JOURNAL OF PLANT REGISTRATIONS

MAPPING POPULATION

Registration of the Rice Diversity Panel 1 for Genomewide Association Studies

Georgia C. Eizenga,* Md. Liakat Ali, Rolfe J. Bryant, Kathleen M. Yeater, Anna M. McClung, and Susan R. McCouch

ABSTRACT

The Rice Diversity Panel 1 (Reg. No. MP-6, NSL 500357 MAP) (RDP1) is a collection of 421 purified, homozygous rice (*Oryza sativa* L.) accessions (GSOR 301001 through GSOR 301421; GSOR 312001 through 312020) representing the broad range of genetic variation within *O. sativa*. The accessions include both landraces and elite rice cultivars, which were classified into five subpopulation groups, including *indica* (95 accessions) and *aus* (60), which belong to the *Indica* varietal group, and *tropical japonica* (106), *temperate japonica* (111), and *aromatic* (Group V) (16) which comprise the *Japonica* varietal group. Thirty-three accessions are classified as admixtures because they shared <60% ancestry with a single group. The seed, with and without the hull, and panicle morphology of each accession were documented with digital images, and the RDP1 was phenotyped for morphological, developmental, and physiological traits. Genotypes for 36,901 SNP loci are publicly available for additional genomewide association mapping studies. In this report, we evaluate three grain quality traits on the RDP1: apparent amylose content (AC), gelatinization temperature as measured by alkali spreading value (ASV), and protein content. Canonical discriminant analysis revealed AC was the quality trait most closely correlated with subpopulation structure, followed by ASV. These traits indicate that *temperate japonica* was the most distinct group, whereas *aus* and *indica* could not be differentiated, and the *aromatic* accessions were closest to *tropical japonica*.

The Rice Diversity Panel 1 (Reg. No. MP-6, NSL 500357 MAP) (RDP1) is a germplasm collection consisting of 421 purified *Oryza sativa* L. accessions collected from 10 geographic regions where rice is grown. Through the use of simple sequence repeats (SSRs) (Ali et al., 2011) and single nucleotide polymorphisms (SNPs) (Zhao et al. (2010, 2011), the panel has been characterized for genetic subpopulation structure. The accessions represent the five major subpopulation groups identified as (i) *indica* and (ii) *aus,* comprising the *Indica* varietal group, and (iii) *tropical japonica,* (iv) *temperate japonica* and (v) *aromatic* (Group V), which collectively comprise the *Japonica* varietal group (Garris et al., 2005; Zhao et al., 2011; Huang et al., 2012). These same subpopulations have been identified in numerous studies using a variety of different collections of germplasm and types of molecular markers (Garris et al., 2005; Caicedo et al., 2007; Gross et al., 2010; Thurber et al., 2010; Zhao et al., 2010, Zhao et al., 2011; Huang et al., 2012).

To date, the RDP1 has been characterized for several agronomic, morphological, developmental, and physiological traits under both field conditions and in controlled environments. Traits that have been evaluated in the field include days to heading, plant height, flag leaf length, flag leaf width, panicle number per plant, panicle length, primary branch number per panicle, spikelets per panicle, number of unfilled grains per panicle, number of filled grains per panicle, seed length, seed width, seed length:width ratio, brown rice length, brown rice width, seed volume, 100-seed weight, and brown rice volume (Ali et al., 2011), as well as grain arsenic (Norton et al., 2012). Traits evaluated under controlled conditions include aluminum tolerance (Famoso et al., 2011), root system architecture (Clark et al., 2011, 2013; P. Benfey, Duke University, personal communication, 2012), root hair development (P. Vejchasarn

G.C. Eizenga, R.J. Bryant, and A.M. McClung, USDA–ARS, Dale Bumpers National Rice Research Center, 2890 Hwy. 130 E., Stuttgart, AR 72160; M.L. Ali, Univ. of Arkansas, Rice Research and Extension Center, 2900 Hwy. 130 E., Stuttgart, AR 72160; K.M. Yeater, USDA–ARS–SPA, 1001 Holleman Dr. East, College Station, TX 77840; S.R. McCouch, Dep. of Plant Breeding and Genetics, 162 Emerson Hall, Cornell Univ., Ithaca, NY 14853-1901. M.L. Ali, current address: Univ. Nebraska–Lincoln, Dep. of Agronomy and Horticulture, 202 Keim Hall, Lincoln, NE 68583-0915.

Abbreviations: AC, apparent amylose content; ASV, alkali spreading value; GRIN, Germplasm Resources Information Network; GSOR, Genetic Stocks-*Oryza*; GWAS, genomewide association studies; NSGC, National Small Grains Collection; PCR, polymerase chain reaction; RDP1, Rice Diversity Panel 1; SNP, single nucleotide polymorphism; SSR, simple sequence repeat.

Copyright © Crop Science Society of America. All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

Journal of Plant Registrations 8:109–116 (2014).
doi: 10.3198/jpr2013.03.0013crmp
Supplemental data file is available online for this article.
Received 13 Mar. 2013. Registration by CSSA.
5585 Guilford Rd., Madison, WI 53711 USA
*Corresponding author (georgia.eizenga@ars.usda.gov)

and K. Brown, Pennsylvania State University, personal communication, 2012), blast disease resistance (G. Wang, The Ohio State University, personal communication, 2012), element concentrations (ionomics) in seeds and vegetative tissues (J. Cobb, Cornell University, personal communication, 2012; D. Salt, University of Aberdeen, personal communication, 2012), and phosphorus and zinc use efficiency (M. Wissuwa, Japan International Center for Agricultural Sciences, personal communication, 2012). In addition, the RDP1 has been evaluated for the three traits that affect the cooking, sensory and nutritional quality of rice, including apparent amylose content (AC), alkali spreading value (ASV), and protein content. The analysis of these biochemical grain quality traits is discussed below.

Genomewide association studies (GWAS) using the RDP1 have successfully identified phenotype–genotype associations using both single SNP approaches and admixture analysis (Zhao et al., 2010, 2011). These studies acknowledge the deep subpopulation structure in rice and highlight the heterogeneity of genetic architecture among subpopulations for virtually all traits analyzed to date.

