Gefördert durch:

Bundesministerium
für Landwirtschaft, Ernährung
und Heimat

aufgrund eines Beschlusses
des Deutschen Bundestages

Projektträger

Bundesanstalt für
Landwirtschaft und Ernährung

Schlussbericht zum Thema „Direkt zu Resistenz, Qualität und Ertrag mit Ackerbohnen-Linien“ (Abo-Direkt)

FKZ: 2820EPS003

Projektnehmer: Georg-August-Universität Göttingen

Gefördert durch das Bundesministerium für Landwirtschaft, Ernährung und Heimat auf Grund eines Beschlusses des Deutschen Bundestages im Rahmen der BMLEH Eiweißpflanzenstrategie.

Gefördert durch:

Bundesministerium
für Landwirtschaft, Ernährung
und Heimat

Projektträger

Bundesanstalt für
Landwirtschaft und Ernährung

aufgrund eines Beschlusses
des Deutschen Bundestages

Gefördert durch das Bundesministerium für Landwirtschaft, Ernährung und Heimat aufgrund eines Beschlusses des Deutschen Bundestages im Rahmen der BMLEH Eiweißpflanzenstrategie“.

Abschlussbericht zum Projekt:
Direkt zu Resistenz, Qualität und Ertrag mit Ackerbohnen-Linien (Abo-Direkt)

Gefördert durch: Bundesministerium für Landwirtschaft, Ernährung und Heimat

Förderkennzeichen: **2820EPS003**

Projektlaufzeit: 01.12.2022 – 31.03.2026

Projektnehmer: Georg-August-Universität Göttingen, Fakultät für Agrarwissenschaften, Abteilung für Zuchtmethodik der Pflanze, apr. Prof. Dr. Wolfgang Link

Projekt-Kooperationspartner: NPZ Lembke KG, Dr. Gregor Welna, Hohenlieth;

Projekt-Unterauftragsnehmer: Naturland, Werner Vogt-Kaute

Autoren dieses Berichtes:

Henri Laugel (henri.laugel@uni-goettingen.de)
Wolfgang Link (wlink@gwdg.de)

1. Juni 2026, Göttingen

Die Resultate des Projektes Abo-Direkt sind auch in der Dissertation von Henri Laugel niedergelegt, teils ähnlich wie in diesem Abschlussbericht, teils detaillierter oder auch mehr auf die wissenschaftlichen Aspekte fokussiert. Diese Dissertation wird in

Kürze unter dem Titel „Novel breeding approaches for direct access to superior inbred lines in faba bean (*Vicia faba* L.)“ online bei https://ediss.uni-goettingen.de/ zu finden sein. Außerdem wurde eine erste Publikation inzwischen zur Veröffentlichung angenommen (Theoretical and Applied Genetics. Faba bean populations already contain the inbreds needed for breeding). Es ist die überarbeitete Version dieser Pre-Print-Fassung: DOI 10.21203/rs.3.rs-8211981/v1. Eine zweite Publikation zur Frage der genomischen Selektion ist in Vorbereitung.

Kurzfassung:

Das Ziel von Projekt „Abo-Direkt“ war, einen neuen, verbesserten Züchtungsansatz für die Ackerbohne (*Vicia faba* L.) zu entwickeln. Es konzentrierte sich auf zwei methodische Lücken:
(1) Direkter Zugang zu Inzuchtlinien der Ackerbohne (Ersatz für fehlende Doppelhaploid-Technologie)
(2) Genomische Vorhersage (Genomic Prediction, GP).

Die Erzeugung von Inzuchtlinien in der Züchtung von Ackerbohnen ist entscheidend, aber wegen der partiellen Allogamie der Pflanze ein langwieriger und arbeitsintensiver Prozess. In genetisch diversen und offen bestäubten Ackerbohnenpopulation führt ihr Fortpflanzungssystem dazu, dass ein Anteil der Individuen ingezüchtet ist. Wir haben einen kosteneffizienten Ansatz entwickelt, um hoch ingezüchteten Individuen aus Populationen mit wenigen KASP-Markern zu identifizieren. Der genaueste Schätzer für den Inzuchtkoeffizienten F bei 10 bis 100 Markern war der sogenannte FL-Schätzer ist. In der „Göttinger Winterbohnenpopulation“ wurden 8,4 % der Individuen (61 von 728) als stark ingezüchtet eingestuft (F ≥ 0,83). Wir bewerteten die Leistungsfähigkeit der Klassifizierung sowohl simulierten und mit empirischen Daten. Mit 50 Markern und einem Schwellenwert von F ≥ 0,9 identifizierten wir 42 Individuen in unserer Population, darunter 36 tatsächlich hoch ingezüchtet (von 61) und sechs unzureichend ingezüchtet (falsch-Positive). Vorfilterung mit einer geringen Anzahl von Markern reduzierte die Genotypisierungskosten um bis zu 46 %.

Aktuelle Studien bewerten die GP vorwiegend durch Kreuzvalidierung, was möglicherweise nur teilweise relevant ist, da Selektion, Drift und Rekombination im Material außer Acht bleiben. Für eine realistischere Bewertung verwendeten wir zwei Genotypensätze (beide aus der „Göttinger Winterbohnenpopulation“). Der A-Satz wurde aus der 9. Erhaltungs-Generation der Population gebildet, der AD-Satz aus der 32. Verwandtschaftsanalyse sollten Divergenz zeigen, die sich über diese 23 Generationen ansammeln konnten.

Für den A-Satz lagen historische Feldversuchs-Daten vor. Mit dem AD-Satz wurden Feldversuche an mehreren Standorten durchgeführt. Unter Verwendung eines genomischen Best-Unbiased-Prediction-Modells führten wir Kreuzvalidierungen innerhalb (generationsintern) beider Sätze durch; generationsübergreifende Vorhersagen sagten den einen Satz anhand (des GP Modells aus Phänotyp- und Genotyp-Daten) des anderen Satzes vorher. Die Kreuzvalidierungen ergaben relativ hohe Vorhersagefähigkeiten. Die genotypischen Daten deckten erhebliche populationsgenetische Unterschiede zwischen beiden Sätzen auf, was wohl zu den relativ geringen Vorhersagefähigkeiten bei den generationsübergreifenden Vorhersagen beitrug. Dieses Ergebnis war zwar enttäuschend, ist jedoch realistisch bei der Bewertung der GP in der Ackerbohnenzüchtung.

Ein höheres Potenzial der GP bei Ackerbohnen entstünde, wenn Vorhersage und Selektion an (DNA von) Ackerbohnen-Samen (vor Aussaat) durchgeführt würde. Die zerstörungsfreie DNA-Extraktion aus Samen ist neu (Probenahme aus Blättern ist Standard). Unser Ziel war es, dieses sog. Chippen von Samen zu optimieren. Es wurden verschiedene Chipping-Verfahren an trockenen Samen als auch an durch Einquellen in Wasser weichen Samen getestet. Als effizienteste Technik erwies sich das Anbohren trockener Samen. Diese Proben konnten mittels KASP-Analysen erfolgreich genotypisiert werden. Anbohren beeinträchtigte die Keimfähigkeit nicht. Es sind weitere Entwicklungen erforderlich, um das Chippen im Hochdurchsatz zu ermöglichen.

Die Identifizierung von Inzuchtlinien aus Populationen bietet eine schnellere, kostengünstigere und weniger arbeitsintensive Alternative zur herkömmlichen Entwicklung von Inzuchtlinien. Die Integration von GP ermöglicht eine rekurrente genomische Selektion in der Zuchtpopulation, aus der dann die Inzucht-Individuen kommen. Saatgut-Chipping beschleunigt den Züchtungsprozess und erweitert das Potenzial der Genomischen Vorhersage und Selektion.

Abstract:

The objective of the “Abo-Direkt” project was to develop a new, improved breeding approach for the faba bean (Vicia faba L.). It focused on two methodological gaps:

(1) Direct access to faba bean inbred lines (as a substitute for the lack of double-haploid technology)

(2) Genomic prediction (GP).

The production of inbred lines in faba bean breeding is crucial, but due to the plant’s partial allogamy, it is a lengthy and labour-intensive process. In genetically diverse and open-pollinated faba bean populations, their reproductive system results in a proportion of individuals being inbred. We have developed a cost-effective approach to identify highly inbred individuals from populations using a small number of KASP markers. The most accurate estimator for the inbreeding coefficient F with 10 to 100 markers was the so-called FL estimator. In the “Göttingen winter faba bean population,” 8.4% of the individuals (61 out of 728) were classified as highly inbred ($F \geq 0.83$). We evaluated the performance of the classification using both simulated and empirical data. Using 50 markers and a threshold of $F \geq 0.9$, we identified 42 individuals in our population, including 36 that were actually highly inbred (out of 61) and six that were insufficiently inbred (false positives). Pre-filtering with a small number of markers reduced genotyping costs by up to 46%.

Current studies evaluate GP primarily through cross-validation, which may be only partially relevant, as selection, drift, and recombination in the material are not taken into account. For a more realistic evaluation, we used two genotype sets (both from the “Göttingen winter bean population”). The A set was derived from the 9th generation of maintaining the population, and the AD set from the 32nd. Analyses of genetic relationship should reveal divergence that could have accumulated over these 23 generations.

Historical field trial data were available for the A set. Field trials were conducted at multiple sites with the AD set. Using a genomic best-unbiased-prediction model, we performed cross-validations within (intra-generation) both sets; cross-generation predictions predicted one set based on the other set (using the GP model derived from

its phenotype and genotype data). The cross-validations yielded relatively high predictive abilities. The genotypic data revealed significant population genetic differences between the two sets, which likely contributed to the relatively low predictive ability in the cross-generation predictions. Although this result was disappointing, it is realistic when evaluating GP in faba bean breeding.

Greater potential for GP in faba beans would arise if prediction and selection were performed on (DNA from) faba bean seeds (prior to sowing). Non-destructive DNA extraction from seeds is a new approach (sampling from leaves is standard). Our goal was to optimize this so-called chipping of seeds. Several chipping approaches were tested on dry seeds as well as on seeds softened by soaking in water. Drilling into dry seeds proved to be the most efficient technique. These samples were successfully genotyped using KASP analyses. Drilling did not impair germination. Further developments are needed to enable high-throughput chipping.

The identification of inbred lines from populations offers a faster, more cost-effective, and less labour-intensive alternative to the conventional development of inbred lines. The integration of GP enables recurrent genomic selection in the breeding population from which the inbred individuals are derived. Seed chipping accelerates the breeding process and expands the potential of genomic prediction and selection.

Inhaltsverzeichnis

1. Einführung

1.1. Gegenstand des Vorhabens

1.2. Ziele und Aufgabenstellung des Projekts, Bezug des Projekts zu den relevanten Zielen der EPS oder zu spezifischen Bekanntmachungen und Ausschreibungen (www.ble.de/eiweisspflanzenstrategie).

1.3. Projektplanung und Ablauf

1.3.1. Planung und Ablauf der Arbeitspakete 1; 2; 3; 7: Kosteneffiziente Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen

1.3.2. Planung und Zeitplan der Arbeitspakete 4; 5; 6; 8: Leistung der genomischen Vorhersage innerhalb und über Generationen hinweg in einer Winterackerbohnenpopulation

1.3.3. Planung und Zeitplan des Arbeitspaketes 9: Optimierung des Seed Chipping für die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen

2. Stand der Wissenschaft und Technik, an den das Projekt anknüpfte

3. Material und Methoden

3.1. Material und Methoden der Arbeitspakete 1; 2; 3; 7: Kosteneffiziente Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen

3.2. Material und Methoden der Arbeitspakete 4; 5; 6; 8: Leistung der genomischen Vorhersage innerhalb und über Generationen hinweg in einer Winterackerbohnenpopulation

3.3. Material und Methoden des Arbeitspakets 9: Optimierung des Seed Chipping für die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen

4. Ausführliche Darstellung der wichtigsten Ergebnisse

4.1.1. Ergebnisse der Teilaufgabe im Arbeitspakets 7: LD-Abbau in Populationen bei partieller Allogamie

4.1.2. Ergebnisse des Arbeitspakets 1; 2; 3; 7 Kosteneffiziente Identifizierung hoch inzüchtiger Individuen aus Ackerbohnenpopulationen

4.2. Ergebnisse der Arbeitspakete 4; 5; 6; 8: Leistungsfähigkeit der genomischen Vorhersage innerhalb und über Generationen hinweg in einer Winteracker-bohnenpopulation

4.3. Ergebnisse des Arbeitspakets 9: Optimierung des Seed Chipping für die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen

5. Diskussion der Ergebnisse

6. Angaben zum voraussichtlichen Nutzen und zur Verwertbarkeit der Ergebnisse.

7. Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen; Hinweise auf weiterführende Fragestellungen

8. Zusammenfassung

9. Verzeichnis der benutzten Fachliteratur

10. Übersicht über alle im Berichtszeitraum vom Projektnehmer realisierten Veröffentlichungen zum Projekt (Printmedien, Newsletter usw.), bisherige und geplante Aktivitäten zur Verbreitung der Ergebnisse

Anhang: Liste agronomisch überlegener Genotypen aus dem Abo-Direkt Projekt

Abkürzungsverzeichnis

ANOVA	Varianzanalyse
A-Set	Assoziations-Satz
AD-Set	Abo-Direkt-Satz
AIS	Alike-in-state
Asco	Ascochyta-Resistenz (*Ascochyta fabae*; Bonitur)
BLUE	Best linear unbiased estimator
Cano	Bestandesentwicklung der Parzellen nach Winter; Bonitur
Choc	Schokoladenflecken-Resistenz (*Botrytis fabae*; Bonitur)
CRD	Vollständig randomisierte Anlage
DAS	Tage nach Saat
Dev1	Pflanzenentwicklung vor Winter, Bonitur
Dev2	Pflanzenentwicklung nach Winter, Bonitur
DH	Doppel-Haploid
F	Inzucht-Koeffizient

Flbeg Blühbeginn, Anzahl Tage bis Blüte (ab 1. Januar)
Flend Blühende, Anzahl Tage bis Blühende (ab 1. Januar)
FLD Fisher's linear discriminant
FP Falsch Positiv
G Keimung (Germination; binäre Daten)
GBLUP Genomic best linear unbiased prediction
GBS Genotyping-by-sequencing
gDAS Keimung (Anzahl Tage nach Saat)
GEBV Genomic-estimated breeding value
GP Genomische Vorhersage, Genomic prediction
GR Keimungs-Rate
GRM Genomic relationship matrix
GVA Allgemeine Sorten-Eignung, General varietal ability
GWAS Genome-wide association study
GWBP Berühmte Göttinger Winterackerbohnen-Population
H^2 Erblichkeit im weiteren Sinn
HFT Vorhandene ('historische') Feld-Versuche
IBD Identical-by-descent
ICT Inbreeding classification threshold
KASP Competitive allele-specific PCR
LD Gametenphasen-Ungleichgewicht, Linkage disequilibrium
Lg1 Frühes Lager, Bonitur
Lg2 Spätes Lager, Bonitur
Ma1 Reifebonitur, frühe Erfassung
Ma2 Reifebonitur, späte Erfassung
PC Polycross
PCA Principal component analysis
PCR Polymerase Chain Reaction
Pero Peronospora-Resistanz (*Peronospora viciae*; Bonitur)
Ph1 Wuchshöhe, frühe Erfassung (in cm)
Ph2 Wuchshöhe, späte Erfassung (in cm)
Pla1C Pflanzenzählung vor Winter
Pla2C Pflanzenzählung nach Winter
S Selbstbefruchtungs-Rate
SL Sämlings-Länge
SNP Single nucleotide polymorphism
SPET Specific Primer Enrichment Technology
SSD Einkorn-Ramsc, Single-seed descent
SVI Sämlings-Wüchsigkeits-Index, Seedling vigor index
TGW Tausendkorngewicht (g)

TP	Wahr Positive,True positive
UPGMA	Unweighted pair group method with arithmetic mean
Yld	Ertrag, Yield (in dt/ha)

Die **Predictive (**aka **Prediction) Ability** wird auch Vorhersagefähigkeit genannt.

Die **Accuracy** wird auch Vorhersagegenauigkeit genannt.

1. Einführung

1.1. Gegenstand des Vorhabens

Das Projekt zielte darauf ab, methodische Verbesserungen für die Züchtung von Ackerbohnen zu entwickeln und umzusetzen. Es konzentrierte sich hauptsächlich auf (1) die Identifizierung *a priori* verfügbarer Inzuchtindividuen innerhalb von Ackerbohnenpopulationen durch Nutzung der natürlichen genetischen Zusammensetzung dieser Populationen unter ‚offen-bestäubten' Bedingungen (Zutritt von Bestäubern), (2) die Implementierung genomischer Vorhersagen in der Ackerbohnenzüchtung und (3) die Optimierung der zerstörungsfreien DNA-Extraktion aus Ackerbohnensamen durch sog. ‚Seed Chipping'. Diese Ansätze werden den Ackerbohnen-Züchterinnen einen direkten Zugang zu neuen und überlegenen Inzuchtlinien verschaffen und bieten eine schnellere und effizientere Alternative zur langsamen und mühsamen konventionellen Entwicklung von Inzuchtlinien. Sie sollen wichtige methodische Lücken im Vergleich zu den wichtigsten Kulturpflanzen schließen. Darüber hinaus wurden im Rahmen der Feldversuche, die darauf abzielten, die Leistungsfähigkeit der genomischen Vorhersage und entsprechend der genomischen Selektion zu bewerten, direkt neue Zuchtlinien entwickelt.

Die im Projekt behandelten Forschungsfragen sind nachfolgend aufgeführt. Sie sind, im Vergleich zum Antrag von Juli 2022 (siehe dort Seite 2-3) weiter entwickelt.

- Gibt es in typischen Ackerbohnenpopulationen die theoretisch zu erwartenden hoch-ingezüchtete Individuen?
- Wie hoch ist ihr Anteil?
- Wie viele SNP-Marker sind optimal, um hoch-ingezüchtete Individuen zu identifizieren?
- Welcher Ansatz eignet sich am ehesten, um den Inzuchtkoeffizienten von Individuen anhand dieser SNP-Daten zu schätzen?
- Welche zerstörungsfreie Chipping-Technik ist effizient, um Gewebe-Proben aus Ackerbohnensamen für die DNA-Extraktion und anschließende Genotypisierung dieser Samen (einzelsamenweise) zu gewinnen?
- Beeinflusst das Chipping der Samen deren Keimung (Keimfähigkeit und Keimlingsvitalität)?

- Welche der neuen Genotypen weisen in der Göttinger Winterbohnenpopulation (GWBP) die beste Leistung auf?
- Wie ist die Genauigkeit der genomischen Vorhersage in verschiedenen Vorhersage-Szenarien?
- Sind die mit Kreuzvalidierung erhaltenen Genauigkeiten realistisch? Um wieviel nimmt die Genauigkeit ab, wenn das genomische Vorhersagemodell mit Daten von Genotypen trainiert wird, die aus einer weiter zurückliegenden Ge-Generation derselben Population stammen, aus der auch das (jüngere) Test- und Zuchtmaterial stammt?
- Erlauben diese untersuchten Ansätze und Techniken eine Verbesserung der Ackerbohnen-züchtung, wenn ja, wie sollte die Umsetzung aussehen?

1.2. Ziele und Aufgabenstellung des Projekts, Bezug des Projekts zu den relevanten Zielen der EPS oder zu spezifischen Bekanntmachungen und Ausschreibungen (www.ble.de/eiweisspflanzenstrategie).

Laut Antrag waren unsere Ziele:
(1) ein direkter, schneller Weg zu homozygoten Linien sollte erschlossen werden und es sollte
(2) Genomische Vorhersage (GV) in die Züchtung implementiert werden.

Im Einzelnen und unter Anpassung an den Stand der verfügbaren Technik sind damit dieses die Ziele:
(1) Entwicklung einer Methode zur Identifizierung und damit zur Selektion der *a priori* verfügbaren hoch-ingezüchteten Individuen in Zuchtpopulationen von Ackerbohnen unter Verwendung einer begrenzten Anzahl von SNP-Markern (KASP-Technik), um den Zeit-, Arbeits- und Kostenaufwand für die Entwicklung von Inzuchtlinien zu reduzieren,
(2) das Potenzial der Genomischen Vorhersage für deren Umsetzung in der Ackerbohnenzüchtung zu bewerten und
(3) die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen (Seed Chipping) zu optimieren, um mittels dieser Technik die Genomische Selektion effizienter durchzuführen.

Das Projekt befasst sich demnach direkt mit mehreren der Ziele, die in der BLE-Förderbekanntmachung Nr. 07/20/33 vom 10. August 2020 (BAnz AT 02.09.2020 B5) genannt wurden.

1.3. Projektplanung und Ablauf

Der Zeitplan des Projekts ist in Tabelle 1 aufgeführt. Die erste Phase des Projekts (bis Anfang 2025) konzentrierte sich auf die Entwicklung einer kosteneffizienten Methode zur Identifizierung von Inzuchtlinien aus Ackerbohnenpopulationen sowie auf die Optimierung der Chipping-Methode zur zerstörungsfreien DNA-

Extraktion aus Ackerbohnensamen. Zudem wurde neues Zuchtmaterial aus unserer Winterackerbohnen-Population entwickelt und durch kontrollierte Selbstbefruchtung in Isolationskäfigen und Gewächshäusern wurde davon Saatgut für Feldversuche vermehrt. Die zweite Phase des Projekts zielte darauf ab, den Hauptfeldversuch des Projekts (im Jahr 2025) unter Verwendung des aufwendig gewonnenen Saatgutes (kontrollierter Selbstbefruchtung) durchzuführen und die Leistungsfähigkeit der genomischen Vorhersage mit diesem neuen Material realistisch zu bewerten.

Die NPZ Lembke KG leistete einen Beitrag zu den Arbeitspaketen 2 und 3, indem sie genotypische Daten von zwei Generationen einer ihrer synthetischen Sorten erhob. Die Auswertung dieser Daten trug zudem zur Erfüllung des Arbeitspakets 7 bei. Im Rahmen des Arbeitspakets 8 führte die NPZ Lembke KG zwei Feldversuche durch (einen im Jahr 2024 und einen im Jahr 2025), und Werner Vogt-Kaute lies im Jahr 2025 einen Feldversuch durchführen. Mehrere Exkursionen nach Hohenlieth und Wartmannsroth sowie Feldbesuche in den Jahren 2024 und 2025 förderten den Austausch mit den Projektpartnern.

Tabelle 1: Überblick über den Zeitplan während der Projektlaufzeit. Die Genotypdaten der Sommer-Ackerbohnen anhand von 43 KASP-Markern (im Rahmen der Arbeitspakete 2 und 3) wurden von der Fa. NPZ erhoben. Das Arbeitspaket 8 (Feldversuche) wurde in Zusammenarbeit mit den Projektpartnern NPZ Lembke KG und Naturland (Werner Vogt-Kaute) durchgeführt.

1.3.1. Planung und Ablauf der Arbeitspakete 1; 2; 3; 7: Kosteneffiziente Identifizierung hoch ingezüchteter Individuen aus Ackerbohnenpopulationen

Aus unserer Winterackerbohnen-Züchtungs-Population (Göttinger Winterbohnenpopulation; GWBP) wurden Einzelpflanzen zufällig entnommen und genotypisiert. Es wurde eine Literaturrecherche zu publizierten Schätzern (Formeln) für den Inzuchtkoeffizienten (F) durchgeführt. Es wurden Simulationen durchgeführt, um die zu erwartende Verteilung der F-Werte der Individuen in einer Ackerbohnenpopulation zu analysieren und um zu ermitteln, welcher der publizierten F-Schätzer am ehesten geeignet war. Der genaueste F-Schätzer wurde dann verwendet, um die F-Werte der beprobten Individuen (aus der GWBP) anhand der gesammelten Markerdaten zu berechnen (zu schätzen). Die genetische Zusammensetzung unserer experimentellen Ackerbohnen-Züchtungs-Population (GWBP) wurde untersucht und mit den Ergebnissen unseren Simulationen verglichen. Die Leistungsfähigkeit der Klassifizierung von Individuen als „hoch ingezüchtet“ wurde anhand von simulierten und der empirischen Daten dargestellt, und die Implikationen für die Umsetzung der Methode in der praktischen Ackerbohnenzüchtung wurden im Detail untersucht.

