*Genotype-by-environment effects on greenup and flowering as functions of environmental cues*

At the continental scale we found no consistent associations for greenup, due to the presence of negative phenotypic and genetic correlations in greenup date between the Texas and North gardens. In contrast, there were some recurring associations for flowering at this scale. A variance components analysis of flowering G and GxE across the eight common gardens also suggested the presence of GxE for rainfall, GDD, and photoperiod cues for flowering, for which variation was more visible outside of each subpopulations’ native range. We first­ explored GxE in greenup as a function of two environmental cues with additional analyses: 1) as a function of Julian date (‘greenup date’); and 2) as a function of cumulative GDD in the ten days before greenup (‘greenup GDD’). We then extended our exploration of GxE in flowering as a function of five environmental cues with additional analyses: 1) as a function of daylength at flowering (‘flowering daylength’); 2) as a function of daylength change from the previous day on the day of flowering (‘flowering daylength change’); 3) as a function of cumulative GDD between greenup and flowering (‘flowering GDD’); 4) as a function of rainfall on the day of flowering (‘flowering rainfall’); and 5) as a function of rainfall between greenup and flowering (‘cumulative rainfall’).

We conducted univariate GWAS at each common garden for these functions of greenup and flowering, then analyzed the allelic effects of unlinked SNPs across common garden sites for the top 5000 or 1000 SNPs, respectively, using mash. When the same SNP set is used in multiple univariate GWAS, a subsequent mash analysis shares information on patterns of effect size and direction for SNPs across these GWAS, improving the power to detect significant, shared results. Importantly, mash allows – indeed, even generates – multiple covariance matrices that particular SNP effects can load on to. Mash then shrinks effect estimates for each SNP towards one of a set of covariance matrices. Just as different genetic subpopulations can have different strengths of connection between phenotypes and genotypes (Korte and Farlow 2013), and can have different genetic covariances between these phenotypes at different gardens (Figure 1B), distinct SNPs can have different patterns of effect, or covariances, on phenotypes at different gardens. These user-specified and data-driven covariance matrices are an important advantage mash offers for studying patterns of GxE.

We first used mash to examine GxE in greenup in the Gulf subpopulation across 16 conditions, representing combinations of garden (eight) and greenup function (two). 10581 of 21697 modeled SNPs had significant effects in one or more condition. The majority of these SNPs (89%) had significant effects in eight or fewer conditions (Figure 3a). Overall, SNP effects were much more varied across gardens than across greenup functions within each garden. SNP effects for greenup exhibited multiple complex types of GxE, demonstrating why attempts to find consistent genetic effects failed for this phenotype. Indeed, only 1.5% of significant SNPs loaded onto a covariance matrix which had equal effects for each condition. Four non-zero covariance matrices loaded 5% or more of total SNPs (Figure 3b). These covariance matrices represent the most common patterns of GxE in SNPs that have significant effects on greenup, and include both data-derived (ED) and canonical matrices. The first of these, ED\_tPCA, loaded 49% of the significant SNPs. 65.3% of the variation in this matrix was captured by two major patterns (eigenvector 1 and 2 from a singular value decomposition explained 65.3% of the total variation in this covariance matrix) (Figure 3c, 3d). The first eigenvector was similar to the first eigenvector of ED\_PCA\_1. ED\_PCA\_1 loaded 21.5% of the significant SNPs, and was characterized by large magnitude (>|0.5|) effects decreasing greenup date and greenup GDD at the Texas gardens and in Oklahoma, with moderate magnitude effects (|0.2| to |0.5|) increasing greenup date and greenup GDD in MO and MI (Figure 3e). The first eigenvectors of both ED\_PCA\_1 and ED\_tPCA have antagonistic pleiotropy for greenup between all four of the southern and two of the northern common gardens. The second eigenvector of ED\_tPCA, explaining 21.7% of the variation in this matrix, showed large magnitudes of effects at the MO and MI gardens, and moderate effects in the TX1, TX3, and SD gardens (Figure 3d). It also showed antagonistic pleiotropy between the SD gardens and the Texas, MO, and MI gardens. The third covariance matrix, ED\_PCA\_3, loaded 8.8% of the significant SNPs, and was characterized by large magnitude effects on both greenup functions at the NE garden, with differentially sensitive effects at the other northern gardens and antagonistically pleiotropic effects at the Texas and OK gardens (Figure 3f). The last, ED\_PCA\_4, loaded 5.4% of the significant SNPs, and was characterized by antagonistic pleiotropy between the gardens in the center of the latitudinal gradient, of large magnitude at OK, relative to gardens at the lowest and highest latitudes, of large magnitude at TX1 and SD (Figure 3g).