Amylose content and gelatinization temperature are the two major determinants of grain cooking and sensory quality in rice. Amylose is a factor responsible for hardness, stickiness, color, gloss, and general acceptability of cooked rice. Rice with a high amylose content (about 230–300 g kg^{-1}) tends to cook firm and dry, whereas rice with an intermediate amylose content (about 180–230 g kg^{-1}) tends to be softer and stickier, and rice with a low amylose content (<180 g kg^{-1}) is generally very soft and sticky or chewy (Bergman et al., 2004). Waxy rice has zero amylose content and is often referred to as glutinous, or sweet rice. Amylose is also responsible for the manner in which cooked rice hardens on cooling due to starch retrogradation. The amylose content of milled rice has been found to correlate positively with hardness values of cooked rice and negatively with the stickiness value (Champagne et al., 2004; Cho et al., 2010).

Gelatinization temperature is the temperature at which the starch granules lose their crystalline structure following hydration and heating. Rice starch usually gelatinizes between 65 and 85°C. Rice that gelatinizes at the lower temperatures often cooks to a softer texture and requires less cooking time than rice with a higher gelatinization temperature (Bhattacharya and Sowbhagya, 1971). Gelatinization temperature is often predicted by the ASV assay measured by degree of digestion of the milled rice grain soaked in potassium hydroxide solution (Little et al., 1958). Protein is an important nutritional component located in both the bran (aleurone) layer and the endosperm. Protein content can affect grain chalkiness, grain color, cooked texture, and starch pasting properties (Champagne et al., 2007; Yadav and Jindal, 2008; Chun et al., 2009).

The public release of the RDP1, in combination with the publicly available genotypic dataset consisting of 36,901 genomewide SNPs (Rice Diversity Project: Data Sets, http://www.ricediversity.org/data/) and 36 SSRs (USDA–ARS, Germplasm Resources Information Network [GRIN], Accession Area Queries, http://www.ars-grin.gov/cgi-bin/npgs/crop/evaluation.pl?494754, then select download at the bottom of the page) will allow this resource to be widely used as the basis for GWAS, fine mapping, and gene discovery. The fact that all the lines have been purified ensures homozygosity and that this eternal GWAS panel can be used reliably for phenotyping diverse traits. Accessions with desirable traits can be further dissected genetically or used as donors by rice breeders to incorporate the trait of interest into new genetic backgrounds and cultivars using marker-assisted selection.

Materials and Methods

Development of the Plant Materials

The RDP1 consists of 421 Asian rice (*O. sativa*) landraces and cultivars from 79 countries representing all the major rice growing regions of the world. The collection includes 139 accessions from East and Southeast Asia, 77 from South Asia, 18 from West and Central Asia, 42 from Africa, 90 from the Americas, 35 from Europe, 6 from Oceania, and 14 of unknown origin (Supplemental Table S1). Of the 421 accessions, 150 were a part of a previous genetic diversity study by Garris et al. (2005) and were obtained from the USDA–ARS Rice Research Unit in Beaumont, TX; 159 accessions were from the rice core collection within the USDA–ARS National Small Grains Collection (NSGC) (Yan et al., 2007); 54 were obtained directly from the USDA–ARS NSGC rice collection (Yan et al., 2007); 18 accessions were part of the *Oryza*SNP project (McNally et al., 2009); and 11 accessions were used as parents for developing the mapping populations evaluated as part of the USDA/CSREES/NRI Applied Plant Genomics Program titled RiceCAP: A Coordinated Research, Education, and Extension Project for the Application of Genomic Discoveries to Improve Rice in the United States (USDA/CSREES grant 2004-35317-14867). Eighteen accessions were parents in other mapping populations, 4 (Spring, Cocodrie, Cybonnet, 93-11) were controls across all field phenotyping environments, and 14 were later found to be incorrectly identified based on genotypic data or comparison with voucher seed specimens by personnel at the International Rice Research Institute (Philippines) and were renamed by the NSF-TV ID number as "NSF-TV no." Additional information regarding the selection of these accessions and field data collection can be found in Ali et al. (2011).

All accessions were purified for two generations via single seed descent to ensure homogeneity and homozygosity before phenotyping and genotyping. DNA used for SSR and SNP analysis was extracted from leaf tissue harvested from single plants. Supplemental Table S1 contains information about the accession name, accession number, original providing country, geographic region of origin, subpopulation group based 36 SSRs (Ali et al., 2011) and on 36,901 SNPs (Zhao et al., 2011). In summary, the RDP1 contains 166 accessions belonging to the *Indica* varietal group including the *indica* (95 accessions), *aus* (60), and admixture of *Indica* (11), and 255 accessions comprising the *Japonica* varietal group including the *tropical japonica* (106), *temperate japonica* (111), admixture of *Japonica* (22), and *aromatic* (Group V) (16).

Evaluation of Grain Quality Traits

For most all grain quality analyses, samples were harvested from three plants grown in one of two replications of field trials conducted in two different years near Stuttgart, AR. For

additional details on the field trials, see Ali et al. (2011). There were 33 accessions that did not produce adequate seed in these field trials and were grown either in the field at Beaumont, TX, in a winter nursery in Puerto Rico, and/or in the greenhouse during the fall and winter months as noted in Supplemental Table S1, with additional details on individual accessions available from coauthors (Eizenga, McCouch, McClung). Approximately 280 seeds were dehulled (lemma and palea removed) using a Kett TR200 Automatic Rice Husker. Subsequently, these seeds were milled using a Kett Pearlest Grain Polisher for AC and ASV. Approximately 20 unbroken seeds were removed for ASV, and the remaining milled seeds were ground using a Cyclotec grinder (Foss North America) with a 0.5-mm mesh screen for AC. Apparent amylose content was determined on rice flour using a simplified version of the Williams et al. (1958) method adapted to an autoanalyzer (Juliano, 1971). A standard curve was constructed from cultivars with known AC. Technical replicates of the samples were analyzed for AC, ASV, and protein content.

Gelatinization temperature was inferred using the method described by Little et al. (1958). Six unbroken milled grains were incubated in 10 mL of 1.7% KOH in deionized water (w/v) at 30°C for 23 h, and the degree of disintegration of the grain was determined using a six-point scale (categories 1 and 2 were considered category 2): (2) grain not affected to slightly swollen; (3) starch collar incomplete and narrow; (4) grain swollen, collar complete and wide; (5) grain split or segmented, collar complete and wide; (6) grain dispersed, merging with collar (displaying cotton center); and (7) grain completely dispersed and intermingled (displaying clear center). Alkali spreading value is inversely related to gelatinization temperature (Juliano, 1985).