Arbeitspakete	Jahr			
	2023	2024	2025	2026
WP 1; 2; 3; 7: Kosteneffiziente Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen				
Genotypisierung mit dem „Vfaba_v2“-60k-Chip (WP 1)	+ + + +	+		
Genotypisierung mit dem 50-KASP-SNP-Markersatz (WP 2.3)	+ + + +	+		
Genotypisierung mit dem 43-KASP-SNP-Markersatz (WP 2.3; NPZ)	+ +			
Schätzung des individuellen Inzuchtkoeffizienten von Ackerbohnenindividuen anhand begrenzter SNP-Daten (WP 7)	+ + + +	+		
Entwicklung einer kosteneffizienten Methode zur Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen (WP 7)		+ + + +	+	
WP 4; 5; 6; 8: Leistungsfähigkeit der genomischen Vorhersage (GP) innerhalb und über Generationen hinweg				
Entwicklung von neuem Material und Vermehrung durch kontrollierte Selbstbefruchtung (WP 6)	+ + + +	+ + +		
Kalibrierung des GP-Modells und 5-fache Kreuzvalidierung auf der Grundlage historischer Felddaten (WP 4)		+ +		
Feldversuche (WP 8; NPZ und Naturland)	+	+ + + +	+ + +	
(Generationsübergreifende) GP des hier neu entwickelten Zuchtmaterials (WP 5)			+	+
WP 9: Optimierung des Seed Chippings für die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen				
DNA-Extraktion und KASP-Analyse (WP 9)		+ + +		
Test verschiedener Chipping-Ansätze (WP 9)		+	+	
Prüfung der Keimfähigkeit nach Chipping (WP 9)			+	

1.3.2. Planung und Ablauf der Arbeitspakete 4; 5; 6; 8: Leistung der genomischen Vorhersage innerhalb und über Generationen hinweg in einer Winterackerbohnenpopulation

Das erwähnte, zufällig der Population entnommene Genotypen-Material zur Realisierung der Arbeitspakete 1, 2, 3 und 7 wurden durch tripping-gestützte, kontrollierte Selbstbefruchtung in Isolationskäfigen und Gewächshäusern vermehrt, um genügend Saatgut für Feldversuche zu gewinnen. Die daraus resultierenden Saatgut-Chargen wurden in den Jahren 2024 und 2025 in insgesamt vier Feldversuchen getestet. Parallel dazu wurde an Kalibrierung (Modellentwicklung für GBLUP) und Kreuzvalidierungen für und mit der genomischen Vorhersage gearbeitet; unter Verwendung vorliegender historischer Feld- und Genotyp-Daten von 185 Inzuchtlinien, die ebenfalls aus unserer GWBP stammten. Diese Daten waren von André Pupkes (M.Sc.-Student unserer Abteilung an der Universität Göttingen) aus Feldbüchern und digitalen Listen aufbereitet worden. Nach der Erfassung der phänotypischen Daten aus den neuen Feldversuchen (Abo-Direkt) wurde die Vorhersagegenauigkeit des genomischen Vorhersagemodells bei der Vorhersage des neuen Materials bewertet. Unsere Ergebnisse ermöglichten es, das Potenzial der genomischen Vorhersage in der praktischen Ackerbohnenzüchtung valide zu bewerten.

1.3.3. Planung und Ablauf des Arbeitspaketes 9: Optimierung des Saatgut-Chippings zur zerstörungsfreien DNA-Extraktion aus Ackerbohnensamen

Die Ziele dieses Arbeitspakets (WP 9) wurde modifiziert, da eine zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen mittlerweile schon einmal erfolgreich durchgeführt worden war. Nach internationaler Rücksprache und geduldiger Adaption eines Extraktions-Protokolls, das im Labor der benachbarten Abteilung für Nutzpflanzengenetik (Team Scholten) vorhanden war, wurden drei Chipping-Techniken getestet und ihre Effizienz sowohl an trockenen als auch an durch Quellen in Wasser aufgeweichten Samen untersucht. Nach Ermittlung der effizientesten Technik wurde deren Einfluss auf die Keimfähigkeit der Samen im Rahmen eines Gewächshausversuchs untersucht.

2. Stand der Wissenschaft und Technik, an den das Projekt anknüpfte

Die Ackerbohne ist eine wichtige Körnerleguminose mit hohem Nährwert (Duc et al. 1999; Karkanis et al. 2018) und zahlreichen agrarökologischen Vorteilen (Karkanis et al. 2018). Ihre Anbaufläche ist jedoch nach wie vor relativ klein. Im Jahr 2024 wurden in Deutschland Ackerbohnen auf etwa 61.700 ha angebaut, weniger als die Hälfte der Erbsenanbaufläche (FAOSTAT 2026). Landwirte zögern wahrscheinlich deswegen mit ihrem Anbau, da ihre Erträge relativ instabil sind, was hauptsächlich auf ihre Anfälligkeit für biotischen (Sillero et al. 2010; Stoddard et al. 2010) und abiotischen Stress (Flores et al. 2012) zurückzuführen ist. Züchtung verspricht eine nachhaltige und umweltfreundliche Lösung zur Verbesserung und Stabilisierung der Erträge.

Inzuchtlinien sind für die Züchtung von Ackerbohnen von entscheidender Bedeutung. Homozygote Inzuchtlinien sind „unsterblich“ und über Generationen hinweg stabil, wenn und weil sie durch Selbstbefruchtung vermehrt werden. Sie werden als Liniensorten oder als Komponenten von Sorten verwendet. Eine große

Herausforderung bei der Züchtung von Ackerbohnen ist die Entwicklung ebensolcher Inzuchtlinien. Dies ist ein langwieriger, mühsamer, kostspieliger Prozess, insbesondere aufgrund des Fehlens einer Doppelhaploid-Technik (DH; Hale et al. 2022). Nicht-ingezüchtete Individuen müssen über Generationen von Selbstbefruchtungen geführt werden, um den für den Aufbau einer Inzuchtlinie erforderlichen Ausgangs-Inzuchtgrad zu erreichen. Aufgrund ihrer teilweise fremd- und teilweise selbstbefruchtenden Fortpflanzungsweise (Brünjes & Link 2025; Duc et al. 2010; Link et al. 1994) erfolgt diese kontrollierte Selbstbefruchtung der Ackerbohne in Isolationskäfigen. Es wird so die Abwesenheit von (kreuz-)bestäubenden Insekten (Honigbienen und Hummeln) sichergestellt. Der Einsatz dieser Isolationskäfige ist ein zusätzlicher Aufwand, den die Züchter von z.B. Erbsen, Sojabohnen, Buschbohnen, Linsen nicht kennen.

Unsere Ackerbohnen zeigen einen sehr schwankenden ‚spontane' Samenansatz bei Bestäuber-Abwesenheit (Brünjes & Link 2025; Thema ‚Autofertilität-Autosterilität'). Daher kommt zusätzlich in diesem Setting der Aufwand einer manuellen, mechanische Stimulation (Bewegen der Blütenblätter; das sogenannte „Tripping"); es wird damit der Samenansatz (bei kontrollierter Selbstbefruchtung) verbessert – bedenke die an sich schon niedrige Vermehrungsrate der Ackerbohne.

Diese partielle Allogamie der Ackerbohne macht wie gezeigt die Züchtung von Linien teuer und langsam. Doch die partielle Allogamie lässt sich nutzen, um die in jeder Saison bereits vorhandene Inzuchtindividuen aufzufinden, anstatt sie mühsam zu entwickeln. In Ackerbohnenpopulationen sind tatsächlich hoch ingezüchtete Individuen vorhanden. Die Populationsgenetik jedenfalls besagt, dass jede genetisch vielfältige und offen bestäubte (d. h. mit Zugang zu Bestäubern) Ackerbohnenpopulation (wegen der partiellen Allogamie) erwartungsgemäß aus Kohorten von Individuen mit unterschiedlichem Inzuchtstatus besteht. Jede solche Inzuchtkohorte enthält Individuen, die auf die gleiche Anzahl aufeinanderfolgender, ununterbrochener Generationen von Selbstbefruchtung zurückgehen. Diese genetische Zusammensetzung partiell allogamer Populationen wurde von Cockerham und Weir (1984) sowie Kelly und Williamson (2000) beschrieben, speziell bei Ackerbohnen von Link et al. (1994) und Brünjes und Link (2021). Unter der Annahme eines Selbstbefruchtungsgrades von 0,5 (ein realistischer Wert) hat jedes Individuum in einer Population beispielsweise eine Wahrscheinlichkeit von $0{,}5^3$ (also 12,5%), aus drei aufeinanderfolgenden Generationen der Selbstbefruchtung zu stammen. Der Selbst- bzw. Fremdbefruchtungsgrad in Ackerbohnenpopulationen liegt Berichten zufolge in Europa oder Kanada niemals bei nahe 0 % oder bei nahe 100 % (Gasim et al. 2004, 2011; Holden & Bond 1960; Marzinzig et al. 2018; McVETTY & Nugent-Rigby 1984; Metz et al. 1992, 1995; Poulsen 1975; Gasim et al. 2002; Suso et al. 2005, 2006; Suso & Maalouf 2010).

Auch die Ackerbohne weist eine deutliche, wenn auch begrenzte Inzuchtdepression auf (offensichtlich begrenzt, weil viele Liniensorten leistungsmäßig bestehen (Abdelmula et al. 1999; Adhikari et al. 2021; Bishnoi & Panchta 2012; Bond 1982; Ebmeyer 1988; Fleck & Ruckenbauer, 1989; Gallais 1992; Ghaouti 2007; Wright 1973, 1977; Zeid et al. 2004). Insgesamt besteht daher die konkrete Erwartung, einen gewissen Anteil solcher hoch-ingezüchteter Individuen zu finden, in unserer Population und ganz allgemein in Zuchtpopulationen.

Bislang wurden diese äußerst wertvollen Individuen nicht in der Zucht eingesetzt, weil sie phänotypisch nicht erkennbar sind. Sie lassen sich jedoch leicht identifizieren, indem man ihren Homozygotiegrad und daraus ihren Inzuchtkoeffizienten (F) mithilfe molekularer Marker bestimmt. Es stehen mehrere Genotypisierungstechniken zur Verfügung: insbesondere Genotyping-By-Sequencing (GBS), Specific Primer Enrichment Technology (SPET) und SNP-Arrays (O'Sullivan et al. 2019). Diese Techniken erfordern jede ein gewisses Budget. Die Genotypisierung mit dem „Vfaba_v2"-Axiom-SNP-Array kostete zwischen 50 und 100 € pro Individuum (im Jahr 2023). F-Abschätzungen individueller Pflanzen lassen sich kostengünstig mit einer begrenzten Anzahl von KASP-Marker (www.biosearchtech.com) ermitteln. KASP-Analysen kosteten weniger als 0,40 € pro Individuum und Marker (im Jahr 2024; siehe auch Antrag Seite 13 die entsprechende Erläuterung zum Arbeitspaket 3).

Die genomische Vorhersage (GP) gilt als die effektivste genomische Methode zur Vorhersage hochpolygener Merkmale (beispielsweise Kornertrag; Lübberstedt et al. 2023). Ihr Potenzial wurde bei allen wichtigen Kulturpflanzen gezeigt (Budhlakoti et al. 2022), aber auch bei anderen Körnerleguminosen wie Sojabohne (Jarquin et al. 2016; Stewart-Brown et al. 2019), Erbsen (Osuna-Caballero et al. 2024), Linsen (Gebremedhin et al. 2024; Haile et al. 2020) oder weiße Lupinen (Annicchiarico et al. 2019). Jüngste Fortschritte bei genomischen Werkzeugen für Ackerbohnen, wie Hochdurchsatz-Genotypisierungstechniken und die Veröffentlichung eines Referenzgenoms (Jayakodi et al. 2023), ermöglichen die Anwendung und Umsetzung von GP. Jüngste Studien haben bereits das Potenzial von GP bei Ackerbohnen aufgezeigt. Abou Khater et al. (2025) untersuchten GP für Hitze- und Herbizidtoleranz bei Ackerbohnen mit einem reduzierten Satz von 192 KASP-Markern. Lippolis et al. (2025) untersuchten GP für ertragsbezogene Merkmale und zeigten relativ hohe Vorhersagefähigkeiten (‚Predictive Ability' für den Ertrag von 0,47 und 0,58). Zuletzt bewertete Windhorst (2025) GP hinsichtlich Überwinterungseigenschaften und lieferte Erkenntnisse zu GP für Merkmale wie Pflanzenhöhe, Ertrag und Tausendkorngewicht (TGW), wobei er das Potenzial von GP bei der Vorhersage wichtiger agronomischer Merkmale bei Ackerbohnen bestätigte. Die letztgenannte Studie wies durch Kreuzvalidierung für die drei oben genannten Merkmale eine moderate Predictive Ability von etwa 0,4 auf. Im Gegensatz zu den beiden zuvor erwähnten Studien (Vorhersage neuer Genotypen in denselben Test-Umwelten oder Vorhersage derselben Genotypen in neuen Umwelten) untersuchte sie zudem ein Szenario der harten Validierung: Vorhersage neuer Genotypen in neuen Umgebungen – nahe an der züchterischen Realität.

Die Gewinnung von DNA und somit von genotypischen Daten aus Ackerbohnensamen ohne diese zu beschädigen (sog. Seed Chipping) ist für die Züchtung von großem Vorteil und würde die Implementierung der Genomischen Vorhersage erheblich begünstigen. Das Szenario ist, dass früh im Zuchtgang (z.B. in F2) oder in der rekurrenten Selektion pro Genotyp nur je ein Individuum existiert, welches im Betrachtungs-Moment als reifer Same vorliegt. Die Auswahl der besten einzelnen Samen (alias Genotypen) bereits vor der Aussaat ermöglicht es, Zeit, Kosten und Ressourcen zu sparen, da die nicht ausgewählten Individuen (Samen, Genotypen) nicht ausgesät werden müssen, geschweige denn für die phänotypische Selektion auf dem Feld vorvermehrt und dort angebaut und geprüft werden müssen. Das Argument gilt insbesondere für Ackerbohnen angesichts ihres geringen

Vermehrungsrate und des sehr aufwendigen Saatgutvermehrungsprozesses mit Einsatz von Isolationskäfigen und Tripping (manuelle und mechanische Stimulation der Blüten; erforderlich, wenn bei Bestäuberabwesenheit ein einigermaßen gleichmäßige und zufriedenstellende Samenansatz angestrebt wird; siehe oben, Kapitel 2).

Die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen wurde nach Projektbeginn bereits andernorts erfolgreich durchgeführt. Nach unserem Kenntnisstand gab es dann zwei Laborprotokolle, die von (1) Annika Kolars, einer Masterstudentin am Lehrstuhl für Pflanzenzüchtung der Justus-Liebig-Universität Gießen (Deutschland), und (2) Dr. Hamid Khazaei (Universität Helsinki, Finnland; persönliche Mitteilung) entwickelt wurden. Trotz der prinzipiellen Durchführbarkeit der Technik wird sie von Ackerbohnenzüchtern bis heute nicht praktisch genutzt. Beide Protokolle empfehlen die Probenahme von trockenen Samen unter Verwendung eines Skalpells, was ein stark manuell geprägtes und umsetzungskritsches Vorgehen ist. Im Vergleich dazu lassen sich Blätter beispielsweise von Sämlingen für die Genotypisierung einfacher entnehmen, der Vorgang ist eher skalierbar und daher Standard.

3. Material und Methoden

3.1. Material und Methoden des Arbeitspakets 1; 2; 3; 7: Kosteneffiziente Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen

In der klassischen Ackerbohnenzüchtung werden Linien-Sorten normalerweise aufgefasst als Sorten, die von einem F4-Individuen abstammen oder von einem höher homozygoten Individuum (Adhikari et al., 2021; persönliche Kommunikation mit erfahrenen Ackerbohnenzüchtern), d. h. die Ausgangspflanze einer Linie stammt aus mindestens drei aufeinanderfolgenden Generationen der Selbstbefruchtung nach der Mendelschen Generation F1. Wir untersuchten die theoretische quantitative Variation der F-Werte pro Inzuchtkohorte, um zu definieren, welche Individuen als „hoch ingezüchtet“ in diesem Sinne zu betrachten sind. Wir nutzten die Binomialverteilung und integrierten dabei die Arbeit von Hillel et al. (1990) zur Verteilung der elterlichen Genome bei Selbstbefruchtung und Rückkreuzung. Die Essenz der Gespräche mit den Praktikern führte zu einem sogenannten Inzuchtschwellenwert von $0{,}8\overline{3}$, was dem 25%-Quantil der F-Verteilung der Individuen aus der F4-Inzuchtkohorte entspricht. Individuen mit einem F-Wert größer oder gleich diesem Schwellenwert wurden hier als „hochgradig ingezüchtet“ definiert, während diejenigen unterhalb dieses Wertes als „unzureichend ingezüchtet“ definiert wurden.

Die Göttinger Winterbohnenpopulation (GWBP) ist ein bekannter akademischer Arbeits-Genpool (sie ist außerdem eine wesentliche Quelle für die aktuelle Züchtung von Winterackerbohnen für Deutschland; s.a. Blatt für Sortenwesen 51(5) Seite 90; Link & Arbaoui 2006). Diese Population wurde 1989 in Göttingen (Stelling, persönliche Mitteilung) durch die gleichmäßige Mischung von 11 Gründer-Inzuchtlinien ins Leben gerufen: sechs aus Deutschland (Webo, Wibo, Hiverna/1, 79/79, L977/88 und L979/S1), zwei aus Frankreich (Côte d’Or/1 und Arrissot) und

drei aus Großbritannien (Banner, Bourdon und Bulldog). Diese Mischung wurde unter offenbestäubten Bedingungen mit Zugang für bestäubende Bienen und Hummeln angebaut und unterlag keiner künstlichen und nur begrenzter natürlicher Selektion. Seitdem wurde diese Population in Göttingen jede Saison erhalten und fortgeführt; die Generation im Jahr 2025 war ihre 37. Generation. Somit ist die GWBP eine frei rekombinierende, partiell allogame Population, die sicherlich nahe am populationsgenetischen Gleichgewicht lebt (minimales LD, nahe im sog. Identitätsgleichgewicht; Gasim 2003; Gasim et al. 2004; Link & Arbaoui 2006). Insgesamt wurden 728 Einzelpflanzen aus der Population zufällig ausgewählt und mithilfe des „Vfaba_v2"-Axiom-SNP-Arrays genotypisiert, welches etwa 60.000 SNP-Marker enthält (O'Sullivan et al., 2019). Die DNA-Extraktion und die Genotypisierung wurden von der SGS – INSTITUT FRESENI-US GmbH, Abteilung TraitGenetics, in Gatersleben durchgeführt. Marker mit >10 % fehlenden Werten, einer Major-Allel-Häufigkeit p > 0,99 und einem Fisher'-schen linearen Diskriminanten (FLD; zur Messung der Clusterqualität eines SNP) von unter 7 wurden eliminiert, was zu N=11.467 validen SNP-Markern führte.

Wir haben in R die Verteilung des individuellen F-Werts von drei Gleichgewichtspopulationen mit unterschiedlichem Fremdbefruchtungsgraden simuliert (Abb. 1), die in dieser Studie wie folgt benannt wurden: (1) die überwiegend auskreuzende Population, (2) die teilweise auskreuzende Population und (3) die überwiegend selbstbefruchtende Population. Geordnet nach absteigendem Fremdbefruchtungsgrad beobachteten wir jeweils einen Populationsmittelwert F von 0,15, 0,34 und 0,60 sowie eine Populationsmittelwert der Selbstbefruchtungsrate $\bar{S}$ von 0,27, 0,53 und 0,78. Der Anteil hoch ingezüchteter Individuen gemäß unserer Schwelle für F (F ≥ 0,8$\bar{3}$) betrug jeweils 1,3 %, 10,6 % und 38,2 %.

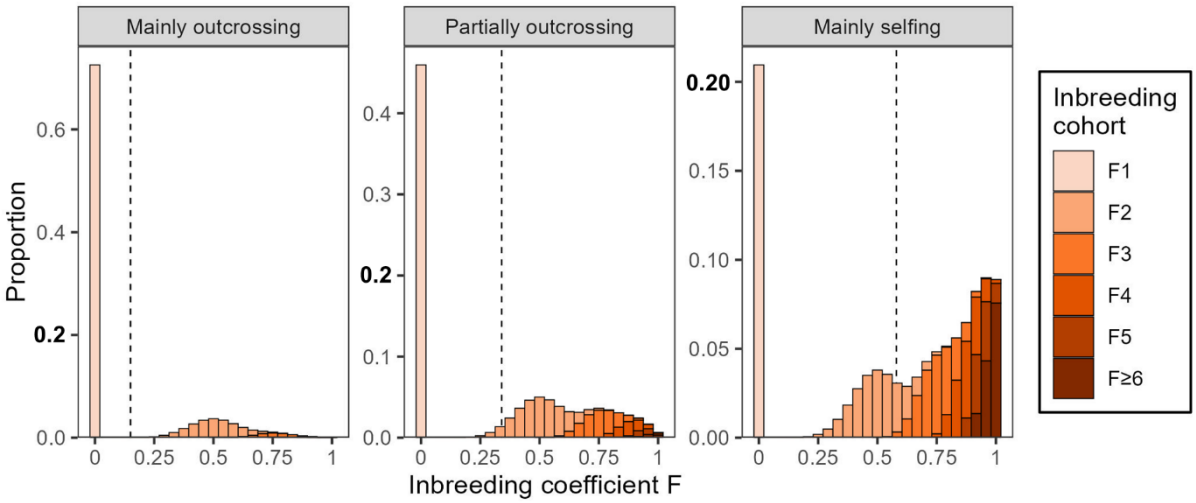


Abb. 1 Gestapeltes Histogramm, das den Anteil der individuellen F-Werte für drei simulierte Populationen entsprechend den Inzuchtkohorten (Farben) darstellt: von links nach rechts die überwiegend auskreuzende Population ($\bar{S}$ = 0,27), die teilweise auskreuzende Population ($\bar{S}$ = 0,53) und die überwiegend selbstbefruchtende Population ($\bar{S}$ = 0,78). Die gestrichelten vertikalen Linien stellen den mittleren F-Wert jeder Population dar.

Um die Genauigkeit der F-Schätzung und die Leistungsfähigkeit bei der Klassifizierung hoch ingezüchteter Individuen aus solchen Ackerbohnenpopulationen zu

bewerten, haben wir F-Schätzungen aus Marker-Teilmengen erstellt. Wir haben Marker-Teilmengen mit einer Markeranzahl von fünf bis 100 in Schritten von 5 generiert; dabei haben wir diese Marker-Teilmengen entweder mit zufälligen Allelfrequenzen oder mit idealen Allelfrequenzen erstellt. „Ideal“ bedeutet: für alle Marker p=0,5, keine fehlenden Werte, Gametenphasen-Gleichgewicht (LD=0). Eine solche ausgeglichene Allelfrequenz macht Marker am aussagekräftigsten für die Quantifizierung des Anteils der Homozygotie, der durch Abstammungsgleichheit (sog. IBD) bedingt ist, wenn dieser inzuchtbedingte Anteil zum Anteil der inzuchtirrelevanten Homozygotie addiert würde (die durch sog. zustandsgleiche Allele – AIS - verursacht ist). Ein LD von Null stellt sicher, dass die Marker jeweils unabhängige Informationen über den Inzuchtstatus des Individuums liefern.

Für die empirischen Daten aus der GWBP wurden zufällige Marker-Teilmengen generiert, indem zunächst 100 zufällige Marker aus unseren 11.467 Markern ausgewählt wurden. Diese Marker wurden dann verwendet, um durch sukzessive Reduzierung um jeweils fünf Marker immer kleinere Mengen zu erzeugen, was zu Teilmengen führte, die von 100 Markern bis hinunter zu fünf Markern reichten. Jede kleinere Menge wurde aus der jeweils nächstgrößeren Menge entnommen. Wir wiederholten den Gesamtvorgang, um für jede Markeranzahl (Teilmengengröße) 50 verschiedene Wiederholungen zu erhalten. Aus unseren 11.467 Markern wählten wir Marker mit einer Major-Allel-Häufigkeit $p \in [0{,}50, 0{,}55]$, mit weniger als 5 % fehlenden Werten und einem LD-Wert r^2 von weniger als 0,2 aus, was insgesamt 372 Marker ergab. Markerpaare mit einer LD von $r^2 > 0{,}1$ machten 1,4 % dieses nahezu idealen Markersatzes aus, gegenüber 13,6 % im Gesamt-Satz von 11.467 Markern. Wir wiederholten das oben beschriebene Verfahren, um Teilmengen mit „idealen“ Markern zu generieren: Aus den 372 Markern wurden 100 zufällig ausgewählt, und die anderen Marker-Teilmengen wurden daraus abgeleitet. Das Verfahren wurde 50 Mal wiederholt, um für jede Markeranzahl (Teilmengengröße) 50 verschiedene „ideale“ Wiederholungen zu erhalten.

Für die simulierten Daten erzeugten wir 50 Teilpopulationen mit jeweils 728 Individuen (das entspricht der Anzahl realer Individuen). Diese Individuen und ihre F-Werte stammen aus den drei theoretischen Verteilungen der individuellen F-Werte der Abbildung 1, um unseren Ansatz mit empirischen Daten hier simulierend nachzubilden. Für jede Teilpopulation erzeugten wir zunächst eine Marker-Teilmenge aus 100 Markern und simulierten in R genotypische Daten für alle Individuen unter Verwendung ihrer zugewiesenen F-Werte. Bei „zufälligen“ Marker-Teilmengen wurden die Allelfrequenzen der Marker zufällig aus einer Gleichverteilung (0,50 bis 0,99) gezogen. Bei „idealen“ Marker-Teilmengen hatten alle Marker $p = 0{,}5$. Wir leiteten durch sequenzielle Reduktionen um jeweils fünf Marker zunehmend kleinere Teilmengen ab.

