Low additive genetic variation in a trait under selection in domesticated rice

Nicholas G. Karavolias[*,†], Anthony J. Greenberg[*,‡], Luz S. Barrero[†,§], Lyza G. Maron[†], Yuxin Shi[†], Eliana Monteverde[†], Miguel A. Piñeros[¶], and Susan R. McCouch[†]

Abstract

Quantitative traits are important targets of both natural and artificial selection. The genetic architecture of these traits and its change during the adaptive process is thus of fundamental interest. The fate of the additive effects of variants underlying a trait receives particular attention because they constitute the genetic variation component that is transferred from parents to offspring and thus governs the response to selection. While estimation of this component of phenotypic variation is challenging, the increasing availability of dense molecular markers puts it within reach. Inbred plant species offer an additional advantage because phenotypes of genetically identical individuals can be measured in replicate. This makes it possible to estimate marker effects separately from the contribution of the genetic background. We focused on root growth in domesticated rice, *Oryza sativa*, under normal and aluminum (Al) stress conditions. This trait has been under long-term selection because root growth correlates with several fitness components. A dense single nucleotide polymorphism (SNP) map is available for all accessions studied. Taking advantage of this map and a set of Bayesian models, we assessed additive marker effects. While total genetic variation accounted for a large proportion of phenotypic variance, marker effects carried little information particularly in the Al tolerant *tropical japonica* population of rice. We were unable to identify any loci associated with root growth in this population. Models estimating the aggregate effects of all measured genotypes likewise produced low estimates of marker heritability and were unable to predict total genetic values accurately. Our results support the long-standing conjecture that additive genetic variation is depleted in traits under selection. We further provide evidence that this depletion is due to the prevalence of low-frequency alleles that underlie the trait.

Introduction

Unraveling the mechanisms of phenotypic evolution under selection is a fundamental task in biology. Among-individual variation in most traits of interest results from a collective action of multiple genes in conjunction with environmental effects (Lynch and Walsh, 1998). To a first approximation, only the component of the genetic variation that is due to the additive contribution of DNA sequence polymorphisms is passed on to the following generations. The fate of this additive genetic variance is thus of particular interest since it is an important factor determining both the speed and direction of adaptation.

Fisher (1930) predicted that selection should deplete this component of phenotypic variance as beneficial alleles become fixed. This intuition has been extensively explored in theoretical studies. Most analyses focus on two scenarios: stabilizing selection, where the population is close to optimum fitness and

[*]These authors contributed equally

[†]Section of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA

[‡]Bayesic Research, Ithaca, NY, USA

[§]Colombian Corporation for Agricultural Research, Agrosavia, Colombia

[¶]Robert W. Holey Center for Agriculture and Health, USDA-ARS, Ithaca, NY, USA

extreme phenotypes are deleterious, and directional selection, where the optimal phenotype is shifted and the population is allowed to evolve towards the new state. While the studies involve different sets of assumptions and employ either analytical or simulation approaches, the general conclusion is that the predictions are very sensitive to model parameters. In some cases (Lande, 1975; Turelli, 1984) mutation-selection balance is expected to maintain a substantial amount of additive genetic variation. In other scenarios, the variance is either reduced or preserved due to segregation of a few large-effect alleles (Turelli, 1984; Barton, 1986; Johnson and Barton, 2005; de Valdar and Barton, 2014; Jain and Stephan, 2017; Stetter *et al.*, 2018). The assumed shape of the mutation effect size distribution appears to play a major role in determining whether substantial additive genetic variation is maintained in the face of selection.

Empirical observations can help us distinguish which theoretical models, and thus genetic mechanisms, are most consistent with data. The difficulty is that we cannot directly observe the components of phenotypic variance, but must rely on statistical models to estimate them. Assuming reliable estimates, the common approach is to compare additive genetic variances in fitness-related (often called life-history) traits to a control group (e.g., seemingly unimportant morphological characteristics). Since measurement scales differ among traits, some scale factor must be used to make comparisons meaningful. A popular scaling factor is the total variance. Its use is well-motivated by theory as the ratio of additive genetic to total variance yields estimates of narrow-sense heritability, a parameter important in predicting outcomes of single-generation selection on phenotypes (Lynch and Walsh, 1998). Most empirical studies (Mousseau and Roff, 1987; Gustafsson, 1986; Kruuk *et al.*, 2000; Johnson and Barton, 2005) find that narrow-sense heritability of traits correlated with fitness is indeed low. However, two traits can have equal levels of additive genetic variance but different heritabilities due to differing amounts of environmental variance (Price and Schluter, 1991). Fisher's prediction applies to the total amounts of additive variation, so arguably estimates of heritability are misleading. There is in fact some evidence (Price and Schluter, 1991; Houle, 1992; Johnson and Barton, 2005) that scaling genetic variances by means reveals that life-history traits, if anything, possess a larger store of additive variation than other traits.

While proper scaling of variance estimates for among-trait comparisons is important, precise estimates of variance magnitudes are required for reliable inference. Observed (natural population pedigrees) or experimental (diallel) crosses have been necessary in the past to partition phenotypic variance into the various genetic and environmental components (Sprague and Tatum, 1942; Lynch and Walsh, 1998; Greenberg *et al.*, 2010). Wide availability of dense DNA sequence markers makes it possible to dispense with crossing and instead fit models that directly incorporate nucleotide polymorphism effects (Meuwissen *et al.*, 2001; Van Raden, 2008). Plant species, as well as some model animals, offer the additional possibility of assaying genetically identical individuals in replicate. This makes it possible to separate the overall genetic contribution into that explained by markers and the rest (we refer to this quasi-residual as the "background" effect). Interpretation of this partitioning is not always clear, however, and depends on the model used for the molecular markers.

Cultivated rice *Oryza sativa* is a selfing species with extensive genomic resources, making it a useful model system to test quantitative genetic theory. It is a major food crop that is mostly grown under irrigated or lowland rain-fed conditions, where standing water and anaerobic soil conditions are at least intermittently present (Khush, 1997; Bernier *et al.*, 2008). However, a minority of area under cultivation is occupied by rice grown under upland (aerobic) rain-fed conditions, characterized by absence of standing water and frequent droughts (De Datta *et al.*, 1975; Khush, 1997; Bernier *et al.*, 2008). Since initial rice domestication is associated with paddy rice (Oka, 1988; Fuller *et al.*, 2010), this arid growth regime has likely imposed a recent and strong selection for survival in the face of water scarcity (De Datta, 1975). An additional difficulty for upland rice farmers is the prevalence of acidic soils (von Uexküll and Mutert, 1995) that promote metal toxicity (e.g., aluminum) and nutrient deficiency (e.g., phospho-

rus). These problems, in turn, exacerbate drought damage. Thus, rice accessions adapted to upland conditions have been under additional selection pressure for tolerance to aluminum toxicity and nutrient deficiency.

