

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Schlussbericht zum Thema

“Genomik-basierte Verbesserung des heimischen Sojazuchtmaterials und Etablierung eines molekularen Screeningsystems für Soja-Pathogene“

FKZ: 2814EPS011; 2814EPS012; 2814EPS013; 2814EPS014

Projektnehmer:

Universität Hohenheim; Julius Kühn-Institut

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

Abschlussbericht

Genomik-basierte Verbesserung des heimischen Sojazuchtmaterials und Etablierung eines molekularen Screeningsystems für Soja-Pathogene

Genomics-Based Improvement of Adapted Soybean Germplasm and Establishment of a Molecular Screening System for Soybean Pathogens

SojaGenoPath

Projektpartner

Universität Hohenheim

Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik (IPSP),
Förderkennzeichen 2814EPS012

Landessaatzuchtanstalt (LSA), Förderkennzeichen 2814EPS011

Institut für Phytomedizin (IPM), Förderkennzeichen 2814EPS013

Julius Kühn-Institut

Bundesforschungsinstitut für Kulturpflanzen, Institut für Resistenzforschung und
Stresstoleranz, Förderkennzeichen 2814EPS014

Berichtszeit: 1. April 2015 – 31. Dezember 2021

Koordinator Dr. Volker Hahn, LSA, Email: volker.hahn@uni-hohenheim.de

Gefördert durch



Bundesministerium
für Ernährung
und Landwirtschaft

Projektträger



Bundesanstalt für
Landwirtschaft und Ernährung

aufgrund eines Beschlusses
des Deutschen Bundestages

KONTAKTINFORMATIONEN

Universität Hohenheim

Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik (IPSP),
Prof. Karl Schmid, Fruwirthstr. 21, 70599 Stuttgart, karl.schmid@uni-hohenheim.de

Landessaatzuchtanstalt (LSA) Dr. Volker Hahn, Fruwirthstr. 21, 70599 Stuttgart,
volker.hahn@uni-hohenheim.de

Institut für Phytomedizin (IPM), Prof. Ralf T. Vögele, Otto-Sander-Str. 5, 70599 Stuttgart,
ralf.voegele@uni-hohenheim.de

Julius Kühn-Institut

Bundesforschungsinstitut für Kulturpflanzen, Institut für Resistenzforschung und
Stresstoleranz, Dr. Christiane Balko, Rudolf-Schick-Platz 3, 18190 Sanitz,
christiane.balko@julius-kuehn.de



KURZFASSUNG

Genomik-basierte Verbesserung des heimischen Sojazuchtmaterials und Etablierung eines molekularen Screeningsystems für Soja-Pathogene

Prof. Karl Schmid, Universität Hohenheim, Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik, karl.schmid@uni-hohenheim.de

Prof. Ralf T. Vögele, Universität Hohenheim, Institut für Phytomedizin, ralf.voegel@uni-hohenheim.de

Dr. Volker Hahn, Universität Hohenheim, Landessaatzuchtanstalt, volker.hahn@uni-hohenheim.de

Dr. Christiane Balko, Julius-Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, Institut für Resistenzforschung und Stresstoleranz, christiane.balko@julius-kuehn.de

Zur Einleitung von Pflanzenschutzmaßnahmen bei Soja ist es essentiell, Informationen über vorhandene Pathogene zu gewinnen. Basierend auf der quantitativen Real Time-PCR wurden Einzelnachweise für die Hauptpathogene der Sojabohne in Deutschland entwickelt. Die Pathogenität der Krankheitserreger wurde geprüft und das Inokulationsverfahren für weitere Untersuchungen etabliert. Erste Mehrfach-Nachweise für verschiedene Erreger und Probenotypen wurden entwickelt, die für eine frühzeitige Detektion von Erregern in Soja-Proben (Saatgut, Pflanze, Boden) verwendet werden können.

Im Rahmen des Projekts wurde zudem durch Kombination von aktuellen und zukünftigen Klimaparametern sowie Genotypisierungs- und Phänotypisierungsdaten eine Kernkollektion von Akzessionen erstellt, die sich durch eine hohe Diversität auszeichnet und sich für den Anbau unter den Bedingungen von Zentraleuropa eignet. Dabei wurden neue Gene für Umweltadaptation identifiziert, für die molekulare Marker für die Züchtung entwickelt werden, und geeignete Akzessionen für die Einkreuzung in aktuelle Züchtungsprogramme wurden selektiert.

Die Ergebnisse zeigen außerdem klar, dass hocheffiziente Linien gezüchtet werden können, um die Sojaanbauregion nach Norden zu erweitern. Zudem bieten die gewonnenen Varianzkomponenten, Heritabilitäten und Merkmalskorrelationen eine solide Grundlage für die Gestaltung von Zuchtprogrammen, insbesondere auch von Speed-Breeding Programmen, die zukünftig mit genomischer und phänotypischer Selektion beschleunigt werden können. Es wurden bereits erste Linien an die private Pflanzenzüchtung abgegeben, die in weiteren Prüfungen ermitteln, ob die Linien als Sorten zugelassen werden können.

Gerade in kühleren Lagen, wie z.B. in Norddeutschland, ist die Gefahr von Ertragseinbußen auf Grund von Kühlestress groß. In diesem Projekt wurden Linien von Kreuzungsnachkommenschaften und ein diverses Set aus Genbankakzessionen hinsichtlich ihres Hülsenansatzes nach einer Kühlestressphase charakterisiert. Das Testsystem führte zu reproduzierbaren Ergebnissen und es wurden erste QTL für den Hülsenansatz unter Stressbedingungen ermittelt.

ABSTRACT

Genomics-Based Improvement of Adapted Soybean Germplasm and Establishment of a Molecular Screening System for Soybean Pathogens

Prof. Karl Schmid, Universität Hohenheim, Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik, karl.schmid@uni-hohenheim.de

Prof. Ralf T. Vögele, Universität Hohenheim, Institut für Phytomedizin, ralf.voegel@uni-hohenheim.de

Dr. Volker Hahn, Universität Hohenheim, Landessaatzuchtanstalt, volker.hahn@uni-hohenheim.de

Dr. Christiane Balko, Julius-Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen, Institut für Resistenzforschung und Stresstoleranz, christiane.balko@julius-kuehn.de

To initiate effective plant protection measures in soybean, it is essential to obtain information on existing pathogens. Therefore, individual detections for the major soybean pathogens in Germany were developed based on quantitative Real Time-PCR. The pathogenicity of the pathogens was tested and inoculation procedures were established for further investigations. After establishing individual detections, first multiplex detections for different pathogens and sample types were established which can be used for early detection of different pathogens in a variety of soybean samples (seeds, plants, soil).

The project also established a core collection of accessions characterized by high diversity and suitable for cultivation under Central European conditions by combining current and future climate parameters with genotyping and phenotyping data. New genes for environmental adaptation were identified, for which molecular markers for breeding are being developed, and suitable accessions were selected for crossing into current breeding programs.

The results also clearly demonstrate that high-yielding lines can be bred to expand the soybean growing region northward. In addition, the variance components, heritabilities and trait correlations obtained provide a solid basis for designing breeding programs, including in particular Speed-Breeding programs, which can be accelerated in the future with genomic and phenomic selection. The first lines have already been released to private plant breeders, who are conducting further trials to determine whether the lines can be approved as varieties.

Particularly in cooler climates, such as in northern Germany, the risk of yield loss due to chilling stress is high. In this project, lines of cross progeny and a diverse set of genebank accessions were characterized with respect to their pod set after a cooling stress period. The test system produced reproducible results and first QTL for pod set under stress conditions were determined.

INHALT

Kontaktinformationen	2
Kurzfassung	3
Abstract	4
Abkürzungsverzeichnis	9
1 Einführung	10
1.1 Gegenstand des Vorhabens	10
1.2 Ziele und Aufgabenstellungen des Projekts, Bezug des Vorhabens zu den einschlägigen Zielen der Bekanntmachung	10
1.3 Planung und Ablauf des Projekts	12
2 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde	14
2.1 IPSP Charakterisierung und Nutzung pflanzengenetischer Ressourcen	14
2.2 LSA Marker- und Genomik-unterstützte Pflanzenzüchtung	15
2.3 IPM Identifizierung und Monitoring von Sojapathogenen	16
2.4 JKI-RS Identifizierung und Nutzung von Kühletoleranzgenen	16
3. Material und Methoden	17
3.1 IPSP	17
3.1a Implementierung einer Datenbank:	17
3.1b Literatursuche nach QTL und Kandidatengenen:	17
3.1c Allelemining und Identifizierung züchterisch interessanter Akzessionen:	18
3.1d Genotypisierung und Charakterisierung von modernen europäischen Sojasorten mit dem SoySNP50K Array:	19
3.1e Methoden zur Genomweiten und Umweltassoziationskartierung	20
3.1f Feldversuche	20
3.2 LSA	20
3.2a QTL-Analysen für Ertrag, Reifezeit, Tausendkornmasse, Protein- und Ölgehalt ..	20
3.2b QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert	22
3.2c Genomische Selektion	24
3.2d,e Kreuzung und Weiterführung Diallel1_neu und Diallel2_neu	24
3.2f QTL-Kartierung der Kühletoleranz	25
3.2g Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus AP1	25
3.2h Sortenprüfungen	25
3.2i Speed-Breeding-System	25
3.2k Ionomics	25
3.2l Genotypisierung	25
3.2m Regionale Genomische Selektion	26
3.3 IPM	27
3.3a,h Literaturarbeit, Primerdesign, Vortests	27

3.3b,i	Etablierung des quantitativen Nachweisverfahrens für verschiedene Pathogene	27
3.3c,k	Etablierung des quantitativen Nachweisverfahrens für verschiedene Probenotypen	28
3.3d,l	Etablierung einer quantitativen Methodik zur Erfassung von Nematoden.....	29
3.3e,m	Etablierung des Multiplex-Systems auf verschiedenen Proben	29
3.3f,n	Testung verschiedener Proben	30
3.4	JKI-RS.....	30
3.4a	Gefäßversuche unter kontrollierten Bedingungen zur Phänotypisierung des Hülsenansatzes unter Kühlestress.....	30
	Gefäßversuche unter kontrollierten Bedingungen zur Erfassung des Hülsenansatzes unter Kühlestress (JKI).....	33
	Genomweite Assoziationsstudie (LSA)	33
4.	Ausführliche Darstellung der wichtigsten Ergebnisse	35
4.1	IPSP	35
4.1c	Allelemining und Identifizierung züchterisch interessanter Akzessionen:.....	35
4.1d	Genotypisierung und Charakterisierung von modernen europäischen Sojasorten mit dem SoySNP50K Array:.....	40
AP1g:	Feinkartierung und Markerentwicklung neuer Reifegene in der CC4CE	46
AP1h:	Aufbau einer Allele Mining Pipeline mit publizierten Daten	46
AP1i:	Assoziationskartierung für neue Merkmale in Genbank-Akzessionen und europäischen Sorten.....	47
4.2	LSA	52
4.2a	QTL-Analysen für Ertrag, Reifezeit, Tausendkornmasse, Protein- und Ölgehalt ..	52
4.2b	QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert.....	57
4.2c	Genomische Selektion	65
4.2d,e	Kreuzung und Weiterführung Diallel1_neu und Diallel2_neu	66
4.2f	QTL-Kartierung der Kühletoleranz.....	66
4.2g	Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus AP1	66
4.2h	Sortenprüfungen	66
4.2i	Speed-Breeding-System	67
4.2k	Ionomics.....	68
4.2l	Genotypisierung	71
4.2m	Regionale Genomische Selektion	74
4.3	IPM.....	75
4.3a	Literaturarbeit, Primerdesign, Vortests	75
4.3b	Etablierung des quantitativen Nachweisverfahrens für verschiedene Pathogene	75

4.3c,k	Etablierung des quantitativen Nachweisverfahrens für verschiedene Probenotypen	77
4.3d,l	Etablierung einer quantitativen Methodik zur Erfassung von Nematoden.....	78
4.3e	Etablierung des Multiplex-Systems auf verschiedenen Proben	79
4.3f,m	Testung verschiedener Proben	79
4.3h	Primerdesign, Vortests.....	79
4.3k	Etablierung des quantitativen Nachweisverfahrens für verschiedene Probenotypen	81
4.3m	Etablierung des Multiplex-Systems auf verschiedenen Proben	82
4.4	JKI-RS.....	83
	Gefäßversuche unter kontrollierten Bedingungen zur Erfassung des Hülsenansatzes unter Kühlestress (JKI).....	91
	Genomweite Assoziationsstudie (GWAS) zum Hülsenansatz unter Kontroll- und Kühlestressbedingungen (HOH)	94
5.	Diskussion der Ergebnisse.....	96
5.1	IPSP	96
5.2	LSA	97
5.2a	QTL-Analysen für Ertrag, Reifezeit, Tausendkornmasse, Protein- und Ölgehalt ..	97
5.2b	QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert	98
5.2c	Genomische Selektion	104
5.2d,e	Kreuzung und Weiterführung Diallel1_neu und Diallel2_neu	105
5.2f	QTL-Kartierung der Kühletoleranz.....	105
5.2g	Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus AP1....	105
5.2h	Sortenprüfungen	105
5.2i	Speed-Breeding System.....	105
5.2k	Ionomics.....	105
5.2l	Genotypisierung	106
5.2m	regionale genomische Selektion.....	106
5.3	IPM.....	107
5.4	JKI-RS.....	108
6.	Angaben zum voraussichtlichen Nutzen und zur Verwertbarkeit der Ergebnisse	111
6.1	IPSP	111
6.2	LSA	111
6.3	IPM.....	111
6.4	JKI-RS.....	112
7	Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen, Hinweise auf weiterführende Fragestellungen.....	113
7.1	IPSP	113

7.2 LSA	114
7.3 IPM.....	115
7.4 JKI-RS.....	116
8 Zusammenfassung.....	117
IPM.....	117
IPSP	117
LSA	117
JKI-RS.....	118
9 References	120
10 Überblick über die im Berichtszeitraum realisierten Veröffentlichungen zum Projekt, bisherige und geplante Aktivitäten zum Projekt.....	130
Publikationen:.....	130
Vorträge	130
Poster.....	131
Anhang.....	132

ABKÜRZUNGSVERZEICHNIS

ASR: Asiatischer Sojabohnen Rost
BLE: Bundesanstalt für Landwirtschaft und Ernährung
BLUE: Best Linear Unbiased Estimates 21
BÖLN: Bundesprogramm Ökologischer Landbau und andere Formen nachhaltiger Bewirtschaftung
CC4CE: Core Collection for Central Europe
CHU: crop heat unit
CIM: composite intervall mapping
Ct: *Colletotrichum truncatum*
DArTseq: Diversity Arrays Sequence Technology
Dp: *Diaporthe phaseolorum*
dpi: Tage nach der Inokulation
EAS Environment Association Study
FIGS Focused Identification of Germplasm Strategy
Fs: *Fusarium* spezies
GBS: Genotyping by Sequencing
GWAS: Genomwide Association Study
IPM: Institut für Phytomedizin
IPSP: Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik
JKI-RS: Julius-Kühn-Institut, Institut für Resistenzforschung und Stresstoleranz
LOD: Logarithm of the Odds
LSA: Landessaatzuchtanstalt
LSD: Least Significant Difference
MARS: Marker Assisted Recurrent Selection
Md: *Microsphaera diffusa*
NGS: Next Generation Sequencing
NIRS: Nahinfrarotspektroskopie
P2: Primus
P3: Protina
P5: Sigalia
PCR: Polymerase Ketten Reaktion
Pm: *Peronospora manshurica*
Ps: *Pythium* spezies
Psg: *Pseudomonas syringae pv. glycinea*
qRT-PCR: quantitative reverse Transkriptase PCR
QTCAT: Quantitative Trait Cluster Association Test
QTL: Quantitative Trait Locus
REML: Restricted Maximum Likelihood
RIL: rekombinante Inzuchtlinien
RR-BLUP: Random Regression Best Linear Unbiased Prediction
Rs: *Rhizoctonia solani*
SCN: Sojabohnen Cysten Nematoden
SNP: Single Nucleotide Polymorphism
Ss: *Sclerotinia sclerotiorum*
SSD: Single Seed Descent
TKG: Tausendkorngewicht
Xag: *Xanthomonas axonopidis pv. glycines*

1 EINFÜHRUNG

1.1 GEGENSTAND DES VORHABENS

Leguminosen haben in Agrarökosystemen vielfältige positive Auswirkungen und sind ein Schlüsselement für einen verantwortungsvollen und nachhaltigen Umgang mit den der Landwirtschaft zur Verfügung stehenden Ressourcen. Insbesondere im Ökologischen Landbau sind Leguminosen, durch ihre Fähigkeit den Luftstickstoff zu nutzen, ein entscheidendes Fruchtfolgeglied. Die Sojabohne spielt dabei als weltweit wichtigste Leguminose eine herausragende Rolle. Allerdings wurde sie in Deutschland nur in geringem Umfang angebaut. Das übergeordnete Ziel des Vorhabens war es daher, die Voraussetzungen zu verbessern, damit der Sojaanbau in Deutschland ausgedehnt werden kann. Um dieses Ziel zu erreichen wurden in einzelnen Arbeitspaketen die folgenden Ziele angestrebt:

Der Aufbau eines Genomik-basierten Zuchtprogramms, die Verbreiterung der Züchtungspopulationen durch Allel-mining genetischer Ressourcen, die Etablierung eines schnellen und sensitiven Testverfahrens auf die Anwesenheit von Pathogenen an Sojabohnen sowie die phänotypische und genetische Erfassung der Reaktion von Sojalinien auf Kühlestress während der Blüte.

1.2 ZIELE UND AUFGABENSTELLUNGEN DES PROJEKTS, BEZUG DES VORHABENS ZU DEN EINSCHLÄGIGEN ZIELEN DER BEKANNTMACHUNG

Dargestellt wird ein Forschungs- und Entwicklungsvorhaben im Rahmen der Bekanntmachung Nr. 18/2012/31 zur „Ausweitung und Verbesserung des Anbaus und der Verwertung von Leguminosen mit Schwerpunkt Sojabohnen und Lupinen in Deutschland“ des Bundesprogramms Ökologischer Landbau und andere Formen nachhaltiger Landwirtschaft (BÖLN).

Im vorgestellten Vorhaben sollten vorhandene Kenntnis- und Erfahrungslücken für die Züchtung und den Anbau von Sojabohnen geschlossen werden. Durch die gezielte Nutzung genetischer und genomischer Ressourcen der Sojabohne sollten der Züchtung neue Allele für die optimierte Nutzung der Sojabohne unter unseren klimatischen Bedingungen zur Verfügung gestellt werden mit dem Ziel einer gesteigerten Ertragsstabilität und damit einer nachhaltigen Eiweißversorgung sowohl direkt für die menschliche Ernährung als auch in der tierischen Erzeugung. Durch den Aufbau eines nachhaltigen, Genomik-basierten Zuchtprogramms sollten mittel- bis langfristig neue Sojasorten entwickelt werden, die besser an unsere Klimaverhältnisse angepasst sind und die notwendigen Qualitäten aufweisen, um in der menschlichen Ernährung bzw. Tierfütterung eingesetzt werden zu können. Gleichzeitig war das Ziel die Kenntnisse zu erweitern, wie eine optimierte Sojazüchtung gestaltet werden muss. Hierfür wurden Erkenntnisse zur Vererbung wichtiger agronomischer Eigenschaften und von Qualitätsmerkmalen gesammelt und in das Zuchtprogramm eingearbeitet. Weitere wichtige Bausteine waren die Erfassung der Kühletoleranz von Sojabohnen während der Blüte und die Erfassung der Qualität zur Herstellung von Tofu. Diese Merkmale sind nur sehr aufwendig zu erfassen, daher sollten die Kenntnisse erweitert werden, wie die Merkmale züchterisch zukünftig einfacher bearbeitet werden können z. B. über die Nutzung molekularer Marker. Das erstellte Zuchtmaterial wird der privaten Pflanzenzüchtung zur Verfügung gestellt. Dadurch ist gewährleistet, dass das Material auch in der breiten Landwirtschaft Verwendung finden wird.

Nachweisverfahren zur Erfassung des Erregerpotentials von Parasiten und Pathogenen sind sowohl für die Einleitung spezifischer Pflanzenschutzmaßnahmen aber vor allem auch zur

Einschätzung von langfristigeren Maßnahmen wie dem Einsatz von Pflanzenstärkungsmitteln oder biologischer Fungizide sowie der phytosanitären Wirkung von Fruchtfolgen von wesentlicher Bedeutung. Mit den im Vorhaben dargestellten Untersuchungen sollten die Voraussetzungen geschaffen werden, um molekulare Nachweismethoden für die in unseren Breiten bedeutenden Sojapathogene zu entwickeln. Damit können zukünftig Aussagen über die Qualität des eingesetzten Saatguts und des geplanten Standortes für den Sojaanbau getroffen werden. Das Verfahren sollte gleichermaßen einsatzfähig sein für Landwirte und Berater und sollte bereits früh auf Praxistauglichkeit durch Untersuchung von Proben von Praxisschlägen getestet werden.

1.3 PLANUNG UND ABLAUF DES PROJEKTS

Das GenoPath-Projekt besteht aus 4 Arbeitspaketen und 2 Projektteilen, dem ursprünglichen Projekt und der Verlängerung

Arbeitspaket 1: IPSP – Verbreiterung der genetischen Züchtungspopulationen durch Allelemining genetischer Ressourcen.

Arbeitspaket 2: LSA – Aufbau eines nachhaltigen Genomik-basierten Soja-Zuchtprogramms.

Arbeitspaket 3: IPM – Quantitative schnelle Erfassung von Soja-Pathogenen.

Arbeitspaket 4: JKI-RS – Kühlestress: Phänotypisierung einer Population aus Eltern mit Unterschieden in der Kühletoleranz hinsichtlich des Hülsenansatzes nach einer Kühlestressphase.

Die Details zu den Arbeitspaketen sind in Abbildung 1 und Abbildung 2 dargestellt. Die Darstellung in den Kapiteln 3-5 orientiert sich an den dargestellten Arbeitspaketen.

AP	Aktivität	2015			2016				2017				2018
		II	III	IV	I	II	III	IV	I	II	III	IV	I
1a	Implementierung einer Datenbank			M1									
1b	Literatursuche nach QTLs und Kandidatengene			M2									
1c	Allelemining und Identifizierung züchterisch interessanter Akzessionen				M3								
1d	Genotypisierung von Genbankakzessionen mit Genotyping by Sequencing								M4				
1e	Integrierung in existierende Züchtungspopulationen												M5
2a	QTL-Analysen für Ertrag, Reifezeit, TKM, Protein- und Ölgehalt			M6						M11			
2b	QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert				M7								M13
2c	Genomische Selektion							M8	M10				
2d	Kreuzung und Weiterführung Diallel1_neu							M9					
2e	Kreuzung und Weiterführung Diallel2_neu												
2f	QTL-Kartierung der Kühletoleranz												M12
2g	Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus 1												
2h	Sortenprüfungen												
3a	Literaturarbeit, Primerdesign und Vortests	M14											
3b	Etablierung des quantitativen Nachweisverfahrens für verschiedene Pathogene				M16								
3c	Etablierung des quantitativen Nachweisverfahrens für verschiedene Probenotypen						M16						
3d	Etablierung einer quantitativen Methodik zur Erfassung von Nematoden											M17	
3e	Etablierung des Multiplex-Systems auf verschiedenen Proben											M18	
3f	Testung verschiedener Proben												M19
3g	Erstellung von Publikationen					M20						M21	M22
4a	Kühlestresstests				M23				M24				
4b	Datentransfer, Berichterstellung											M25	

Abbildung 1 Arbeitspakete und geplante Zeitrahmen, Teil 1.

AP	Aktivität	2019			2020				2021
		II	III	IV	I	II	III	IV	I
1f	Resequenzierung		M1-1		M1-2				
1g	Feinkartierung und Markerentwicklung				M1-3				
1h	Aufbau einer Allele Mining Pipeline				M1-5				
1i	Assoziationskartierung			M1-4		M1-6			M1-7
2i	Speed-Breeding-System			M2-1	M2-2			M2-3	
2k	Ionomics		M2-4	M2-5		M2-6			
2l	Genotypisierung			M2-7	M2-8			M2-9	
2m	Regionale Genomische Selektion				M2-10		M2-11		
3h	Primerdesigns, Vortests								
3i	Etablierung eines quantitativen Nachweisverfahrens für verschiedene Pathogene		M3-1						
3k	Etablierung eines quantitativen Nachweisverfahrens für verschiedene Probentypen			M3-2					
3l	Etablierung einer quantitativen Methodik zur Erfassung von Nematoden				M3-3				
3m	Etablierung des Multiplex-Systems auf verschiedene Proben				M3-4				
3n	Testung verschiedener Proben						M3-5		
3o	Publikation		M3-6						
4c	Wdh1				M4-1				
4d	Wdh2								M4-2
4e	Datentransfer, Bericht								

Abbildung 2: Arbeitspakete und geplante Zeitrahmen, Teil 2.

2 WISSENSCHAFTLICHER UND TECHNISCHER STAND, AN DEN ANGEKNÜPFT WURDE

2.1 IPSP CHARAKTERISIERUNG UND NUTZUNG PFLANZENGENETISCHER RESSOURCEN

Die Züchtung von an europäische Bedingungen angepassten Sorten basiert derzeit auf wenigen Ausgangslinien mit einer engen genetischen Variation, die mit den hierzulande herrschenden Umweltbedingungen zurechtkommen. Dem gegenüber steht eine große genetische Vielfalt von Soja weltweit, die als mögliche Quellen nutzbarer genetischer Variation für die Ertragssteigerung, die Resistenzzüchtung und die Verbesserung von Qualitätseigenschaften verwendet werden können. Die Nutzung dieser genetischen Ressourcen ist notwendig, um die derzeit in Europa genutzten Züchtungspopulationen genetisch zu verbreitern, damit die Sojabohne nachhaltig zu einer konkurrenzfähigen Kulturart für den Anbau in Europa entwickelt werden kann.

Aufgrund der großen Bedeutung von Soja gibt es umfangreiche genetische und genomische Ressourcen. In der US-amerikanischen Genbank sind mehr als 10.000 Akzessionen registriert. Dazu kommen noch mehrere Tausend Akzessionen in asiatischen Genbanken, die sich in wichtigen Ertrags-, Anpassungs- und Qualitätsmerkmalen unterscheiden. Zusätzlich gibt es eine große Menge genomischer Daten, die öffentlich verfügbar sind. Nach der Sequenzierung des Soja-Genoms (Schmutz et al. 2010) werden derzeit Re-Sequenzierungen und Transkriptomanalysen zahlreicher Sojalinien durchgeführt. Die umfangreichen genetischen Ressourcen der Sojabohne werden derzeit sowohl genetisch als auch phänotypisch charakterisiert. Vor allem die Genomsequenzierung macht sehr schnelle Fortschritte, so dass es bereits umfangreiche genomische Ressourcen wie zum Beispiel eine sogenannte HapMap aus tausenden imputierter Genome beziehungsweise ein Pangenom vollständiger Genbankakzessionen gibt. Diese Informationen sind eine wertvolle Ressource, um effizient Ausgangsmaterial mit potentiell nützlicher genetischer Variation zu identifizieren.

Die Verbindung von genomischen Daten, Herkunftsinformationen, Klimadaten und phänotypischer Charakterisierung in Genotypenscreens in Zielregionen ermöglicht die Identifizierung von Kandidatengen und genomischen Regionen in genetischen Ressourcen, die mit einer hohen Wahrscheinlichkeit nützliche genetische Variation für die Züchtung enthalten (Bari et al. 2012). Die für die Identifizierung dieser Variation verwendeten Ansätze werden aus der quantitativen Genetik, z. B. Genom-weite Assoziationsstudien (Gawenda et al. 2015) oder der Populationsgenetik (z.B. Selektionstests) abgeleitet (Cao et al. 2011; Günther und Schmid 2011). Die Vielfalt genetischer Variation in genetischen Ressourcen von Soja wurden zusammen mit klimatischen und phänotypischen Daten (insbesondere zum Reifegrad) untersucht. Mittels Focused Identification of Germplasm Strategy (FIGS) wurde eine genetisch hoch diverse Gruppe von Genbankakzessionen (Core Collection for Central Europe [CC4CE]) etabliert, die sich aufgrund des Klimas in der Herkunftsregion für den Anbau in Zentraleuropa unter den aktuellen und vom IPCC prognostizierten Klimabedingungen eignen könnten (Haupt and Schmid 2020). Damit ist es möglich, Genotypen aus den genetischen Ressourcen auszuwählen, die sich für eine Einkreuzung in aktuelles Zuchtmaterial eignen.

2.2 LSA MARKER- UND GENOMIK-UNTERSTÜTZTE PFLANZENZÜCHTUNG

Die Pflanzenzüchtung spielt eine Schlüsselrolle bei der Entwicklung von Sojasorten, die sich für den Anbau unter unseren klimatischen Bedingungen eignen und zudem den hohen Qualitätsansprüchen der Industrie entsprechen. Nachdem die Selektion basierend auf genomischen Daten in der Tierzucht breit angewendet wird, hat dieses Verfahren in jüngster Zeit Einzug in die Pflanzenzüchtung gehalten (Würschum et al. 2014). Komplexe, aber ökonomisch wichtige Eigenschaften wie der Ertrag oder manche Qualitätseigenschaften sind züchterisch schwierig zu bearbeiten, da sie von vielen Genen gleichzeitig beeinflusst werden. Die Genomik-basierte Züchtung ist eine Form der markergestützten Selektion, die überlegene Individuen nicht basierend auf in Feldversuchen erhobenen Daten, sondern aufgrund ihres genomisch geschätzten Zuchtwertes selektiert. Diese Art der Selektion ist als äußerst vielversprechende Methode der Pflanzenzüchtung anzusehen, vor allem aufgrund der stetig sinkenden Kosten für die molekularen Marker (Zhao et al. 2012). Allerdings ist die Anwendung in Pflanzenzuchtprogrammen noch längst keine Routine und es gibt noch einen hohen Forschungsbedarf beispielsweise bei der Entwicklung und Umsetzung der komplexen Algorithmen, die zur Zuchtwertschätzung herangezogen werden. Eine notwendige Voraussetzung für eine genomische Selektion ist das Vorhandensein einer großen Anzahl molekularer Marker in der verwendeten Kulturart. In jüngster Zeit werden daher verstärkt Next-Generation-Sequencing-Methoden verwendet, die eine hohe Zahl an Markern generieren. Eine dieser Methoden ist die Diversity Arrays Sequence Technology (DArTseq) deren Anwendung seit jüngster Zeit auch für Sojabohnen möglich ist (Hang Vu et al. 2012) und bereits in der Sojaforschung eingesetzt wird (Hahn and Würschum 2014). Um darüber hinaus auch wissensbasierte Pflanzenzüchtung (breeding by design) wie beispielsweise das MARS-Verfahren (marker assisted recurrent selection) durchführen zu können (Bohra 2013) ist es wichtig, Untersuchungen zur Anzahl und Lokalisation der den wichtigen agronomischen Merkmalen wie Ertrag, Ölgehalt, Proteingehalt und Tausendkornmasse zugrundeliegenden Genen oder Genbereichen (QTL) durchzuführen. Da Aussagen über solche Bereiche jeweils stark vom verwendeten Untersuchungsmaterial abhängen (Würschum 2012), ist man in der letzten Zeit dazu übergegangen, anstatt biparentaler Populationen (also von Populationen, die auf zwei Eltern zurückgehen) Mehrlinienkreuzungen zu verwenden (Steinhoff et al. 2011). Dies sind verbundene Populationen, die auf mehreren Elternlinien basieren. Damit können die Vorteile einer biparentalen Population mit den Vorteilen der Assoziationskartierung verbunden werden. QTL-Analysen basierend auf zwei verbundenen Populationen wurden kürzlich für Protein- und Ölgehalte für chinesisches (spätreifendes) Soja-Zuchtmaterial publiziert und zeigen die Vorteile des Verfahrens auf (Mao et al. 2013). Für frühreifes Sojamaterial lagen solche Studien nicht vor.

Die Tofuabbeute und -qualität sind wichtige Qualitätskriterien von Sojasorten, die für die Tofuproduktion eingesetzt werden sollen. In einem vom BÖLN geförderten Kooperationsprojekt hat das Taifun-Tofu-Werk in Freiburg eine Labortofufabrik entwickelt, die diese Merkmale messen kann und dabei nur ca. 80 g Sojasaatgut benötigt. Damit ist es erstmals möglich, Untersuchungen zur Tofuqualität von europäischem Tofu (der sich in einigen Eigenschaften, wie z.B. Festigkeit, von asiatischem Tofu unterscheidet) durchzuführen. Bisher werden für die Tofuproduktion die Sojabohnen über Nacht eingeweicht, anschließend wird die Sojamilch ausgepresst um damit Tofu zu erstellen. Neuere Technologien verwenden jedoch wesentlich kürzere Einweichzeiten. Dies führt zu deutlichen Einsparungen an benötigtem Platz und an Produktionszeit. Bisher gibt es jedoch keine Untersuchungen darüber, ob sich Sojabohnen genetisch darin unterscheiden, wie schnell und in welchem Ausmaß sie in der Lage sind, Wasser aufzunehmen zu können.

2.3 IPM IDENTIFIZIERUNG UND MONITORING VON SOJAPATHOGENEN

Der Sojaanbau in Deutschland ist grundsätzlich mit zwei großen Problemen konfrontiert: den relativ niedrigen Tagesdurchschnittstemperaturen und Ertrags- und/oder Qualitätsverluste verursacht durch verschiedene Pathogene.

Für die Sojabohne sind mehr als 90 verschiedene Pathogene beschrieben (mindestens fünf bakterielle und rund 50 pilzliche Erreger, 15 Viren und 20 Nematoden). Die Bedeutung der einzelnen Erreger hängt stark von der Anbauregion und den Umweltbedingungen ab. Weltweit stellt der Asiatische Sojabohnen Rost (ASR), verursacht durch *Phakopsora pachyrhizi*, das bedeutendste Pathogen der Sojabohne dar. In unseren Breiten spielt ASR aufgrund der Witterung bisher allerdings keine Rolle. Neben Auflaufkrankheiten, verursacht durch verschiedene *Pythium* (Ps) und *Fusarium* Spezies (Fs), durch *Diaporthe phaseolorum* (Dp) und durch *Rhizoctonia solani* (Rs), sind bei uns vor allem der Falsche Mehltau (*Peronospora manshurica*, Pm), der Echte Mehltau (*Microsphaera diffusa*, Md) und die Anthraknose (*Colletotrichum truncatum*, Ct), sowie Bakteriosen verursacht durch *Xanthomonas axonopodis* pv. *glycines* (Xag) und *Pseudomonas syringae* pv. *glycinea* (Psg) von Bedeutung.

Das wirtschaftlich bedeutendste Sojapathogen in unseren Breiten ist allerdings *Sclerotinia sclerotiorum* (Ss). Besondere Probleme bei Ss sind die Überdauerungsformen (Sklerotien), die über ein Jahrzehnt hinweg infektiös bleiben können, sowie das mit 400 verschiedenen Wirtspflanzen extrem breite Wirtsspektrum, zu dem neben vielen anderen Leguminosen auch Solanaceen und Raps gehören.

Eine Umfrage unter Sojaproduzenten im Jahr 2013 identifizierte Ss als bedeutendstes Pathogen, gefolgt von den bakteriellen Erkrankungen hervorgerufen durch Xag und Psg, sowie Krankheiten, die durch die Pilze Pm, Md, Dp, Ps, Fs und Rs hervorgerufen werden. Während bei der Umfrage 2013 Nematoden kein größeres Gefährdungspotential zugeordnet wurde, scheinen erste Befunde aus 2014 ein verstärktes Auftreten vor allem von Sojabohnen Cysten Nematoden (SCN) anzudeuten. Aus diesem Grund sollten SCN zusätzlich in die Analysen mit einbezogen werden.

Zur Einleitung effizienter Pflanzenschutzmaßnahmen ist es essentiell, über die Art und Menge möglicherweise vorhandener Pathogene rechtzeitig Aufschluss zu gewinnen. Das IPM arbeitet seit Jahren an den molekularen Grundlagen der Interaktion ASR – Sojabohne. Im Zuge der Forschungsarbeiten wurden auch für das ASR Pathosystem molekulare Nachweise entwickelt, die eine quantitative Erfassung des Pathogens ermöglichen.

2.4 JKI-RS IDENTIFIZIERUNG UND NUTZUNG VON KÜHLETOLERANZGENEN

Für einen erfolgreichen Sojaanbau werden Sorten mit hoher Ertragsstabilität benötigt. Solche Sorten sind bislang nur für klimatisch sehr günstige Lagen wie die Oberrheinebene verfügbar. In kälteren Lagen ist die Gefahr von Ertragseinbußen und damit auch Ertragsschwankungen sehr groß, da die Sojabohne in zwei Entwicklungsstadien besonders empfindlich auf Kühlestress reagiert: (I) während der Keimung und des Sämlingsstadiums mit Auswirkung auf die Bestandesdichte und (II) während der Blüte und des Hülsenansatzes mit Auswirkung auf den Ertrag (verminderter Hülsenansatz und Kornansatz in der Hülse) (Kurosaki et al. 2003) und die Qualität (Verfärbungen, Risse durch Kühleschäden). Entsprechend wurde im Freiland (Hokkaido, Japan) eine stark negative Korrelation zwischen den Temperaturen im Juli zur Blüte der Sojabohnen und dem Kornertrag nachgewiesen (Horiguchi 1980). Dabei ist der Hülsenansatz die Ertragskomponente, die bei Kühlestress am deutlichsten beeinflusst wird (Kurosaki et al. 2003). Über Ausfälle im Hülsenansatz, verursacht durch Kühlestress in der Blüte, berichten auch Hume und Jackson (1981), Lawn and Hume (1985) sowie Michailov et al. (1989). Tolerante Genotypen zeichnen sich schon während der Stressphase durch einen anhaltenden Hülsenansatz aus (Kurosaki et al. 2003). Dagegen können empfindliche

Genotypen schon bei < 18 °C mit vermindertem Hülsenansatz reagieren, während andere Temperaturen bis zu 15 °C oder geringer noch tolerieren. Eigene Untersuchungen bestätigten eine gute Differenzierbarkeit von Genotypen bei einem Tag-/Nachtregime von 15 °C/10 °C (Balko et al. 2014). In Untersuchungen mit einer Nachttemperatur von nur 5 °C wurde kein anhaltender Hülsenansatz beobachtet und Unterschiede in der Kühletoleranz waren erst 6 Wochen nach der Stressphase erkennbar (Cober et al. 2013). Dabei korreliert die besonders sensitive Phase im Hülsenansatz im Wesentlichen mit der Blütenentwicklung an den Nodien 4-9 (Gass et al. 1996).

Eine genotypische Variabilität für Kühletoleranz – eine wichtige Voraussetzung, um die Sojabohne an die hiesigen klimatischen Bedingungen anzupassen – ist sowohl für Kühletoleranz per se als auch für die Steigerung der Kühletoleranz durch Härtung beschrieben worden (Funatsuki et al. 2003). Für die Anpassung an niedrige Temperaturen sind mindestens sechs verschiedene Gene verantwortlich.

Auch Cober et al. (2013) konnten, basierend auf der Analyse von zwei Populationen, 6 chromosomale Regionen mit potenziellem Einfluss auf die Kühletoleranz identifizieren.

Mit dem Hülsenansatz nach der Stressphase war von einer Reihe untersuchter physiologischer Merkmale vor allem der Chlorophyllgehalt positiv korreliert. Ein Abfall im Chlorophyllgehalt wird unter anderem auf Störungen in der symbiontischen N-Fixierung und eine deutliche Reduktion der Nitrogenase Aktivität zurückgeführt (van Heerden et al. 2008).

3. MATERIAL UND METHODEN

3.1 IPSP

3.1a Implementierung einer Datenbank:

In der Datenbank wurden insbesondere alle öffentlich verfügbaren Informationen über die Sojaakzessionen der USDA Genbank (Nelson 2011) zusammengefasst. Der Datensatz umfasst Passport-Informationen (z.B. Herkunft) und phänotypische Informationen, sowie klimatischen Daten (Fick und Hijmans 2017; Hijmans et al. 2005) zur Charakterisierung der Umwelt aus der Genbankmaterial kommt.

Phänotypische Daten lagen für ~20.000 Akzessionen vor. Diese umfassen qualitative (morphologische Merkmale wie z.B. Blüten- / Behaarungs- / Samenfarbe, Samen- / Blattform, Resistenzen, Reifegruppenzugehörigkeit etc.) und quantitative Merkmale (z.B. Ertrag, 100-Korngewicht, Protein- / Ölgehalt, Blüh- / Reifezeitpunkt etc.).

~7.000 Akzessionen verfügten in den Passportdaten der Genbank über einen georeferenzierten Sammlungsort und ermöglichten die Erhebung von Umweltdaten für die Herkunftsregionen der entsprechenden Akzessionen. Hierbei wurde auf *WorldClim* Datensätze (www.worldclim.org) zurückgegriffen. Klima- und Umweltdaten wurden für den Zeitraum Mai bis September berücksichtigt, was in etwa dem Anbauzeitraum für Soja in der nördlichen Hemisphäre entspricht.

Die ebenfalls verfügbaren genotypischen Daten für das Genbankmaterial wurden nicht explizit in die Datenbank eingepflegt, aber wir nutzten sie im Rahmen der Arbeitspakete AP1c und AP1d. Diese Daten umfassen ~50.000 SNP Marker für die ~20.000 Genbankakzessionen charakterisiert wurden. Bei Letzteren handelt es sich zu 2/3 um Landrassen und den wilden Vorläufer der Sojabohne und zu 1/3 um moderneres Material.

3.1b Literatursuche nach QTL und Kandidatengen:

Eine explizite Erfassung von publizierten QTL und Kandidatengen ist nicht erfolgt, da mit der *Soybase* Ressource (www.soybase.org) zu Bearbeitungsbeginn des Projekts bereits eine

entsprechende öffentlich verfügbare Datenbank zur Verfügung steht. Hier sind und werden funktionelle genetische Informationen für die Sojabohne hinterlegt. Wir nutzten Soybase zum Abgleich mit Ergebnissen aus AP1c.

3.1c Allelemining und Identifizierung züchterisch interessanter Akzessionen:

Zum Zwecke der Identifizierung von möglichst vielversprechenden genetischen Ressourcen zur Einkreuzung in aktuelles Zuchtmaterial wurde die USDA Sojabohnen Genbank (Nelson 2011) unter Zuhilfenahme öffentlich zugänglicher Daten wie der vermerkten Passport Informationen (Herkunft, phänotypische Informationen etc.), genotypischer Informationen (Song et al. 2015) sowie von klimatischen Daten (Fick und Hijmans 2017; Hijmans et al. 2005) systematisch durchsucht. Am Beginn dieses Unterfangens stand die klimatische Charakterisierung der derzeitigen und zukünftigen Zielumwelt für den mitteleuropäischen Sojaanbau, sowie der Spenderumwelt (d.h. des ostasiatischen Herkunftsgebiets der Sojabohne). Zur Strukturierung der erzeugten multivariaten Klimadaten wurden Hauptkomponentenanalysen und ein Vergleich von Ziel- und Spenderumwelt durchgeführt. Dieser Vergleich und die zusätzliche Berücksichtigung der Reifegruppenevaluierung aus den Passport-Informationen identifizierte vielversprechendes Genbankmaterial mit potentieller abiotischer Anpassung an vergleichsweise kühle Anbauregionen höherer geografischer Breite, wie sie für die mitteleuropäische Sojazüchtung notwendig ist.

Um Genombereiche mit potentieller Umwelanpassung zu identifizieren, wurde zwischen der Gruppe des ausgewählten Materials und dem Rest der Genbanksammlung ein Genomscan mit dem BAYPASS Programm durchgeführt (Gautier 2015). Die genetischen Grundlagen für potentiellen Umwelanpassungen zwischen den Gruppen wurden hierbei mithilfe von über die generelle Populationsstruktur hinausgehenden genetischen Differenzierungsstatistiken kartiert (Günther und Coop 2013).

