Location and field experiment

Evaluations were conducted for 14 traits in two field locations, USDA-ARS Dale Bumpers National Rice Research Center near Stuttgart, Arkansas and USDA-ARS Rice Research Unit near Beaumont, Texas during the 2009 growing season. The Stuttgart test site is located at N 34°27′44″ and W 91°24′59″, representing a temperate climate with a 243 d frost free period and average
temperature of 23.9°C during the growing season. The Beaumont test site is located at N 30°03'47" and W 94°17'45", representing a subtropical climate with a 253 d frost free period and an average temperature of 26.1°C during the growing season. The experiments at both locations utilized a randomized complete block design having three replications with nine plants spaced 0.3×0.6 m in each plot. Three seeds were sown in each of nine hills in a plot using a Hege 1000 grain drill planter on April 23 and May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was a plot using a Hege 1000 grain drill planter on April 23 and May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively. Each hill was thinned to a single plant right after the permanent flood was achieved on May 6 of 2009 at Stuttgart and Beaumont, respectively.
were compared using the LSMEANS and pdiff option in the MIXED procedure, using Saxton’s PDMIX000 SAS macro [53].

Supporting Information

Table S1 A. Spearman correlation for each pair of 14 traits evaluated at Stuttgart, Arkansas in 2009. B. Spearman correlation for each pair of 14 traits evaluated at Beaumont, Texas in 2009. (DOC)

Table S2 Fitness analysis of mapping model for harvest index traits using Bayesian information criterion (BIC) in both Arkansas and Texas. (DOC)

Table S3 The marker loci associated with harvest index traits at Stuttgart, Arkansas and Beaumont, Texas in 2009. (DOC)