Die Genauigkeit verschiedener F-Schätzer in Abhängigkeit von der Anzahl der verwendeten Marker (fünf bis 100) wurde also anhand simulierter Daten in R bewertet. Die Genauigkeit wurde mittels RMSE (Root Mean Square Error) beurteilt, der auf der Grundlage der Abweichung der 728 geschätzten F-Werte von den wahren 728 F-Wert berechnet wurde. Wir untersuchten den RMSE nur für die teilweise auskreuzenden Population (vergl. Abb. 1), da deren Selbstbefruchtungsrate ($\overline{S} = 0{,}53$) am nächsten zu dem in unserem GWBP erwarteten Wert sowie zu anderen in der Literatur berichteten Werten passte (Brünjes & Link 2021).

Caballero et al. (2022) haben mehrere Methoden zur Schätzung des Inzuchtkoeffizien-ten F innerhalb einer Population für Individuen mit biallelischen Markerdaten untersucht, von denen mehrere in unserer Studie verwendet wurden (Tabelle 2). Wenn die Allelfrequenz aller Marker p = 0,5 beträgt, lassen sich die algebraischen Formeln aller dargestellten Schätzer (Tabelle 2) auf dieselbe Gleichung zurückführen:

$$F = 1 - \frac{2}{m}\sum_{i=1}^{m} x_i(2 - x_i) = 1 - \frac{2H}{m}, \quad (1)$$

wobei m die Gesamtzahl der polymorphen Marker und H die Anzahl der beobachteten heterozygoten Loci ist.

F-Werte müssen im Bereich von 0 bis 1 liegen, doch können Schätzungen von F Werte außerhalb dieses Bereichs annehmen. Wir haben die geschätzten F-Werte für alle Schätzer auf den Bereich von 0 bis 1 beschränkt. Alle Schätzungen unter 0 wurden auf 0 gesetzt, während Werte über 1 auf 1 gesetzt wurden.

Wir wollten Individuen aus Ackerbohnenpopulationen entweder als unzureichend ingezüchtet oder als (ausreichend) hoch-ingezüchtet klassifizieren. Die Individuen wurden anhand (1) des Inzuchtschwellenwerts F = 0,8$\overline{3}$ als tatsächlich unzureichend ingezüchtet oder hoch-ingezüchtet definiert. In den empirischen Daten wurden die tatsächlich hoch-ingezüchteten Individuen (Gesamtzahl Positiver) und die tatsächlich unzureichend ingezüchteten Individuen (Gesamtzahl Negativer) auf der Grundlage von F-Werten definiert, die aus dem 11467-Marker-Satz unter Verwendung des genauesten F-Schätzers ermittelt wurden. Wenn aus begrenzten also kleinen Anzahlen von Markern geschätzten wurde, dann wurde dieser Inzuchtschwellenwert (F=0,8$\overline{3}$) nicht direkt angewendet. Stattdessen wurden Individuen unter Verwendung (2) der sogenannten Inbreeding Classification Threshold (Inzuchtklassifizierungs-Schwellenwert; ICT) als hoch-ingezüchtet eingestuft. Bei begrenzter Markeranzahl wurden Individuen mit einem geschätzten F-Wert, der größer oder gleich diesem ICT war (F ≥ ICT), als „hoch-ingezüchtet“ klassifiziert, während diejenigen mit einem niedrigeren Wert (F < ICT) als „unzureichend ingezüchtet“ eingestuft wurden. Es wurde mehr Wert gelegt auf die Sicherheit bei der Einstufung als hoch-ingezüchtet als auf die vollständige Ausschöpfung aller tatsächlich vorhandener Hoch-ingezüchteten.

Um die Leistungsfähigkeit unserer Klassifizierung zu bewerten, untersuchten wir (1) die Anzahl der echten Positiven (TP): die Anzahl der korrekt klassifizierten, tatsächlich hoch ingezüchteten Individuen, und (2) die Anzahl der falschen Positiven (FP): die Anzahl der tatsächlich unzureichend ingezüchteten Individuen, die fälschlicherweise als hoch ingezüchtet klassifiziert wurden. Eine effektive Klassifizierung führte zu einer hohen Anzahl an TP und einer geringen Anzahl an FP. Eine Klassifizierung war optimal, wenn alle tatsächlich hoch-ingezüchteten Individuen (Gesamtzahl der Positiven) als TP klassifiziert wurden und dennoch keine FP auftraten.

Tabelle 2: Liste der 7 getesteten Schätzer zur Schätzung von F anhand einer begrenzten Anzahl von Markern; siehe die jeweilige Formel. In den Formeln steht m für die Gesamtzahl der Marker, x_i für die Anzahl der Minor-Allele am Locus i, p_i für die Major-Allelfrequenz am Locus i und g_i für den Genotyp am Locus i.

Schätzer (Estimator)	Formel	Literaturquellen
F_{LH1}	$F_{LH1} = 1 - \frac{\sum_{i=1}^{m} x_i(2-x_i)}{\sum_{i=1}^{m} 2p_i(1-p_i)}$	Li & Horvitz 1953
F_{LH2}	$F_{LH2} = 1 - \frac{1}{m}\sum_{i=1}^{m} \frac{x_i(2-x_i)}{2p_i(1-p_i)}$	Caballero et al. 2022
F_{VR1}	$F_{VR1} = \frac{\sum_{i=1}^{m}(x_i-2p_i)^2}{\sum_{i=1}^{m} 2p_i(1-p_i)} - 1$	VanRaden 2008
F_{VR2}	$F_{VR2} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{(x_i-2p_i)^2}{2p_i(1-p_i)} - 1\right)$	VanRaden 2008
F_{YA1}	$F_{YA1} = \frac{\sum_{i=1}^{m} {x_i}^2 - (1+2p_i)x_i + 2{p_i}^2}{\sum_{i=1}^{m} 2p_i(1-p_i)}$	Zhang et al. 2022
F_{YA2}	$F_{YA2} = \frac{1}{m}\sum_{i=1}^{m} \frac{{x_i}^2 - (1+2p_i)x_i + 2{p_i}^2}{2p_i(1-p_i)}$	Yang et al. 2010
F_L	$F_L = \text{argmax } l(F)$ $l(F) = \sum_{i=1}^{m} \log\left(P(g_i\|F,p_i)\right)$ $P(g_i\|F,p_i) = \begin{cases} {p_i}^2 + p_i(1-p_i)F, & g_i = A_iA_i \\ 2p_i(1-p_i)(1-F), & g_i = A_ia_i \\ (1-p_i)^2 + p_i(1-p_i)F, & g_i = a_ia_i \end{cases}$	Hall et al. 2012; Wang 2011; Wright 1921

Wir untersuchten die Auswirkungen von (1) der Populationszusammensetzung, (2) der Genauigkeit der F-Schätzung und (3) des Klassifizierungsschwellenwerts ICT auf die Anzahl (bzw. Anteil) der TP und FP. Wir verwendeten simulierte Daten, um den Einfluss dieser drei Faktoren auf die Klassifizierungsleistung zu untersuchen. Für die (1) Populationszusammensetzung verglichen wir die Ergebnisse aus den drei simulierten Populationen (siehe Abbildung 1; eine überwiegend fremdbefruchtende, eine partiell allogame und eine überwiegend autogame Population). Der Einfluss der (2) Genauigkeit der F-Schätzung wurde durch Variation der Anzahl der verwendeten Marker (fünf bis 100) und ihrer Allelfrequenz (zufällig oder ideal) untersucht. Wir geben die Anzahl der TP und FP basierend auf (3) vier ICT-Werten an: 0,8$\overline{3}$; 0,90; 0,95 und 1,00. Wir haben einen ICT von unter 0,8$\overline{3}$ vermieden, da in diesem Fall die Klassifizierung selbst bei maximaler

Genauigkeit der F-Schätzung Falsch-Positive erzeugte, was unserer grundsätzlichen Maxime zuwiderlief, Falsch-Positive zu vermeiden (siehe oben).

Wir haben zusätzlich dann noch den Einsatz eines vorläufigen Klassifizierungsschritts getestet, der darauf abzielt, die Genotypisierungs-Kosten weiter zu verringern. Dieser Schritt sollte in einem ersten (kostengünstigen) Schritt die meisten der (offensichtlich) nicht ausreichend ingezüchteten Individuen schon mit einer recht geringen Anzahl an Markern aussortieren. Die hier gewonnenen DNA-Proben und Markerdaten (aus diesen wenigen Markern) der hier nicht aussortierten Individuen werden für den nächsten Schritt wiederverwendet. Da die Anzahl der verbleibenden Individuen geringer ist, ist die Analyse mit einer größeren Anzahl an Markern (zweiter Schritt) kostengünstiger als ohne diesen Vor-Filter. Die Individuen werden schließlich in diesem zweiten Schritt (Hauptklassifizierungsschritt) eingestuft: als hoch ingezüchtet oder eben nicht.

Um die Kosteneffizienz bei Hinzunehme des vorläufigen Klassifizierungsschritts zu bewerten, haben wir die Kosten des Ansatzes mit und ohne diesen Schritt verglichen, entsprechend der (Kosten der) Anzahl der in beiden Schritten verwendeten Marker. Die Kosten wurden nicht in Euro berechnet, sondern als Multiplikator der Kosten der KASP-Analyse pro Individuum und Marker, dem sogenannten KASP-Kostenfaktor, wobei die Kosten für die Einrichtung des KASP-Assays und die DNA-Extraktion (die unabhängig davon, ob der vorläufige Klassifizierungsschritt verwendet wird oder nicht, konstant sind) unberücksichtigt blieben.

3.2. Material und Methoden der Arbeitspakete 4; 5; 6; 8: Leistungsfähigkeit der Genomischen Vorhersage innerhalb und über Generationen hinweg in einer Winterackerbohnen-Population

Wir verwendeten zwei Genotypensätze: den sogenannten (1) Assoziations-Set (A-Set) und (2) Abo-Direkt-Set (AD-Set). Beide stammten aus der Göttinger Winterbohnenpopulation (GWBP; siehe oben). Diese Population wurde seit 1999 jede Saison unter offenbestäubten Bedingungen angebaut, mit einer mittleren Fremdbefruchtungsrate von etwa 50-60 % (Gasim et al. 2004), und sie unterlag keiner künstlichen und nur begrenzter natürlicher Selektion. Das A-Set (N = 185 Linien) wurde aus zufällig ausgewählten Individuen der GWBP in ihrer 9. Generation gebildet. Diese Ausgangs-Individuen wurden seither mittels Single-Seed-Descent (SSD) weiterentwickelt (erhalten, ingezücht; F ≈ 1.0; Ali et al. 2016). Der AD-Satz (N = 489) wurde ebenfalls aus zufällig gewählten Individuen der GWBP gebildet, diese Entnahme geschah jedoch aus in der 32. und 34. Generation der GWAB (das sind 23 und 25 Generationen nach der Entnahme der Ausgangsindividuen des A-Satzes). Die Abo-Direkt-Genotypen wiesen wie zu erwarten unterschiedliche Inzuchtgrade auf und waren nach Lage der Dinge (s.o.) in diesem Aspekt repräsentativ für die genetische Zusammensetzung einer genetisch diversen, langfristig offen bestäubten Ackerbohnenpopulation (Brünjes & Link 2021; Laugel & Link 2025; Link et al. 1994).

Die für das A-Set verwendeten phänotypischen Daten wurden von Windhorst (2025) ausführlich beschrieben. Diese 185 Linien wurden von 2005 bis 2022 in 18 historischen Feldversuchen (HFT) umfassend getestet. Die Versuche fanden

hauptsächlich auf der Versuchsstation Reinshof der Universität Göttingen statt (Niedersachsen; 51,48° N, 9,93° O); der zweite Standort war Hohenlieth bei der Norddeutschen Pflanzenzucht (NPZ) Hans-Georg Lembke KG (Schleswig-Holstein; 54,41° N, 9,89° O). Diese Linien wurden entweder als Inzuchtlinien *per se* (Feldversuchs-Saatgut aus kontrollierter Selbstbefruchtung) oder als Nachkommen von A-Set-Linien (sogenanntes Ex-Trial-Saatgut aus dem jeweils vorangegangenen Feldversuch mit natürlichem Zugang von Bestäubern) getestet. Die in den HFT bewerteten Merkmale sind in Tabelle 3 aufgeführt. Die A-Set-Genotypen wurden in 11 Versuchen als Linien *per se* und in sieben Versuchen als Ex-Trial Nachkommen getestet.

Das AD-Set wurde 2024 und 2025 in vier Feldversuchen getestet. Die Genotypen, die zur Saatgutvermehrung eine Selbstungsgeneration durchliefen, wurden in den kleinen, sogenannten „WEE“-Versuchen mit Parzellen von nur 30 Pflanzen (~1 m² im Jahr 2024 und ~1,5 m² im Jahr 2025) getestet. Die Genotypen, die zwei aufeinanderfolgende Generationen der Selbstbefruchtung durchlaufen hatten und somit über mehr Saatgut verfügten, wurden im sogenannten „BIG“-Versuch mit Parzellen von 200 Pflanzen getestet. Wir führten drei „WEE“-Versuche und einen „BIG“-Versuch durch (Tabelle 3). Im Jahr 2024 wurden 168 Genotypen im sogenannten „WEE2024“-Versuch in Göttingen und Hohenlieth getestet. In Göttingen wurden von den 168 Genotypen 47 zweimal und 14 dreimal wiederholt, während sie in Hohenlieth nicht (r=1) wiederholt wurden. Im selben Jahr wurden 102 weitere Genotypen in einem ergänzenden Versuch in Göttingen getestet, der als „WEE2024+“ bezeichnet wurde und über nicht replizierte Parzellen verfügte. Beide Versuche wurden durch den Einsatz der Sorten „Augusta“, „Curlew“ und „Pantani“ als Kontrollen ergänzt, um die Feldheterogenität zu bewerten. Im Jahr 2025 wurden 76 weitere Genotypen im sogenannten „WEE2025“-Versuch in Göttingen in einem p-rep-Design getestet (10 der 76 Genotypen wurden zweimal wiederholt, zusammen mit den Sorten „Nairobi“, „Patagonia“, „Curlew“, „GL Alice“ und „GL Arabella“, die als Kontrollen dienten). Im selben Jahr wurden 143 weitere Genotypen im sogenannten „BIG2025“-Versuch an drei Standorten getestet: Göttingen, Hohenlieth und Wartmannsroth (Bayern; 50,17° N, 9,79° O). Die Parzellen waren in Göttingen größer (~7,8 m²) als in Hohenlieth (~5.6m^2). In Wartmannsroth (bei Werner von Kaute) wurden die Genotypen unter ökologischen Bedingungen geprüft. Die Versuchsanlage an diesen drei Orten war eine α-Gitter mit zwei Wiederholungen.

Die Gitterversuche in Göttingen und Hohenlieth wurden durch einen zusätzlichen, angrenzenden Feldversuch im vollständig randomisierten Design (CRD) ergänzt, um für alle Genotypen die Saatgutverfügbarkeit bestens zu nutzen. Die Anzahl der Wiederholungen pro Genotyp lag so zwischen einer und elf, im Durchschnitt bei vier. In Göttingen wurden alle 143 Genotypen getestet: 121 im Gitterversuch; der CRD enthielt 57 Genotypen. In Hohenlieth wurden 132 der 143 Genotypen getestet (Gitterversuche 64; CRD 74). In Wartmannsroth wurden 64 Genotypen im dortigen Gitterversuch getestet (51 davon standen auch in Hohenlieth). Alle Versuche wurden an jedem Standort Mitte Oktober ausgesät und Mitte August geerntet, mit der Ausnahme, dass der WEE2024 in Hohenlieth Mitte Januar ausgesät und der BIG2025 in Wartmannsroth Ende Juli geerntet wurde. Die in den vier Versuchen bewerteten Merkmale sind in Tabelle 3 aufgeführt. Im Jahr 2025 hatten wir in Göttingen die Möglichkeit, die Toleranz gegenüber Schokoladen-

fleckigkeit und Rost zu bewerten. Außerdem wurde der Proteingehalt (Prot) der in Hohenlieth getesteten Genotypen durch das NPZ bestimmt.

Für jeden Versuch der HFT wurden die Genotyp-Mittelwerte unter Berücksichtigung des Versuchsdesigns (RCBD oder Gitterdesign) mit PLABSTAT ermittelt. Ausreißer wurden anhand des in PLABSTAT implementierten Ausreißertests entfernt. Anhand der Genotyp-Mittelwerte pro Versuch wurde eine Varianzanalyse (ANOVA) über die Versuche hinweg durchgeführt, und die BLUEs für jeden Genotyp wurden mittels eines linearen gemischten Modells generiert. Für die Abo-Direkt-Versuche wurde eine ANOVA durchgeführt und BLUEs für jeden getesteten Genotyp generiert, indem ein lineares gemischtes Modell angepasst wurde. Der BIG2025-Versuch war in der Analyse die Kombination aus Gitterdesign und CRD.

Um die phänotypischen Daten aus und über alle Versuche zu kombinieren, haben wir die phänotypischen Roh-Daten der Gitteranlagen unter Berücksichtigung der Teilblockeffekte für jedes Merkmal angepasst (Schätzung mittels eines linearen gemischten Modells). Nach dieser Anpassung betrachteten wir die beiden Wiederholungen der Gitter und das CRD als drei separate Blöcke (in Göttingen und in Hohenlieth). Es wurde entsprechend eine ANOVA durchgeführt und die BLUEs wurden für jeden Standort separat unter Verwendung von Modell (2) sowie standortübergreifend durch Anpassung eines weiteren linearen gemischten Modells ermittelt. Alle Analysen wurden mit dem Paket „lme4“ (v.1.1-37) in R durchgeführt. Für alle Versuche wurden die merkmalsbezogenen Heritabilitäten im weiteren Sinne (H^2) oder die Wiederholbarkeiten (wenn nur Daten für einen Standort verfügbar waren) nach der Formel von Piepho & Möhring (2007) berechnet. Die Varianzkomponenten der linearen gemischten Modelle wurden geschätzt, indem alle Effekte als zufällig modelliert wurden. Die phänotypischen Daten der Kontrollen (Checks) wurden verwendet, um die Schätzung der Fehlervarianz zu verbessern, insbesondere für die WEE-Versuche (die keine oder nur wenige Wiederholungen pro Genotyp aufwiesen). Um zu überprüfen, ob sich die aus den linearen gemischten Modellen erhaltene genetische Varianz (σ^2_G) signifikant von Null unterschied, wurden Likelihood-Ratio-Tests zwischen dem vollständigen Modell und einem Nullmodell (ohne den genetischen Effekt) durchgeführt.

Die Genotypisierung und die Verarbeitung der Genotypdaten des A-Sets wurden von Windhorst et al. (2024) ausführlich beschrieben und ergaben 17.227 SNPs. Die 471 Genotypen des AD-Sets wurden mit dem „Vfaba_v2“-Axiom-SNP-Array analysiert (O’Sullivan et al. 2019; dasselbe Array war auch für das A-Set verwendet worden). Die DNA-Extraktion und die eigentliche Genotypisierung wurden von SGS – INSTITUT FRESENIUS GmbH, Abteilung TraitGenetics in Gatersleben, durchgeführt. Marker mit >10 % fehlenden Werten, mit einer Major-Allel-Häufigkeit $p > 0,95$ und mit einem FLD < 5 wurden eliminiert. Anschließend haben wir für die verbliebenen Marker mittels der Referenzgenomsequenz von V. faba Hedin/2 v.1 (Jayakodi et al. 2023) fehlende Daten aufgefüllt, dieses Imputieren wurde mit dem LD-kNNi-Algorithmus in TASSEL 5.0 unter Verwendung der Standardeinstellungen durchgeführt (Bradbury et al. 2007). Dies ergab insgesamt 16.283 SNP-Marker. Der kombinierte Datensatz aus dem A-Set (17.227 SNPs) und dem AD-Set (16.283 SNPs) umfasste 13.913 gemeinsame SNPs. In einem abschließenden Filterungsschritt wählten wir Marker mit einer Major-Allel-Häufigkeit von $p \leq 0,95$ aus, was final zu 13.840 SNP-Markern führte.

Tabelle 3. Verfügbarkeit von Daten aus Umwelten (Jahre mal Standorte) pro Merkmal in jedem Versuch. Bei den historischen Feldversuchen (HFT) ist „L“ die Anzahl der Umwelten, in denen die A-Set-Genotypen als Linien getestet wurden, „P“ die Anzahl der Umgebungen, in denen sie als Ex-Trial Nachkommen von Linien getestet wurden, und „L+P“ die Gesamtzahl beider Kombinationen. Pla1C ist Pflanzenanzahl vor Winter, Pla2C ist Pflanzenanzahl nach Winter. Dev1 ist Pflanzenentwicklung vor Winter (Bonitur), Dev2 ist Pflanzenentwicklung nach Winter (Bonitur); Cano ist Bestandes-Entwicklung nach Winter (Bonitur), Flbeg ist Blühbeginn, Flend ist Blühende (beides als Tage nach dem 1. Januar). Ma1 ist der Wert bei erster und Ma2 bei zweiter Reifebonitur. Asco ist die Bonitur der Anthraknose-Symptomatik, Pero der Peronospora-Symptomatik, Choc der *Botrytis*-Symptomatik, Rust der *Uromyces viciae-fabae*-Symptomatik. Ph1 gibt Resultate der ersten Pflanzenwuchshöhenmessung (in cm)und Ph2 der zweiten, Lg1 die erste Lagerbonitur und Lg2 die zweite Bonitur. Yld ist der Ertrag (yield in dt/ha) und TGW ist das Tausendkorngewicht (in g)

Merkmals-gruppe	**Merk-mal**	**H**ist.**F**.**T**rials			**WEE2024**	**WEE2024+**	**WEE2025**	**BIG2025**
		L+P	**L**	**P**				
Überwinterung	Pla1C	8	5	3	2	1	1	2
	Pla2C	9	5	4	2	1	1	2
	Dev1	3	1	2	0	0	0	1
	Dev2	9	4	5	2	1	0	0
	Cano	13	8	5	0	0	0	0
Phänologische Merkmale	Flbeg	18	11	7	2	1	1	3
	Flend	4	3	1	2	1	1	3
	Ma1	9	5	4	0	0	1	3
	Ma2	12	9	3	0	0	0	1
Krankheitsresistenz	Asco	7	7	0	0	0	0	0
	Pero	8	2	6	0	0	1	1
	Choc	0	0	0	0	0	0	1
	Rust	0	0	0	0	0	1	1
Agronomische Merkmale	Ph1	13	10	3	2	1	1	3
	Ph2	13	7	6	0	0	1	3
	Lg1	13	7	6	0	0	1	2
	Lg2	15	10	5	1	1	0	1
	Yld	17	10	7	1	1	1	2
	TGW	9	4	5	1	1	1	3
Qualität	Prot	0	0	0	0	0	0	1

Die genetische Ähnlichkeit des A-Sets mit dem AD-Set wurde anhand ihrer Genotypdaten (13.840 SNPs) untersucht. Dazu wurde eine Hauptkomponentenanalyse

(PCA) mit dem Paket „stats“ (v. 4.5.1) in R durchgeführt. Die genomische Verwandtschaftsmatrix (GRM) wurde nach der Methode I von VanRaden (2008) berechnet. Um zu verhindern, dass die Schätzungen der Allelfrequenzen unverhältnismäßig stark durch das AD-Set beeinflusst werden (welches mehr Genotypen enthielt als das A-Set), folgten wir dem Ansatz von Wientjes et al. (2015). Anstatt also die Allelfrequenzen über alle 674 Genotypen (A-Set plus AD-Set) zu berechnen, wurden die Allelfrequenzen zunächst separat innerhalb jedes Sets geschätzt. Diese beiden Schätzungen wurden gemittelt. Mittelwert und Varianz der paarweisen Beziehungen außerhalb der Diagonalen zwischen den Genotypen innerhalb und zwischen beiden Sets wurden angegeben. Die Genotypen wurden auf der Grundlage paarweiser genomischer Beziehungen unter Verwendung UPGMA Methode gruppiert (Dendrogramm), die mit dem „stats“-Paket (v.4.3.2) in R berechnet wurde. Das auf GRM und UPGMA basierende Dendrogramm wurde mit ComplexHeatmap (v.2.15.4; Gu 2022) dargestellt. Die Allelfrequenzen beider Sätze wurden verglichen und hierfür die Spearman-Korrelation zwischen A-Set und AD-Set angegeben. Das paarweise Gametenphasen-Ungleichgewicht (LD; r) wurde in R mit einem gleitenden Rahmen 50 SNP berechnet (für jeden Marker wurde das LD mit den 50 physikalisch nächstgelegenen Markern berechnet). Die 692.000 (13.840*50) paarweisen LD (also r)-Werte beider Datensätze wurden verglichen, und ihre Spearman-Korrelation zwischen A-set und AD-set wurde dargestellt.

Die genomische Vorhersage erfolgte als GBLUP-Modell mithilfe der Funktion „kin.blup“ des Pakets „rrBLUP“ (v. 4.6.3) in R. Die GRM wurde wie oben beschrieben berechnet, wobei wie beschrieben beiden Datensätzen gleichstark die benutzten Allelfrequenzen bestimmten. Die Vorhersagefähigkeiten (Predictive Ability) wurden als Pearson-Korrelation zwischen den beobachteten BLUEs und den genomisch geschätzten Zuchtwerten (GEBVs) ausgedrückt. Die sog. Accuracy wurden berechnet, indem die Predictive Ability durch die Quadratwurzel der Heritabilität der BLUEs aus dem Testdatensatz dividiert wurden.

