

## Master Thesis Topics on FHB Genetics and Genomics

### Introduction:

Diseases and rapid climate change cause challenges in present day agriculture. Here, in our case, Fusarium Head Blight (FHB) is one of the most threatening wheat diseases in Europe and North America. FHB epidemics have increased dramatically due to climate changes, reduced tillage, lack of adequate crop rotation and cultivation of susceptible cultivars. FHB epidemics regularly cause both yield and quality losses in major wheat growing regions and is of serious concern for human and animal nutrition due to the production of mycotoxins. Climate change and increasing food demand give rise to challenges in wheat breeding – new cultivars have to be developed quickly to accommodate new conditions – both ecological and economical. Traditional breeding processes are time and resource-consuming – typically it takes roughly 10 to 12 years to develop a new cultivar. Moreover, it is also labour intensive and costly. The important fusarium species causing FHB is the DON-producing *F. graminearum* for which a broad genetic variation in resistance can be found both in elite and exotic wheat germplasm. FHB resistance is of quantitative nature with a complex genetic architecture. Our group has been studying the genetics of FHB resistance over many years. Genome-wide association study (GWAS) is a widely used methodology to explore the genetic variation associated with traits of interest. We also use GWAS to identify associated markers and candidate genes that will be used to make Genomic Predictions.



Genomic Selection (GS) is a modern genomic tool in plant breeding to make the selection of new cultivars faster and more accurate. Genomic prediction aims to utilise genetic resources (such as SNP markers) to enable *in silico* prediction of the breeding values of new breeding lines based on their genetic makeup. Coupled with state-of-the-art data analysis, including machine learning and multi-trait prediction models, GS can prove to be a potent tool for breeding, decreasing the necessity of manual evaluation at great costs. Using GS, estimated prediction accuracies are very low for complex diseases such as FHB. For this reason, in order to improve the prediction accuracies of disease resistance, we use the huge established knowledgebase of biologically relevant data, quantitative trait loci (QTL) and marker-trait relationships. This prior knowledge will be included into the GS models to improve prediction accuracy, which in combination with speed breeding will bring out a faster and less costly production of improved resistant cultivars.

### Topic 1: To build multi-trait genomic prediction models for FHB resistance incorporating known markers from GWAS as fixed effects

FHB resistance is of quantitative nature with a complex genetic architecture and has broad genetic variation. The most consisted loci that contribute the largest effects can be identified using GWAS. Here the idea is to use these known loci as fixed effects in the genomic prediction models to improve their prediction accuracy. The models will be trained on data from the MASBASIS spring wheat panel (consisting of 300 cultivars and breeding lines tested over many years) then validated on new breeding lines from Graminor (in field evaluation from 2020).

#### Tasks:

- Identify the markers to be included as fixed effects from the GWAS results and other relevant literature
- Compare single-trait models for FHB and DON with multi-trait models incorporating, anther extrusion (AE), days to heading (DH) and plant height (PH) as co-variables
- Compare the above models with and without including known markers as fixed effects

#### Learning outcomes:

- Data handling and management in R and TASSEL
- Statistical analysis and visualisation in R
- Experience in scoring of FHB-related traits in field trials
- Scientific writing

### **Topic 2: Evaluating the potential of machine learning models over the statistical methods in achieving promising prediction abilities (accuracies) for FHB disease resistance in wheat**

This topic will assess the potential of machine learning methodology in achieving good prediction abilities (accuracies) by making use of the genomic-wide SNP markers and environmental covariates (GxE interactions) in genomic prediction modeling, and compare if they can replace the widely used statistical models for genomic prediction like GBLUP and Bayesian models. The models will be trained on the MASBASIS data (this consists of 300 elite and exotic lines the group is working on so far) and then validated on new breeding lines from Graminor (in field evaluation from 2020).

#### Tasks:

- Involvement in field trials and evaluation
- Comparing efficiency of classical genomic prediction models and machine learning models in terms of capturing GxE interactions, better estimation of SNP effects, in order to evaluate if the machine learning methods improve prediction abilities (provided information of environmental covariates and genome-wide SNP markers)

#### Learning outcomes:

- Data handling and management in R/python
- Statistical analysis and visualisation in R/python
- Genomic Prediction Modeling and Machine learning
- Experience in scoring of FHB-related traits in field trials
- Scientific writing

#### **For further information please contact:**



Supervisor – Prof. Dr. Morten Lillemo, [morten.lillemo@nmbu.no](mailto:morten.lillemo@nmbu.no), +4767232775

Co-supervisor – Vinay Kumar Reddy Nannuru, [vinay.nannuru@nmbu.no](mailto:vinay.nannuru@nmbu.no), +47409691483

Faculty of Biosciences (Department of Plant Sciences - Genetics and plant breeding)

Norwegian University of Life Sciences (NMBU)