

Regional association scanning

Jinliang Yang

May 15, 2015

Results

With the control of the population structure and polygenic effects of background QTLs, we conducted the regional association scanning around the ZmSWEET4c gene. As show in Figure 1, two SNPs near the ZmSWEET4c gene were significantly ($FDR < 0.01$) associated with the traits of ten kernel weight and total kernel weight, which are two important yield indexes. The most significant SNPs could explained 4.3% and 4.3% of the total phenotypic variations for ten kernel weight and total kernel weight, respectively. And about 7.8% and 10.8% of the heritability of the traits could be accounted by the most significant markers for ten kernel weight and total kernel weight. Importantly, non-B73 like genotype has the magtitude of effects of 0.2g (9.1%) and 4.4g (12%) for ten kernel weight and total kernel weight, respectively.

Materials and Methods

Regional association study

A maize diversity panel composed of 282 inbred lines was employed for the regional association study. To conduct the analysis, we obtained Genotype-By-Sequencing (GBS) data from panzea (www.panzea.org) and obtained phenotypic data from Flint-Garcia et al., 2009 ¹. The SNPs data were filtered with minor allele frequency (MAF) > 0.05 and allele missing rate $< 50\%$. After filtering, a total of 306,190 SNPs were remaining, including 79 SNPs in a 10-Mb region surrounding the ZmSWEET4c gene.

Association study with the mixed-model method was conducted using an R ² add-on package “GenABEL” ³. First of all, a kinship matrix was estimated from the genomic data to control population structure. Secondly, genome-wide polygenic effects were computed with the function “polygenic” for controlling background QTLs. Finally, the 79 SNPs near zmSWEET4c gene were tested one-by-one as the fixed effect and polygenic QTL effects derived from previous step were fitted as random effects using the function of “mmscore”.

Figure legend

Figure: Regional association scanning results. Vertical dashed blue lines indicate the center of the ZmSWEET4c gene and horizontal red lines indicate the threshold of $FDR=0.01$. Q-value represents the corrected p-values with FDR control.

¹Flint-Garcia, Sherry A., et al. “Heterosis is prevalent for multiple traits in diverse maize germplasm.” PloS one 4.10 (2009): e7433.

²Team, R. Core. “R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012.” (2012).

³Aulchenko, Yuri, et al. “GenABEL: genome-wide SNP association analysis.” R package version (2010): 1-6.