Die genomische Vorhersage innerhalb einer Generation wurde fünffach mit 100 Iterationen kreuzvalidiert; sowohl für den A-Satz als auch für den AD-Satz. Für das A-Set wurden die 185 Genotypen verwendet (148 Genotypen für das Training und 37 für den Test). Für das AD-Set wurden die 143 im BIG2025-Versuch feldgetesteten Genotypen verwendet (114–115 Genotypen als Trainingsset und 28–29 Genotypen als Testset), da deren phänotypische Daten unter den AD-Set-Resultaten am zuverlässigsten waren (große Parzellen, drei Standorte und mehrere Wiederholungen pro Genotyp).

Wir verwendeten zwei Modelle, die mit (1) dem A-Set (185 Genotypen) beziehungsweise (2) dem BIG2025 (143 Genotypen) trainiert wurden, um die GEBVs des jeweils anderen Datensatzes vorherzusagen. Das erste Modell diente zur Bewertung der genomischen Vorhersagefähigkeiten (Prediction Ability) und -genauigkeiten (Accuracy) über Generationen hinweg (Training mit A-Satz, Vorhersage von AD-Genotypen). Das zweite Modell demonstrierte ebenso die genomischen Vorhersagefähigkeiten und -genauigkeiten für die generationsübergreifende Vorhersage (BIG2025 zur Vorhersage des A-Sets); zusätzlich jedoch die generationsinterne Vorhersage (Training mit den BIG2025-Daten zur Vorhersage der Resultate der in den WEE-Versuchen getesteten Genotypen; die Genotypen in BIG2025 und in den WEE-Versuchen waren ja gleichermaßen spät, in der 32.

und 34. Erhaltungsgeneration der GWBP aus dieser entnommen worden, siehe Kapitel 3.2).

Um den Einfluss einer genomischen Differenzierung zwischen beiden Sätzen eingehender zu betrachten, wurde wiederum ein GP-Modell mit den Daten der A-Satz-Linien trainiert. Nun jedoch wurden nur phänotypischen Daten verwendet, die bei den Feldprüfung der Linien *per se* (Prüfung als Inzuchtlinien; L) erhoben wurden. Die Daten aus Prüfungen als P (also aus Ex-Trial Saatgut; siehe Tabelle 3) wurden hier nun nicht benutzt. Wir verglichen demnach die Vorhersagegenauigkeit (Accuracy) zweier Vorhersage-Szenarien: (1) Vorhersage neuer (also eben nicht für das Training eingesetzter) Genotypen aus dem A-Set (innerhalb derselben Generation) in neuen Umwelten und (2) Vorhersage neuer Genotypen aus dem AD-Set (andere Generation) in neuen Umwelten.

Für das erste Szenario führten wir eine fünffache Kreuzvalidierung (100 Iterationen) durch, bei der die für die 37 verbleibenden Genotypen aus dem A-Set (20 %) geschätzten GEBVs mit den phänotypischen Daten korreliert wurden, die aus der Prüfung dieser 37 Linien anhand ihrer Ex-Trial Nachkommen stammten (in – *nota bene* - anderen Umwelten; die nicht für das Modell-Training zur Verfügung standen). Dies war möglich, da diese Ex-Trial Nachkommen dieser Linien in den HFT eben nicht gemeinsam (in derselben Umwelt) mit den Linien *per se* getestet worden waren (siehe Tabelle 3). Im zweiten Szenario wurden die 148 Genotypen des A-Sets (L), die (in jeder Iteration) als Trainingsset dienten, verwendet, um GEBVs der 143 BIG2025-Genotypen zu schätzen, die dann mit deren phänotypischen Daten korreliert wurden. Jede sich daraus ergebende Lücke in der Accuracy zwischen den beiden Szenarien in diesem Vergleich sollte fast vollständig durch eine genomische Differenzierung zwischen beiden Material-Sets erklärt werden können (bzw. durch die genomische Differenzierung der Generationen der GWBP, aus der diese Sets genommen und entwickelt worden waren).

3.3. Material und Methoden des Arbeitspakets 9: Optimierung des Saatgut-Chipping-Verfahrens zur zerstörungsfreien DNA-Extraktion aus Ackerbohnensamen

Wir erhielten drei KASP-Asays von SGS – INSTITUT FRESENIUS GmbH (Gatersleben; Probeset-ID: AX-416739140, AX-416825417 und AX-181153608). Es handelt sich um SNP-Marker aus dem „Vfaba_v2“ Axiom SNP Array. Die DNA-Extraktion und die KASP-Analyse wurden an Saatgutproben durchgeführt, abweichend von der Standard-Beprobung aus jungen Blättern. Es wurden Proben der Sorte „Allison“ verwendet. Wir haben das von Dr. Hamid Khazaei (Natural Resources Institute Finland LUKE; persönliche Mitteilung) bereitgestellte DNA-Extraktionsprotokoll unter Verwendung einer NaOH-Behandlung getestet. Die Samenproben von den trockenen Samen wurden entnommen, indem ein kleiner Bereich der Samenschale abgekratzt wurde. Das somit freigelegte Keimblatt wurde anschließend mit einer sterilen Pinzette abgeschabt. Der Probenbereich befand sich auf der dem Hilum gegenüberliegenden Seite der Ackerbohnensamen, um sicherzustellen, dass die für die Keimung notwendige Keimwurzel nicht beschädigt wird. Für jede Probe wurden drei technische Replikate als Kontrolle analysiert, zusammen mit drei NTC (no-template-control) und mit drei Wasserkontrollen.

Es wurden im Weiteren dann drei Entnahmeverfahren getestet: an trockenen oder an feucht-gequollenen Samen; mit Schneiden, Stanzen oder Bohren, somit mit 6 Behandlungskombinationen. Bei der Probenahme an trockenen Samen wurde ein Schraubstock zur Fixierung der Samen verwendet (Abb. 2). Das Schneiden erfolgte mit einem Skalpell, das Stanzen mit einem Stanzer mit 1 mm Schnittdurchmesser (Rapid-Core Biopsy & Sampling Punch mit Kolbensystem, Micro to Nano) und das Bohren mit einem Bohrer mit 1 mm Durchmesser (Velleman Multifunktions-Präzisionsbohrer, VTHD04 135 W, 10000–32000 U/min; Von Post et al. 2003).

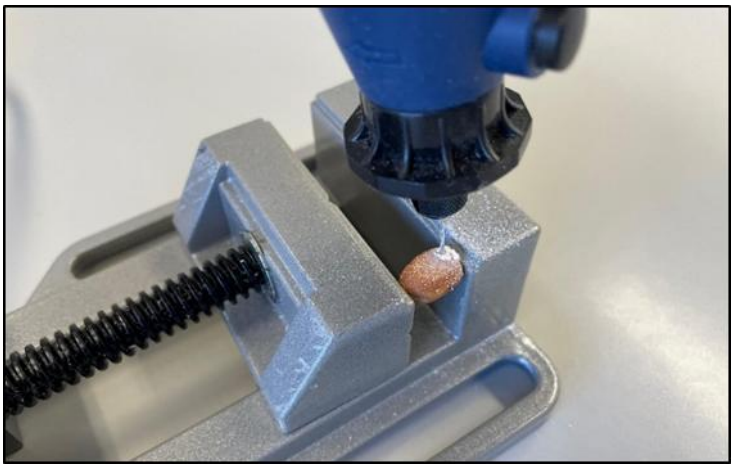


Abb. 2: Vorrichtung zur Gewinnung von Gewebeproben aus reifen Ackerbohnensamen mittels Anbohrens, wobei der Samen mit einem Schraubstock fixiert und ein Bohrer mit 1 mm Durchmesser verwendet wird.

Für die Probenahme an feuchten Samen wurde untersucht, wie lange die Samen in Wasser eingequollen werden müssen, damit sie weich genug für die Probenahme waren. Es wurden Samen der Genotypen „Côte d'Or", „Allison" und „GL Alice" in Standard-Petrischalen mit unterschiedlichen Wassermengen pro Schale (30, 50 und 65 ml) untersucht.

Beim Schneiden und Stanzen war die Reinigung des Probenahmegerätes (mit Ethanol) einfach. Die Reinigung der Bohrmaschine war jedoch eher mühsam, da dazu jedes Mal (Sicherheit) der Bohrer vom Strom getrennt werden muss. Von Post et al. (2001) reinigten ihre Bohrmaschine durch Eintauchen in Wasser. Diese Methode wurde nicht angewendet, da eine unsachgemäße Trocknung des Bohrers das Bohrergebnis beeinträchtigt. Die Verwendung von Ethanol, wie sie von Kamiya & Kiguchi (2003) praktiziert wurde, war in der Bedienungsanleitung der Bohrmaschine (Sicherheit) untersagt. Daher testeten wir, ob man die Reinigung des Bohrers zwischen den Probenahmen unterlassen könnte, also inwiefern dieses zu einer Kreuzkontamination von einer zur nächsten Probe führen würde. Genotypen, die für verschiedene Allele homozygot waren, wurden hierfür abwechselnd und ohne Reinigung des Bohrers beprobt. Insgesamt wurden 28 Proben verwendet, jeweils 14 Proben von „Côte d'Or" und „Hiverna". Für jede Probe wurden drei technische Replikate zur Kontrolle analysiert, zusammen mit drei NTC- und drei Wasserkontrollen. Als Kreuzkontamination der Proben würde deklarieren, wenn eine der Proben auf dem Allel-Diskriminierungsdiagramm der KASP-Analyse nicht ordnungsgemäß (nicht richtig) gruppiert wurde.

Insgesamt 140 Samen des Genotyps „Allison“ wurden mit der hier als optimal befundenen Chipping-Technik bearbeitet. Diese wurden in vier Kunststoffschalen mit 35 Vertiefungen in einer 50/50-Mischung aus Kompost und Fruhstorfer Erde (Spezialsubstrat, Typ Profi, Hawita) ausgesät, zusammen mit 140 unbehandelten „Allison“-Samen aus derselben Charge (als Kontrolle) in vier weiteren Schalen. Die Keimung wurde bis 19 Tage nach der Aussaat (DAS) überwacht. Am 19. Tag nach Aussaat wurden Daten zur Keimung erhoben (Keimung als binäres Merkmal; 0 = nicht gekeimt; 1 = gekeimt; Keimrate pro Schale als GR; Keimlingslänge als SL in cm; daraus abgeleitet der Keimlingsvitalitätsindex als SVI = GR * SL / 100).

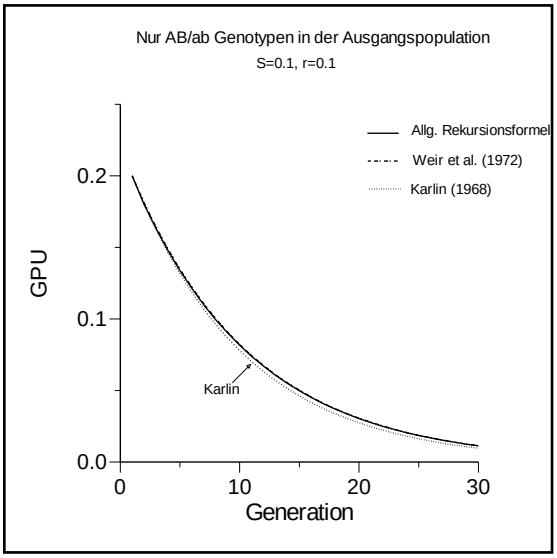
4. Ausführliche Darstellung der wichtigsten Ergebnisse

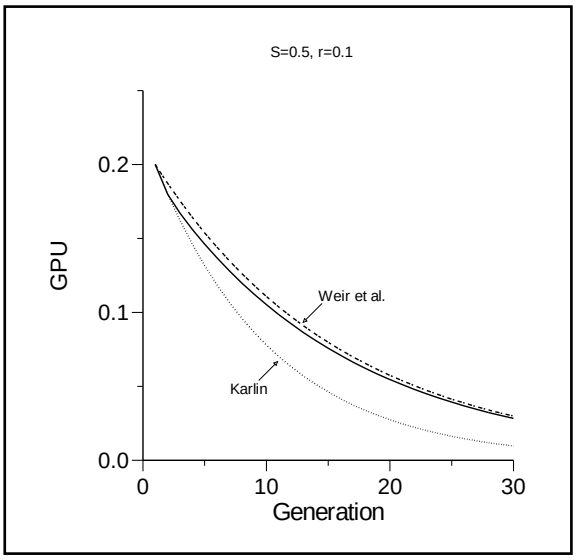
4.1.1. Ergebnisse der Teilaufgabe im Arbeitspakets 7: LD-Abbau in Populationen bei partieller Allogamie

Der Abbau des Gametenphasen-Ungleichgewichtes (GPU; aka LD) im Falle von partieller Allogamie wurde modelliert. Die Bewertungen und Schlussfolgerungen ergaben sich unter Rückgriff auf Arbeiten und unveröffentlichte Ausführungen von Gumpert, F., 2007, die zu der unserer Projektfragestellung in Bezug gesetzt wurden. Zum Vergleich mit diesen Resultaten liegen aus der Literatur zwei Publikationen vor (Karlin, 1968; Weir et al., 1972). Als Ausgangs-Szenario wurde mit maximalem GPU gestartet. Wir zeigen hier, wie dieses GPU sich bei einer Rekombinationsrate von c=10% abbaut wird (entspricht nach Haldane 11,16cM oder nach Kosambi 10,14cM). In den Grafiken beginnt die Darstellung nach der ersten Rekombination, als der Einstiegswert von D=0,25 in der ersten Meiose erstmals verringert wurde. Wegen enger Kopplung ergibt sich dort ein GPU von D=0,20. Wir ziegen den Abbau für partielle Allogamie: Selbstbefruchtungsraten von 10% oder 50% oder 90% angenommen.

Die hier vorgelegten Ergebnisse werden denen der o.g. Publikationen gegenbergestellt (Abb. 3). Karlin (1968) und Weir et al., (1972) haben unter Umgehung der Genotypenfrequenzen ihre Rekursionsformeln direkt für das GPU der Gametischen Reihe gegeben (Formeln sind unten angegeben).

Die Grafiken zeigen, dass bei einem kleinen Selbstbefruchtungsgrad von 10% alle 3 Kurven sehr gut übereinstimmen. Bei einem Selbstungsanteil von 50% gibt nur die Formel von Weir et al. (1972) annähernd verlässliche Werte. Bei einem höhen Selbstungsgrad von 90% weichen die Rekursionsformeln von unseren Ergebnissen erheblich ab.





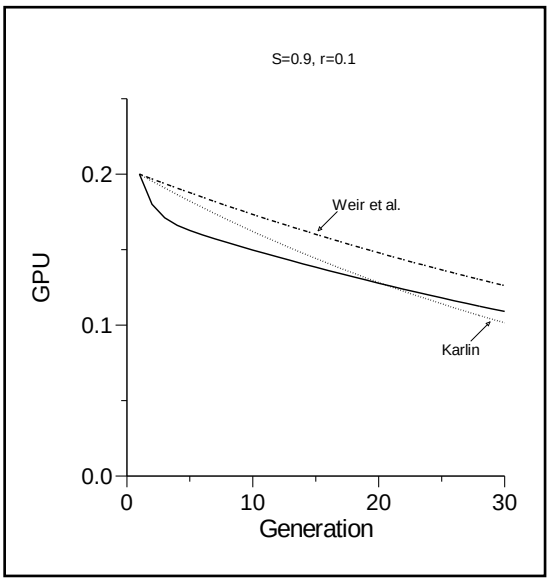


Abb. 3. Abbau des LD (GPU) nach eigenen Berechnungen und nach der Literatur. Der hier angenommene Selbstbefruchtungsgrad ist jeweils in den Legenden ausgewiesen, S=0,1 bzw. S=0,5 bzw. S=0,9. Die Resultate der Rekursionsformeln von Karlin (1968) und Weir et al., (1972) sind kenntlich gemacht. Deren Rekursionsformeln sind hier ebenfalls angegeben.

Autor	Rekursionsformel zur Darstellung des Abbaus des GPU (Gametenphasenungleichgewicht; LD)
Karlin (1968)	$GPU_i = \frac{1}{2}\left\{1-r+\frac{S}{2}+\sqrt{\left(1-r+2rS-\frac{S}{2}\right)^2-4rS(1-S)(1-r)}\right\}GPU_{i-1}$
Weir et al. (1972)	$GPU_i = \frac{1}{2}\left\{1-r+\frac{S}{2}+\sqrt{\left(1-r+\frac{S}{2}\right)^2-2S(1-2r)}\right\}GPU_{i-1}$

Der Unterschied liegt teilweise daran, dass wir in das GPU per Kreuzung einsteigen. In unserem Setting sind also auch nach den ersten Generationen (bei hohem Selbstungsanteil) die doppelt-Heterozygoten noch häufig; es sind ja nur die Meiosen der Doppelt-Heterozygoten, die zum LD-Abbau führen. Der bei S=90% starke Abbau des LD in den ersten Generationen rührt daher. Die Formeln von Weir et al. (1972) und Karlin (1968) starten mit einer Population, die schon im Identitätsgleichgewicht ist und daher einen kleineren Anteil von Doppelt-Heterozy-

goten aufweisen. Unsere Ergebnisse liegen insgesamt, vom Muster her gesehen, näher an denen von Weir et al. (1972).

Der Abbau eines im Zuchtmaterials vorliegenden LD (aka GPU) ist nicht der einzige Grund, aus dem heraus die generationsübergreifende GP zu merklich schwächeren Vorhersagen führt als innerhalb derselben Generation. Es entsteht während der Erhaltung einer Population (bei partieller Allogamie) in der Realität durch natürliche Selektion, durch Mutation, durch Drift, durch Migration in gewissem Umfang auch neues LD, durchaus auch mit der jeweils anderen Kopplungs-Phase und mit zusätzlichen Allelen. Das hier bearbeitete Topic (Abbau eines vorhandenen GPU) für den speziellen Fall der partiellen Allogamie ist somit nur eines von mehreren noch weiter zu bearbeitenden Themen.

4.1.2. Ergebnisse des Arbeitspakets 1; 2; 3: Kosteneffiziente Identifizierung hochingezüchteter Individuen aus Ackerbohnenpopulationen

Wenn wir Teilmengen mit den sog. „zufälligen“ Markern einsetzten (Marker, deren Allelfrequenz nicht konstant p=0,5 betrug, sondern variierende Allelfrequenzen aufweisen; siehe Kapitel 3.1), dann beobachteten wir Unterschiede in der Genauigkeit zwischen den F-Schätzern. Wie erwartet stieg die Genauigkeit der F-Schätzung mit der Anzahl der verwendeten Marker. Die Verbesserung war zwischen fünf und 25 Markern am größten. Der Likelihood-Schätzer FL war bei allen verwendeten Markeranzahlen (außer bei fünf Markern) der genaueste. Bei Verwendung von mindestens 20 Markern war der Schätzer FLH1 der zweitgenaueste Schätzer. Um jedoch die Genauigkeit des Schätzers FL bei Verwendung von 50 Markern zu erreichen, waren beim Schätzer FLH1 60 Marker erforderlich. Bei Verwendung von 50 Markern betrug die Pearson-Korrelation zwischen den geschätzten FL-Werten und den wahren F-Werten im Durchschnitt 0,94. Bei Marker-Teilmengen von sog. „idealen“ Markern (deren Allelfrequenz war nun konstant p=0,5) wiesen alle Schätzer (mit Ausnahme von HOM) die gleiche Genauigkeit auf (Gleichung 1). Der RMSE der Schätzungen mit idealen Markern lag im Durchschnitt um ~15 % unter den Schätzungen des Likelihood-Schätzers FL mit zufälligen Markern. Der RMSE war bei Verwendung von fünf Markern um 19 % niedriger, bei 50 Markern um 15 % niedriger und bei 100 Markern auf 13 % gesunken. Der Genauigkeitsgewinn der idealen Marker war somit bei einer geringeren Anzahl von Markern größer.

Unter Verwendung des genauesten Schätzers (FL) schätzten wir die Verteilung der individuellen F-Werte für (1) die 728 Individuen aus dem GWBP, die mit dem Axiom-SNP-Array genotypisiert wurden (11.467 Marker; Abb. 4a), (2) der zusätzlichen 624 Individuen aus dem GWBP, die mit den 50 nahezu idealen KASP-Markern genotypisiert wurden (Abb. 4b), und (3) der beiden Generationen der synthetischen Sommerackerbohnen-Sorte (Syn-4 und Syn-8; Fa. NPZ), die mit 43 KASP-Markern genotypisiert wurden (Abb. 4c und d).

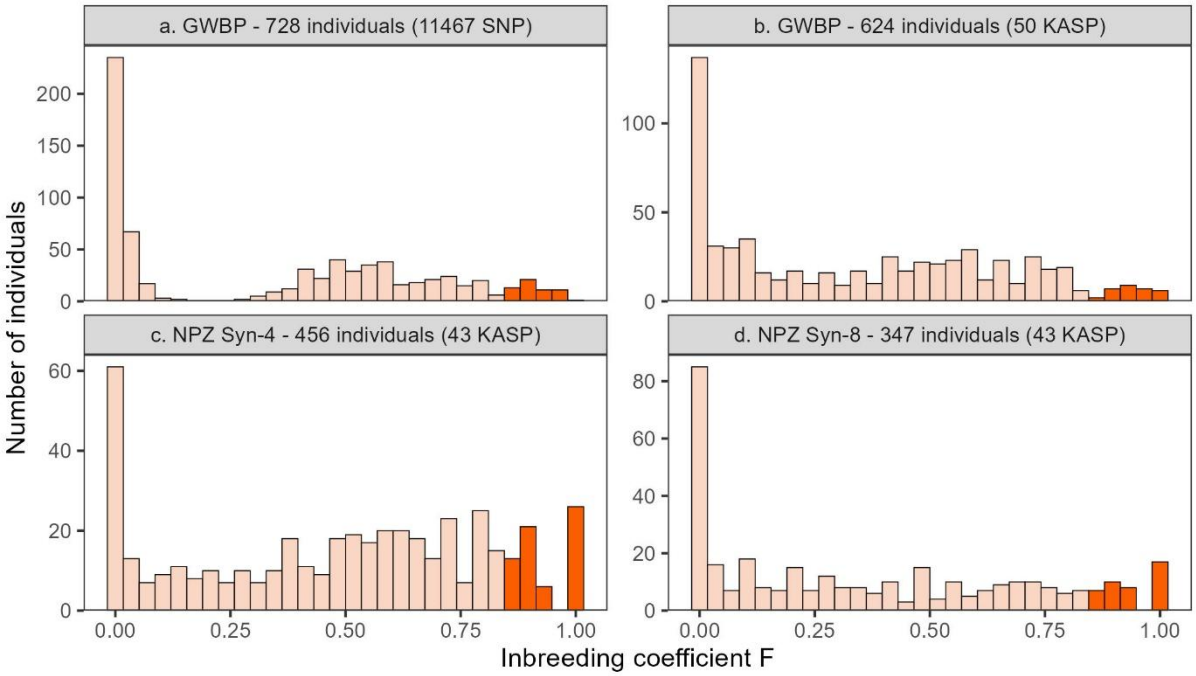


Abb. 4 Verteilung der geschätzten individuellen F-Werte in der Göttinger Winterbohnenpopulation (GWBP) a) bei Individuen, die mit dem Axiom-SNP-Array (11.467 SNP-Marker) genotypisiert wurden, b) bei Individuen, die nur mit 50 jedoch nahezu idealen (p ≈ 0,5; LD ≈ 0) KASP-SNP-Markern genotypisiert wurden, sowie in der synthetischen Sorte des Züchtungsunternehmens NPZ Hans-Georg Lembke KG in Generationen Syn-4 (c) und in Generation Syn-8 (d). In dunklerem Orange hervorgehoben ist der Anteil hoch ingezüchteter Individuen, basierend auf unserem Inzuchtschwellenwert (F ≥ 0,8$\overline{3}$).

Insgesamt wurde bei 61 von 728 Individuen aus der GWBP (etwa 8,4 %) ein F-Wert von mindestens unserem Inzuchtschwellenwert (FL ≥ 0,8$\overline{3}$) geschätzt, sodass diese als hoch-ingezüchtet eingestuft wurden (Abb. 4a). Dieser Anteil war im Vergleich zur simulierten, partiell allogamen Population etwas geringer (8,4 % < 10,6 %). In der GWBP beobachteten wir einen mittleren Inzuchtwert aller Individuen von Populations-$\overline{F}$ = 0,35. Daraus lässt sich ableiten, dass diese Population unter idealen Bedingungen einen Selbstbefruchtungsgrad von $\overline{S}$ = 0,52 realisierte. Die F-Schätzungen anhand von 50 Markern (Abb. 4b) ergaben für beide Parameter identische Werte ($\overline{F}$=0,35; $\overline{S}$ =0,52). Von den 624 Individuen wurden 34 als hoch-ingezüchtete Individuen eingestuft (5,4 % mit FL ≥ 0,8$\overline{3}$). In den beiden Generationen der synthetischen Sorte aus NPZ (Abb. 4c und d) fanden wir einen Populations-$\overline{F}$ von 0,49 in Syn-4 und von 0,37 in Syn-8 und leiteten daraus ab, dass $\overline{S}$ =0,65 bzw. 0,54 beträgt. Die Anzahl der hoch ingezüchteten Individuen in dieser Sommerbohnen-Population (FL ≥ 0,8$\overline{3}$) betrug in Syn-4 N=70 (15,4 %) und in Syn-8 N=46 (13,3 %).