Root architecture is an important determinant of drought tolerance in rice. Accessions successfully grown under rain-fed (rather than irrigated) conditions tend to have relatively long and thick roots (De Datta *et al.*, 1975; Champoux *et al.*, 1995; Bernier *et al.*, 2008). There is evidence of a genetic correlation between root length, survival, and grain yield under arid conditions (Champoux *et al.*, 1995; Bernier *et al.*, 2009; Cairns *et al.*, 2009; Lyu *et al.*, 2014). Aluminum tolerance is also assessed by comparing root length of treated and control plants (von Uexküll and Mutert, 1995; Famoso *et al.*, 2010). While rice is relatively aluminum tolerant overall, *japonica* cultivars are particularly resistant (Famoso *et al.*, 2010, 2011). Accessions from this population are often grown on acidic, upland, rain-fed soils in Africa, Latin America, and South-East Asia (Khush, 1997; Lyu *et al.*, 2014). Root length, including under aluminum treatment, is thus a good candidate for a character undergoing strong directional selection. Examining its genetic architecture thus may shed light on the evolution of polygenic traits. In addition, any information gained can be used to explore the adaptive potential of rice root length in water-limited environments.

Using a world-wide panel of *O. sativa* accessions and a dense set of single-nucleotide polymorphisms (SNP), Famoso *et al.* (2011) performed a genome-wide association study (GWAS) to look for associations between genotyped loci and relative root growth under aluminum treatment. While they found associations in the *aus* and *indica* subpopulations, they did not detect any signal in *tropical japonica*. To determine if this result reflects a depletion of additive genetic variation under selection, we set out to build on this work. We first recapitulated the previous analyses using a new, ten-fold denser, genotyping platform (McCouch *et al.*, 2016). We also introduced a more robust approach to modeling replicated data. Despite these improvements, we still failed to find any GWA signal in *tropical japonica*. Adding more accessions, increasing the duration of aluminum treatment, and measurement of additional traits did not produce a higher power to detect loci associated with aluminum tolerance. Furthermore, while total genetic variation accounted for a large portion of overall phenotypic variance, additive genetic variation captured by SNP genotypes was consistently low. We provide evidence that this is due to a low frequency of alleles conferring susceptibility or tolerance. These results are consistent with a set of theoretical expectations and allow us to deduce a possible genetic architecture of root growth in *tropical japonica* varieties adapted to upland cultivation. Our inferences provide testable predictions for future experiments and point to breeding strategies that are more likely to improve resistance to drought and cultivation on acid soils in this population.

Materials and Methods

Root measurements and accession sets

We expanded the *tropical japonica* accession set analyzed by Famoso *et al.* (2011) by adding 77 new accessions to the previously-measured 92. These were treated and measured the same protocol as before (Famoso *et al.*, 2010, 2011). We also chose a new set of 119 *tropical japonica* accessions (including some overlap with the original 169). The Al treatment in the latter experiment was conducted as before, but the duration was extended to six days. The root systems of seedlings were imaged and total and longest root length measured using RootReader 2D, a digital imaging and quantification system (Clark *et al.*, 2013).

Multi-trait modeling of replicated experiments

We built hierarchical Bayesian models (Gelman *et al.*, 2004; Greenberg *et al.*, 2010, 2011) to estimate parameters from our replicated experiments. The detailed models are listed in the documentation provided in the Supplemental File 1. This supplemental file also includes the raw phenotypic data, scripts we used to manage the raw data, Markov chain conver-

gence diagnostics, and post-processing and plotting of the results. The analyses were implemented in C++. The source code is included in the Supplemental File 1. The underlying library is also available from `https://github.com/tonymugen/MuGen`. Briefly, the location parameters of the model used to re-analyze data from Famoso *et al.* (2011) are

$$\begin{aligned}
\boldsymbol{y}_{i\cdot} &\sim t_{\nu_e,d}\left(\boldsymbol{\mu}^{acc}_{j[i]\cdot} + \boldsymbol{x}^{\text{EXP}}_{i\cdot}\boldsymbol{B}^{\text{EXP}}; \boldsymbol{\Sigma}_e\right) \\
\boldsymbol{\mu}^{acc}_{j\cdot} &\sim \text{N}_d\left(\boldsymbol{\mu} + \boldsymbol{u}_{j\cdot}\boldsymbol{\Gamma}; \boldsymbol{\Sigma}_s\right) \\
\boldsymbol{\gamma}_{j\cdot} &\sim t_{\nu_g,d}\left(\boldsymbol{0}_d; \boldsymbol{\Sigma}_a\right)
\end{aligned}$$

The location parameter vectors of the form $\boldsymbol{y_{i\cdot}}$ are row-vectors of the corresponding parameter matrices that have data points as rows and traits as columns. The models for Famoso *et al.* (2011) and additional *tropical japonica* accessions have the treated and control total root lengths as the two traits. In these cases, the error covariance matrix is diagonal with error variances and no between-residual correlations because the root lengths were measured on separate plants. All other covariance matrices $\boldsymbol{\Sigma_x}$ include covariances. The model for the third experiment analyses the treated and control data sets separately. In this case the two traits are length of the total root system and that of the longest root. Since these were measured on the same plant, the error covariances are non-zero. The corrections for the year effects were done at the level of accessions in the second experiment. The third experiment was performed in one batch and therefore no such correction was required. The errors were modeled using a multivariate Student-t distribution with three degrees of freedom to dampen the effects of outliers (Greenberg *et al.*, 2010, 2011) for the last experiment. Our set-up did not allow for Student-t models for block-diagonal error matrices. Therefore, the first two data sets were analyzed with Gaussian models for errors. The marker effects enter through the eigenvectors $\boldsymbol{U}$ of the relationship matrix, weighted by square roots of their eigenvalues. They are also modeled using a multivariate Student-t distribution. This is equivalent to a multivariate version of the BayesA model (Meuwissen *et al.*, 2001). The results presented in the main text were obtained by setting the degrees of freedom to three (the minimal number that still results in defined moments of the Student-t). We repeated all analyses with the degrees of freedom set to 1000 (making the prior close to Gaussian) with the same results.