Nitrogen concentration was determined on brown rice flour (0.2 g) using a LECO FP-2000 Nitrogen Analyzer. Approximately 280 seeds were dehulled; then the brown rice (not milled) was used to produce flour as previously described for AC evaluation. Nitrogen concentration was reported as percent protein using 5.95 as a conversion factor (AACC International, 2000).

Two polymerase chain reaction (PCR)-based markers, Intron 1 of the *waxy* gene (Ayres et al., 1997) and RM190 (Temnykh et al., 2000), both known to be associated with AC (Chen et al., 2008), were used to identify putative AC genotypes. Similarly, the PCR-based *Alk* gene marker was used to discriminate between the high or intermediate gelatinization temperatures and low gelatinization temperature using the GC (92 bp) and TT (94 bp) alleles, respectively, for the fourth and fifth SNPs reported by Umemoto and Aoki (2005).

Statistical Analysis

The SAS software (SAS Institute, 2008) was used for all statistical analyses of the grain quality data. The differences in grain quality among accessions and subpopulations (as assigned in Ali et al., 2011; Supplemental Table S1) were assessed via canonical discriminant analysis using the CANDISC procedure of SAS software. Canonical discriminant analysis is a combination of principal component and canonical correlation analyses (Vaylay and van Santen, 2002). From this analysis, the Mahalanobis distance (D^2) statistic was calculated as an indicator of the difference between populations (Loos, 1993). These analyses are further described in Ali et al. (2011). Correlation analysis among trait means was performed using the CORR procedure with the Pearson correlation coefficient and between the trait means and marker (Intron 1, RM190, *Alk*) classes using the Spearman correlation coefficient.

Results

Grain Quality Trait Variation between the Subpopulations

The variation observed among the five subpopulations and two admixture groups was highly significant for the three grain quality traits: AC, ASV, and protein content (Table 1). The highest mean AC levels were observed in the *aus* (250.9 g kg^{-1}) and *indica* (230.1 g kg^{-1}) subpopulations and the admixture of *Indica* accessions (249.6 g kg^{-1}), whereas the accessions containing the lowest AC were the *temperate japonica* subpopulation (160.2 g kg^{-1}) and the admixture of *Japonica* group (146.0 g kg^{-1}). The accessions classified as *aromatic* and *tropical japonica* contained intermediate AC, with means of 198.0 and 196.0 g kg^{-1}, respectively.

Table 1. Subpopulation means and standard deviations for amylose content, alkali spreading value (ASV), and protein content. Intron 1 and RM190 alleles are used to predict amylose content; *Alk* alleles are used to predict gelatinization temperature. The most prevalent allele for the aforementioned markers in a given subpopulation is listed along with the percentage of accessions with the most prevalent allele.

Subpopulation	No. accessions	Apparent amylose content	Intron 1 alleles†	RM190 alleles‡ (% accessions)	ASV§	*Alk* alleles¶	Protein content
		g kg^{-1}	% accessions		2–7	% accessions	g kg^{-1}
Aus	58	250.9 ± 11.5	145/147 bp (100%)	101/105 bp (95%)	5.49 ± 0.66	92 bp (100%)	81.9 ± 56
Indica	89	230.1 ± 43.5	145/147 bp (89%)	101/105/107 bp (84%)	6.20 ± 0.85	92 bp (76%)	82.9 ± 9.1
Admixture of *Indica*	11	249.6 ± 20.5	145/147 bp (100%)	101/105/107 bp (91%)	6.15 ± 0.50	92 bp (100%)	84.8 ± 8.0
Aromatic (Group V)	16	198.0 ± 21.0	147 bp (100%)	118/120/122/124 bp (100%)	5.81 ± 0.89	92 bp (100%)	92.6 ± 10.9
Temperate japonica	110	160.2 ± 37.2	150 bp (73%)	120/122/124 bp (96%)	6.50 ± 0.73	94 bp (58%)	86.8 ± 10.9
Tropical japonica	106	196.0 ± 40.4	147 bp (79%)	113/118/126 bp (61%)	5.43 ± 1.17	92 bp (88%)	87.5 ± 7.5
Admixture of *Japonica*	22	146.0 ± 80.1	150 bp (57%)	120/122/124 bp (73%)	6.05 ± 1.31	92 bp (68%)	91.4 ± 9.9

† The Intron 1 alleles are: 145/147 bp for high or intermediate amylose and 150 bp for low amylose (Ayres et al., 1997; Chen et al., 2008).

‡ The RM190 alleles are: 101/105/107 bp for high amylose, 126 bp for high or intermediate amylose, 113/118 bp for intermediate amylose, and 120/122/124 bp for intermediate or low amylose (Temnykh et al., 2000; Chen et al., 2008).

§ The ASV is a scoring of 2 (milled grain not affected to slightly swollen) to 7 (milled grain completely dispersed), which is inversely related to gelatinization temperature (Juliano 1985); thus, samples with a low ASV score have a high gelatinization temperature.

¶ The *Alk* alleles are 92 bp for high or intermediate gelatinization temperature and 94 bp for low gelatinization temperature (Umemoto and Aoki, 2005).

The *waxy* Intron 1 and RM190 markers were highly correlated with AC ($r = -0.74$ and $r = -0.71$, respectively; $p = 0.0001$). Whereas accessions in the *Indica* varietal group (*aus*, *indica*, admixture of *Indica*) tended to be uniform for Intron 1 alleles associated with high or intermediate AC, and >88% of these accessions were associated with the high amylose alleles for RM190, there was greater allelic variability at these two markers among the three *japonica* subpopulations (Table 1). The *aromatic* (Group V) accessions in the RDP1 are characterized by intermediate AC, which is confirmed by 100% of the accessions having the Intron 1 allele associated with intermediate (or high) AC and 100% of these carrying RM190 alleles, which are associated with intermediate (or low) AC. Although these two markers explain only a portion of the variability in AC, these results illustrate that the markers can be used to characterize differences between some of the *O. sativa* subpopulations.