References

1. Koutroubas SD, Ntanos DA (2003) Genotype differences for grain yield and nitrogen utilization in indica and japonica rice under Mediterranean conditions. Field Crops Res 83: 251–260.
2. Raes D, Steduto P, Hsiao TC, Fereres E (2009) AquaCrop-The FAO Crop Model to Simulate Yield Response to Water: II. Main Algorithms and Software Description. Agron J 101: 483–497.
3. Donald CM, Hamblin J (1976) The biological yield and harvest index of cereals as agronomic and plant breeding criteria. Adv Agron 28: 361–405.
4. Hay RKM (1995) Harvest index: a review of its use in plant breeding and crop physiology. Annu Appl Biol 126: 197–216.
5. Sinclair TR (1989) Historical changes in harvest index and crop nitrogen accumulation. Crop Sci 38: 638–643.
6. Gepts P (2004) Crop domestication as a long-term selection experiment. Plant Physiol 135: 55–71.
7. Tyagi A, Khurana JP, Khurana P, Raghuvanshi S, Gaur A, et al. (2004) Structural and functional analysis of rice genome. J Genet 83: 79–99.
8. Terao T, Nagata K, Morino K, Hirosue T (2010) A gene controlling the number of proembryos branches also controls the vascular bundle formation and hence is responsible to increase the harvest index and grain yield in rice. Theor Appl Genet 120: 875–893.
9. Marri PR, Sarla N, Reddy LV, Siddiqi EA (2005) Identification and mapping of yield and yield related QTLs from an Indian accession of Oryza sativa L. BMC Genet 13: 33–39.
10. Sabouri H, Sabouri A, Reza DA (1999) Genetic dissection of biomass production, harvest index and panicle characteristics in indica-indica crosses of Iranian rice (Oryza sativa L.) cultivars. Aust J Crop Sci 3: 155–166.
11. Can ND, Yoshida T (1999) Genotypic and phenotypic variances and covariances in early maturing grain sorghum in a double cropping. Pl Prod Sci 2: 67–70.
12. Mohammad D, Coo PB, Posley GL, Kirkham MB, Hussain A, et al. (1993) Correlation of characters contributing to grain and forage yields and forage quality in sorghum (Sorghum bicolor). Indian J Agric Sci 63: 92–95.
13. Soltan A, Rezaei AM, Pour MRK (2001) Genetic variability of some physiological and agronomic traits in grain sorghum (Sorghum bicolor L.). J Sci Tech Agric Nat Resources 5: 127–137.
14. Shrotoria PK, Singh R (1968) Harvest index-A useful selection criteria in sorghum. Sorghum-Newsletter Uttar Pardesh India 31: 4.
15. Yoshida S (1981) Fundamentals of rice crop science. Los BanosPhilippines: International Rice Research Institute. 109 p.
16. Dalling MJ (1985) The physiological basis of nitrogen redistribution during filling in cereals. In: Harper JE, Schrader LE, Howell HW, eds. Exploitation of physiological and genetic variability to enhance crop productivity. Rockville MD: American Society of Plant Physiology. pp 55–71.
17. Prasad PVV, Boullard JLH, Sheehy JE, Thomas JMG (2006) Species, ecotype and cultivar differences in spikelet fertility and harvest index of rice in response to high temperature stress. Field Crops Res 95: 398–411.
18. Peng S, Huang J, Sheehy JE, Laza RC, Vipperan RM, et al. (2004) Rice yields decline with higher night temperature from global warming. Proc Natl Acad Sci USA 101: 9971–9975.
19. Duan F (1997) Formation of harvest index in rice and its improvement. Crop Res 2: 1–3.
20. Mao B-B, Cai WJ, Zhang Z-h, Hu Z-L, Li P, et al. (2003) Characterization of QTLs for Harvest Index and Source-sink Characters in a DH Population of Rice (Oryza sativa L.). Acta Genetica Sinica 30: 1119–1126.
21. Lanceras JC, Pantuvan CP, Jongdeel B, Toojinda T (2004) Quantitative trait loci associated with drought tolerance at reproductive stage in rice. Plant Physiol 135: 384–399.
22. Roug JF, Felus FA, Waghmare VN, Pierce GJ, Chee PW, et al. (2007) Meta-analysis of polyphylly cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics 176: 2577–2588.
23. Hao Z, Li X, Liu X, Xie C, Li M, et al. (2010) Meta-analysis of constitutive and adaptive QTL for drought tolerance in maize. Euphytica 174: 165–177.
24. Zhang YM, Mao Y, Xie C, Smith H, Liao L, et al. (2005) Mapping quantitative trait loci using naturally occurring genetic variance among commercial inbred lines of maize (Zea mays L.). Genetica 126: 2267–2275.
25. Yu JM, Pressoir G, Briggs WH, Bi F, Yamazaki M, et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38: 203–208.
26. Kim S, Zhao K, Jiang R, Molinot J, Boorevitz JO, et al. (2006) Association mapping with single-locus polymorphism. Genetics 173: 1125–1133.
27. Huang X, Wei X, Sang T, Zhao Q, Feng Q, et al. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genetics 42: 961–969.
28. González-Martínez SC, Wheeler NC, Eroso E, Nelson CD, Neale DB (2007) Association genetics in Pinnus taeda L. I. Wood property traits. Genetics 175: 399–409.
29. Kang HM, Zainif NA, Wade CM, Kirby A, Heckerman D, et al. (2008) Efficient control of population structure in model organism association mapping. Genetics 178: 1709–1729.
30. Li X, Yan W, Agrawana H, Jia L, Shen X, et al. (2011) Mapping QTLs for improving grain yield using the USDA rice mini-core collection. PLoS 234: 347–361.
31. Pariseaux B, Bernardo R (2004) In silico mapping of quantitative trait loci in maize. Theor Appl Genet 109: 508–514.
32. Zhao K, Aranzana MJ, Kim S, Later C, Shindo C, et al. (2007) Arabidopsis example of association mapping in structured samples. PLoS Genet 19: 51e; e4.
33. Li X, Yan W, Agrawana H, Hu B, Jia L, et al. (2010) Genotypic and phenotypic characterization of genetic differentiation and diversity in the USDA rice mini-core collection. Genetics 138: 1221–1230.
34. Agrawana HA, Yan WG, Lee FN, Fjellstrom R, Chen MH, et al. (2009) Genetic assessment of a mini-core developed from the USDA rice genebank. Crop Sci 49: 1356–1366.
35. Yan WG, Rutger JR, Bryant RJ, Beckelman HE, Fjellstrom RG, et al. (2007) Development and evaluation of a core subset of the USDA rice (Oryza sativa L.) germplasm collection. Crop Sci 47: 869–878.
36. Nei M, Takezaki N (1994) Estimation of genetic distances and phylogenetic trees from DNA analysis. Proc 3rd World Cong Genet Appl Livestock Prod 24: 405–412.
37. Garris AJ, Tai TH, Coburn J, Kressovich S, McCouch SR (2005) Genetic structure and diversity in Oryza sativa L. Genetics 169: 1631–1638.
38. Agrawana HA, Ezeenga GC (2000) Molecular diversity and genome-wide linkage disequilibrium pattern in worldwide rice and its wild relatives. Euphytica 160: 339–355.
39. Pinto RS, Reynolds MP, Mathews KL, McIntyre CL, Olivares-Villegas JJ, et al. (2010) Heat and drought adaptive QTL in a wheat population to minimize confounding agronomic effects. Theor Appl Genet 121: 1001–1021.
40. Cho YG, Ishii T, Tenshuk S, Chen X, Lipowich L, et al. (2000) Diversity of microsatellites derived from genomic libraries and genbank sequences in rice (Oryza sativa L.). Theor Appl Genet 100: 713–722.
41. Jain S, Jain RK, McCouch SR (2004) Genetic analysis of Indian aromatic and quality rice (Oryza sativa L.) germplasm using panels of fluorescently-labeled microsatellite markers. Theor Appl Genet 109: 963–977.