We used the posterior modes of accession means as predictors in GWA, with mixed-model correction for population structure (Kang *et al.*, 2010) using our own R package (Greenberg, 2019), available from `https://github.com/tonymugen/GWAlikeMeth`.

Genome prediction

We used the posterior modes of accession means (see above) to fit a set of genome prediction models. The general form is

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{1}_N\mu + \boldsymbol{u} + \boldsymbol{e} \\
\boldsymbol{e} &\sim \text{N}\left(\boldsymbol{0}; \boldsymbol{I}\sigma^2_e\right) \\
\boldsymbol{u} &\sim \text{N}\left(\boldsymbol{0}; \boldsymbol{K}\sigma^2_u\right),
\end{aligned}$$

where $\boldsymbol{y}$ is the response vector of posterior accession modes, $\boldsymbol{1}_N$ is a column vector of ones of order, μ is the overall mean, $\boldsymbol{u}$ is the column vector of random genetic values, σ^2_u is the among-$\boldsymbol{u}$ variance, $\boldsymbol{K}$ is the among-genotype covariance matrix, $\boldsymbol{e}$ is the column vector of random errors, and σ^2_e is the error variance.

To verify results obtained using the hierarchical Bayesian model outlined above, we used a linear kernel $\boldsymbol{K} = \boldsymbol{X}\boldsymbol{X}^{\text{T}}/p$, where $\boldsymbol{X}$ is the centered and standardized matrix of p genotypes (originally coded as -1, 0, and 1). This is the genome best linear unbiased predictor (GBLUP) model of additive effects. To test whether modeling among-locus epistatic interactions improves prediction, we repeated the analyses with a Gaussian kernel for $\boldsymbol{K}$, setting its elements to $\boldsymbol{K}(x_j, x_j) = \exp\left(-hd^2_{ij}\right)$. Here, d_{ij} is the Euclidian distance between genotypes i and j, h is the bandwidth parameter that controls the speed of decay of $\boldsymbol{K}$ elements with genetic distance. Following Crossa *et al.* (2010), we set $h = 2d^{-2}_m$, where d_m is the sample median among all d_ij.

We used the R package Bayesian generalized linear regression (BGLR, Pérez and de los Campos, 2014) to fit the models. The package imple-

ments a Gibbs sampler. We ran 55,000 iterations including 5,000 burn-in steps and 5-fold thinning. We assessed prediction accuracy using 50 training-validation random partitions. Pearson's correlations (r) between the predicted and the observed values were employed as accuracy measurements. To calculate summary statistics, we transformed raw correlation values using Fisher's z transformation ($z = \frac{1}{2}\left[\ln(1+r) - \ln(1-r)\right]$). The mean and interval values presented in Table 1 were back-transformed to the original scale.

Introgression mapping

We used RFMix (Maples *et al.*, 2013) to identify introgressions from *aus* and *indica* in the six admixed *tropical japonica* accessions. We used 10 training accessions from each of the three populations of interest (see the Supplemental File 2 for the full list). We ran Beagle v4.1 (Browning and Browning, 2016) with default parameters to phase genotypes, after filtering loci with minor allele frequencies below 0.08 and more than 30% missing genotypes. Our final data set included 304,488 SNPs. We then ran RFMix with the option "TrioPhased" on the phased genotypes from 36 lines: three sets of 10 training pure accessions plus the six admixed accessions. The results were plotted using an R script included in the Supplemental File 1.

Data and reagent availability statement

Supplemental files are available at FigShare. Supplemental File 1 contains detailed models, raw phenotypic data, scripts to manage the raw data, Markov chain convergence diagnostics, post-processing and plotting of the results, and source code. The raw phenotype data include those generated by Famoso *et al.* (2011). Genotype data are from McCouch *et al.* (2016) and available at `ricediversity.org/data/index.cfm`. Supplemental File 2 lists the training accessions we used for admixture mapping.

Results

The initial study (Famoso *et al.*, 2011) only examined relative root growth between Al treated and control genotypes. To measure Al tolerance, separate individuals from the same accession are subjected to aluminum treatment and compared to plants grown under control conditions. The relative root growth is then estimated by calculating the ratio of the means among the treated and control plants within each genotype. Partitioning the phenotypic variance of this compound trait into genetic and environmental components is thus impossible. Therefore, we focused on estimating genetic parameters for root length itself under the two experimental regimes. This approach also fits better with results from previous studies that suggest a role for root length in water stress tolerance. The data are more sparse and ambiguous for relative root growth.

To establish that our analyses can recapitulate previous results, we started with the original Famoso *et al.* (2011) data set. It is possible that markers in the original study (44,000 SNPs) were not dense enough. Some causal polymorphisms may have existed in the *tropical japonica* data set, but they were not closely linked to any assayed SNPs. In addition, the experiments were performed across several years. While there is a clear year effect, and population representation among experiments is not even (Supplemental Fig. 1), this was not accounted for in the initial analyses (Famoso *et al.*, 2011). Therefore, we took advantage of the high density rice array (HDRA, 700,000 SNPs, McCouch *et al.*, 2016) that already includes genotypes for the accessions used in the original experiment. We also built a Bayesian hierarchical model (Gelman *et al.*, 2004) to account for the replication and year effect. We considered root length under control and aluminum treatment conditions as separate traits and used a multivariate hierarchical model (Greenberg *et al.*, 2011, see Methods for details) that included a SNP-derived relationship matrix to account for marker effects. Replication of identical genotypes, possible because rice is a selfing species with almost completely homozygous accessions (McCouch *et al.*, 2016), allowed us to separate additive marker-based variance from the total

genetic variation and residual environmental effects. We used the posterior modes of accession effects (as well as $\log_{10}$ of the control to treated ratio, $-\lg RRG$) as predictors in a standard single-trait mixed-model GWA (Kang *et al.*, 2010, see Methods for details).

We were able to reproduce the GWA peak Famoso *et al.* (2011) found for untransformed *RRG* in the *aus* subpopulation (Fig. 1A), but did not detect any new associations. We still did not unearth any associations in *tropical japonica* (Fig. 1B), and no associations of SNPs with root length traits under Al treated or control conditions in either population. Furthermore, we find that the heritability contributed by all markers considered together accounts for less than half of overall genetic variance in the *japonica* population (Fig. 2A). This does not seem to be due to a lack of power resulting from a small sample size, since marker heritability is higher in *aus* under stress (Fig. 2A) despite a smaller number of accessions in that population (55 in *aus vs* 92 in *tropical japonica*). Nevertheless, we opted to directly address the potential sample size limitation by adding more accessions to our *tropical japonica* sample, bringing the total to 169. These lines were already genotyped using HDRA (McCouch *et al.*, 2016). We administered the same treatment protocol as Famoso *et al.* (2011). The data were analyzed as described above. Despite the increased power, we did not find any significant GWA peaks (Supplemental Fig. 3A). While the measurements of each accession were highly reproducible, yielding high estimates of broad-sense heritability, the proportion of this genetic variation accounted for by the sum of marker effects was much smaller than in the initial limited sample (Fig. 2B).