Accessions displaying the highest ASV, which indicates the lowest gelatinization temperatures, were found in the *temperate japonica, indica*, admixture of *Indica*, and admixture of *Japonica* subpopulations with mean ASVs of 6.50, 6.20, 6.15, and 6.05, respectively, whereas accessions with the lowest ASVs (highest gelatinization temperatures) belonged to the *aus* and *tropical japonica* subpopulations, with mean ASVs of 5.49 and 5.43, respectively. The *aromatic* (Group V) subpopulation had intermediate ASV of 5.81 (Table 1). Alleles at the *Alk* locus were correlated with ASV ($r = 0.53$; $p = 0.0001$) in the RDP1 accessions. This low correlation is because ASV ranged from 2.0 to 7.0 for the accessions with the 92-bp allele. To separate these accessions, a third functional SNP (A/G) in the rice *starch synthase IIa* (*SSIIa*) gene needs to be evaluated (Umemoto et al., 2004; Umemoto and Aoki, 2005). This gene produces the *SSIIa* enzyme, which has a major effect on rice gelatinization temperature; and because it was not available as a PCR-based marker, the RDP1 accessions were not screened for this marker. The range in ASV for those accessions with the 94-bp allele was 5.5 to 7.0 with a mean of 6.75, reflective of a low gelatinization temperature.

The *aromatic* (Group V) accessions had the highest protein content, with a mean of 92.6 g kg^{-1}, whereas *aus* (81.9 g kg^{-1}) and *indica* (82.9 g kg^{-1}) accessions had the lowest protein content (Table 1). *Temperate japonica* and *tropical japonica* had intermediate protein content, with mean values of 86.8 and 87.5 g kg^{-1}, respectively.

Correlations between the biochemical and morphological grain quality traits, including grain (brown rice) length and width, were evaluated for each subpopulation. Although there were several significant correlations, most were nonsignificant, reflecting the diversity that is present within the subpopulations. However, for the *aromatic* (Group V) accessions, grain width was positively correlated with protein content ($r = 0.54$, $p = 0.05$) and negatively associated with AC ($r = -0.44$, $p = 0.08$). Among the *temperate japonica* accessions, a negative association between grain length and ASV ($r = -0.48$, $p = 0.0001$) and a positive association between grain width and ASV ($r = 0.28$, $p = 0.004$) was observed. This reflects the preference for sticky rice with shorter and wider grains in the regions where *temperate japonica* accessions are grown. Among the *tropical japonica* accessions, there was a negative correlation between grain width and AC, ($r = -0.28$, $p = 0.003$), and ASV was negatively correlated with grain length ($r = -0.27$, $p = 0.005$). This demonstrates a trend within this subpopulation for genotypes that have long, slender grain shape to have intermediate to high amylose content and gelatinization temperature. The strongest correlations observed within the *Indica* subpopulations were in the *aus* subpopulation where grain width was negatively correlated with AC ($r = -0.39$, $p = 0.003$); in the *indica* subpopulation where grain width was negatively correlated with ASV ($r = -0.22$, $p = 0.03$); and the admixture of *Indica* group, where ASV was positively associated with grain length ($r = 0.60$, $p = 0.05$). In addition, in the admixture of *Indica* group and the *temperate japonica* subpopulation, positive correlations were observed between protein content and grain length ($r = 0.77$, $p = 0.006$; and $r = 0.27$, $p = 0.004$, respectively). These observations again underscore the genetic heterogeneity of complex trait variation in *O. sativa* and the deep differences that exist among the subpopulations.

Canonical Discriminant Analysis

The canonical loadings (coefficients) showing the contribution of each of the three measured grain quality traits are presented in Table 2. The first two canonical variates together accounted for 99% of the variance in grain quality observed among the RDP1 accessions. The significant ($p < 0.0001$) canonical correlations between the accessions, and the first canonical variate ($r_c = 0.71$) and the second canonical variate ($r_c = 0.40$) indicate that the canonical variates explain virtually all of the phenotypic differences between the accessions. The first canonical discriminant function (canonical variate 1) explained 83% of the variance and was dominated by the singularly large loading from AC. Canonical variate 2 accounted for 16% of the variation and was dominated by a large loading from ASV, followed by protein content. These results indicate that among these quality traits, mean AC plays a major discriminatory role among subpopulations.

The centroid values of the first two canonical discriminant functions for each accession were plotted as the x and y axis and coded as to their subpopulation membership to visualize the subpopulation clusters (Fig. 1). The extent of separation between the subpopulation groups was measured by D^2 (Table 3). Most of the pairwise distances were significant ($p < 0.0001$, $p < 0.001$, or $p < 0.01$). The greatest phenotypic distance was observed between the *aus* subpopulation and accessions classified as admixture within the *Japonica* varietal group ($D^2 = 9.77$), followed by the admixture within the *Japonica* and admixture within the *Indica* varietal groups ($D^2 = 8.02$). The distance between the *temperate* and *tropical japonica* groups ($D^2 = 2.46$) was statistically significant, with the *temperate japonica* group

Table 2. The canonical loadings of grain quality traits on the first two canonical variables of the RDP1.

Quality trait	Canonical variate	
	1	2
Apparent amylose content	0.89	0.28
Alkali spreading value (gelatinization temp.)	−0.29	0.88
Protein content	−0.23	−0.49
Canonical correlation	0.71	0.40
p level of significance	0.0001	0.0001
Variance accounted for, %	0.83	0.16