Die Kandidatengruppe wurde in der Folge zu zwei Kernkollektionen weiterentwickelt, um als Grundlage für zukünftige Prebreeding Aktivitäten und für Trait-mining Studien zu dienen. Unter einer Kernkollektion ist hierbei eine in der Anzahl reduzierte Auswahl an Akzessionen zu verstehen, die das Diversitätslevel der Ausgangsgruppe weitestgehend erhalten und Redundanz in Form von einander sehr ähnlichen Akzessionen vermeiden soll. Die eingehende genotypische und phänotypische Charakterisierung von Kernkollektionen soll eine effiziente züchterische Nutzung gewährleisten (Frankel 1984). Die Bildung der Kernkollektionen basierte auf den verfügbaren genotypischen Daten und wurde mit R CoreHunter durchgeführt (Beukelaer et al. 2017; Thachuk et al. 2009). Der CoreHunter Algorithmus gewährleistet die effiziente Identifizierung von optimalen Kernen aus der sehr großen Population potentiell möglicher Akzessionskombinationen unter Erfüllung definierbarer Optimierungsparameter. Unterschiedliche Optimierungs- (Beukelaer et al. 2012; Thachuk et al. 2009) und Samplingstrategien (siehe Tabelle 1) wurden evaluiert und dokumentiert, um die beste Strategie für das vorliegende Genbank- und Datenmaterial zu identifizieren. Beide Kernkollektionen wurden auf ihre Repräsentationsleistung für die Kandidatengruppe mittels genotypischer Hauptkomponentenanalysen sowie anhand der vorliegenden phenotypischen Daten mittels Brown-Forsythe Levene Test und Welch's t-Test geprüft (Hui et al. 2008).

Tabelle 1: Übersicht über die evaluierten Samplingstrategien zur Bildung von Kernkollektionen.

Samplingstrategie	
Ohne Stratifizierung	Sampling aus ganzer Gruppe, keine Untergruppierung
Klassische Stratifizierung	Sampling in drei Gruppen: 1.1 Reifegruppen 000 – 0 1.2 Reifegruppe I 1.3 Reifegruppen II – X 2. Fusionierung von 1.1 – 1.3 zur Kernkollektion
2-fache Pseudostratifizierung	1. Sampling innerhalb der Reifegruppen 000 – 0 2. Sampling innerhalb der Reifegruppen 000 – X wobei Ergebnisse aus 1. schon festgelegt sind
3-fache Pseudostratifizierung	1. Sampling innerhalb der Reifegruppen 000 – 0 2. Sampling innerhalb der Reifegruppen 000 – I wobei Ergebnisse aus 1. schon festgelegt sind 3. Sampling innerhalb der Reifegruppen 000 – X wobei Ergebnisse aus 2. schon festgelegt sind

Die eingehende phänotypische Charakterisierung der erstellten Kernkollektionen durch Wissenschaft und praktische Sojazüchter ist die Voraussetzung für eine effiziente züchterische Nutzung dieses Genbankmaterials in der Zukunft. In einer Pilotstudie wurde in einem einortigen Feldversuch am Standort Hohenheim im Jahr 2018 anhand einer weiteren Kernkollektion evaluiert, wie eine solche Charakterisierung effizient durchgeführt werden kann. 210 Akzessionen wurden in einem mit zwei mitteleuropäischen Schecks (die Sorten Merlin und Regina) angereicherten Versuchsdesign evaluiert. Neben klassischen Merkmalen wie Ertrag und Abreife wurde der Schwerpunkt hier auf Merkmale gelegt, die über den gesamten Vegetationsverlauf non-invasiv erhoben werden können, und/oder die Aufschluss über etwaige Umwelтанpassungen geben können. Beispiele wären hier die Bodendeckung mit Soja und die Erhebung der Pflanzenhöhe in wöchentlichen Abständen. Die Quantifizierung dieser Merkmale erfolgte auf Grundlage von drohnenbasierten RGB Aufnahmen. Verwendet wurde die Drohne *Phantom 4 Pro* des Herstellers DJI. Die Verarbeitung der Bildaufnahmen (insbesondere Stitching, 3D Modellierung, Erstellung von Digital Elevation Model und Orthomosaic) wurde mit der Software *Agisoft Photoscan* durchgeführt.

3.1d Genotypisierung und Charakterisierung von modernen europäischen Sojasorten mit dem SoySNP50K Array:

Vormals „Genotypisierung von Genbankakzessionen mit Genotyping by Sequencing (GBS)“. Die Refokussierung der Genotypisierung von Genbankmaterial auf modernes europäisches Material ist erfolgt, da zu Projektbeginn mit der Veröffentlichung genotypischer Daten für die globale Sojasammlung der USDA Genbank bereits sehr gute genotypische Informationen für Genbankakzessionen in Form von SNP Markern zur Verfügung standen (Song et al. 2015). Eine Kontextualisierung von modernem europäischem Material mit diesem Genbankmaterial ist jedoch potentiell sehr informativ. Entsprechende genotypische Daten standen für europäisches Material nicht zur Verfügung. Daher genotypisierten wir 160 moderne europäische Sorten mit dem SoySNP50K Array (Song et al. 2013). Das Material der USDA

Genbank wurde mit der gleichen Ressource genotypisiert, was eine gute Datengrundlage für Vergleiche gewährleisten sollte.

In der Folge analysierten wir die Populationsstruktur im gesamten Datensatz. Zur Untersuchung der grundlegenden genetischen Struktur des globalen Sojasamples nutzten wir Hauptkomponentenanalysen und berechneten die genetische Distanz sowohl paarweise zwischen Individuen als auch zwischen Materialgruppen. Als Maß für die genetische Distanz nutzten wir die modifizierte Rogers Distanz (z.B. Reif et al. 2005). Das Kopplungsungleichgewicht wurde mittels r^2 ermittelt (Weir 1996). Zur Ermittlung grundlegender Abstammungsverhältnisse nutzten wir ADMIXTURE (Alexander et al. 2009), welches mittels modellbasierter Maximum-Likelihood Schätzverfahren die individuelle Abstammung von Ursprungspopulationen detektiert.

3.1e Methoden zur Genomweiten und Umweltassoziationskartierung

Von den zahlreichen Methoden für GWAS (genomwide association study) und EAS (environment association study) verwendeten wir zum einen BAYPASS (Gautier 2015), welche es erlaubt genetische Differenzierungsstatistiken mit Umweltparametern oder phänotypischen Merkmalen zu korrelieren. Daneben verwendeten wir den Quantitative Trait Cluster Association Test (QTCAT; Klasen et al. 2016), welche eine Clusterung von Allelfrequenzen entlang des Genoms durchführt, um für Populationsstruktur zu korrigieren und mit einer hohen Sensitivität und Spezifität Marker-Merkmal-Assoziationen durchführen kann. Für die Identifizierung von epistatische Selektion zwischen Genen wurde das Programm GWESS (Boyrie et al. 2021) verwendet.

3.1f Feldversuche

Für die Untersuchung wichtiger phänologischer Merkmale wurden Feldversuche an den beiden Standorten Eckartsweier (Baden-Württemberg) und Groß-Lüsewitz (Mecklenburg-Vorpommern) durchgeführt. In diesen Versuchen wurden sowohl die CC4EU Core collection und Checks in einem "Augmented Block Design" angebaut und die Parameter Tage bis zur Blüte und Tage bis zur Reife bestimmt.

3.2 LSA

3.2a QTL-Analysen für Ertrag, Reifezeit, Tausendkornmasse, Protein- und Ölgehalt

Pflanzenmaterial, Feldversuche und statistische Auswertung

Für dieses Arbeitspaket wurden phänotypische Daten genutzt, die in einem von der BLE geförderten Vorhaben in dreierartigen Feldversuchen im Jahr 2014 ermittelt wurden. Die Details sind publiziert in Kurasch et al. (2017). Dabei wurden neun spaltende Populationen mit insgesamt 1008 rekombinanten Inzuchtlinien (RIL) sowie die fünf Eltern in Eckartsweier, Hohenheim und Neuenstein angebaut. Die Elternlinien sind Sorten, die gute agronomische Eigenschaften in Deutschland aufwiesen. In den Feldexperimenten wurden die Linien in drei Versuchen angebaut, abhängig von ihrer Reifezeit auf der Basis von Felddaten aus 2013. Die Parzellengröße betrug 9 m² und es wurden 65 Bohnen je m² ausgesät. Die Versuche wurden teilweise wiederholt angebaut mit einem Wiederholungsfaktor von 1,2. Die folgenden Merkmale wurden erfasst: Kornertrag, Tausendkornmasse, Pflanzenhöhe, Proteingehalt und Ölgehalt.

Die Varianzkomponenten wurden mit der Methode der eingeschränkten maximalen Wahrscheinlichkeit (REML) in einem vollständigen Zufallsmodell geschätzt, und die Bedeutung der Varianzkomponentenschätzungen wurde durch Wahrscheinlichkeits-Verhältnisprüfungen bestimmt. Die Heritabilität (h^2) wurde nach dem von Piepho und Möhring (2007) vorgeschlagenen Ansatz mit einem festen genotypischen Effekt geschätzt. Die besten

linearen unvoreingenommenen Schätzungen (BLUEs) wurden unter der Annahme fester genotypischer Effekte berechnet. Phänotypische Korrelationen zwischen den Merkmalen wurden als Pearson-Korrelationskoeffizienten basierend auf den BLUEs geschätzt. Die statistische Analyse wurde mit dem Programm R 3.2.0 (R Core Team 2015) unter Verwendung des ASReml 3.0 Pakets durchgeführt (Gilmour et al. 2009).

Genotypisierung

DNA-Proben der Sojalinien wurden an die Cornell University zur Genotypisierung durch Sequenzierung (GBS; 192-fach) geschickt, wie von (Elshire et al. 2011) beschrieben. Die Elternlinien wurden siebenmal untersucht, um die Sequenziergenauigkeit zu erhöhen, während die Nachkommen nicht repliziert wurden. Die in dem Softwarepaket Tassel Version 5.0 implementierte GBS-Analysepipeline wurde für das Auffinden von SNPs verwendet (Glaubitz et al. 2014). Das Alignment der Reads auf das Referenzgenom Gmax Wm82.a2.v1 (Schmutz et al. 2010) erfolgte mit dem Burros Wheeler Alignment (BWA) Paket (Li und Durbin 2010).

Alle Marker, die in den Elternlinien heterozygot waren, wurden entfernt. Innerhalb jeder Population wurden SNPs mit mehr als 85 % fehlenden Werten entfernt. Die Berechnung der restlichen fehlenden Werte wurde populationsweise mit LinkImpute durchgeführt (Money et al. 2015). Schließlich wurden monomorphe Stellen innerhalb jeder Population entfernt.

Zusätzlich zu den genomweiten Markerprofilen wurden auch noch die wichtigsten vier Reifegene und der Hauptdeterminismus-Locus analysiert. Basierend auf den Publikationen von (Tsubokura et al. 2014) und Xu et al. (2013) wurden 17 PCR Protokolle inkl. Restriktionsenzymverdau für die Reifegene *E1-E4* und den Determinismus-Locus *Dt1* etabliert. Für die Allele *E1-as* und *E2-del* wurden außerdem KASP Marker entwickelt, welche eine schnelle Untersuchung dieser zwei entscheidenden Reifeloci erlauben. Nach erfolgreicher Etablierung aller 19 Marker wurden zuerst alle Eltern der neun spaltenden Populationen untersucht. Basierend auf diesen Ergebnissen wurden familienspezifisch die polymorphen Marker in 1850 Nachkommen detektiert.

Des Weiteren wurden 69 europäische Sorten, inklusive aller Elternlinien mittels whole-genome-sequencing mit einer 20-fachen Abdeckung sequenziert. Hierzu wurde DNA an die Firma Novogene geschickt, welche 350bp Illumina-Bibliotheken anfertigte und diese auf einem HiSeq-X Gerät sequenzierte. Pro Linie wurden mindestens 20GB rohe Sequenz-Daten generiert. Insgesamt wurden somit 1.4TB Daten generiert. Zur Datenanalyse wurde eine neue Next-Generation-Sequencing Analyse Pipeline an der LSA etabliert. Die Pipeline beinhaltet eine bioinformatische Bereinigung der rohen Daten (Adapter-Trimming, Duplicate removal), die Anordnung der sauberen Daten an das Soja Referenz-Genom (Alignment), die Detektion von SNP und InDel Polymorphismen (SNP Calling), die Aufarbeitung der Polymorphismen (Filtering) und die Schätzung von Mutationseffekten (SNP prediction). Es wurden für alle Schritte unterschiedliche Linux Software Pakete verwendet.

Kartenkonstruktion und QTL-Analyse

Die Erstellung der genetischen Karten für die QTL-Kartierungen erfolgte mit dem ASMap-Paket (Taylor and Butler 2017), das in dem Softwareprogramm R 3.2.0 (R Core Team 2015) implementiert ist. In einem ersten Schritt wurden falsch aufgerufene Allele in den Daten basierend auf flankierenden Markern mit der Funktion `correctStretches()` aus dem R-Paket `ABHgenotypeR` (Reuscher and Furuta 2016) korrigiert. Die maximale Länge der Bereiche, die von nicht-heterozygoten Stellen flankiert werden und verändert wurden, wurde auf fünf festgelegt. Vor der eigentlichen Kartenerstellung wurden die Datensätze auf fehlende Allele, genetische Klone, Marker mit übermäßiger Segregationsverzerrung und Co-Lokalisierungsmarker überprüft. Nach der Kartenerstellung mit der Funktion `mstmap()` wurde

die genetische Abfolge gegen die physikalische Abfolge aufgetragen und Marker, die ein stark abweichendes Muster zeigten, entfernt.

Die Quantitative Trait Locus (QTL)-Kartierung wurde mit einem Composite Intervall Mapping (CIM)-Ansatz mit einem in der Software PlabMQTL (Utz 2011) implementierten additiven genetischen Modell durchgeführt. Für die Co-Faktorauswahl wurde das mBIC (modifiziertes Bayes'sches Informationskriterium) verwendet. Wir führten Permutationstests für jedes Merkmal durch, um den empirischen Logarithmus der ungeraden (LOD) Schwellenwerte zu bestimmen, die einer genomweiten Fehlerrate von $\alpha \leq 0.10$ (Churchill und Doerge 1994) entsprechen.

Der Anteil der genotypischen Varianz (p_G), der durch die identifizierten QTL erklärt wird, wurde berechnet als $p_G = R_{adj}^2/h^2$ wobei R_{adj}^2 der angepasste Anteil der phänotypischen Varianz ist, der durch das Modell erklärt wird. Es wurde eine fünffache Kreuzvalidierung verwendet (Utz et al. 2000).

3.2b QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert

Pflanzenmaterial

Die Details sind ausführlich beschrieben in Kurasch et al. (2018a) und Kurasch et al. (2018b). Zur Ermittlung der Tofueigenschaften wurden 215 rekombinante Inzuchtlinien (RILs) mit ihren drei Elternsorten im Tofulabor der Taifun Tofu GmbH getestet. Die Linien stammen von zwei Kreuzungen der Tofusorte "Primus" (P2) mit entweder der proteinreichen (42 %) Sorte "Protina" (P3) oder der proteinarmen (38 %) Sorte "Sigalia" (P5). Die Familie P2×P3 bestand aus 129 Linien und die Familie P2×P5 aus 86 Linien. Zur Untersuchung des Quellvermögens wurde zusätzlich die Kreuzung zwischen P3×P5 einbezogen, diese bestand aus 103 Linien.

Quellvermögen

Die Sojaprobe von ca. 20 g wurden in einem Kunststoffbecher benetzt und durch ein Sieb gegossen. Das Sieb wurde dann dreimal am Rand der Spüle ausgeklopft, um eine gleichmäßige Benetzung der Proben zu gewährleisten. Die Proben wurden nach diesem Verfahren gewogen, um das Anfangsgewicht zu erhalten. Anschließend wurden sie dann wieder in ein mit Leitungswasser gefülltes Becherglas gelegt und für eine Stunde eingeweicht. Die Proben wurden auf die gleiche Weise wie oben beschrieben erneut durch ein Sieb gegossen und gewogen, um das Gewicht nach einer Stunde Einweichen zu erfassen. Dieser Vorgang wurde wiederholt und die Proben wurden 23 bis 25 Stunden lang eingeweicht und gewogen, nachdem sie durch das Sieb gegossen wurden. Die Gewichtszunahmen in g nach einer Stunde Einweichen (1h), die Gewichtszunahme nach weiteren 24 Stunden Einweichen (insgesamt 25h) und die Differenz zwischen dem Gewicht nach 25 h und 1 h wurden gemessen (24h) und für die nachfolgende Analyse verwendet.

Tofuqualität

Die Sojamilch- und Tofuzubereitung wurde 2014 im Tofulabor der Taifun Tofu GmbH in Freiburg durchgeführt. Die Laboranalyse wurde von vier verschiedenen Mitarbeiterteams an 21 Tagen durchgeführt. Die Sojaprobe von 80 g wurden 23 bis 25 h bei 8 °C in Wasser eingeweicht (Abbildung 3). Die eingeweichten Sojabohnen wurden entwässert und gewogen, um den Einweichfaktor zu berechnen und anschließend mit 480 ml Wasser gemahlen. Die produzierte Maische wurde dann durch einen Filterschlauch mit einer Maschenweite von 0,177 mm gefiltert, um die Sojamilch vom Rückstand zu trennen. Der Filterschlauch wurde von Hand geglättet und anschließend mit einer Ringmaschine gewrungen, wobei der Filterschlauch mit einem Drehmomentschlüssel auf ein Drehmoment von 3 Nm gedreht wurde. Die erhaltene Sojamilch wurde 3 Minuten lang bei 98,5 °C gekocht. Unmittelbar danach wurde die Dichte der Sojamilchproben mit einem Refraktometer (PR32 α , ATAGO Co., LTD, Tokyo,

Japan) in °Brix gemessen und durch Zugabe von Wasser auf 9 °Brix eingestellt. Die angepasste Sojamilch wurde gewichtet, um den Sojamilchertrag in g zu beurteilen (Abbildung 3). Zur Herstellung des Tofus wurde die gleiche Menge Sojamilch pro Probe verwendet (540 g), um eine gleichmäßige Menge an Koagulierungsmittel in jeder Probe zu gewährleisten. Das Koagulierungsmittel wurde im Voraus durch Lösen von 0,75 g dihydratisiertem CaSO₄ in 20,55 ml Lösung von 2,9545 % Nigari (MgCl₂ 6 H₂O) mit Leitungswasser hergestellt. Nach dem Abkühlen der Sojamilch auf 85 °C wurde das Koagulierungsmittel zugegeben und für 6 Sekunden homogenisiert. Nach 20 Minuten wurde der Bruch mit einem Quarkbrecher geschnitten, um den festen Anteil des Tofuquarks zu homogenisieren und die Molke aus dem Bruch auslaufen zu lassen. Zehn Minuten später wurde die Molke mit einem Sieb entfernt. Der restliche Tofuquark wurde in Formen mit einem Durchmesser von 7 cm mit Perforationen gegeben und 20 Minuten lang mit steigendem Druck von 0 bis 400 kPa gepresst. Nach dem Pressen wurde der Tofu für 23-25 Stunden bei 8 °C in kaltem Wasser gelagert. Der endgültige Tofu wurde entwässert und gewogen, um die Menge an Tofu in g zu bestimmen. Die Tofuausbeute wurde nach folgender Gleichung berechnet, die das Massenverhältnis von Tofu zu Sojabohnen korrigiert um den Anteil der eingesetzten Sojamilch:

$$Tofuausbeute = \frac{Tofu (g)}{trockene Bohnen (g)} \times \frac{Sojamilch (g)}{540 (g)}$$

Die Tofufestigkeit wurde mit einem Texturometer (Zwick AG, Ulm, Deutschland) in Newton mit einem Stempel von 100 mm Durchmesser und Proben mit einer Höhe von 27 mm, bei denen Ober- und Unterseite abgeschnitten wurden, gemessen und bis zu 30 % der Probenhöhe verformt.

Ein allgemeiner Tofuwert wurde mit der folgenden Gleichung berechnet:

$$Tofuwert = Tofuertrag^2 \times Tofufestigkeit (N)$$

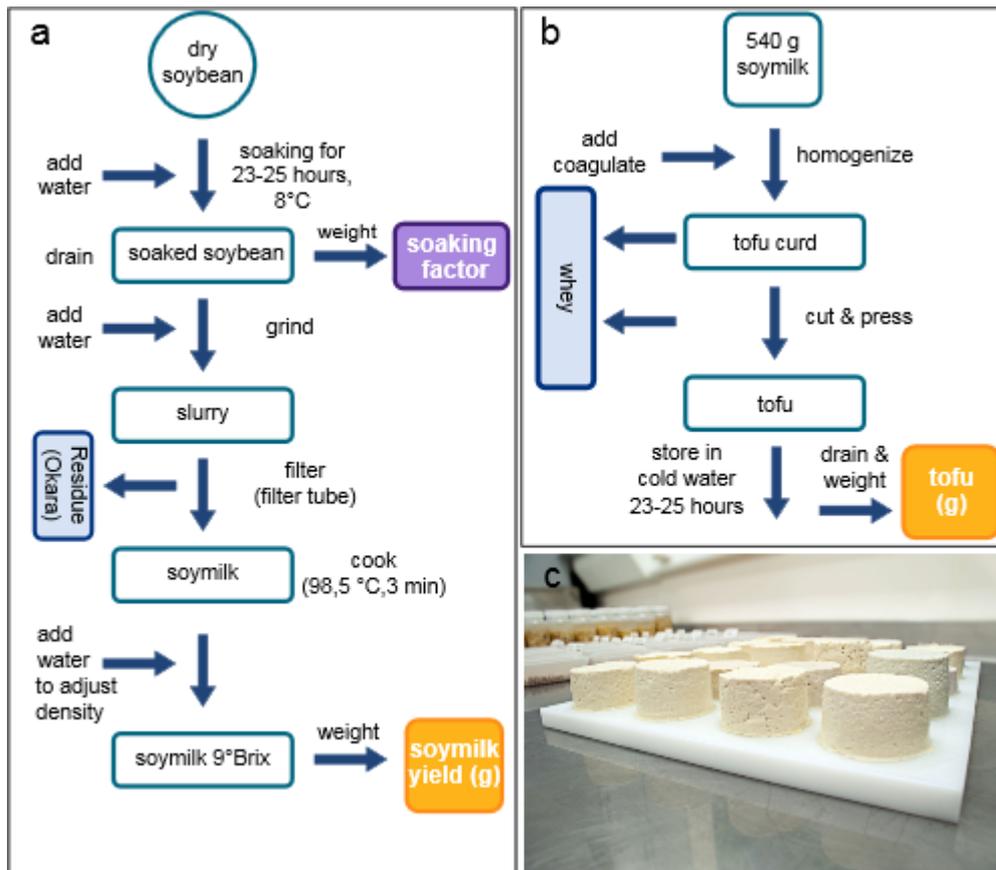


Abbildung 3: a) Schema der Produktionsschritte der Sojamilch und des endgültigen Tofus (b). c) Bild der im Tofulabor der Life Food GmbH erzeugten Tofuproben.

Genotypisierung, Kartenkonstruktion und QTL-Analyse

Siehe Kapitel 3.2a

3.2c Genomische Selektion

Die genomische Vorhersage wurde durch Ridge-Regression BLUP (RR-BLUP) mit dem in R (Endelman 2011) implementierten rrBLUP-Paket durchgeführt. Wir führten genomische Vorhersagen sowohl innerhalb einzelner Populationen für die Tofu-Eigenschaften und über Populationen hinweg für die Merkmale Ertrag, Proteingehalt, Ölgehalt, Pflanzenhöhe und Tausendkorngewicht durch. In beiden Szenarien wurde eine fünffache Kreuzvalidierung mit >100 Durchläufen durchgeführt, wobei die Datensätze in Schätz- (80%) und Vorhersagemengen (20%) unterteilt wurden. Bei den Merkmalen Protein- und Ölgehalt wurde zusätzlich zu den Vorhersagen innerhalb der 1008 Genotypen mit den Daten aus 2014 auch Schätzungen über Jahre und in weiteren 800 neuen Genotypen durchgeführt. Hierbei sollte ermittelt werden, wie gut die Vorhersagemodelle der 1008 Genotypen für unterschiedliche Jahre (2014 vs. 2016) und neue Nachkommen verwendbar sind.

Die Genauigkeit der genomischen Selektion wurde als $r_{GP} = r_{MP}/h$ ausgedrückt (Dekkers 2007; Lande und Thompson 1990), wobei r_{MP} sich auf die Korrelation zwischen beobachteten und vorhergesagten Phänotypen und h auf die Quadratwurzel der Erblichkeit bezieht.

3.2d,e Kreuzung und Weiterführung Diallel1_neu und Diallel2_neu

Die Kreuzungen wurden jeweils im Sommer im Feld oder in Folienhäusern durchgeführt, anschließend wurde das Material in den Winterzuchtgarten nach Costa Rica geschickt. Dort wurde es eine Generation weitergeführt und nach Deutschland zurückgesendet. Der Winterzuchtgarten 2016/17 wurde durch Hurrikan Otto stark geschädigt, so dass die Materialentwicklung nicht wie geplant durchgeführt werden konnte.

3.2f QTL-Kartierung der Kühletoleranz

Ist dargestellt in Kapitel 3.4

3.2g Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus AP1

Die Kreuzungen wurden im Sommer im Feld oder in Folienhäusern durchgeführt.

3.2h Sortenprüfungen

Die Auswahl der zu prüfenden Sorten wurde mit dem Sojaförderring abgesprochen. Die Leistungsprüfungen wurden in Hohenheim und Eckartsweier durchgeführt. Die Versuche wurden 4-reihig in Parzellen von 1,5m Breite und 6m Länge angebaut.

In 2015 wurden 25 Sorten geprüft. Diese wurde als 5x5 Gitter mit 2 Wiederholungen angebaut.

In 2016 wurden eine frühe und eine späte Prüfung durchgeführt. In der frühen Prüfung standen 25 Sorten, diese wurden als 5x5 Gitter mit 2 Wiederholungen angebaut. In der späten Prüfung wurden 16 Sorten geprüft. Diese Prüfung wurde als 4x4 Gitter angelegt.

In 2017 wurden 35 Sorten als 7x5 Gitter mit 2 Wiederholungen geprüft.

In 2018 wurden eine frühe Prüfung mit 20 Sorten als 5x4 Gitter und eine späte Prüfung mit 30 Sorten als 6x5 Gitter durchgeführt.

3.2i Speed-Breeding-System

Bei dem entwickelten Container handelt es sich um einen handelsüblichen, gebrauchten 40 Fuß Isoliercontainer. Dieser wurde mit einem Regalsystem ausgestattet, so dass auf 3 Ebenen Pflanzen angezogen werden können. Die Beleuchtung besteht aus BluRail 120 Lampen der Firma Growking aus Leinfelden-Echterdingen. Anfangs wurde die im Container verbaute Klimaanlage verwendet, um die Temperatur und Luftfeuchtigkeit einzustellen. Später wurde eine Klimaanlage des Typs SRK 71 ZR-W der Firma Mitsubishi Heavy Industries eingebaut. Als Erds substrat wird eine Mischung aus 60% Torf, 25% Mutterboden und 15% Perlit verwendet. Der pH-Wert des Bodens wird auf 7 eingestellt. Gegossen wird von Hand. Die Pflanzen für das Single Seed Descent- (ssd) Verfahren werden in 96er Platten angezogen. Die Pflanzen zur Durchführung von Kreuzungen werden in 2l Töpfen angezogen. Die Pflanzen werden für 8 Stunden beleuchtet, dabei beträgt die Temperatur 25°C, während der Dunkelzeit wird die Temperatur auf 20°C abgesenkt.

3.2k Ionomics

Zur Durchführung der energiedispersiven Röntgenfluoreszenzanalysen wurde ein S2 PUMA Röntgenspektrometer der Firma Bruker, Karlsruhe, verwendet. Die notwendige Röntgenstrahlung wird in einer Röntgenröhre erzeugt durch das Anlegen einer Hochspannung zwischen der Kathode und der Anode von 10 – 60 kV. Die so beschleunigten Elektronen treten aus einer aufgeheizten Wolframwendel, der Kathode, aus. Sie werden durch die angelegte Hochspannung beschleunigt und kollidieren mit der Palladium-Anode. Dadurch interagieren sie mit dem Palladium und geben einen Teil der Energie als Photonen ab. Diese elektromagnetischen Wellen werden mit einem energiedispersiven Röntgenhalbleiterdetektor erfasst. Anfangs wurde mit einer sogenannten standardlosen Messung gemessen. Diese wurde nach Erhalt der Referenzwerte durch Messungen mit einer Kalibration ersetzt. Die Referenzwerte wurden an der Core Facility der Universität Hohenheim ermittelt.

3.2l Genotypisierung

Für die Genotypisierung der Sojalinien wurde ein modifiziertes Protokoll basierend auf Elshire et al. (2011) entwickelt. Die wesentlichen Unterschiede zum Protokoll von Elshire et al. (2011) sind 1) die Verwendung von Sequenzierprimern mit Index um mehr als 96 Linien gleichzeitig zu sequenzieren, 2) eine zusätzliche Library-Size selection und 3) eine zusätzliche Aufreinigung des PCR-Produkts mit dem Enzym GenUP™ Exo SAP Kit von Biozym.

Durch die Verwendung von zusätzlichen Indexen an den Sequenzierprimern konnten mehr als 96 Linien gleichzeitig sequenziert werden. Für jede Platte mit 96 Proben wurde ein Index verwendet, diese wurden gepoolt, wodurch ein flexibles Multiplexing von bis zu 96 mal 94 Proben möglich ist, um bis zu 9024 Proben parallel sequenzieren zu können. Die Library-Size selection sowie die zusätzliche Enzymatische Aufreinigung wurden eingeführt, da das Protokoll von Elshire et al. (2011) für SE Sequencing optimiert ist. Da der zweite Read beim PE Sequencing jedoch eine relativ schlechte Qualität hat, wurden zuerst verschiedene Methoden getestet um die Größe der Sequenzierbibliothek einzuschränken. Die Size selection der Sequenzierbibliothek wurde 1) mit Magnetic Beads (ProNex von Promega), 2) mit Gel-Cutting und 3) dem BluePippin Gerät durchgeführt. Da keine dieser Techniken die Qualität des zweiten Reads stark verbesserte, wurde ein zusätzlicher enzymatischer Aufreinigungsschritt mit dem Enzym GenUP™ ExoSAP Kit von Biozym eingeführt. Der Verdau mit ExoSAP sollte noch vorhandene einzelsträngige DNA, welche die Library-Size selection beeinflussen kann, entfernen.

Die Sequenzierbibliotheken wurden bei der Firma Novogene (Cambridge, United Kingdom) auf einem Novaseq Sequenzierer mit PE150 sequenziert.

Da der zweite Read immer noch nicht die gewünschte Qualität hatte, wurde das aktuelle Protokoll nochmals mit anderen Protokollen verglichen um mögliche Ansatzpunkte für eine bessere Qualität beim PE Sequencing zu finden. Eine SE Sequenzierung ist nicht optimal, da bei gleichen Kosten weniger Daten generiert werden.

Für die bioinformatische Analyse wurde die Pipeline GB-eaSy (Wickland et al. 2017) für GBS Daten angewendet. Diese Pipeline beinhaltet übliche Tools für die Bearbeitung von NGS (next generation sequencing) Daten und ist sehr flexibel anwendbar. Das Demultiplexing und Adapter Trimmen der Reads erfolgte mit dem Programm GBSX (Herten et al. 2015). Das Mapping der Reads gegen das Referenzgenom Gmax Wm82.a2.v1 (Schmutz et al. 2010) erfolgte mit dem Programm BWA-MEM (Li 2013). Anschließend wurde das SNP Calling mit BCFtools (Li 2011) gemacht und die SNPs anhand von einer minimalen Read Depth sowie einem Grenzwert von fehlenden Daten (--max-missing) mit VCFtools (Danecek et al. 2011) gefiltert. Die einzelnen Schritte liefen parallel mit GNU um die volle Leistungsfähigkeit des Clusters zu nutzen und die Analyse zu beschleunigen.

3.2m Regionale Genomische Selektion

Für die regionale genomische Selektion war geplant, Ergebnisse der Landessortenversuche sowie Wertprüfungsergebnisse zu verwenden. Dies konnte nicht in ausreichendem Umfang durchgeführt werden, da die Daten nicht zur Verfügung gestellt wurden. Zwar sind einige Ergebnisse im Internet zu finden, hierbei handelt es sich jedoch immer um Mittelwerte, nicht um die Daten je Parzelle. Diese wären jedoch notwendig gewesen, um die statistischen Auswertungen durchführen zu können. Daher konnte nur auf eigene Ergebnisse zurückgegriffen werden. In Tabelle 2 sind die Umfänge an Leistungsprüfungen aufgeführt, die 2019 – 2021 an den verschiedenen Orten durchgeführt wurden. In 2020 war es durch die Corona-Restriktionen nicht möglich, Versuche außerhalb der Versuchsflächen der Universität Hohenheim durchzuführen.

Tabelle 2: Anzahl Prüfglieder und Parzellen an den angegebenen Orten in den Jahren 2019 bis 2021.

Jahr	Ort	Prüfglieder	Parzellen
2019	Eckartsweier	1050	1275
	Hohenheim	736	943
	Ihinger Hof	147	294
	Moosburg	49	98
	Niedertraubling	49	98
	Reichersberg	49	98
2020	Eckartsweier	1041	1264
	Hohenheim	937	1119
	Ihinger Hof	245	340
2021	Eckartsweier	790	1325
	Hohenheim	715	1175
	Ihinger Hof	275	350
	Moosburg	50	100
	Niedertraubling	25	50
	Reichersberg	100	200

Um die Grundlagen der phenomischen Selektion näher zu beleuchten, wurden einerseits Daten aus Versuchen aus 2014 erneut verrechnet. Zudem wurden Daten aus Leistungsprüfungen von 2017 bis 2020 verwendet.

3.3 IPM

3.3a,h Literaturarbeit, Primerdesign, Vortests

Die im Antrag gelisteten Pilze und Bakterien wurden mit der Hohenheimer Stammsammlung abgeglichen und nicht vorhandene Kulturen ergänzt. Zusätzlich wurden Pathogenproben aus infizierten Pflanzen isoliert. Die daraus extrahierte DNA wurde partiell sequenziert und die Sequenzen mit NCBI BLAST abgeglichen. Für das Primer- und Sondendesign wurden die Software BioEdit (BioEdit Sequence Alignment Editor Copyright 1997-2013 Tom Hall), der Server Primer3 (<https://primer3.ut.ee/>) und GeneRunner (Version 6.5.52.Beta, Copyright 1992-2019 Frank Buquicchio) verwendet. Bevor die Primer und Sonden bestellt wurden, wurden sie außerdem mit PrimerBLAST auf mögliche unerwünschte Targets überprüft (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome).

Vortests wurden mit den Pathogenen Psg, Ss, Fs, Ct, Ps und Rs durchgeführt. Für die Anpassung der pilzlichen Pathogene an *in planta* Versuche wurde zunächst die Virulenz der einzelnen Isolate überprüft. Hierzu wurden Blätter von Sojapflanzen in Petrischalen gelegt und mit zwei Ss Stämmen einem *S. minor* Stamm oder einem *Stromatinia gladioli* Stamm inokuliert. Der Pathogenitätstest wurde weiterhin für *Pythium debaryanum*, *P. ultimum*, *P. arrhenomanes* und *P. phragmitis* mit Stängelstücken und Blättern von Sojapflanzen durchgeführt.

3.3b,i Etablierung des quantitativen Nachweisverfahrens für verschiedene Pathogene

Zur Etablierung des qRT-PCR Detektionsverfahrens für *Sclerotinia* spp. (*Sclerotinia sclerotiorum*, *S. minor*, *Stromatinia gladioli*), wurde das Primerpaar qSsCL2fw, qSsCL2rv, welches auf der Calmodulin-Gensequenz basiert und ein 212 bp Amplikon erzeugt (Kim und

Knudsen 2008), verwendet. Der Nachweis wurde zunächst als Einzelnachweis über SYBR Green I realisiert. Dafür wurden unterschiedliche Annealingtemperaturen über einen Temperaturgradienten von 58,5 °C bis 68,5 °C getestet. Zusätzlich zu den Cq-Wert Auswertungen wurden die PCR Produkte per Gelelektrophorese auf Primer Dimere überprüft. Für *Rhizoctonia solani* und *Pythium ultimum* wurden das Primerpaar RsTqF1/RsTqR1 (Lees et al. 2002) und das Primerpaar qPyu92f/qPyu166r verwendet (Cullen et al. 2007). Die optimale Annealingtemperatur der Primerpaare wurde durch eine Gradienten-qPCR ermittelt. Es wurden acht verschiedene Temperaturen zwischen 58,5 °C und 68,5 °C getestet. Zusätzlich wurden die Produkte über die Agarose-Gelelektrophorese nach ihrer Größe aufgetrennt und auch hier wurde überprüft, ob es unspezifische Amplikons oder ein vermehrtes Vorkommen von Primer-Artefakten gab. Die Spezifität der Primer für *P. ultimum* wurde mit anderen *Pythium*-Arten (*P. debaryanum*, *P. arrhenomanes* und *P. phragmitis*) überprüft.

Auch für den Erreger des Bakterienbrandes *Pseudomonas syringae* pv. *glycinea* wurde ein Nachweisverfahren über SYBR Green I etabliert. Dafür wurde das spezifische Primerpaar (qPscyo2fw, qPscyo2rv) für die verschiedenen Pathovaren (*Pseudomonas syringae* pv. *glycinea*, *P. syringae* pv. *tomato*, *P. syringae* pv. *tabaci* und *P. syringae* pv. *savastanoi*) verwendet. Annealingtemperaturen von 58,5 °C bis 68,5 °C wurden getestet.

Nachweisverfahren wurden für *Sclerotinia* spp. und *Pseudomonas* spp., sowie *Fusarium* spp. entwickelt. Die bereits für SYBR Green etablierten Real-time Nachweise für *Sclerotinia* spp. und *Pseudomonas* spp. wurden dabei auf die TaqMan-Technologie umgestellt, um Multiplex Nachweise zu ermöglichen. Hierbei wurden für die Nachweise von *Sclerotinia* spp. und *Pseudomonas* spp. Primer und Sonden aus bestehenden Publikationen verwendet (Kim und Knudsen 2008; Xu und Tambong 2011).

Die Primer-/Sonden-Kombination (qSsCL2fw, qSsCL2rv, TMSCL2P) für Ss basiert auf der Calmodulin-Gensequenz und erzeugt ein 212 bp Amplikon (Kim und Knudsen 2008). Über eine Temperaturgradienten-Real-time PCR wurde die passende Annealingtemperatur bestimmt. Zudem wurde die Spezifität gegenüber *Diaporthe* spp, *Colletotrichum* spp., Fs und Rs überprüft.

Die Primer-/Sonden-Kombination für *Pseudomonas* spp. (qPscyo2fw, qPscyo2rv, Psy_cyoII-pb) basiert auf der Sequenz des Gens Cytochrom O Ubiquinol Oxidase Untereinheit II und erzeugt ein 176 bp Amplikon (Xu und Tambong 2011). Die Spezifität gegenüber *P. syringae* pv. *tomato*, Psg, *P. savastanoi* und andere *Pseudomonas*-Arten wie *P. phaseolicola* und *P. fluorescens*, sowie *Xanthomonas campestris* pv. *phaseoli* wurde getestet.

Um einen gattungsspezifischen TaqMan Nachweis für Fs zu etablieren, wurde DNA von zehn *Fusarium*-Isolaten extrahiert. Für die Gene β -Tubulin, ITS und Translations-Elongationsfaktor 1 α wurden PCRs durchgeführt und die gereinigten PCR-Produkte sequenziert. Die Gensequenzabschnitte wurden zusammen mit Sequenzen von *Diaporthe* spp. aligniert und so auf Bereiche untersucht, die innerhalb der Gattung *Fusarium* aber nicht zu anderen Gattungen konserviert sind. Die Effizienz der Primer und Sonden wurde bestimmt.

3.3c,k Etablierung des quantitativen Nachweisverfahrens für verschiedene Probenarten

Das im ersten Projektzeitraum etablierte qRT-PCR System für *Pseudomonas syringae* pv. *glycinea* wurde weiterentwickelt und für *in planta* Proben angepasst. Hierzu wurden zum einen die RNA symptomatischer Blätter extrahiert und zum anderen Blattstücke Oberflächen-desinfiziert und auf KB Medium gegeben. Die daraus wachsenden bakteriellen Kolonien wurden isoliert und die RNA extrahiert. Die RNA wurde zu cDNA umgeschrieben und mittels qRT-PCR getestet.

Die für Reinkulturen etablierten Nachweise von Ss, Fs und *Pseudomonas* spp. wurden für die Beprobung von Blättern, Stängeln und Saatgut angepasst.

Für die Anpassung an Blattproben wurden jeweils fünf 0,5 x 0,5 cm² große Blattstücke von natürlich infizierten und künstlich inokulierten Blättern entnommen und mit zwei Metallkugeln in Schraubdeckelgefäßen in einem FastPrep-24 Tissue and Cell Homogenizer (MP Biomedicals, Solon, USA) homogenisiert. Anschließend wurde mit dem DNeasy Plant Mini Kit (QIAGEN, Hilden) die DNA extrahiert und bis zur weiteren Verwendung bei -20 °C gelagert.

Für die Anpassung der pilzlichen Pathogene an *in planta* Versuche wurde zunächst die Virulenz der einzelnen Isolate überprüft. Hierzu wurden Blätter von Sojapflanzen in Petrischalen gelegt und mit zwei Ss Stämmen (DSM, IZS), einem *S. minor* Stamm oder einem *Stromatinia gladioli* Stamm inokuliert.

Der Pathogenitätstest wurde weiterhin für *P. debaryanum*, *P. ultimum*, *P. arrhenomanes* und *P. phragmitis* mit Stängelstücken und Blättern von Sojapflanzen durchgeführt.

Um die Funktionalität des qRT-PCR-Nachweissystems für Ss an Sojapflanzen zu prüfen, wurde ein Gewächshausversuch durchgeführt. Dazu wurden Sojapflanzen im 3. Laubblattstadium jeweils ein mit Ss bewachsener Zahnstocher in die Blattachsel des zweiten Laubblattes gesteckt. Um eine hohe Feuchtigkeit zu gewährleisten wurden die Pflanzen in Polyethylen-Folie eingehüllt. Anschließend erfolgte die Probenahme (1 cm Abschnitte vom Stängel jeweils unter- und oberhalb der Inokulationsstelle) nach 0, 3, 6 und 9 dpi (Abbildung 4). Die Proben wurden in einem

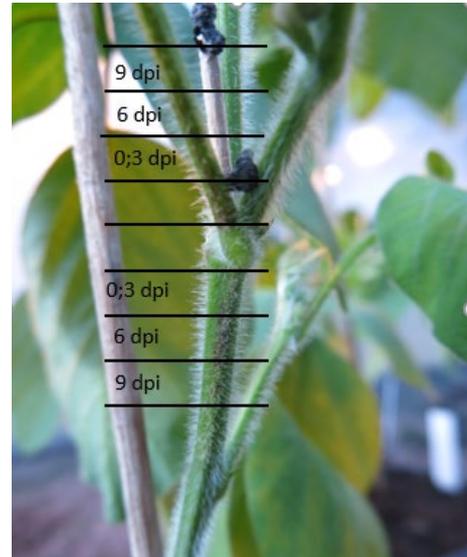


Abbildung 4: Inokulationsverfahren und Probenahme bei mit *S. sclerotiorum* inokulierten Sojapflanzen.

FastPrep-24 homogenisiert und die DNA wie oben beschrieben extrahiert. Anschließend wurden die Proben bis zur weiteren Verwendung bei -20 °C gelagert.