Although the increased sample size made us more confident that the limited number of accessions is not the reason our estimates of genome estimated heritability was low, we still had two concerns with our experiment. First, because we added lines to a preexisting data set, most accessions were measured only in one year (albeit with replication). Thus, the year was confounded with the line effect and thus genotypes were also partially confounded with year. We thus performed the year correction at the accession level in our model (see Methods for details). This still leaves the possibility that the adjustment leads to an incorrect estimate of the genome-estimated breeding values (GEBV). Second, we noticed that the total genetic correlation between root length under control and aluminum treated conditions was very high (mode: 0.886; 95% highest posterior density (HPD) interval: [0.803, 0.945]). This implies that there is no differential response among *tropical japonica* accessions to our aluminum treatment regime, resulting in very little gene by environment interaction. This is not surprising, since the treatment was originally developed for all *O. sativa* populations and *tropical japonica* is the most resistant group (Famoso *et al.*, 2010).

To address these limitations, we selected a new panel of 119 *tropical japonica* accessions that represented the full diversity of the representatives of this population genotyped using HDRA (Supplemental Fig. 2). In addition, we performed more replications per line (eight on average). We performed all experiments at the same time, eliminating the year effect as a potential confounding variable. We measured two root traits: total root and longest root length (see Methods for details). We also increased the duration of aluminum treatment to put extra stress on the plants in an effort to identify genotypes that perform differently in the presence and absence of aluminum stress. We analyzed the data the same way as before but considered the treated and control data sets separately because replication numbers were not the same between treatment categories for some accessions. The two root traits were modeled together.

The longer treatment did reduce the genetic relationships between treatment and control, although the correlations remained fairly high: 0.770 (0.733, 0.811) for total root length (the trait measured in the previous set of experiments) and 0.816 (0.777, 0.844) for longest root length. We still did not find any peaks approaching significance with false discovery rate ≤ 0.1 (Supplemental Fig. 3B). Marker heritability likewise remained low despite the high fraction of the total phenotypic variance explained by accession effects (Fig. 2C).

These results give us confidence that small sample sizes, biased accession selection, or inadequate treatment regimes are not the primary explanations for the small fraction of phenotypic variance accounted

for by additive effects of SNPs. However, artifacts of model fitting may result in an over-shrinking of marker effects and thus lead to their underestimation. Lack of individual GWA peaks is evidence against this possibility but is not conclusive because a sum of small (and thus unidentifiable by GWA) SNP contributions can still lead to relatively high marker heritability. We therefore took an additional approach to assessing the predictive power of SNPs. We used the posterior modes of accession means generated in the last experiment described above. Given the high total heritability, we are fairly confident in the robustness of these estimates. We also calculated the negative log of relative root growth ($-\lg RRG$) for both root traits. We separated the data into two sets. The first set was used to train a model that used all markers in a linear penalized regression (Meuwissen *et al.*, 2001; Pérez and de los Campos, 2014). We used a separate model fitting implementation (Pérez and de los Campos, 2014) as an additional check on our previous results. Given the model parameter estimates from the training set, we attempted to predict the remaining accession values. We repeated this process 50 times and generated empirical distributions of prediction accuracy (see Methods for details). Consistent with the low additive marker heritability estimates presented above, genome-enabled prediction accuracy was low, with most distributions containing zero (Table 1).

All lines of evidence presented so far point to the conclusion that additive marker effects do not explain much of the total genetic variance of the root growth traits measured in *tropical japonica*. Given the high broad-sense heritability, we wondered if non-additive SNP effects can explain our results. Given the very high homozygosity of individual lines (McCouch *et al.*, 2016), dominance is highly unlikely to play a measurable role. We are therefore left with epistatic interactions among loci as the remaining possibility. However, the parameter space of all interactions is prohibitively large, making it statistically and computationally unfeasible to estimate individual effects. However, models that rely on summaries of interactions are available. It has been shown that using the Hadamard (element-wise) product of the SNP-based relationship matrix (Van Raden, 2008) with itself in a genomic model gives approximate estimates of the sum of all pairwise epistatic interactions (Jiang and Reif, 2015). Similarly, use of a Gaussian kernel (see Methods) approximates all by all interactions (Jiang and Reif, 2015). We used the quadratic in place of the additive kernel in our hierarchical multi-trait model and the Gaussian kernel for genome prediction. We did not observe any measurable improvement in explanatory ability (Supplemental Fig. 4) or prediction accuracy (Table 1).

Since we cannot find any evidence that epistatic interactions account for much genetic variance, we are left with the possibility that low-frequency DNA sequence variants underlie our root traits. Because such alleles are not shared widely among accessions in our panel, and are not in high enough linkage disequilibrium with the SNPs we genotyped, we do not detect their effects in our models. If this conjecture is correct, we would expect many regions of the genome to make small contributions to the genetic variance. We see some evidence of this when we look at accessions that carry some introgressions from the less Al tolerant *indica* and *aus* populations into the *tropical japonica* background (Fig. 3). We found six such lines in the Famoso *et al.* (2011) data set. All of them have shorter roots than the average *tropical japonica* accession in the sample, particularly under aluminum stress (Fig. 3A,B). When we looked at the genomic regions that come from *indica* or *aus* in these lines using RFMix (Maples *et al.*, 2013) we found an idiosyncratic collection of introgressed loci (Fig. 3C). This is expected if each *tropical japonica* accession carries a different set of rare causal variants.

Discussion

Informative tests of the effects of selection on the genetic architecture of quantitative traits require partitioning of the total genetic variation into components. This can be achieved in selfing plants, *O. sativa* among them, because replicated measurements can be performed on genetically identical individuals. Relying on genotyping rather than diallel crosses saves on phenotyping costs. It also allows us to directly model additive and various non-additive effects

of variants in our genotyping panel as well as loci that were not assayed but are in LD with those we assayed.

