Biometry and Population Genetics

Thesis topics
16. Oktober 2024

Note that all thesis topics are described in English but you can of course write your thesis in German as well if nothing else is stated in the description. All our theses require that you have basic knowledge in R programming from our core modules. Please contact the potential supervisor for more information on the topics. You can start anytime if not stated otherwise.

Master thesis: Genome-wide association study of stem length in carnation

Stem length is a determinant factor in cut flower quality control. Important traits in cut flower production, like harvesting time and vase life are directly influenced by stem length (Varu *et al.*, 2010). Therefore, longer, stronger and faster growing stems are part of the breeding goals for cut flower breeders.

Carnation is part of the top economically important ornamental crops in the cut flower trade. Despite this, and similar to other important cut flower cultures like Rose and Chrysanthemum, research based on genomic resources is still underutilized in ornamental plant breeding (Schultz *et al.*, 2016). However, this is changing with the increasing availability of genetic marker and genomic data.

A well know method that exploits genetic marker data is genome-wide association studies (GWAS). Here, the idea is to pair phenotypic data of a specific trait (e.g. stem length) with genetic marker data from a given set of individuals. Differences in the phenotype and genotype between individuals are used to detect which, if any, of the evaluated markers shows a strong relation (association) to the trait under observation. This approach is useful for understanding which and how many regions of the subject species' genome can be influencing a given trait.

In your thesis, you will conduct a genome-wide association study with a cut carnation data set to investigate underlying genetic factors of stem length as a trait.

Target group: Master students of Nutzpflanzenwissenschaften or Agrobiotechnology.
Supervisor: Hugo Tavera (`hugo.tavera@agrar.uni-giessen.de`).
Requirements: MK-002 or MK-002-EN and either MK-119-EN or MP-236-EN completed successfully.
Literature to start with: For GWAS in ornamentals, Schultz et al., 2016. Frontiers in Plant Science 7:1798. For the trait "stem length", Varu, D. K., & Barad, A. V., 2010. Journal of horticultural sciences, 5(1), 42-47.

Bachelor thesis: Grain protein deviation in wheat

The price that a farmer gets for one dt of wheat strongly depends not only on the grain moisture but also on the protein content. Unfortunately, grain yield and grain protein concentration are negatively correlated: When one goes up, the other on goes down. "However, some wheat genotypes consistently deviate from this relationship, a phenomenon known as Grain Protein Deviation (GPD) (Monaghan et al., 2001). GPD is defined as the deviation from the negative regression between grain protein and yield as described by Monaghan et al. (2001). GPD is negative in wheats developed to have high contents of starch and low contents of protein for distilling or livestock feed, and positive in some high protein wheat genotypes bred for breadmaking" (Mosleth et. al. 2020). This means that a high PPD is desirable in many elite wheat lines.

In your thesis, you will characterize the GPD of two wheat populations from ongoing breeding programs and examine its relationship to other yield parameters.

Target group: Bachelor students of Agrarwissenschaften
Supervisor: Dr. Carola Zenke-Philippi (`carola.zenke@agrar.uni-giessen.de`)
Requirements: BK-005 completed successfully, BP-041 is recommended but not mandatory.
Literature to start with: Monaghan et al. 2001 Euphytica 122:309–317, Mosleth et al. 2020 Field Crops Research 255:107896

Master thesis: Different measures for heritability in wheat and/or oat

There are different definitions for the heritability. They all aim to somehow quantify how much of the observed phenotypic variability in a field trial can be attributed to genetics. However, this can be difficult to measure, particularly if the field trials under consideration are unbalanced. Different measures for the heritability have been proposed as a solution to this problem.

In your thesis, you will investigate different measures for the heritability, calculate them for data from ongoing breeding programs and wheat and/or oat and compare their properties.

