

Whole-genome-based breeding in legumes

Gentzbittel L.

Project Center for Agro Technologies – Skolkovo Institute of Science and Technology, Moscow, Russia
l.gentzbittel@skoltech.ru

Key words: GWAS, Genomic Prediction, quantitative genetics, Machine Learning, disease resistance, yield

Legumes (*Fabaceae*) are a nutritious staple of diets around the world. Grain legumes (e.g. soybean, chickpea, bean, pea, etc...) are mostly used as source of edible proteins and oil. Forage legumes (alfalfa, *Vicia* spp.) are among the most effective sources of protein for animal husbandry.

The *Fabaceae* comprise annual and perennial species, also including some tree species (e.g. *Acacia*, *Prosopis*). The different species also evolved a large diversity of reproductive strategies, from self-pollination (e.g. soybean, pea, peanut) to full outcrossing (e.g. alfalfa) and some intermediate cases (e.g. faba bean). The complexities of the genomes are vast, from small sized diploid genomes (e.g. *Medicago truncatula*, $2n = 2x = 16$, genome size of 500Mb) to diploidized allopolyploids (eg. Soybean, $2n = 2x = 40$, genome size = 1.15Gb), to allotetraploids (cultivated peanut, $2n = 4x = 40$, genome size = 2.7Gb) and to very large genomes (faba bean, $2n = 2x = 12$, genome size = approx. 13Gb).

The reproductive strategies of the different legume crops imposed the creation and the use of very diverse breeding schemes to improve quantity, quality and resilience to biotic and abiotic stresses. The complexities of the genomes also imposed various molecular and bioinformatics strategies to be developed. We will present how Genome Wide Association Studies (GWAS), Marker Assisted selection (MAS) and Genomic Selection (GS) are implemented in some legume crops, reviewing strategies for molecular characterization, high-throughput, high quality phenotyping and computational methods.