3.3d,l Etablierung einer quantitativen Methodik zur Erfassung von Nematoden

Für die Arbeit mit Nematoden wurden Prof. Dr. Johannes Hallmann und Dr. Holger Heuer vom JKI kontaktiert und bei einem gemeinsamen Treffen beschlossen, zunächst ein allgemeines Detektionssystem für den Nematoden *Pratylenchus penetrans* zu etablieren.

In einem ersten Arbeitsschritt wurden Bodenproben von sieben landwirtschaftlichen Betrieben genommen. Dabei handelte es sich um zwei Standorte der Versuchsstation Agrarwissenschaften der Universität Hohenheim, drei konventionelle Betriebe und zwei ökologisch wirtschaftende Betriebe. Beprobte wurde für jeden Betrieb jeweils ein Schlag, auf dem im Jahr 2015 Soja angebaut wurde und ein Schlag, auf dem 2016 Soja angebaut werden sollte. Die Proben wurden zur konventionellen Quantifizierung zum JKI nach Münster geschickt.

3.3e,m Etablierung des Multiplex-Systems auf verschiedenen Proben

Ausgehend von den bereits etablierten Einzelnachweisen (siehe 3.3.b.) soll idealerweise ein universell für alle Probentypen einsetzbarer Multiplexnachweis etabliert werden, der die Pathogene Psg, Ss, Rs und *Pythium ultimum* umfasst. Um den Multiplexnachweis zu etablieren, sollten die Pflanzen mit den bestimmten Pathogenen infiziert werden, dazu waren Optimierungsversuche zur erfolgreichen Inokulation mit verschiedenen Pathogenen notwendig (siehe 3.3c). Zusätzlich zu den etablierten Einzelnachweisen wurden

Duplexnachweise für Ss und Fs, sowie für Ss und *Pseudomonas* spp. etabliert. Die Duplexnachweise wurden mit 1:10 Verdünnungen und unverdünnten Reinkulturen getestet.

3.3f,n Testung verschiedener Proben

Es wurden verschiedene Probentypen untersucht. Natürlich bzw. künstlich infizierten Blätter-, Stängel- und Saatgutproben wurden mit den bereits etablierten Einzelnachweisen auf möglichen Pathogenbefall überprüft.

Zusätzlich zur Testung verschiedener Probentypen mittels Einzelnachweise, wurde eine Testung von Saatgutproben im Duplex-Verfahren von Ss und Fs durchgeführt. Hierzu wurden 19 DNA-Proben des Kultivars Primus 6 und 20 DNA-Proben des Kultivars Sultana 94 mittels TaqMan Real-time PCR analysiert.

3.4 JKI-RS

3.4a Gefäßversuche unter kontrollierten Bedingungen zur Phänotypisierung des Hülsenansatzes unter Kühlestress

Im Gefäßversuch wurden 2014 bis 2016 122 Linien aus der Kreuzung Merlin x Sigalia, die Eltern sowie der kühesensitive Standard Malaga geprüft.

Zur Untersuchung des Hülsenansatzes unter Kühlestress wurden je 6 Pflanzen pro Genotyp in 5 unabhängigen Wiederholungen bei 25 °C/ 20 °C Tag/Nacht in der Klimakammer bei 10 h Licht, Lichtstärke ca. 20 klux /14 h Dunkelphase bis zum Blühbeginn angezogen (ca. 30 Tage nach der Aussaat). In diesem Zeitraum wurden die nicht inokulierten Pflanzen 2 x im wöchentlichen Abstand mit 30 ml einer 0,3 % Wuxallösung und dann mit 20 ml einer 0,175 % Hakaphos-Rot Lösung gedüngt. Mit Blühbeginn wurden 4 Pflanzen je Genotyp und Wiederholung einer 3-wöchigen Kühlestressphase bei 15 °C / 10 °C Tag/Nacht und 12 h Licht ausgesetzt. Zwei Pflanzen je Genotyp und Wiederholung wurden ebenfalls bei 12 h Licht unter Kontrollbedingungen bei 25 °C/ 20 °C Tag/Nacht weiter kultiviert. Dreimal wöchentlich wurden die Pflanzen auf 70 % der maximalen Wasserkapazität des Bodens gegossen, um den Einfluss einer unterschiedlichen Wasserversorgung auf die Kühletoleranz auszuschließen. Die Messung des Chlorophyllgehalts mittels SPAD Meter (SPAD 502Plus, Konica-Minolta) wurde am Ende der Stressphase für Kontrolle und Stressbehandlung am jüngsten voll ausgewachsenen Blatt jeder Pflanze vorgenommen. Jeder Messwert ist der Mittelwert aus 15 Einzelmesspunkten des gemessenen Blattes. Ebenfalls am Ende der Stressphase wurde der Hülsenansatz am Nodium erfasst.

Feldversuch – Hülsenansatz und Prüfanbau

Ein Feldanbau der Linien und Eltern wurde 2016 auf dem Versuchsfeld des Standortes Groß Lüsewitz des JKI durchgeführt. Dabei handelt es sich um einen diluvialen Standort mit lehmigem Sand und einer Bodenzahl von 47, 692 mm Niederschlag im langjährigen Mittel und einer durchschnittlichen Jahrestemperatur von 8,3 °C.

Wärmesumme, Niederschlag und Globalstrahlung für die Vegetationszeit 2016 wurden für die für den Anbau von Sojabohnen relativ sichere nutzbare Vegetationszeit vom 01. Mai bis 30. September berücksichtigt (Daten von der Universität Rostock, Agrarwissenschaftliche Fakultät, Institut für Umweltingenieurwesen). Temperatur und Niederschlag sind in Abbildung 5 wiedergegeben. Die Wärmesumme wurden mittels der kanadischen Methode der Crop Heat Units (CHU) (Brown und Bootsma 1993) ermittelt. Das CHU-System berücksichtigt im Gegensatz zu der sonst gebräuchlichen Wärmesumme auf Basis 6 °C auch die für Soja u.U. kritischen Nachttemperaturen und reduziert das Gewicht der Maximaltemperaturen von Hitzetagen bis hin zu negativen Werten. Die CHU Wärmesumme für 2016 betrug 2971 °C und

die Niederschlagssummen im gleichen Zeitraum 175 mm, damit handelte es sich um einen warmen und sehr trockenen Sommer. Die Globalstrahlung belief sich auf 32726 W/m². Die Anlage des Feldanbaus erfolgte im Blockdesign mit 2 Wiederholungen und Einzelreihenparzellen a 1,25 m. Da über die Abreife der Linien keine Informationen vorlagen, wurden die Blöcke vollständig randomisiert.

Ein Überblick über die wichtigsten agronomischen und phänologischen Daten findet sich in Tabelle 3. Erfasst wurden Blühbeginn und Blühdauer sowie Abreife, der Chlorophyllgehalt indirekt als SPAD Wert vor Beginn der Abreife Anfang August, Blattform (nur Typ) sowie Wuchshöhe zur Abreife. An 10 Pflanzen pro Parzelle wurden der Hülsenansatz am Nodium, der Ertrag sowie das TKG ermittelt.

Die Messung des Chlorophyllgehalts mittels SPAD Meter (SPAD 502Plus, Konica-Minolta) wurde am jüngsten voll ausgewachsenen Blatt jeder Pflanze vorgenommen. Der Messwert der Parzelle setzt sich aus 30 Messpunkten an 6 Blättern (Pflanzen) zusammen. Die Reifebonitur erfolgte ab den ersten Seneszenzerscheinungen Ende August bis zum Beginn der Ernte wöchentlich auf einer Skala von 1 (komplett grün) bis 9 (komplett braun und druschreif). Die Ernte der 10 Einzelpflanzen erfolgte manuell nach Abreife der Linien. Alle Erträge und Tausendkorngewichte sind auf der Bezugsbasis von 14 % Feuchte angegeben. Saatgut zur Bestimmung des Protein- und Ölgehaltes mittels NIRS wurde an die Landessaatzuchtanstalt der Universität Stuttgart Hohenheim weitergegeben.

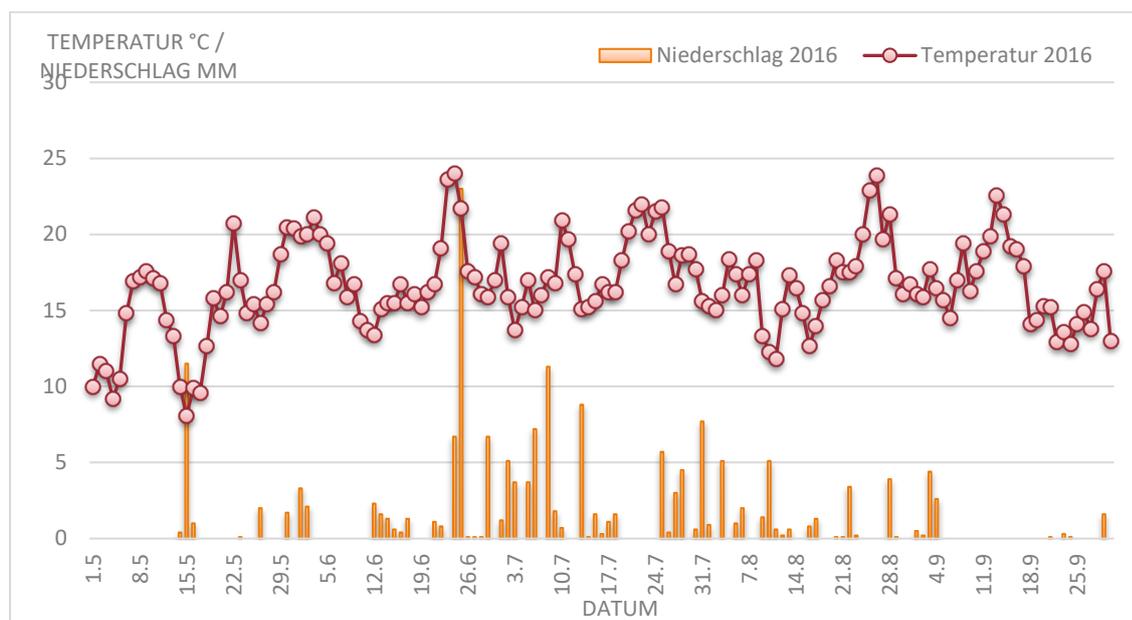


Abbildung 5: Temperatur und Niederschlag für die Hauptvegetationszeit vom 1.5. bis 30.9.2016 am Standort Groß Lüsewitz.

Tabelle 3: Überblick über die agronomischen Eckdaten 2016.

2016	
Inokulat	Rhizoliq Top S
Aussaat	10.5.
Herbizidbehandlung	Stomp Aqua VA 11.5.
Auflauf	20.5.
Blühbeginn (Spanne)	20.6. - 7.7.
Abreife (Spanne)	7.9. - 19.10.
Ernte	individuell nach Abreife

Statistik (JKI)

Alle Daten wurden mit SAS 9.4 (SAS Institute Inc.) varianzanalytisch verrechnet (Prozedur GLM), die Berechnung der kleinsten signifikanten Differenz (LSD) erfolgte nach Newman-Keuls. Für die Korrelationen wurde der Korrelationskoeffizient nach Pearson bestimmt. Die Daten für die Berechnungen der Heritabilität wurden mit der Prozedur VARCOMP errechnet.

QTL Analyse (LSA)

Die Nachkommenschaft der ‚Sigalia x Merlin‘-Population sowie die beiden Eltern wurden mithilfe eines Genotype-by-sequencing-Ansatzes genotypisiert. Der daraus generierte SNP (single-nucleotide-polymorphism)-Datensatz wurde mit TASSEL (Version 5.2.13, (Bradbury et al. 2007) imputiert und monomorphe Marker entfernt. Die genetische Karte wurde mit dem R-package ASMap (Taylor und Butler 2017) aus den verbleibenden 7711 polymorphen Markern generiert.

Das Modell, welches für die die statistische Analyse des Datensatzes genutzt wurde lautet:

$$Y_{ijkl} = \mu + g_i + t_j + (gt)_{ij} + sb_k + sb_k:pb_l + \varepsilon_{ijkl} \quad [1]$$

dabei gilt

Y_{ijkl} =	beobachteter Wert des Nachkommens i während der Behandlung j im Superblock k im Teilblock l
μ =	allgemeiner Mittelwert
g_i =	Effekt des Nachkommens i
t_j =	Effekt der Behandlung j
$(gt)_{ij}$ =	Interaktionseffekt des Nachkommens i in der Behandlung j
sb_k =	Effekt des Superblocks k
$sb_k:pb_l$ =	Effekt des Teilblocks l im Superblock k
ε_{ijkl} =	Fehler von Y_{ijkl} .

Alle statistischen Analysen wurden mit dem ASReml 3.0 (Gilmour et al. 2009) package, welches mithilfe von R (R Core Team 2015) genutzt wurde, durchgeführt. Varianzkomponenten und die BLUEs der Nachkommen (Best Linear Unbiased Estimates) wurden getrennt für die beiden Behandlungen errechnet, wobei sich das Modell [1] dann dahingehend modifiziert, dass der Effekt der Behandlung t_j sowie der Interaktionseffekt $(gt)_{ij}$ entfallen und alle Modellkomponenten als zufällig eingestuft wurden. Für die Berechnungen wurde die REML-Methode (restricted maximum likelihood) benutzt. Die Heritabilitäten wurden nach Hallauer und Miranda (1981) bestimmt:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{R}} \quad [2]$$

dabei gilt:

σ_g^2 = genotypische Varianz

σ_e^2 = Fehlervarianz

R = Anzahl der kompletten Wiederholungen

Die QTL-Analyse erfolgte mit PLABQTL (Utz, H.F., Melchinger, A.E. 1996) und wurde 100fach kreuzvalidiert.

Gefäßversuche unter kontrollierten Bedingungen zur Erfassung des Hülsenansatzes unter Kühlestress (JKI)

Die vorliegenden Untersuchungen basieren auf einem Test, der im Rahmen des Projektes 'Ausweitung des Sojaanbaus in Deutschland durch züchterische Anpassung und pflanzenbauliche Optimierung' (2811NA007) entwickelt und beschrieben wurde. Im Gefäßversuch wurden 2019 bis 2021 mehr als 400 Akzessionen eines breiten Sortiments (LSA) für die Tests angezogen.

Geplant war, 300 Prüfglieder unter Kontroll- und Stressbedingungen und 100 weitere aus Kapazitätsgründen nur unter Stressbedingungen zu prüfen. Auf Grund von Problemen mit der Keimfähigkeit und der Qualität von Jungpflanzen wurden in der Bilanz 419 Prüfglieder (auch mehrfach) angezogen, von denen 72 nicht oder unzureichend keimten, so dass 347 Akzessionen erfolgreich in 2 Wiederholungen getestet wurden, davon 250 mit Kontrolle und 97 ohne Kontrolle. Eine Liste des geprüften Materials befindet sich im Anhang 1.

Für die erste Variante wurden je 5 Pflanzen (3 Stress/2 Kontrolle) pro Akzession, für die zweite Variante 3 Pflanzen (3 Stress) in jeweils 2 unabhängigen Wiederholungen bei 25 °C/20 °C Tag/Nacht in der Klimakammer bei 10 h Licht, Lichtstärke ca. 20 klux/14 h Dunkelphase bis zum Blühbeginn (ca. 30 Tage nach der Aussaat) angezogen. In diesem Zeitraum wurden die nicht mit Bradyrhizobien inokulierten Pflanzen 2 x im wöchentlichen Abstand mit 30 ml einer 0,3 % Wuxallösung und dann mit 20 ml einer 0,175 % Hakaphos-Rot Lösung, gedüngt. Mit Blühbeginn wurden 3 Pflanzen je Akzession und Wiederholung einer 3-wöchigen Kühlestressphase bei 15 °C/10 °C Tag/Nacht und 12 h Licht ausgesetzt. Zwei Pflanzen je Akzession und Wiederholung wurden – soweit eine Kontrolle vorgesehen war – ebenfalls bei 12 h Licht unter Kontrollbedingungen bei 25 °C/ 20 °C Tag/Nacht weiter kultiviert. Die Pflanzen wurden täglich gegossen, dabei mindestens dreimal wöchentlich auf 70 % der maximalen Wasserkapazität des Bodens eingestellt, um den Einfluss einer unterschiedlichen Wasserversorgung auf die Kühletoleranz zu minimieren.

Am Ende der Stressphase wurde der Hülsenansatz am jeweiligen Nodium erfasst.

Datenauswertung (JKI)

Alle Daten wurden mit SAS 9.4 (SAS Institute Inc.) varianzanalytisch verrechnet (Prozedur GLM); die Berechnung der kleinsten signifikanten Differenz (LSD) erfolgte nach Newman-Keuls. Für die Korrelationen wurde der Korrelationskoeffizient nach Pearson bestimmt. Die Minimum Variance Quadratic Unbiased Estimators (MIVQUEs) der Varianzkomponenten wurden für die Berechnung der Heritabilität nach Becker (2011) genutzt.

Genomweite Assoziationsstudie (LSA)

Die verwendeten Genotypen wurden resequenziert. Für die GWAS-Analyse wurden 1,45 Mio SNPs verwendet. Die in dem Softwarepaket Tassel implementierte GBS-Analysepipeline wurde für das Auffinden von SNPs verwendet (Glaubitz et al. 2014). Die Assoziationsanalyse

zwischen SNP-Markern und Merkmalen erfolgte mit GenABEL unter Verwendung eines linearen gemischten Modells. Alle signifikant assoziierten Marker eines Merkmals wurden in ein lineares Modell in der Reihenfolge der Stärke ihrer Assoziation eingepasst, d. h. der am stärksten assoziierte Marker wurde zuerst eingepasst. Der Anteil der genotypischen Varianz, der durch alle signifikant assoziierten SNPs erklärt wird, wurde als das Verhältnis zwischen dem bereinigten R^2 dieses linearen Modells und der Heritabilität des Merkmals $pG=R^2_{adj}/h^2$ geschätzt.

4. AUSFÜHRLICHE DARSTELLUNG DER WICHTIGSTEN ERGEBNISSE

4.1 IPSP

Die Darstellung konzentriert sich hier auf die Arbeitspakete AP1c, AP1d, AP1g, AP1h und AP1i.

4.1c Allelemining und Identifizierung züchterisch interessanter Akzessionen:

Die Charakterisierung der derzeitigen und zukünftigen (im Jahr 2070) klimatischen Bedingungen in mitteleuropäischen Sojaanbauumwelten (Zielumwelt) zeigte, dass der Sojaanbau in Zukunft als wesentliche Veränderung mit einem höheren Temperaturangebot bei niedrigeren Niederschlagsmengen im Anbauzeitraum konfrontiert werden wird (siehe Abbildung 6).

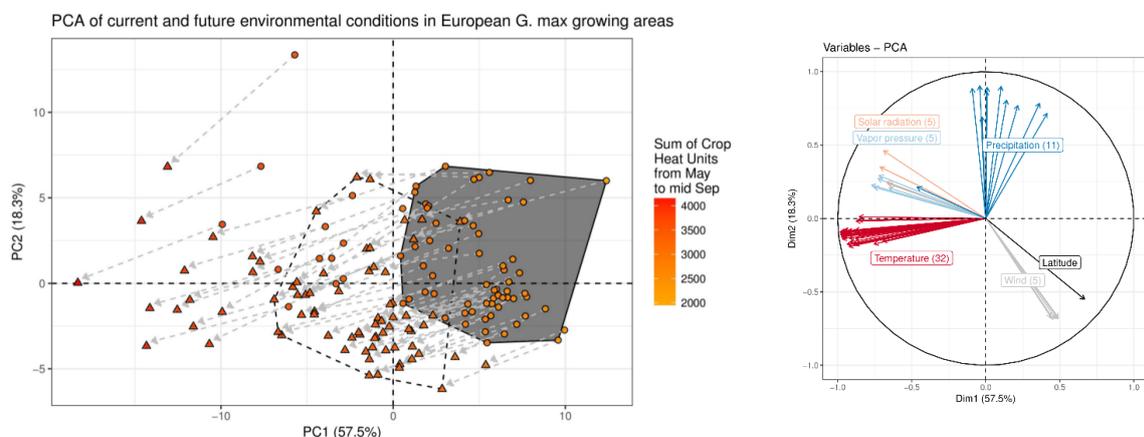


Abbildung 6: Hauptkomponentenanalyse für europäische Sojaumwelten mit Umweltdaten für heutige (Kreise) und zukünftige (Dreiecke) Bedingungen. Pfeile kennzeichnen die Verschiebungen aufgrund sich ändernder Klimabedingungen. Polygone kennzeichnen die Bandbreite mittel europäischer Anbauumwelten (heute und in 2070). Symbole außerhalb der Polygone kennzeichnen südeuropäische Umwelten. **Rechts:** Korrelationszirkel der Umweltvariablen für die erstellte Hauptkomponentenanalyse.

Der Abgleich der mitteleuropäischen Umweltprofile mit den Daten für ostasiatische Sojaherkünfte (Spenderumwelt) ergab, dass sich der Großteil der 689 georeferenzierten Herkünfte deutlich von den mitteleuropäischen Verhältnissen unterscheidet. 123 Herkünfte waren jedoch in Hinblick auf Temperatur- und andere Klima-/Umweltparameter hinreichend vergleichbar mit der derzeitigen und/oder zukünftigen Zielumwelt und umfassten 523 Akzessionen mit potentieller Umwelthanpassung an Mitteleuropa (siehe Abbildung 7 und Abbildung 8). Diese stammen vornehmlich aus dem Nordosten Chinas, dem Norden Japans, dem Himalaya, sowie aus Russland.

Die Daten bestätigten außerdem die Hypothese, dass frühreifes Material aus Umwelten stammt, die der Zielumwelt ähneln. Die bereits erfolgte Vorauswahl züchterisch interessanter Akzessionen wurde daher zusätzlich durch als frühreif klassifiziertes Material (Reifegruppe 000 – I) ergänzt, auch wenn keine Herkunftsinformationen vorlagen. Insgesamt wurden somit aus 17.017 Sojaakzessionen 3.663 Akzessionen aufgrund ihrer potentiellen Umwelteignung

und / oder Reifegruppenzugehörigkeit als vielversprechende Kandidaten für die mitteleuropäische Sojazüchtung ausgewählt.

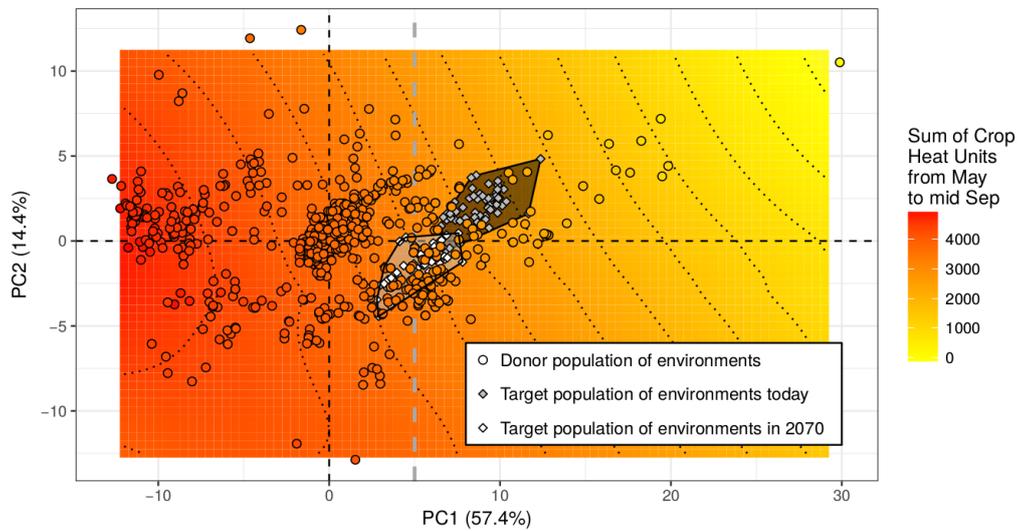


Abbildung 7: Hauptkomponentenanalyse der Sojaherkünfte mit Umweltdaten. Die erste Hauptkomponente separiert Herkünfte entlang eines Temperaturgradienten. Derzeitige (graue Rauten) und zukünftige (weiße Rauten) klimatische Bedingungen in mitteleuropäischen Anbaumwelten sind zur Illustration mit dargestellt. Polygone kennzeichnen die Bandbreite mitteleuropäischer Anbaumwelten (heute und in 2070).

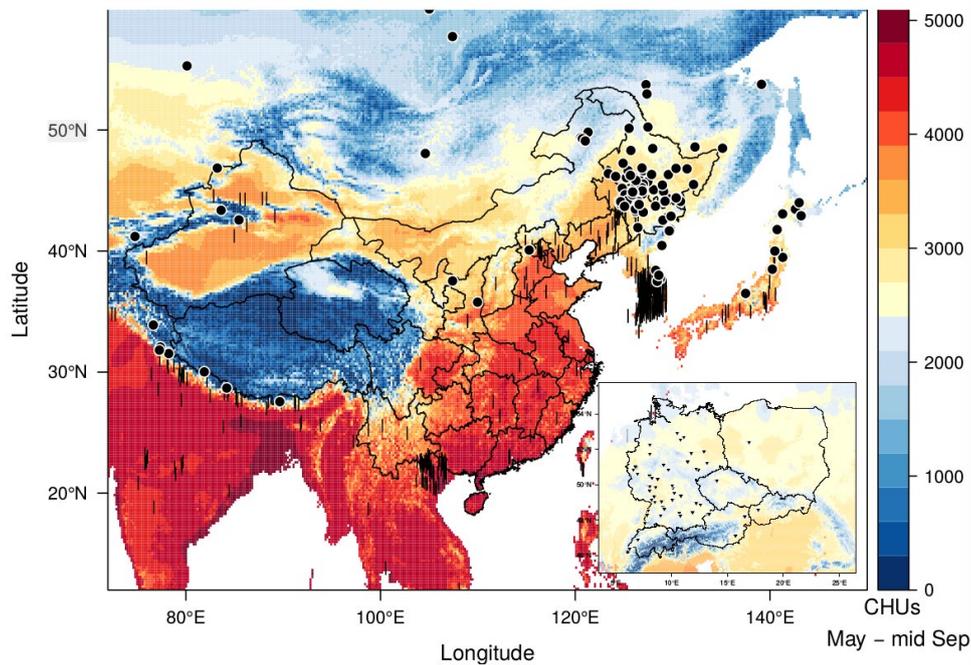


Abbildung 8: Herkünfte des Sojamaterials in der USDA Genbank in der Spenderumwelt. Punkte kennzeichnen Herkünfte die als vielversprechende Spender von Genbankmaterial für die mitteleuropäische Sojazüchtung identifiziert wurden. Vertikale Striche kennzeichnen Herkünfte, die nicht ausgewählt wurden. Farbskala repräsentiert die Summe an Crop Heat Units die im Zeitraum Mai bis Mitte September erreicht wird. **Einsatz:** Derzeitige Zielumwelt zum Vergleich. Dreiecke zeigen die Ausdehnung des Sojaanbaus basierend auf den Standortdaten für Sortenversuche.

Eine genomweite Differenzierungsanalyse in der die 3.663 Kandidaten den nicht ausgewählten Akzessionen gegenübergestellt wurden, offenbarte Genombereiche mit deutlichen Unterschieden zwischen den beiden Gruppen. Einige der deutlichsten Signale entfielen hier auf bereits bekannte Gene die für Umwelanpassung verantwortlich zeichnen (siehe Abbildung 9). Dies zeigt, dass unsere Auswahl insbesondere mit Hinblick auf die Auswahl von Akzessionen mit potentieller Umwelanpassung an mitteleuropäische Bedingungen erfolgreich war. Eine Vielzahl von signifikanten Signalen entfiel darüber hinaus auf weitere Genombereiche die in der Sojabohne bisher weniger prominent untersucht worden sind, die aber beispielsweise Transkriptionsfaktoren aus Gruppen enthalten, für die eine Rolle in der Anpassung an abiotische Bedingungen in Soja bereits berichtet wurde.

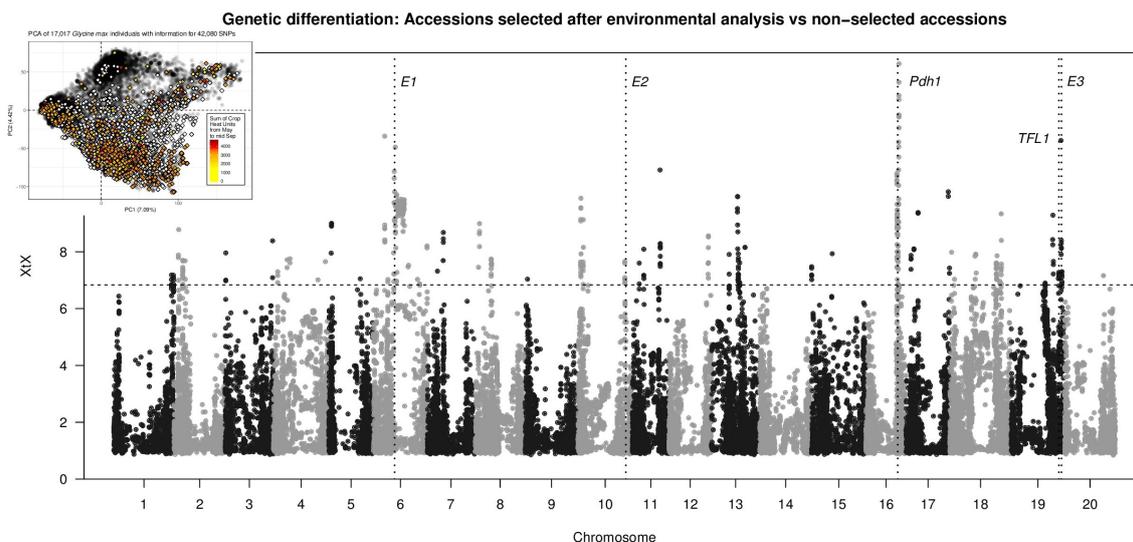


Abbildung 9: Die BAYPASS Analyse identifizierte Bereiche mit genetischer Differenzierung zwischen den 3.663 Akzessionen, die nach der Umweltanalyse als vielversprechend für Mitteleuropa angesehen werden, und dem übrigen Genbankmaterial. Selektionssignale entfallen u.a. auf bekannte Loci für Umwelanpassung (E1-E3, Pdh1, TFL1). Einsatz: Hauptkomponentenanalyse mit genetischen Daten. Die Umweltauswahl der 3.663 Akzessionen ist farblich gekennzeichnet, nicht ausgewähltes Material in grau.

Die Kandidatengruppe der 3.663 Akzessionen wurde als Grundstock zur Bildung einer Kernkollektion genutzt. Diese soll unter Wahrung größtmöglicher genetischer und phänotypischer Diversität eine Repräsentation des potentiell an mitteleuropäische Verhältnisse angepassten Sojapools darstellen. Die Untersuchung verschiedener Optimierungsstrategien zeigte, dass die Maximierung der genetischen Distanz zwischen den Kernakzessionen für unsere Zielvorgabe und unsere Datengrundlage die besten Ergebnisse lieferte und sehr erfolgreich darin war, redundante Einträge zu entfernen um diverse Kerne zu erhalten (siehe Abbildung 10). Daher etablierten wir unsere Kernkollektion(-en) in der Folge nur noch auf dem Parameter genetische Distanz zwischen Akzessionen.

Das Sampling von Akzessionen zu Kernkollektionen ohne Stratifizierung, d.h. der Behandlung der Kandidatengruppe als eine Gruppe gleichberechtigter Kernkandidaten, zeigte, dass insbesondere frühreifes Material in entsprechenden Kollektionen unterrepräsentiert wurde (siehe Abbildung 11A), da frühreife Akzessionen im Durchschnitt eine geringere genetische Distanz zueinander aufweisen als dies in anderen Reifegruppen der Fall ist. Frühreife spielt für die mitteleuropäische Sojazüchtung allerdings eine besondere Rolle. Daher wurden alternative Samplingstrategien erprobt, bei denen frühes Material bevorzugt wurde.

Unter diesen war insbesondere die 3-fache Pseudostratifizierung erfolgreich darin, alle Reifegruppen im Verhältnis zu den Proportionen in der Ausgangsgruppe zu konservieren (siehe Abbildung 11A), ohne Abstriche bei der Maximierung der genetischen Distanz zu machen (siehe Abbildung 11B).

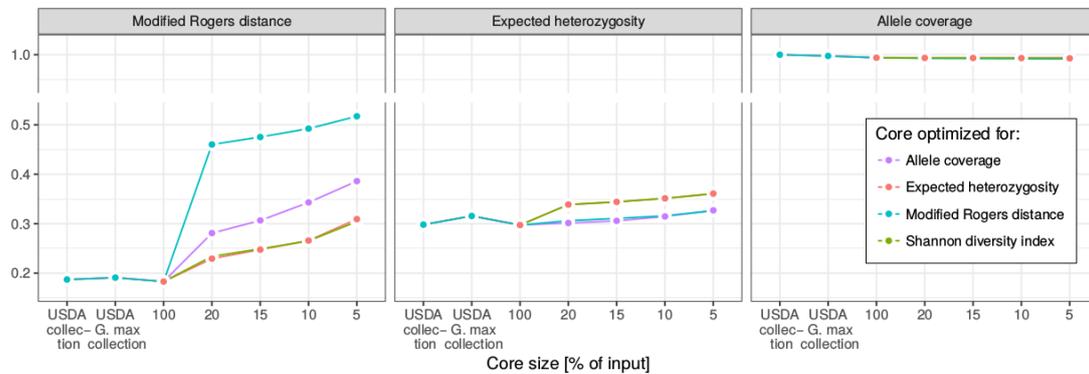


Abbildung 10: Genetische Distanz, Allelische Diversität und Allel Abdeckung in Kernkollektionen die mit verschiedenen Optimierungsstrategien gebildet wurden. „100“ ist die Ausgangsgruppe der 3.663 Akzessionen. Die Optimierung der genetischen Distanz zwischen Akzessionen auf Grundlage der Modifizierten Rogers Distanz gewährleistete die Vermeidung von ähnlichem Material in Kernlösungen am effektivsten.

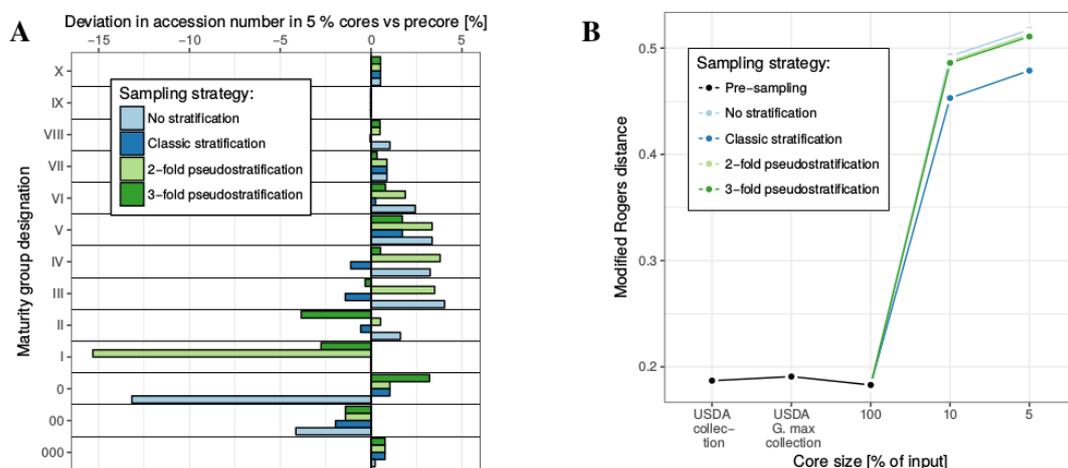


Abbildung 11: A: Evaluierung verschiedener Samplingstrategien für die Konservierung der Reifegruppenproportionen bei der Zusammenstellung von Kernkollektionen B: und die Maximierung der genetischen Distanz zwischen Kernakzessionen. Im Vergleich erzielte die 3 fache Pseudostratifizierung insbesondere durch die Konservierung frühreifen Materials bei hoher Maximierungsleistung das beste Ergebnis.

Für die Bildung von möglichst diversen Kernkollektionen bestehend aus Akzessionen mit Nutzungspotential für die mitteleuropäische Sojazüchtung wurde so nach umfangreicher Abwägung verschiedener Optimierungs- und Samplingstrategien die Optimierung (d.h. Maximierung) der genetischen Distanz zwischen Kernakzessionen und das Sampling dieser gemäß der 3-fachen Pseudostratifizierung als beste Vorgehensweise identifiziert. Zwei Kernkollektionen, die jeweils fünf bzw. 10 % der Akzessionen der Kandidatengruppe

umfassten (d.h. 183 bzw. 366 Akzessionen), wurden zusammengestellt. Die Analyse der vorliegenden genotypischen und phänotypischen Daten ergab, dass beide Kernkollektionen die Kandidatengruppe gut repräsentieren (siehe Abbildung 12 und Abbildung 13) und können von Wissenschaftlern und praktischen Züchtern für eine effiziente Suche nach nützlicher Variation (z.B. Kühletoleranz) genutzt werden.

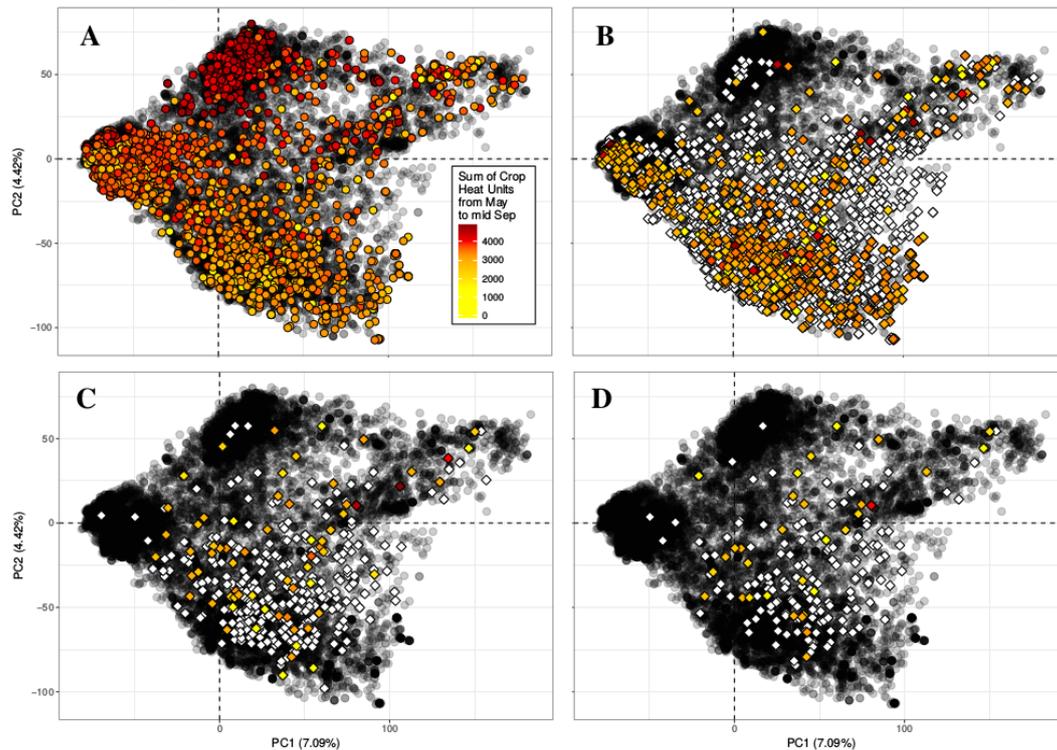
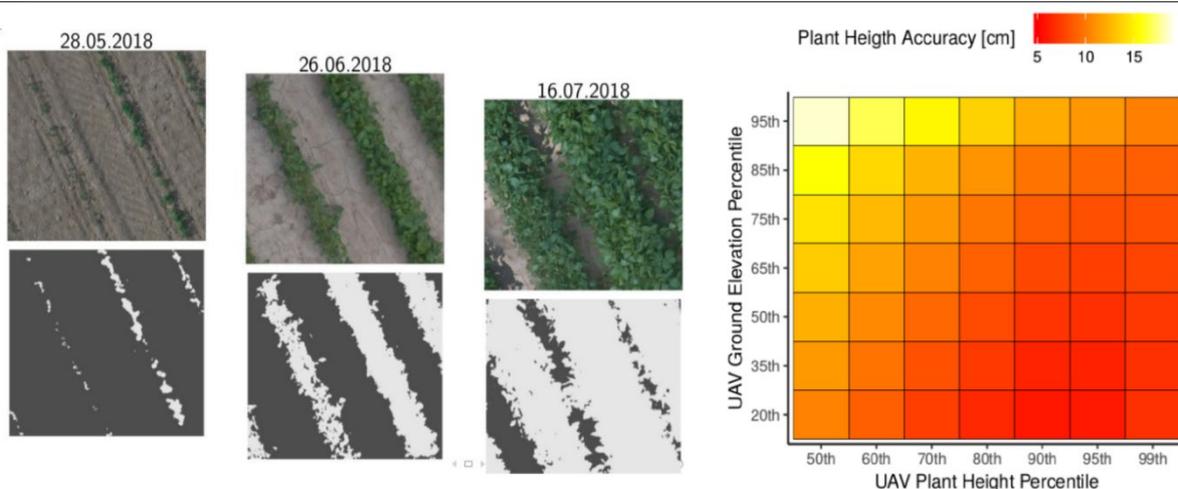
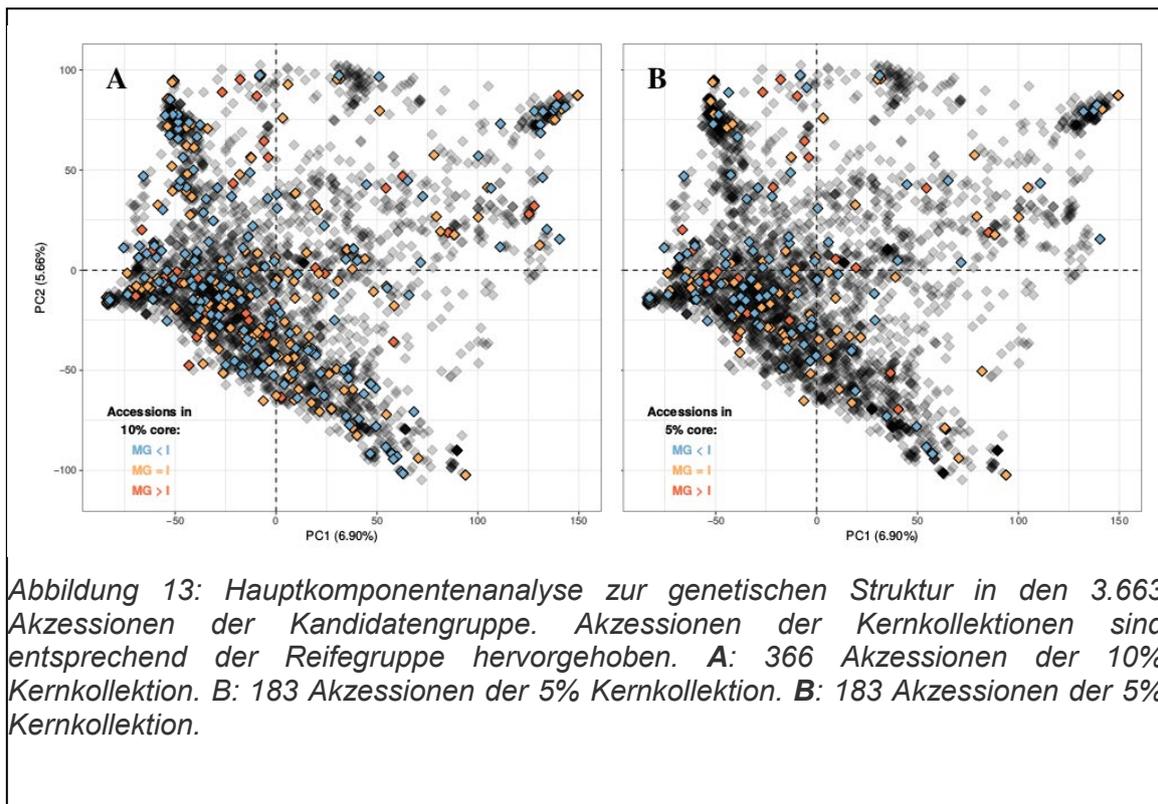


Abbildung 12: Hauptkomponentenanalyse zur genetischen Struktur in > 17.000 Sojaakzessionen die in der USDA Genbank erhalten werden. Jeder Punkt repräsentiert eine Akzession. **A:** Akzessionen mit Passport Informationen zu ihrer geografischen Herkunft sind entsprechend dem örtlichen Temperaturaufkommen farblich dargestellt. **B:** 3.663 Akzessionen der Kandidatengruppe sind hervorgehoben (farblich bzw. in Abwesenheit der Herkunftsinformationen in weiß). **C:** 366 Akzessionen der 10% Kernkollektion sind hervorgehoben. **D:** 183 Akzessionen der 5% Kernkollektion sind hervorgehoben.