We employed this strategy to study the genetic architecture of root length in the *tropical japonica* population of rice. A significant fraction of accessions from this group are adapted to upland rain-fed conditions (Khush, 1997; Lyu *et al.*, 2014) where root length is under strong selection because it underlies drought and nutrient toxicity tolerance (De Datta, 1975; Champoux *et al.*, 1995; Bernier *et al.*, 2009; Cairns *et al.*, 2009; Lyu *et al.*, 2014). We find high broad-sense heritability using two experimental setups and three sets of accessions. However, we find no individual SNP effects that approach genome-wide significance in *tropical japonica*. Aggregate contributions of all genotyped loci, assessed using relationship matrices in the context of Bayesian hierarchical models, are also very small. This is in contrast to the situation in the *aus* population, where we see moderate marker heritability overall and a significant effect of the *Nrat1* locus despite a smaller sample of accessions. Alternative modeling approaches, including genome-enabled prediction, yielded the same results.

Given at best a weak signal from additive SNP effects, we wondered if non-additive contributions can account for more of the genetic variance. Because inbred (pure line) rice accessions are highly homozygous (McCouch *et al.*, 2016), epistatic interactions can be the only source of non-additive variance. Since the number of such interactions is much larger than the number of individual effects, we cannot model individual epistatic events with any confidence. However, aggregate contributions can be approximated by building relationship matrix analogs that account for interactions (Jiang and Reif, 2015). Using these approaches did not increase our estimates of marker effect contributions or prediction accuracy of overall genotypic values. The methods to assess the total contribution of non-additive effects are still in development. It is possible that better approaches would uncover some epistatic effects we missed in the current analyses. Given the current evidence and state of the art modeling techniques, however, we did not find any measurable contribution of densely genotyped SNPs to the overall genetic variation in root length in *tropical japonica* under control or aluminum stress conditions.

We are then left with the possibility that the alleles underlying these traits are at low frequency. The distribution of minor allele frequencies of the SNPs genotyped by our array is skewed towards the low prevalence compared to the neutral expectation (Supplemental Fig. 5). These variants are thus unlikely to be in linkage disequilibrium with ungenotyped low-frequency causal polymorphisms and would not account for any phenotypic variance. A genetic architecture dominated by small-effect loci with low-frequency alleles is predicted by models of directional selection on quantitative traits (Jain and Stephan, 2017; Stetter *et al.*, 2018). The models envision a shift in the optimal phenotype value, followed by a change in the adaptive direction and finally stabilizing selection at the new optimum. This scenario fits with our understanding of the role of root architecture in the relatively recent adaptation of rice to cultivation under the arid upland rain-fed conditions (De Datta *et al.*, 1975; Oka, 1988; Khush, 1997; Fuller *et al.*, 2010). Genomic analyses of shifts in cultivation practices of domesticated plants may thus be a promising avenue for empirical studies of genetic architectures of traits under selection.

If our conjecture is correct that small-effect low-frequency alleles underlie root traits in *tropical japonica*, each accession should have a unique set of causal variants. Some would have a positive effect, some would be negative. The balance would then determine the overall genotypic value of the individual line. This, in turn, implies that most regions of the genome can contribute to phenotypic variance. We indeed see some evidence of this: introgressions from *indica* and *aus* populations reduce total root length of *tropical japonica* accessions, particularly under aluminum stress (Fig. 3), and the introgressions are evenly distributed throughout the genome. Conversely, Famoso *et al.* (2011) report that introgressions from *tropical japonica* into *indica* increase relative root length under aluminum stress, consistent with our hypothesis.

Although the low-frequency alleles may not individually contribute much to within-population variance, they should be detectable in biparental crosses. This is because allele frequencies of the causal variants present in these lines will be 50% and thus at

high enough frequency to measure. Indeed, multiple quantitative trait locus mapping studies found loci contributing to root length under various conditions (Champoux *et al.*, 1995; Bernier *et al.*, 2009; Famoso *et al.*, 2011; Arbelaez *et al.*, 2017). However, subsequent attempts to introgress favorable alleles into a variety of genetic backgrounds have been largely unsuccessful (Bernier *et al.*, 2008). This is expected under our model, since every pair of parents should have a different set of causal loci. Furthermore, the predicted presence of at least some negative-effect alleles even in the high-value lines (Stetter *et al.*, 2018) should lead to transgressive segregation in the progeny. This is indeed often observed (Champoux *et al.*, 1995; Arbelaez *et al.*, 2017).

Our results shed some light on a fundamental question in quantitative genetics. They also should be taken into consideration when designing breeding strategies to improve productivity of rice grown in aerobic soils under water-limited conditions. The upland cultivation method has been historically low-yielding (De Datta *et al.*, 1975; De Datta, 1975; Bernier *et al.*, 2008) and has resisted efforts at genetic improvement (Bernier *et al.*, 2008). If our model is correct, progress is possible as long as the number of founding accessions is not too high. All causal alleles would be at moderate frequency and thus visible to selection. Total genetic, rather than breeding, values of the founders should be used to select them from the base population. GWAS, marker assisted selection, and even multi-parental mapping populations (Yu *et al.*, 2008) are not expected to be effective to determine the best founders. However, genome prediction can still be used if the focus is within breeding populations. Finally, periodic addition of new founders can introduce additional beneficial alleles even if the new lines themselves are not top performers in the environments of interest.

Our study highlights the importance of detailed genome-enabled investigations of polygenic traits under selection to shed light on fundamental evolutionary questions and guide practical artificial selection decisions. We hope that it will stimulate further analyses of such traits in agriculturally important plants, taking advantage of sophisticated experimental designs and increasing availability of deep whole-genome DNA sequence data.

Acknowledgments

We thank Sandra Harrington for seed and accession maintenance, Eric Craft for technical assistance with Al treatment experiments, and Elsa de Becker for her assistance with setting up the GWAS pipeline and training NK on it. This work was supported by a grants from NSF-PGRP (award #1444511) and USDA-NIFA (award #2013-67013-21379) and salary support for LSB from Agrosavia in Colombia.

References

Arbelaez, J. D., L. G. Maron, T. O. Jobe, M. A. Pinẽros, A. N. Famoso, *et al.*, 2017 ALUMINUM RESISTANCE TRANSCRIPTION FACTOR 1 (ART1) contributes to natural variation in aluminum resistance in diverse genetic backgrounds of rice (*O. sativa*). Plant Direct **1**: 1–19.

Barton, N. H., 1986 The maintenance of polygenic variation through a balance between mutation and stabilizing selection. Genet. Res. **47**: 209–216.

Bernier, J., G. N. Atlin, A. Kumar, and D. Spaner, 2008 Breeding upland rice for drought resistance. J. Sci. Food Agric. **88**: 927–939.