Die Klassifizierungsleistung wurde stark vom Auskreuzungsgrad der Population beeinflusst (Abb. 5). Eine höhere Genauigkeit bei der F-Schätzung (d. h. eine größere Anzahl von Markern mit idealen statt zufälligen Allelfrequenzen) und ein höherer ICT führten stets zu einer Verringerung der Anzahl der FP. Dies galt jedoch nicht für die Anzahl der TP. Wenn der gewählte ICT nahe am Inzuchtschwellen-

wert lag (F = 0,8$\overline{3}$), erhöhte eine höhere Genauigkeit bei der F-Schätzung die Anzahl der TP, da dadurch die Anzahl derjenigen hoch-ingezüchteten Individuen sank, die fälschlicherweise als unzureichend ingezüchtet klassifiziert wurden (falsch-Negative). Im Gegensatz dazu erhöhte eine geringere Genauigkeit der F-Schätzung die Anzahl der TP, wenn der gewählte ICT nahe bei 1 lag, da tatsächlich hoch-ingezüchtete Individuen mit einem tatsächlichen F-Wert unter 1 möglicherweise als vollständig ingezüchtet (F=1) eingeschätzt wurden.

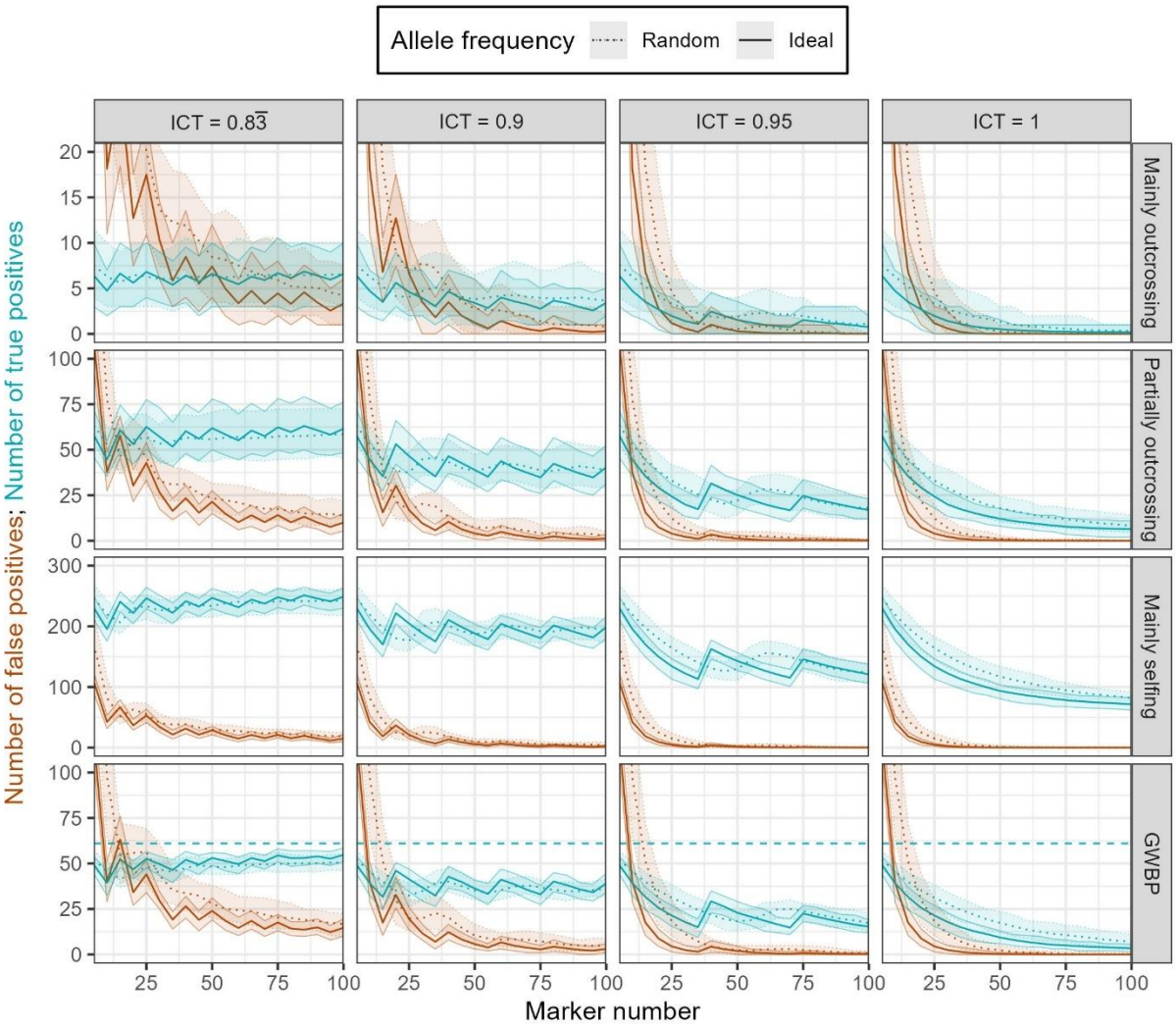


Abb. 5 Anzahl der wahren Positiven und der falsch Positiven (unter 728 Individuen) bei der Klassifizierung hoch ingezüchteter Individuen in Abhängigkeit von der Anzahl der für die F-Schätzung verwendeten Marker, deren Allelfrequenz (zufällig oder ideal, gestrichelte oder durchgezogen Linie), dem verwendeten Schwellenwert für die Inzuchtklassifizierung (ICT; vier vertikale Facetten) in unseren drei simulierten Populationen (überwiegend auskreuzend, partiel allogam, überwiegend autogam) sowie dem GWBP (vier horizontale Facetten). Beachte, dass die Skala der y-Achse je an die Gesamtzahl der in der Population verfügbaren positiven Ergebnisse angepasst ist. Für die GWBP-Ergebnisse haben wir die Gesamtzahl der positiven Ergebnisse (61), d. h. die maximale Anzahl echter positiver Ergebnisse, durch eine horizontale blaue gestrichelte Linie gekennzeichnet.

Die Linien stellen den Mittelwert der 50 Iterationen dar, die Bänder stellen das 90-%-Konfidenzintervall über die 50-Marker-Teilmengen dar.

Die aus der simulierten, partiell allogamen Population gewonnenen Ergebnisse ähneln deutlich denen, die mit der GWBP beobachtet wurden (Abb. 5), was eine gute Übertragbarkeit der simulierten Ergebnisse auf eine bestehende Population erwarten lässt. Simulationen ermöglichten robuste Vorhersagen zur Anzahl der TP und FP in der GWBP. Angenommen, wir hätten beispielsweise eine ideale Marker-Teilmenge von 50 Markern mit ICT=0,9, dann ergab die Klassifizierung in den Simulationen durchschnittlich 39 TP und vier FP, während es in unserer realen Population 36 TP und sechs FP gab.

Durch die Verwendung der vorläufigen Klassifizierung konnten die Genotypisierungs-Kosten deutlich gesenkt werden. Die größten Einsparungen würden bei einer geringen Anzahl von Markern erzielt (zwischen 15 und 30 Markern). So konnten wir beispielsweise bei unserer Population von 728 Individuen bis zu 46 % der berechneten Genotypisierungs-Kosten einsparen, wenn im vorläufigen Klassifizierungsschritt 20 Marker (M1=20) und im Hauptklassifizierungsschritt 80 weitere Marker verwendet wurden (M2=20+80=100).

4.2. Ergebnisse des Arbeitspakets 4; 5; 6; 8: Leistung der genomischen Vorhersage innerhalb und über Generationen hinweg in einer Winterackerbohnenpopulation

Die Spearman-Korrelation zwischen der Allelfrequenz der 13.840 SNP-Marker im A-Set und im AD-Set betrug r_S=0,75. Die durchschnittliche Differenz der Allelfrequenz eines Markers im Vergleich von A-Set und AD-Set betrug Δp=0,09. Die Spearman-Korrelation zwischen den LD Werten von Marker-Paaren (erfasst als r; Marker-Paare in gleitendem Rahmen von 50 physisch benachbarten Markern) in A-Set und AD-Set (r_S = 0,70) war geringer als die Korrelation zwischen den Allelfrequenzen. Die durchschnittliche LD-Veränderung betrug (absolut) Δp=0,16. Markerpaare, deren LD-Phase sich sogar von Coupling zu Repulsion und umgekehrt zwischen den Sets unterschieden, machten 23 % aller Markerpaare aus. Die ersten beiden Hauptkomponenten der PCA erklärten einen sehr geringen Anteil der SNP-basierten Varianz (0,02 %). Dies war angesichts der hohen Anzahl von 674 analysierten Genotypen zu erwarten. Wir beobachteten eine klare Trennung der beiden Sätze in der Projektionsfläche (Laugel, 2026; Seite 47, Abbildung 3.1).

Die mittlere paarweise genomische Verwandtschaft innerhalb des A-Satzes und des AD-Satzes betrug 0,0218 bzw. 0,0297, und zwischen den Sätzen lag der Wert bei -0,0324. Die Variation der paarweisen genomischen Verwandtschaft war im A-Set (σ=0,11) höher als im AD-Set (σ=0,08) und zwischen den beiden Sätzen noch geringer (σ=0,04). Die UPGMA-Ergebnisse verstärkten den Eindruck der PCA-Ergebnisse - die allermeisten Linien des A-Sets (181 von 185) waren eindeutig zusammen-gruppiert waren.

Die genetischen Varianzen (σ^2_G) aller Merkmale in allen Versuchen erwiesen sich als signifikant, mit Ausnahme von (Merkmalsabkürzungen siehe Tabelle 3):

Dev1 im HFT;

Flbeg im WEE2024+;

Pla1C, Pla2C, Ph2 und Rust im WEE2025;

Pla1C, Pla2C im BIG2025.

In den HFT- und im BIG2025-Versuch (Daten, die für das Training des GP-Modells verwendet wurden) waren die Heritabilitäten mäßig hoch bis hoch; sowohl für agronomische als auch für phänologische Merkmale (Ph1, Ph2, Lg1, Lg2, Yld, TGW hatten einen mittleren Wert in den HFT von H^2 = 0,83 und in BIG2025 von H^2 = 0,73; Flbeg, Flend, Ma1, Ma2 hatten in den HFT einen mittleren Wert von H^2 = 0,80 und in BIG2025 von H^2 = 0,62; Tabelle 4). Interessanterweise beobachteten wir eine höhere Heritabilität für den Ertrag im BIG2025 (H^2 = 0,77; was höher war als H^2 = 0,75 in den HFT), obwohl die Anzahl der Umgebungen für den Ertrag geringer war (2; deutlich niedriger als 17 im HFT). Die Heritabilitäten waren für Überwinterungseigenschaften nur moderat (durchschnittlicher H^2-Wert von 0,42 im HFT) und Krankheitstoleranzen (durchschnittlicher H^2-Wert von 0,33 im HFT). Die für die Merkmale der WEE-Versuche beobachteten Fehlervarianzen (σ^2_e) zeigten die geringere Zuverlässigkeit der gesammelten phänotypischen Daten für diese Versuche (insbesondere für Merkmale, die erfahrungsgemäß stark von einer geringen Parzellengröße beeinflusst wurden, wie beispielsweise Pflanzenhöhe und Ertrag).

Die Predictive Ability der Kreuzvalidierung innerhalb des A-Sets waren relativ hoch, insbesondere für diejenigen agronomische und phänologische Merkmale, die höhere Heritabilitäten aufwiesen. Im Durchschnitt zeigten Überwinterungsmerkmale und Krankheitsresistenz etwas geringere Werte. Der Blühbeginn war das Merkmal mit der höchsten Predictive Ability (r = 0,59), und selbst der Ertrag wies noch eine fast moderate Vorhersagefähigkeit von r = 0,39 auf. Erwartungsgemäß zeigte die Entwicklung der Pflanzen vor Winter (Dev1) die geringsten Vorhersagefähigkeiten, was mit der fehlenden signifikanten genetischen Varianz übereinstimmte (Tabelle 4).

Insgesamt korrelierten die Predictive Ability Werte positiv mit den Heritabilitäten der phänotypischen Daten (diese Korrelation betrug r = 0,75 für das A-Set und r = 0,59 für den BIG2025-Datensatz). Die mit dem BIG2025-Datensatz erzielten Ergebnisse waren denjenigen des A-Sets ähnlich, was aufgrund der vergleichbaren Heritabilitäten in den HFT- und BIG2025-Versuchen zu erwarten war. Die Merkmale Pflanzenhöhe und Tage bis Blühende, die im BIG2025-Versuch niedrigere Heritabilitäten gezeigt hatten (Tabelle 4), wiesen tatsächlich auch geringere Werte für Predictive Ability auf.

Tabelle 4: Zusammenfassung der BLUEs, die aus der phänotypischen Datenanalyse der Feldversuche gewonnen wurden. Genotypen aus dem A-Set (historischen Feldversuchen; HFT) wurden in einigen Versuchsstandorten als Linien (L *per se*) und in weiteren Versuchsstandorten als Ex-Trial-Nachkommen (P) dieser Reinzuchtlinien getestet (ergibt insgesamt L+P). Genotypen aus dem AD-Set wurden in den Versuchen WEE2024, WEE2024+, WEE2025 und BIG2025 getestet. Die Merkmale sind in der Legende zu Tabelle 3 beschrieben. Für Merkmale, die in nur einer Umwelt erfasst wurden, ist (in Klammern) die *Wiederholbarkeit statt der dann nicht darstellbaren Heritabilität H^2 angegeben.

Merkmal	Versuch	Mittel	Min.	Max.	SD	CV (%)	H^2 (*)
Pla1C	HFT (L+P)	49.29	36.90	54.53	2.83	5.74	0.48
	WEE2024	0.86	0.37	1.10	0.12	13.80	0.20
	WEE2024+	0.86	0.43	1.00	0.12	13.47	(0.11)
	WEE2025	0.91	0.67	1.00	0.06	6.46	(0.24)
	BIG2025	0.81	0.52	0.98	0.08	9.70	0.00
Pla2C	HFT (L+P)	54.17	30.07	61.37	4.12	7.60	0.64
	WEE2024	0.63	0.22	1.05	0.14	22.31	0.24
	WEE2024+	0.81	0.27	1.03	0.14	17.26	(0.51)
	WEE2025	0.96	0.80	1.02	0.05	5.02	(0.08)
	BIG2025	0.75	0.43	0.91	0.08	10.50	0.00
Dev1	HFT (L+P)	7.81	5.70	8.73	0.57	7.31	0.01
	BIG2025	8.02	3.99	9.04	0.77	9.59	(0.73)
Dev2	HFT (L+P)	6.59	1.60	8.52	0.71	10.72	0.43
	WEE2024	6.77	2.00	9.00	1.10	16.20	0.43
	WEE2024+	6.39	1.00	9.00	1.63	25.46	(0.69)
Cano	HFT (L+P)	6.52	2.37	7.67	0.58	8.90	0.54
Flbeg	HFT (L+P)	132.06	125.82	139.79	3.48	2.64	0.94
	WEE2024	126.81	121.78	136.00	2.27	1.79	0.68
	WEE2024+	127.09	121.00	137.00	2.51	1.97	(0.14)
	WEE2025	123.70	119.00	131.00	3.45	2.79	(0.94)
	BIG2025	124.29	117.81	131.27	3.58	2.88	0.91
Flend	HFT (L+P)	160.03	155.52	165.58	1.92	1.20	0.78
	WEE2024	168.59	160.27	176.27	2.55	1.51	0.37
	WEE2024+	168.78	162.00	176.00	2.93	1.74	(0.34)
	WEE2025	160.93	153.00	165.00	3.05	1.89	(0.66)
	BIG2025	159.58	153.53	165.77	2.91	1.82	0.43
Ma1	HFT (L+P)	5.28	3.44	7.26	0.74	13.98	0.67
	WEE2025	5.54	1.00	9.00	1.71	30.83	(0.81)
	BIG2025	5.51	1.98	8.78	1.42	25.78	0.68
Ma2	HFT (L+P)	7.25	3.54	8.89	0.72	9.97	0.82
	BIG2025	6.19	2.23	9.08	1.38	22.37	(0.66)
Asco	HFT (L+P)	2.32	1.76	4.22	0.34	14.86	0.19
Pero	HFT (L+P)	2.38	1.68	3.79	0.35	14.80	0.46
	WEE2025	3.52	1.00	7.00	1.48	42.06	(0.83)
	BIG2025	3.65	0.78	9.44	1.74	47.77	(0.89)
Choc	BIG2025	3.73	1.00	7.87	1.13	30.35	(0.48)

Tabelle 4: Fortsetzung

Merkmal	Versuch	Mittel	Min.	Max.	SD	CV (%)	H^2 (*)
Rust	WEE2025	6.19	4.00	9.00	1.30	21.05	(0.34)
	BIG2025	6.27	2.18	10.01	1.39	22.12	(0.49)
Ph1	HFT (L+P)	94.47	68.98	108.41	6.08	6.44	0.88
Lg1	HFT (L+P)	3.58	2.37	5.72	0.69	19.40	0.78
	WEE2025	5.75	1.00	9.00	2.40	41.80	(0.79)
	BIG2025	3.20	0.50	8.90	2.09	65.19	0.77
Lg2	HFT (L+P)	3.60	1.93	7.00	0.95	26.43	0.86
	WEE2024	4.19	0.94	8.94	1.56	37.18	(0.5)
	WEE2024+	4.03	1.00	9.00	1.61	40.02	(0.53)
	BIG2025	5.25	0.99	8.61	2.03	38.60	(0.74)
Yld	HFT (L+P)	44.27	23.34	54.88	4.92	11.11	0.75
	WEE2024	62.64	3.57	131.16	25.40	40.54	(0.4)
	WEE2024+	64.76	0.00	187.96	32.12	49.60	(0.37)
	WEE2025	39.39	14.07	70.18	12.64	32.10	(0.61)
	BIG2025	35.47	11.97	52.91	6.95	19.58	0.77
TGW	HFT (L+P)	509.86	379.57	620.84	50.43	9.89	0.92
	WEE2024	406.63	272.28	585.68	47.90	11.78	(0.64)
	WEE2024+	413.53	297.53	604.20	51.11	12.36	(0.51)
	WEE2025	428.52	330.07	556.67	49.70	11.60	(0.67)
	BIG2025	427.48	339.83	556.50	45.63	10.67	0.77
Prot	BIG2025	23.77	20.47	26.78	1.29	5.41	(0.81)

Überraschenderweise war das Tausendkorngewicht (TKG) im BIG2025-Datensatz das am besten vorhersagbare Merkmal (r = 0,45) und übertraf damit sogar die Predictive Ability im A-Set (r = 0,39), obwohl seine Heritabilität hier niedriger war als im A-Set. Im Gegensatz dazu zeigte der Blühbeginn, dessen Heritabilität in beiden Datensätzen ähnlich war, im BIG2025-Datensatz eine deutlich geringere Vorhersagefähigkeit (r = 0,39).

Die Botrytis-Resistenz zeigte die geringste Vorhersagefähigkeit (r = 0,04), während die Rost- und Peronospora-Resistenz mit r = 0,43 bzw. r = 0,40 höhere Vorhersagefähigkeiten aufwies. Die hohen Werte dieser beiden letztgenannten Merkmale könnten durch die Ähnlichkeit der Genotyp-Umwelt-Interaktionen zwischen Trainings- und Testdatensätzen erklärt werden, da beide an einem gemeinsamen Standort geprüft wurden.

Die Predictive Ability und Accuracy der generationsübergreifenden genomischen Vorhersage waren relativ gering (Abb. 6). Ähnliche Ergebnisse wurden auch für die genomische Vorhersage innerhalb derselben Generation beobachtet, bei der

Genotypen aus den WEE-Versuchen anhand des BIG2025-Datensatzes vorhergesagt wurden (Abb. 6b).

Bei den generationsübergreifenden Vorhersagen war das Tausendkorngewicht (TKG) das Merkmal mit der höchsten Vorhersagefähigkeit im BIG2025-Versuch (Abb. 6a), was mit seiner hohen Vorhersagbarkeit in den Kreuzvalidierungen innerhalb desselben Datensatzes übereinstimmt. Ebenso zeigte im A-Set der Blühbeginn sowohl in den Kreuzvalidierungen als auch in den generationsübergreifenden Vorhersagen die höchste Vorhersagefähigkeit (Abb. 6b).

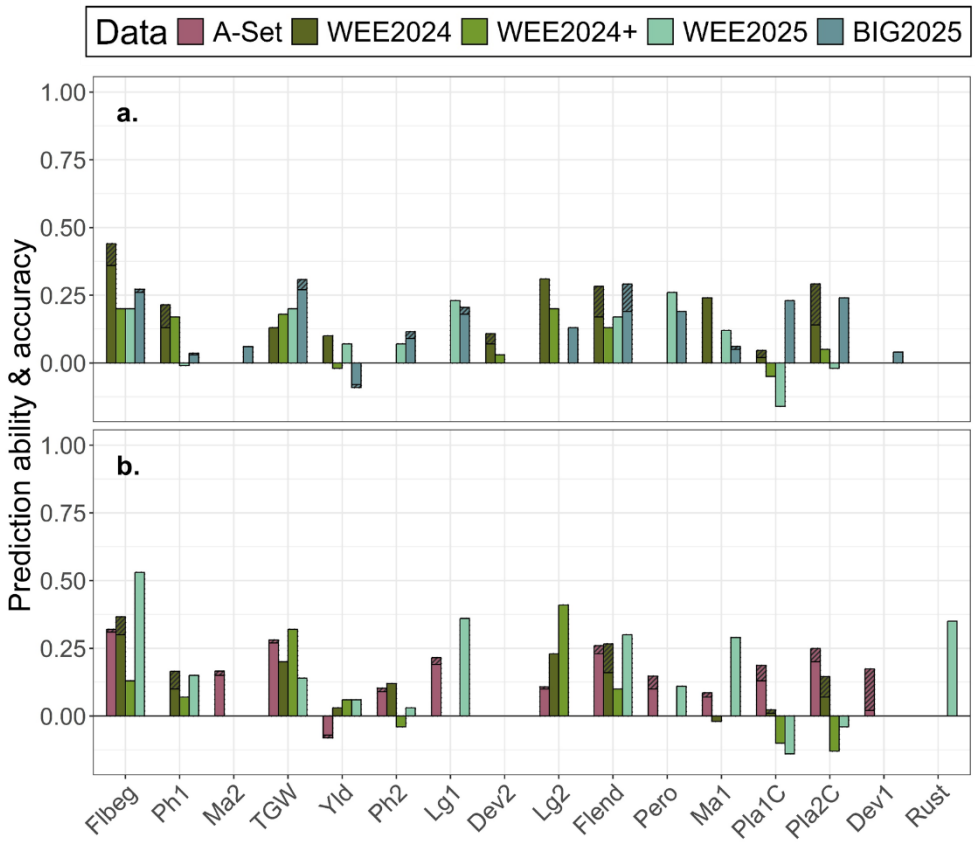


Abb. 6 Vorhersagefähigkeiten (Predictive ability) und Vorhersagegenauigkeiten (Accuracy; schraffierte Verlängerung der Balken) aller untersuchten Merkmale mit dem GBLUP-Modell: **a**. trainiert mit dem A-Set (185 Genotypen) zur Vorhersage der Abo-Direkt-Feldversuche, und **b**. trainiert mit den im BIG2025-Versuch getesteten Genotypen (143 Genotypen) zur Vorhersage des A-Sets oder der Genotypen aus dem AD-Set, die in den anderen-als BIG2025-Feldversuchen geprüft wurden. Die Merkmale wurden entsprechend der Reihenfolge in Abb. 2 angeordnet (nach absteigender Predictive Ability innerhalb derselben Generation). Die Streudiagramme veranschaulichen für Merkmal „Tage bis Blühbeginn“ (Flbeg) die Beziehung zwischen den beobachteten phänotypischen Werten aus den Feldversuchen und den mit dem Modell vorhergesagten Werten (GEBVs). Alle Merkmale sind schon in der Legende von Tabelle 3 beschrieben.

Im Fall des Kornertrags standen die relativ hohen Vorhersagefähigkeiten aus den Kreuzvalidierungen jedoch in starkem Gegensatz zu den sogar negativen Vorhersagefähigkeiten, die für BIG2025 (r = -0,08) und das A-Set (r = -0,07) erhalten

wurden. Unabhängig vom gewählten Datensatz waren die Vorhersagefähigkeiten für alle untersuchten Merkmale relativkonsistent – sowohl bei der Vorhersage des älteren Gentoypen-Satzes (Vorhersage des A-Sets anhand von BIG2025-trainiertem Modell; Abb. 6b) als auch bei der Vorhersage des neueren Satzes (Vorhersage der Leistungen der BIG2025-Genotypen anhand des A-Set-trainierten Modells; Abb. 6a).

