Georg-August Universität Göttingen, DNPW, Abteilung für Zuchmethodik der Pflanze

apl. Prof. Dr. Wolfgang Link wlink@gwdg.de   28. Mai 2024
MSc Henri Laugel, Wissenschaftlicher Mitarbeiter, Doktorand

Zwischenbericht / Interim Report

Project name: **Abo-Direkt**

Duration of the project: 01.12.2022 – 30.11.2025

Reporting period: 30.04.2023 – 30.04.2024

1. Most important scientific and technical results

Work package P1 to P3: Genotyping of faba bean individuals

A set of 50 KASP markers (SNP) was designed to estimate the inbreeding coefficient of faba bean individuals (individual plants) taken from the famous Göttingen Winter Bean Population (GWBP; docplayer.org/72619769-Neues-von-der-goettinger-winter-ackerbohne.html). These 50 markers originate from the 60k "Vfaba_v2" Axiom chip (DOI 10.1002/leg3.75). Their selection was based on their localisation and distribution along faba bean chromosomes, their missingness and minor allele frequency in pre-acquired genotypic data from 79 random individuals from that GWBP.

A total of 627 individuals were genotyped using this 50-KASP set; 614 plants were derived from the GWBP-generation as harvested in 2020, and 13 from the generation of 2022. Another 765 individuals were genotyped using the "Vfaba_v2" Axiom chip; 355 were derived from the generation harvested in 2020, and 410 from 2022.

Different from the initial plan, a 20-KASP set was not developed; we realized from the data obtained from the 50-KASP set that 20 markers are not enough to get a reasonably meaningful estimate of the inbreeding status of an individual. The optimal number of markers to be used in this approach still needs to be determined, this is in progress.

Two generations of the synthetic faba bean cultivar Ghengis were genotyped by NPZ using another set of 43 KASP markers (designed by the company NPZ). 458 individuals from the Syn 4 generation of that cultivar, and 349 individuals from its Syn 8 generation were genotyped (the two generations differ mainly in their LD, which was the purpose of taking both generations).

Several algebraic formulas for the estimation of the inbreeding coefficient of individual plants based on SNP data, especially listed in Caballero et al. 2022 (DOI 10.1186/s12711-022-00772-0), were tested. Based on our simulations, the most cer-

tain method appeared to be a likelihood estimator, also used by Wang (2011). Its associated likelihood function is based on the genotypic frequencies in a population at equilibrium considering inbreeding. With the information of the marker-genotype of the focal individual and allele frequency in the population of each locus, we obtained a curve of likelihood of observing the focal genotype according to inbreeding coefficient for this focal individual. The estimated inbreeding coefficient of an individual is taken as that very value that maximizes this likelihood.

The inbreeding generation of each individual (such as F2, F3, F4 or the like; referring to the number of successive (uninterrupted) generations of self-fertilizations prior to that individual), was also estimated, in addition to its presumed inbreeding coefficient; We used the expected proportion of individuals in the population as a 'prior' to do so (Bayesian approach). This expected proportion is itself based on the average outcrossing rate in the faba bean population, indirectly obtained from the average individual inbreeding coefficient of the population. We also accounted for the difference in the self-fertilization rate between inbred and non-inbred individuals, as observed in several previous studies (DOI 10.1007/BF00226888).

In each sub-population, highly inbred individuals were indeed detected (Fig. 1; F>0.75). From the individuals which were Axiom-chip genotyped, 22 and 19 individuals were born from 4 or more generations of successive selfing, in respectively the generation 2020 and 2022 of the GWBP.
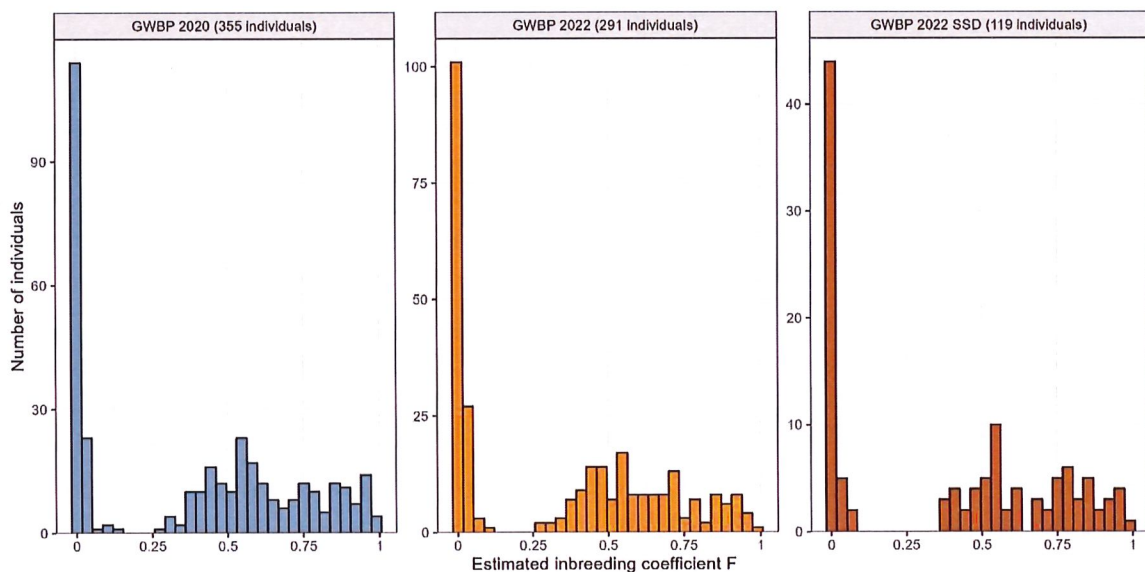