Bernier, J., R. Serraj, A. Kumar, R. Venuprasad, S. Impa, *et al.*, 2009 The large-effect drought-resistance QTL *qtl12.1* increases water uptake in upland rice. Field Crops Res. **110**: 139–146.

Browning, B. L. and S. R. Browning, 2016 Genotype imputation with millions of reference samples. Am. J. Hum. Genet. **98**: 116–126.

Cairns, J., A. Audebert, C. Mullins, and A. Price, 2009 Mapping quantitative trait loci associated with root growth in upland rice (*Oryza sativa* l.) exposed to soil water-deficit in fields with contrasting soil properties. Field Crop. Res. **114**: 108–118.

Champoux, M. C., G. Wang, S. Sarkarung, D. J. Mackill, J. C. O'Toole, *et al.*, 1995 Locating genes

associated with root morphology and drought avoidance in rice via linkage to molecular markers. Theor. Appl. Genet. **90**: 969–981.

Clark, R. T., A. N. Famoso, K. Zhao, J. E. Shaff, E. J. Craft, *et al.*, 2013 High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development. Plant Cell Environ. **36**: 454–466.

Crossa, J., G. de los Campos, P. Pérez, D. Gianola, J. Burgueño, *et al.*, 2010 Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics **186**: 713–724.

De Datta, S. K., 1975 Upland rice around the world. In *Major Research in Upland Rice*, pp. 2–11, The International Rice Research Institute, Los Baños, Laguna, Philippines.

De Datta, S. K., T.-T. Chang, and S. Yoshida, 1975 Drought tolerance in upland rice. In *Major Research in Upland Rice*, pp. 101–116, The International Rice Research Institute, Los Baños, Laguna, Philippines.

de Valdar, H. P. and N. H. Barton, 2014 Stability and response of polygenic traits to stabilizing selection and mutation. Genetics **197**: 749–767.

Famoso, A. N., R. T. Clark, J. E. Shaff, E. Craft, S. R. McCouch, *et al.*, 2010 Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal aluminum tolerance and investigations into rice aluminum tolerance mechanisms. Plant Physiol. **153**: 1678.

Famoso, A. N., K. Zhao, R. T. Clark, C.-W. Tung, M. H. Wright, *et al.*, 2011 Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. PLoS Genet. **7**: e1002221.

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon, Oxford, UK.

Fuller, D. Q., Y.-I. Sato, C. Castillo, L. Qin, A. R. Weisskopf, *et al.*, 2010 Consilience of genetics and archaeobotany in the entangled history of rice. Archaeol. Anthropol. Sci. **2**: 115–131.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004 *Bayesian Data Analysis*. CRC Press, London, second edition.

Greenberg, A. J., 2019 GWAlikeMeth: mixed models for genome-wide associtations (doi: `10.5281/zenodo.2658787`).

Greenberg, A. J., S. R. Hackett, L. G. Harshman, and A. G. Clark, 2010 A hierarchical Bayesian model for a novel sparse partial diallel crossing design. Genetics **185**: 361–373.

Greenberg, A. J., S. R. Hackett, L. G. Harshman, and A. G. Clark, 2011 Environmental and genetic perturbations reveal different networks of metabolic regulation. Mol. Syst. Biol. **7**: 563.

Gustafsson, L., 1986 Lifetime reproductive success and heritability: Empirical support for Fisher's fundamental theorem. Am. Nat. **128**: 761–764.

Houle, D., 1992 Comparing evolvability and variability of quantitative traits. Genetics **130**: 195–204.

Jain, K. and W. Stephan, 2017 Rapid adaptation of a polygenic trait after a sudden environmental shift. Genetics **206**: 389–406.

Jiang, Y. and J. Reif, 2015 Modeling epistasis in genomic selection. Genetics **201**: 759–768.

Johnson, T. and N. H. Barton, 2005 Theoretical models of selection and mutation on quantitative traits. Phil. Trans. R. Soc. B **360**: 1411–1425.

Kang, H. M., J. H. Sul, S. K. Service, N. A. Zaitlen, S.-Y. Kong, *et al.*, 2010 Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. **42**: 348–354.

Khush, G. S., 1997 Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. **35**: 25–34.

Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, and S. Brotherstone, 2000 Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci. USA **97**: 698–703.

Lande, R., 1975 The maintenance of genetic variation by mutation in a polygenic character with linked loci. Genet. Res. **26**: 221–235.

Lynch, M. and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

Lyu, J., B. Li, W. He, S. Zhang, Z. Gou, *et al.*, 2014 A genomic perspective on the important genetic mechanisms of upland adaptation of rice. BMC Plant Biol. **14**: 160.

Maples, B., S. Gravel, E. Kenny, and C. Bustamante, 2013 RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. Am. J. Hum. Genet. **93**: 278–288.

McCouch, S. R., M. H. Wright, C.-W. Tung, L. G. Maron, K. L. McNally, *et al.*, 2016 Open access resources for genome-wide association mapping in rice. Nat. Commun. **7**: 10532.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157**: 1819–1829.

Mousseau, T. A. and D. A. Roff, 1987 Natural selection and the heritability of fitness components. Heredity **59**: 181–197.

Oka, H.-I., 1988 *Origin of cultivated rice*. Japan Scentific Societies Press and Elsevier, Tokyo, Amsterdam.

Pérez, P. and G. de los Campos, 2014 Genome-wide regression and prediction with the BGLR statistical package. Genetics **198**: 483–495.

Price, T. and D. Schluter, 1991 On the low heritability of life-history traits. Evolution **45**: 853–861.

Sprague, G. F. and L. A. Tatum, 1942 General vs. specific combining ability in single crosses of corn. J. Amer. Soc. Agron. **34**: 923–932.

Stetter, M., K. Thornton, and J. Ross-Ibarra, 2018 Genetic architecture and selective sweeps after polygenic adaptation to distant trait optima. PLoS Genet. **14**: e1007794.

Turelli, M., 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Pop. Biol. **25**: 138–193.

Van Raden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci. **91**: 4414–4423.

von Uexküll, H. and E. Mutert, 1995 Global extent, development and economic impact of acid soils. Plant Soil **171**: 1–15.

Yu, J., J. B. Holland, M. D. McMullen, and E. S. Buckler, 2008 Genetic design and statistical power of nested association mapping in maize. Genetics **178**: 539–551.

Table 1: **Genome prediction accuracy.** Mean (lower 95%, upper 95% confidence interval).