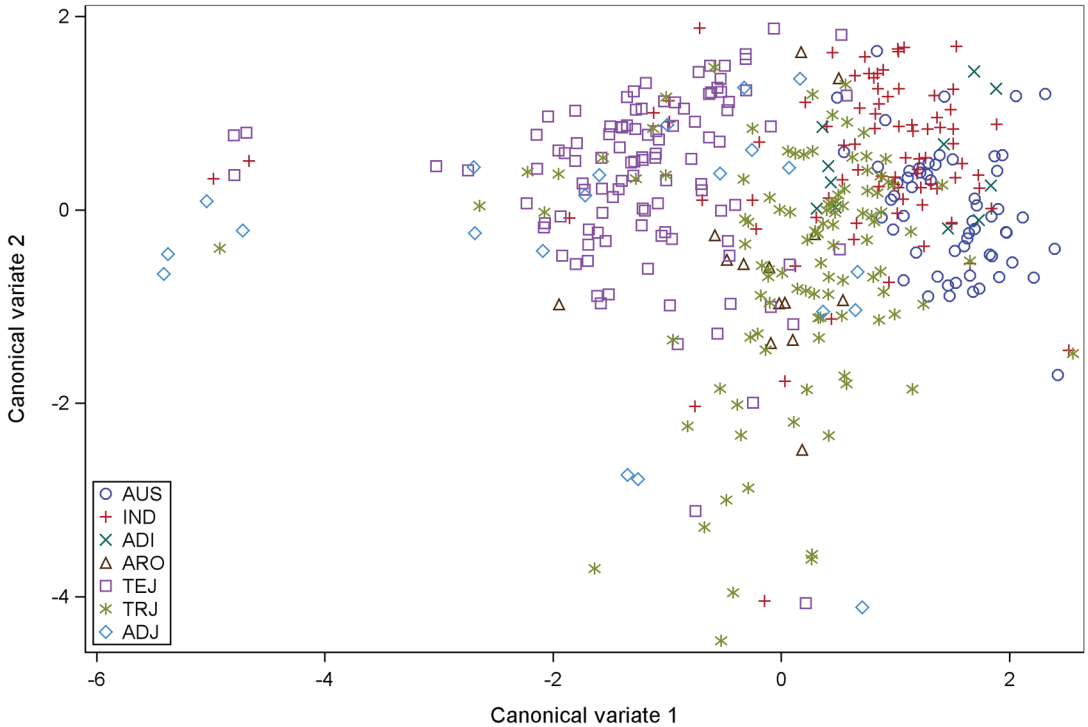

Fig. 1. Scatterplot of the rice accessions belonging to different subpopulations based on the first two canonical discriminant functions (variates) for apparent amylose content, alkali spreading value, and protein content. Accessions are labeled with the abbreviations of their subpopulation names: ARO, *aromatic* (Group V) with 16 accessions; AUS, *aus* (60); IND, *indica* (95); TEJ, *temperate japonica* (111); TRJ, *tropical japonica* (106); ADI, admixture of *Indica* (11); and ADJ, admixture of *Japonica* (22).

being more distant from both the *indica* (D^2 = 3.55) and *aus* (D^2 = 7.39) subpopulations than the *tropical japonica* subpopulation was from these subpopulations (*indica* [D^2 = 1.44], *aus* [D^2 = 2.31]). Phenotypically, the *aromatic* subpopulation (Group V) was quite distant from *temperate japonica* (D^2 = 2.38), *aus* (D^2 = 3.48) and *indica* (D^2 = 2.03) but close to *tropical japonica* (D^2 = 0.47) as noted in the AC and ASV means for these two subpopulations (Table 3).

Variation in Grain Quality Traits between Geographic Regions

The accessions of this panel were divided into 10 groups based on geographic region of origin (Table 4). Fourteen accessions were of unknown geographical origins and were excluded from this analysis. Phenotypic variation among the 10 geographic regions for the three grain quality traits was highly significant (Table 4). Accessions containing the highest AC were from South Asia with a mean value of 241.2 g kg^{-1}, whereas the accessions with lowest AC were from Central Asia with a mean AC of 160.1 g kg^{-1}. Accessions originating from Europe and East Asia also had fairly low AC with a mean of 169.5 and 180.0 g kg^{-1}, respectively. Intermediate and highly variable levels of AC were observed in the accessions from Oceania, North America and the Caribbean, Africa, South and Central America, Southeast Asia, and West Asia, with means ranging from 188.4 to 216.4 g kg^{-1}.

Accessions that originated from Europe and East Asia displayed the lowest gelatinization temperature, with a mean ASV of 6.42 and 6.40, respectively, whereas accessions originating from Africa and Central Asia revealed the highest gelatinization temperatures, with mean ASVs of 5.36 and 5.38, respectively. Accessions originating from other regions, such as Oceania, South and Central America, West Asia, South

Table 3. Pairwise squared distance, a measure of the phenotypic divergence of grain quality traits between the subpopulation groups as calculated by Mahalanobis Distance (D^2).

Subpopulation	*Indica*	Admixture of *Indica*	*Aromatic* (Group V)	*Temperate japonica*	*Tropical japonica*	Admixture of *Japonica*
Aus	0.93***	0.67	3.48***	7.39***	2.31***	9.77***
Indica		0.31	2.03***	3.55***	1.44***	5.90***
Admixture of *Indica*			2.58*	5.51***	2.28***	8.02***
Aromatic (Group V)				2.38***	0.47	2.41**
Temperate japonica					2.46***	0.69*
Tropical japonica						2.91***

* Significant at the 0.01 level.

** Significant at the 0.001 level.

*** Significant at the 0.0001 level.

Asia, North America and the Caribbean, and Southeast Asia, displayed mean ASVs ranging from 5.54 to 6.17.

The highest mean protein content was found in accessions sampled from Central Asia (99.3 g kg^{-1}), followed by accessions sampled from Europe (90.7 g kg^{-1}) and West Asia (90.0 g kg^{-1}) Accessions from South Asia (82.5 g kg^{-1}), Oceania (83.6 g kg^{-1}), and East Asia (84.5 g kg^{-1}) had the lowest mean protein content. Accessions originating from West Asia, North America and the Caribbean, South and Central America, Southeast Asia, and Africa were intermediate in protein content compared to the other regions.

Discussion

The five subpopulations of *O. sativa* can be differentiated on the basis of grain quality traits. Different cooking textures and sensory qualities are preferred in different regions of the world, where cuisines have coevolved with the regional climatic adaptation of the different subpopulations. It is not surprising, therefore, that the landraces and cultivars included in the RDP1 have a predominance of certain biochemical grain quality traits in the subpopulations that are adapted to the regions where those grain quality characteristics are preferred.

The canonical discriminant analysis demonstrated that among the grain quality traits evaluated, the subpopulations were primarily differentiated by AC. Although significant differences existed among accessions in the RDP1 for ASV and protein content, these phenotypes did not discriminate subpopulations as well as AC. This is not surprising considering the importance that AC has on determining the texture of cooked milled rice (Champagne et al., 2004).

Analysis of the grain quality traits in this study and agro-morphological traits analyzed by Ali et al. (2011) confirm that although there is significant variability within each of the five rice subpopulations, they can be differentiated on the basis of numerous traits, including plant height, number of panicles per plant, panicle length, and seed shape, as well as on cooking and eating qualities. This is visualized in the scatterplot for the first two canonical discriminant functions (variates) from the canonical discriminant analysis of the subpopulations using the 18 agro-morphological traits (Ali et al., 2011) and three grain quality traits (Supplemental Fig. S1), as well as the canonical loadings of these variables (Supplemental Table S2). Genotypic analysis has demonstrated that although accessions in the RDP1 share many alleles, some of the distinct phenotypic features of each subpopulation can be traced to selection events associated with domestication and cultivar differentiation (Zhao et al., 2010; Huang et al., 2012; McCouch et al., 2012).

