

# Genomic Prediction for Phenology, Yield & Forage Quality Traits on a Large Diversity of Alfalfa

Marie Pégard, INRAE  
Julien Leuenberger, INRAE  
Bernadette Julier, INRAE  
Philippe Barre, INRAE

China and Europe dependency on imported proteins is a problem that could be solved by growing more legumes. This requires productive and good quality cultivars. The current breeding programmes could be improved by using a large genetic diversity combined with a genomic evaluation such as in many crop or animal species, including forage species. Some studies have already been published on genomic evaluation of alfalfa, they have used between 75 and 274 accessions or populations, with a number of SNP from 8 K to 44 K. They have reported variable predictive accuracies (0-0.65) for yield and quality traits.

We investigated the ability of genomic prediction to predict various genetic material on a large number of traits. We used a set of 395 alfalfa accessions mostly coming from Europe, North and South America and China with a fall dormancy between 3 to 7. This material was genotyped by genotyping-by-sequencing (GBS) method at the accession level. After filtering, 228 568 markers (SNP) covering the genome were obtained. These genomic data were used to identify genetic clusters of accessions with the Discriminant Analysis of Principal Components (DAPC) method. From 2018 to 2020, the accessions were phenotyped in four locations (Lusignan in France, Novi Sad in Serbia, Aberystwyth in England, Store Heddinge in Denmark) for three categories of traits: Forage yield, Forage quality, and, Phenology (winter dormancy and flowering date), ending to 187 traits representing the combination of a trait in a location and a year or integrated traits (combined between year and/or locations). The phenotypes were adjusted for the micro-environmental variation inside each trial, the year and the location effects. The QTL were detected by Multi-Locus Mixed Model (mlmm) and retained if they pass the Bonferonni threshold. Then, we assessed the quality of the genomic prediction for each trait. A cross-validation was used to assess the prediction quality with a test of several models (GBLUP, Ridge-regression (RR), Bayesian lasso). We also tested the integration of QTL in the model as fixed effects (QGBLUP) compared to the classical GBLUP model.

A genetic structure with seven clusters, mostly related to the geographical origin of the accessions, was found. Two groups were clearly separated from the other accessions: one with only Chinese accessions and the other with some subsp. *falcata* accessions. The other five groups, respectively composed by the accessions of (1) France and Northern Europe, (2) Southern and Eastern Europe, (3) a mix from Europe, North and South America, (4) USA and 1 of China, and (5) North-American only, formed a continuum. Many QTLs were found: 438 for forage yield, 79 for phenology, and 564 for quality. Each QTL explained on average 0.17, 0.09, 0.13 and 0.07 of the phenotype for forage yield, phenology, quality and seed yield, respectively. The evaluation of the genomic prediction reveal that, we reached high predicting abilities. The infinitesimals methods showed higher quality of prediction (GBLUP :  $0.42 \pm 0.16$ , and RR :  $0.44 \pm 0.16$ ) than the bayesian lasso (lasso :  $0.38 \pm 0.16$ ) or the QGBLUP (QGBLUP :  $0.27 \pm 0.15$ ). Although less efficient, the QGBLUP showed that with a few QTLs, it is possible to predict between 40% and 80% of the prediction quality obtained GBLUP. The traits linked to phenology tended to be better-predicted ( $0.62 \pm 0.17$ ) than forage yield ( $0.45 \pm 0.19$ ) or quality ( $0.43 \pm 0.14$ ). Our results showed that it was possible to achieve high genomic prediction qualities in a large number of traits; this is encouraging for alfalfa breeding.