Fig. 1 Distribution of faba bean individuals (three samples derived from the GWBP); according to their inbreeding coefficient F, estimated based on SNP data and via likelihood estimator (further details in text).

This result confirmed our hypothesis that highly inbred individuals are *a priori* (without breeding actions) present in a faba bean population – as expected by theory and in case of being propagated with natural access of pollinators for several generations (*cf.* https://www.uni-goettingen.de/de/document/download/f75c1efcd662b3e296478dc0f5c0789d.pdf/PSSC2023_poster_Henri_Laugel.pdf)

.

Work package P4 to P6: calibration and validation of genomic prediction

We use our so-called association set of 189 highly inbred winter faba bean lines, derived from the famous GWBP, as training to calibrate our genomic prediction model. The MSc student André Mario Pupkes retrieved data from field trials for the lines and their ex-trial progenies; this field-data originates from field trials carried out over a period of 16 years. Data of a total of 28 traits is available. Pupkes generated best linear unbiased estimators of each line for each trait using both the line and progeny dataset. A poster was created on the GWAS he conducted using this data and presented at the GPZ German plant breeding conference 2024 organized at the University of Geisenheim.
https://www.uni-goettingen.de/de/document/download/f27f281282aebbec9fcc3ba6ccc7515b.pdf/GPZ24_poster_vffff-Geisenheim%20Pupkes%20Award.pdf.
The poster earned one of the three best-poster awards of the conference (cf. https://www.uni-goettingen.de/de/preistr%c3%a4ger+gpz-konferenz+2024+in+geisenheim/667273.html).

With help of the PhD student Alex Windhorst, a first genomic prediction model was created, and field performance of pertinent traits of the chip genotyped individuals from the GWBP were predicted and prediction ability found to be useful for application.

Selfed offspring of 70 highly inbred individuals, detected in the GWBP via chip genotyping, are currently propagated through controlled, tripping-assisted self-fertilization in isolation houses and isolation cages for multiplication. The seeds obtained from this multiplication step will be used to test the performance of the genotypes (more precisely, of their selfed offspring) in the field next season across 3 locations (sowing fall 2024). Along with genotypic data of the ancestral highly inbred individual, the phenotypic data of the selfed offspring obtained from this multi-location field trial will allow to check and validate the prediction ability of the genomic prediction model.

We aim to validate the prediction for plant height, beginning of flowering, lodging, maturity, thousand grain weight and yield. Depending on circumstances, especially depending on whether we can recruit a suitable BSc and MSc student, we will also validate drought stress tolerance, winter hardiness, resistance against *Ascochyta* and Vicin-Convicine wild-type variation. We do have the pertinent data as prerequisite for such scientific theses (DOI 10.1111/pbr.12918; DOI 10.5073/&JfK.2022.01-02.01).

In addition, we decided to initiate trials to study the so-called Mendelian sampling variance among the selfed progenies of non-inbred individuals from the GWBP by phenotyping a fair number of individual offsprings from such non-inbred plants. The objective is to study the relation between the variance obtained from segregation and the heterozygosity proportion in the parent (so-called 'usefulness concept'; DOI 10.1534/genetics.117.300403). The main motivation to start this additional (in the application not mentioned) task is to estimate the dominance deviation of genotypes by comparing the phenotypic values of each genotypic group for the genotyped loci. For one trial, the progenies were Axiom chip genotyped and propagated through selfing. Their progenies will be then tested in the field next season (2025). In another trial, the progenies will be assessed in semi-controlled conditions and will be genotyped for few SNP markers this season (2024). The markers that we will use were found to be significantly associated with plant height, beginning of flowering, thousand grain weight and yield,

in the GWAS of André Mario Pupkes (exploiting our historical data as mentioned earlier).

Estimating dominance should improve the accuracy of genomic prediction and may help the breeders to optimize the choice of components for synthetic cultivars and the choice of parents for crossings by maximizing genetic variation for specific traits.

Work package P7: Theoretical work

We searched for algebra to determine the theoretical proportion of individuals according to inbreeding generations in faba bean populations. It needs to be further investigated.

The optimization of the newly developed breeding methodological concepts was mainly discussed during meetings, in which the project was presented. It will be further discussed with breeders in the next months, and further studied via the simulation of breeding programs.