Table S4 Comparison of allelic effect of four constitutive marker loci at two locations, Stuttgart, Arkansas and Beaumont, Texas. RM431 for Plant height, RM471 and Rsal2 for Plant weight, and RM24011 for Panicle length. (DOC)

Acknowledgments

The authors thank two anonymous reviewers and Ellen McWhirter for critical review. Tiffany Sookaserm, Tony Beaty, Yao Zhou, Biaolin Hu, Melissa Jia, LalDuska Simpson, Curtis Kerns, Sarah Hendrix, Bill Lubcke, Jodie Cammack, Kip Landry, Carl Henry, Jason Bonnette, and Piper Roberts for technical assistance.

Author Contributions

Conceived and designed the experiments: WY AM DW KM. Performed the experiments: XL IJ AY WV. Analyzed the data: XL IJ KY HA. Contributed reagents/materials/analysis tools: HA AJ. Wrote the paper: XL WV AM DW KM.
42. Thomason MJ, Polato NR, Praseityono J, Trijatmiko KR, Silontonga TS, et al. (2009) Genetic diversity of isolated populations of Indonesian landraces of rice (Oryza sativa L.) collected in East Kalimantan on the island of Borneo. Rice 2: 80–92.

43. Borba TCO, Brondani RVP, Rangel PHN, Brondani C (2009) Microsatellite marker-mediated analysis of the EMBRAPA Rice Core Collection genetic diversity. Genetica 137: 293–304.

44. Thomason MJ, Sut smirkh M, Sunarto T, Sintoso TS, et al. (2007) Genetic diversity analysis of traditional and improved Indonesian rice (Oryza sativa L.) germplasm using microsatellite markers. Theor Appl Genet 114: 559–568.

45. Xu YB, Beachell H, McCouch SR (2004) A marker-based approach to broadening the genetic base of rice in the USA. Crop Sci 44: 1947–1959.

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

47. Xin Z, Velten JP, Oliver MJ, Burke JJ (2003) High throughput DNA extraction method suitable for PCR. Biotechniques 34: 826–829.

48. Din RU, Subhani GM, Ahmad N, Hussain M, Rehman AU (2010) Effect of temperature on development and grain formation in spring wheat. Pak J Bot 42: 899–906.

49. Hommmna K, Horie T, Shiraawi T, Strepoldk S, Supapae N (2004) Delay of heading date as an index of water stress in rainfed rice in mini-watersheds in Northeast Thailand field. Crops Res 88: 11–19.

50. Shpiler I, Bham A (1991) Heat tolerance for yield and its components in different wheat cultivars. Euphytica 51: 257–263.

51. Din K, Singh RM (2005) Grain filling duration: An important trait in wheat cultivars. Euphytica 51: 257–263.

52. Vaughan DA, Lu BR, Tomasina N (2008) The evolving story of rice evolution. Plant Sci 174: 394–408.

53. Xin Z, Velten JP, Oliver MJ, Burke JJ (2003) High throughput DNA extraction method suitable for PCR. Biotechniques 34: 826–829.

54. Din RU, Subhani GM, Ahmad N, Hussain M, Rehman AU (2010) Effect of temperature on development and grain formation in spring wheat. Pak J Bot 42: 899–906.

55. Hommmna K, Horie T, Shiraawi T, Strepoldk S, Supapae N (2004) Delay of heading date as an index of water stress in rainfed rice in mini-watersheds in Northeast Thailand field. Crops Res 88: 11–19.

56. Shpiler I, Bham A (1991) Heat tolerance for yield and its components in different wheat cultivars. Euphytica 51: 257–263.

57. Din K, Singh RM (2005) Grain filling duration: An important trait in wheat cultivars. Euphytica 51: 257–263.

58. Vaughan DA, Lu BR, Tomasina N (2008) The evolving story of rice evolution. Plant Sci 174: 394–408.

59. Dini R, Subhani GM, Ahmad N, Hussain M, Rehman AU (2010) Effect of temperature on development and grain formation in spring wheat. Pak J Bot 42: 899–906.

