

M-170

A versatile approach to define gene-centric haplotypes using dense SNP data  
*Aurélie Tardivel*, University of Laval, Quebec, Canada

Genes controlling economic traits in soybean have been abundantly reported and natural variation at these genes contributes to variation in phenotype. For such known genes, it is useful to assess the allelic diversity captured in germplasm collections and to identify individuals carrying favorable alleles at these genes. High-throughput genotyping (e.g. Genotyping By Sequencing (GBS) or SNP arrays) has made it possible to genotype large numbers of individuals at thousands of SNPs at low cost, thus providing extensive SNP coverage and making it possible to define major haplotypes. In many cases, these haplotypes correspond to alleles of a given gene. In this study, we have developed a versatile and systematic approach to facilitate the process of identifying informative SNPs that help to define haplotypes for a given gene from a set of dense SNP markers. In contrast to other haplotype block definition methods, our approach is centered on a gene of interest and aims to select only markers in the vicinity of a gene that are assumed to be in LD with allelic variation at this gene. The objectives were to determine the potential and limits of this tool by testing its performance on four different maturity (*E*) genes located in various regions of the soybean genome in a collection of 67 inbred Canadian lines using different genotyping datasets. Our approach proved equally successful for three of the four genes studied using both SNP datasets (GBS and array). Nonetheless, while we established the feasibility of this approach, we also illustrate its limits. Indeed, the success of this strategy is largely dependent on the ability to genotype markers in high LD with each allele.