Work package P8: Field trials

A field trial is conducted this season (2024, sowing in Oct. 2023) to check the relation between vicine-convicine (VC) content and tolerance to legume fatigue (threat to faba bean by soil-borne pathogens such as *Fusarium sp.*). Bulk of low VC and high VC lines from the association set are currently tested in a field with legume fatigue (pathogen pressure at that field spot was accumulated via cultivated with only faba bean for more than 7 seasons (orgprints.org/id/eprint/39470/1/2815EPS004_Abschlussbericht.pdf).

A trial is currently (season 2024, sowing in Oct. 2023) conducted at Göttingen-Reinshof, to assess the field performance of the progenies of individuals derived from the GWBP, which were estimated to be non- or partially inbred. For validation of our genomic prediction model, we propagated only the highly inbred plants - so the other individuals, being non-inbred, along with the expensively propagated selfed seeds were not simply left un-used, but used in that actual trial for breeding purposes. Also, the trial gives already first, crucial hints on the accuracy of the prediction, since we acquired chip genotypic data from the parents of sample of the tested progenies. This current field trial is replicated in Hohenlieth in collaboration with NPZ. Autumn-sowing at Hohenlieth (planned for October 2023) could not be carried out due to rain and moist soil then. Instead, that trial was sown in January 2024; the impact of such untimely sowing has still to be observed.

Work package P9: Chipped-seed experiments

Nota bene: Axiom chip genotypes are the ones for which SNP data were gathered via a SNP array (Axiom chip) – and DNA was extracted from leaf tissue. Here, we talk about chipped seed: this means that DNA was not collected from leaves (the default DNA source) but from the seed (cotyledons).

4

The objective of the chipped-seed experiments changed a bit from the beginning of the project. Some other institutions already experimented the DNA extraction from faba bean seeds (www.uni-giessen.de/de/fbz/fb09/institute/plantbreeding/ipz/staff/snow-don-r). After meeting and discussing the technique, we could refine the objective of our project. The chipped-seed experiments aim in testing several seed priming and chipping techniques beyond of what is known so far, to determine an efficient method for KASP genotyping from faba bean seeds, plus ease of technique, plus ease of storing chipped seeds for later selection and sowing.

We are finalizing the preparation of this project. Several specific tools and chemicals are meanwhile available.

We will first test the convenience of chipping techniques and make sure no sample is contaminated by seed coat tissue. Then, we will analyse whether soaking the seeds in water ('imbibtion') eases the chipping. If so, we will test the effect of several such priming treatments on ease of chipping and seed storage. Finally, we will test whether the chipping influences seed germination rate and seed viability.

2. Comparison of the project status with the original work, time and cost planning

Work package P1 to P3: The tasks were completed. However, the planned theoretical studies need to be finalized to determine the cost-efficient set of KASP markers for estimation of inbreeding coefficient. Moreover, details of the application of the identification of highly inbred individuals from faba bean population needs to be further studied and discussed in collaboration with faba bean breeders.

Work package P4 to P6: The task for calibration of genomic prediction started recently with the availability of the data from the association set of lines. Investigation to find the best suited model for faba bean prediction is ongoing. On the other hand, further multiplication of the highly inbred individuals from the GWBP is needed to have the necessary amount of seed, for the field-validation of genomic prediction, and this in progress.

Work package P9: The chipped-seed experiments were only started now (May 2024). Our new objective is to answer our research questions before this summer season and work on the implementation of the technique, so that it could be applied from the following season on.

3. Change in prospects for achieving the objectives of the project compared to the original application

Disregarding the delay in the chipped-seed experiment, the further tasks as originally planned could be completed. The team at Göttingen or at NPZ (Hohenlieth) observed no change in prospects for achieving the objectives of the project.

4. Third-party results relevant to the implementation of the project

Work package P1 to P3: Among both generation of the synthetic cultivar genotyped, several individuals were found to be homozygous for all markers tested. This result supports the presence of highly inbred individuals.

Work package P9: Experiences at the University of Giessen and the University of Helsinki on DNA extraction of chipped-seed of faba bean allowed to better-plan our experiment. Both successfully achieved DNA extraction from the seeds and developed protocols. However, they used specific techniques and we conclude that further chipping techniques can be tested for their convenience and avoid contamination with tissues from the seed coat. In addition, we want to test priming the seeds and the effect of the technique on germination rate and seed viability.

5. Necessity of changes to the objectives

The objectives of the project did not change. Highly inbred individuals were as expected, present in our population. Only the objectives of the chipped-seed experiments were not changed, but adapted regarding the work running at other institutions.

6. Annual update of the utilisation plan

There was no granted property right made or claimed by the applicant or parties involved in the project.

The prospects for a scientifically successful implementation of the project are very high. It still exists some risks related to weather damage (hail, wintering out, vandalism) to the field trials, but reduced by replicating the trials across several locations. All genotypic data necessary for the success of the project is already acquired.

apl. Prof. Dr. Wolfgang Link	Göttingen, 24.05.2024	
	Location, Date	Signature