We characterized patterns of differential sensitivity and antagonistic pleiotropy between all SNPs and all pairs of gardens. Between pairs of gardens, the fraction of antagonistically pleiotropic loci varied between 0 and 53%, with the most antagonistic pleiotropy between the Texas and OK gardens and the MO and MI gardens (Figure 3h). The fraction of differentially sensitive loci varied between 0 and 44%, with the most differential sensitivity occurring between the NE garden and other gardens, particularly the Texas, OK, and SD gardens (Figure 3i).

We looked for enrichments of genes that could potentially affect greenup in the 20kb genomic windows surrounding the 10581 SNPs with significant effects on greenup. *Were these effects enriched for genes in rice with presumably similar functions? What were the GO enrichments for genes from regions that loaded onto each of these covariance matrices? Distinct GO enrichments in SNP sets that load onto different covariance matrices, maybe?*

Diagram

Description automatically generated**Figure 3 (rough)**

Chart, line chart

Description automatically generated

Figure 3a and Figure 3b. A) Number of significant SNPs per condition. Conditions represent all combinations of garden (eight) and greenup functions (two). B) Loading of SNP effect patterns onto the covariance matrices included in the mash model. X-axis represents the fraction of SNPs that loaded onto each matrix. Matrices with loadings of 0.05 or greater are shown below.

Latitudinal gradient/autocorrelated spatial pattern; PLOS Biology paper lengths? Check. Talk about types of covariance matrices you explored, without necessarily including them.

Make an effort to explain what the covariance matrices are – and they’re hypotheses – and name them more clearly

Think about Q’s for Stephens

Chart, waterfall chart

Description automatically generated

Chart, waterfall chart

Description automatically generated

Figure 3C, 3D. First and second eigenvectors from a SVD of ED\_tPCA, which loaded 49% of the significant SNPs for greenup.

Chart, waterfall chart

Description automatically generatedFig 3E. First eigenvector from a SVD of ED\_PCA\_1, which loaded 21.5% of the significant SNPs for greenup.

Chart, bar chart

Description automatically generated

Fig 3F. First eigenvector from a SVD of ED\_PCA\_3, which loaded 8.8% of the significant SNPs for greenup.

Chart, waterfall chart

Description automatically generatedFig 3G. First eigenvector from a SVD of ED\_PCA\_4, which loaded 5.4% of the significant SNPs for greenup.

A picture containing scatter chart

Description automatically generated

Figure 3H Fraction of SNPs significant in both conditions that exhibit antagonistic pleiotropy between these conditions. Pairwise fractions for each pair of conditions are shown.

A picture containing text

Description automatically generated

Figure 3I Fraction of SNPs significant in both conditions that exhibit differential sensitivity between these conditions. Pairwise fractions for each pair of conditions are shown.

A picture containing box and whisker chart

Description automatically generated

Figure 3J. Example effect plot for a single SNP – the most significant SNP for greenup effects in the Gulf subpopulation. This SNP has 99.9% probability of loading on the ED\_tPCA covariance matrix. It shows antagonistic pleiotropy between SD and the other gardens, and larger effect sizes for the Texas and OK gardens. (More plots like this can be generated on R-shinyapp).