Die Vorhersage neuer Genotypen aus demselben Set wie das Training-Set (in neuen Umwelten) wurde mit der Vorhersage neuer Genotypen aus dem späteren Set (ebenso in neuen Umwelten) verglichen. Es wurde wie oben schon beschrieben das Modell mit Daten der HFT (L) trainiert und ein Validationssatz weiterer also neuer Linien (A-Set; HFT; P) wurde benutzt. Und es wurde ein Modell mit den Daten der BIG2025-Versuche trainiert, die Validation war dann mit A-Satz-Linien (generationsübergreifen) oder mit anderen AD-Linien (innerhalb derselben Generation; Abb. 5; vergl. Tabelle 3). Es wurde dies für die folgende Merkmale durchgeführt; Blühbeginn, Reife, Pflanzenhöhe, Lageranfälligkeit, Ertrag und TKG.

Im Durchschnitt betrug die Vorhersagegenauigkeit der untersuchten Merkmale 0,39 für den HFT (P) und 0,12 für den BIG2025. Die mit den genetischen Unterschieden zwischen den zwei Sets einhergehende Differenz der Genauigkeiten (Vorhersage von Genotypen aus demselben bzw. aus dem anderen Set) betrug somit im Mittel 0,27 über alle Merkmale hinweg.

Pflanzenhöhe und Reife waren die am stärksten betroffenen Merkmale mit einer Vorhersagegenauigkeit von 0,46 bzw. 0,38 in den HFT (P), die im BIG2025 auf 0,05 bzw. 0,00 abfiel. Ebenso sank die Genauigkeit beim Ertrag von einem Wert von r= 0,31 auf den Wert von r=-0,01. Der Blühbeginn und das TKG waren die einzigen Merkmale, die noch eine irgendwie nutzbare Vorhersagegenauigkeit zeigten, wenn das AD-Set mit dem A-Set(L)-trainierten Modell vorhergesagt wurde (r = 0,28 bzw. r = 0,30).

4.3. Ergebnisse des Arbeitspakets 9: Optimierung des Seed-Chipping für die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen

Das von uns verwendete Protokoll funktionierte zunächst für unsere Proben deutlich nicht zufriedenstellend. Tastversuche ergaben, dass die PCR-Zyklenzahl nicht für unsere Ackerbohnensamen optimiert waren. Es waren schlussendlich fünf zusätzliche PCR-Zyklen erforderlich, damit die Proben im Allel-Diskriminierungs-Plot der üblichen KASP-Analyse korrekt gruppiert wurden. Diese Vorgehensweise wurde von Post et al. (2001) für die DNA-Extraktion aus Gerstensamen empfohlen. Der endgültige Allel-Diskriminierungs-Plot nach den zusätzlichen Zyklen zeigte eine korrekte Clusterbildung der DNA-Proben.

Bei trockenen Samen zeigte sich: (1) Schneiden funktionierte gut, war jedoch sehr zeitaufwendig; das Abschaben der Keimblätter mit einer Pinzette war recht mühsam und erlaubte keine ausreichend konsistente Probenmengen. (2) Stanzen konnte erwartungsgemäß nicht verwendet werden, da die trockenen Samen zu hart waren. (3) Bohren erwies sich als weniger zeitaufwendig und weniger anstrengend als das Schneiden und ermöglichte gleichzeitig eine relativ

gleichmäßige Probenentnahme. Allerdings führte das Bohren bei trockenen Samen zu einer Kontamination der Arbeitsumgebung mit Gewebepulver (sichtbar in Abb. 2).

Bei feuchten Samen zeigte sich: (1) Schneiden war sehr einfach, jedoch konnten die feuchten Keimblätter nicht abgeschabt werden, sodass statt Pulver größere Keimblattstücke als Proben entnommen werden mussten, was eine konsistente Probenentnahme erschwerte. (2) Stanzen war sehr einfach und stellte die Methode mit der höchsten Konsistenz bei der Probenentnahme dar. (3) Bohren war in den aufgeweichten Keimblättern nicht praktikabel.

Insgesamt zeigte sich, dass das Stanzen bei feuchten Samen und das Bohren bei trockenen Samen bis dahin die besten Chipping-Techniken darstellen. Einige KASP-Analysen wurden mit Proben durchgeführt, die mit diesen beiden Methoden gewonnen wurden, diese Analysen waren alle erfolgreich.

Im Vor-Quell-Experiment (Einweichen) benötigten die Samen des Genotyps, der im Isolierhaus angebaut worden war, länger zum Quellen und Aufweichen als Samen der beiden anderen Genotypen, deren Erntegut aus dem Feld stammte. Zudem wurde innerhalb derselben Kombination aus Genotyp und Wassermenge eine erhebliche Variation im Einweichverhalten der Samen beobachtet: Einige Samen waren bereits nach wenigen Stunden weich und begannen sogar zu keimen, während andere deutlich länger hart blieben.

Einzelne Samen benötigen somit unterschiedliche Einweichzeiten, bevor sie ausreichend weich sind, um beprobt zu werden. Aus unserer Sicht spricht diese Beobachtung gegen den Einsatz jeglicher Chipping-Technik an vorgequollenen Samen, da die Einweichdauer individuell für jeden einzelnen Samen angepasst werden müsste. Möglicherweise könnte dieses Problem durch weitere Maßnahmen überwunden werden, solche Versuche wären in weiteren Arbeiten vorzusehen.

Darüber hinaus lag der Feuchtigkeitsgehalt der Samen – ausgehend von einem Ausgangsfeuchtegehalt der trockenen Samen von etwa 12 % (GRDC, 2018) - nach dem Einweichen, als sie gerade weich genug zur Probenahme waren - zwischen 43 % und 54 %. Wir gehen davon aus, dass diese hohe Feuchte die Fähigkeit der Samen beeinträchtigt, anschließend wieder rückgetrocknet und gelagert zu werden, damit dann zu einem späteren Zeitpunkt die GP-basiert Selektierten erfolgreich keimen (GRDC, 2017). Auch hier könnten Anschluss-Arbeiten sinnvoll sein.

Es wurde keine Kreuzkontamination zwischen den trocken angebohrten Samenproben der Sorten „Côte d'Or“ und „Hiverna“ beobachtet, obwohl diese abwechselnd nacheinander und ohne Reinigung des Bohrers bearbeitet wurden. Insgesamt sprechen die Ergebnisse für das Bohren an trockenen Samen als derzeit einfachste und effizienteste Technik zur zerstörungsfreien DNA-Extraktion aus Ackerbohnensamen.

Zwischen der Kontrollgruppe und der behandelten (gechippten) Gruppe wurde hinsichtlich der im Experiment erhobenen Keimungs-Merkmale kein signifikanter Unterschied festgestellt (Tabelle 5). Die t-Test-p-Werte für die Keimzeit (gDAS), die Sämlingslänge (SL) und den SVI betrugen 0,74, 0,66 bzw. 0,63.

Tabelle 5: Zusammenfassung der phänotypischen Datenanalyse für die Kontroll- (nicht gechippt) und die Chip-Behandlung (jeweils 140 einzelne Samen) während des Gewächshausexperiments zur Untersuchung des Einflusses des Seed-Chippings auf die Samenviabilität. SD steht für die Standardabweichung und CV (%) für den Variationskoeffizienten. gDAS bezeichnet die Keimzeit im Verlauf des Experiments (19 Tage nach der Aussaat), G das binäre Merkmal für Keimung (0 = nicht gekeimt; 1 = gekeimt), SL die Sämlingslänge (in cm) und SVI den Keimkraftindex (Seedling Vigor Index).

Merkmal	Kontrolle					Chipped				
	Mittel	Min	Max	SD	CV (%)	Mittel	Min	Max	SD	CV (%)
gDAS	10.06	8.00	17.00	1.75	17.40	9.98	8.00	19.00	2.25	22.51
G	0.97	0.00	1.00	0.17	17.21	0.94	0.00	1.00	0.23	24.71
SL	13.46	3.00	22.00	3.26	24.24	13.65	3.00	20.00	3.81	27.94
SVI	0.13	0.03	0.22	0.03	24.42	0.13	0.03	0.20	0.04	28.51

Die Samenkeimung wurde durch das Chipping nur sehr geringfügig reduziert. Da die Keimung ein binäres Merkmal ist, wurde eine logistische Regression durchgeführt, um die Signifikanz dieser durch das Chipping verursachten Reduktion zu prüfen. Auch hierbei wurde kein signifikanter Unterschied festgestellt, da die Anpassung des Gesamtmodells nicht signifikant besser war als die des Nullmodells ($\chi^2(1) = 1{,}418$, $p = 0{,}23$).

5. Diskussion der Ergebnisse

Der hier vorgestellte Ansatz, Homozygote rasch und direkt aus einer segregierenden Population zu entnehmen, soll im Vergleich zur Ramschzüchtung (methodische Variante der Linienzüchtung bei Selbstbefruchtern) diskutiert werden. Für die Ramschzüchtung führt man eine Anzahl Kreuzungen durch, die weniger auf die unmittelbaren, sondern mehr auf spätere und spekulative Zuchtziele ausgerichtet sind. Die Kreuzungen werden mit minimalem Aufwand zur Homozygotie geführt (‚Ramsch'), im Wesentlichen ohne Selektion und Drift; Ramschzüchtung ist für Selbstbefruchter (selbsten sich ‚von selbst'). Wenn man dann Generationen später in das Material hineingreift, ist das gesamte Material hochgradig homozygot. Das Potenzial der Individuen kann zwar durch die Eltern der jeweiligen Kreuzung eingeschätzt werden kann, im Einzelnen aber sind die Genotypen unbekannt. Wir ersetzen hier sozusagen die Vielzahl von Einzelkreuzungen durch eine einzige, diverse Population. Von den hier detektierten homozygoten Individuen sind keine Kreuzungseltern bekannt. Wir müssen jedoch Kreuzungen gar nicht durchführen, weil die partielle Allogamie in jeder Saison Kreuzungen in der Population ‚von selbst' realisiert. Das gilt insbesondere, wenn wir der Population Elitematerial zuführen – in der Ramschzüchtung von Selbstbefruchtern ist dieser Weg einer stetigen Populations(nach)verbesserung versperrt. Anders als in der Ramschzüchtung wissen wir in unserem Vorschlag nicht, zu welcher Generation ein konkretes Individuum aus der Population gehört, F1 oder F2 oder F3 etc.. Wir

müssen dies per Markeranalyse feststellen. Wenn jedoch absehbar ‚sowieso' für die Genomische Vorhersage Individuen in großer Anzahl genotypisiert werden, dann erhalten wir diese Information als Nebenprodukt: Außerdem erhalten wir über die Einschätzung der Filial-Generation hinaus auch einen Wert für den konkreten, Individuums-spezifischen Homozygotiegrad.

Im Unterschied zu DH-Züchtung erhalten wir – wie in der Ramschzüchtung – leider nicht Homozygote von Kreuzungen, die wir erst ganz kürzlich durchgeführt haben, also von Material, welches allgemein oder merkmalsspezifisch das allerjüngste Niveau der Kreuzungseltern abbildet. Im Unterschied zur Ramsch- oder auch Stammbaum-Züchtung aber können wird solches Material stetig in die Population hinein-migrieren lassen. Wir sind auch nicht gezwungen, nur ausreichend Homozygote (deren Genese schon mehrere Generationen zurückliegt) zu verwenden, sondern wir können auch Individuen der jüngeren Generationen (näher an der jüngsten Elite) zur Sortenzüchtung einsetzen (nach Einstufung durch die GP). Prinzipiell könnte dieses ab der Generation F2 auch in der Ramschzüchtung geschehen. Die Ramsch- und Linienzüchtung jedoch kann nicht wie unsere Population die notwendige Vielzahl von Zuchtzielen in der Rekombination realisieren, selbst wenn man Dreiwege- oder Doppelkreuzungen an ihren Beginn setzte. Eine MAGIC-Population als Beginn der Linienzüchtung hat ihre eigenen Parallelen zu unserem Vorschlag (hohe Vielfalt), dort aber ist es auch so, dass nach den Einstiegs-Kreuzungen keine weitere Ein- oder Durchkreuzung mehr geschieht, die Rekombination und die Vielfältigkeit ist eingeschränkt, anders als in unserer Population, die sich stets inniglich weiter-rekombiniert. Dieser Aspekt ist in unserer Population wie bei der Rekurrenten Selektion von Fremdbefruchtern – mit dem Unterschied, dass diese keine Homozygoten als Koppelprodukt anbieten.

Wie nach der Theorie zu erwarten, war in der Göttinger Winter-Ackerbohnen-Population (GWBP) ein signifikanter Anteil hoch-ingezüchteter Individuen vorhanden. Diese Individuen können durch die Schätzung ihrer individuellen F-Werte identifiziert werden, selbst bei einer begrenzten Anzahl eingesetzter Marker (insbesondre über den Likelihood-Schätzer FL). Hoch-ingezüchtete Individuen lassen sich effektiv mithilfe des ICT klassifizieren, der gezielt auf das für die F-Schätzung verwendete Markerset sowie auf die erwartete inzucht-mäßige Zusammensetzung der untersuchten Population abgestimmt ist.

Die Klassifikation kann kostengünstiger gestaltet werden, indem ein vorgelagerter Klassifizierungsschritt integriert wird, ohne die Leistungsfähigkeit der Klassifikation zu beeinträchtigen. Anhand unserer Ergebnisse können Züchter eine geeignete Anzahl von Markern finden, um Individuen ihrer eigenen Ackerbohnenpopulation kosteneffizient zu identifizieren. Die Ackerbohnenzüchtung kann erheblich von diesem Ansatz profitieren, da er einen schnellen, kostengünstigen und unkomplizierten Zugang zu ingezüchteten Ackerbohnenlinien ermöglicht – im Vergleich zur konventionell aufwendigen Erzeugung von Inzuchtlinien und gewissermaßen als Alternative zum nicht verfügbaren DH-Verfahren (doppelte Haploiden).

Die populationsgenetischen Unterschiede zwischen dem A-Set und dem AD-Set war erheblich, insbesondere hinsichtlich des Mittelwerts und der Varianz der genomischen Verwandtschaft sowie sichtbar an den Unterschieden im GPU zwischen beiden Sets – zwei zentrale Parameter, die der genomischen Vorhersage (GP) zugrunde liegen.

Die genomische Vorhersage innerhalb derselben Generation zeigte ein hohes Potenzial für den Einsatz von GP in der Ackerbohnenzüchtung, mit Vorhersagefähigkeiten (Predictive Ability) von bis zu 0,58 beim Merkmal Blühbeginn. Dieses Potenzial wurde jedoch in generationsübergreifenden Vorhersagen nicht realisiert, da Predictive Ability und Accuracy dann sehr deutlich reduziert waren.

Unter den untersuchten Merkmalen zeigte das GP Modell für Blühbeginn und Tausendkorngewicht unter diesem strengen Validierungsszenario noch eine gewisse Vorhersagekraft.

Die Studie vertieft somit das Verständnis für Potenzial und Grenzen der genomischen Vorhersage bei Ackerbohnen. In praktischen Züchtungsprogrammen sind tatsächlich aufgrund züchterischer Selektion, Drift und Populationsstruktur (Kreuzungen, Familien, separate Zuchtprogramme) ausgeprägtere genetische Unterschiede zwischen Trainingspopulationen und Zuchtmaterial zu erwarten. Züchterinnen und Züchter sollten daher vorsichtig sein, sie sollten nicht die Vorhersagegenauigkeiten aus Kreuzvalidierungen naiv als real einschätzen; sie sollten sehr regelmäßig ihre spezifischeTrainingspopulation aktualisieren, um deren genetischen Match zur Zuchtpopulation möglichst hoch zu halten.

Das Anbohren trockener Samen erwies sich hier als die beste der drei getesteten Chipping-Techniken. Diese Technik wird auch bei anderen Kulturarten wie Gerste oder Sojabohne eingesetzt (Kamiya & Kiguchi 2003; von Post et al. 2003). Offenbar hatte das Anbohren praktisch keinen Einfluss auf die Samenviabilität. Dieses Ergebnis war nicht vollständig überraschend, da die Keimrate von Ackerbohnensamen, die vom Bohnenkäfer (*Bruchus rufimanus*) befallen sind, ebenfalls relativ hoch bleibt (81,4 % laut Almogdad et al. 2023; der Käfer-Befall führt zu Löchern in den Samen, die deutlich größer sind als die hier durch das Anbohren erzeugten).

Der von uns verwendete Bohrer erzeugte im Vergleich sehr kleine Löcher (1 mm) und zielte auf die Keimblätter ab (wodurch der Embryo offensichtlich unversehrt blieb). Das von uns eingesetzte Setup zum Anbohren der Ackerbohnensamen (Abb. 2) funktionierte gut für eine geringe Anzahl an Samen. Für die 140 Samen, die zur Prüfung der Samenviabilität untersucht wurden, brauchten wir etwa eine Stunde für das Bohren.

Weitere Arbeiten sind erforderlich, es müssen noch technische Lösungen gefunden werden, um eine Hochdurchsatzanwendung in der praktischen Ackerbohnenzüchtung zu entwickeln.

Alle hier untersuchten Ansätze ermöglichen eine Beschleunigung der Ackerbohnenzüchtung. Insbesondere die Identifikation ingezüchteter Individuen aus Populationen führt zu einem Zeitgewinn beim Zugriff auf Inzuchtlinien. Der konventionelle Linienzüchtungsprozess dauert in der Regel mindestens zehn Jahre bis zur Entwicklung einer neuen Sortenlinie, vor allem aufgrund der Selbstungsgenerationen, die zur Erzeugung von hochgradig Homozygoten dienen.

Mit dem von uns vorgeschlagenen Ansatz können selektierte Zucht-Linien bereits in nur drei Generationen erzeugt werden (drei Jahre bei einer Generation pro Jahr; Abb. 7). Im ersten Jahr können die identifizierten ingezüchteten Individuen direkt durch kontrollierte Selbstbestäubung vermehrt werden. Im zweiten Jahr kann auf Basis der erzeugten Selbstungsnachkommen ein weiterer Vermehrungsschritt in Isolationskäfigen durchgeführt werden, und parallele kann ein

Polycross (PC) kann angelegt werden. Im dritten Jahr werden die Kandidaten im Feld getestet, *per se* unter Verwendung der Selbstungs-Samen sowie zusätzlich unter Verwendung der Polycross-Nachkommen (diese Daten entsprechen dann der Erfassung der Allgemeinen Sorteneignung alias General Varietal Ability, GVA; nach Link & Ederer 1993).

Die erhobenen Daten zur *per se* Leistung und zur allgemeinen Sorteneignung dienen anschließend zur Selektion der besten Kandidaten, entweder als zukünftige Liniensorte oder als Komponenten eines Syn-0 für eine zukünftige synthetische Sorte. Übrig gebliebene Selbstungs-Samen aus dem ersten oder zweiten Vermehrungsschritt können zur weiteren Vermehrung der ausgewählten Kandidaten genutzt werden.

Die Genotypen mit den besten GVA-Werten werden als Komponenten zu einer (oder mehreren) Syn-0 zusammengestellt. Dieses kann im vierten Jahr geschehen. Nach zwei oder drei Jahren der Vermehrung über Syn-1, Syn-2 und Syn-3 ist die synthetische Experimental-Sorte dann bereit für die offizielle Prüfung und gegebenenfalls für die Zulassung.

Unsere Studie ermöglicht die Anwendung der von Adhikari et al. (2021) vorgeschlagenen zweistufigen Strategie (Abb. 8). Die Identifizierung hoch ingezüchteter Individuen aus Populationen kann genutzt werden, um das Fehlen einer Doppelhaploid-Technologie bei Ackerbohnen zu kompensieren. Die zerstörungsfreie Genotypisierung von Samen mittels Seed Chipping kann genutzt werden, um genotypische Daten von geernteten Samen zu erheben. Anhand dieser Daten kann anschließend eine genomische Selektion an den Samen durchgeführt und die Zuchtpopulation durch wiederholte genomische Selektion verbessert werden. Das Vorhersagemodell kann anhand der Feldversuche in der Sortenproduktionsphase aktualisiert werden.

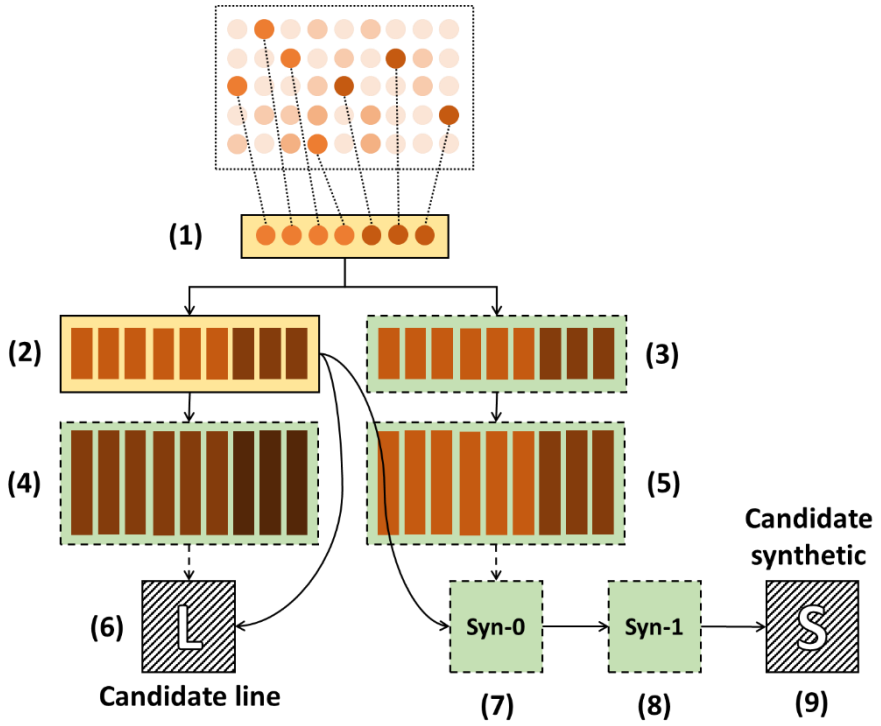


Abb. 7 Anwendung der Identifikation ingezüchteter Individuen aus Ackerbohnenpopulationen in einem vorgeschlagenen beschleunigten Züchtungsschema.

(1) Identifikation ingezüchteter Individuen aus einer offen bestäubenden Zuchtpopulation (Individuen werden vorzugsweise Same sonst als Keimling DNA-beprobt) und Vermehrung durch kontrollierte Selbstbestäubung.

(2) Vermehrung der Genotypen durch kontrollierte Selbstbestäubung.

(3) Parallel dazu Anlage eines Polycrosses (PC).

(4) Feldprüfung der Genotypen mit Selbstungssamen (*per se* Leistung).

(5) Parallel: Feldprüfung der Genotypen als Polycross-Nachkommenschaften (Erfassung der allgemeinen Sorteneignung; General Varietal Ability, GVA).

(6) Identifikation und dann Vermehrung der besten Genotypen aus Schritt (4) über verbliebene Samen aus Schritt (2) und Anmeldung als Linien-Sorten.

(7) Anbau des (der) bestvorhergesagten Syn-0 aus den besten Genotypen aus Schritt (5).

(8) Anbau von Syn-1.

(9) Vermehrung und Anmeldung als synthetische Sorte.

Die Kreise repräsentieren einzelne Genotypen. Die Rechtecke stellen Pflanzenreihen eines Genotyps dar. Die Farben der Kreise und Rechtecke zeigen den Inzuchtgrad an – von hellorange (nicht ingezüchtet) bis dunkelbraun (vollständig ingezüchtet).

Die gelb hinterlegten Blöcke stehen für Isolatierkäfige; die grün markierten Bereiche mit gestrichelten Linien stehen für Pflanzen im offenen Feld (Bestäuber-Zutritt).

Durchgezogene Pfeile zeigen an, dass die in einem Schritt erzeugten Samen im folgenden Schritt verwendet werden. Gestrichelte Pfeile zeigen an, dass in einem Schritt erhobene Daten für Selektionsentscheidungen genutzt werden.

Für kleine Züchtungsprogramme für Ackerbohnen bleibt die genomische Vorhersage (GP) eine kostspielige Technik und ist gegebenenfalls nicht zugänglich. Die Verbesserung des Zuchtmaterials kann alternativ zur direkten rekurrenten

Populationsverbesserung auch durch stete Zuführung von neuem Elite-Material stattfinden, welches auf anderen Wegen erzeugt wurde oder erlangt hat. Insbesondere können die besten synthetischen Kandidaten, die in der Sortenproduktionsphase erzeugt wurden, hier eingeführt werden und die so die Zuchtpopulation verbessern.