60. Hommmna K, Horie T, Shiraawi T, Strepoldk S, Supapae N (2004) Delay of heading date as an index of water stress in rainfed rice in mini-watersheds in Northeast Thailand field. Crops Res 88: 11–19.

61. Shpiler I, Bham A (1991) Heat tolerance for yield and its components in different wheat cultivars. Euphytica 51: 257–263.

62. Din K, Singh RM (2005) Grain filling duration: An important trait in wheat cultivars. Euphytica 51: 257–263.

63. Vaughan DA, Lu BR, Tomasina N (2008) The evolving story of rice evolution. Plant Sci 174: 394–408.

64. Dini R, Subhani GM, Ahmad N, Hussain M, Rehman AU (2010) Effect of temperature on development and grain formation in spring wheat. Pak J Bot 42: 899–906.

65. Hemamalini GS, Shashidhar HE, Hittalmani S (2000) Molecular marker assisted tagging of morphological and physiological traits under two contrasting moisture regimes at peak vegetative stage in rice (Oryza sativa L.). Euphytica 112: 69–70.

66. Ando T, Yamamoto T, Shimizu T, Ma XF, Shomura A, et al. (2000) Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. Theor Appl Genet 116: 681–692.

67. Hittalmani S, Huang N, Courtou B, Vemuprasad R, Shashidhar HE, et al. (2003) Identification of QTL for growth- and grain yield-related traits in rice across nine locations of Asia. Theor Appl Genet 107: 679–690.

68. Xu W, Xing Y, Weng X, Zhao Y, Tang W, et al. (2008) Natural variation in GhHIs is an important regulator of heading date and yield potential in rice. Nat Genet 41: 1–7.

69. Peng J, Richards DE, Hartley NM, Murphy GP, Deivos KM, et al. (1999) ‘Green revolution’ genes encode mutant gibberellic response modulators. Nature 400: 296–298.

70. Fu Q, Zhang P, Tan L, Zhu Z, Ma D, et al. (2010) Analysis of QTLs for yield-related traits in Yuanjiang common wild rice (Oryza rufipogon Griff.). J Genet Genomics 37: 147–157.

71. Hedden P (2003) The genes of the Green Revolution. Trends Genet 19: 5–9.

72. Yan WG, Rutger JN, Bockelman HE, Tai TH (2005) Agronomic evaluation and seed stock establishment of the USDA rice core collection. In: Norman RJ, Meullenet JF, Mohnenharau KAK, eds. BR Wells Rice Research Studies. Stuttgart: University of Arkansas, Agri Exp Sta Res Ser. pp 63–68.

73. Yan WG, Rutger JN, Bockelman HE, Tai TH (2005) Evaluation of kernel characteristics of the USDA rice core collection. In: Norman RJ, Meullenet JF, Mohnenharau KAK, eds. BR Wells Rice Research Studies. Stuttgart: University of Arkansas, Agricultural Experimentation Station Research Serie. Agri Exp Sta Res Ser. pp 69–74.

74. Brondani C, Borba TCO, Rangel PHN, Brondani RVP (2006) Determination of traditional varieties of Brazilian rice using microsatellite markers. Genet Mol Biol 29: 676–684.

75. Xin Z, Velten JP, Oliver MJ, Burke JJ (2003) High throughput DNA extraction method suitable for PCR. Biotechniques 34: 826–829.

76. Borba TCO, Brondani RVP, Rangel PHN, Brondani C (2005) Evaluation of the number and information content of fluorescence-labeled SSR for rice germplasm characterization. Crop Breed Appl Biotechnol 2: 157–165.

77. Liu K, Mue SV (2005) Powermarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128–2129.

78. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Bio Evol 24: 1596–1599.

79. Wang LQ, Liu WJ, Xu Y, He YQ, Luo LJ, et al. (2007) Genetic basis of 17 traits and vescosity parameters characterizing the eating and cooking quality of rice grain. Theor Appl Genet 115: 463–476.

80. Rohlf F (2000) FNTSYS-PC numerical taxonomy and multivariate analysis system version 2.11L. Applied Biostatistics, NY.

81. Prichard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. AM J Hum Genet 67: 170–181.

82. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611–2620.

83. 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.

84. Wang ML, Hu G, Barkley NA, Chen Z, Erspelding JE, et al. (2009) Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. Theor Appl Genet 120: 13–23.

85. Saxton AM (1998) A macro for converting mean separation output to letter groupings in Proc Mixed. In: Proc 23rd SAS Users Group Cary NC. SAS Institute. pp 1243–1246.