Rice Diversity Panel as a Genetic Resource

The RDP1 is a genetic resource that can be used to conduct genomewide association analysis for virtually any trait of interest. Because all accessions were genetically purified via two generations of single seed descent before DNA extraction and phenotyping, the RDP1 is an eternal, genetically stable population of pure lines, allowing comparisons to be drawn across present and future generations of seed production that are used in a variety of studies. To date, the accessions have been evaluated for more than 60 phenotypic traits by a diverse group of researchers from around the world. Means for the quality traits, traits analyzed by Ali et al. (2011), plant type, hull color, pericarp color, field leaf blast disease rating, straighthead disease rating, and alleles for blast disease markers (RM144, RM224, AP5659_1) and quality markers (RM190, Intron 1, *Alk*) collected from each accession in the RDP1 are available at the USDA–ARS GRIN Web site (by querying Rice Diversity Panel, selecting any individual accession, observations, a single trait, the study or evaluation, and selecting download at the bottom of the page [http://www.ars-grin.gov/npgs/acc/acc_queries.html]) and the Genetic stocks–*Oryza* (GSOR) collection Web site (http://www.ars.usda.gov/Main/Docs.htm?docid=18825). In addition, the panel has been characterized with 36 SSRs (Ali et al., 2011), available through GRIN (http://www.ars-grin.gov/cgi-bin/npgs/crop/evaluation.pl?494754; select download at the bottom of the page), an Illumina 1536-SNP OPA (1311 high-quality SNPs were identified) (Zhao et al., 2010), and an Affymetrix custom-designed genotyping array providing 36,901 high quality SNP genotypes (Zhao et al., 2011) available at the Rice Diversity Project Web site at Data Sets (http://www.ricediversity.org/data). With genotypic data from genotyping-by-sequencing (GBS) technologies and a high-density SNP genotyping array becoming available in the near future, these genetic tools will facilitate further exploration of both genotypic and phenotypic diversity of the rice genome and rapid discovery of genes and genetic elements that condition complex trait variation in rice.

A Seed Photo Library (Fig. 2) provides digital images of four seeds from each accession with and without the hull (lemma and palea), aligned to a ruler showing size (Rice Diversity Project Seed Photo Library, http://www.ricediversity.org/

Table 4. Means and standard deviations for three grain quality traits based on the ten geographic regions from which the accessions originated.

Geographic region	No. accessions	Apparent amylose content	Alkali spreading value	Protein content
		g kg^{-1}	2–7	g kg^{-1}
East Asia	83	180.0 ± 72.4	6.40 ± 0.78	84.5 ± 10.8
Southeast Asia	51	203.5 ± 55.4	6.17 ± 0.78	85.9 ± 9.6
South Asia	73	241.2 ± 25.5	5.86 ± 0.78	82.5 ± 7.6
West Asia	12	216.4 ± 34.3	5.65 ± 0.85	90.0 ± 13.0
Central Asia	6	160.1 ± 10.2	5.38 ± 1.56	99.3 ± 12.5
Africa	41	200.9 ± 47.1	5.36 ± 1.38	85.5 ± 7.8
Oceania	6	188.4 ± 57.7	5.54 ± 1.31	83.6 ± 07.5
Europe	35	169.5 ± 26.2	6.42 ± 0.50	90.7 ± 9.3
North America & Caribbean	52	192.8 ± 42.4	5.91 ± 1.16	86.6 ± 7.3
Central & South America	38	202.3 ± 37.6	5.64 ± 1.13	86.5 ± 7.7

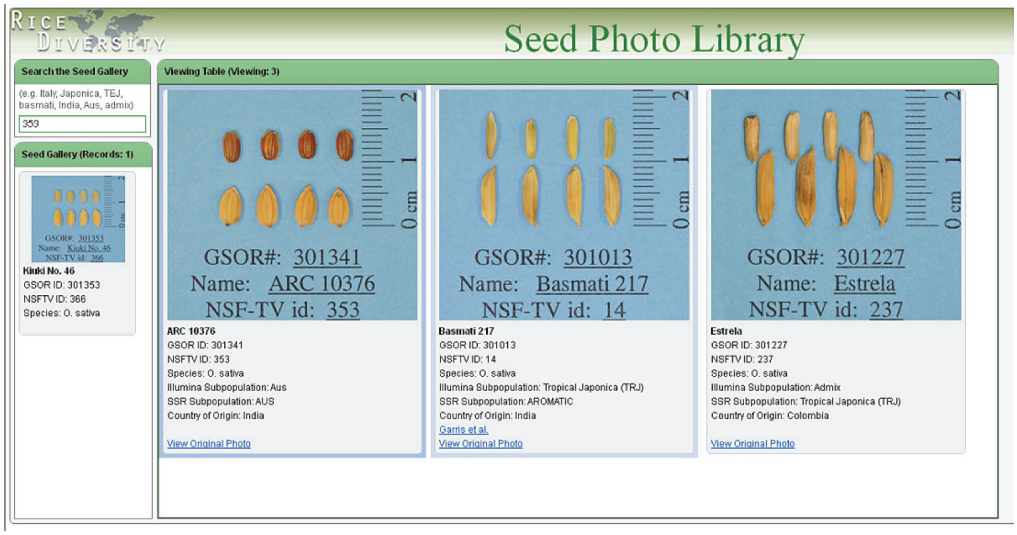

Fig. 2. Seed pictures of the varieties ARC10376, Basmati 217 and Estrela illustrating the differences in seed size as viewed in the Rice Diversity Project Seed Photo Library (http://www.ricediversity.org/photolibrary/).

photolibrary/). A Panicle Photo Library at the Genetic Stocks-*Oryza* (GSOR) Web site (http://www.ars.usda.gov/Main/Docs.htm?docid=18825) provides images of individual panicles for each accession by clicking on the accession number (Fig. 3). These photos will be useful for quality assurance in maintaining the genetic stocks true-to-type, as well as for visualizing seed and panicle characteristics.