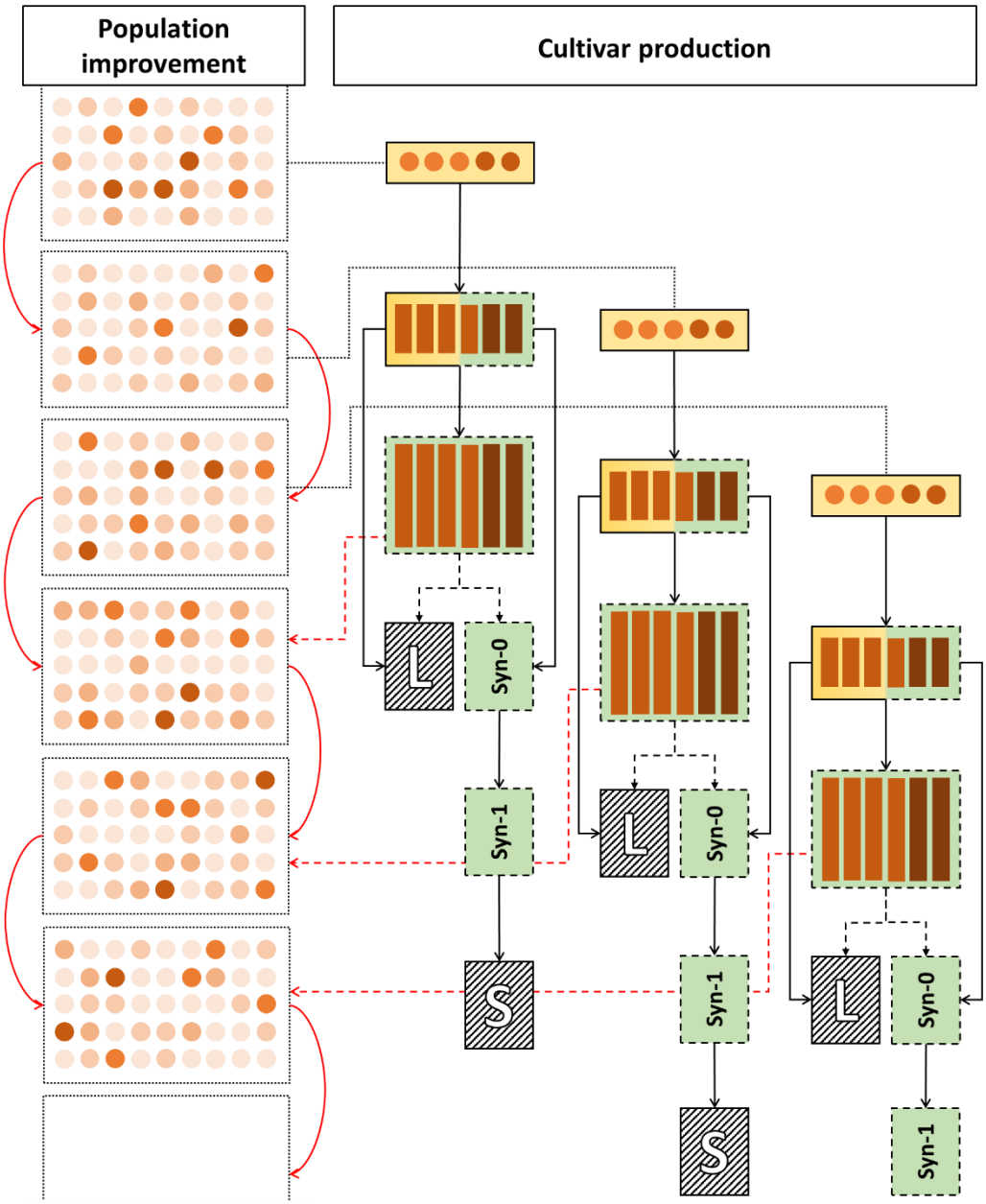


Abb. 8 Unsere vorgeschlagene zweiteilige Strategie, bestehend aus Populationsverbesserung (rekurrente Selektion, Vorlauf-Züchtung; links) und der Phase der Sorten-Züchtung (rechts). Die Phase der Sorten-Züchtung folgt dem in Abb. 7 dargestellten Züchtungsschema. Der Block mit gelbem und grünem Hintergrund stellt die Schritte (2) und (3) aus Abb. 7 zusammen dar. Der darunterliegende Block mit grünem Hintergrund stellt die Schritte (4) und (5) aus Abb. 7 zusammen dar. Rote Pfeile beziehen sich auf die Verwendung der

genomischen Vorhersage. Der durchgezogene rote Pfeil zeigt an, dass die in einem Schritt produzierten Samen mithilfe der genomischen Vorhersage ausgewählt werden (genomische rekurrente Selektion). Der gestrichelte rote Pfeil zeigt an, dass die in einem Schritt gesammelten Daten zum Trainieren des Modells für die genomische Vorhersage verwendet werden, das in dem Schritt eingesetzt wird, auf den er zeigt. Siehe auch die Erklärung zu Abb. 7.

6. Angaben zum voraussichtlichen Nutzen und zur Verwertbarkeit der Ergebnisse.

Es wird erwartet, dass die hier erzielten Ergebnisse der Züchtung von Ackerbohnen in hohem Maße zugutekommen werden. Das vorgeschlagene beschleunigte Züchtungskonzept (Abb. 7) und die Pipeline (Abb. 8; einschließlich rekurrenter genomischer Vorhersage und Selektion) zeigen den Zeit- und Effizienzgewinn, den die untersuchten Methoden für die Züchtung bringen können. Die Vorteile der Umsetzung der untersuchten Methoden werden mittelfristig (in 5 bis 10 Jahren) sichtbar werden, wenn unter Verwendung dieser methodischen Verbesserungen neue, überlegene Sorten, schneller auf den Markt kommen als bisher.

Die Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen lässt sich leicht umsetzen und verspricht eine deutliche Erleichterung bei der Produktion von Inzuchtlinien. Der entwickelte Ansatz wurde kosteneffizient konzipiert, um seine Umsetzung zu erleichtern. Die Anwender sollten ihre konkrete Balance zwischen der Genauigkeit der F-Schätzung (d. h. der Klassifizierungsleistung) und den Kosten der Genotypisierung bestimmen. Unserer Ansicht nach stellt die Verwendung von 50 nahezu idealen Markern (also mit Allelfrequenzen nahe p=0,5) einen solchen Kompromiss zwischen Genauigkeit und Kosten dar, obwohl das optimale Gleichgewicht letztlich von den spezifischen Zielen und Ressourcenbeschränkungen der Anwender abhängt. Die Genauigkeit der F-Schätzung wird maximiert, wenn der Ansatz in Kombination mit rekurrenter genomischer Vorhersage verwendet wird, weil für die genomische Vorhersage „sowieso“ mit großer Marker-Anzahl gearbeitet wird.

Die Anwendung der genomischen Vorhersage hängt größtenteils vom verfügbaren Budget ab, da sie nach wie vor kostspielig ist (große Anzahl Marker), insbesondere für kleine Zuchtprogramme. Unsere Ergebnisse deuten darauf hin, dass eine sehr regelmäßige und prompte, unverzügliche Aktualisierung des Trainingsdatensatzes erforderlich ist, um relativ hohe und nutzbare Vorhersagefähigkeiten aufrechtzuerhalten, was die Kosten im Vergleich zur Anwendung der Methode eher erhöht. Ansonsten bestätigte sich der erwartete Zusammenhang zwischen Erblichkeit und Predictive Ability auch bei Ackerbohnen. Für die Berechnung der Accuracy muss die Erblichkeit der Feldversuche genommen werden, die zur Validation genutzt werden (nicht die der Trainingsdaten - es sei denn es handelt sich um Kreuzvalidierung, also Cross-Validation; das naive Moment der Cross-Validation wurde in diesem Bericht deutlich herausgestellt). Nach unseren Vorarbeiten fehlt nun für die Umsetzung des zerstörungsfreien Seed-Chippings eine Hochdurchsatz-Chipping-Plattform, um in Kombination mit der genomischen Vorhersage das Potenzial dieses Ansatzes zu erschließen.

Die acht ertragreichsten Genotypen des Abo-Direkt-Feldversuchs BIG2025 sind hier im Anhang gelistet. Unser Material und Daten stehen auf Nachfrage zur

Verfügung. Die berühmte Göttinger Winterackerbohnenpopulation wurde schon im September 2024 zur Erhaltung der Genbank Gatersleben übergeben (Kontakt: IPK; Nils Stein, Ulrike Lohwasser).

7. Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen; Hinweise auf weiterführende Fragestellungen

In der GWBP wurden wie theoretisch erwartet Inzuchtindividuen gefunden. Simulationen der für die Inzucht relevanten Zusammensetzung von Ackerbohnenpopulationen entsprachen sehr gut den Beobachtungen in dieser Versuchspopulation. Das erste Teilziel wurde erreicht. Die genomische Vorhersage wurde anhand historischer Feld- und Genotypdaten von 185 Inzuchtlinien kalibriert. Darüber hinaus wurde ein GP-Modell aus den Daten eins AD-Satzes von Genotypen entwickelt und eingesetzt. Feldversuche, die darauf abzielten, die Vorhersagefähigkeiten der genomischen Vorhersagemodelle zu bewerten, führten zur Identifizierung von Genotypen mit genetisch höherem Ertrag, Resistenz, Standfestigkeit. Das zweite Teilziel wurde erfüllt, mit der Ausnahme, dass eine Züchtung auf Trockenheits- und Frosttoleranz nicht möglich war, da solche Belastungen während der Feldversuche nicht auftraten. Einige Genotypen mit höherer Frosttoleranz und genomische Grundlagen für deren Züchtung sind in unserer Abteilung allerdings als Ergebnis des 2023 beendeten Profaba-Projektes vorhanden (DOI 10.21203/rs.3.rs-5356723/v1; ediss.uni-goettingen.de/handle/11858/16041; www.biorxiv.org/content/10.1101/2024.11.25.625124v11;). Zur Trockenstress-Toleranz wurde in Göttingen bis 2014 (Ali et al., 2026) und nun wieder seit 2025 gearbeitet (Team Brünjes; https://s.gwdg.de/6AGJPE).

Die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen erwies sich als durchführbar. Das dritte Teilziel wurde erreicht und auf die Optimierung der verwendeten Chipping-Technik ausgeweitet. Die konzeptionelle Umsetzung der untersuchten Methoden in der praktischen Züchtung in Anlehnung an die zweiteilige (‚two-part') Strategie von Gaynor et al. (2017) wurde untersucht. Das vierte Teilziel wurde erreicht. Es wurden theoretische Untersuchungen zur genetischen Zusammensetzung von Populationen unter partieller Allogamie inclusive LD-Abbau sowie zur Bestimmung individueller Inzuchtgrade und zur Optimierung des neuen züchterischen Methodik-Konzepts unter Nutzung natürlicher partieller Allogamie durchgeführt, da diese für die Anwendung unseres kosteneffizienten Ansatzes (zur Identifizierung von Inzuchtlinien aus Populationen) von großer Bedeutung waren. Das fünfte Teilziel wurde erreicht.

Faba-Bohnen-Linien mit verbesserter phänotypischer Ausprägung für die Merkmale, die in unseren Feldversuchen nicht bewertet werden konnten, lassen sich trotz dieser Lücke mit den im Projekt untersuchten Ansätzen schneller, kostengünstiger und weniger aufwendig züchten als bislang.

Zukünftige Forschung sollte darauf abzielen, das Potenzial der genomischen Vorhersage in der Ackerbohnenzüchtung weiter zu untersuchen. Unsere Studie verwendet kleine Trainings- und Validierungssätze (<200 Genotypen), und es wäre

daher von großem Interesse, den Gewinn an Vorhersagegenauigkeit mit einem größeren Satz (zum Beispiel 1000 Genotypen; wie von Resende 2024 empfohlen) zu bewerten. Der Einfluss der genetischen Architektur der Merkmale auf die Vorhersagefähigkeiten verschiedener GP-Modelle sollte gründlich analysiert werden (Daetwyler et al. 2013); grundsätzlich stehen dafür inzwischen Grunddaten zur Verfügung (Nature, 615(7953), 652–659; Theoretical and Applied Genetics, 136 (5), 114).

Dies würde es den Züchtern ermöglichen, merkmalsspezifische GP-Modelle anzuwenden, um die Genauigkeit der Vorhersage zu erhöhen. Marker, die durch genomweite Assoziationsstudien (GWAS) entdeckt und validiert wurden, können als Ko-Variable in die GP-Modelle integriert werden. Es zeigte sich, dass dieser Ansatz die Vorhersagefähigkeiten bei Merkmalen der Frosttoleranz deutlich erhöht (Zhang et al. 2026). Da auch unsere Gruppe über solche Daten verfügt (Windhorst 2025), bietet sich hier eine Möglichkeit für Folgeuntersuchungen. Zukünftige Studien sollten zudem der Bewertung der Vorhersagegenauigkeit in stark strukturierten Populationen Vorrang einräumen, da diese die praktische Anwendung der genomischen Vorhersage in vielen Pflanzenzüchtungsprogrammen besser widerspiegeln.

Das zerstörungsfreie Chipping von Samen (und die anschließende Genotypisierung) erwies sich in unserem vorgeschlagenen Schema und Pipeline als eine Technik zur Beschleunigung und Erleichterung der Züchtung. Weitere Forschungen könnten zusätzliche innovative Chipping-Techniken untersuchen, um die Praktikabilität und Effizienz des Ansatzes zu verbessern. Weitere Ansätze könnten in Betracht gezogen werden, beispielsweise das Chipping unreifer Samen, falls aufgrund anderer Entwicklungen das Speed-Breeding und schnelle Generationszyklen für die Ackerbohne (hier spezifisch für die Winterackerbohne) umgesetzt werden kann. Die Keimfähigkeit der Samen wurde durch das Chipping nicht beeinträchtigt, allerdings wurden die Samen in unserem Experiment unmittelbar nach dem Chipping zur Probe ausgesät. Der Einfluss einer Saatgutlagerung auf die Keimfähigkeit könnte untersucht werden. Die Entwicklung eines Hochdurchsatzes für das Chipping von Ackerbohnensaatgut würde die praktische Anwendung und großflächige Umsetzung der Technik erheblich verbessern. Darüber hinaus könnten Fortschritte beim Chipping von Saatgut anderer Kulturen helfen, insbesondere bei Körnerleguminosen wie Erbsen. Idealerweise sollte ein Hochdurchsatzsystem so konzipiert sein, dass es für mehrere Arten geeignet ist.

8. Zusammenfassung

Die Ackerbohne (*Vicia faba* L.) ist teilweise fremdbefruchtend (das Erntegut einer Pflanze oder auch einer Hülse besteht sowohl aus Selbstungskörnern als auch – sofern die Bestäuber Zutritt hatten und Genotypen mit Fremdpollen in der Nähe zur Verfügung standen - aus Samen aus Fremdbefruchtung), was ihre Züchtung zu einem aufwendigen Prozess macht, insbesondere bei der Erzeugung von Inzucht-linien. Durch diesen Umstand sind Züchter gezwungen, das Material in Isolierhäusern (um Fremdbefruchtung zu verhindern) durch eine Reihe aufeinanderfolgender Generationen kontrollierte Selbstung zu führen. Zudem sollten die Blüten der Pflanze

manuell zur Selbstung angeregt werden - das sog. Trippen - um auch bei Genotypen mit geringer sog. Autofertilität einen zufriedenstellenden Samenanfall zu gewährleisten. Da in der Züchtung stets neue Genotypen entstehen und autosterile Typen häufig sind, kann das Unterlassen des Trippens zu problematisch starker Variation im Samenansatz des Zuchtmaterials führen.

Diese teilweise Fremdbefruchtung bietet jedoch auch eine außergewöhnliche Chance. In offen bestäubten Populationen führt sie zu einer besonderen genetischen Zusammensetzung, bei der sich die Pflanzen der offen bestäubten Population hinsichtlich ihres Inzuchtkoeffizienten unterscheiden. Es sind in solchen Populationen stets *a priori* einige hoch ingezüchtete Individuen vorhanden, die aus mehreren aufeinanderfolgenden Generationen der Selbstbefruchtung resultieren, und diese können direkt in der Züchtung genutzt werden (als Quelle für eine Linie, also für einen Liniensorten-Kandidat; oder als Komponente eines Synthetiks). Diese hoch ingezüchteten Individuen lassen sich zwar nicht phänotypisch, aber durch Genotypisierung und die Schätzung ihres Inzuchtkoeffizienten identifizieren und von den anderen Individuen unterscheiden. Mit hochwertigen genotypischen Daten können für Individuen auch (Genomische Vorhersage) deren agronomischer Wert vorhergesagt werden, als Grundlage für eine Vor-Auslese. Durch die zerstörungsfreie DNA-Extraktion aus Ackerbohnensamen (Seed Chipping) kann die genomische Vorhersage auf Genotypen angewendet werden während diese erst nur als ein Same pro Genotyp vorliegen (typischerweise in Generation F2), also noch bevor diese ausgesät werden (Zeit- und Geldersparnis).

Unser erstes Ziel war es, einen Ansatz zu entwickeln, um hoch ingezüchtete Individuen aus offen bestäubten und genetisch vielfältigen Ackerbohnenpopulationen zu identifizieren. Wir schlagen vor, hoch ingezüchtete Individuen anhand einer begrenzten Anzahl von KASP-Markern zu identifizieren, um so eine kostengünstige Umsetzung in der Ackerbohnenzüchtung darzustellen. Durch umfangreiche Simulationen stellten wir fest, dass der genaueste algebraische Schätzer für den Inzuchtkoeffizienten beim Einsatz einer begrenzten Anzahl von Markern auf einer Likelihood-Methode basierte (der sogenannte Likelihood-Schätzer FL). In der Göttinger Winterbohnenpopulation fanden wir 61 hoch ingezüchtete Individuen unter 728 zufällig ausgewählten Pflanzen. Die Identifizierung hoch ingezüchteter Individuen erfolgte am kosteneffizientesten, indem die Individuen vorab anhand von 20 bis 30 KASP-Markern klassifiziert wurden, bevor eine Hauptklassifizierung mit den verbleibenden Individuen durchgeführt wurde. Die Hauptklassifizierung mit 50 Markern erwies sich als guter Kompromiss zwischen Kosten und Klassifizierungsleistung. Anhand von Simulationen und empirischen Daten wurden den Züchterinnen Empfehlungen zum Schwellenwert des Inzuchtkoeffizienten gegeben, der zur Deklarierung von Individuen als hoch ingezüchtet herangezogen werden sollte, abhängig von der Genauigkeit der Inzuchtkoeffizientenschätzung (also abhängig von der Anzahl der verwendeten Marker und deren Allelfrequenz) und auch von der inzuchtmäßigen Zusammensetzung der untersuchten Population (definiert durch ihren Fremd- bzw. Selbstbefruchtungsgrad).

Unser zweites Ziel war es, die Leistungsfähigkeit der genomischen Vorhersage (GP) innerhalb (Kreuzvalidierung) und zwischen (‚harte' Validierung) zwei Genotypensätzen zu bewerten. Beide Sätze stammen aus der Göttinger Winterbohnenpopulation (GWBP; unsere Referenz-Population). Jedoch wurde der erste Satz (A-Set) aus einer früheren Generation dieser Population entnommen als der zweite Satz (AD-Set9. Der im Abstand betrug 24 Jahre (24 Erhaltungsgeneration der GWBP).

Das Trainieren eines Modells (Algorithmus) für die GP mit Daten des einen Genotyp-Satzes und daraus dann die Vorhersage der Leistung des anderen Genotyp-Satzes (Validierung) nennen wir harte Validierung oder Generations-übergreifende Vorhersage, weil nicht nur die agronomische Leistung ‚anderer' Genotypen an ‚anderen' Umwelten gefragt ist, sondern weil sich diese anderen Genotypen auch Gruppe populationsgenetisch von der ‚einen' Gruppe unterscheiden. Dieses harte, praxisnaher Testen steht in gewissem Kontrast zu den etwas blauäugigen Resultaten des oft genutzten Kreuzvalidierens.

In diesem Sinne generationsübergreifende genomische Vorhersagen werden durch Selektion, Drift und Rekombinationen beeinflusst, die zu Unterschieden im Laufe der Zeit in Populationen auftreten. Im Vergleich zur Vorhersage innerhalb einer Generation war es unser Ziel, den Einfluss dieser Effekte auf die Vorhersagefähigkeiten zu quantifizieren. Der alte Datensatz bestand aus 185 Inzuchtlinien (abgeleitet aus Einzelkornabkommen), die in 18 Feldversuchen phänotypisiert wurden. Der neue Datensatz bestand aus 489 Genotypen (vermehrt durch aufeinanderfolgende Generationen kontrollierter Selbstbefruchtung), die in vier Feldversuchen phänotypisiert wurden, die in den Jahren 2024 und 2025 durchgeführt wurden. Insbesondere verfügten 143 Genotypen des neuen Datensatzes über hochwertige phänotypische Daten von drei Standorten und 200-Pflanzen-Parzellen mit mehreren Wiederholungen. Die genomischen Vorhersagefähigkeiten innerhalb einer Generation erwiesen sich als relativ hoch. Im Gegensatz dazu waren die genomischen Vorhersagefähigkeiten über Generationen hinweg relativ gering und erreichten bei der Vorhersage der 143 Genotypen mit hochwertigen phänotypischen Daten einen Höchstwert von 0,30 für das Tausendkorngewicht. Anhand genotypischer Daten konnten wir eine erhebliche genomische Differenzierung zwischen dem alten und dem neuen Datensatz feststellen. Der durchschnittliche Verlust an Vorhersagegenauigkeit betrug 0,23 für agronomische Merkmale (0,32 für den Ertrag) und 0,34 für phänologische Merkmale. Es wurden Strategien besprochen, um die genomische Vorhersage bei generationenübergreifenden Vorhersagen im Falle geringer Genauigkeiten zu nutzen.

Wir haben drei Chipping-Techniken getestet, die zur zerstörungsfreien Probenentnahme aus Ackerbohnensamen für die DNA-Extraktion und die anschließende KASP-Analyse eingesetzt werden können. Wir stellten fest, dass Proben aus Ackerbohnensamen fünf zusätzliche PCR-Zyklen benötigten, um eine korrekte Clusterbildung der Proben in den für die KASP-Analyse typischen Allel-Diskriminierungsdiagrammen zu ermöglichen. Es war nicht möglich, das Chippen der Samen zu vereinfachen, indem die trockenen Samen durch Einweichen in Wasser aufgeweicht wurden, da sich herausstellte, dass die Einweichdauer stark vom jeweiligen Samen abhängt und die Keimfähigkeit der Samen beeinträchtigt. Die effizienteste Technik war das Anbohren trockener Samen, was eine schnelle und konsistente Probenahme ermöglichte. Wir konnten zudem nachweisen, dass die Keimfähigkeit der angebohrten Samen nicht beeinträchtigt wurde.

Zusammenfassend lässt sich sagen, dass mittels des Abo-Direkt-Projektes mehrere methodischer Verbesserungen für die Ackerbohnenzüchtung entwickelt wurden, wobei der Schwerpunkt vor allem darauf lag, Züchtern einen einfachen, schnellen und kostengünstigen Zugang zu überlegenen Inzuchtgenotypen der Ackerbohne zu ermöglichen und damit letztlich den genetischen Fortschritt zu beschleunigen.

9. Verzeichnis der benutzten Fachliteratur

Abdelmula, A., Link, W., Kittlitz, E., & Stelling, D. (1999). Heterosis and inheritance of drought tolerance in faba bean*, Vicia faba* L. Plant Breeding, 118, 458–490.

Abou Khater, L., Joukhadar, R., Maalouf, F., Alsamman, A. M., Babiker, Z., Balech, R., Hu, J., Ma, Y., Dunham, A., Sanchez, M., Jighly, A., & Kumar, S. (2025). Genomic prediction for heat and herbicide tolerance in faba bean. The Plant Genome, 18(3), e70076. https://doi.org/10.1002/tpg2.70076

Adhikari, K. N., Khazaei, H., Ghaouti, L., Maalouf, F., Vandenberg, A., Link, W., & O'Sullivan, D. M. (2021). Conventional and Molecular Breeding Tools for Accelerating Genetic Gain in Faba Bean (*Vicia Faba* L.). Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.744259

Almogdad, M., Jonavičienė, A., & Semaškienė, R. (2023). Bruchus rufimanus Boh. Effect on Broad Bean Seed Quality and the Infection Level of Seed-Borne Fungal Pathogens. Plants, 12(9), 1825. https://doi.org/10.3390/plants12091825

Annicchiarico, P., Nazzicari, N., Ferrari, B., Harzic, N., Carroni, A. M., Romani, M., & Pecetti, L. (2019). Genomic prediction of grain yield in contrasting environments for white lupin genetic resources. Molecular Breeding, 39(10), 142–142.

Bishnoi, S., & Panchta, R. (2012). Advances on heterosis and hybrid breeding in faba bean (*Vicia faba* L.). Forage Research, 38, 65–73.