Target group: Master students of Nutzpflanzenwissenschaften or Agrobiotechnology
Supervisor: Dr. Carola Zenke-Philippi (`carola.zenke@agrar.uni-giessen.de`)
Requirements: MK-119-EN and either MK-002 or MK-002-EN completed successfully.
Literature to start with: Visscher et al. 2008 Nature Reviews Genetics 9: 255–266, Piepho and Möhring 2007 Genetics 177: 1881–1888

Master thesis: Genomic prediction of the segregation variance - approaches, improvements, applications and perspectives

In recent years, several analytical derivations of the segregation variance σ_g of a cross based on genotypic data have been published. The concept has since then been employed for selection of crossing partners and planning crosses with high performance and diversity. However, critical studies have highlighted current problems with prediction accuracy and suggested improvements for existing approaches.

In your thesis, you will write a comprehensive literature review on the topic of genomic prediction of the segregation variance and its applications in breeding programs. You should also discuss suggested

modifications and future perspectives of the concept.

Target group: master students of Crop Sciences (Nutzpflanzenwissenschaften) or Agrobiotechnology
Supervisor: Dr. Eva Herzog (`eva.herzog@agrar.uni-giessen.de`)
Requirements: MK-002 or MK-002-EN and either MK-119-EN or MP-236-EN completed successfully, basics of plant breeding, good skills in maths
Literature to start with: Osthushenrich et al. 2017 PLoS One 12.12: e0188839; Lehermeier et al. 2017 Genetics 207.4 : 1651-1661.

Master or bachelor thesis: Development of models for conversion of physical maps to linkage maps in wheat

For many applications in population genetics such as simulation studies and cross planning, recombination frequencies between genotypic loci are required. Information on the recombination frequency is usually presented in the form of a linkage map in centiMorgans (cM). However, while physical marker positions are often available for many DNA marker assays, linkage maps are rare.

In your thesis, you will collect different publically available consensus maps of the wheat genome, prepare the data and fit different prediction models (KNN, sliding windows, splines etc.) that take physical positions in bP as inputs and return a linkage map suitable for population genetical simulation.

Target group: Bachelor students of Agricultural Science, and master students of Crop Sciences (Nutzpflanzenwissenschaften) or Agrobiotechnology
Supervisor: Dr. Eva Herzog (`eva.herzog@agrar.uni-giessen.de`)
Requirements: Bachelor level: BP-041 Biostatistics, master level: MK-002 or MK-002-EN and either MK-119-EN or MP-236-EN completed successfully, all: R programming skills, willingness to learn use of new R packages.
Literature to start with: Qu et al. 2021 Frontiers in Plant Science 12:727077 and supplementary material (get an idea of the type of data set), potentially chapters on *k nearest neighbors* and splines in the following book: `https://www.statlearning.com/`

Master thesis: Representative subset selection in an elite winter wheat panel

From a panel of 191 elite winter wheat lines representing 50 years of breeding progress, a representative subset is to be selected as a training set for genomic selection in a cross-validation study. The main objective of this thesis is to compare different methods of representative subset selection such as partitioning around medoids (PAM) and maximum connectedness and diversity (MaxCD) and others, and their effects on prediction accuracy of genomic selection. Scenarios which will serve as benchmark for the performance of the new methods are a) random selection and b) an established selection method based on principal component analysis.

Your task will be to implement the the different methods of representative subset selection in R code, mostly using available functions from R packages, to program and run the cross-validation routine for different sizes of the training and the validation set, and to evaluate and compare the methods by different criteria (e.g. prediction accuracy, mean square error, ranking of selection candidates).

Target group: Master students of Crop Sciences (Nutzpflanzenwissenschaften) or Agrobiotechnology

Supervisor: Dr. Eva Herzog (`eva.herzog@agrar.uni-giessen.de`)

Requirements: MK-002 or MK-002-EN and either MK-119-EN or MP-236-EN completed successfully, R programming skills, willingness to learn use of new R packages.

Literature to start with: Voss-Fels et al. 2019 Nature Plants 5(7) 706-714 (dataset), Guo et al. 2019 Molecular Plant 12, 390-401 (method)