Availability

Seed of the Rice Diversity Panel 1 (RDP1) may be obtained from the USDA–ARS GSOR Collection at the USDA–ARS Dale Bumpers National Rice Research Center in Stuttgart, AR. These lines are identified by GSOR accession numbers 301001 through 301421 and GSOR312001 through 312020. The complete list of accessions is found in the GSOR Collection Catalog under Miscellaneous Collections (http://www.ars.usda.gov/Main/docs.htm?docid=18992&page=7) and selecting the NSF Rice Diversity Panel and *Oryza*SNP Set. The GSOR general policy is to distribute five seeds per sample for each accession. Instructions for requesting seed can be found at https://www.ars.usda.gov/Main/site_main.htm?docid=8324). All accessions not bound by intellectual property have been shared with the International Rice Research Institute (IRRI) in the Philippines, where they have been assigned IRGC accession numbers and now form part of the International Rice Germplasm Collection for distribution outside of the United States. We ask that appropriate recognition be given when accessions from the RDP1 and the associated genotypic or phenotypic data are used in future research by citing this article. For use of the SSR or SNP genotypic data associated with these lines, please cite Ali et al. (2011) or Zhao et al. (2011), respectively.

Fig. 3. Panicle pictures of five varieties showing variability in panicle sizes and grain number. Images for each accession may be viewed on the GSOR Web site (http://www.ars.usda.gov/Main/Docs.htm?docid=18825) by clicking on the accession number.

Acknowledgments

This research was supported by National Science Foundation- Plant Genome Project: "The Genetic Basis of Transgressive Variation in Rice" (Award nos. 0606461 and 1026555). We gratefully acknowledge the superb technical assistance of Teresa Hancock, Daniel Wood, and Heather Maupin; the assistance of student interns Patricia Archambeau, Chance Jackson, and T. Grant Brock; USDA–ARS personnel Lorie Bernhardt (GSOR technician), Dr. Wengui Yan (rice germplasm geneticist), and Dr. Harold Bockleman (NSGC curator) for their contributions to managing the rice seed stocks; and Dr. Kenneth McNally, Ma. Elizabeth Naredo, and other staff members at the International Rice Research Institute, Philippines, for their collaboration during validation and quality control of the seed stocks. USDA–ARS scientists Dr. Ming-Hsuan Chen and Dr. Shannon R.M. Pinson are acknowledged for providing excellent reviews of this manuscript.

References

Ali, M.L., A.M. McClung, M.H. Jia, J.A. Kimball, S.R. McCouch, and G.C. Eizenga. 2011. A rice diversity panel evaluated for genetic and agro-morphological diversity between subpopulations and its geographic distribution. Crop Sci. 51:2021–2035. doi:10.2135/cropsci2010.11.0641

AACC International. 2000. Method 46-30: Crude protein-combustion method. In: Approved methods of the American association of cereal chemists. Vol. 2. 10th ed. American Association of Cereal Chemists, Int., St. Paul, MN.

Ayres, N.M., A.M. McClung, P.D. Larkin, H.F.J. Bligh, C.A. Jones, and W.D. Park. 1997. Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germplasm. Theor. Appl. Genet. 94:773–781. doi:10.1007/s001220050477

Bhattacharya, K.R., and C.M. Sowbhagya. 1971. Water uptake by rice during cooking. Cereal Sci. Today 16:420–424.

Bergman, C.J., K.R. Bhattacharya, and K. Ohtsubo. 2004. Rice end-use quality analysis. In: E.T. Champagne, editor, Rice chemistry and technology. Am. Assoc. Cereal Chemists, St. Paul, MN. p. 415–772.

Caicedo, A.L., S.H. Williamson, R.D. Hernandez, A. Boyko, A. Fledel-Alon, T.L. York, N.R. Polato, K.M. Olsen, R. Nielsen, S.R. McCouch, C.D. Bustamante, and M.D. Purugganan. 2007. Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet. 3:e163. doi:10.1371/journal.pgen.0030163

Champagne, E.T., K.L. Bett-Garber, A.M. McClung, and C. Bergman. 2004. Sensory characteristics of diverse rice cultivars as influenced by genetic and environmental factors. Cereal Chem. 81:237–243. doi:10.1094/CCHEM.2004.81.2.237

Champagne, E.T., K.L. Bett-Garber, A.M. McClung, and C.C. Grimm. 2007. Effects of organic fertility on physicochemical properties and sensory quality of diverse rice cultivars. Cereal Chem. 84:320–327. doi:10.1094/CCHEM-84-4-0320

Chen, M.-H., C.J. Bergman, S.R.M. Pinson, and R.G. Fjellstrom. 2008. *Waxy* gene haplotypes: Associations with apparent amylose content and the effect by the environment in an international rice germplasm collection. J. Cereal Sci. 47:536–545. doi:10.1016/j.jcs.2007.06.013

Cho, Y.G., H.J. Kang, Y.T. Lee, S.K. Jong, M.Y. Eun, and S.R. McCouch. 2010. Identification of quantitative trait loci for physical and chemical properties of rice grain. Plant Biotechnol. Rep. 4:61–73.

Chun, A., J. Song, K. Kim, and H. Lee. 2009. Quality of head and chalky rice and deterioration of eating quality by chalky rice. J. Crop Sci. Biotechnol. 12:239–244. doi:10.1007/s12892-009-0142-4

Clark, R.T., A.N. Famoso, K. Zhao, J.E. Shaff, J.E. Craft, C.D. Bustamante, S.R. McCouch, D.J. Aneshansley, and L.V. Kochian. 2013. High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development. Plant Cell Environ. 36:454–466. doi:10.1111/j.1365-3040.2012.02587.x

Clark, R.T., R.B. MacCurdy, J.K. Jung, J.E. Shaff, S.R. McCouch, D.J. Aneshansley, and L.V. Kochian. 2011. Three-dimensional root phenotyping with a novel imaging and software platform. Plant Physiol. 156:455–465. doi:10.1104/pp.110.169102

Famoso, A.N., K. Zhao, R.T. Clark, C.-W. Tung, M.H. Wright, C. Bustamante, L.V. Kochian, and S.R. McCouch. 2011. Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. PLoS Genet. 7(8):E1002221. doi:10.1371/journal.pgen.1002221

Garris, A.J., T.H. Tai, J. Coburn, S. Kresovich, and S. McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. Genetics 169:1631–1638. doi:10.1534/genetics.104.035642

