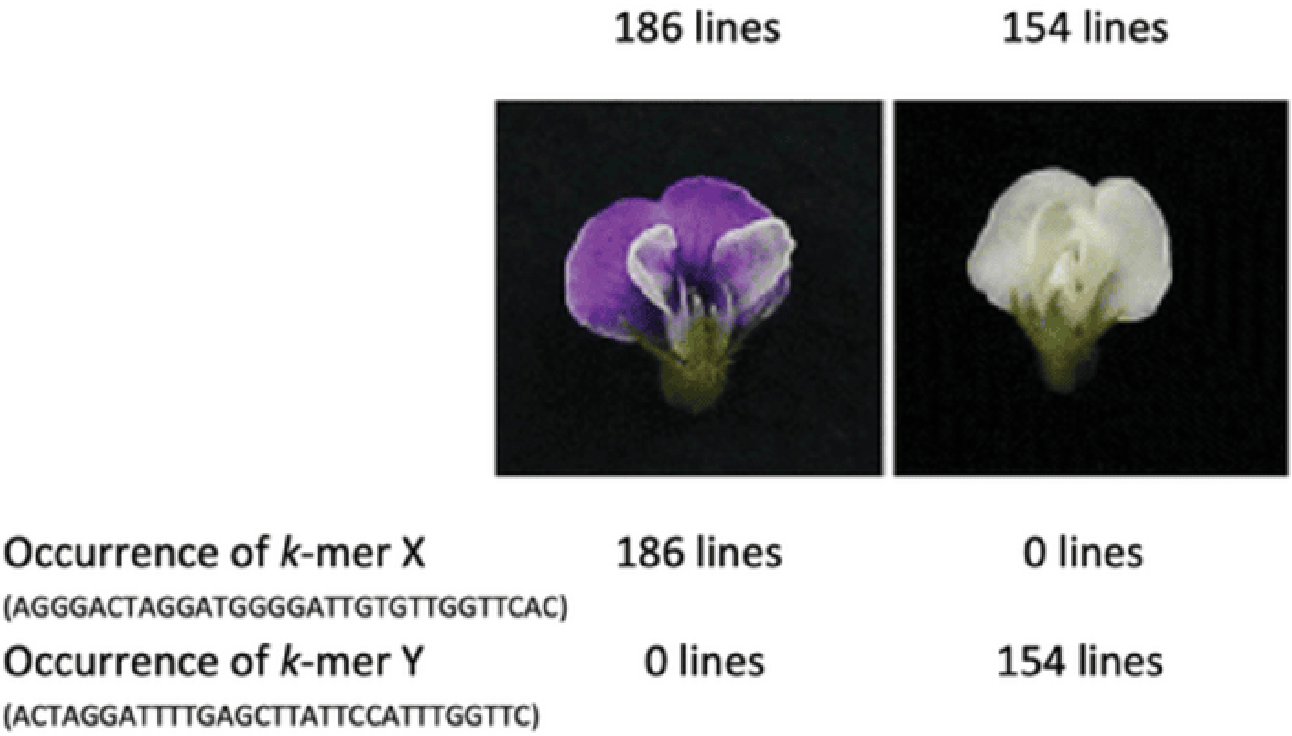
Science
Societies

k-mer-Based GWAS Unlock Causal Variant Discovery in Soybean

December 15, 2023



The occurrence of short sequence motifs, pinpointed from the re- sults of a GWAS, distinguished between white- and purple-flowered soybean accessions. Figure reproduced with permission from the following article by Sundaramoorthy et al. published in the Journal of the Korean Society for Applied Biological Chemistry: https://doi. org/10.1007/s13765-015-0077-z.

Genome-wide association studies (GWAS) allow the discovery of loci associated with traits of interest in crops by identifying statistical associations between molecular markers and traits. Many studies use single-nucleotide polymorphisms (SNPs) as molecular markers, but observed variation in traits can be due to other types of genetic changes such as structural variants (SVs). GWAS frequently fail to identify them.

Researchers tested if short sequence motifs, known as *k*-mers, could identify causal variants regardless of if they are SNPs or SVs. Using a population of 363 cultivated soybean lines, they applied SNP-, SV-, and *k*-mer-based GWAS to 13 traits. The researchers found that *k*-mers could pinpoint known causal variants at four loci, while identifying promising causal genes for several other traits. These analyses can assess all types of variants in a single analytic framework.

Such analyses hold promise for speeding up the application of genomics to plant breeding. The framework and computational tools developed by the authors for downstream analysis of *k*-mer-based GWAS will facilitate the adoption of this method.

Adapted from Lemay, M.-A., de Ronne, M., Bélanger, R., & Belzile, F. (2023). k-mer-based GWAS enhances the discovery of causal variants and candidate genes in soybean. The Plant Genome, e20374. https://doi.org/10.1002/tpg2.20374