Methods

Mash

We analyzed SNP effects on 4 panicle traits in 3 gardens in mash, using effect estimates from univariate GWAS. GWAS were conducted on 381 genotypes present at all three gardens. Mash was then run in three stages following mash documentation: first, 100K SNPs unlinked at r^2 of 0.2 were used as a 'random' set to learn the correlation structure among null tests; second, 5K SNPs with large effects in a single GWAS were used to construct data-driven covariance matrices, third, the random set was used to fit the mashr model; fourth, posterior summaries using the model fit on the random set were computed on 18.7M SNPs present at a MAF of 5% at all three gardens.

Enrichment tests

To determine if SNPs with significant trait effects in mash were enriched in panicle QTL intervals, both SNP clumping by mash Bayes Factor significance level and permutation tests of QTL intervals were used.

First, all 18.7M SNPs used in mash were clumped to keep only the most significant SNP in each LD block, using a linkage threshold of $r^2 > 0.2$. Significance was assessed using the log10(Bayes Factor) computed by mash, which measures the overall significance of a SNP on the trait effects included in mash. SNP clumping resulted in ~2.7M 'relatively unlinked' regions of the genome, with 6149 SNPs significant at a log10BayesFactor of 1.3 (equivalent to a FDR-adjusted p-value of 0.05). As these SNPs represent the most significant SNP per LD block, the remaining analyses were run on this SNP subset.

Second, 1000 permutations of the QTL regions were created, which each had 18 genomic regions of the same size as the 18 QTL found for panicle traits. For each of these permuted QTL intervals, we found the number of these permuted regions that had significant enrichments of mash SNPs using hypergeometric tests.

Results

Significant overlap in genetic effects in two mapping populations

Our mash model indicated that 6149 (0.23%) of LD blocks unlinked at $r^2 = 0.2$ ('unlinked mash SNPs') had significant effects on at least one trait x garden condition. Using SNPs in the top 1% quartile of significance by -log10BayesFactor, all 18 QTL had 46 or more unlinked mash SNPs with significant effects on one or more panicle trait in mash. In addition, 10 of the 18 QTL regions had a significant enrichment of unlinked mash SNPs (p hypergeometric test < 0.05, Table S#1). 0.2% of random genomic intervals had as many or more regions enriched for unlinked mash SNPs (permutation p = 0.002), while no random genomic intervals had more regions significantly enriched for unlinked mash SNPs. Thus, even with the very different population makeup of the GWAS panel, we could confirm a higher than expected overlap between SNP effects and QTL effects on panicle traits.

All QTL regions also had significant unlinked mash SNPs within 20kb of genes with functionally validated roles in panicle, spikelet, or grain traits in rice (Table S#2). Interestingly, 76.1% of significant unlinked mash SNPs had significant effects on all nine combinations of panicle trait and garden. Most SNPs (56.2%) had high mash model weights on equal effects in all combinations of trait and garden; most remaining SNPs (43.7%) had high model weights on having no effects in any condition. Thus, the SNP data supported our QTL findings that there is little GxE for panicle traits in switchgrass.

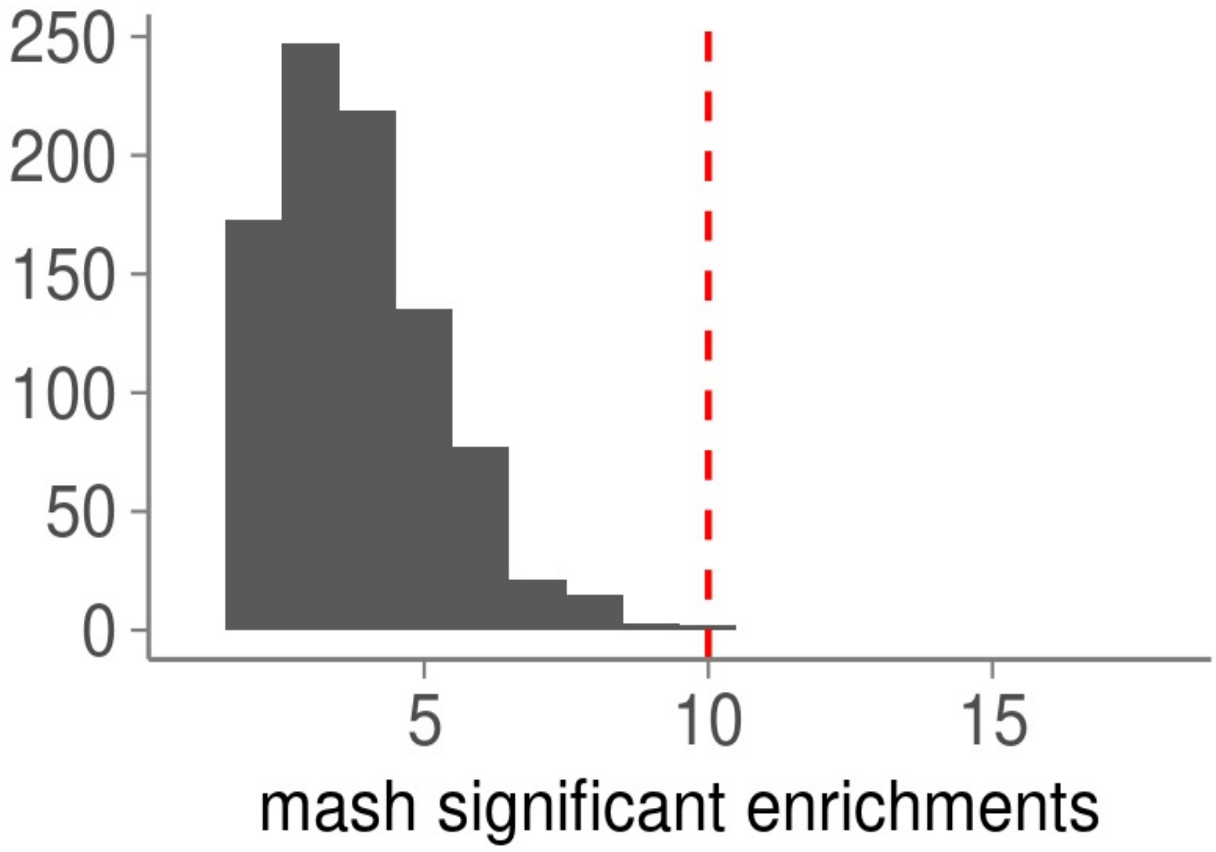

Table S#1 mash_QTL_enrichments_pvdiv_panicles_2019_BLUPs_PLANT_ID_geno_subset_All.csv

Table S#2

Mash_annotations_in_QTL_regions_with_functional_validation_pvdiv_panicles_2019_BLUPs_PLANT_ID_geno_subset_All.csv