Gross, B.L., M. Reagon, S.H. Hsu, A.L. Caicedo, Y. Jia, and K.M. Olsen. 2010. Seeing red: The origin of pigmentation in US weedy rice. Mol. Ecol. 19:3380–3393. doi:10.1111/j.1365-294X.2010.04707.x

Huang, X., N. Kurata, X. Wei, Z.X. Wang, A. Wang, Q. Zhao, Y. Zhao, K. Liu, H. Lu, W. Li, Y. Guo, Y. Lu, C. Zhou, D. Fan, Q. Weng, C. Zhu, T. Huang, L. Zhang, Y. Wang, L. Feng, H. Furuumi, T. Kubo, T. Miyabayashi, X. Yuan, Q. Xu, G. Dong, Q. Zhan, C. Li, A. Fujiyama, A. Toyoda, T. Lu, Q. Feng, Q. Qian, J. Li, and B. Han. 2012. A map of rice genome variation reveals the origin of cultivated rice. Nature 490:497–501. doi:10.1038/nature11532

Juliano, B.O. 1971. A simplified assay of milled-rice amylose. Cereal Sci. Today 16:334–360.

Juliano, B.O. 1985. Criteria and tests for rice grain qualities. In: B.O. Juliano, editor, Rice chemistry and technology. 2nd ed. Am. Assoc. Cereal Chemists, St. Paul, MN. p. 443–524.

Little, R.R., G.B. Hilder, and E.H. Dawson. 1958. Differential effect of dilute alkali on 25 varieties of milled rice. Cereal Chem. 35:111–126.

Loos, B.P. 1993. Morphological variation in *Lolium* (Poaceae) as a measure of species relationships. Plant Syst. Evol. 188:87–89.

McCouch, S.R., M.J. Kovach, M. Sweeney, H. Jiang, and M. Semon. 2012. The dynamics of rice domestication: A balance between gene flow and genetic isolation. In: P. Gepts, T.R. Famula, R.L. Bettinger, S.B. Brush, A.B. Damania, P.E. McGuire, and C.O. Qualset, editors, Biodiversity in agriculture: Domestication, evolution, and sustainability. Cambridge Univ. Press, New York. p. 311–329.

McNally, K.L., K.L. Childs, R. Bohnert, R.M. Davidson, K. Zhao, V.J. Ulat, G. Zeller, R.M. Clark, D.R. Hoen, T.E. Bureau, R. Stokowski, D.G. Ballinger, K.A. Frazer, D.R. Cox, B. Padhukasahasram, C.D. Bustamante, D. Weigel, D.J. Mackill, R.M. Bruskiewich, G. Rätsch, C.R. Buell, H. Leung, and J.E. Leach. 2009. Genome-wide SNP variation reveals relationships among landraces and modern varieties of rice. Proc. Natl. Acad. Sci. USA 106:12273–12278. doi:10.1073/pnas.0900992106

Norton, G.J., J. Alexander, S. Mckay, H. Hansen, G. Duan, M.R. Islam, S. Islam, J.L. Stroud, F.-J. Zhao, S.P. McGrath, Y. Zhu, B. Lahner, E. Yakubova, M.L. Guerinot, S.R.M. Pinson, L. Tarpley, G.C. Eizenga, D.E. Salt, A.A. Meharg, and A.H. Price. 2012. Variation in grain arsenic assessed in a diverse panel of rice grown in multiple sites. New Phytol. 193:650–664. doi:10.1111/j.1469-8137.2011.03983.x

SAS Institute. 2008. The SAS system for Windows. Release 9.2. SAS Inst., Cary, NC.

Temnykh, S., W.D. Park, N. Ayres, S. Cartinhour, N. Hauck, L. Lipovich, Y.G. Cho, T. Ishii, and S.R. McCouch. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet. 100:697–712. doi:10.1007/s001220051342

Thurber, C.S., M. Reagon, B.L. Gross, K.M. Olsen, Y. Jia, and A.L. Caicedo. 2010. Molecular evolution of shattering loci in U.S. weedy rice. Mol. Ecol. 19:3271–3284. doi:10.1111/j.1365-294X.2010.04708.x

Umemoto, T., and N. Aoki. 2005. Single-nucleotide polymorphisms in rice *starch synthase IIa* that alter starch gelatinization and starch association of the enzyme. Funct. Plant Biol. 32:763–768. doi:10.1071/FP04214

Umemoto, T., N. Aoki, H. Lin, Y. Nakamura, N. Inouchi, Y. Sato, M. Yano, H. Hirabayashi, and S. Maruyama. 2004. Natural variation in rice *starch synthase IIa* affects enzyme and starch properties. Funct. Plant Biol. 31:671–684. doi:10.1071/FP04009

Vaylay, R.E., and E. van Santen. 2002. Application of canonical discriminant analysis for the assessment of genetic variation in tall fescue. Crop Sci. 42:534–539. doi:10.2135/cropsci2002.0534

Williams, V.R., W.-T. Wu, H.Y. Tsai, and H.G. Bates. 1958. Varietal differences in amylose content of rice starch. J. Agric. Food Chem. 6:47–48. doi:10.1021/jf60083a009

Yadav, B.K., and V.K. Jindal. 2008. Changes in head rice yield and whiteness during milling of rough rice (*Oryza sativa* L.). J. Food Eng. 86:113–121. doi:10.1016/j.jfoodeng.2007.09.025

Yan, W.G., J.N. Rutger, R.J. Bryant, H.E. Bockelman, R.G. Fjellstrom, M.G. Chen, T.H. Tai, and A.M. McClung. 2007. Development and evaluation of a core subset of the USDA rice (*Oryza sativa* L.) germplasm collection. Crop Sci. 47:869–878. doi:10.2135/cropsci2006.07.0444

Zhao, K., M. Wright, M. Kovach, A. Reynolds, W. Tyagi, J. Kimball, G. Eizenga, A. McClung, M.L. Ali, C.D. Bustamante, and S. McCouch. 2010. Genomic diversity and introgression in *O. sativa* reveal impact of domestication and breeding on the rice genome. PLoS ONE 5(5):E10780. doi:10.1371/journal.pone.0010780

Zhao, K., C.-W. Tung, G.C. Eizenga, M.H. Wright, M.L. Ali, A.H. Price, G.J. Norton, M.R. Islam, A. Reynolds, J. Mezey, A.M. McClung, C.D. Bustamante, and S.R. McCouch. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nat. Commun. 2:467. doi:10.1038/ncomms1467