Ein erster Feldversuch mit einer Kernkollektion bestehend aus 210 Akzessionen wurde 2018 in Hohenheim durchgeführt. Diese wurde im Wesentlichen ebenfalls gemäß der obigen Darstellung zusammengestellt. Einen Unterschied stellte hier jedoch die Nichtberücksichtigung von Material ohne Herkunftsinformation dar. Das wichtigste Versuchsergebnis war, dass eine kontinuierliche und non-invasive Datenerhebung über den gesamten Vegetationszeitraum mittels Drohnenaufnahmen qualitativ und quantitativ möglich ist. Beispielsweise konnte die Bodendeckung durch Sojapflanzen bis zum Reihenschluss Mitte Juli plot-spezifisch und akkurat quantifiziert werden (siehe Abbildung 14, links) und eignet sich somit als Merkmal um in Versuchen Reaktionen auf abiotischen Stress zu quantifizieren. Auch die Pflanzenhöhe konnte mittels Drohnenaufnahmen schnell und mit hoher Genauigkeit ermittelt werden. Die Abweichungen zu manuell ermittelten Werten lag hier im Mittel bei nur 5 cm (siehe Abbildung 14, rechts). Damit kommt der Drohnenphänotypisierung bei der Einschätzung der Umweltanpassung von Genbankmaterial durch die unmittelbare Abschätzung von Pflanzenreaktionen im Vegetationsverlauf potentiell eine Schlüsselrolle zu.



4.1d Genotypisierung und Charakterisierung von modernen europäischen Sojasorten mit dem SoySNP50K Array:

Die erfolgte Genotypisierung moderner europäischer und insbesondere mitteleuropäischer Sojasorten und die sehr gute Datenqualität ermöglichten es erstmals, diese systematisch mit amerikanischem und asiatischem Material zu kontextualisieren. Hierfür standen ~35.000 genomweite SNP-Marker mit einer Minor Allele Frequency von mindestens 0,05 zur Verfügung, die sowohl im europäischen Material als auch in den o.g. Vergleichsgruppen durch Genotypisierungsanstrengungen der USDA Genbank charakterisiert wurden. Analysen zum

Kopplungsungleichgewicht zeigten, dass dieses in modernem Material erwartungsgemäß ausgeprägter war als in älteren Sorten, welche wiederum im Vergleich zu Landrassen ein erhöhtes Kopplungsungleichgewicht aufwiesen. Die Berechnung des hintergründigen Kopplungsungleichgewichts ermöglichte die Auswahl eines Markersets bestehend aus ~5000 entkoppelten Markern. Die Ermittlung der Populationsstruktur mittels Hauptkomponentenanalyse zeigte, dass moderne europäische und insbesondere frühreife nordamerikanische Sorten ein Cluster bildeten (siehe Abbildung 15). Diese Gruppen wiesen auch eine vergleichsweise geringe genetische Distanz zueinander auf (siehe Abbildung 20). Älteres und insbesondere spätes amerikanisches Material zeigte sich deutlich differenziert von den europäischen Sorten. Generell lässt sich die beprobte globale Sojapopulation im Zusammenhang mit der Kurztagsnatur der Sojabohne als lose strukturiert gemäß der Reifegruppenzugehörigkeit beschreiben.

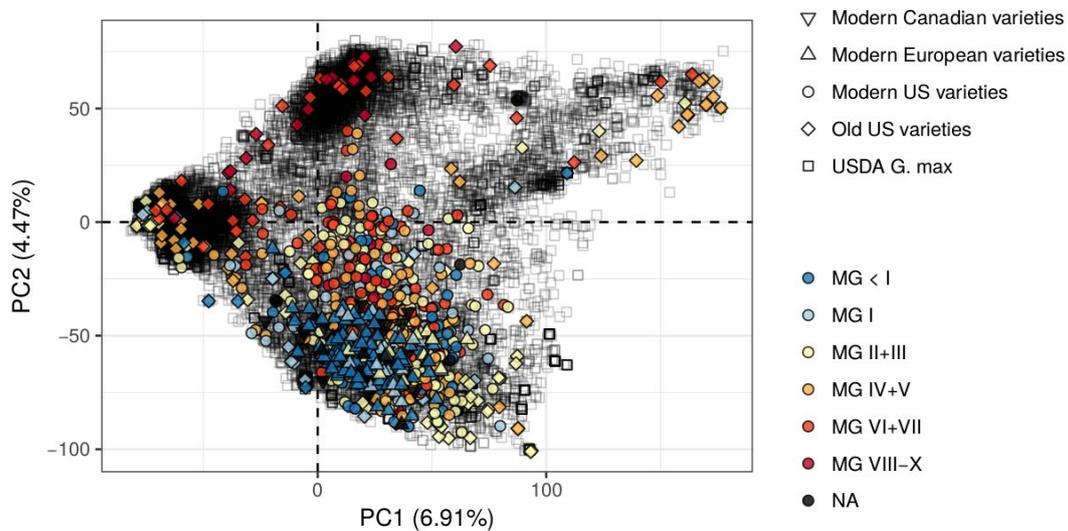


Abbildung 15: Hauptkomponentenanalyse zur genetischen Struktur in > 17.000 Sojaakzessionen die in der USDA Genbank erhalten werden, inklusive alte und moderne nordamerikanische Sorten sowie 160 moderne europäische Sorten. Jedes Zeichen repräsentiert eine Akzession. Zur besseren Übersicht ist Genbankmaterial (USDA G. Max) nicht entsprechend der Reifegruppeninformation eingefärbt.

Die Inferenz von Abstammungsverhältnissen mittels ADMIXTURE zeigte eine deutliche Durchmischung im Pool der Sojaakzessionen (siehe Abbildung 16), die von einem regen historischen Materialaustausch in der Sojazüchtung zeugen. Die Kreuzvalidierung identifizierte Modelle mit $K = 8$ bzw. $K = 10$ als datenstabile Szenarien. Auch hier zeigte sich bei den Genomanteilen verschiedener Ursprungspopulationen in den Individuen ein der Reifegruppenklassifizierung lose entsprechender Gradient. Die Fokussierung auf Material mit konservierter Abstammung und die Einbeziehung von Herkunftsinformationen ermöglichte darüber hinaus, diese Ursprungspopulationen trotz der historischen Durchmischung geografisch zu verorten (siehe Abbildung 17). Für das Szenario $K = 8$ wurden

17,019 accessions of USDA 'Introduced G. max collection'

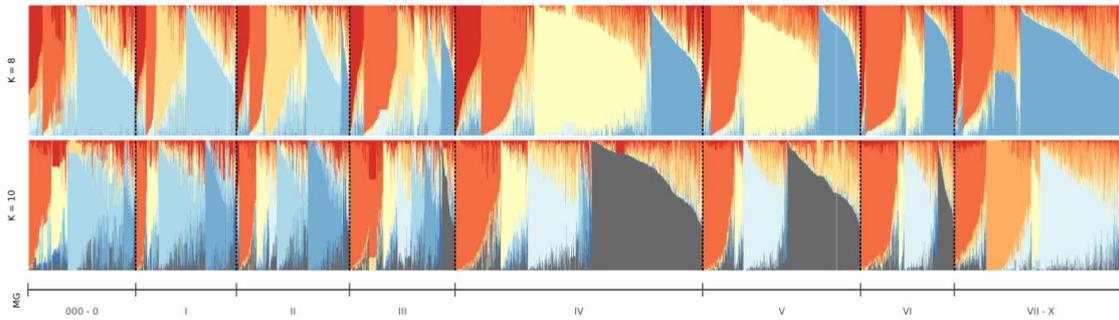


Abbildung 16: ADMIXTURE Ergebnis für die modelbasierte Inferenz von Abstammungsverhältnissen für $K = 8$ und $K = 10$ Ursprungspopulationen. Jeder vertikale Balken stellt ein Individuum dar. Die Farbgebung eines Individuums entspricht den Abstammungsproportionen von verschiedenen Ursprungspopulationen. Farbschlüssel wie in Abb. 17-19 angegeben. Dargestellt sind Akzessionen der USDA Genbank.

sechs von acht Ursprungspopulationen hinreichend konservierter Materialgruppen in Größenordnungen von 291 (K8.1) bis zu 2.138 (K8.2) Akzessionen identifiziert. Japanisches (K8.2) und koreanisches (K8.5) Material treten hier als voneinander abgegrenzte Populationen hervor, die sich wiederum deutlich von chinesischem Material unterscheiden. Insgesamt spiegelt die Populationsstruktur genetisch konservierter Akzessionen ihre geografische Abstammung wieder (siehe Abbildung 13). Das chinesische Material untergliederte sich in vier Populationen, die hinsichtlich ihrer Abstammung in etwa der Verteilung der von chinesischen Autoren beschriebenen vier Ökotypen entsprechen: nordöstliches und nördliches Soja zur Frühlingsaussaat, Soja zum Anbau im Huanghuai Tal (Gelber Fluss) zur Frühjahres-/Sommeraussaat, sowie südchinesisches Soja mit Aussaatmöglichkeit bis zum Herbst (Htwe et al. 2015; Wang et al. 2006). Diese Gruppierung blieb auch für ein Modell mit zehn Ursprungspopulationen ($K = 10$) stabil. Hier trat zusätzlich eine Gruppe von Akzessionen mit vornehmlich indonesischer und indischer Herkunft hervor (K10.3), sowie eine Gruppe US-amerikanischen Ursprungs (K10.9). Alte US-amerikanische Sorten und auch altes europäisches Zuchtmaterial stellen sich hinsichtlich ihrer Zusammensetzung aus den asiatischen Ursprungspopulationen und ihrer Durchmischung vergleichbar zur Gesamtheit des Genbankmaterials dar (Vgl. Abbildung 16 mit Abbildung 17 und Abbildung 18).

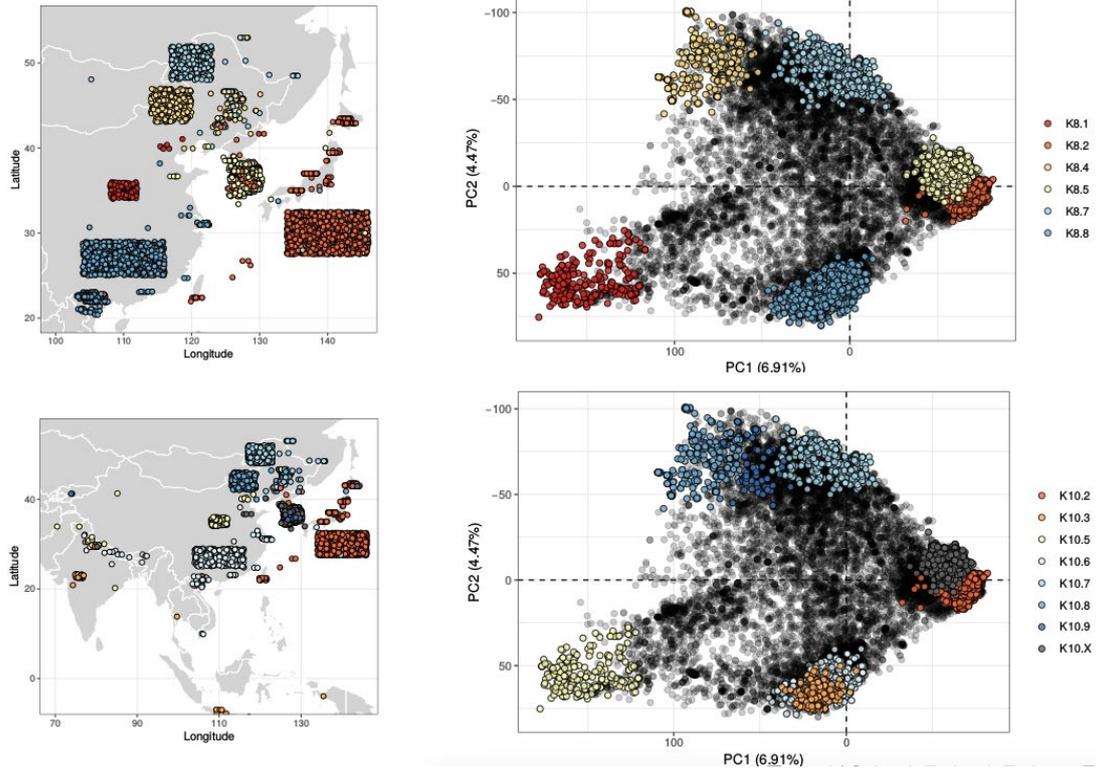
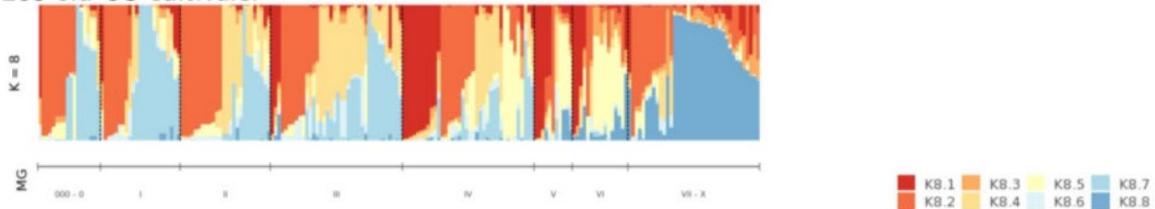
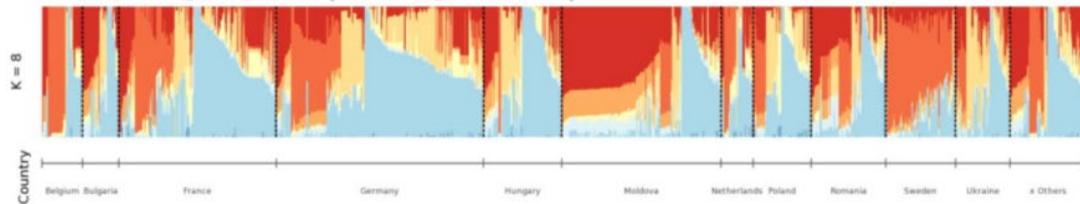


Abbildung 17: Geografische Herkunft und Position in der Hauptkomponentenanalyse für Akzessionen mit hohem Anteil an konservierter Abstammung (Genom zu mindestens 80% zugehörig zu einer Ursprungspopulation). Akzessionen ohne positionsgenaue Herkunftsangaben sind in Blöcken dargestellt. Die Position der Blöcke entspricht den Angaben zur nationalen Herkunft und der Reifegruppenzugehörigkeit des Materials.

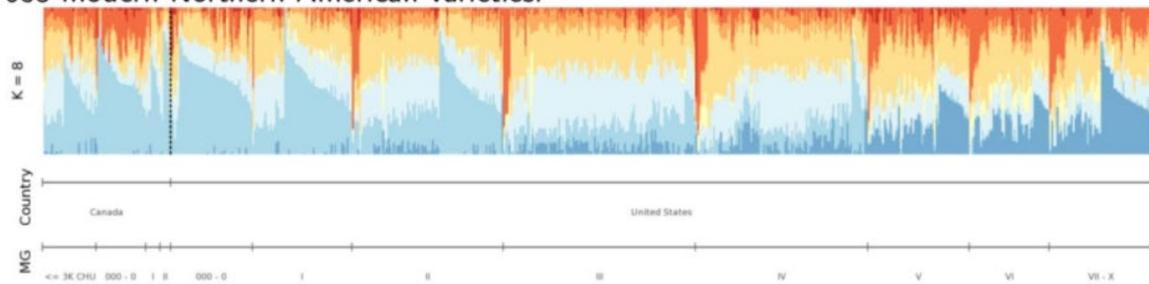
208 old US cultivars:



Old European germplasm (MG range 000 - III):



635 modern Northern American varieties:



160 modern European varieties:

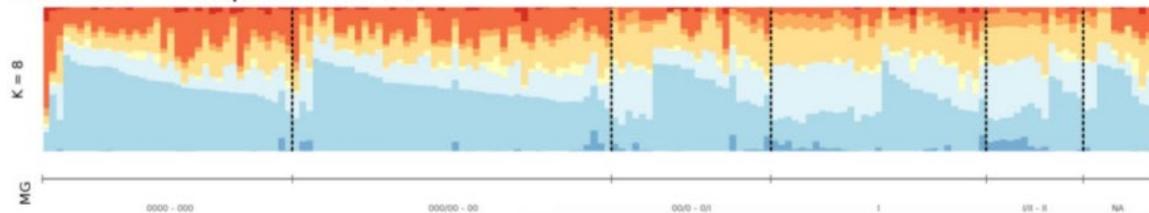
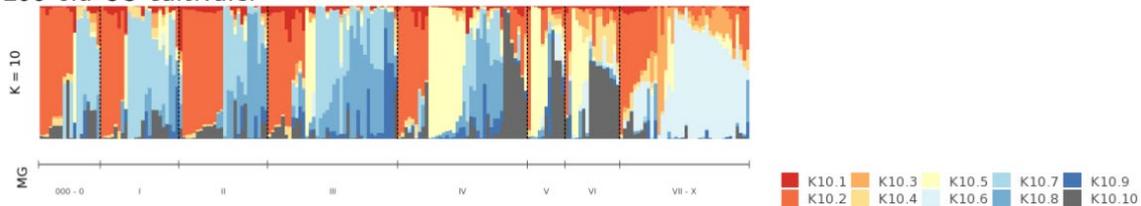


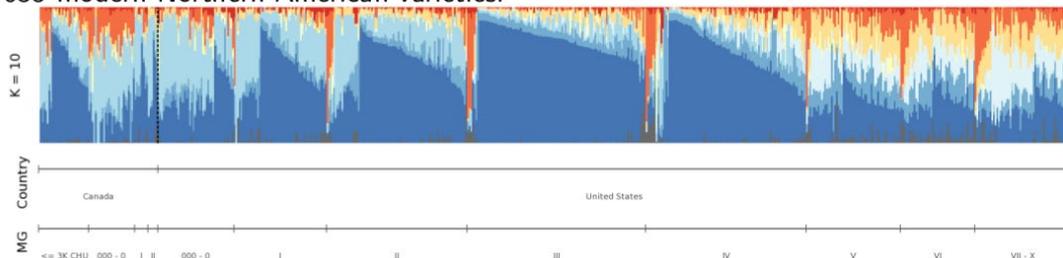
Abbildung 18: ADMIXTURE Ergebnis für die modellbasierte Inferenz von Abstammungsverhältnissen für $K = 8$ Ursprungspopulationen. Jeder vertikale Balken stellt ein Individuum dar. Die Farbgebung eines Individuums entspricht den Abstammungsproportionen von verschiedenen Ursprungspopulationen. Dargestellt sind alte US Sorten, alte europäische Sorten und Zuchtmaterial, moderne nordamerikanische Sorten und moderne europäische Sorten.

Erst in modernem Material kam es offensichtlich zu Differenzierungen. So spielen das genetische Erbe der Ursprungspopulation aus dem Huanghuai Tal (K8.1 bzw. K10.5) weder in modernen europäischen noch nordamerikanischen Sorten eine große Rolle. Auch das koreanische Erbe ist marginal. Stattdessen stellen die modernen Sojasorten im Wesentlichen eine Mischung aus japanischem (K8.2 bzw. K10.2), nordost-chinesischem (K8.7 bzw. K10.7) und nordchinesischem (K8.4 bzw. K10.8) Erbe dar. In spätreifem Material sind außerdem Anteile aus Südchina (K8.8 bzw. K10.6) sowie Indonesien/Indien (K10.3) vertreten. Darüber hinaus enthalten moderne Sorten je nach Szenario deutlich Anteile der Ursprungspopulationen K8.6 bzw. K10.9. Diese stellen keine Ursprungspopulationen im eigentlichen Sinne dar und kommen in asiatischem Material quasi nicht vor. Material mit hohen Anteilen an K10.9 kommt vornehmlich aus den USA. Entsprechend scheint diese Population ein Ergebnis westlicher Sojazüchtung zu sein. Die Hauptkomponentenanalyse legt nahe, dass diese Zuchtpopulation (K10.9) aus Kreuzungen von nordost-chinesischem (K8.7 bzw. K10.7) und nordchinesischem (K8.4 bzw. K10.8) Material hervorgegangen ist (siehe Abbildung 17, rechts unten). Vergleichsweise geringe genetische Distanzen zwischen den Individuen dieser Gruppen stützen diese Interpretation (siehe Abbildung 20).

208 old US cultivars:



635 modern Northern American varieties:



160 modern European varieties:

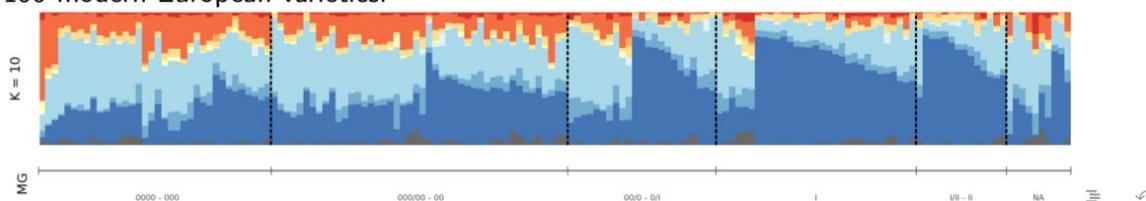


Abbildung 19: ADMIXTURE Ergebnis für die modellbasierte Inferenz von Abstammungsverhältnissen für $K = 10$ Ursprungspopulationen. Jeder vertikale Balken stellt ein Individuum dar. Die Farbgebung eines Individuums entspricht den Abstammungsproportionen von verschiedenen Ursprungspopulationen. Dargestellt sind alte US Sorten, moderne nordamerikanische Sorten und moderne europäische Sorten.

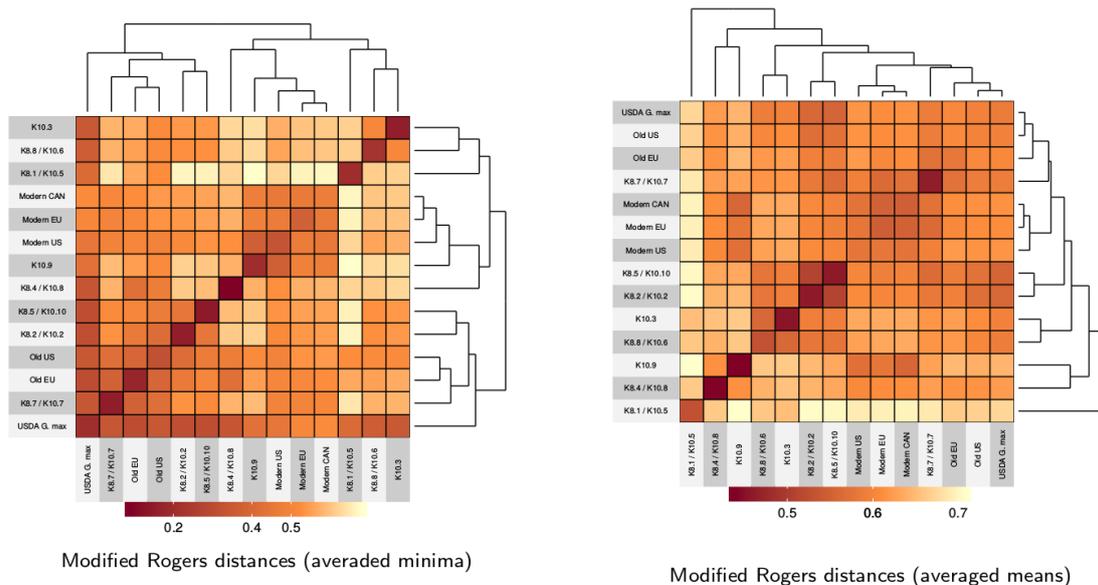


Abbildung 20: Durchschnittliche genetische Distanz zwischen Individuen verschiedener Gruppen. **Links:** Durchschnitt der Minimaldistanz zwischen Individuen. **Rechts:** Durchschnitt der mittleren Distanz zwischen Individuen.

AP1g: Feinkartierung und Markerentwicklung neuer Reifegene in der CC4CE

Für die Feinkartierung von neuen Reife- und Anpassungsgenen wurden Genotypisierungsdaten sowie die öffentlichen HapMap-Daten der USDA Genbankkollektion verwendet, um die in Projektphase 1 gefundenen Reifegene in hoher Auflösung zu kartieren bis auf Ebene der Kandidatengene zu identifizieren und Marker für die markergestützte Selektion zu entwickeln. Die Feinkartierung wurde mit etablierten Methoden der genom-weiten Assoziationskartierung wie gemischten Linearen Modellen und Statistiken der Populationsdifferenzierung (F_{st}) durchgeführt, um eine möglichst genaue Kartierung von Zielgenen zu erzielen. Diese Informationen sind in der Publikation von Haupt and Schmid (2020) und im unten beschriebenen Arbeitspaket AP1i und der daraus resultierenden Publikation (Haupt and Schmid 2022) enthalten, so dass ohne großen Aufwand Marker für die Genotypisierung zur Nutzung in der markergestützten Züchtung entwickelt werden können.

AP1h: Aufbau einer Allele Mining Pipeline mit publizierten Daten

Neben der Reifegruppe sind zahlreiche andere Merkmale wie Krankheitsresistenzen, Kältetoleranz oder Produktqualität für die Sojazüchtung von Bedeutung. Pflanzengenetische Ressourcen bieten auch für diese, in der Regel quantitativen Merkmale, eine Möglichkeit, neue genetische Variation zu identifizieren und durch Introgression eine züchterische Verbesserung von Elitematerial zu erzielen. Allele mining ist ein effizienter Ansatz, um neue genetische Variation (z.B. neue Allele an einem Resistenzgen) zu identifizieren und in der Züchtung zu verwenden. In diesem Arbeitspaket wurde eine bestehende Allele Mining Pipeline weiterentwickelt, um die gesamten USDA Kollektion mit Hilfe der öffentlich verfügbaren phänotypischen, d. h. über <https://www.ars-grin.gov/> abrufbare phänotypischen und genetischen Informationen für das Allele Mining nutzen zu können. Zu diesem Zweck werden die mittlerweile zahlreichen publizierten genom-weiten Assoziationsstudien (GWAS) bei Soja herangezogen. Zusätzlich wurde mit Hilfe von öffentlichen Whole Genome-Sequenzen (GmHapMap; (Torkamaneh et al. 2020) und den SNP-Genotypisierungsdaten eine Imputierung aller in SNP-genotypisierten Akzessionen durchgeführt. Anschließend wurden Literaturdatenbanken nach Publikationen mit beschriebenen SNP-

Merkmalsakzessionen durchsucht, um signifikante SNP-Marker-Assoziationen zu identifizieren. Die SNP Informationen wurden verwendet, um in den entsprechenden Zielregionen nach zusätzlicher allelischer Variation in Genbankakzessionen zu suchen. Mit entsprechenden Datenbankabfragen werden die allelische oder Haplotypen-Diversität extrahiert und Akzessionen identifiziert, die in den Zielregionen unterschiedliche Allele beziehungsweise Haplotypen aufweisen. Die in dieser Analyse und zusätzlich in der Literatur identifizierte Marker-Merkmals-Assoziationen wurden mit den in AP1i erhaltenen QTL verglichen und sind in der Publikation von Haupt und Schmid (2022) beschrieben.

AP1i: Assoziationskartierung für neue Merkmale in Genbank-Akzessionen und europäischen Sorten.

In diesem Arbeitspaket wurden phänotypische Daten von Genbankakzessionen aus der USDA Genbank und von in Hohenheim durchgeführten Feldversuchen verwendet, um GWAS Studien und genomische Vorhersagen für Merkmale durchzuführen, die in bisher publizierten GWAS Analysen noch nicht berücksichtigt wurden, aber für die Züchtung europäischer Sorten von Interesse sind. Ziel war es, QTL in genetischen Ressourcen zu identifizieren, die über markergestützte oder genomische Selektion durch rekurrente Selektion für die Verbesserung europäischer Sorten genutzt werden können. Die Identifizierung von neuen Kandidatengenen wurde über zwei verschiedene Ansätze durchgeführt:

Assoziationskartierung mit Umweltdaten (Environmental Association Mapping, EAM): Die Identifizierung von neuen Frühreife-QTL mit der auf Umweltdaten beruhenden Core Collection zeigt, dass mit EAM in genetischen Ressourcen von Soja Adaptationsgene identifiziert werden können. Mit weiteren Umweltvariablen wurden EAM Analysen durchgeführt, um Kandidatengene für Trocken- und Kältetoleranz zu identifizieren. Die geographischen Herkunftsdaten der Akzessionen wurden mit den entsprechenden Klimavariablen der WorldClim Datenbank kombiniert und mit den imputierten Genomdaten eine Assoziationsstudie durchgeführt. Der Fokus lag dabei auf Umweltvariablen, die für Merkmale wie Trocken- und Kältetoleranz von Bedeutung sind.

Die in Projektphase 1 gewonnenen Erkenntnisse zur Populationsstruktur im Material der *USDA Soybean Germplasm Collection* wurden genutzt, um die Verbindung von Genbankmaterial und seiner ursprünglichen Herkunft zu rekonstruieren. Insgesamt spiegelt die Populationsstruktur genetisch konservierter Akzessionen ihre geografische Abstammung wieder (siehe Abbildung 21B und C). Unsere weitergehende Analyse fokussierte sich auf chinesisches Material, welches sich in vier Populationen untergliedert. Diese entsprechen hinsichtlich ihrer Abstammung in etwa der Verteilung der von chinesischen Autoren beschriebenen vier Ökotypen: nordöstliches (K10.7) und nördliches Soja (K10.8) zur Frühljahrsaussaat, Soja zum Anbau im Huanghuai Tal (K10.5, Gelber Fluss) zur Frühljahres-/Sommeraussaat, sowie südchinesisches Soja (K10.6) mit Aussaatmöglichkeit bis zum Herbst.

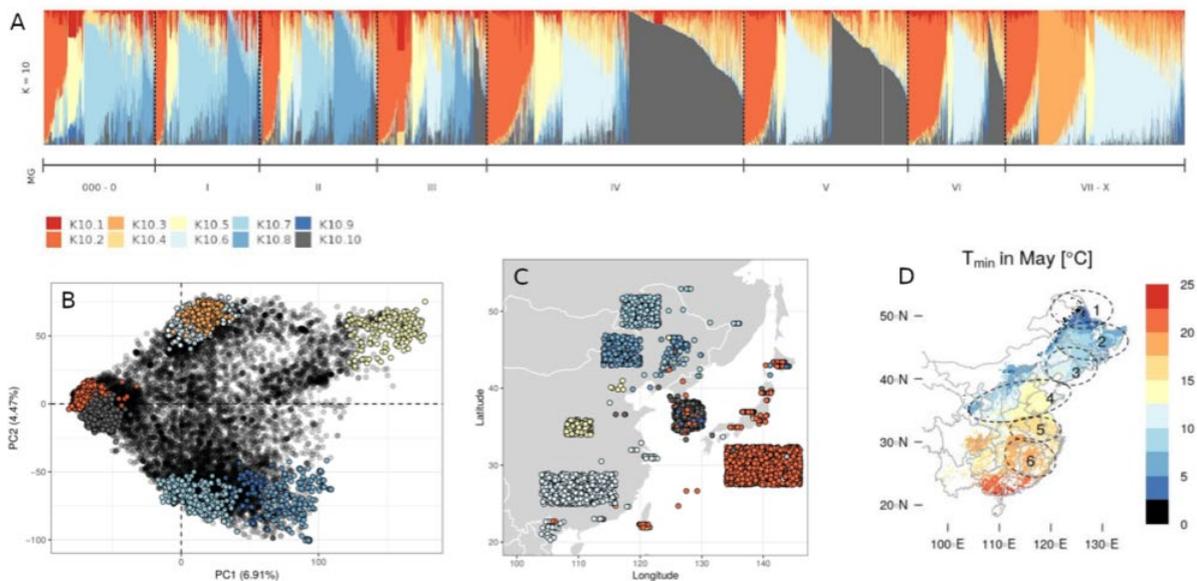


Abbildung 21: **A** ADMIXTURE Ergebnis für die modelbasierte Inferenz von Abstammungsverhältnissen in USDA Genbank Material für $K = 10$ Ursprungspopulationen. Jeder vertikale Balken stellt ein Individuum dar. Die Farbgebung eines Individuums entspricht den Abstammungsproportionen von verschiedenen Ursprungspopulationen. **B** Position in der Hauptkomponentenanalyse und **C** geografische Herkunft für Akzessionen mit hohem Anteil an konservierter Abstammung (Genom zu mindestens 80% zugehörig zu einer Ursprungspopulation). Akzessionen ohne positionsgenaue Herkunftsangaben sind in Blöcken dargestellt. Die Position der Blöcke entspricht den Angaben zur nationalen Herkunft und der Reifegruppenzugehörigkeit des Materials. **D** Regionale Untergliederung Chinas in Sojaanbauregionen mit spezifischer Agrarökologie.

Diese vier Populationen beinhalten jeweils Akzessionen mehrerer Reifegruppenkategorien, was den Schluss auf Anpassung an diverse agrarökologische Bedingungen in den jeweiligen Herkunftsgebieten nahelegt. Die Populationen wurden daher anhand von Reifegruppenzugehörigkeit und ihrer verzeichneten Herkunftsprovinzen weiter in Untergruppen unterteilt, um ein detaillierteres Matchmaking mit der jeweiligen potentiellen Umwelanpassung an spezifische Anbauregionen zu erreichen (siehe Abbildung 21D). Diese Anbauregionen wurden hinsichtlich ihrer Umweltbedingungen mittels WorldClim2 Daten charakterisiert. Die Umweltdaten wurden mit einer Hauptkomponentenanalyse zusammengefasst und die ersten beiden Hauptkomponenten wurden für das EAM herangezogen. Das EAM wurde mit dem Programm BAYPASS durchgeführt, welches über Allelfrequenzvergleiche zwischen Populationen sowohl genomische Regionen die sich durch genetische Differenzierung, als auch Assoziationen mit spezifischen Umweltgradienten auszeichnen, identifiziert. Genetische Differentiation und signifikante Genotyp–Umwelt Assoziationen können Genombereiche, die für Umwelanpassung verantwortlich sind, identifizieren. Wir konnten entsprechende Selektionssignaturen auf allen 20 Chromosomen identifizieren (siehe Abbildung 22). Diese umfassten Regionen bis zu einer Länge von 826kb und beinhalteten insgesamt 6.314 Gene. 90% dieser Signaturen waren jedoch kürzer als 150kb und beinhalteten weniger als fünf Kandidatengene, was die anschließende Inspektion der funktionellen Annotation begünstigt.

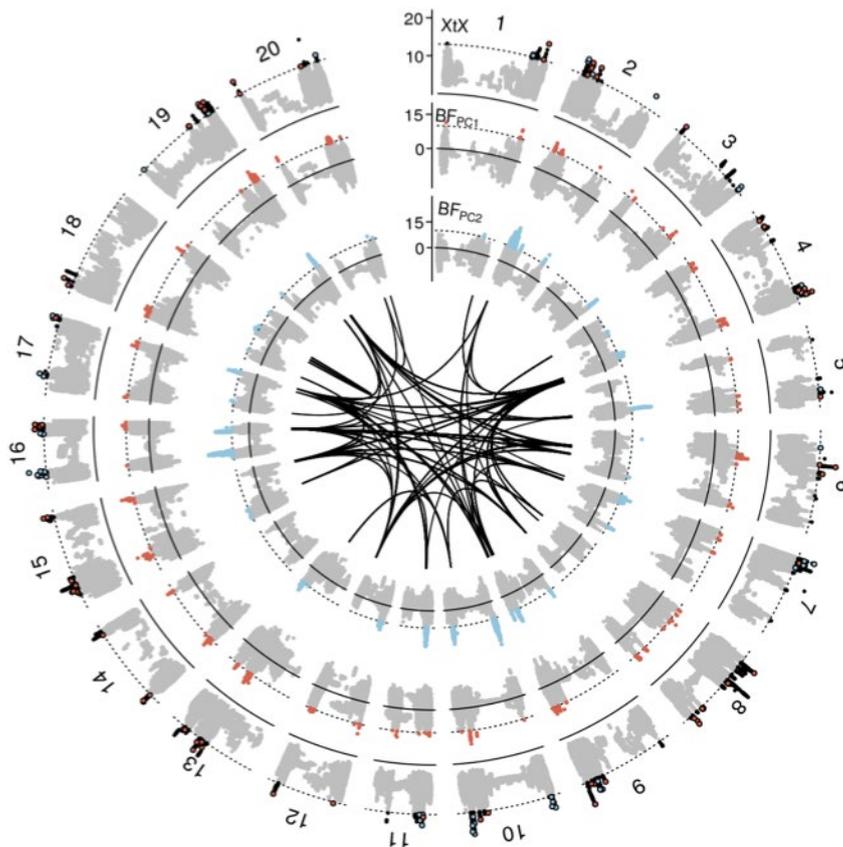


Abbildung 22: Genetische Differenzierungssignale (XtX) und Genotyp-Umwelt Assoziationen (BF PC1 und BF PC2) im Sojagenom für den Vergleich von Sojapopulationen aus den Anbauregionen von Nordost- bis Südchina. Gestrichelte Linien stellen die jeweilige Signifikanzschwelle dar. Rote und blaue Punkte im äußerem Ring stellen signifikante Genotyp-Umwelt Assoziationen dar, welche auch eine signifikante genetische Differenzierung zwischen den untersuchten Populationen aufweisen. Schwarze Verbindungslinien im Zentrum der Abbildung stellen Fälle von erhöhtem Kopplungsungleichgewicht zwischen detektierten Regionen dar (LD > 0.6, entspricht > 3fachem Background LD).

Insgesamt war zu erwarten, dass Umwelthanpassung als komplexes polygenisches Merkmal eine komplexe genetische Architektur aufweist. So ergaben sich für die gemeinsame Analyse der Materialgruppen von K10.7, K10.5 und K10.6, 330 signifikant differenzierte Regionen. Genombereiche mit signifikanter Genotyp-Umweltassoziation sind in ähnlicher Größenordnung vorhanden und überlappen teilweise mit den Differenzierungssignalen. Dies stützt die Hypothese, dass die betreffenden Regionen tatsächlich in Anpassung an verschiedene Umweltbedingungen in den untersuchten Anbauregionen involviert sind. Auch sind die detektierten Signale für Gene angereichert, die im Zusammenhang mit Blüte, Photoperiode und der Reaktion auf abiotischen und biotischen Stress stehen. Die detektierten Regionen enthalten zahlreiche Gene, deren Relevanz für Umwelthanpassung beispielsweise in *Arabidopsis* nachgewiesen ist. Interessanterweise weist ein Großteil der detektierten Signale ein erhöhtes Kopplungsungleichgewicht zueinander auf, welches deutlich das Kopplungsungleichgewicht in Stichproben zufällig ausgewählter Marker übersteigt und trotz großer räumlicher Distanz auf dem gleichen Chromosom und chromosomenübergreifend besteht. Dies ist ein Hinweis darauf, dass sich die genetische Zusammensetzung in den detektierten Regionen in den untersuchten Populationen nicht unabhängig voneinander

entwickelt. Eine mögliche Interpretation ist, dass entsprechende Signale genomische Regionen kennzeichnen, welche gemeinsam, z.B. in Form von Reaktionskaskaden, Anpassung an spezifische Umwelteinflüsse bedingen.

Um zu testen, ob Allele, die in den EAM als für die Umwelthanpassung in den genetischen Ressourcen aus China vorteilhaft erschienen, bereits in amerikanischen europäischen Sorten vorhanden sind, wurden die 50k SoySNP Genotypisierungsdaten mit Hilfe der GmHapMap imputiert und die entsprechenden Bereiche verglichen. Es zeigte sich, dass für eine Reihe von Maturitätsgenen die Allele für frühe Reife bereits im modernen US- und europäischen Zuchtmaterial vorhanden sind, viele andere genomische Regionen mit bisher unbekanntem vorteilhaften Allelen aber noch nicht fixiert sind.

Um genomische Regionen mit Selektionssignalen und mutmaßlich adaptiver genetischer Variation in genetischen Ressourcen für die Züchtung von Elitesorten zu nutzen, haben wir Haplotyp-Blöcke identifiziert, die in Keimplasma aus China und modernem Zuchtmaterial aus Nordamerika und Europa enthalten sind. Kanadische und mitteleuropäische Sojazuchtprogramme zielen auf eine verbesserte Anpassung an kühlere Regionen in höheren Breitengraden, für die derzeit nur wenig genetische Variation verfügbar ist. Neue Variation für die Kälteanpassung könnte daher für Haplotyp-Blöcke in Landrassen gefunden werden, die an den hohen Norden Nordostchinas und die angrenzenden Regionen in Russland angepasst sind (Haupt und Schmid 2020). Wir haben daher untersucht, ob Haplotyp-Blöcke, die in genomischen Regionen mit Selektionssignalen in Keimplasma aus China vorhanden sind in modernen Sorten aus Nordamerika und Mitteleuropa vorhanden sind und fanden variable Anteile von Haplotyp-Blöcken mit Abstammung aus Nordostchina in diesen Sorten. Zum Beispiel war die genomische Region, die das Reifegen *E9* enthält, in den am frühesten reifenden modernen nordamerikanischen und europäischen Sorten (MG 000-0) für den Haplotyp-Block, der in Material aus dem Nordosten Chinas vorherrscht, nahezu fixiert, während die Häufigkeit von Haplotyp-Blöcken mit Abstammung von Material aus Südchina in Sorten der späteren Reifegruppe rapide zunimmt. Ähnliche Variationsmuster bei der Haplotyp-Häufigkeit wurden auch in anderen genomischen Regionen beobachtet und scheinen die Selektion für eine bessere Anpassung an bestimmte Breitengrade widerzuspiegeln. Diese Interpretation steht im Einklang mit einer Zunahme von Haplotyp-Blöcken aus Nordostchina stammenden Haplotyp-Blöcken in modernen früh reifenden Sorten im Vergleich zu älteren Sorten. Insgesamt waren Haplotyp-Blöcke nordöstlicher Abstammung mit Selektionssignaturen im Allgemeinen häufiger in früh reifenden Sorten als in später reifenden Sorten, aber sie sind immer noch in vielen genomischen Regionen unterrepräsentiert. Diese Beobachtungen deuten auf ein großes Potential für weitere Introgression von potenziell adaptiven Allelen für den Anbau in kühlen Regionen hoher Breitengrade. Haplotyp-Block-Profile moderner, früh reifender nordamerikanischer und europäischer Sorten zeigen, dass eine solche Introgression innerhalb des modernen Keimplasma-Pools durch gezielte Kreuzungen von Sorten mit komplementären Haplotyp-Blockprofilen und anschließender markergestützter Selektion erreicht werden könnte, um Zuchtmaterial zu entwickeln, das mit Haplotyp-Blöcken mit Abstammung aus dem Nordosten Chinas angereichert ist.