Trait name	GBLUP	RKHS
Selected root length, Al treated	0.034 (-0.147, 0.214)	0.130 (-0.052, 0.304)
Total root length, Al treated	0.118 (-0.065, 0.292)	0.102 (-0.081, 0.277)
Selected root length, control	0.138 (-0.043, 0.311)	0.025 (-0.156, 0.205)
Total root length, control	0.036 (-0.146, 0.216)	-0.205 (-0.372, -0.025)
Selected root length, lg(RRG)	-0.223 (-0.388, -0.044)	-0.200 (-0.368, -0.020)
Total root length, lg(RRG)	-0.194 (-0.362, -0.014)	0.0385 (-0.143, 0.218)

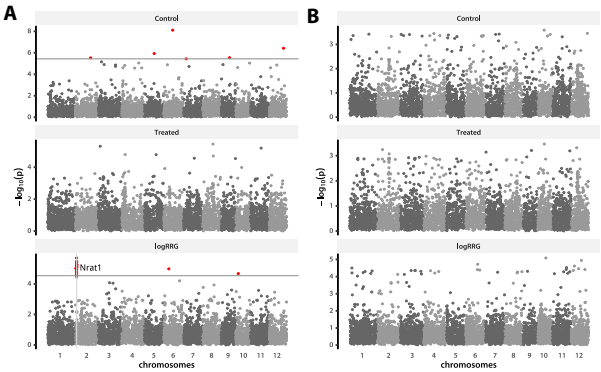


Figure 1: **Genome-wide associations in the aus and tropical japonica populations.** The Manhattan plots depict single-SNP associations (see Methods) of root length under control and Al treated conditions, as well as the logarithm of the treated to control ratio. The horizontal lines mark the $-\lg(p)$ values corresponding to the FDR cut-off at 0.1. (A) Associations within the *aus* population, with the *Nrat1* locus marked on the lower panel; (B) associations within *tropical japonica*.

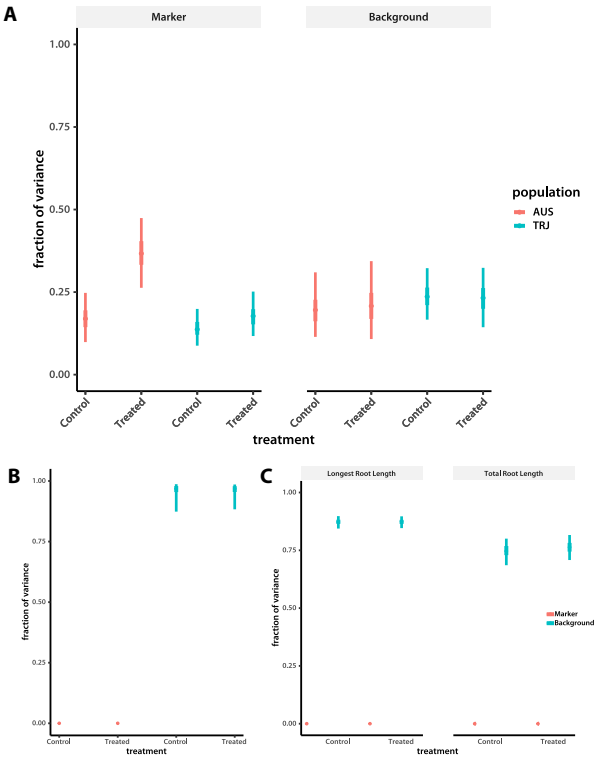


Figure 2: **Fraction of variance explained by markers and genetic background.** The plots depict regions of highest posterior density (HPD; thin lines are 95% and thick lines are 50% intervals, the points are modes) of separate estimates of variance fractions explained by markers and the rest of the genetic background. (A) Results of the re-analysis of the (Famoso *et al.*, 2011) data. (B) The larger data set of *tropical japonica* accessions. (C) Tropical *japonica* accessions subjected to new measurements and experimental conditions.

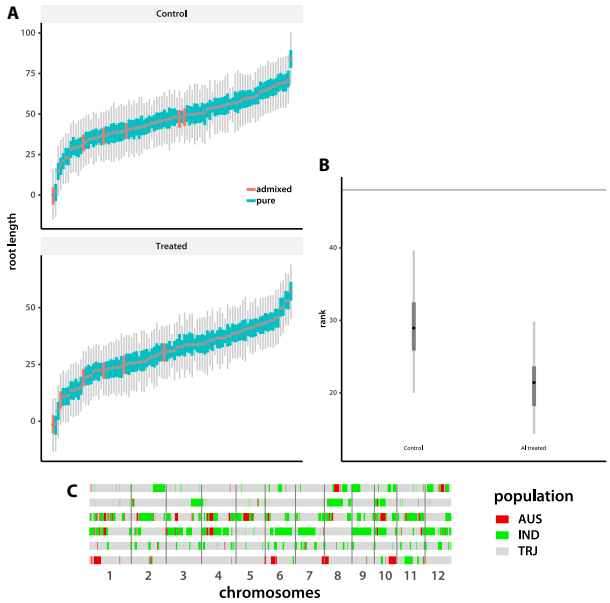
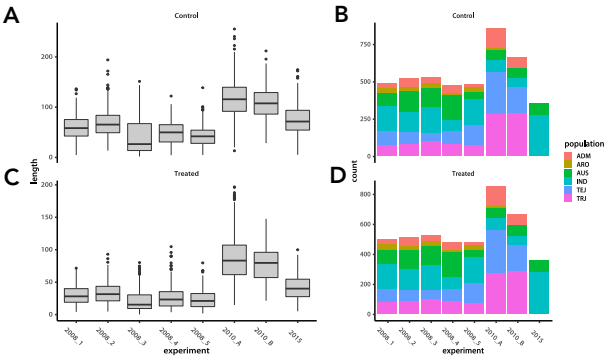
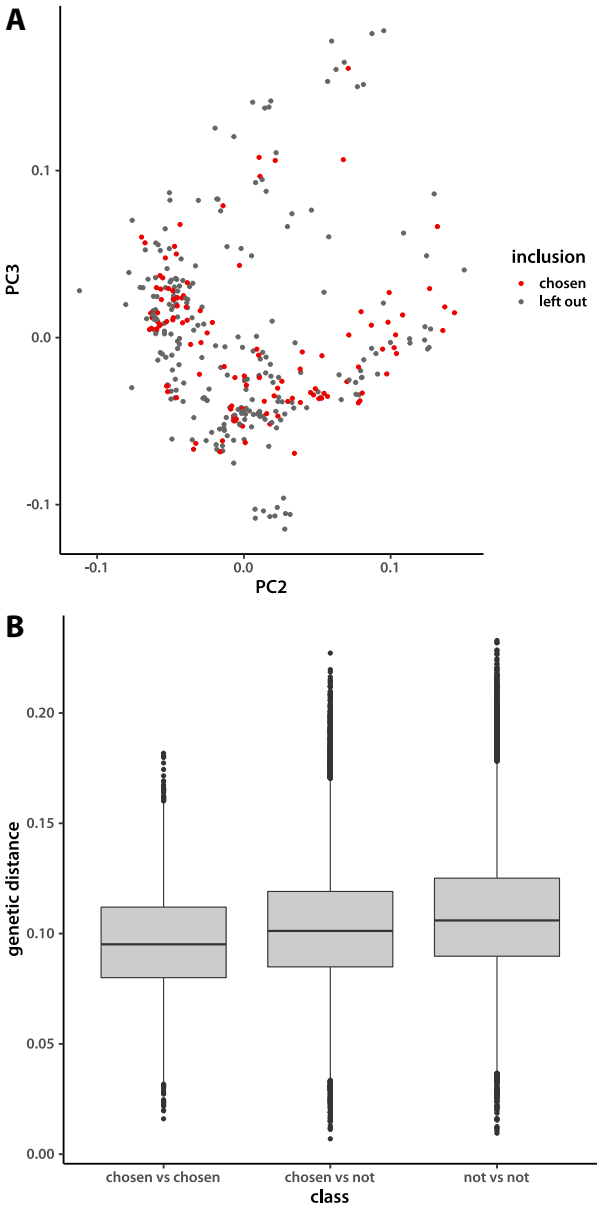