Bond, D. A. (1982). The Development and Performance of Synthetic Varieties of *Vicia faba* L. In G. Hawtin & C. Webb (Eds), Faba Bean Improvement: Proceedings of the Faba Bean Conference held in Cairo, Egypt, March 7–11, 1981 (pp. 41–51). Springer Netherlands. https://doi.org/10.1007/978-94-009-7499-9_4

Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ramdoss, Y., & Buckler, E. (2007). TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. Bioinformatics (Oxford, England), 23, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brünjes, L., & Link, W. (2021). Paternal outcrossing success differs among faba bean genotypes and impacts breeding of synthetic cultivars. Theoretical and Applied Genetics, 134(8), 2411–2427. https://doi.org/10.1007/s00122-021-03832-z

Brünjes, L., & Link, W. (2025). Assessment of Pollen Production, Pollen Viability, Autofertility and Paternal Success in Faba Bean (*Vicia faba* L.) for Better Exploitation of Grain Yield Heterosis. Plant Breeding. https://doi.org/10.1111/pbr.70033

Budhlakoti, N., Kushwaha, A. K., Rai, A., Chaturvedi, K. K., Kumar, A., Pradhan, A. K., Kumar, U., Kumar, R. R., Juliana, P., Mishra, D. C., & Kumar, S. (2022). Genomic Selection: A Tool for Accelerating the Efficiency of

Molecular Breeding for Development of Climate-Resilient Crops. Frontiers in Genetics, 13. https://doi.org/10.3389/fgene.2022.832153

Caballero, A., Fernández, A., Villanueva, B., & Toro, M. A. (2022). A comparison of marker-based estimators of inbreeding and inbreeding depression. Genetics Selection Evolution, 54(1), 82. https://doi.org/10.1186/s12711-022-00772-0

Cockerham, C. C., & Weir, B. S. (1984). Covariances of relatives stemming from a population undergoing mixed self and random mating. Biometrics, 40(1), 157–164.

Daetwyler, H. D., Calus, M. P. L., Pong-Wong, R., de Los Campos, G., & Hickey, J. M. (2013). Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. Genetics, 193(2), 347–365. https://doi.org/10.1534/genetics.112.147983

Duc, G., Bao, S., Baum, M., Redden, B., Sadiki, M., Suso, M. J., Vishniakova, M., & Zong, X. (2010). Diversity maintenance and use of *Vicia faba* L. genetic resources. Field Crops Research, Faba Beans in Sustainable Agriculture, 115(3), 270–278. https://doi.org/10.1016/j.fcr.2008.10.003

Duc, G., Marget, P., Esnault, R., Guen, J., & Bastianelli, D. (1999). Genetic variability for feeding value of faba bean seeds (*Vicia faba*): Comparative chemical composition of isogenics involving zero-tannin and zero-vicine genes. The Journal of Agricultural Science, 133, 185–196. https://doi.org/10.1017/S0021859699006905

Ebmeyer, E. (1988). Heterosis and Genetic Variances and their Implications for Breeding Improved Varieties of Spring Beans (*Vicia faba* L.). Plant Breeding, 101(3), 200–207. https://doi.org/10.1111/j.1439-0523.1988.tb00288.x

FAOSTAT. (2026). https://www.fao.org/faostat/en/#home

Fleck, A., & Ruckenbauer, P. (1989). Der Polycrosstest als methodischer Schritt in der Fababohnenzüchtung (experimentelle Ergebnisse). Der Polycrosstest Als Methodischer Schritt in Der Fababohnenzüchtung (Experimentelle Ergebnisse), 40(1), 61–72.

Flores, F., Nadal, S., Solis, I., Winkler, J., Sass, O., Stoddard, F., Link, W., Raffiot, B., Muel, F., & Rubiales, D. (2012). Faba bean adaptation to autumn sowing under European climates. Agronomy for Sustainable Development, 32, 727–734. https://doi.org/10.1007/s13593-012-0082-0

Gallais, A. (1992). Pourquoi faire des variétés synthétiques ? Agronomie, 12(8), 601–609. https://doi.org/10.1051/agro:19920803

Gasim, S. (2003). Winter-hardy faba bean: Applied genetic research on the reproductive mode of the European gene pool [doctoral Thesis; University of Göttingen]. https://cuvillier.de/de/shop/publications/3382-winter-hardy-faba-bean-applied-genetic-research-on-the-reproductive-mode-of-the-european-gene-pool

Gasim, S., Abdelmula, A., & Khalifa, J. (2011). Analysis of the degree of cross-fertilization and auto- fertility and their impact on breeding faba bean (*Vicia faba* L.) cultivars grown in Sudan. African Journal of Agricultural Research Vol. 6(30), Pp. 6387-6390., 6, 6387–6390.

Gasim, S., Abel, S., & Link, W. (2002). Ein Beitrag zur Züchtungsforschung an Winterbohnen (*Vicia faba* L.). Vorträge Für Pflanzenzüchtung, 469–472.

Gasim, S., Abel, S., & Link, W. (2004a). Extent, variation and breeding impact of natural cross-fertilization in German winter faba beans using hilum colour as marker. Euphytica, 136(2), 193–200. https://doi.org/10.1023/B:EUPH.0000030672.56206.f0

Gasim, S., Abel, S., & Link, W. (2004b). Extent, variation and breeding impact of natural cross-fertilization in German winter faba beans using hilum colour as marker. Euphytica, 136(2), 193–200. https://doi.org/10.1023/B:EUPH.0000030672.56206.f0

Gaynor, R. C., Gorjanc, G., Bentley, A. R., Ober, E. S., Howell, P., Jackson, R., Mackay, I. J., & Hickey, J. M. (2017). A Two-Part Strategy for Using Genomic Selection to Develop Inbred Lines. Crop Science, 57(5), 2372–2386. https://doi.org/10.2135/cropsci2016.09.0742

Gebremedhin, A., Li, Y., Shunmugam, A. S. K., Sudheesh, S., Valipour-Kahrood, H., Hayden, M. J., Rosewarne, G. M., & Kaur, S. (2024). Genomic selection for target traits in the Australian lentil breeding program. Frontiers in Plant Science, 14. https://doi.org/10.3389/fpls.2023.1284781

Ghaouti, L. (2007). Comparison of pure line cultivars with synthetic cultivars in local breeding of faba bean (*Vicia faba* L.) for organic farming [doctoral Thesis; University of Göttingen]. https://doi.org/10.53846/goediss-3609

GRDC. (2017). Faba bean plant growth and physiology. GRDC Grownotes.

GRDC. (2018). Faba bean storage. GRDC Grownotes.

Gu, Z. (2022). Complex heatmap visualization. iMeta, 1(3), e43. https://doi.org/10.1002/imt2.43

Haile, T. A., Heidecker, T., Wright, D., Neupane, S., Ramsay, L., Vandenberg, A., & Bett, K. E. (2020). Genomic selection for lentil breeding: Empirical evidence. The Plant Genome, 13(1), e20002. https://doi.org/10.1002/tpg2.20002

Hale, B., Ferrie, A. M. R., Chellamma, S., Samuel, J. P., & Phillips, G. C. (2022). Androgenesis-Based Doubled Haploidy: Past, Present, and Future Perspectives. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.751230

Hall, N., Mercer, L., Phillips, D., Shaw, J., & Anderson, A. D. (2012). Maximum likelihood estimation of individual inbreeding coefficients and null allele frequencies. Genetics Research, 94(3), 151–161. https://doi.org/10.1017/S0016672312000341

Hillel, J., Schaap, T., Haberfeld, A., Jeffreys, A. J., Plotzky, Y., Cahaner, A., & Lavi, U. (1990). DNA fingerprints applied to gene introgression in breeding programs. Genetics, 124(3), 783–789. https://doi.org/10.1093/genetics/124.3.783

Holden, J. H. W., & Bond, D. A. (1960). Studies on the breeding system of the field bean, *Vicia faba* (L.). Heredity, 15(2), 175–192. https://doi.org/10.1038/hdy.1960.74

Jarquin, D., Specht, J., & Lorenz, A. (2016). Prospects of Genomic Prediction in the USDA Soybean Germplasm Collection: Historical Data Creates Robust Models for Enhancing Selection of Accessions. G3 Genes|Genomes|Genetics, 6(8), 2329–2341. https://doi.org/10.1534/g3.116.031443

Jayakodi, M., Golicz, A. A., Kreplak, J., Fechete, L. I., Angra, D., Bednář, P., Bornhofen, E., Zhang, H., Boussageon, R., Kaur, S., Cheung, K., Čížková, J., Gundlach, H., Hallab, A., Imbert, B., Keeble-Gagnère, G., Koblížková, A., Kobrlová, L., Krejčí, P., … Andersen, S. U. (2023). The giant diploid faba genome unlocks variation in a global protein crop. Nature, 615(7953), 652–659. https://doi.org/10.1038/s41586-023-05791-5

Kamiya, M., & Kiguchi, T. (2003). Rapid DNA Extraction Method from Soybean Seeds. Breeding Science, 53(3), 277–279. https://doi.org/10.1270/jsbbs.53.277

Karkanis, A., Ntatsi, G., Lepse, L., Fernández, J. A., Vågen, I. M., Rewald, B., Alsiņa, I., Kronberga, A., Balliu, A., Olle, M., Bodner, G., Dubova, L., Rosa, E., & Savvas, D. (2018). Faba Bean Cultivation – Revealing Novel Managing Practices for More Sustainable and Competitive European Cropping Systems. Frontiers in Plant Science, 9. https://doi.org/10.3389/fpls.2018.01115

Karlin, S., 1968: Equilibrium behaviour of population genetic models with non-random mating. Part I and Part II. Vol. 5, No. 2, pp. 231- 313 und Vol . 5(3): 487-566. doi:10.2307/3211920

Kelly, J. K., & Williamson, S. (2000). Predicting Response to Selection on a Quantitative Trait: A Comparison between Models for Mixed-mating Populations. Journal of Theoretical Biology, 207(1), 37–56. https://doi.org/10.1006/jtbi.2000.2154

Laugel, H., & Link, W. (2025). Faba bean populations already contain the inbreds needed for breeding. In Review. https://doi.org/10.21203/rs.3.rs-8211981/v1

Laugel, H., 2026. Novel breeding approaches for direct access to superior inbred lines in faba bean (Vicia faba L.). Dissertation, Georg-August-Universität Göttingen. Siehe https://ediss.uni-goettingen.de/.

Li, C. C., & Horvitz, D. G. (1953). Some methods of estimating the inbreeding coefficient. American Journal of Human Genetics, 5(2), 107–117.

Link, W., & Arbaoui, M. (2006). Neues von der Göttingen Winter-Ackerbohne (pp. 31–38) [Bericht über die 56. Tagung 2005 der Vereinigung der Pflanzenzüchter und Saatgutkaufl eute Österreichs].

Link, W., & Ederer, W. (1993). The Concept of Varietal Ability for Partially Allogamous Crops. Plant Breeding, 110(1), 1–8. https://doi.org/10.1111/j.1439-0523.1993.tb00561.x

Link, W., Ederer, W., & von Kittlitz, E. (1994). Zuchtmethodische Entwicklungen: Nutzung von Heterosis bei Fababohnen. Vorträge Für Pflanzenzüchtung, 30, 201–230.

Lippolis, A., Gezan, S. A., Zuidgeest, J., Cafaro, V., van Dinter, B.-J., Elzes, G., Paulo, M.-J., & Trindade, L. M. (2025). Targeted genotyping (90K-SPET)

facilitates genome-wide association studies and the prediction of yield-related traits in faba bean (*Vicia faba* L.). BMC Plant Biology, 25(1), 558. https://doi.org/10.1186/s12870-025-06546-2

Lübberstedt, T., Beavis, W., & Suza, W. (2023). Chapter 7: Marker Assisted Selection and Genomic Selection. https://iastate.pressbooks.pub/molecular-plantbreeding/chapter/marker-assisted-selection-and-genomic-selection/

Marzinzig, B., Brünjes, L., Biagioni, S., Behling, H., Link, W., & Westphal, C. (2018). Bee pollinators of faba bean (*Vicia faba* L.) differ in their foraging behaviour and pollination efficiency. Agriculture, Ecosystems & Environment, 264, 24–33. https://doi.org/10.1016/j.agee.2018.05.003

McVetty, P. B. E., & Nugent-Rigby, J. (1984). NATURAL CROSS POLLINATION OF FABA BEANS (*Vicia faba* L.) GROWN IN MANITOBA. Canadian Journal of Plant Science, 64(1), 43–46. https://doi.org/10.4141/cjps84-007

Metz, P. L. J., Buiel, A. a. M., Norel, A. van, & Helsper, J. P. F. G. (1995). Genetic factors controlling outcrossing in faba bean (*Vicia faba*): Effects of pollen donor and receptor genotypes and estimates of inter-receptor cross-fertilization. The Journal of Agricultural Science, 122(2), 249–254.

Metz, P. L. J., Buiel, A. A. M., van Norel, A., & Helsper, J. P. F. G. (1992). Rate and inheritance of cross-fertilization in faba bean (*Vicia faba* L.). Euphytica, 66(1), 127–133. https://doi.org/10.1007/BF00023517

O'Sullivan DM, Angra D, Harvie T, Tagkouli V, & Warsame A. (2019). A genetic toolbox for *Vicia faba* improvement.

Osuna-Caballero, S., Rubiales, D., Annicchiarico, P., Nazzicari, N., & Rispail, N. (2024). Genomic prediction for rust resistance in pea. Frontiers in Plant Science, 15. https://doi.org/10.3389/fpls.2024.1429802

Poulsen, M. H. (1975). Pollination, seed setting, cross-fertilization and inbreeding in *Vicia faba* L. Zeitschrift Fur Pflanzenzuchtung, 74(2), 97–118.

Resende, R. T. (2024). Balancing genomic selection efforts for allogamous plant breeding programs. Journal of Crop Science and Biotechnology, 27(3), 349–358. https://doi.org/10.1007/s12892-024-00234-7

Sillero, J. C., Villegas-Fernández, A. M., Thomas, J., Rojas-Molina, M. M., Emeran, A. A., Fernández-Aparicio, M., & Rubiales, D. (2010). Faba bean breeding for disease resistance. Field Crops Research, 115(3), 297–307. https://doi.org/10.1016/j.fcr.2009.09.012

Stewart-Brown, B. B., Song, Q., Vaughn, J. N., & Li, Z. (2019). Genomic Selection for Yield and Seed Composition Traits Within an Applied Soybean Breeding Program. G3, 9(7), 2253–2265. https://doi.org/10.1534/g3.118.200917

Stoddard, F. L., Nicholas, A. H., Rubiales, D., Thomas, J., & Villegas-Fernández, A. M. (2010). Integrated pest management in faba bean. Field Crops Research, 115(3), 308–318. https://doi.org/10.1016/j.fcr.2009.07.002

Suso, M. J., Gilsanz, S., Duc, G., Marget, P., & Moreno, M. T. (2006). Germplasm Management of Faba Bean (*Vicia faba* L.): Monitoring Intercrossing between Accessions with Inter-plot Barriers. Genetic Resources and Crop Evolution, 53(7), 1427–1439. https://doi.org/10.1007/s10722-005-6844-7

Suso, M. J., Harder, L., Moreno, M. T., & Maalouf, F. (2005). New strategies for increasing heterozygosity in crops: *Vicia faba* mating system as a study case. Euphytica, 143(1), 51–65. https://doi.org/10.1007/s10681-005-2526-y

Suso, M. J., & Maalouf, F. (2010). Direct and correlated responses to upward and downward selection for outcrossing in *Vicia faba*. Field Crops Research, 116(1), 116–126. https://doi.org/10.1016/j.fcr.2009.12.001

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. Journal of Dairy Science, 91(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Von Post, R., Von Post, L., Dayteg, C., Nilsson, M., Forster, B. P., & Tuvesson, S. (2003). A high-throughput DNA extraction method for barley seed. Euphytica, 130(2), 255–260. https://doi.org/10.1023/A:1022863006134

Wang, J. (2011). A new likelihood estimator and its comparison with moment estimators of individual genome-wide diversity. Heredity, 107(5), 433–443. https://doi.org/10.1038/hdy.2011.30

Weir, B.S., R.W. Allard, and A. L. Kahler, 1972. Analysis of complex allozyme polymorphisms in a barley population. Genetics 72, 505-523

Wientjes, Y. C., Veerkamp, R. F., Bijma, P., Bovenhuis, H., Schrooten, C., & Calus, M. P. (2015). Empirical and deterministic accuracies of across-population genomic prediction. Genetics, Selection, Evolution : GSE, 47(1), 5. https://doi.org/10.1186/s12711-014-0086-0

Windhorst, A. (2025). Winter Hardiness and Freezing Tolerance of Winter Faba Bean (*Vicia faba* L.): From QTL Identification and Validation to Genome-Wide Predictions of Genotypic Values [doctoral Thesis; University of Göttingen]. https://doi.org/10.53846/goediss-11296

Windhorst, A., Skovbjerg, C. K., Angra, D., O'Sullivan, D. M., Andersen, S. U., & Wolfgang, L. (2024). Genome-wide association studies identify promising QTL for freezing tolerance in winter and early spring as a basis for in-depth genetic analysis and implementation in winter faba bean (*Vicia faba* L.) breeding (p. 2024.11.25.625124). bioRxiv. https://doi.org/10.1101/2024.11.25.625124

Wright, A. J. (1973). The selection of parents for synthetic varieties of outbreeding diploid crops. TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik, 43(2), 79–82. https://doi.org/10.1007/BF00274961

Wright, A. J. (1977). Inbreeding in synthetic varieties of field beans (*Vicia faba* L.). The Journal of Agricultural Science, 89(2), 495–501. https://doi.org/10.1017/S0021859600028410

Wright, S. (1921). Systems of mating. I. The biometric relations between parent and offspring. Genetics, 6(2), 111–123. https://doi.org/10.1093/genetics/6.2.111

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. Nature Genetics, 42(7), 565–569. https://doi.org/10.1038/ng.608

Zeid, M., Schön, C.-C., & Link, W. (2004). Hybrid performance and AFLP- based genetic similarity in faba bean. Euphytica, 139(3), 207–216. https://doi.org/10.1007/s10681-004-3156-5

Zhang, Q. S., Goudet, J., & Weir, B. S. (2022). Rank-invariant estimation of inbreeding coefficients. Heredity, 128(1), 1–10. https://doi.org/10.1038/s41437-021-00471-4

10. Übersicht über alle im Berichtszeitraum vom Projektnehmer realisierten Veröffentlichungen zum Projekt (Printmedien, Newsletter usw.), bisherige und geplante Aktivitäten zur Verbreitung der Ergebnisse

Im Rahmen des Projekts „Abo-Direkt“ wurden zwei Informationsblätter mit den Titeln „Direkter Zugang zu Inzuchtlinien der Ackerbohne“ und „Vergleich der Vorhersagekraft eines GBLUP-Modells vor und nach 23 Generationen der Erhaltung der Göttinger Winterackerbohnenpopulation“ erstellt. Beide ‚Factsheets‘ richten sich an Züchterinnen von Ackerbohnen. Wie die jeweiligen Titel vermuten lassen, enthält das erste Factsheet Empfehlungen zur Identifizierung von Inzuchtlinien aus Ackerbohnenpopulationen. Das zweite konzentriert sich auf die Anwendung der genomischen Vorhersage in der Ackerbohnenzüchtung und hebt insbesondere die Bedeutung der Aktualisierung des als Trainingsdaten verwendeten Genotypensatzes hervor, um einen generationsbedingten Rückgang der Vorhersagefähigkeiten zu vermeiden.

Das Projekt wurde der wissenschaftlichen Gemeinschaft erstmals als Poster auf der Plant Science Student Conference (Gatersleben, Deutschland; 2023; ausgezeichnet mit dem „Best Poster“-Preis; DOI 10.5281/zenodo.8154322) vorgestellt; dazu auf der International Legume Society Conference (Granada, Spanien; 2023) sowie auf dem 5. CiBreed-Herbstworkshop (Göttingen; 2023). Die Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen wurde erwähnt in „Brünjes L., Ecke W., Windhorst A., Laugel H., Link W. (2023). Faba bean breeding for global mandate: Status and future progress. Legume Perspectives: Faba Beans Take the Spotlight (Ausgabe 24)“. Die historischen Felddaten, die zur Kalibrierung des genomischen Vorhersagemodells verwendet wurden, wurden auf der deutschen Pflanzenzüchtungskonferenz „Accelerating crop genetic gain“ (Geisenheim, Deutschland; 2024; ausgezeichnet mit dem „Best Poster“-Preis) vorgestellt. Die Ergebnisse zum kosteneffizienten Ansatz zur Identifizierung von Inzuchtlinien aus Populationen wurden auf der GPZ AG4-Konferenz: Genomanalyse (Kiel, Deutschland; 2025) und auf dem 6. CiBreed-Herbstworkshop (Göttingen; 2025) vorgestellt. Derselbe Ansatz wurde zusammen mit vorläufigen Ergebnissen der genomischen Vorhersage auf der GFPi-Jahreskonferenz (Bonn, Deutschland; 2025) präsentiert und diskutiert. Die Ergebnisse der genomischen Vorhersage wurden auf dem Ackerbohnen-Workshop (organisiert von Prof. Dr. Andreas Stahl; Julius-Kühn-Institut; Quedlinburg; 2025) vorgestellt, in Verbindung mit unseren Ergebnissen zum zerstörungsfreien Chippen von Ackerbohnensamen für die anschließende DNA-Extraktion und Genotypisierung. Siehe dazu auch https://s.gwdg.de/Ynwuwh; und https://s.gwdg.de/j7SfZe.

Werner Vogt-Kaute führte in seiner Funktion für das LEGUNET (www.legunet.de) den Wissens- und Materialtransfer an dieses Netzwerk durch. Im Jahr 2025 wurde in Göttingen ein Feldbesuch mit dem Netzwerk organisiert, bei dem der Hauptfeldversuch des Projekts, seine Ziele und vorläufigen Ergebnisse vorgestellt wurden. Im selben Jahr wurde ein weiterer Feldbesuch organisiert, um den Naturland-Biobauern die Wiederholung desselben Feldversuchs in Wartmannsroth (Bioland) vorzustellen.

Die Studie zur Identifizierung hoch-ingezüchteter Individuen aus Ackerbohnenpopulationen wurde kürzlich zur Veröffentlichung in der Zeitschrift Theoretical and Applied Genetics (Impact-Faktor: 4,2; 2024) angenommen und ist als Preprint verfügbar: Laugel, H., und W. Link, 2025: Faba bean populations already contain the inbreds needed for breeding. DOI 10.21203/rs.3.rs-8211981/v1

Wir planen, die Ergebnisse unserer Studie zur Leistungsfähigkeit der genomischen Vorhersage innerhalb und über Generationen hinweg im Rahmen des GWBP sehr bald in einem einschlägigen wissenschaftlichen Artikel zu veröffentlichen.

Die in Abo-Direkt entwickelten ertragreichsten Genotypen sowie ihr Kornertrag und weitere Merkmale. Weitere Daten zu weiteren Genotypen sowie aus den weniger intensiven Versuchen (WEE2024, WEE2024+, WEE2025) und dem HFT sind auf Anfrage erhältlich. F(S0) ist der Inzuchtkoeffizient des Ausgangs-Individuums. Bonituren von 1-9: 9 heißt hohes Lager bzw. starker Befall bzw. Späte Reife

Trial	Rang	Genotyp	F (S0)	Ertrag (dt/ha)	Blüh-beginn	Wuchs-höhe (Mai)	Lager (Juli)	Falscher Mehltau	Botrytis	Rost	Reife	TKWG (in g)	Protein-gehalt (in %)
BIG2025 (143 Prüfglieder plus Checks)	1	DPB20S0_**084**	0,00	52,9	118,6	95,8	3,5	3,4	2,0	4,2	7,5	485,4	25,7
	2	DPB20S0_071	0,11	51,1	118,6	107,2	4,4	3,3	3,4	4,9	4,6	483,1	20,8
	3	SSD22S0_087	0,86	49,3	118,5	90,1	3,3	5,2	3,2	4,7	7,8	425,5	21,5
	4	SSD22S0_043	0,91	48,8	119,7	99,6	3,4	5,2	3,7	6,2	5,4	444,3	22,2
	5	DPB20S0_082	0,00	47,6	126,8	98,9	4,0	4,3	4,1	4,0	4,6	431,2	22,0
	6	DPB20S0_243	0,85	47,2	127,9	81,6	3,2	3,6	2,6	6,8	5,5	442,6	23,6
	7	DPB20S0_410	0,87	46,8	119,2	105,5	2,3	2,7	2,7	6,4	5,8	431,5	23,1
	8	DPB20S0_010	0,05	46,5	119,6	91,1	4,1	4,3	4,8	7,1	3,8	451,9	23,4
	Kontroll-Genotypen												
	Nairobi			52,6	Diese vier Prüfglieder wiesen den höchsten Proteingehalt im Samen auf: DPB20S0_872; DPB20S0_025; SSD22S0_054; DPB20S0_076. Diese drei Prüfglieder wiesen die geringsten Peronospora-Symptome auf: DPB20S0_300; DPB20S0_157; DPB20S0_873. Diese fünf Prüfglieder wiesen die geringsten Botrytis-Symptome auf: DPB20S0_044; DPB20S0_**856**; DPB20S0_804; DPB20S0_084; DPB20S0_031. Beachte, dass DPB20S0_**084** beim Kornertrag den höchsten Rang einnahm (52,9 dt/ha). Diese beiden Prüfglieder wiesen die geringsten Rostsymptome auf: DPB20S0_**856**; DPB20S0_509. Beachte, dass DPB20S0_**856** auch bei Botrytis recht günstig abschnitt und beim Ertrag den 11. Platz belegte (45,7 dt/ha).							
	Patagonia			53,2									
	Curlew			60,3									
	GL Arabella			46,6									
	GL Alice			49,7									
	Nebraska			49,2									
	Mean of Checks			51,9									