



**Figure S9** Genome-wide association study (GWAS) for total  $\beta$ -xanthophyll content in maize grain. Scatter plot of association results from a unified mixed model analysis of total  $\beta$ -xanthophyll and linkage disequilibrium (LD) estimates ( $r^2$ ) across the *lcyE* chromosome region. Negative  $\log_{10}$ -transformed  $P$ -values (left y-axis) from a GWAS for total  $\beta$ -xanthophyll and  $r^2$  values (right y-axis) are plotted against physical position (B73 RefGen\_v2) for a 12 Mb region on chromosome 8 that encompasses *lcyE*. The blue vertical lines are  $-\log_{10} P$ -values for SNPs that are statistically significant for total  $\beta$ -xanthophyll at 5% false discovery rate (FDR), while the gray vertical lines are  $-\log_{10} P$ -values for SNPs that are non-significant at 5% FDR. Triangles are the  $r^2$  values of each SNP relative to the peak SNP (indicated in red) at 138,883,206 bp. The black horizontal dashed line indicates the  $-\log_{10} P$ -value of the least statistically significant SNP at 5% FDR. The black vertical dashed lines indicate the start and stop positions of *lcyE* (GRMZM2G12966). (B) Scatter plot of association results from a conditional unified mixed model analysis of total  $\beta$ -xanthophyll and LD estimates ( $r^2$ ) across the *lcyE* chromosome region, as in (A). The two SNPs (*lcyE* SNP216 and S8\_138882897) from the optimal multi-locus mixed model (MLMM) model were included as covariates in the unified mixed model to control for the *lcyE* effect.