Assoziationskartierung mit phänotypischen Daten: Die in AP1h in der Datenbank abgespeicherten phänotypischen Daten aus der USDA Genbank und den Versuchen in Hohenheim und Groß-Lüsewitz (JKI) wurden mit den imputierten Polymorphismusdaten kombiniert und eine GWAS mit dem QTCAT Programm durchgeführt. Um genomische Regionen zu identifizieren, die bei der Reife eine Rolle spielen, wurden historische Daten aus der USDA Genbank von 1.688 mit Feldversuchsdaten von 160 Akzessionen der Core Collection aus Phase 1 des Projekts (Haupt and Schmid 2020) für eine Assoziationsstudie verwendet. Die Feldversuche wurde in 2020 an den Standorten Eckartsweier (Baden-Württemberg) und Groß-Lüsewitz (Mecklenburg-Vorpommern) durchgeführt. Um die

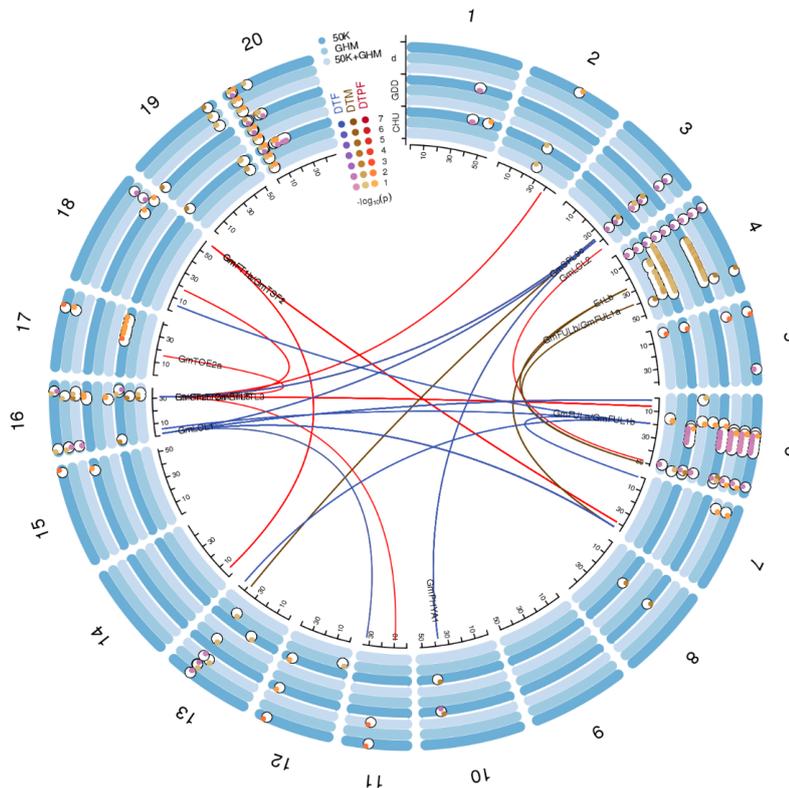


Abbildung 23: Ergebnis der GWAS für die Merkmale Tage bis zur Blüte, Tage bis zur Reife und Tage bis zur Schotenfüllung mit historischen Daten aus den USA für 1,688 Akzessionen. Es wurden drei verschiedene Markersätze verwendet und die QTCAT Methode für die Assoziationen. Die Linien zwischen den Genen zeigen signifikante epistatische Assoziationen, die mit dem Programm GWESS identifiziert wurden.

Robustheit der Assoziationen zu überprüfen, wurden drei verschiedene Markerdatensätze verwendet: 50k SoySNP Marker (41,084 SNPs), die GmHapMap Daten (ca. 900,000 SNPs) und die Kombination aus beiden Markersätzen. Für die Assoziationsstudie wurde die Methode QTCAT verwendet. Für die Merkmale Tage bis zur Blüte, Tage bis zur Reife und Tage bis zur Hülsenfüllung wurden jeweils signifikante Assoziationen gefunden, die z.T. sehr robust über die verwendeten Markerdaten waren (Abbildung 23).

Eine weitere Analyse von epistatischen Interaktionen von bekannten Blühzeitpunktgenen wie E9 mit anderen Loci mit dem Programm GWESS (Abbildung 23) unter Verwendung der historischen Daten ergab weitere Assoziationen. Dabei wurden je nach Merkmal 37-53 Gene für die drei Merkmale jeweils mit QTCAT und GWESS identifiziert.

Arbeitspaket AP1i zeigt, dass sowohl EAM als auch klassische GWAS mit historischen und aktuellen phänotypischen Daten für die Identifizierung von Adaptationsgenen verwendet werden können. Signifikante Assoziationsgene aus den Untersuchungen von Genbankakzessionen aus China sind den Analysen zufolge auch in Europäischen Sorten funktionell und können für die Züchtung verwendet werden. Ein Vergleich der Assoziationsmethoden und eine Analyse auf Epistasie identifizierte bekannte und bisher unbekannte Loci und zeigen die Netzwerkstruktur der Regulation von Blühzeitpunkt und Reife.

4.2 LSA

4.2a QTL-Analysen für Ertrag, Reifezeit, Tausendkornmasse, Protein- und Ölgehalt

Phänotypisierung

Die Ergebnisse sind ausführlich beschrieben in Kurasch et al. (2017).

Diese Studie basierte auf 1.008 $F_{5:8}$ rekombinanten Inzuchtlinien (RIL), die aus neun Familien aus einem halbdiallelen Kreuzungsdesign mit fünf Elternlinien stammen, die alle an Mitteleuropa angepasst sind, sich aber in Bezug auf die untersuchten Merkmale unterscheiden. Alle RILs und ihre Eltern wurden 2014 in Feldversuchen evaluiert und entsprechend ihrer im Vorjahr festgelegten Reife in drei verschiedene Versuche mit unterschiedlichen Erntezeitpunkten unterteilt. Mittel mit Bandbreiten, Varianzkomponenten und Heritabilität wurden für jede Studie und in den Versuchsreihen übergreifend bewertet (Tabelle 4). Die beobachteten Mittelwerte für Proteingehalt (39,15 % - 40,13 %), Ölgehalt (18,55 % - 18,73 %) und Tausendkorngewicht (199,3 - 208,4) waren in jeder Studie ähnlich und zeigten in den Einzelstudien ähnliche Bereiche wie in der Serie. Der mittlere Bohnenertrag unterschied sich dagegen zwischen den drei Studien mit durchschnittlich 2,6 t/ha in den frühreifen Linien von Versuch 1, 2,5 t/ha in Versuch 2 und einem etwas höheren Mittelwert von 2,8 t/ha in den spätreifen RILs von Versuch 3. Ein fast ähnliches Bild wurde für den Proteinertrag gefunden, bei der die RILs von Versuch 3 mit 1,14 t/ha einen etwas höheren Mittelwert hatten als die RILs von Versuch 1 und Versuch 2 mit 1,05 t/ha bzw. 1,04 t/ha. Ebenso wurde ein ähnlicher Trend für die Pflanzenhöhe beobachtet, wobei Versuch 3 generell größere Pflanzen aufwies. Die Heritabilitäten in der Serie über die Versuche hinweg waren für alle Merkmale hoch und lagen zwischen 0,77 für den Bohnenertrag und 0,94 für das Tausendkorngewicht. Die Heritabilitäten in jedem Versuch und in der Serie waren nahezu identisch, mit Ausnahme des Bohnenertrags und des Proteinertrags, für die sie zwischen 0,58 und 0,86 bzw. 0,56 und 0,82 lagen. Aufgrund der Struktur des Datensatzes mit neun segregierenden Familien haben wir die genotypische Varianz unterteilt in die Varianz zwischen und innerhalb der Familien. Die genotypische Varianzkomponente war für alle Merkmale in jeder Studie und in der Serie hochsignifikant, mit Ausnahme der genotypischen Varianz des Bohnenertrags und des Proteinertrags zwischen den Familien. Für den Bohnenertrag war diese Varianzkomponente in Versuch 1 signifikant, in Versuch 2 hoch signifikant, erklärte aber keine Variation in Versuch 3 und in der Serie. Für die in unserer Studie untersuchten Merkmale war die gesamte genotypische Variation hauptsächlich auf eine signifikante genotypische Varianz innerhalb der Familien und in geringerem Maße auf die genotypische Varianz zwischen den Familien zurückzuführen. Es muss darauf hingewiesen werden, dass dies nicht ausschließt, dass es erhebliche Unterschiede in der Merkmalsleistung zwischen den neun Familien gibt. Für den Proteingehalt zum Beispiel zeigten die beiden Familien P2×P3 und P1×P3 durchweg höhere Werte als die anderen Familien(siehe Abbildung 24).

Obwohl die Variation aufgrund der Interaktion von Genotyp zu Ort für alle sechs Merkmale signifikant war, war sie viel kleiner als die genotypische Variation. Der Versuchseffekt, der zwischen den Orten in der Serie modelliert wurde, war nur für die Pflanzengröße und das Tausendkorngewicht signifikant, was auf eine gewisse Beziehung dieser Merkmale zur Reife hinweist, während der Versuchseffekt für Bohnenertrag, Proteingehalt, Ölgehalt und Proteinertrag nicht signifikant war. Diese Unterschiede zwischen den Studien zeigen sich auch in der Verteilung der Merkmalswerte aus den Einzelstudien. Insbesondere bei der Pflanzengröße ist der Versuchseffekt insofern erkennbar, als nicht nur der Mittelwert von Versuch 3 höher ist, sondern auch die Verteilung der Genotypen deutlich in Richtung größerer Pflanzen verschoben wird.

Tabelle 4: Zusammenfassende Statistiken für die Merkmale Bohnenertrag (BE), Proteingehalt (PG), Ölgehalt (OG), Proteinertrag (PE), Tausendkorngewicht (TKG) und Pflanzenhöhe (PH) jeder Studie und der Serie über die Versuche hinweg (Kurasch et al. 2017).

Quelle †	BE [Mg ha ⁻¹]	PG [%]	OG [%]	PE [Mg ha ⁻¹]	TKG [g]	PH [cm]
Versuch 1 (N=366)						
Mittel	2,64	39,70	18,73	1,05	199,50	72,51
(Min; Max)	(1,94; 3,58)	(35,37; 44,47)	(16,04; 20,40)	(0,80; 1,32)	(149,20; 275,10)	(42,91; 108,15)
σ_G^2	0,47 ***	1,73 ***	0,37 ***	0,06 ***	241,35 ***	81,41 ***
σ_{Fam}^2	0,10 *	0,63 ***	0,19 ***	0,01	245,49 ***	36,90 **
σ_L^2	0,77	1,72	0,09	0,21	771,41 *	9,63
$\sigma_{G \times L}^2$	0,12 **	0,21 **	0,03	0,01*	15,11 *	17,77 ***
$\sigma_{Fam \times L}^2$	0,07 ***	0,08 ***	0,02 ***	0,01 ***	3,30 **	12,93 ***
σ_e^2	0,46	0,69	0,15	0,09	63,10	25,10
h^2	0,73	0,87	0,91	0,65	0,95	0,85
Versuch 2 (N=315)						
Mittel	2,52	40,13	18,55	1,04	208,4	64,44
(Min; Max)	(1,41; 3,62)	(36,49; 43,63)	(16,73; 20,48)	(0,57; 1,43)	(162,30; 260,10)	(38,12; 109,47)
σ_G^2	1,22 ***	1,39 ***	0,27 ***	0,14 ***	286,63 ***	130,29 ***
σ_{Fam}^2	0,47 ***	0,37 **	0,14 ***	0,07 ***	158,25 ***	97,23 ***
σ_L^2	0,11	2,11	0,38 *	0,08	500,81	6,54
$\sigma_{G \times L}^2$	0,27 ***	0,52 ***	0,09 ***	0,01	45,03 ***	23,85 ***
$\sigma_{Fam \times L}^2$	0,09 ***	0,21 ***	0,02 ***	0,02 ***	12,84 ***	9,69 ***
σ_e^2	0,36	0,67	0,15	0,14	59,48	32,00
h^2	0,86	0,78	0,81	0,82	0,92	0,91
Versuch 3 (N=446)						
Mittel	2,88	39,15	18,57	1,14	199,3	102,22
(Min; Max)	(1,95; 3,83)	(35,30; 44,44)	(15,97; 20,64)	(0,75; 1,56)	(138,00; 267,60)	(65,98; 138,47)
σ_G^2	0,44 ***	1,7 ***	0,33 ***	0,06 ***	316,97 ***	127,61 ***
σ_{Fam}^2	0	0,43 **	0,14 ***	0	127,46 ***	28,07 **
σ_L^2	2,39 **	3,93	0,14	0,64 **	887,57 **	6,66
$\sigma_{G \times L}^2$	0,13 *	0,51 ***	0,08 ***	0,01	43,47 ***	40,26 ***
$\sigma_{Fam \times L}^2$	0,32 ***	0,23 ***	0,05 ***	0,05 ***	4,35 ***	16,58 ***
σ_e^2	0,93	0,68	0,16	0,18	51,98	48,54
h^2	0,58	0,82	0,83	0,56	0,93	0,82
Serie (N=1,013)						
Mittel	2,70	39,38	18,58	1,06	202,00	81,74
(Min; Max)	(1,35; 3,91)	(35,30; 44,55)	(15,76; 20,72)	(0,5; 1,56)	(145,10; 274,20)	(40,81; 132,97)
σ_G^2	0,81 ***	1,71 ***	0,35 ***	0,09 ***	296,28 ***	214,22 ***
σ_{Fam}^2	0	0,4 ***	0,12 ***	0	153,91 ***	15,25 *
σ_L^2	0,96 *	2,5 **	0,09	0,25 *	668,07 ***	0
$\sigma_{G \times L}^2$	0,1 ***	0,33 ***	0,04 ***	0,01 *	26,92 ***	13,16 ***
$\sigma_{Fam \times L}^2$	0,1 ***	0,1 ***	0,02 ***	0,01 ***	3,47 ***	12,01 ***
$\sigma_{L \times Trial}^2$	0,32	0,11	0,12	0,09	56,90 *	85,88 **
σ_e^2	0,74	0,82	0,18	0,17	67,82	52,56
h^2	0,77	0,85	0,87	0,70	0,94	0,91

† σ_G^2 : genotypische Varianz innerhalb der Familien, σ_{Fam}^2 : genotypische Varianz zwischen Familien, σ_L^2 : Ortsvarianz, $\sigma_{G \times L}^2$: genotypische Interaktionsvarianz innerhalb Familien x Ort, $\sigma_{Fam \times L}^2$: genotypische Interaktionsvarianz zwischen Familien x Ort, $\sigma_{L \times Trial}^2$: Interaktionsvarianz Ort x Versuch, σ_e^2 : Fehlervarianz, h^2 : Heritabilität.

*, **, *** signifikant unterschiedlich von 0 beim Wahrscheinlichkeitsgrad von 0,05, 0,01 und 0,001.

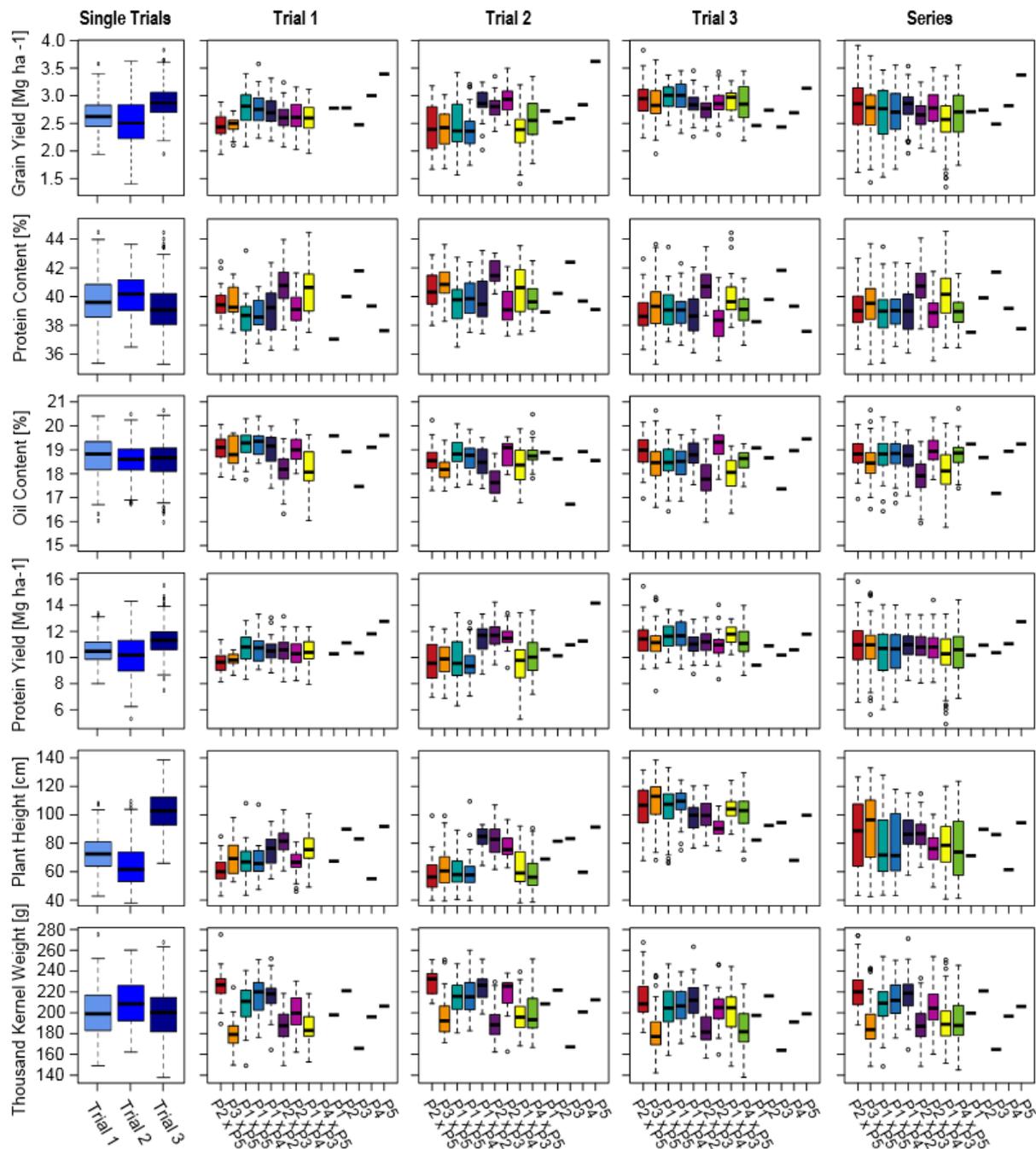


Abbildung 24: Boxplots der BLUES der 6 untersuchten Merkmale für die drei Versuche (Trial), die 9 Familien je Versuch und der Serie über die Versuche (Kurasch et al. 2017).

Genotyping by Sequencing

In den 1850 Sojalinien, welche mittels Genotyping-by-Sequencing genotypisiert wurden, konnten 98307 SNPs detektiert werden. Diese SNPs wurden imputiert und dann basierend

auf Minorallelfrequenz und Heterozygotie der Eltern gefiltert. Letztendlich konnten somit 24211 hochqualitative SNP Marker für die weiterführenden Analysen verwendet werden. Mittels dieser SNPs wurde überprüft, ob alle Individuen der einzelnen Familien korrekt zugeordnet waren und ob die Familien erwartungsgemäß genetisch zwischen ihren Eltern lagen. Abbildung 25 zeigt, dass es nur wenige Ausreißer in den Familien gab und nur die Population Gallec x Sultana wohl eine Fehlkreuzung war, diese Individuen aber mittels der Marker zur Familie Gallec x Sigalia zugeordnet werden konnten. Die GBS basierte Genotypisierung erwies sich als sehr gute Methode und wird auch weiterhin eingesetzt.

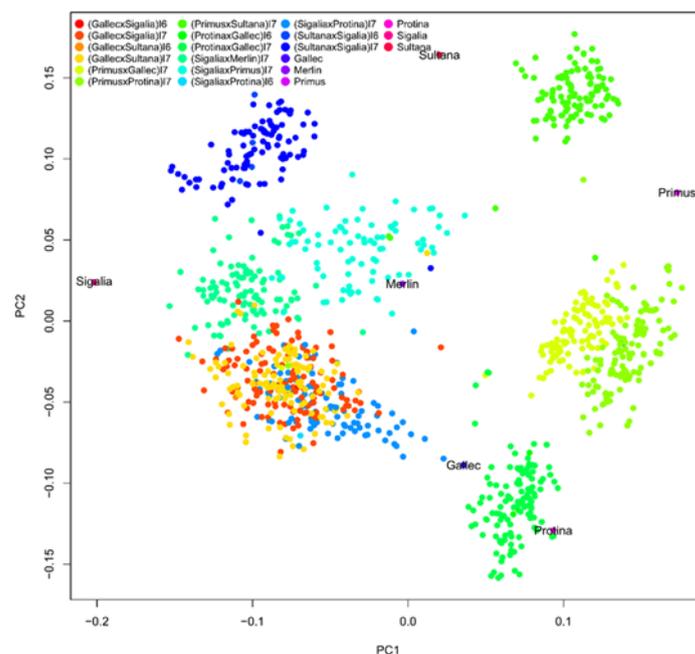


Abbildung 25: Principal-Coordinate-Analyse der 1008 Genotypen basierend auf 24K SNP Marker.

In der genomweiten Re-Sequenzierung der 69 Sorten wurden die Polymorphismen mit zwei verschiedenen Softwareprogrammen detektiert. Die Software Samtools detektierte 9,5 Millionen und die Software Freebayes 7,7 Millionen. Da es keine absolute Genauigkeit in der SNP-Detektion gibt, wurden die übereinstimmenden Marker zwischen beiden Methoden als robuste hoch qualitative Marker angesehen. Es konnten somit 2,1 Millionen hochqualitative Marker detektiert werden. Um die 69 Sojasorten mit einer globalen Sojakollektion zu vergleichen, wurden die 50K SNP Daten der 20087 weltweiten Sojalinien der US Genbank heruntergeladen und mit diesen 69 Sorten verglichen. Es wurden 11238 überlappende SNP Marker gefunden, welche zur Berechnung einer Principal Coordinate Analyse verwendet wurden. In Abbildung 26 ist zu sehen, dass die europäischen Sorten genetisch zwischen den amerikanischen und asiatischen Genotypen liegen. Diese Information ist sehr wichtig für eine weitere Diversifizierung des europäischen Zuchtpools.

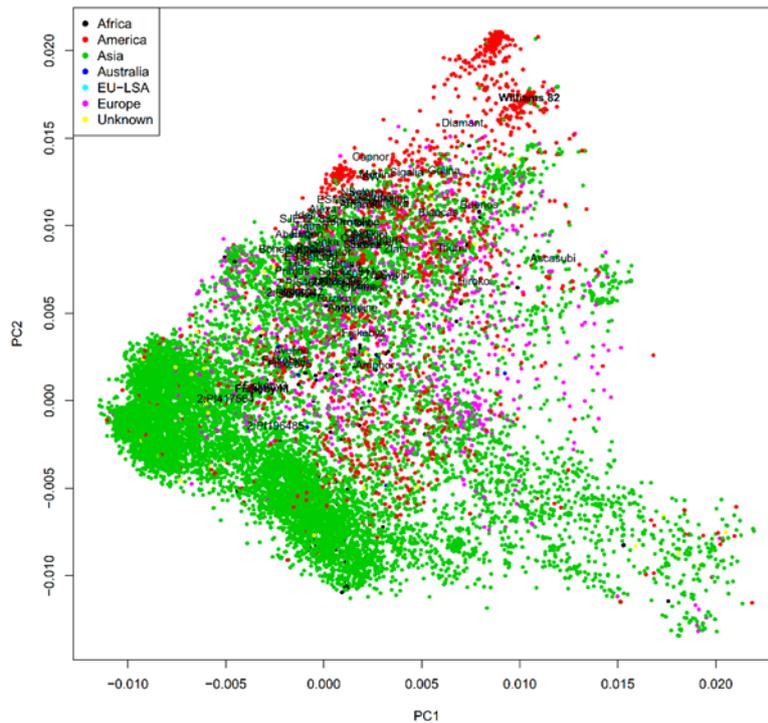


Abbildung 26: Principal-Coordinate-Analyse von ~20K weltweiten Sojalinien mit 69 europäischen Sojasorten (mit Namen) basierend auf ~11K SNPs.

QTL-Kartierung

Für alle fünf agronomischen Merkmale konnten signifikante QTL-Regionen in einer genomweiten Assoziationsstudie detektiert werden (Tabelle 5, Abbildung 27). Speziell für die Merkmale Wuchshöhe, Protein- und Ölgehalt wurden höchst signifikante genomische Regionen entdeckt. Die besten Marker für diese Merkmale erklärten bis zu 30% der phänotypischen Varianz und sind somit vielversprechende Marker für eine markergestützte Selektion. Die detektierten genomischen Regionen für Ertrag waren übereinstimmend mit den Markern für Wuchshöhe. Dies zeigt, dass Ertrag u.a. stark von der Wuchshöhe beeinflusst ist.

Tabelle 5: Höchst assoziierte Marker und ihre erklärte phänotypische Varianz in 1008 Linien für die agronomischen Merkmale in 2014.

Merkmals	SNP	Chr	Pos	$-\log_{10}$ P-value	R ²
Ertrag	S19_48096195	19	48096195	7.9	0.08
Ertrag	S06_20735680	6	20735680	7.3	0.11
Protein	S20_31924120	20	31924120	19.0	0.29
Protein	S15_4569252	15	4569252	6.8	0.16
Öl	S20_31924120	20	31924120	16.4	0.24
Öl	S15_4569252	15	4569252	7.6	0.19
Wuchshöhe	S06_20735680	6	20735680	15.0	0.24
Wuchshöhe	S11_10950383	11	10950383	8.2	0.02
Wuchshöhe	S18_55791681	18	55791681	7.8	0.09
Wuchshöhe	S19_48203425	19	48203425	7.6	0.01
TKG	S20_31656055	20	31656055	9.2	0.10
TKG	S17_40153761	17	40153761	8.6	0.21
TKG	S08_47630142	8	47630142	5.9	0.11

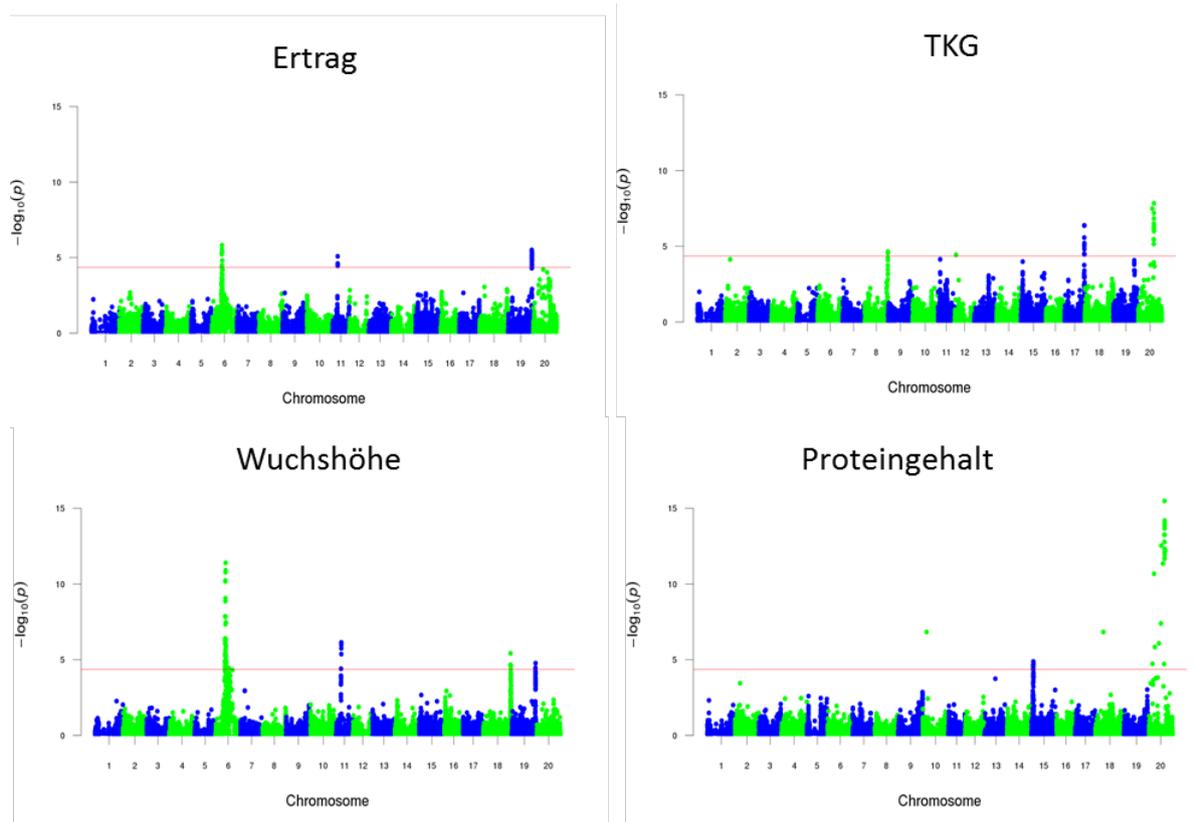


Abbildung 27: Manhattan-Plots für die Merkmale Ertrag, Wuchshöhe, Proteingehalt und TKG basierend auf den Feldversuchen von 2014. Auf der Abszisse sind die Chromosomen aufgereiht und die Ordinate zeigt den Grad der Assoziation an. Jeder Punkt repräsentiert einen SNP, Punkte über dem eingezeichneten Schwellenwert zeigen eine signifikante Assoziation zwischen den SNPs und dem Merkmal.

4.2b QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert

Die Ergebnisse sind ausführlich dargestellt in den Publikationen Kurasch et al. (2018a) und Kurasch et al. (2018b).

Phänotypisierung

Die Studie basierte auf 215 Sojabohnen-RILs aus zwei Familien, die durch Kreuzung der Tofusorte Primus (P2) mit einer Hoch- (Protina, P3) und einer Niedrig-Protein-Sorte (Sigalia, P5) gewonnen wurden. Alle tofubezogenen Merkmale zeigten in etwa eine Normalverteilung in beiden Familien (Abbildung 28). Allerdings unterschieden sich die beiden Familien in ihren Mittelwerten für alle Merkmale mit Ausnahme der Tofuhärte signifikant (siehe Tabelle 6). Mit Ausnahme des Einweichfaktors hatte die Familie P2×P5 höhere mittlere Merkmalswerte. Über alle RILs hinweg reichte die Menge der gewonnenen Sojamilch von 591,7 bis 619,2 Gramm, der Tofuertrag von 136,8 g bis 159,4 g und die Tofuausbeute von 1,89 bis 2,17. Der Gesamtmittelwert für die Tofuhärte betrug 99,30 N, variierend von 84,04 bis 121,90 N. Die beiden Familien unterschieden sich signifikant für den Tofuwert mit einem Durchschnittswert von 420,1 für die Familie P2×P5 und 412,5 für die Familie P2×P3. Die genotypische Varianzkomponente war für alle Merkmale hoch signifikant ($p < 0,001$). Die Ortsvarianz war nur für die Tofufestigkeit und den Tofuwert hochsignifikant ($p < 0,001$), für den Einweichfaktor und die Sojamilch signifikant ($p < 0,01$) und gering signifikant für die Tofuausbeute ($p < 0,05$) und nicht signifikant für den Tofuertrag. Insbesondere jedoch war die Ortsvarianz für alle Merkmale außer Tofuertrag viel höher als die Varianz, die durch den Genotyp erklärt wurde. Die Varianz der Interaktion zwischen Genotyp und Ort war nur für den Einweichfaktor signifikant. Die Versuchsvarianz, die die Reife mit abdeckt, war signifikant für den

Einweichfaktor, Sojamilch, Tofufestigkeit und Tofuwert. Die Laboreffekte des Teams, das die Analyse durchführte, sowie der Tag und die Zeit, zu der die Tofuanalysen durchgeführt wurden, waren für die meisten Merkmale von großer Bedeutung. Die Erbllichkeit war hoch und reichte von 0,65 für Sojamilch bis 0,79 für den Einweichfaktor.

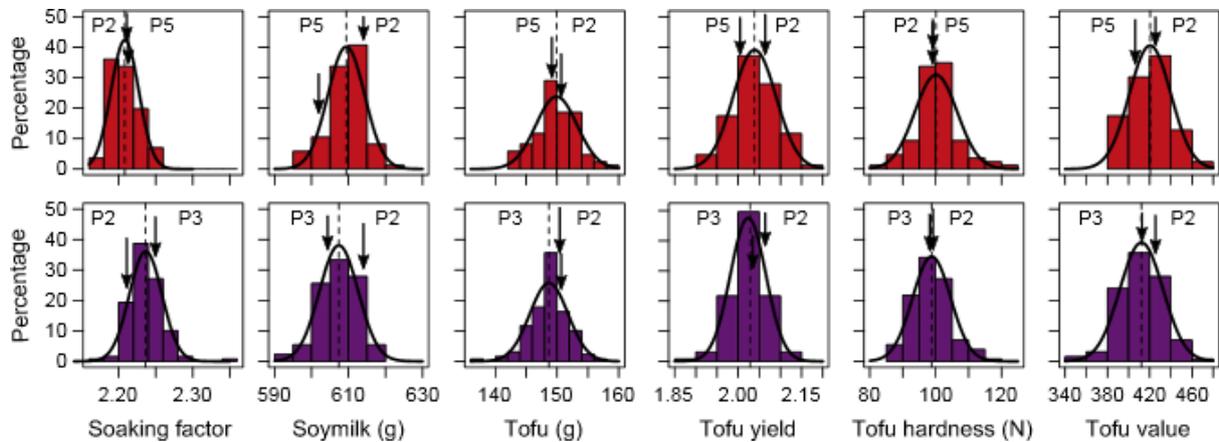


Abbildung 28: Histogramme der beiden Familien, die für die Merkmale Einweichfaktor (Soaking factor), Sojamilch (Soymilk), Tofuertrag (Tofu), Tofuausbeute (Tofu yield), Tofufestigkeit (Tofu hardness) und Tofuwert (Tofu value) dargestellt sind, wobei die Eltern durch Pfeile gekennzeichnet sind (Kurasch et al. 2018a).

Tabelle 6: Zusammenfassende Statistiken für tofubezogene Merkmale in RILs aus zwei Familien (Kurasch et al. 2017).

Quelle†	Quellfaktor	Sojamilch (g)	Tofuertrag(g)	Tofuausbeute (g/g)	Tofufestigkeit (N)	Tofuwert
Mittel _{total}	2,23	608,3	149,2	2,03	99,30	415,4
(Min; Max)	(2,17; 2,34)	(591,7; 619,2)	(136,8; 159,4)	(1,89; 2,17)	(84,04; 121,90)	(347,6; 476,3)
Mittel _{P2×P5}	2,21 ^{a ‡}	609,4 ^a	150,0 ^a	2,04 ^a	100,06 ^a	420,1 ^a
Mittel _{P2×P3}	2,24 ^b	607,7 ^b	148,7 ^b	2,03 ^b	98,85 ^a	412,5 ^b
$\sigma_{Genotyp}^2$	0,0011 ^{***}	32,15 ^{***}	14,94 ^{***}	0,0033 ^{***}	60,94 ^{***}	637,55 ^{***}
σ_{Ort}^2	0,0073 ^{**}	446,22 ^{**}	0,21	0,0045 [*]	1354,58 ^{***}	31195,90 ^{***}
$\sigma_{G \times O}^2$	0,0004 ^{***}	0,00	3,43	0,0016	11,08	88,67
$\sigma_{Versuch}^2$	0,0007 ^{**}	44,55 [*]	0,00	0,0000	50,78 ^{**}	1045,52 ^{**}
σ_{Team}^2	0,0014 ^{***}	417,38 ^{***}	0,45	0,0050 ^{***}	1,52	253,59 ^{***}
σ_{Tag}^2	0,0004 ^{***}	65,79 ^{***}	9,33 ^{***}	0,0027 ^{***}	41,17 ^{***}	3332,23 ^{***}
σ_{Zeit}^2	0,0003 ^{**}	79,12 ^{***}	2,67	0,0016 ^{***}	12,11 [*]	281,50 ^{***}
σ_{Fehler}^2	0,0013	199,15	58,79	0,0094	181,39	1981,50
h^2	0,79	0,65	0,69	0,71	0,72	0,71

† $\sigma_{Genotyp}^2$: genotypische Varianz, σ_{Ort}^2 : Ortsvarianz, $\sigma_{G \times O}^2$: Genotyp-Ort-Interaktionsvarianz, $\sigma_{Versuch}^2$: Reife-Versuch-Varianz, σ_{Team}^2 : Mitarbeitervarianz, σ_{Tag}^2 : Tag der Tofubewertung-Varianz, σ_{Zeit}^2 : Zeit der Tofubewertung-Varianz, σ_{Fehler}^2 : Fehlervarianz und h^2 : Heritabilität.

‡ Werte mit dem gleichen Buchstaben innerhalb der Spalte sind nicht signifikant unterschiedlich, Student's t-Test mit Signifikanzniveau bei $p < 0,05$.

*, **, *** unterscheiden sich signifikant von Null bei der Wahrscheinlichkeitsstufe von 0,05, 0,01 und 0,001.

Der Einweichversuch basiert auf drei Familien, P2×P3, P2×P5 und P3×P5, mit insgesamt 318 RILs. Das Gewicht der Bohnen wurde nach einer Stunde Einweichen in Wasser (1h), nach weiteren 24 Stunden (insgesamt 25h) gemessen und dann die Differenz berechnet (24h). Die Gewichtszunahme nach 1h lag zwischen 6,52 g und 11,2 g und in den folgenden 24 h stieg das Gewicht im Durchschnitt um 18,61 g, zwischen 12,74 g und 23,05 g. Die maximale Gewichtszunahme nach 25 h betrug 33,03 g und die minimale 23,31 g. Die genotypische Varianzkomponente war für alle drei Zeitpunkte hoch signifikant, ebenso wie der Genotyp-Ortseffekt für 1h und 24h. Die Heritabilität für 1h war mit 0,4 moderat, aber höher für 24h und 25h mit 0,56 bzw. 0,57. Die drei Familien unterschieden sich signifikant, wobei die Familie P3×P5 die höchste Gewichtszunahme nach 1 h zeigte, aber die niedrigste für 24 h. Die RILs der Familie P2×P3 zeigten im Durchschnitt eine signifikant höhere Gewichtszunahme nach 24 und 25 Stunden Einweichen als die Linien der beiden anderen Familien. Die drei Elternteile variierten stark in ihren Quellungseigenschaften. P2 nahm das geringste Wasser innerhalb der ersten Stunde auf, aber am meisten nach 24 Stunden. P3 quoll in der ersten Stunde am

stärksten und danach nur mäßig an, zeigte aber insgesamt die höchste Gewichtszunahme. P5 war vergleichbar mit P3 in der ersten Stunde, zeigte aber nach 25 Stunden die geringste Gewichtszunahme insgesamt. P2 und die Nachkommen dieser Sorte zeigten einen Anteil an der Endgewichtszunahme von weniger als 30 % Gewichtszunahme innerhalb der ersten Stunde und mehr als 70 % Gewichtszunahme in den folgenden 24 Stunden. Im Gegensatz dazu wiesen die beiden anderen Eltern und die daraus resultierende Familie P3×P5 einen Anteil von etwa 35 % Gewichtszunahme in der ersten Stunde und etwa 65 % in den folgenden 24 Stunden auf. 1h Einweichen und 24h Einweichen waren moderat, aber signifikant negativ korreliert ($r = -0,27$), während 1h und insgesamt 25h signifikant positiv korreliert waren ($r = 0,25$). Signifikant positiv und stark korreliert waren 24h und die gesamten 25h mit $r = 0,86$.

Die analysierten Merkmale können in vier Merkmalsgruppen gruppiert werden: die agronomischen Merkmale wie Bohnenertrag, Pflanzenhöhe, Tausendkorngewicht und Proteinertrag, Qualitätsmerkmale wie Öl- und Proteingehalt, Tofueigenschaften einschließlich Sojamilch, Tofuertrag- und Tofuausbeute und Einweichmerkmale mit Einweichfaktor nach 1h, 24h und insgesamt 25 h Einweichen. Die Merkmalswerte wurden transformiert, um ihre Verteilung trotz der unterschiedlichen absoluten Merkmalswerte zu vergleichen (Abbildung 29a). Die beiden Qualitätsmerkmale Öl- und Proteingehalt zeigten eine recht breite Streuung über das gesamte Merkmalspektrum. Die agronomischen Merkmale Getreideertrag und Proteinertrag waren sehr ähnlich mit einer höheren Dichte im Mittel, während Pflanzenhöhe und Tausendkorngewicht breiter und gleichmäßiger verteilt waren. Der Einweichfaktor zeigte eine engere, etwas bimodale Verteilung. Sojamilch hingegen zeigte eine viel breitere Streuung, während der Tofuertrag ebenfalls schmal war mit einem deutlichen Peak um den Mittelwert. Die Tofufestigkeit zeigte eine Verteilung mit einer hohen Dichte um den Mittelwert. Der Tofuwert war vergleichbar mit der Tofuausbeute und der Sojamilch mit einer gleichmäßigen Verteilung. Die Einweichmerkmale unterscheiden sich dadurch, dass die 1h Einweichung die breiteste Verteilung zeigte und die 24h eine engere Verteilung mit einer höheren Dichte um den Mittelwert.

Die Netzwerkanalyse basiert auf der Korrelationsmatrix der Merkmale und visualisiert Merkmale oder Merkmalsgruppen, die miteinander verbunden sind, d.h. höhere absolute Korrelationen untereinander aufweisen (Abbildung 29b). Die drei agronomischen Merkmale Proteinertrag, Bohnenertrag und Pflanzenhöhe waren bei allen Individuen sowie in den einzelnen Familien stark miteinander verbunden. Darüber hinaus zeigten diese Eigenschaften keine starke Beziehung zu den anderen Eigenschaften. Tausendkorngewicht hingegen war positiv mit dem Bohnenertrag und der Proteinausbeute in der Familie P2×P3 verbunden, aber in beiden Familien auch mit den Tofueigenschaften Tofuwert und Tofufestigkeit. In der Familie P2×P5 war das Tausendkorngewicht zusätzlich negativ mit dem Einweichfaktor korreliert. Die beiden Qualitätsmerkmale Proteingehalt und Ölgehalt waren stark negativ korreliert. Der Proteingehalt war stark positiv korreliert mit dem Tofuqualitätsmerkmal Tofuwert in der Familie P2×P5 und schwächer mit der Tofuhärte in der Familie P2×P3. Tofufestigkeit und Tofuwert waren sehr stark positiv korreliert. Die Tofufestigkeit war negativ korreliert mit der Tofuausbeute und dem Tofuertrag, der in der Familie P2×P5 stärker ausgeprägt war als in P2×P3. Das Merkmal Sojamilch gruppierte sich nahe dem Tofuertrag und war schwächer, aber auch positiv korreliert mit Tofuwert, 24h Einweichen und Proteingehalt, jedoch nur in der Familie P2×P5. Im Gegensatz dazu war Sojamilch in der Familie P2×P3 mit keinem der anderen Merkmale stark korreliert und wird daher getrennt platziert. Die Einweichmerkmale gruppierten sich dicht beieinander und etwas auseinander von den anderen Merkmalen in der Familie P2×P3, aber weniger in der Familie P2×P5. Die Einweichmerkmale zeigten keinen Zusammenhang mit dem Tausendkorngewicht in den beiden Familien, jedoch war das Tausendkorngewicht über alle Individuen hinweg signifikant negativ korreliert mit dem Einweichfaktor und 25h Einweichen.

