

# The Alfalfa Breeder's Toolbox

Deborah Samac, USDA-ARS

Modern-day travelers rely on maps to navigate to places they need to go. Likewise, modern-day plant breeders rely on DNA maps in their crops to breed improved cultivars. This technology has been slow to develop in alfalfa, mostly due to the tetraploid genetics (four sets of chromosomes) and high level of genetic diversity. However, needed tools are rapidly being constructed and tested to assemble an Alfalfa Breeder's Toolbox that will be available to public and private alfalfa breeders for crop improvement.

The basis of mapping in plants is a marker or unique signpost in the map. The earliest maps were created using crop traits such as seed color, flower color, and leaf shape. A giant leap forward was made when DNA sequencing technologies were developed in the 1980s enabling identification of unique DNA sequences in chromosomes. Repeated sequences, small insertions and deletions, and single nucleotide differences within a specific DNA sequence could be identified. Mapping these markers relied on identifying the different markers in two parent plants, then crossing them to generate populations that were segregating for the maternal and paternal markers. Computer programs, mostly based on diploid plants, were developed that could place the markers on linkage groups roughly corresponding to chromosomes. Plant breeders now had "mile markers" along the maps. When combined with phenotypic traits of the progeny plants, markers could be associated with those traits. Thus, in many crops, markers associated with resistance to specific diseases and other important traits were identified and could be used to track the trait in a process called marker-assisted selection. In alfalfa, traits were mapped for aluminum tolerance in acid soil, resistance to several diseases, water-use efficiency, and biomass yield. However, in most cases, the markers were only useful in the population in which they were developed.

With the development of next-generation DNA sequencing, it became possible to rapidly sequence whole genomes and assemble the sequences into long scaffolds that covered most of each chromosome. Additionally, the protein coding sequences could be identified and annotated to identify the putative function of the genes along the chromosomes. Now breeders had the physical maps of the markers and genes, analogous to populating a highway map with places of interest like gas stations, restaurants, and other businesses. Sequencing whole genomes also identified numerous new DNA markers based on single nucleotide differences or polymorphisms, usually called SNPs or "snips," that could be associated with genes or traits of interest. Many different platforms were developed to determine the specific SNPs in different plants and to use the SNP markers to identify plants with the desired traits for breeding improved crops. Although genome sequencing and SNP marker identification were done in the annual relative of alfalfa, *Medicago truncatula*, most of the new mapping technology was not applied to alfalfa because it was too expensive to identify enough markers linked to traits of interest and because of the lack of a genome sequence for alfalfa.

One newer technology used to map numerous traits in alfalfa is Genotyping-by-Sequencing (GBS). This method randomly sequences portions of genomes of segregating populations to generate massive numbers of genome-wide SNPs. Methods for analyzing the data in a tetraploid were developed and first used the *M. truncatula* genome sequence as a reference with more recent use of alfalfa genome sequences to place the markers onto DNA maps. The drawbacks of GBS are the relatively high cost for analysis of each plant and the random nature of the sequencing. GBS cannot target the same set of SNPs in individuals or populations and the SNPs identified in one mapping population may not transfer to another population. Breeders need stable targeted markers to follow the transmission of traits through different breeding cycles.

In 2018, USDA-ARS initiated the Breeding Insight program at Cornell University to make genomics-enabled breeding a reality for crops such as alfalfa. The goal is to develop universal SNP markers that can be used for genotyping any alfalfa population at low cost with a rapid turn-around time. Additionally, the aim is to reduce the burden on breeders to learn new methods for analysis of genome data by providing genotypic data associated with phenotypic data from the breeding program for making plant selections and to accelerate breeding efforts. To develop the SNP markers, a panel of 40 diverse alfalfa plants was assembled including elite parents from public and private breeding

programs. The genome of each plant was sequenced and mapped to a reference alfalfa genome sequence. An initial set of 3,000 markers was identified that was evenly distributed across the chromosomes, and an additional set of 3,000 markers targeting regions with low marker coverage from the initial set is currently under development. The platform for detecting the SNPs is the DArTag technology in which the regions of DNA containing the SNPs are amplified by PCR and then sequenced to identify the polymorphisms present.

The markers were tested with a very broad array of alfalfa plants from breeding programs, perennial *Medicago* species, and *M. sativa* subspecies. The markers successfully identified different alfalfa populations selected for salt tolerance and drought tolerance from non-selected plants (Figure 1). The markers were also used to map resistance to *Aphanomyces* root rot, and the results agreed with those obtained using GBS, narrowing the chromosomal region containing the genes involved in disease resistance.

To support use of the SNP markers, Breeding Insight is developing analysis tools to generate the genotypic calls and estimate marker and trait associations based on phenotyping and genotyping results. The full Alfalfa Breeder's Toolbox will encompass alfalfa genome sequence data, SNP marker information, gene annotation and conditions under which those genes are expressed, and identification of chromosomal regions associated with specific traits. These tools will allow a breeder to select plants with the most favorable set of genes while avoiding unfavorable genes having negative impacts on desired traits such as forage yield and quality. The ability to use genomic data will shorten the response time for developing new cultivars with new traits or when new market needs emerge.

**Figure 1.** SNP markers distinguish plants selected for salt tolerance and drought tolerance from non-selected plants.

