

Development of an Alfalfa Breeder's Toolbox

Maria J. Monteros¹, Patrick X. Zhao¹, Shreyartha Mukherjee¹, Nick Krom¹, E. Charles Brummer²

¹The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

²University of California, Davis, CA 95616 USA

Advances in DNA sequencing technologies and current efforts in alfalfa transcriptome sequencing, SNP discovery and germplasm evaluation generate useful datasets that can be used to inform breeding strategies. The challenge is to organize, store, integrate and analyze these complex datasets and represent them in a meaningful way. We aim to develop an Alfalfa Breeder's Toolbox as a comprehensive resource of breeder-friendly tools that integrates basic and applied knowledge. The goal is to integrate all relevant information from existing and emerging databases into a user-friendly resource that will enable breeders to make meaningful inferences regarding genes, sequence variation (molecular markers), gene expression patterns and multi-location phenotypic data that would be useful for improving target traits. Consistency in phenotypic data can be maintained through the use of Trait Ontologies, a controlled vocabulary to describe each trait as a distinguishable feature of a developing or mature plant. Ontologies for alfalfa growth, development and phenotypic traits can be integrated with the Plant Ontology Consortium, as well as with links to genes and relevant genomic regions thus enabling searches by trait, gene or other desirable query fields. The considerable genetic and genomic resources in other legumes, particularly *M. truncatula*, could add further value, and hence we propose to integrate these resources with the Alfalfa Breeder's Toolbox. Relevant databases include those developed for the model legumes *M. truncatula*, *Lotus japonicus*, *Glycine max* and other crop legumes via comparative mapping approaches to integrate alfalfa's growing genomic and transcriptomic resources. The Alfalfa Breeder's Toolbox can enhance efficiencies, facilitate identification of meaningful information from a range of datasets and result in shorter timelines to develop new cultivars that address current and future needs.