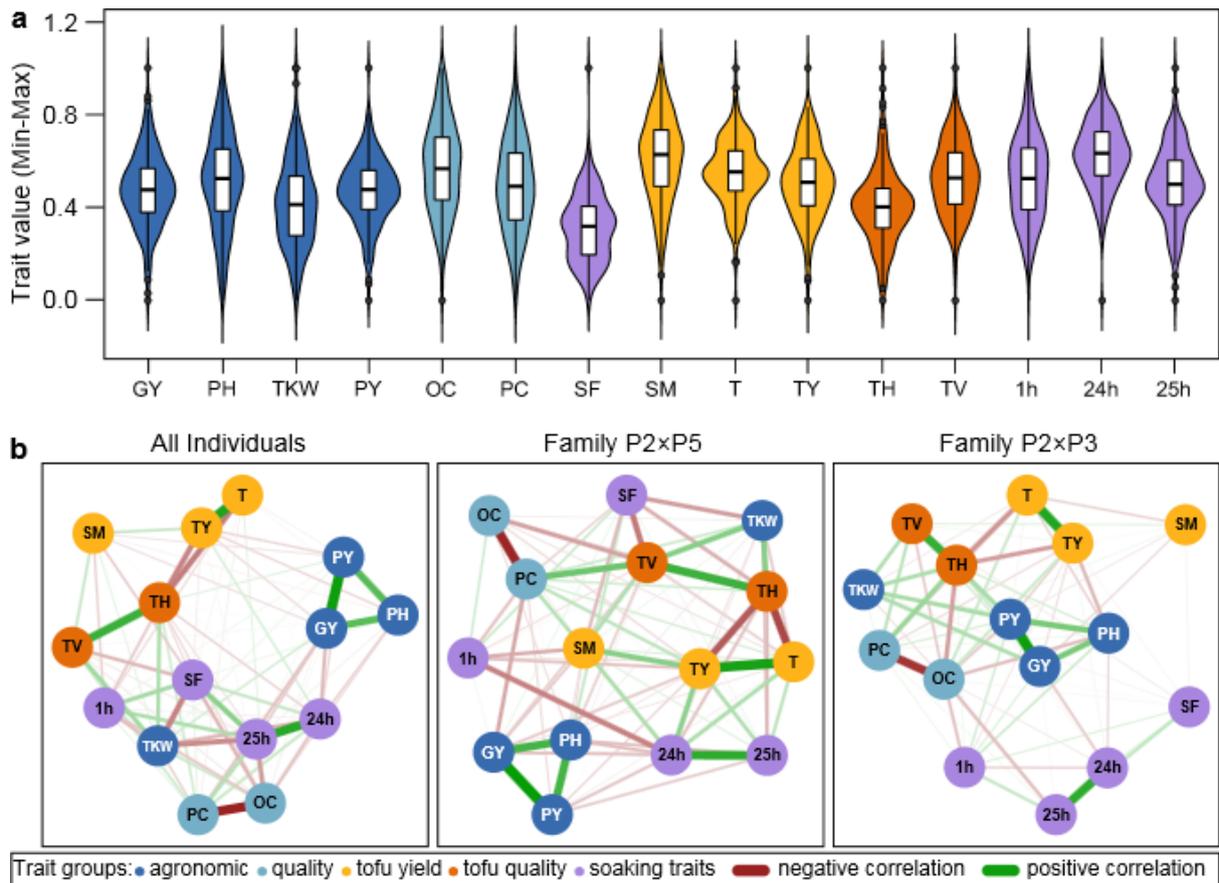


Abbildung 29: a) Violinplots mit transformierten Werten von Bohnerertrag (GY), Pflanzenhöhe (PH), Tausendkorngewicht (TKW), Proteinertrag (PY), Ölgehalt (OC), Proteingehalt (PC), Einweichfaktor (SF), Sojamilch (SM), Tofuertrag (T), Tofuausbeute (TY), Tofufestigkeit (TH), Tofuwert (TV), Einweichen für eine Stunde (1h), Einweichen von einer bis 25 Stunden (24h) und insgesamt 25 Stunden Einweichen (25h). b) Netzwerkanalyse aller bewerteten Merkmale für alle Individuen und für jede Familie separat (Kurasch et al. 2018a)

QTL-Kartierung

Wir führten eine Hauptkoordinatenanalyse basierend auf 26.333 SNP Markern durch. Die ersten beiden Hauptkoordinaten erklärten 34,89 % der Gesamtvariation und die Nachkommen clusterten zwischen ihren Eltern (Abbildung 30).

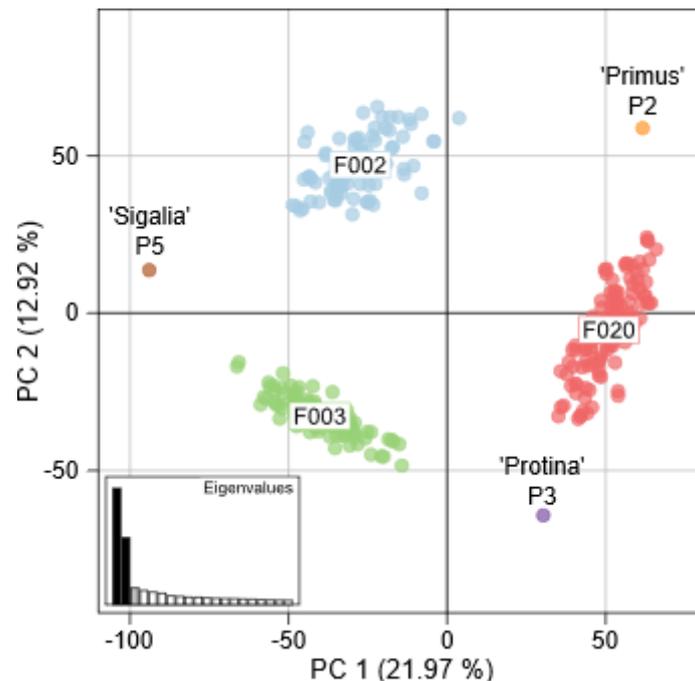


Abbildung 30: Hauptkoordinatenanalyse, die die genetische Verwandtschaft zwischen den Elternlinien P2 'Primus', P3 'Protina' und P5 'Sigalia' und ihren Nachkommen F020, F002, F003 zeigt. Die Zahlen in Klammern beziehen sich auf den Anteil der Varianz, der durch die ersten beiden Hauptkoordinaten erklärt wird (Kurasch et al. 2018b).

Der vollständige Genomscan nach dem Haupteffekt-QTL identifizierte maximal einen QTL pro Merkmal und Population (Tabelle 7, Abbildung 27). Bezüglich der Tofueigenschaften wurde kein QTL für Tofuertrag und Tofuausbeute identifiziert. Nur in der Population "Primus × Protina" wurde ein einziges QTL für das Sojamilchgewicht und ein weiteres für die Tofufestigkeit und den Tofuwert gefunden (Abbildung 31). Einzelne QTL wurden auch für die Einweichmerkmale in jeder Population gefunden, aber sie waren sowohl zwischen den drei Einweichmerkmalen als auch zwischen den Populationen unterschiedlich. In der Population "Primus × Protina" befand sich der QTL für die Proteinkonzentration auf Chromosom 15, während in den beiden anderen Populationen ein einziger QTL für die Proteinkonzentration auf Chromosom 20 gefunden wurde. Der QTL für das Tausendkorngewicht in den Populationen "Primus × Sigalia" und "Protina × Sigalia" wurde auf Chromosom 15 gefunden. Der Anteil der genotypischen Varianz, der durch einen einzigen QTL erklärt wurde, war hoch für die Proteinkonzentration und das Tausendkorngewicht und reichte von 23,6 % bis 27,4 % bzw. 21,6 % bis 29,9 %. Der für S1H identifizierte einzelne QTL zeigte auch sehr hohe Werte für den Anteil der erklärten genotypischen Varianz, die von 42,4 % bis 74,1 % für die drei Populationen reichten. In der Population "Primus × Sigalia" stammten die QTL-Allele mit den positiven Wirkungen alle von "Primus". In der Population "Primus × Protina" war der Donor des günstigen Allels für Proteinkonzentration, Tofufestigkeit und Tofuwert "Protina", während es für das Tausendkorngewicht, das Sojamilchgewicht und die Einweichmerkmale S1h und S25h "Primus" war. In der Population "Protina × Sigalia" lieferte "Protina" das günstige Allel für alle fünf Merkmale mit Ausnahme von S25h.

Tabelle 7: QTL-Positionen auf Chromosomen (Pos.) und deren unterstützende Intervalle (SI), LOD-Wert (LOD), Effektgröße, erklärte genotypische Varianz (p_G) und Frequenz aus der Kreuzvalidierung (Freq. CV) für Proteinkonzentration (PC), Tausendkorngewicht (TKW), Sojamilchgewicht (SM), Tofuertrag (T), Tofuausbeute (TY), Tofufestigkeit (TH), Tofuwert (TV), Gewichtszunahme nach einer Stunde Einweichen (S1H), nach weiteren 24 Stunden (S24H) und Gesamtgewichtszunahme nach 25 Stunden (S25H) (Kurasch et al. 2018b).

Merkmal	Chr.‡	Pos. (SI)‡ [cM]	LOD‡	Donor fav. Allele	Add Eff.	p_G [%]‡	Freq. CV
‘Primus × Protina’ (F020)							
PC	15	20 (18-22)	9,77	Protina	0,60	24,77	0,67
TKW	17	246 (244-248)	10,75	Primus	7,92	22,19	0,45
SM	5	92 (84-100)	3,73	Primus	2,08	18,84	0,38
T	-	-	-	-	-	-	-
TY	-	-	-	-	-	-	-
TH	8	62 (58-66)	3,55	Protina	2,15	16,12	0,34
TV	8	64 (62-66)	3,29	Protina	7,28	15,07	0,23
S1H	1	62 (60-66)	5,12	Primus	0,28	42,40	0,74
S24H	-	-	-	-	-	-	-
S25H	4	68 (60-78)	3,06	Primus	0,41	10,9	0,04
‘Primus × Sigalia’ (F002)							
PC	20	80 (74-88)	3,64	Primus	0,70	23,57	0,21
TKW	15	156 (154-158)	6,86	Primus	8,53	23,19	0,48
SM	-	-	-	-	-	-	-
T	-	-	-	-	-	-	-
TY	-	-	-	-	-	-	-
TH	-	-	-	-	-	-	-
TV	-	-	-	-	-	-	-
S1H	14	198 (196-200)	5,23	Primus	0,36	42,56	0,19
S24H	20	112 (102-118)	3,90	Primus	0,74	32,86	0,24
S25H	20	158 (152-160)	3,35	P2	0,58	27,61	0,13
‘Protina × Sigalia’ (F003)							
PC	20	90 (86-92)	5,80	Protina	0,92	27,41	0,84
TKW	15	348 (344-350)	5,97	Protina	9,25	21,59	0,53
S1H	15	208 (206-214)	7,55	Protina	0,42	74,18	0,67
S24H	6	174 (168-184)	3,78	Protina	0,42	18,82	0,10
S25H	11	96 (94-100)	4,66	Sigalia	0,48	30,32	0,31

‡ Chr. = Chromosom; Pos.= genetische Position mit unterstützenden Intervallen (SI); LOD (Log10 des Likelihood Odds Ratios) die Wahrscheinlichkeit, die mit der wahrscheinlichsten Position des detektierten QTL verbunden ist; p_G = Anteil der erklärten genetischen Varianz; Freq. CV = Häufigkeiten des QTL-Vorkommens in der fünffachen Kreuzvalidierung mit 1000 Läufen.

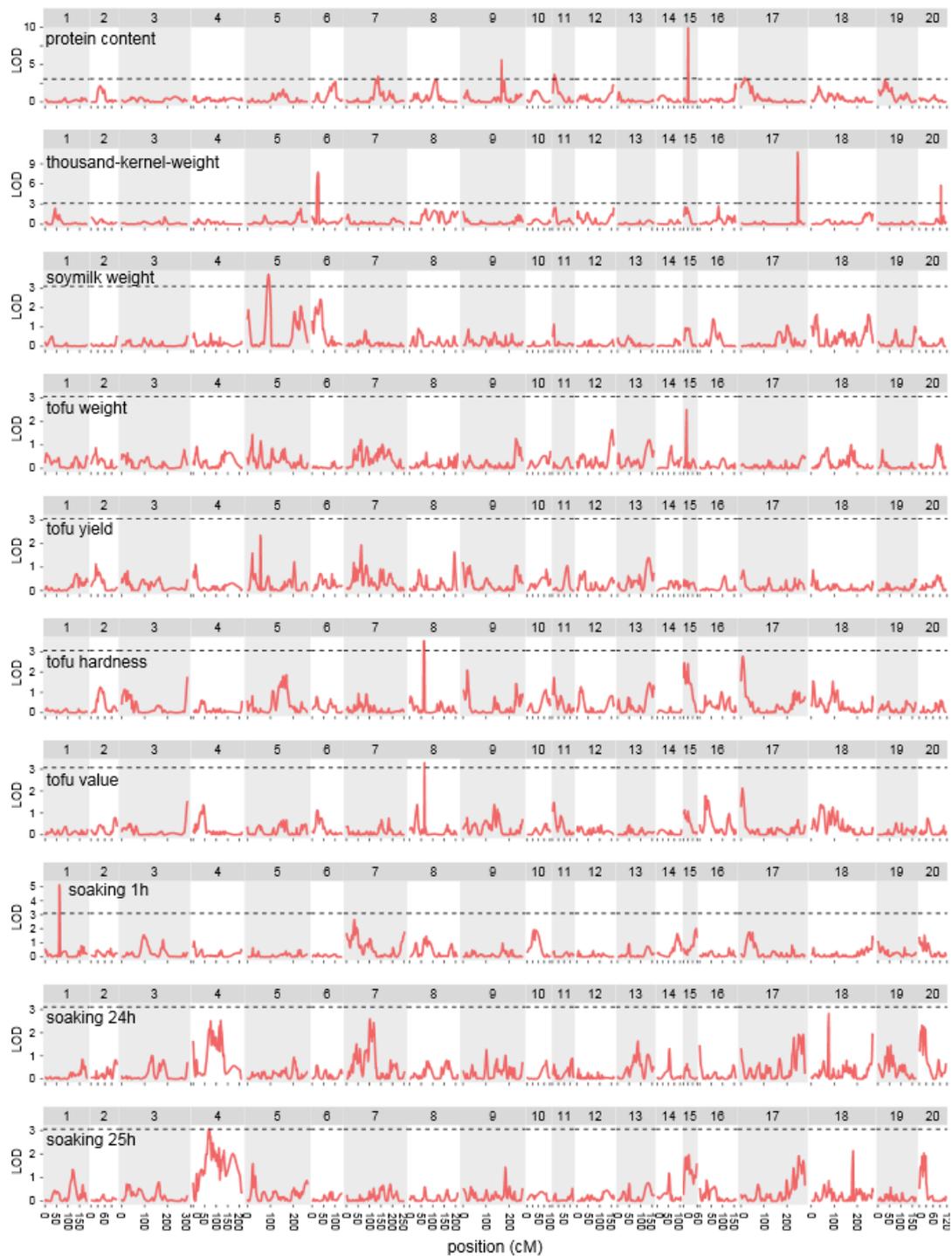


Abbildung 31: LOD-Kurven für F020 'Primus × Protina' und die untersuchten Merkmale Proteinkonzentration (%), Tausendkorngewicht (g), Sojamilchgewicht (g), Tofuertrag (g), Tofuausbeute, Tofufestigkeit (N), Tofuwert, Gewichtszunahme nach Einweichen 1h (g), Gewichtszunahme nach Einweichen 24h (g) und Gesamtgewichtszunahme nach Einweichen 25h (g). Der signifikante LOD-Schwellenwert wird durch die gestrichelte Linie angezeigt und wurde durch einen Permutationstest mit 5000 Läufen pro Merkmal bestimmt (Kurasch et al. 2018b).

4.2c Genomische Selektion

Agronomische Merkmale

Es wurde ein genomisches Vorhersagemodell über alle Familien für die agronomischen Merkmale Ertrag, Proteingehalt, Ölgehalt, Pflanzenhöhe und Tausendkorngewicht erstellt. Dieses Modell konnte sehr gute Vorhersagefähigkeiten (Predictive ability) für alle fünf Merkmale erreichen. Die beiden Merkmale Pflanzenhöhe und Tausendkorngewicht erreichten hierbei die höchsten Werte mit einer Genauigkeit über 0.8 (Abbildung 32). Jedoch konnte auch für den Ertrag eine sehr gute Vorhersage mit >0.65 erreicht werden.

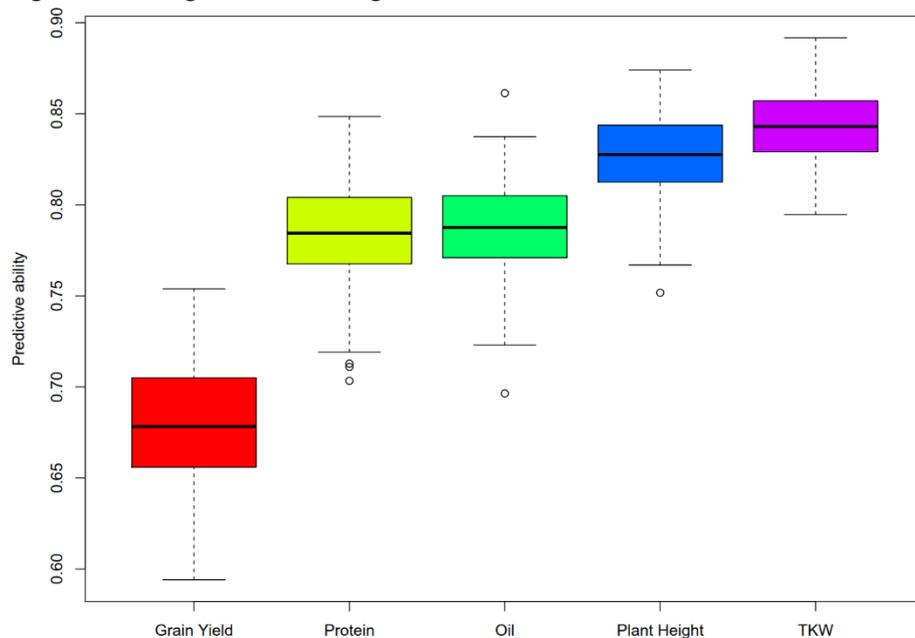


Abbildung 32: Genomische Vorhersagefähigkeit für fünf agronomische Merkmale evaluiert im Jahr 2014 für 1008 RIL.

Zusätzlich zur Vorhersage innerhalb des Jahres 2014, wurde auch die Vorhersagefähigkeit des Modells von 2014 auf die Daten von 2016 überprüft. Hierfür wurden die Öl- und Proteinwerte mit einer Korrelation von 0,6 bzw. 0,58 innerhalb der 1008 Genotypen vorhergesagt. Als weitere Überprüfung des Modells basierend auf den 1008 Genotypen aus 2014, wurde die Leistung der zusätzlich genotypisierten 800 Linien für das Jahr 2016 mittels dieses Modells vorhergesagt. Auch hier konnten für die Merkmale Öl und Proteingehalt sehr gute Korrelationen zwischen wahrem und vorhergesagtem Wert (~ 0.55) erreicht werden. Diese Ergebnisse zeigen, dass die genomische Selektion über Jahre und über Halbgeschwister-Familien in Soja ein sehr großes Potential hat, welches effizient in der Züchtung eingesetzt werden kann.

Tofueigenschaften

Die Genauigkeit der genomischen Vorhersage wurde für alle Merkmale in jeder Population mittels fünffacher Kreuzvalidierung bewertet (Abbildung 33). Wir fanden hohe Vorhersagegenauigkeiten für Proteinkonzentration, Tausendkorngewicht und Gewichtszunahme nach 1 Stunde Einweichen und moderate Vorhersagegenauigkeiten für die beiden anderen Einweichmerkmale. Im Gegensatz dazu zeigten die fünf Tofueigenschaften nur eine geringe bis mittlere Vorhersagegenauigkeit. Die erhaltenen Vorhersagegenauigkeiten waren im Allgemeinen zwischen den Populationen vergleichbar. Zusätzlich zur Vorhersage innerhalb der Populationen haben wir die Vorhersagegenauigkeit über die Populationen hinweg und von einer Population zur anderen bewertet. Die Zusammenführung

der Populationen führte zu einer leichten Erhöhung der Vorhersagegenauigkeit für alle Merkmale und zu weniger Schwankungen, während die Vorhersagegenauigkeiten für die Vorhersage von einer Population zur anderen eher gering waren.

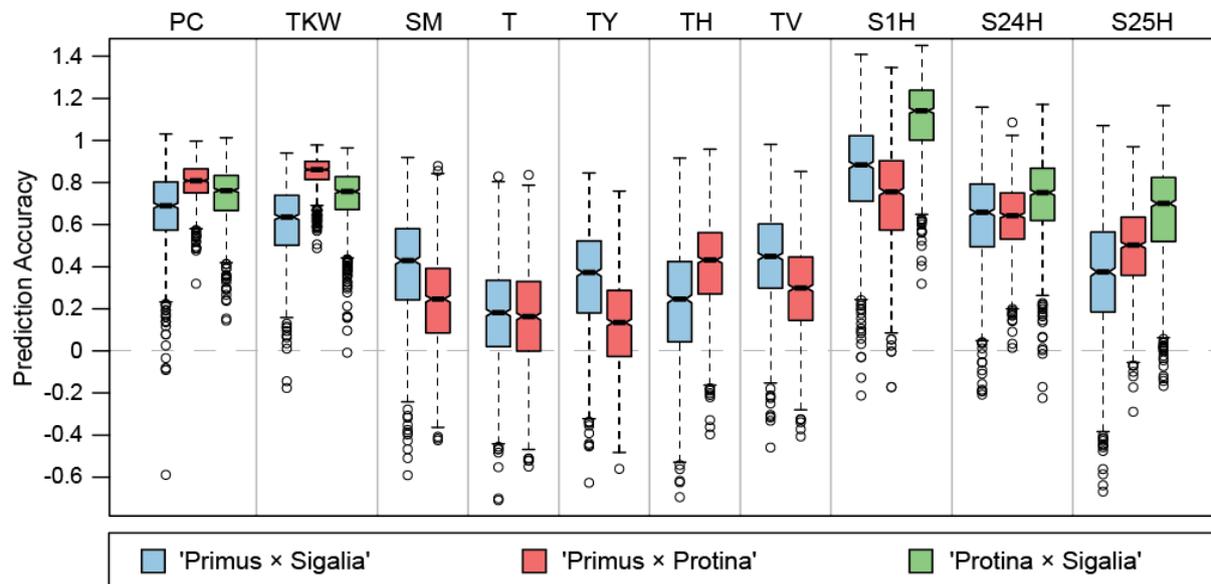


Abbildung 33: Gekerbte Boxplots für die Vorhersagegenauigkeit innerhalb von Populationen für die Proteinkonzentration (PC), Tausendkorngewicht (TKW), Sojamilch-Gewicht (SM), Tofuertrag (T), Tofuausbeute (TY), Tofufestigkeit (TH), Gewichtszunahme nach Einweichen für 1 Stunde (S1H), Gewichtszunahme nach Einweichen für 24 Stunden (S24H) und Gesamtgewichtszunahme nach Einweichen für 25 Stunden (S25H) (Kurasch et al. 2018b).

4.2d,e Kreuzung und Weiterführung Diallel1_neu und Diallel2_neu

Wie geplant wurden Kreuzungen durchgeführt und die Nachkommen wurden selektiert und weiterentwickelt. Mehrere Stämme wurden an private Zuchtunternehmen abgegeben.

4.2f QTL-Kartierung der Kühletoleranz

Ist dargestellt in Kapitel 4.4

4.2g Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus AP1

Die Kreuzungen wurden im Sommer im Feld oder in Folienhäusern durchgeführt. Kreuzungen sind bei Sojabohnen generell sehr schwer durchzuführen, da der Kreuzungsansatz nur sehr niedrig ist. Noch schwieriger ist dies mit genetischen Ressourcen, da dort die Blüten oft nur sehr klein sind und es häufig schwierig ist, die zu kreuzenden Pflanzen synchron zur Blüte zu bringen. Daher konnten nur wenige Kreuzungen erfolgreich weitergeführt werden. Ausgewählt wurden genetische Ressourcen, die sich hinsichtlich der Samenqualität unterschieden (z.B. verändertes Fettsäuremuster oder weniger allergenes Potential). Diese wurde mit Sorten bzw. Elitelinien gekreuzt. Die Nachkommen werden weitergeführt.

4.2h Sortenprüfungen

In den Sortenprüfungen wurden jeweils aktuelle Sorten auf ihren Ertrag und ihren Proteingehalt geprüft. In Tabelle 8 sind die 5 ertragreichsten Sorten pro Jahr und Standort dargestellt. Das ertragreichste Jahr war 2017 mit Höchsterträgen bis über 47 dt/ha.

Tabelle 8: Ergebnisse der Sortenprüfungen. Dargestellt sind der Ertrag und der Proteingehalt je bei 100% Trockensubstanz der jeweils 5 ertragreichsten Sorten an den Standorten Hohenheim (HOH) und Eckartsweier (EWE).

Ort	Name	Ertrag (dt/ha)	Protein-gehalt (%)	Ort	Name	Ertrag (dt/ha)	Protein-gehalt (%)
2015				2017			
HOH	Cardiff	25,1		HOH	ESMentor	47,4	43,1
	OACWallace	23,7			Stumpa	47,0	42,3
	Sultana	22,6			Marquise	44,6	43,3
	Soprana	22,2			Sultana	42,7	42,8
	Primus	21,9			Sirelia	42,5	41,8
EWE	Sigalia	40,3	40,9	EWE	ESMentor	47,3	41,9
	OACWallace	37,2	39,0		Stumpa	46,7	39,8
	SYElliot	35,9	40,9		Marquise	46,9	41,2
	ESMentor	34,4	41,7		Regina	45,7	42,5
	Soprana	32,9	40,7		SilviaPZO	45,2	39,7
2016 früh				2018 früh			
HOH	Gallice	33,5	35,5	HOH	Tourmaline	35,8	37,2
	Viola	33,3	38,1		ESComandor	35,0	39,6
	Sirelia	32,7	35,2		Sculptor	35,5	40,2
	RGTShouna	32,7	36,0		Alexa	35,0	38,0
	Lissabon	32,2	35,8		Acardia	34,8	36,2
EWE	Meteor	44,9	43,8	EWE	Obelix	35,6	39,8
	Tourmaline	41,4	38,8		Marquise	33,8	40,7
	Coraline	40,4	38,5		Lissabon	32,3	38,6
	SYLivius	39,9	42,1		Solena	31,5	40,1
	Sirelia	39,9	39,5		Acardia	30,9	38,9
2016 spät				2018 spät			
HOH	Soprana	29,5	37,1	HOH	Atacama	37,2	36,9
	SilviaPZO	29	33,4		ESMentor	35	40,6
	Galice	28,2	35,4		Yakari	34,7	38,9
	SYElliot	27,6	36,2		SilviaPZO	34,4	33,1
	Tequila	26,8	36,1		Soprana	33,8	38,7
EWE	SYElliot	41,8	40,7	EWE	RGTSicroca	33,3	41
	Naya	40,8	40,4		Idefix	33	41
	Tequila	39,4	39,1		Angelica	32,9	40,1
	ESComandor	39,1	40,5		Yakari	32,6	41,2
	SilviaPZO	38,8	38,6		Svela	31,9	40,6

4.2i Speed-Breeding-System

Die Speed-Breeding-Kammer wurde etabliert (Abbildung 34). Das dazugehörige Protokoll wurde publiziert (Jähne et al. 2020), die Details können in der Publikation nachgelesen werden. In der Kammer wurden Kreuzungen erstellt und es wurde Material durch ssd weitergeführt. Es hat sich jedoch herausgestellt, dass für die Durchführung von Kreuzungen und die

Erstellung schneller Generationen unterschiedliche Bedingungen in der Kammer herrschen müssen. Um möglichst schnell zu sein, sollte die Temperatur beim ssd-Verfahren ca. 30°C betragen. Dies führt jedoch zu sehr kleinen Blüten und einer geringen Pollenschüttung und macht eine Durchführung von Kreuzungen unmöglich. Um Pflanzen mit größeren Blüten zu erzeugen, muss die Temperatur abgesenkt werden. Da nur eine Kammer zur Verfügung steht, wurde die Temperatur zur Durchführung von Kreuzungen angepasst. Dadurch werden jedoch maximal 4 Generationen ssd pro Jahr möglich.

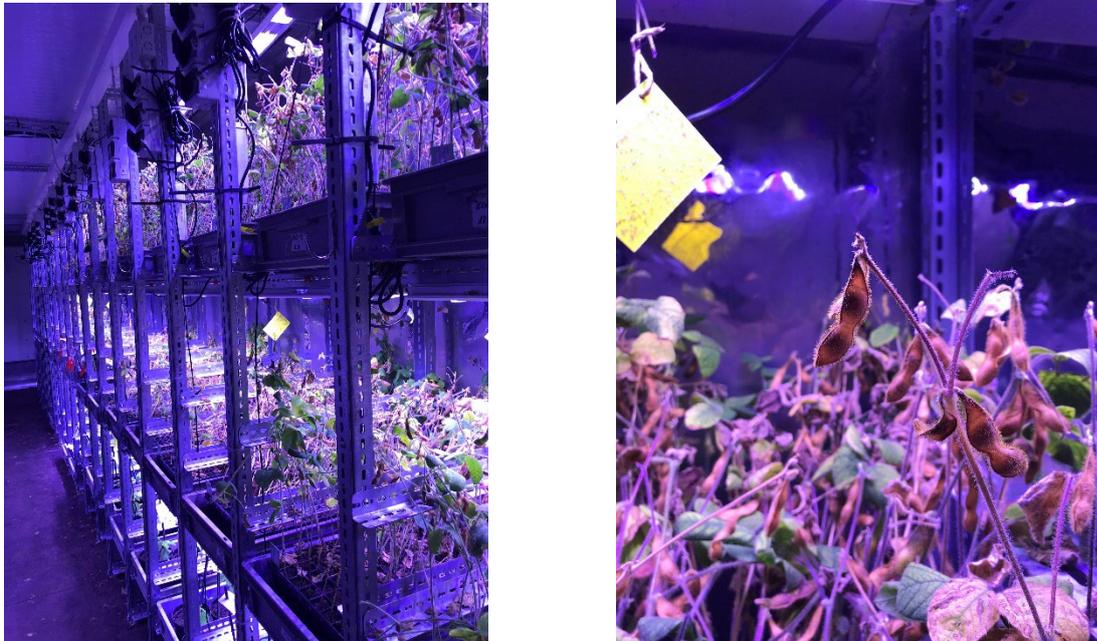


Abbildung 34: links: Speed-Breeding- Kammer mit 3 Regalebene, rechts: abreifende Sojabohnen in der Kammer.

Im Sommer 2020 wurden 1900 Reihen im Zuchtgarten ausgesät. Davon wurden 200 Genotypen selektiert und in Leistungsprüfungen 2021 geprüft. Die Ergebnisse wurden den Züchterhäusern zur Verfügung gestellt. Neun Genotypen wurden von den Züchtern abgenommen. Diese werden jetzt von den Züchtern weiterentwickelt und gegebenenfalls für die Wertprüfung angemeldet. Weitere Genotypen werden in 2022 in mehrortige Leistungsprüfungen gestellt und anschließend wieder den Züchterhäusern angeboten.

In der Speed-Breeding Kammer wurden ca. 6000 Pflanzen weitergeführt. Diese kommen 2022 auf das Feld um eine Einzelpflanzenselektion durchzuführen. Ein großer Teil des Winterzuchtgartens 2021/22 wurde in der Kammer durchgeführt und nicht wie sonst in Costa Rica. Die Kammer ist daher inzwischen voll in das Zuchtprogramm integriert.

4.2k Ionomics

Mit dem für das Projekt beschafften XRF-Gerät wurden gemahlene und gepresste Sojabohnen untersucht. Für die Elemente Magnesium, Phosphor, Schwefel, Kalium, Kalzium, Mangan, Eisen, Nickel, Kupfer, Zinn und Molybdän wurden Kalibrationen erstellt.

Für den Standort Eckartsweier 2020 wurden von Proben von fünf Leistungsprüfungen, die mit zwei Wiederholungen angebaut wurden, die Konzentrationen der messbaren Elemente untersucht. Bei den Prüfgliedern der Prüfungen handelte es sich um fortgeschrittenes Zuchtmaterial kurz vor der Abgabe oder fertige Sorten.

Es zeigte sich, dass in der Mehrzahl der Leistungsprüfungen die 2020 durchgeführt wurden, die Konzentration der untersuchten Elemente signifikante Unterschiede zwischen den Genotypen aufwies, mit Ausnahme von Molybdän, hier wurden keine signifikanten Unterschiede zwischen den Genotypen nachgewiesen (Tabelle 9). Die Mittelwerte und

Streubreiten der einzelnen Elemente waren auch im Erntematerial der Leistungsprüfungen aus 2021 zu finden (Tabelle 10). Nur für Molybdän wurden 2021 deutlich höhere Konzentrationen ermittelt.

Tabelle 9: Mittel-, Maximum- und Minimumwerte sowie Signifikanzniveau der genetischen Varianz der Proben der angegebenen Leistungsprüfungen in 2020 für die angegebenen Merkmale.

LP		Ertrag (dt/ha)	Protein- gehalt (%)	Öl- gehalt (%)	K	P	S	Ca ppm	Mg	Fe	Zn	Mn	Ni	Cu	Mo
601	Mittel	25,5	39,9	19,8	16590	6049	3158	2128	2109	77	52	27	12	12	3
n = 25	Max	33	44,1	21,2	17998	6590	3415	2489	2242	86	57	30	14	14	3
	Min	19,1	36,6	17,8	14573	5568	2865	1921	1988	68	46	23	11	11	2
	gen. Var.	ns	**	**	**	**	**	**	**	**	**	**	**	**	ns
602	Mittel	24,6	41,2	19,2	17146	4379	3124	2052	2105	74	50	26	11	12	3
n = 50	Max	32,3	46,1	20,3	18268	6117	3423	2503	2330	87	58	33	15	14	3
	Min	14,5	38,7	17	15900	2794	2794	1525	1971	63	39	22	6	10	2
	gen. Var.	**	**	**	**	**	**	**	**	**	**	**	**	**	ns
603	Mittel	27,4	39,5	19,9	19870	5600	3078	2139	2075	75	50	26	13	11	3
n = 25	Max	32,7	45,4	21,1	17698	5997	3407	2595	2209	81	57	31	15	13	3
	Min	21,6	37,2	17,4	15743	5200	2843	1849	1936	68	42	24	11	11	2
	gen. Var.	**	**	**	**	**	**	**	**	**	**	**	**	**	ns
604	Mittel	31,2	40,9	19,2	17242	5429	3117	2111	2068	71	49	26	12	11	3
n = 50	Max	40,8	46,0	20,9	18894	6427	3483	2671	2235	85	56	32	15	13	3
	Min	21,7	35,4	17,4	16057	4994	2810	1770	1872	62	41	22	10	10	2
	gen. Var.	**	**	**	**	**	**	**	**	**	**	**	**	**	ns
711	Mittel	24,6	39,9	19,7	16581	5857	3103	2294	2069	74	51	26	10	12	3
n = 36	Max	36,3	42,9	21,3	17833	6496	3391	2737	2228	82	56	32	14	14	4
	Min	12,9	36,0	18,5	14726	5326	2888	1789	1897	68	45	23	8	10	2
	gen. Var.	*	ns	ns	ns	ns	*	*	**	ns	ns	*	ns	ns	ns

*, ** = signifikant bei P = 0,05 und 0,01, ns = nicht signifikant

Tabelle 10: Mittel-, Maximum- und Minimumwerte von Proben aus der Ernte 2021 für die angegebenen Merkmale und die aufgeführten Elemente (in ppm).

	Ertrag (dt/ha)	Protein- gehalt (%)	Öl- gehalt (%)	Höhe (cm)	Reife (d)	K	P	S	Ca	Mg	Fe	Zn	Mn	Ni	Cu	Mo
Mittel	38,3	45,3	16,8	112	266	16515	5659	2944	1552	2001	68	20	37	4	13	9
Max	55,4	50,2	19,1	150	274	17988	6530	3379	2021	2354	82	25	47	6	16	13
Min	23,7	39,0	14,8	90	246	14592	5009	2462	1100	1733	53	15	30	2	11	6

Über alle Untersuchungen zusammengefasst wurden die Korrelationen zwischen den erhobenen Merkmalen und der Konzentration der einzelnen Elemente ermittelt (Abbildung 35).

Hierbei wurde ersichtlich, dass der Ertrag negativ korreliert ist mit dem Zink-, Calcium-, Phosphor- und Schwefelgehalt und positiv mit dem Molybdän- und Mangangehalt. Der Proteingehalt war positiv mit dem Molybdän-, Mangan-, Kupfer- und Schwefelgehalt und negativ mit dem Kalzium-, Zink- und Nickelgehalt korreliert.

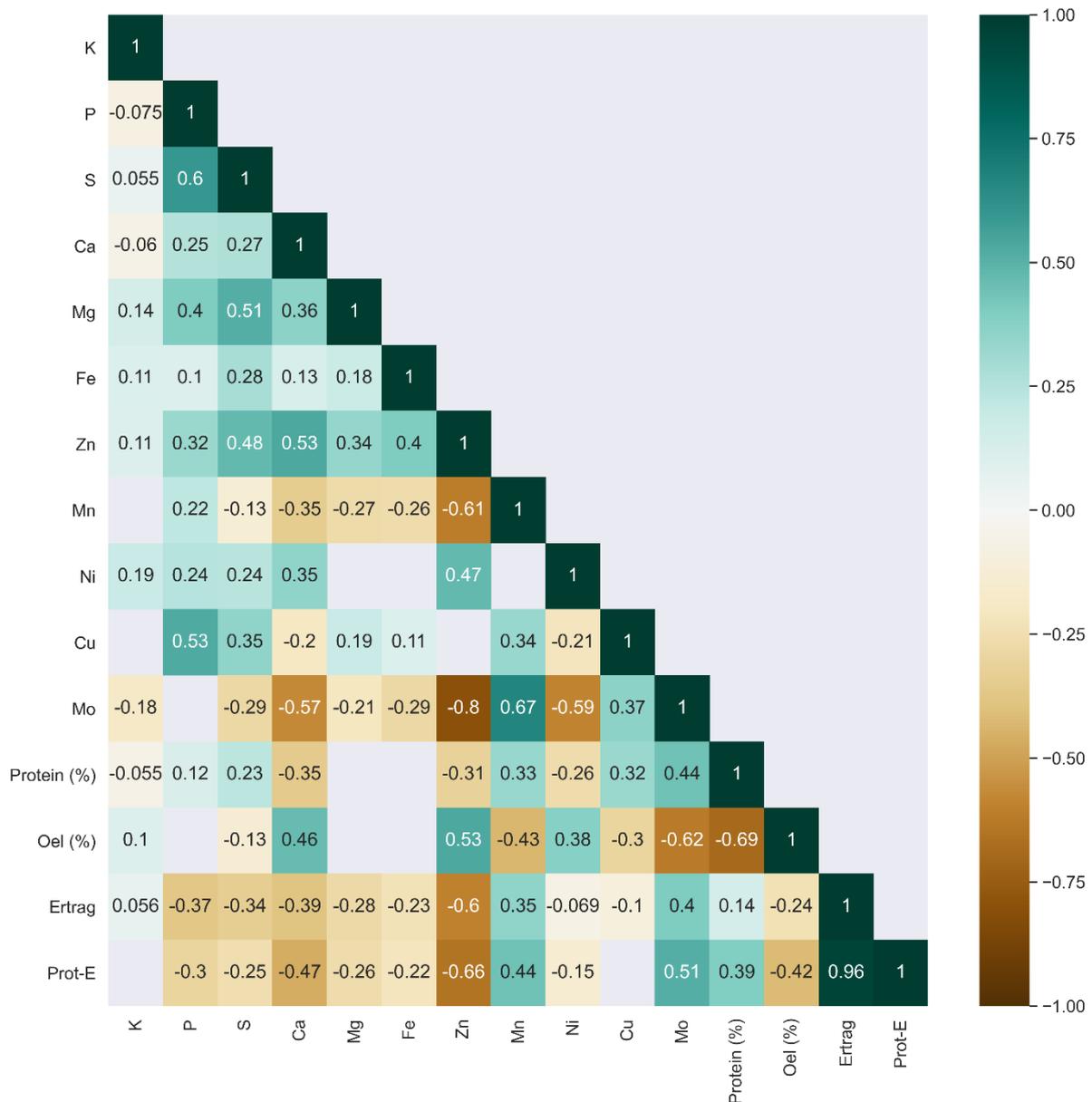


Abbildung 35: Heatmap der Korrelationen zwischen den Konzentrationen (%) der angegebenen Elemente und den Merkmalen Ölgehalt, Proteingehalt, Ertrag und Proteinertag zusammengefasst aus dem Anbau 2020 und 2021. Dargestellt sind nur signifikante ($P < 0,05$) Korrelationen.

Insgesamt war der Ertrag meist negativ mit den Konzentrationen der verschiedenen Elemente korreliert. Dies lässt darauf schließen, dass es sich um einen Verdünnungseffekt handelt. Je höher der Ertrag ist, umso geringer werden die Konzentrationen der einzelnen Elemente. Da der Ertrag sich zu einem großen Teil aus Öl und Kohlenhydraten zusammensetzt, die nur C-, H- und O-Atome enthalten lässt sich dies gut erklären. Anders sieht es beim Proteingehalt aus. Die unterschiedlichen Aminosäuren enthalten unterschiedliche Anteile an Schwefel, so dass bei einem höheren Proteingehalt auch die Schwefelkonzentration steigen kann. Hohe

Phosphorgehalte sind nicht erwünscht, da diese teilweise negative Effekte auf die tierische und menschliche Ernährung aufweisen. Das Korrelationsnetzwerk der untersuchten Elemente (Abbildung 36) zeigt, dass die Kaliumkonzentration im Samen nicht von den anderen Elementen beeinflusst wird. Bei den anderen Elementen gibt es unterschiedliche Beziehungen. So sind beispielsweise der Schwefel- und Phosphorgehalt positiv korreliert, die Molybdän- und Zinkgehalte dagegen negativ. Weitere Untersuchungen müssen in den nächsten Jahren zeigen, wie sich dies an unterschiedlichen Orten und unter unterschiedlichen klimatischen Bedingungen verhält. Wenn man die ermittelten Konzentrationen mit dem Ertrag multipliziert, erhält man die Mengen, die je Hektar aufgenommen wurden. Diese sind nahezu alle positiv korreliert, das heißt, bei einem höheren Samenertrag wird auch von den einzelnen Elementen mehr je Hektar aufgenommen, die Konzentration je Samen sinkt jedoch. Der Proteingehalt der Samen wird positiv vom Molybdängehalt und gering positiv vom Schwefelgehalt beeinflusst. Beim Ölgehalt gibt es positive Beziehungen zum Zink- und Calciumgehalt, stark negative Beziehungen zum Molybdängehalt und gering negative zum Mangan- und Kupfergehalt.

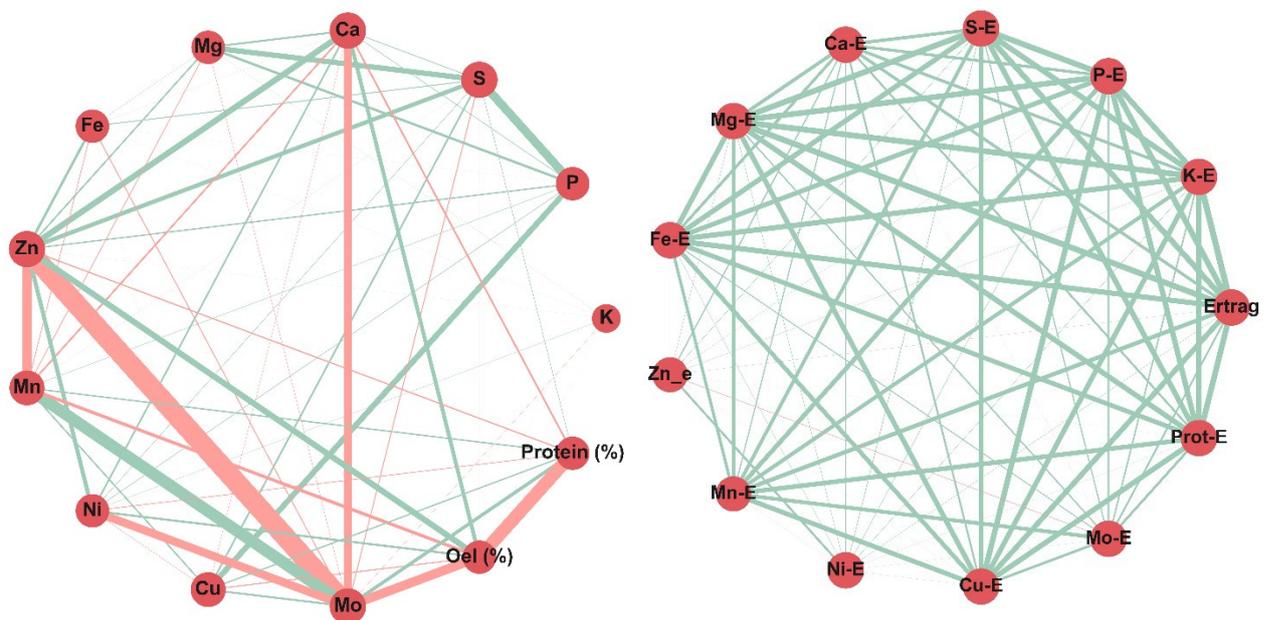


Abbildung 36: Korrelationsnetzwerk der untersuchten Elemente aller untersuchten Proben (links, Konzentration in den Samen, rechts Erträge) mit dem Öl- und Proteingehalt (links) und dem Ertrag und Proteinertrag (rechts), grün = positive, rot = negative Beziehungen.