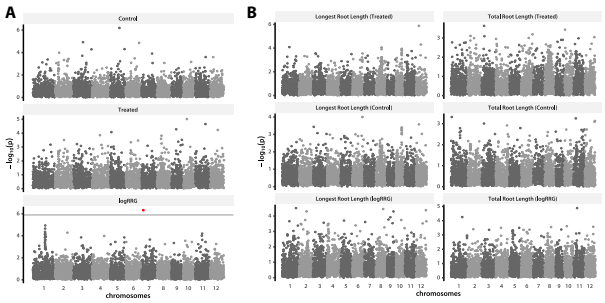
Figure 3: **Rank of accessions with aus admixture.** (A) Posterior highest-density intervals of accession means, sorted from smallest mode to largest. Tropical *japonica* accessions with *aus* admixture are marked with orange (see plot legend). (B) Posterior distributions of mean rank of admixed accessions under control and treatment conditions, with the expected value shown by the horizontal line. (C) Introgression locations in the genomes of the six admixed accessions. The six horizontal lines represent the genomes, with the accession with the shortest roots under aluminum stress on the bottom and the longest on top. Colored blocks mark the regions of introgression from *aus* ("AUS") or *indica* ("IND"), as indicated on the legend. The *tropical japonica* ("TRJ") background is in gray.



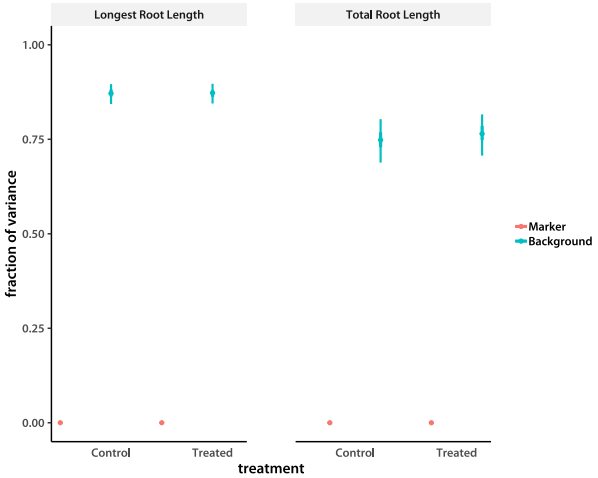
Supplemental Figure 1: **Experiment effect and accession distribution.** The box plots (A, C) show distributions of raw replicate data from each experiment in Famoso *et al.* (2011) data. The bar plots (B, D) depict the total number of replicates coming from each subpopulation within each experiment.



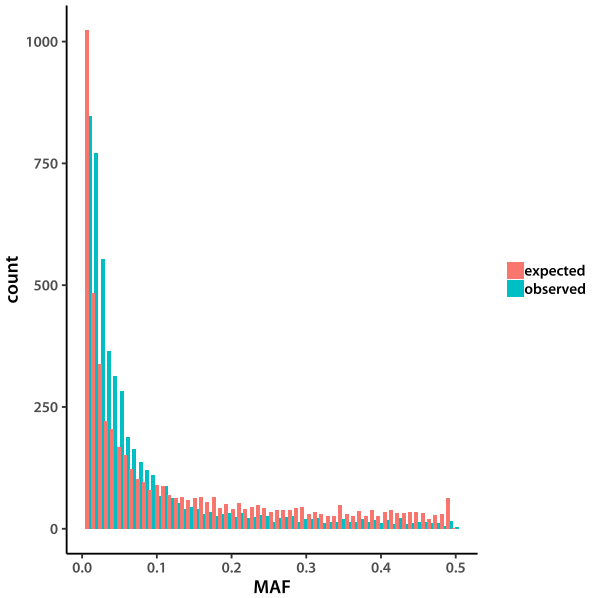
Supplemental Figure 2: **Genetic diversity of the chosen tropical japonica accessions.** (A) Principal component plots of TRJ accessions with HDRA genotypes. Components 2 and 3 were chosen because outlier genotypes make PC1 less informative. (B) Distributions of genetic distances among chosen, chosen *vs* left out, and among left-out accessions.



Supplemental Figure 3: **GWA in tropical japonica.** (A) The population examined in Famoso *et al.* (2011) with additional accessions measured under the same conditions. (B) The new set of accessions measured under harsher aluminum treatment.



Supplemental Figure 4: **Fraction of variance explained by the quadratic kernel.** Distributions of the fraction of variance explained, plotted as in Fig. 2. The markers enter the model via a quadratic kernel, so the "marker" variance reflects all pairwise epistatic interactions.



Supplemental Figure 5: **Allele frequency spectrum of tropical japonica SNPs.** Folded site frequency spectrum of the 119 accessions genotyped with HDRA, compared to the equilibrium neutral expectation. Minor allele frequency is on the x axis.