Across all eight common gardens, the combined estimates for G and GxE for greenup as a function of Julian date were less than 7% for each set of subpopulations (Supplementary Figure/Table). G and GxE estimates for greenup date were much higher when the sites were restricted to either Texas or the Northern four sites (Supplemental Figure). G and GxE estimates for greenup were higher for the Gulf subpopulation than for the Midwest, and higher outside of each subpopulations native range than within its native range (Supplementary Figure). Across all sites for the Gulf subpopulation, GDD 10 days before greenup had the most variance partitioned to G and GxE; for the Midwest subpopulation, GDD from January to greenup had the most variance partitioned to G and GxE. At subsets of sites, only GDD 5 days before greenup explained more variance for the Midwest subpopulation at the Northern five sites. For other combinations of subpopulations and site subsets, greenup date outperformed greenup as functions of weather. This result likely indicates that the weather functions we chose are not cuing greenup. Additional signals such as soil temperatures or chilling days may influence greenup for each subpopulation within its native range; however, we did not have good proxies for these values for this experiment.

G and GxE estimates for flowering date for the Gulf subpopulation were 24% and 30% (Figure 1D).

In contrast, flowering as a Julian date had moderate heritability within subpopulations (h2 = 0.19 +/- 0.061 Gulf; 0.097 +/- 0.045 Midwest) and moderate heritability across both subpopulations (h2 = 0.18 +/- 0.061).For the Gulf subpopulation, daylength (h2 = 0.336 +/- 0.073) was the only cue that had higher heritability than Julian date. Two environmental functions had higher heritabilities than Julian date for the Midwestern subpopulation: cumulative GDD (h2 = 0.368 +/- 0.052) and daylength (h2 = 0.144 +/- 0.052). Across both subpopulations, both daylength and cumulative GDD had higher heritabilities than Julian date (0.287 +/- 0.032; 0.201 +/- 0.041).

To explore this further, we tested whether subpopulation flowering as a function of GDD varied significantly by latitude of origin. The Midwest subpopulation response to GDD varied significantly by latitude of origin, with plants from the northernmost 20% of the range flowering at an average GDD of 568, and plants from the southernmost 20% of the range flowering at an average GDD of 779. In contrast, the southernmost and northernmost 20% of the Gulf plants differed only slightly in GDD (1008 vs 1090, \*statistically significant).

Because no greenup cue as a function of weather had h2 > 10% across all sites, we did not evaluate genetic effects of greenup across all sites. Rather, we focused on genetic effects at the Texas sites and at the Northern four sites, and on genotype by environment effects of SNPs on greenup.