4.2I Genotypisierung

Insgesamt wurden 540 Sojalinien mit 192,8 G an Rohdaten generiert. In diesen genotypisierten Linien wurden insgesamt 54376 SNPs gefunden.

Des Weiteren wurden publizierte Marker im Labor etabliert sowie KASP Marker für Kreuzungserfolg anhand von SNP Frequenzen entwickelt (Tabelle 11). SNPs für Kreuzungsmarker sollten eine Major Allel Frequenz von etwa 0.5 haben um für eine große Anzahl an Linien anwendbar zu sein.

Ein Beispiel für einen etablierten Marker ist der Längenpolymorphismus im Gen Glyma.20G085100 (Fliege et al. 2022). Dieser Längenpolymorphismus kann nach einer PCR auf einem Agarosegel visualisiert werden und ermöglicht es Individuen zu identifizieren die das Allel für einen hohen Proteingehalt tragen (Abbildung 37). Die KASP Kreuzungsmarker, welche im Labor anhand von SNP Daten entwickelt wurden, erlauben es den Erfolg von

Kreuzungen zwischen verschiedenen Linien in einem frühen Wachstumsstadium der F1 Pflanzen eindeutig zu bestimmen (Abbildung 38).

Tabelle 11: Etablierte KASP und PCR Marker. Bei einigen PCR Markern wird zusätzlich ein enzymatischer Verdau benötigt um Allele zu unterscheiden (PCR + RV).

Marker	Lokus	Merkmal	Technik	Referenz
dt1-ab	Dt1	determinate stem	PCR + RV	Xu et al. 2013
dt1-bb	Dt1	determinate stem	PCR + RV	Xu et al. 2013
dt1-tb	Dt1	determinate stem	PCR + RV	Xu et al. 2013
e1-as	E1	Blüte / Reife	PCR + RV	Xu et al. 2013
e1-fs/e1-nl	E1	Blüte / Reife	PCR + RV	Xu et al. 2013
KASP E1-as	E1	Blüte / Reife	KASP	LSA *
e2	E2	Blüte / Reife	PCR + RV	Xu et al. 2013
E2_InDel	E2	Blüte / Reife	PCR	Tsubokura et al. 2014
KSP_E2_ns	E2	Blüte / Reife	KASP	LSA *
E3_Mse1	E3	Blüte / Reife	PCR + RV	Tsubokura et al. 2014
e3-fs	E3	Blüte / Reife	PCR + RV	Xu et al. 2013
e3-ns	E3	Blüte / Reife	PCR + RV	Xu et al. 2013
e3-tr	E3	Blüte / Reife	PCR	Xu et al. 2013
e4-kam	E4	Blüte / Reife	PCR + RV	Xu et al. 2013
e4-kes	E4	Blüte / Reife	PCR + RV	Xu et al. 2013
e4-oto	E4	Blüte / Reife	PCR + RV	Xu et al. 2013
e4-SORE-1	E4	Blüte / Reife	PCR	Xu et al. 2013
e4-tsu	E4	Blüte / Reife	PCR + RV	Xu et al. 2013
KSP_SoyCdm	GmHMA3	Cadmiumgehalt	KASP	Wang et al. 2012
GmHS1-1	Glyma02g43700	Hard seeds	KASP	Sun et al. 2015
Glyma.20G085100	Glyma.20G085100	High protein	PCR	Fliege et al. 2022
GM01_53333405	NA	Kreuzung	KASP	LSA
GM02_39084542	NA	Kreuzung	KASP	LSA
GM02_48352600	NA	Kreuzung	KASP	LSA
GM02_707754	NA	Kreuzung	KASP	LSA
GM03_2449715	NA	Kreuzung	KASP	LSA
GM03_34395833	NA	Kreuzung	KASP	LSA
GM04_18565313	NA	Kreuzung	KASP	LSA
GM04_50871044	NA	Kreuzung	KASP	LSA
GM05_35179258	NA	Kreuzung	KASP	LSA
GM05_35908463	NA	Kreuzung	KASP	LSA
GM06_47429496	NA	Kreuzung	KASP	LSA
GM06_49090708	NA	Kreuzung	KASP	LSA
GM07_36241553	NA	Kreuzung	KASP	LSA
GM07_6996323	NA	Kreuzung	KASP	LSA
GM08_46378251	NA	Kreuzung	KASP	LSA
GM08_7517453	NA	Kreuzung	KASP	LSA
GM09_47335027	NA	Kreuzung	KASP	LSA

GM09_47367366	NA	Kreuzung	KASP	LSA
GM10_51069308	NA	Kreuzung	KASP	LSA
GM10_9793830	NA	Kreuzung	KASP	LSA
GM11_4259267	NA	Kreuzung	KASP	LSA
GM12_35893204	NA	Kreuzung	KASP	LSA
GM13_10578721	NA	Kreuzung	KASP	LSA
GM14_1185455	NA	Kreuzung	KASP	LSA
GM15_7914545	NA	Kreuzung	KASP	LSA
GM16_3175860	NA	Kreuzung	KASP	LSA
GM17_11134233	NA	Kreuzung	KASP	LSA
GM20_44443165	NA	Kreuzung	KASP	LSA
FAD2.1A_545	Glyma.10G278000	Oleic acid	KASP	Pham et al. 2011
FAD2.1B_Rev	Glyma.20G111000	Oleic acid	KASP	Pham et al. 2011
Pdh1	Pdh1	Shattering	KASP	Funatsuki et al. 2014

* basierend auf (Tsubokura et al. 2014), LSA = Landessaatzuchtanstalt.



Abbildung 37: Genotypisierung von verschiedenen Sojalinien für den Marker Glyma.20G085100, die 321 bp Insertion / Deletion in Glyma.20G085100 ist mit einem Unterschied im Proteingehalt assoziiert und wird mit Hilfe eines Agarosegel visualisiert. Linien mit dem kürzeren Fragment tragen das Allel welches mit einem höheren Proteingehalt assoziiert ist.

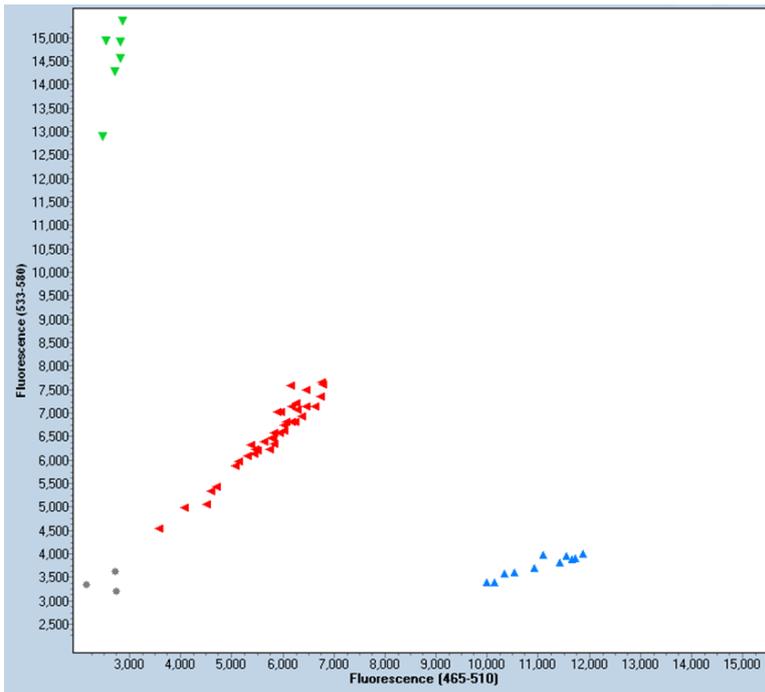


Abbildung 38: Kompetitive Allel spezifische PCR (KASP) mit dem Marker GM09_47367366 um Kreuzungen zu identifizieren. Die SNPs auf denen die Kreuzungsmarker basieren haben eine Major Allel Frequenz von etwa 0,5, wodurch sie sich gut eignen um Unterschiede zwischen Elternlinien zu sehen. Der Scatterplot zeigt die Fluoreszenz bei 465-510 nm (X-Achse) gegen die Fluoreszenz bei 533-580 nm (Y-Achse) auf. Die Elternlinien sind entweder homozygot für das Allel welches mit dem FAM

Primer (blau) oder dem HEX Primer (grün) amplifiziert wird. Erfolgreiche Kreuzungen tragen beide Allele (rot). Die grauen Punkte sind Negativkontrollen.

4.2m Regionale Genomische Selektion

Um weitere Informationen zur genomischen Selektion von Sojabohnen zu erhalten und dies mit der phänomischen Selektion zu vergleichen, wurden Daten aus 2014 herangezogen und weiter analysiert. Die ausführliche Analyse ist publiziert (Zhu et al. 2021b). Hierfür wurden 944 rekombinante Inzuchtlinien von Sojabohnen [*Glycine max* (L.) Merr.] aus acht Familien verwendet, die durch ein partiell dialleles Paarungsdesign zwischen fünf Elternlinien entstanden. Die kreuzvalidierte Vorhersagegenauigkeit für die sechs Merkmale Samenertrag, 1.000-Samen-Gewicht, Proteinertrag, Pflanzenhöhe, Proteingehalt und Ölgehalt war hoch und lag zwischen 0,79 und 0,87. Wir untersuchten die Vorhersagen zwischen den Familien, indem wir das spezielle Paarungsdesign mit verschiedenen Verwandtschaftsgraden zwischen den Familien nutzten. Im Allgemeinen nahm die Vorhersagegenauigkeit von Vollgeschwisterfamilien über Halbgeschwisterfamilien bis hin zu nicht verwandten Familien ab. Allerdings wiesen auch Halbgeschwisterfamilien und nicht verwandte Familien erhebliche Unterschiede in ihrer Vorhersagegenauigkeit für eine bestimmte Familie auf, was zumindest teilweise durch die gemeinsame Segregation quantitativer Merkmalsloci sowohl in den Trainings- als auch in den Vorhersagesätzen verursacht zu sein scheint. Die Kombination mehrerer Halbgeschwisterfamilien in zusammengesetzten Trainingsätzen führte im Allgemeinen zu einer Erhöhung der Vorhersagegenauigkeit im Vergleich zur besten Familie allein. Die Vorhersagegenauigkeit nahm mit der Größe des Trainingssets zu, aber für eine vergleichbare Vorhersagegenauigkeit waren wesentlich mehr Halbgeschwister als Vollgeschwister erforderlich. Insgesamt unterstreichen unsere Ergebnisse das Potenzial der genomischen Selektion für die Sojabohnenzüchtung und verdeutlichen in einem breiteren Kontext die Bedeutung einer gezielten Gestaltung des Trainingssets. Für die regionale genomische Selektion bedeutet dies, dass diese erst sinnvoll eingesetzt werden kann, wenn ein ausreichend umfangreicher Trainingsatz zur Verfügung steht.

Allerdings wurde in der Zwischenzeit die sogenannte phänomische Selektion entwickelt. Daher analysierten wir erneut die Daten aus 2014 und bewerteten die phänomische Vorhersage, die ähnlich wie ihr genomisches Gegenstück darauf abzielt, die Leistung von nicht getesteten Individuen vorherzusagen, jedoch unter Verwendung von

Nahinfrarotspektroskopie-Daten (NIRS). Anhand des oben dargestellten Panels von 944 rekombinanten Inzuchtlinien für Sojabohnen zeigten wir, dass die phänomischen Vorhersagefähigkeiten hoch und mit denen der genomischen Vorhersage vergleichbar sind. Auch diese Arbeiten sind publiziert (Zhu et al. 2021a). Wir haben herausgefunden, dass die Ridge-Regression als beste lineare, unvoreingenommene Vorhersage für die phänomische Vorhersage gut funktioniert und dass die Anzahl der Wellenlängen reduziert werden kann, ohne dass sich die Vorhersagefähigkeit verringert. Für die Vorhersage an verschiedenen Orten können NIRS-Daten von einem einzigen Ort verwendet werden. Allerdings sollten NIRS-Daten aus verschiedenen Umgebungen, z. B. aus verschiedenen Jahren, durch gemeinsame Genotypen in Trainings- und Vorhersagesätzen verbunden werden. Die phänomische Vorhersage scheint weniger anfällig für die Verwandtschaft zwischen den Individuen in den Trainings- und Vorhersagesätzen zu sein als die genomische Vorhersage, da im Allgemeinen Halbgeschwisterfamilien, aber auch nicht verwandte Familien hohe Vorhersagefähigkeiten erreichten. Außerdem führte die phänomische Vorhersage bei gleichem Trainingsumfang zu einer höheren Vorhersagekraft als die genomische Vorhersage. Die phänomische Vorhersage kann in verschiedenen Phasen eines Züchtungsprogramms angewandt werden, und insgesamt unterstreichen unsere Ergebnisse das Potenzial dieses Ansatzes zur Steigerung des genetischen Gewinns in der Pflanzenzucht. Für einen großen Teil der in Tabelle 2 dargestellten Leistungsprüfungen wurden die NIRS-Spektren ermittelt. Diese dienen als Grundlage (Trainingsatz) für eine Selektion des jüngeren Zuchtmaterials. Von diesem Zuchtmaterial wurden von der Ernte 2021 NIRS-Spektren von 3500 F4-Pflanzen erhoben. Anhand dieser Spektren wird der Zuchtwert geschätzt. Dies kann für verschiedene Umwelten geschehen. Die Zuchtwerte fließen in die weitere Selektion des Materials ein. Damit wird ein erster Anfang für eine regionale phänomische Selektion gemacht.

4.3 IPM

4.3a Literaturarbeit, Primerdesign, Vortests

Die Grundlagen der quantitativen und schnellen Erfassung von Sojapathogenen durch PCR wurden erarbeitet.

4.3b Etablierung des quantitativen Nachweisverfahrens für verschiedene Pathogene

Eine Annealingtemperatur von 61,5°C Zur Etablierung des qRt-PCR Nachweisverfahrens für *Sclerotinia* spp. wurde als optimal determiniert. Die Primereffizienz wurde erfasst und die Spezifität der Primer geprüft (Abbildung 39).

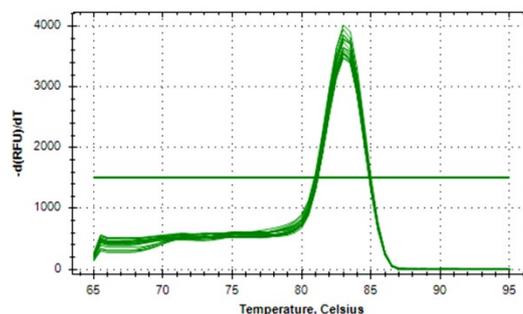


Abbildung 39: Schmelzkurve des Calmodulin qRT-PCR-Produktes von *S. minor*.

Die optimale Annealingtemperatur der Primerpaare für *R. solani* und *P. ultimum* wurde durch eine Gradienten-qPCR ermittelt und lag zwischen 64,9°C und 66,9°C. Die qPCR Produkte wurden über die Erstellung einer Schmelzkurve auf ihre Reinheit überprüft.

Die Überprüfung der Spezifität der Primer für *R. solani* ergab, dass diese nur ein Produkt mit der DNA des Isolates aus der Anastomose Gruppe 3 (AG-3) amplifizieren. Bei Verwendung der DNA eines Isolates aus der AG-5 wurde kein Produkt amplifiziert. Der Nachweis für *R. solani* ist somit bisher nur für die AG-3 realisiert.

Bei Verwendung der DNA von *P. ultimum* wurde ein Produkt amplifiziert. Die optimale Annealingtemperatur lag zwischen 58,5 °C und 62,5 °C. Auch hier gab es eine klare Schmelztemperatur von 79 °C (Abbildung 40b). Durch die Gelelektrophorese wurden keine weiteren Produkte sichtbar gemacht. Das Primerpaar qPyu92f/qPyu166r wies nur *P. ultimum* nach und kann somit als spezifisch angesehen werden.

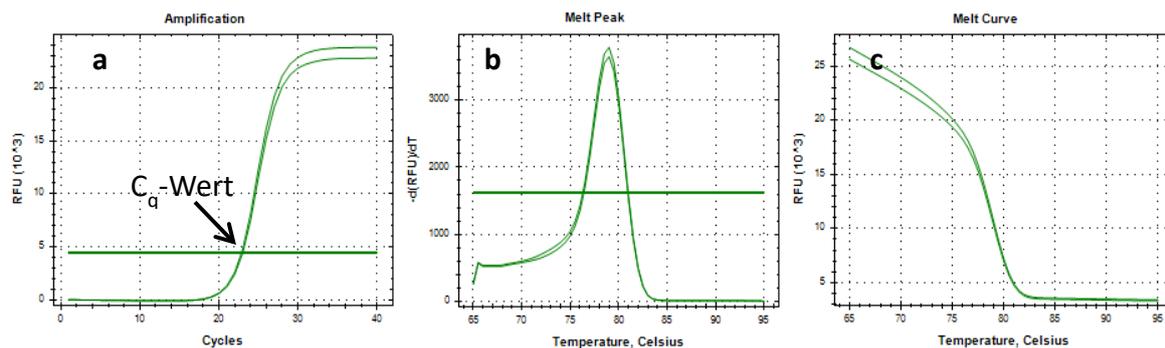


Abbildung 40: Relative Fluoreszenz mit Darstellung des C_q-Wertes (a), der Schmelztemperatur (b) und der Schmelzkurve (c) für *P. ultimum* bei Verwendung der Annealingtemperatur von 62,5°C.

Des Weiteren wurde eine Annealingtemperatur von 61,5°C der Primärpaare für Pathovare von *P. syringae* für weitere Versuche etabliert. Die Primereffizienz (98 %) wurde erfasst und die Spezifität getestet. Zusätzlich zu den C_q-Wert Auswertungen wurden die qRT-PCR Produkte per Gelelektrophorese auf Primer-Artefakte überprüft. Demnach können diverse Pathovare von *P. syringae* nachgewiesen werden.

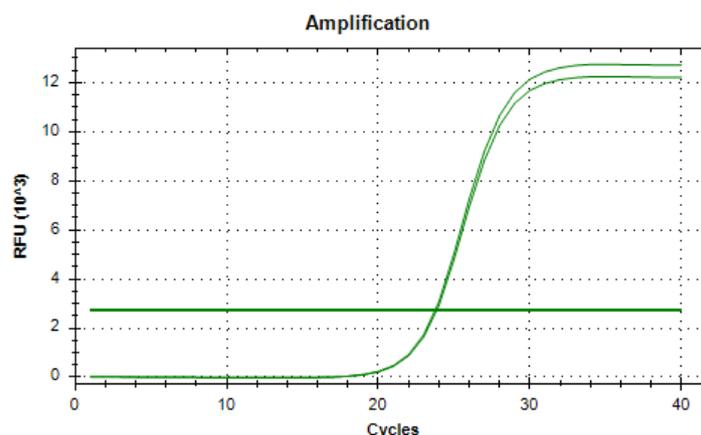


Abbildung 41: Amplifikation der Cytochrome O Ubiquinol Oxidase Untereinheit II von *P. syringae*.

Dieses AP kann als abgeschlossen angesehen werden, da der Einzelnachweis zumindest für vier Hauptpathogene (*S. sclerotiorum*, *P. ultimum*, *R. solani* und *P. syringae* pv. *glycinea*) etabliert wurde.

4.3c,k Etablierung des quantitativen Nachweisverfahrens für verschiedene Probenotypen

Das Nachweisverfahren für *Pseudomonas syringae* pv. *glycinea* wurde weiterentwickelt und für *in planta* Proben angepasst. Die aus befallenen Blätter wachsenden bakteriellen Kolonien konnten anhand der physiologischen und morphologischen Eigenschaften als *Pseudomonas* sp. eingeordnet werden. Bei der Inokulation gesunder Sojabohnenpflanzen mit zwei Bakterienstämmen, zeigten sich auf den inokulierten Blättern ähnliche Symptome. Zur Überprüfung, ob es sich bei den aus den Blättern nachgewiesenen Organismen um die selben handelt, mit denen inokuliert wurde, wurde eine qRT-PCR mit dem für *Pseudomonas syringae* spezifischen Primerpaar und eine Agarose-Gelelektrophorese mit anschließender optischer Überprüfung unter UV-Licht durchgeführt. Hierbei wurde für alle getesteten Proben jeweils eine Bande bei 176 bp sichtbar (Abbildung 42). Dies entspricht der erwarteten Größe der amplifizierten Gensequenz der Cytochrom O Ubiquinol Oxidase Untereinheit II (Xu und Tambong 2011). Zur weiteren Überprüfung der Spezifität der Primer wurden sechs *Xanthomonas axonopodis* Stämme getestet. Auch hier fand eine Amplifikation statt. Bei der optischen Überprüfung der qRT-PCR Produkte wurde jedoch ersichtlich, dass für jede Probe zwei Produkte amplifiziert wurden. Eines bei 176 bp und eines bei ca. 500 bp (Abbildung 43). Dementsprechend liegt es nahe, für weitere Versuche andere Primer zu erstellen.

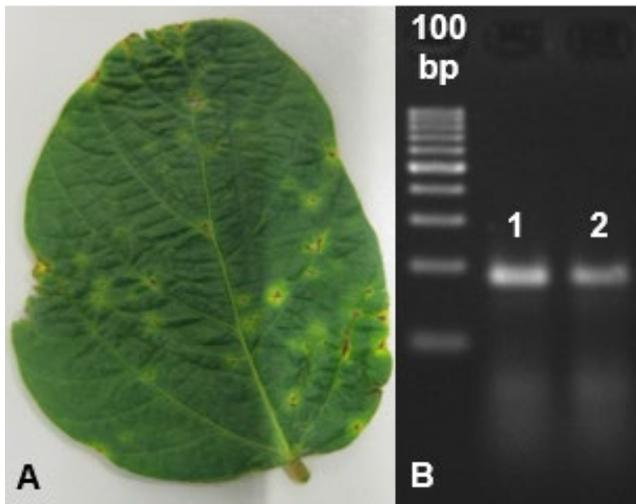


Abbildung 42 A: Fröhe Symptome von *P. syringae* pv. *glycinea*. B: TAE-Gel (2 %) von qRT-PCR Produkte von zwei Bakterienstämmen (1, 2) die aus zwei infizierten Blättern isoliert wurden.

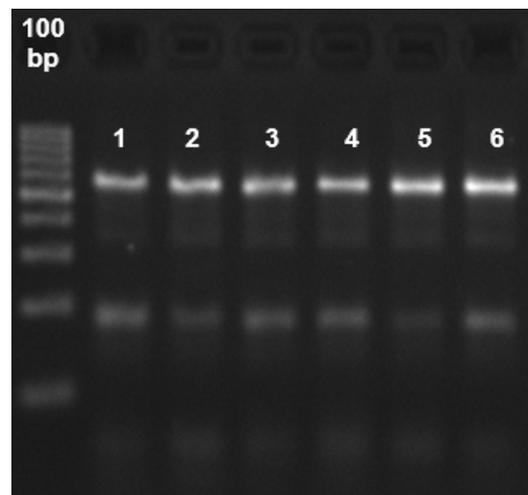


Abbildung 43: TAE-Gel (2 %) der qRT-PCR Produkte von sechs verschiedenen *Xanthomonas axonopodis* Stämmen.

Zur Anpassung der Pathogenitäts-/Virulenztest der pilzlichen Pathogene wurden Blätter von Sojabohnenpflanzen mit zwei *S. sclerotiorum* Stämmen (DSM, IZS), einem *S. minor* Stamm, einem *Stromatinia gladioli* Stamm und verschiedene *Pythium*-Arten inokuliert. Dabei zeigte für *S. sclerotiorum* der IZS Stamm eine höhere Virulenz als der DSM Stamm. Der *S. minor* und der *S. gladioli* Stamm führten zu keiner sichtbaren Infektion der Blätter. Desweiteren konnten Infektionen der Blätter durch die verschiedenen *Pythium*-Arten beobachtet werden.

Unter Gewächshausbedingungen konnten die Pflanzen mit *S. sclerotiorum* nach künstlicher Inokulation mit Hilfe der Zahnstocher-Methode erfolgreich befallen werden. Symptomatische Unterschiede waren erst nach 6 dpi auf den Stängel und Blätter sichtbar (Abbildung 4). Die qRT-PCR-Probenanalyse steht noch aus.

Bei der Umstellung auf die TaqMan Technologie wurden für die Real-time PCR eine Denaturierungstemperatur von 95°C und eine Annealingtemperatur von 60°C gewählt. Der Nachweis ist spezifisch für *P. syringae* pv. *tomato*, Psg und *P. savastanoi*. Andere getestete *Pseudomonas*-Arten wie *P. phaseolicola* und *P. fluorescens*, sowie *Xanthomonas campestris* pv. *phaseoli* wurden nicht nachgewiesen. Der Nachweis wurde an Blattproben angepasst.

Auch für für Ss wurde für die TaqMan Real-time PCR eine Denaturierungstemperatur von 95°C und einer Annealingtemperatur von 60°C durchgeführt. Der Nachweis wurde durch Tests mit den ebenfalls relevanten *Diaporthe* spp, *Colletotrichum* spp. und Rs überprüft und ergab hier keinen Nachweis, was auf eine gute Spezifität hinweist.

Die Primer und Sonden für den gattungsspezifischen TaqMan Nachweis für Fs, wurden außerdem mit DNA von Rs, *Diaporthe* spp. und *Sclerotinia* spp. getestet, wobei sich ein negatives Ergebnis ergab. Die Spezifität der Primer für die Gattung *Fusarium* ist somit weitgehend nachgewiesen. In weiteren Schritten wurde eine Primer- und Sondeneffizienz von 95,1% ermittelt (Abbildung 44).

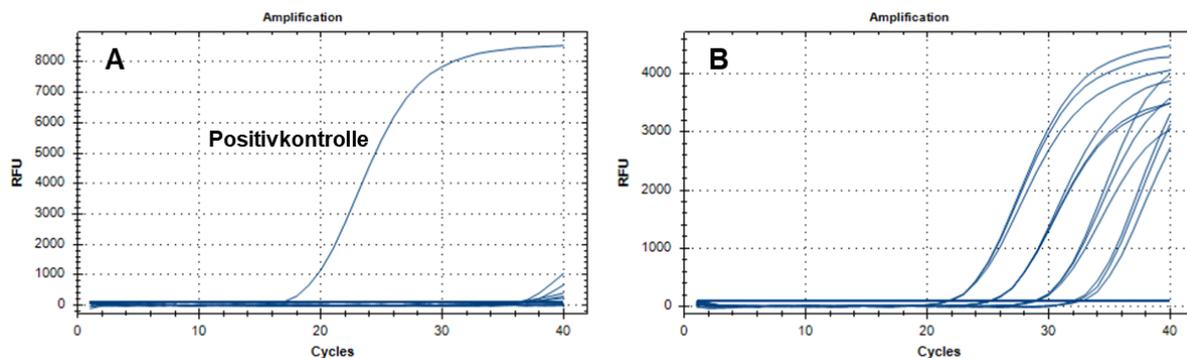


Abbildung 44: **A** Testung des *Fusarium* sp. Nachweises auf die Spezifität gegen andere pilzliche Pathogene. Als Positivkontrolle diene DNA von *Fusarium* sp.. **B** Verdünnungsreihe zur Effizienzbestimmung der Primer und Sonden von *Fusarium* sp.. Die Verdünnungsreihe wurde in 1:10 Verdünnungsschritten in Triplikaten getestet mit drei unterschiedlichen Isolaten wiederholt.

4.3d,I Etablierung einer quantitativen Methodik zur Erfassung von Nematoden

Einheimische Nematodenarten, die in anderen Kulturen zu Schäden führen können, kommen auch im Sojaanbau vor. Diese Art der Schädigung spielt jedoch im deutschen Sojaanbau kaum eine Rolle, da die hier erreichten Populationsdichten zu gering sind, um schwerwiegende Wuchsdepressionen an Sojabohnenpflanzen hervorrufen zu können (Elhady et al. 2018). Von wesentlich größerer Bedeutung sind die durch Nematoden hervorgerufenen Sekundärschäden, wozu das Schaffen von Eintrittspforten für Krankheitserreger wie Pilze und Bakterien zählt.

Während dieses APs wurden Bodenproben von sieben landwirtschaftlichen Betrieben genommen. Beprobt wurde für jeden Betrieb jeweils ein Schlag, auf dem im Jahr 2015 Soja angebaut wurde und ein Schlag, auf dem 2016 Soja angebaut werden sollte. In allen Schlägen waren in unterschiedlicher Anzahl Nematoden der Gattung *Pratylenchus* (zw. 20 und 688 pro 100 ml Boden). In zwölf Schlägen wurden *Paratylenchus* spp. (zw. 4 und 160 pro 100 ml Boden) nachgewiesen. Diese Anzahl befindet sich im Normbereich (Tabelle 12).

Auffallend bei den Untersuchungen des JKI war die insgesamt hohe Vermehrungsrate bei *P. penetrans* an der Sojasorte Merlin in 1 m² großen Mikroplots. Die durchschnittliche Vermehrungsrate betrug 66,6 % was den sehr guten Wirtsstatus von Soja für diese Nematodenart bestätigt. Die maximale Nematodendichte betrug 2.442 *P. penetrans* / 100 ml

Boden, was angesichts der Schadschwelle in anderen Kulturen sehr hoch ist (Elhady et al. 2018).

Tabelle 12: Anzahl nachgewiesener Nematoden im Boden. Dargestellt wird die Anzahl Tiere / 100 ml Boden für *Meloidogyne*, *Paratylenchus* spp. und *Pratylenchus* spp. (*P. penetrans*, *P. neglectus*, *P. thornei*, *P. crenatus*) beispielhaft für zwei Betriebe.

Gattung/Probe	Tiere / 100 ml Boden			
	Heidfeldhof	Heidfeldhof	Buck Riedlingen	Buck Riedlingen
	Soja 2015	Soja 2016	Soja 2015	Soja 2016
<i>Meloidogyne</i>	-	-	-	-
<i>Paratylenchus</i> spp.	60	8	160	8
<i>Pratylenchus</i> spp.	396	104	40	96

4.3e Etablierung des Multiplex-Systems auf verschiedenen Proben

Da die bisher erbrachten Einzelnachweise alle mit spezifischen Primern kombiniert mit SYBR Green basierten qPCR erbracht wurden, ist es hier zunächst notwendig, für die Amplikons der verschiedenen Nachweise spezifische Sonden zu erstellen. Dies sollte recht problemlos funktionieren, da die Spezifität für die verschiedenen Arten bereits durch die Primer gewährleistet wird und somit die Sonden nur zwischen den im Multiplex kombinierten Amplikons unterscheiden müssen. Diese wiederum sind stark unterschiedlich, da ja verschiedene Gene in den Nachweisen verwendet werden.

4.3f,m Testung verschiedener Proben

Die bisherigen Ergebnisse hinsichtlich der verschiedenen Probenotypen sind sehr vielversprechend. Idealerweise müssen keine separaten Multiplexnachweise für die verschiedenen Probenotypen etabliert werden. Es wäre aber noch zu entscheiden, ob dennoch verschiedene Multiplexverfahren angestrebt werden sollten, da verschiedene Erreger mit verschiedenen Proben verstärkt assoziiert zu sein scheinen, was gegebenenfalls den Nachweis in einer Probenart überflüssig machen würde. Als wichtigster Probenotyp erscheinen *a priori* Samenproben, da sich die qRT-PCR insbesondere für die Effizienzsteigerung der Saatgutkontrolle eignet.

Der Duplexnachweis von für Ss und Fs wurde an Samenproben getestet, an denen bereits Fs im Einzelnachweis festgestellt werden konnte. Im Duplexnachweis wurde dabei kein Ss nachgewiesen. Dies kann einer fehlenden Infektion zugrunde liegen und sollte mit künstlich inokuliertem Saatgut bei Bedarf wiederholt werden.

4.3h Primerdesign, Vortests

Nach eingehender Literaturrecherche wurden die in den Publikationen Kim und Knudsen (2008) sowie Xu und Tambong (2011) vorgestellten Erkenntnisse als Basis in dieser Arbeit gewählt und betrachtet. Die daraus verwendeten Primer und Sonden werden in Tabelle 13 gelistet.

Tabelle 13: Verwendete Primer und Sonden für die Nachweise von Ss und Psg.

Bezeichnung	Sequenz
qSsCL2fw	5'-CCCAGTTCGACTCTCCTCTTTTAT-3'
qSsCL2rv	5'-AACTCAGACTCGGAAGGGTTTTG-3'
TMSCL2P	5'-Cy5-AGACATCTTGACCGACACCGCCCC-IBFQ-3
qPscyo2fw	5'-TCGAGCAGCGGAACCTGATC-3'
qPscyo2rv	5'-TGGGTACGCCCCAGACTGCGA-3'
Psy_cyoII-pb	5'-FAM-GCCAAGTACACGCCGGACTGGTC-BHQ1-3'

Zur Anpassung der Pathogenitäts-/Virulenztest der pilzlichen Pathogene wurden Blätter von Sojabohnenpflanzen mit zwei Ss Stämmen (DSM, IZS), einem S. minor Stamm, einem *Stromatinia gladioli* Stamm und verschiedene *Pythium*-Arten inokuliert. Dabei zeigte für Ss IZS Stamm eine höhere Virulenz als der DSM Stamm. Der S. minor und der S. gladioli Stamm führten zu keiner sichtbaren Infektion der Blätter. Desweiteren konnten Infektionen der Blätter durch die verschiedenen *Pythium*-Arten beobachtet werden.

Um einen gattungsspezifischen TaqMan Nachweis für Fs, den Verursacher der *Fusarium* Wurzelfäule zu etablieren, wurde DNA von zehn *Fusarium*-Isolaten extrahiert. Für die Gene β -Tubulin, ITS und Translations-Elongationsfaktor 1 α wurden PCRs durchgeführt und die gereinigten PCR-Produkte sequenziert. Die sequenzierten PCR-Produkte wurden per NCBI BLASTn abgeglichen (Tabelle 14).

Tabelle 14: Identifizierung der *Fusarium* Spezies laut NCBI BLASTn.

Nr.	Spezies laut NCBI BLASTn	Herkunft
1	<i>Fusarium avenaceum</i>	Auflauf Stängel
2	<i>Fusarium solani</i>	Auflauf Wurzel
3	<i>Fusarium avenaceum</i>	Stammsammlung
4	<i>F. oxysporum</i> f. sp. <i>rapae</i> or <i>raphani</i>	Auflauf Wurzel
5	<i>Fusarium graminearum</i>	Stammsammlung
6	<i>Fusarium equiseti</i>	Auflauf Stängel
7	<i>Fusarium solani</i>	Auflauf Wurzel
8	<i>Fusarium fabeceaum/pisi/glycines</i>	Stammsammlung
9	<i>Fusarium tricinctum</i>	Samen
10	<i>Fusarium tricinctum</i>	Samen

Die Gensequenzabschnitte wurden zusammen mit Sequenzen von *Diaporthe* spp. in der Software BioEdit align und so auf Bereiche untersucht, die innerhalb der Gattung *Fusarium* aber nicht zu anderen Gattungen konserviert sind. In diesen Bereichen wurden zwei Primer (FusTub_GsP1f, FusTub_GsP1r) und zwei Sonden (FusTub_Gs1aP, FusTub_Gs1bP) design, die im Test alle zehn Isolate nachwiesen (Tabelle 15).

Tabelle 15: Primer und Sonden für den Nachweis von Fs.

Bezeichnung	Sequenz
FusTub_GsP1f	5'-CAT CTC TGG CGA GCA [C/T]GG-3'
FusTub_GsP1r	5'-CA[A/T] ACC TCG TTG AAG TA[A/G] AC-3'
FusTub_Gs1aP	5'-6-FAM- A CAG CAA TGG TGT C TA CAA CGG-BMN-Q535-3'
FusTub_Gs1bP	5'-6-FAM-GA CAG CAA TGG TGT T TA CAA CGG-BMN-Q535-3'

4.3k Etablierung des quantitativen Nachweisverfahrens für verschiedene Probentypen

Das Nachweisverfahren für Psg wurde weiterentwickelt und für *in planta* Proben angepasst. Die aus per Real-Time-PCR als befallen bestimmten Blättern wachsenden bakteriellen Kolonien konnten anhand der physiologischen und morphologischen Eigenschaften als *Pseudomonas* sp. eingeordnet werden. Somit ließ sich klar und reproduzierbar das Ergebnis der Real-Time-PCR als korrekt bestätigen.

Für die Anpassung an Stängelproben wurde die DNA der mit Ss inokulierten Sojapflanzen extrahiert und per TaqMan-Real-time PCR analysiert (Abbildung 45). Dabei konnte Ss an 3, 6 und 9 Tagen nach der Inokulation nachgewiesen werden.

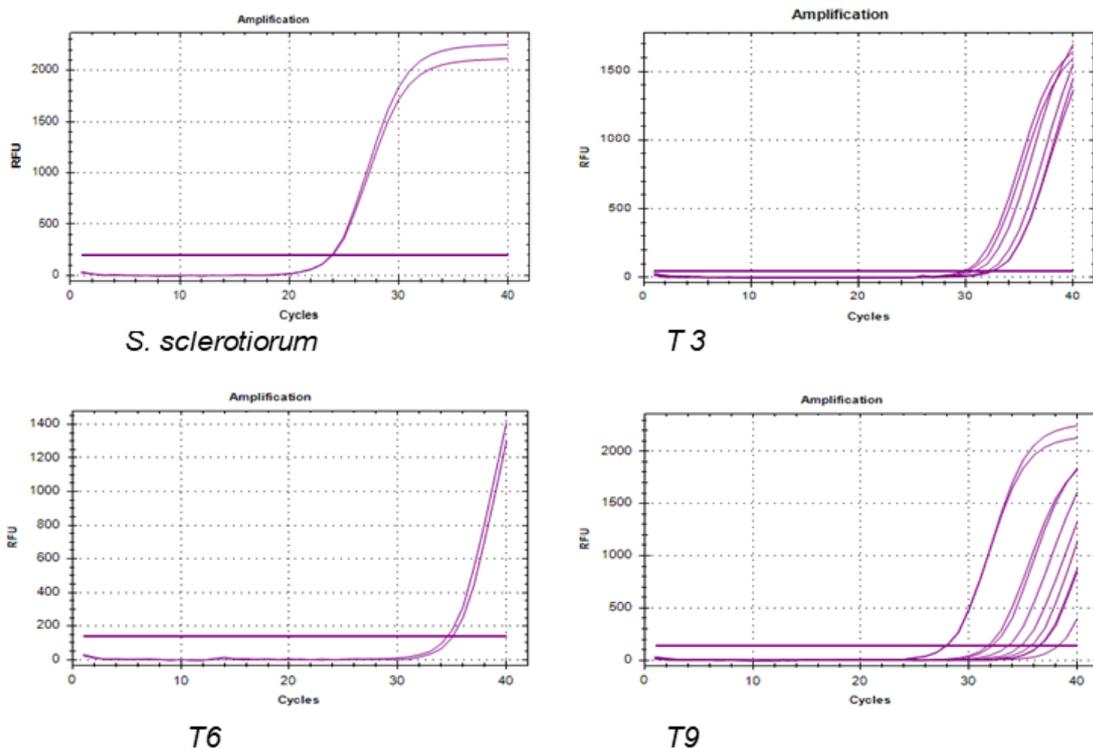


Abbildung 45: Nachweis von *S. sclerotiorum* an inokulierten Soja-Pflanzen. Positivkontrolle *S. sclerotiorum*, Proben T3, T6 und T9 entsprechen 3, 6 und 9 dpi.

Beim Nachweis von Fs an Saatgutproben konnte in drei der sieben getesteten Kultivare Fs nachgewiesen werden (Abbildung 46). Als Positivkontrolle diente hierbei DNA von Fs, als Negativkontrolle extrahierte DNA eines gesunden Samens und Wasser. Es konnte somit gezeigt werden, dass die von uns gewählte Vorgehensweise, DNA-Präparation aus der Probe zu präparieren und hierauf dann die Real-Time-PCR durchzuführen prinzipiell auf die verschiedenen Probentypen übertragbar ist.

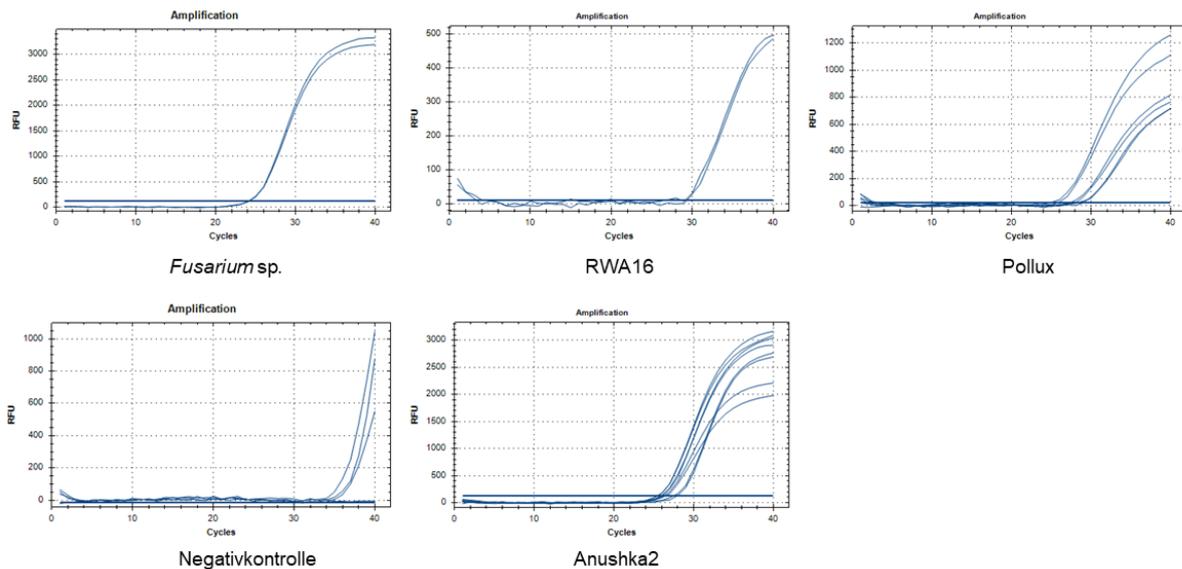


Abbildung 46: Nachweis von *Fusarium sp.* an verschiedenen Saatgutproben (RWA16, Pollux, Anushka2). *Fusarium sp.* diente hierbei als Positivkontrolle, extrahierte DNA eines gesunden Samens und Wasser als Negativkontrolle.

4.3m Etablierung des Multiplex-Systems auf verschiedenen Proben

Der Duplexnachweis wurde für Ss und Fs etabliert (Abbildung 47). Die Duplexnachweise wurden mit 1:10 Verdünnungen und unverdünnten Reinkulturen getestet.

