

# Biometry and Population Genetics

Thesis topics  
16. Mai 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.

## 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@agrار.uni-giessen.de](mailto:hugo.tavera@agrار.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.

## Marker imputation with machine learning

With the reduced cost and increased prevalence of genotyping, companies are accumulating large amounts of marker data. The ongoing development of SNP arrays has resulted in different arrays with a limited number of common markers. Therefore, one challenge is to use datasets generated with different SNP arrays in a common genomic selection approach.

In your thesis, you will start by creating a small literature overview of marker imputation. The core of the thesis will be the application, comparison and evaluation of different imputation methods, both traditional and based on machine learning.

**Target group:** Master students of Nutzpflanzenwissenschaften or Agrobiotechnology

**Supervisor:** Philipp Heilmann ([philipp.g.heilmann@agrار.uni-giessen.de](mailto:philipp.g.heilmann@agrار.uni-giessen.de))

**Requirements:** MK-002 or MK-002-EN and either MK-119-EN or MP-236-EN completed successfully, strong interest in programming, ideally already some experience with machine learning.

## 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 one 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@agrار.uni-giessen.de](mailto:carola.zenke@agrار.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

## 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](mailto: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