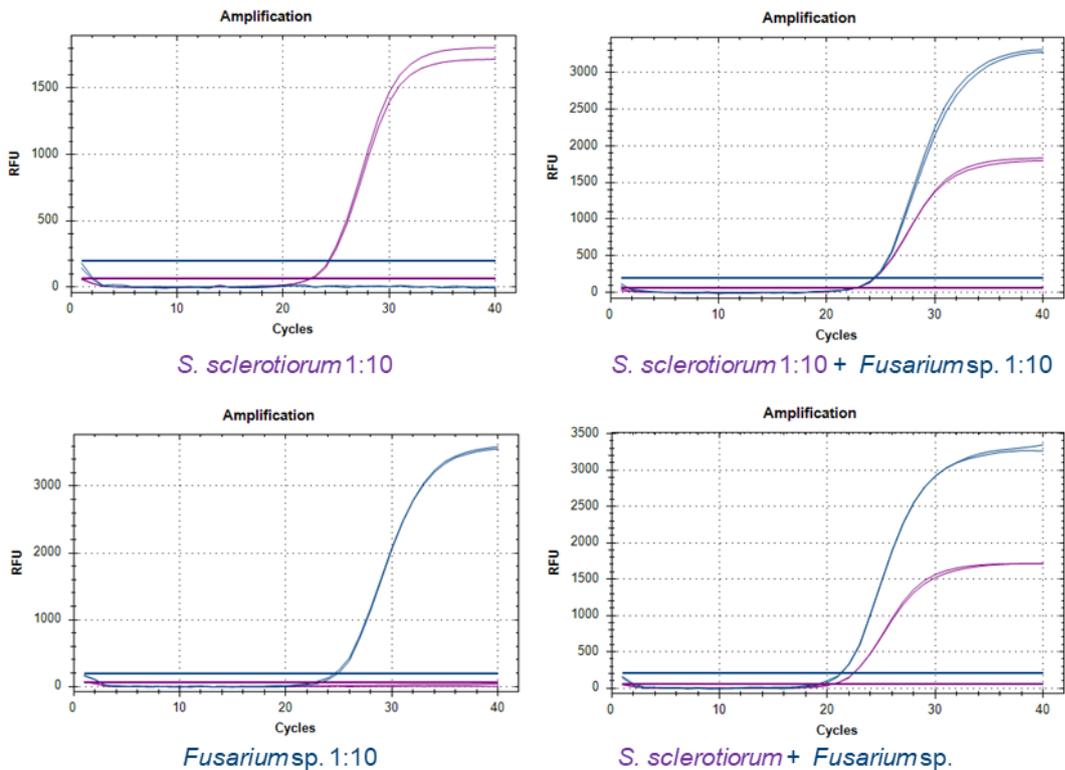


Abbildung 47: Multiplex-Nachweis von *S. sclerotiorum* und *Fusarium spp.*

Es wurde außerdem ein Duplexnachweis für *Ss* und *Pseudomonas* spp. Etabliert (Abbildung 48). Die Duplexnachweise wurden mit 1:10 Verdünnungen und unverdünnten Reinkulturen getestet.

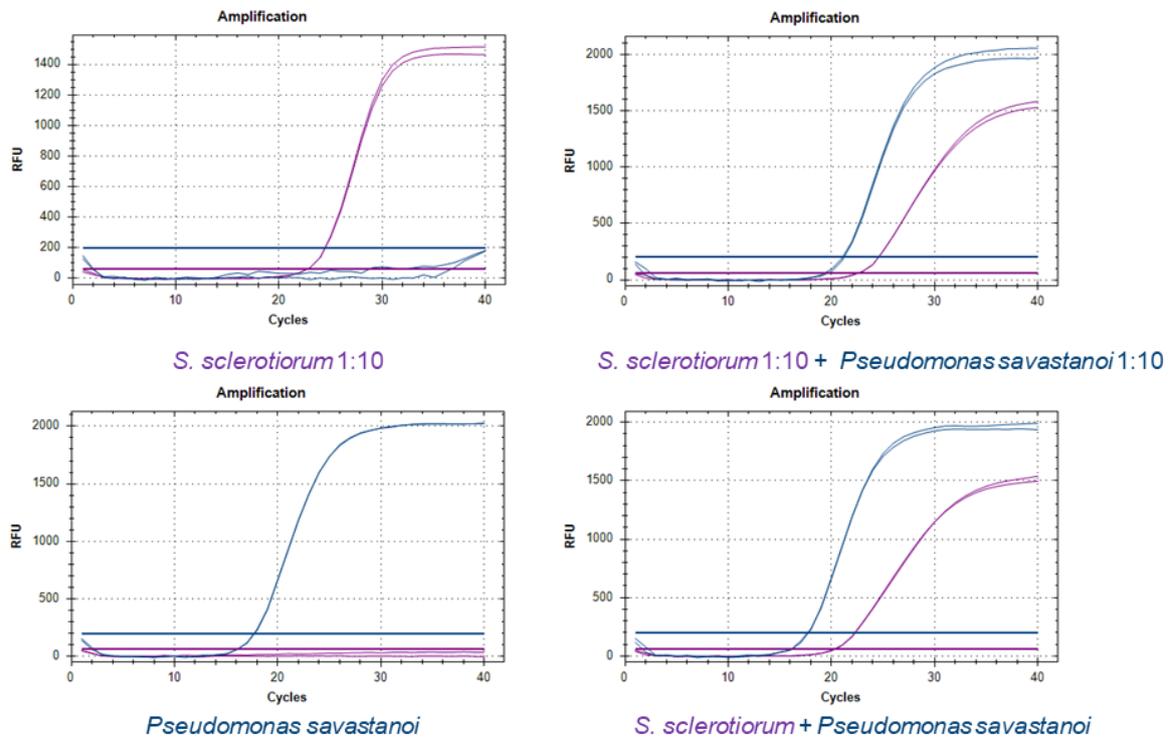


Abbildung 48: Multiplex-Nachweis von *S. sclerotiorum* und *Pseudomonas* spp.

Beide getesteten Duplex-Zusammensetzungen waren erfolgreich. Damit ist auch geklärt, dass sich die drei Arten prinzipiell auch im Triplex nachweisen lassen. Es muss dann jedoch entweder für Fs oder Ps ein anderes Fluorophor für die Sonde gewählt werden.

4.4 JKI-RS

Gefäßversuche unter kontrollierten Bedingungen zur Phänotypisierung des Hülsenansatzes unter Kühlestress (JKI)

Im Berichtszeitraum wurde die Phänotypisierung von 122 Linien aus einer Kreuzungsnachkommenschaft Merlin (kühletolerant) x Sigalia (mittelgradig kühesensitiv), der Eltern sowie eines kühesensitiven Standards bezüglich ihrer Kühletoleranz zum Zeitpunkt des frühen Hülsenansatzes unter kontrollierten Bedingungen durchgeführt. Diese Untersuchungen basierten auf einem Test, der im Rahmen des Projektes 'Ausweitung des Sojaanbaus in Deutschland durch züchterische Anpassung und pflanzenbauliche Optimierung' (2811NA007) entwickelt wurde und dort auch beschrieben ist. Im Ergebnis zeigen sowohl der Hülsenansatz nach Stress wie auch der Chlorophyllgehalt (indirekt als SPAD Werte erfasst) geringere Werte infolge des Kühlestresses (Abbildung 49). Darüber hinaus ist sowohl im Hülsenansatz als auch bei den SPAD Werten neben signifikanten Effekten von Genotyp und Variante eine signifikante Genotyp x Variante Wechselwirkung nachweisbar. Diese steht für eine Variabilität in der Reaktion der Genotypen auf den Stressfaktor und ist Voraussetzung für eine Selektion auf eine verbesserte Toleranz. Die grundlegenden statistischen Parameter sind in Tabelle 16 dargestellt.

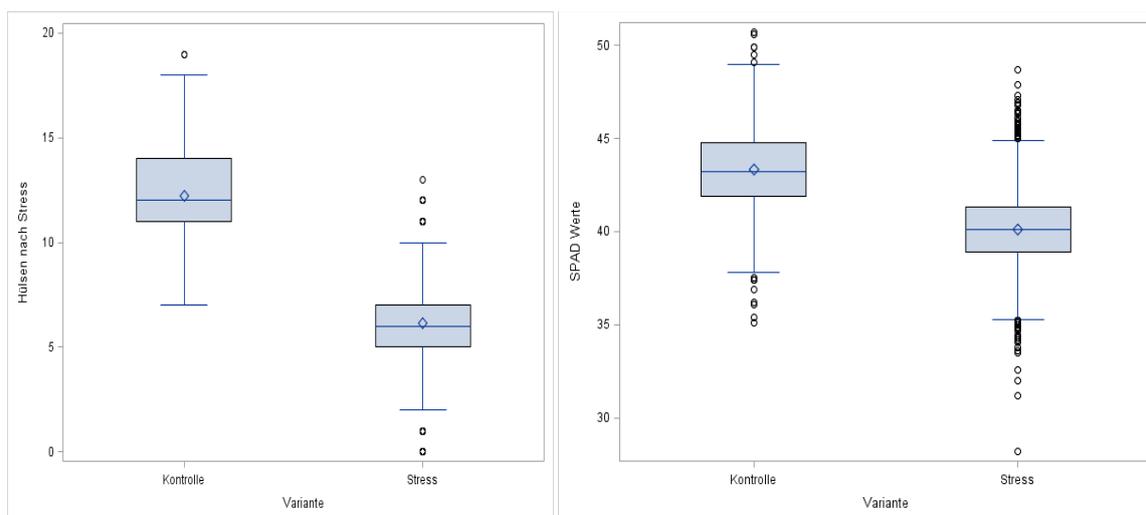


Abbildung 49: Hülsenansatz nach einer 3 wöchigen Kühlestressphase sowie SPAD Werte für die Kontroll- und Stressvariante basierend auf 5 unabhängigen Wiederholungen.

Tabelle 16: Statistische Analyse der untersuchten Merkmale.

Merkmal		Gesamt	Kontrolle	Kühlestress
Hülsenansatz Anzahl	Genotyp	***	***	***
	Variante	***	12,2	6,1
	G x V	***	(15,6 – 8,4) LSD 3,1	(9,0 – 1,4) LSD 2,1
SPAD SPAD Einheiten	Genotyp	***	***	***
	Variante	***	43,3	40,1
	G x V	***	(47,1 - 39,2) LSD 3,1	(43,1 – 36,0) LSD 2,15

(*** signifikant für $\alpha \leq 0,001$; LSD kleinste signifikante Differenz für $\alpha \leq 0,05$)

Die Heritabilität der Merkmale als Maß für die genetische Stabilität fiel vergleichsweise hoch aus (Tabelle 17). Das ist indirekt ein Zeichen dafür, dass das Testsystem reproduzierbare Ergebnisse ermöglicht hat.

Tabelle 17: Heritabilitäten der untersuchten Merkmale.

Heritabilität (h^2)	Kontrolle	Kühlestress
Hülsenansatz	0,74	0,75
SPAD	0,83	0,76

Daraufhin wurde untersucht, in welchem Umfang sich die Heritabilitäten des Merkmals unter Stress- und Kontrollbedingungen verändern, wenn nur eine Kombination von 2,3 oder 4 der 5 verfügbaren Superblöcke (komplette Blöcke) in die Analyse mit einfließen (Abbildung 50). Die Daten lieferten sehr robuste Werte, sodass selbst bei nur 2 Wiederholungen noch eine Wiederholbarkeit von 0,6 errechnet wurde. Bei 4 Wiederholungen lag die Wiederholbarkeit mit 0,7 nur leicht unter der errechneten Wiederholbarkeit von 0,75 bzw. 0,74. Die Reproduzierbarkeit war bei 3 oder 4 Superblöcken unter Kontrollbedingungen nicht signifikant unterschiedlich zu

der Heritabilität wenn alle 5 Superblöcke verwendet wurden. Lediglich bei den Kombinationen aus 2 Superblöcken war das Mittel der Heritabilitäten signifikant niedriger zur Ausgangslage, allerdings mit 0,6 noch immer relativ hoch (Abbildung 50). Etwas anders sah es unter Stressbedingungen aus, dort sind allerdings die Unterschiede der Mittelwerte bei 3 und 4 Replikaten bei einem Signifikanzniveau von $\alpha = 0,05$ nur marginal signifikant (p -Wert des Mittelwertvergleichs = 0,047) und davon abgesehen sind die Wiederholbarkeiten bei nur 3 Wiederholungen mit im Mittel 0,65 noch relativ hoch.

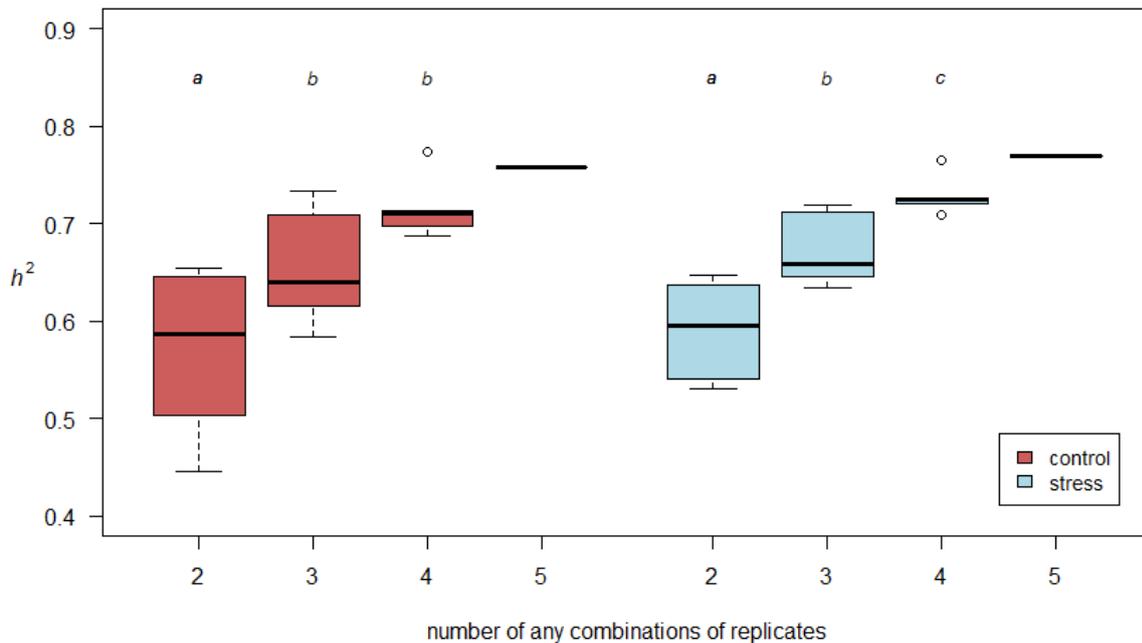


Abbildung 50: Veränderung der Heritabilität des Hülsenansatzes unter Kontroll- und Stressbedingungen, wenn nur Kombinationen aus 2, 3 oder 4 statt 5 Wiederholungen verwendet wurden. Die Boxplots für 2, 3 und 4 Wiederholungen setzen sich aus den errechneten Heritabilitäten aller möglichen Kombinationen von diesen Wiederholungen zusammen. Bei 5 Replikaten gab es nur eine Kombination, weshalb dort nur ein Wert vorkommt. Gleiche Buchstaben über den Boxen zeigen keinen signifikanten Unterschied der Mittelwerte (Signifikanzniveau $\alpha=0,05$).

Ein Blick auf die mit untersuchten Eltern Merlin und Sigalia sowie den sensitiven Standard Malaga zeigt, dass diese Sorten beim Hülsenansatz wie erwartet reagieren. Bei den SPAD Werten ist zwar bei allen drei Sorten ein Abfall in Folge des Kühlestresses erkennbar, allerdings keine signifikanten Sortenunterschiede in der Reaktion auf den Stress (Abbildung 51).

Das bestätigt sich auch für die Linien. Während die SPAD Werte über beide Varianten mit dem Hülsenansatz noch mit $r = 0,63$ korreliert sind, was aus dem Abfall der Werte beider Merkmale unter Stress resultiert, zeigt der Abfall der SPAD- Werte unter Stress zur Differenz im Hülsenansatz mit $r = 0,03$ keine Korrelation.

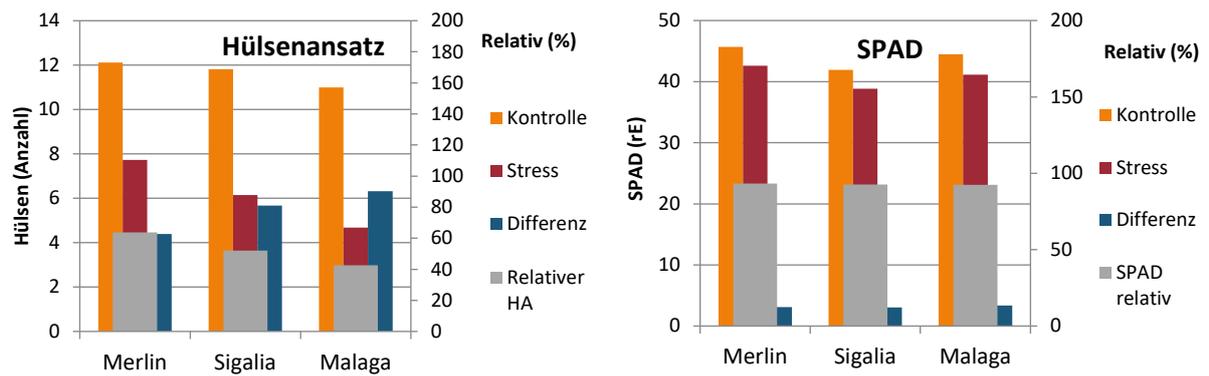


Abbildung 51: Hülsenansatz und SPAD Werte für die Eltern Merlin und Sigalia sowie die sensitive Sorte Malaga im Vergleich Kontrolle-Kühlestress (LSDs siehe Tabelle 16).

QTL Analyse zum Hülsenansatz der biparentalen Population Merlin x Sigalia unter Kontroll- und Kühlestressbedingungen (UHOH)

Mittels eines ‚Genotype-by-sequencing‘ (GBS) -Ansatzes und durch die anschließende QTL-Analyse der SNP-Marker auf Einzelpflanzenbasis wurden im Umfang dieses Projektes für die ‚Merlin x Sigalia‘-Population 3 QTL für den Hülsenansatz unter Kontroll- (Chromosom 4,7,13), sowie 2 QTL unter Stressbedingungen (Chromosom 7,13) gefunden und durch Kreuzvalidierungen bestätigt (Abbildung 52A und D). Betrachtet man die Alleleffekte der SNPs bei denen die LOD-Werte spitzen, fällt auf, dass Allele von beiden Eltern einen positiven Effekt auf den Hülsenansatz haben und nicht nur Allele vom Merlin-Elter, der in den Vortests als weniger Kühlesensitiv eingestuft wurde (Abbildung 52B und C). Dies untermalt, zusätzlich zu dem Fund, dass es nicht den einen großen QTL gibt, den quantitativen Charakter des Merkmals. Züchterisch gilt es also weitere QTL (wenn möglich populationsübergreifend) zu finden und diese positiven Allele in Zielpflanzen anzureichern.

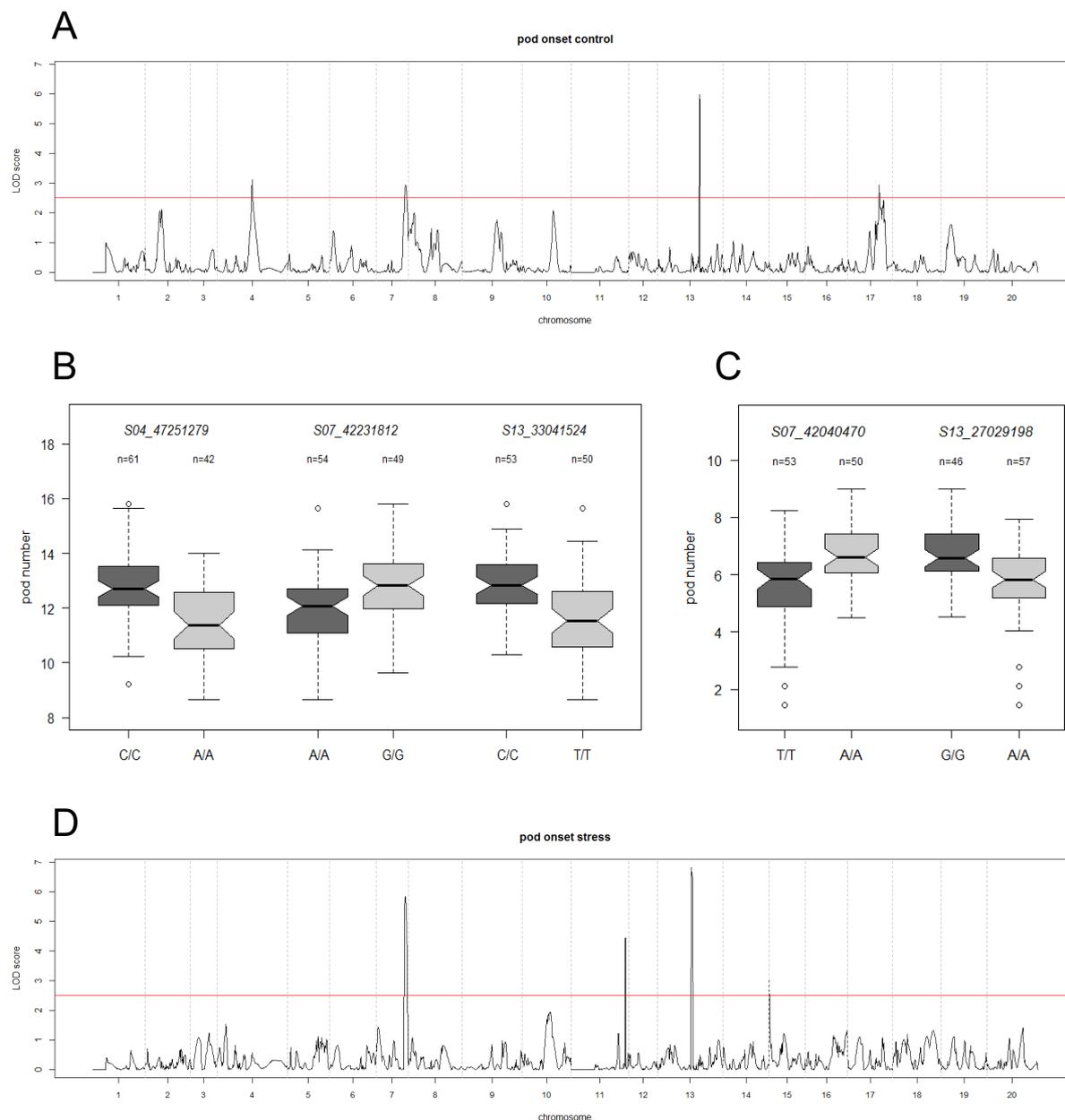


Abbildung 52: QTL-Mapping und Effekt der signifikanten Marker, A und D) QTL-Kartierung des Hülsenansatzes unter Kontrollbedingungen (A) und Stressbedingungen (D), die rote Linie zeigt den LOD-Grenzwert von 2.5; B zu A zugehörige Effekte der ‚peak‘-Marker auf Chromosom 4, 7 und 13, Peak auf Chromosom 17 konnte durch Kreuzvalidierung nicht bestätigt werden, C) zu D zugehörige Effekte der ‚peak‘-Marker auf Chromosom 7 und 13, Peak auf Chromosom 11 und 15 konnte durch Kreuzvalidierung nicht bestätigt werden, dunkelgraue Färbung der Boxen bedeutet, dass das zugehörige Allel dieses Markers vom Sigalia-Elter stammt, hellgrau bedeutet Abstammung des Allels vom Merlin-Elter.

Feldversuch – Hülsenansatz und Prüfanbau (JKI)

In Ergänzung der Ergebnisse zur Phänotypisierung der Kreuzungsnachkommenschaft unter kontrollierten Bedingungen für die QTL Kartierung und - bei Vorhandensein entsprechender Stressbedingungen im Feld - Verifizierung der Ergebnisse zur Kühletoleranz, wurden die Eltern, die 122 Linien sowie die kühlesensitive Sorte Malaga 2016 in Groß Lüsewitz im Feld angebaut. Aus Gründen der begrenzten Saatgutverfügbarkeit erfolgte dies in Einzelreihen mit 1,25 m Reihelänge in zweifacher Wiederholung.

Erfasst wurden Blühbeginn und Blühdauer sowie Abreife, der Chlorophyllgehalt indirekt als SPAD Wert vor Beginn der Abreife Anfang August, Blattform (nur Typ) sowie Wuchshöhe zur Abreife. Bei allen ausgewerteten Merkmalen zeigte sich ein signifikanter Effekt des Genotyps mit einer breiten Variabilität zwischen den Linien (Tabelle 18). Für die SPAD-Werte ist die Abstufung der Eltern und des Standards (Abbildung 53) vergleichbar zum Gefäßversuch (Abbildung 51), die Nachkommenschaft zeigt eine deutliche Transgression der Werte für dieses Merkmal. An 10 Pflanzen pro Parzelle wurden der Hülsenansatz am Nodium, der Ertrag sowie das TKG erfasst. Abbildung 54 zeigt die Verteilung des Hülsenansatzes, Abbildung 55 die der Einzelpflanzenenerträge. Beim Hülsenansatz entspricht die Abstufung der Eltern und des Standards noch dem Hülsenansatz im Gefäßversuch (siehe

Abbildung 51). Das liegt hauptsächlich darin begründet, dass der Gefäßversuch im Kurztag durchgeführt wurde, der Feldversuch aber unter typisch norddeutschen Langtagbedingungen abreifte. Hier hätten die Versuchsansteller eine Verschiebung der Höhe des Hülsenansatzes zu der im Langtag später abreifenden Sorte Sigalia erwartet, die noch an höheren Nodien Hülsen angesetzt hat. Das war allerdings nicht der Fall, wahrscheinlich auf Grund der fortschreitenden Trockenheit in der Vegetationszeit. Der geringere Ertrag von Merlin ist in der Frühzeitigkeit dieser Sorte im Vergleich zu Sigalia begründet. Für den Hülsenansatz, den Ertrag wie auch das TKG zeigt die Nachkommenschaft eine deutliche Transgression der Werte.

Der Hülsenansatz und Anzahl der Bohnen waren mit $r = 0,78$, Hülsenansatz und Ertrag an der Einzelpflanze mit $r = 0,72$ signifikant korreliert. Noch enger war die Korrelation der Anzahl der Bohnen je Einzelpflanze mit dem Ertrag mit $r = 0,88$. Zwischen Ertragsmerkmalen und SPAD Werten gab es keine signifikante Korrelation. Die grundlegenden statistischen Werte für alle Merkmale können Tabelle 16 entnommen werden.

Tabelle 18: Statistische Analyse der im Freiland erfassten Merkmale.

Merkmal	Variationsquelle	Signifikanz	Mittelwert	Spannweite	LSD
Blühbeginn	Genotyp	***	48,3	42 – 59	6,7
Tage nach Aussaat	Block	n.s.			
Blühdauer	Genotyp	***	45,5	36 – 57	12,7
Tage	Block	n.s.			
SPAD	Genotyp	***	38,2	31,0 - 46,4	5,9
SPAD Einheiten	Block	n.s.			
Wuchshöhe	Genotyp	***	85,1	52,5 - 112,5	28,9
cm	Block	**			
Abreife	Genotyp	***	141,0	118 - 163	25,4
Tage nach Aussaat	Block	***			
Hülsen/Einzelpflanze	Genotyp	***	21,1	12,4 – 33,8	15,4
Anzahl	Block	n.s.			
Ertrag/Einzelpflanze	Genotyp	***	6,9	3,2 - 11,5	6,4
g 86 % TS/EP	Block	n.s.			
TKG	Genotyp	***	219,3	167,0 - 265,4	65,8
g 86 % TS	Block	n.s.			

*** signifikant für $\alpha \leq 0,001$; ** signifikant für $\alpha \leq 0,01$; n.s. nicht signifikant, LSD kleinste signifikante Differenz für $\alpha \leq 0,05$; die Merkmale Hülsen und Ertrag je Einzelpflanze und das TKG wurden an 10 Einzelpflanzen/Wiederholung ermittelt.

Für die Bestimmung des Protein- und Ölgehaltes mittels NIRS wurden dem Projektpartner an der Landessaatgutstation der Universität Hohenheim die Körner der 10 Einzelpflanzen zur Verfügung gestellt.

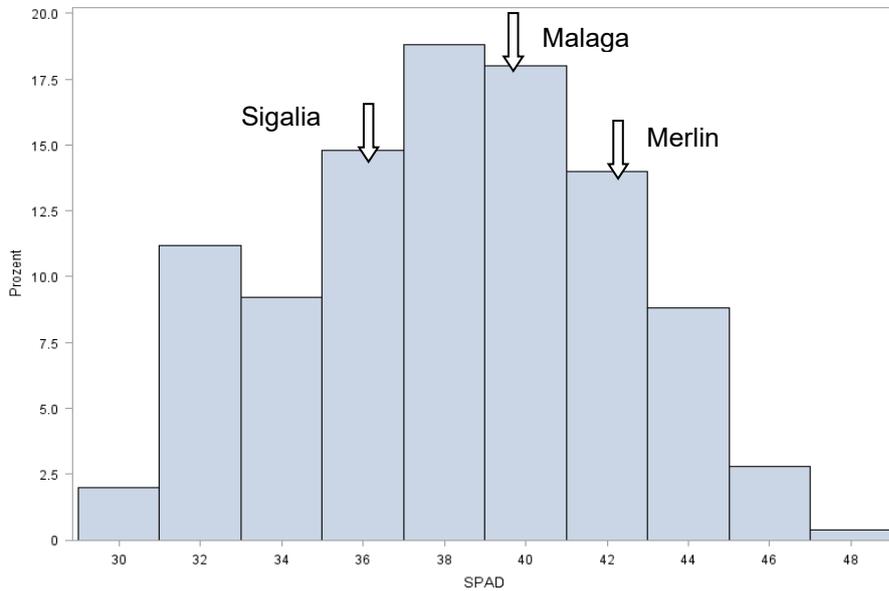


Abbildung 53: Häufigkeitsverteilung der SPAD Werte von Eltern und Nachkommenschaft sowie der kühesensitiven Sorte Malaga im Feld.

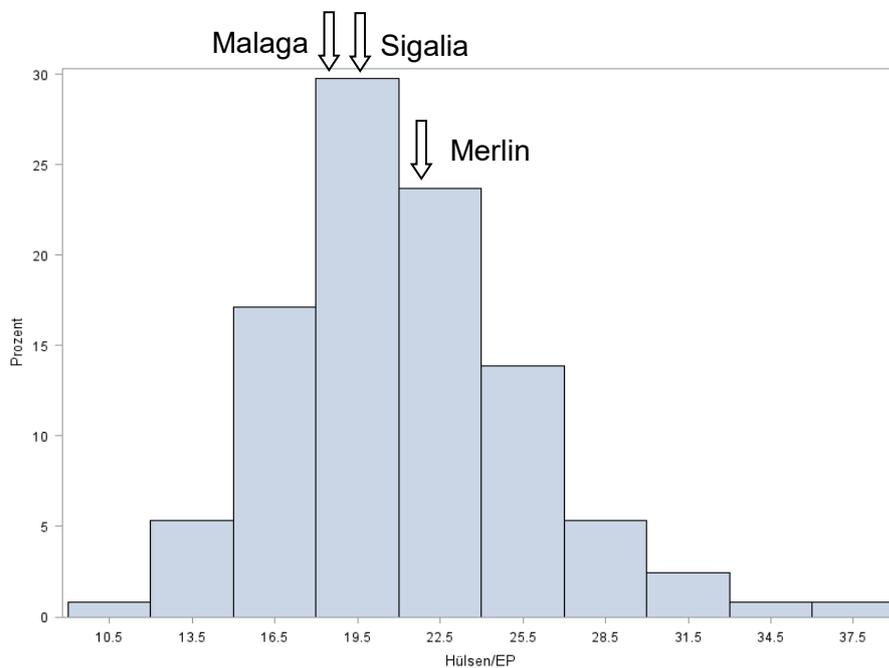


Abbildung 54: Häufigkeitsverteilung des Hülsenansatzes an der Einzelpflanze von Eltern und Nachkommenschaft sowie der kühesensitiven Sorte Malaga im Feld.

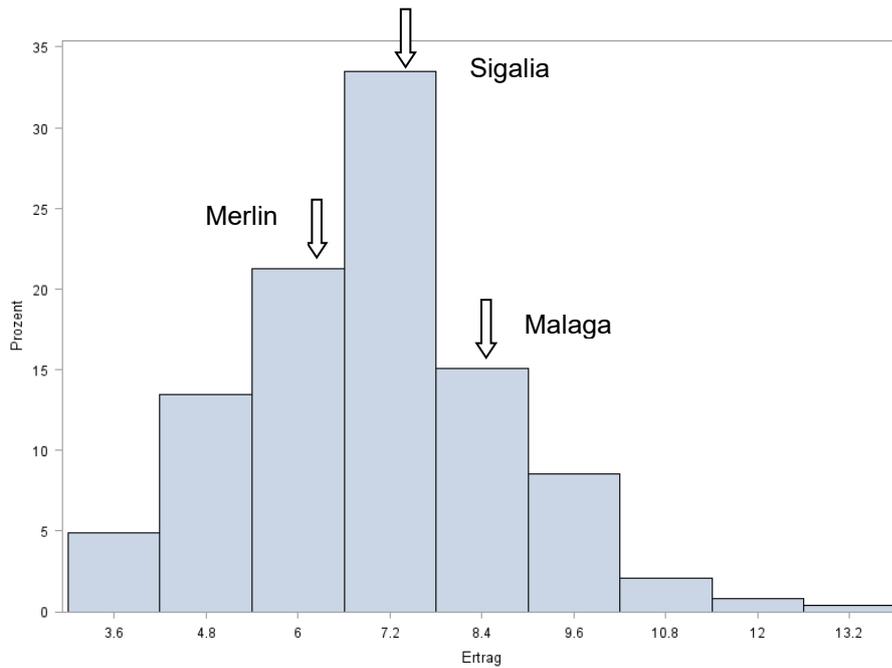


Abbildung 55: Häufigkeitsverteilung des Einzelpflanzenenertrages von Eltern und Nachkommenschaft sowie der kühesensitiven Sorte Malaga im Feld.

Um beim Eintreten entsprechender Witterungsbedingungen Unterschiede in der Kühletoleranz der Genotypen im Feld verifizieren zu können, wurde von allen Linien an 10 Einzelpflanzen pro Parzelle nicht nur der gesamte Hülsenansatz, sondern auch der Hülsenansatz am Einzelnodium erfasst. Leider war 2016 im Norden ein sehr warmes (CHU Wärmesumme 2971) und trockenes Jahr (175 mm Niederschlag, beides von Mai bis September) ohne erkennbaren Kühleeinbruch in der Hülsenansatzphase. Entsprechend zeigten die Genotypen auch keine Unregelmäßigkeiten im Hülsenansatz am Nodium, wie in Abbildung 56 exemplarisch für den sensitiven Standard Malaga dargestellt ist.

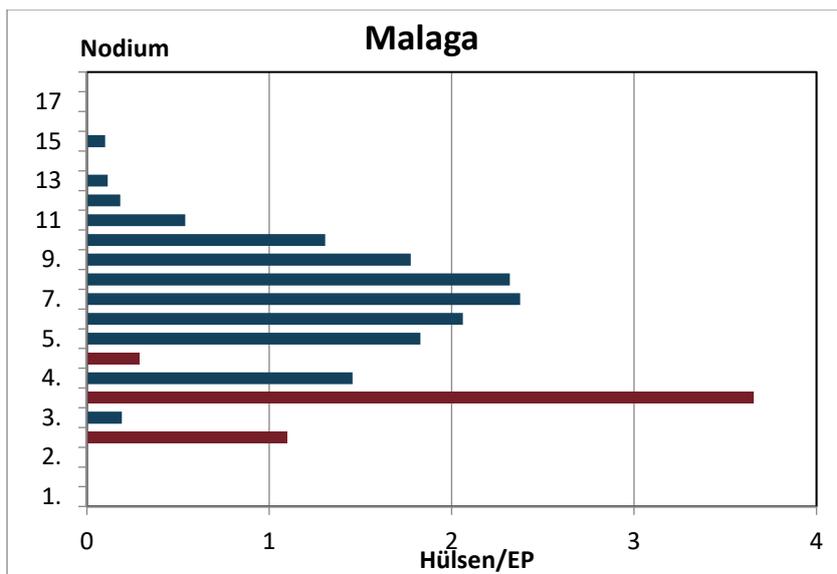


Abbildung 56: Hülsenansatz am Nodium für die kühesensitive Sorte Malaga im Feld 2016. Blau: Hülsenansatz am Hauptsproß, rot: Hülsenansatz an den entsprechenden Seitensprossen.

Gefäßversuche unter kontrollierten Bedingungen zur Erfassung des Hülsenansatzes unter Kühlestress (JKI)

Im Mittel weist der Hülsenansatz nach Stress einen signifikanten Abfall gegenüber der Kontrolle auf (Abbildung 57). Darüber hinaus ist neben signifikanten Effekten von Akzession und Stressvariante eine signifikante Akzession x Variante-Interaktion nachweisbar. Diese steht für Unterschiede in der Reaktion der Prüfglieder auf den Stressfaktor und ist Voraussetzung für eine Selektion auf eine verbesserte Toleranz.

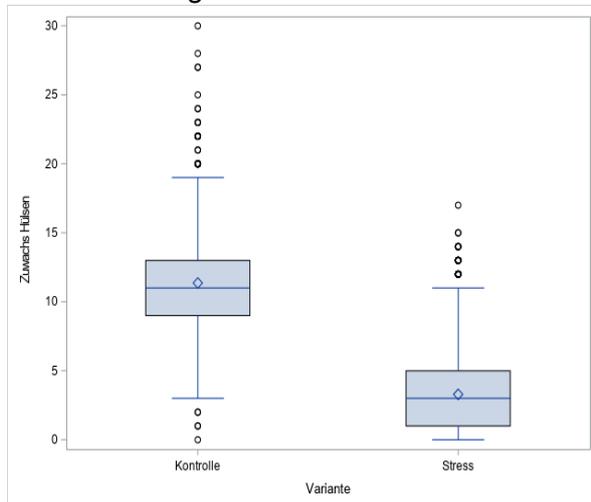


Abbildung 57: Hülsenansatz nach einer 3-wöchigen Kühlestressphase für die Kontroll- und Stressvariante, basierend auf 2 unabhängigen Wiederholungen.

Die grundlegenden statistischen Parameter sind in Tabelle 19 dargestellt.

Tabelle 19: Statistische Analyse des Merkmals „Hülsenansatz“ (***) signifikant für $\alpha \leq 0,001$; LSD für $\alpha \leq 0,05$).

Merkmals Einheit	Quelle der Varianz	Gesamt	Kontrolle	Kühlestress
Hülsenansatz (Anzahl)	Akzession	***	***	***
	Variante	***	Mittelwert 11,4	Mittelwert 3,3
	A x V	***	Spannweite 20,2 – 3,0 LSD 8,7	Spannweite 11,6 – 0,0 LSD 5,2

Die Verteilung des Hülsenansatzes über die Nodien ist über alle Prüfglieder für Kontrolle und Stressvariante grundsätzlich vergleichbar. Dabei werden unter Stressbedingungen im Mittel an allen Nodien weniger Hülsen angesetzt; das betrifft etwas vermehrt die unteren Seitentriebe (Nodium 1-3) und die Nodien ab Nodium 7 aufwärts (Abbildung 58).

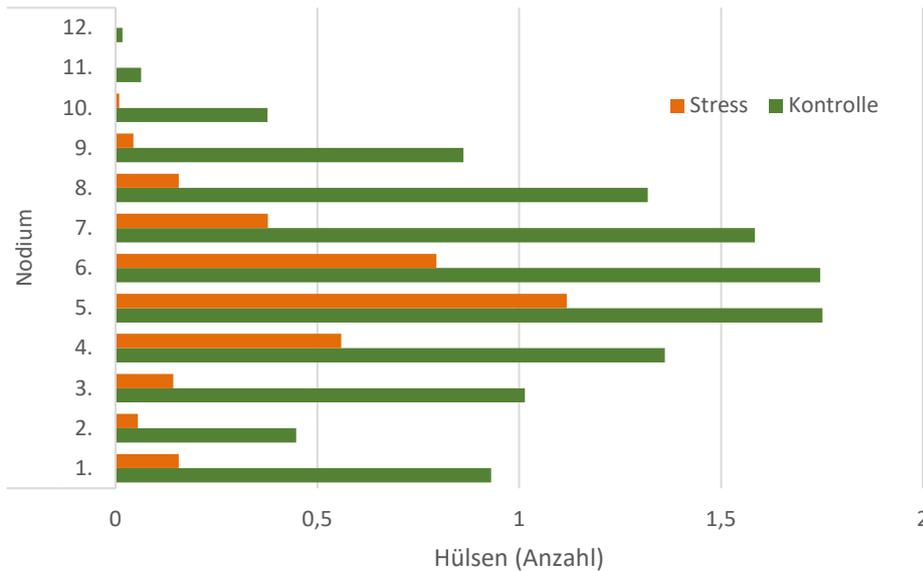


Abbildung 58: Mittlerer Hülsenansatz an den Nodien in der Kontroll- und Stressvariante unter Berücksichtigung aller Prüfglieder.

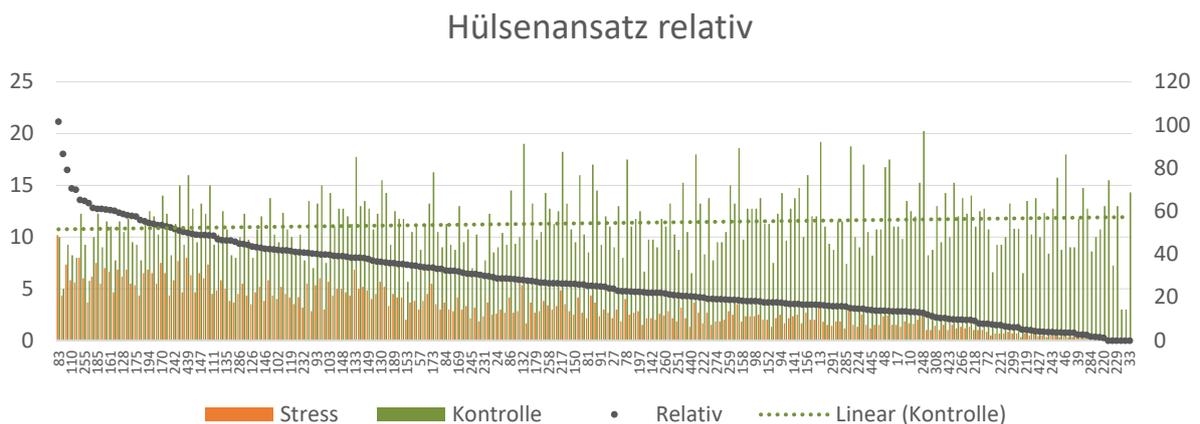


Abbildung 59: Hülsenansatz unter Kontroll- und Stressbedingungen sowie Stress/Kontrolle relativ in den 250 mit Kontrolle geprüften Prüfgliedern.

Die Relativwerte Stress/Kontrolle (%) zeigen wie erwartet, dass die Prüfglieder mit den höheren relativen Hülsenansätzen eher im niedrigeren Bereich der Hülsenansätze der Kontrolle liegen. Das bestätigt auch die tendenziell negative Regressionsgerade der Hülsen in der Kontrolle, wenn die Akzessionen nach absteigenden Relativansätzen sortiert werden (Abbildung 59). Abbildung 60 bestätigt diese Aussage. Trägt man den Hülsenansatz unter Stress gegen den der Kontrolle in einem Diagramm auf, entspricht die halbierende Linie einer Stabilität des Hvon 100 %. Auch hier zeigt sich, dass mit steigenden Hülsenzahlen in der Kontrolle die Abweichungen zu dieser Linie immer größer werden und somit die Stabilität im Hülsenansatz durch den Kühlestress bedingt sinkt. Allerdings ist die Variabilität im vorliegenden Material sehr hoch.

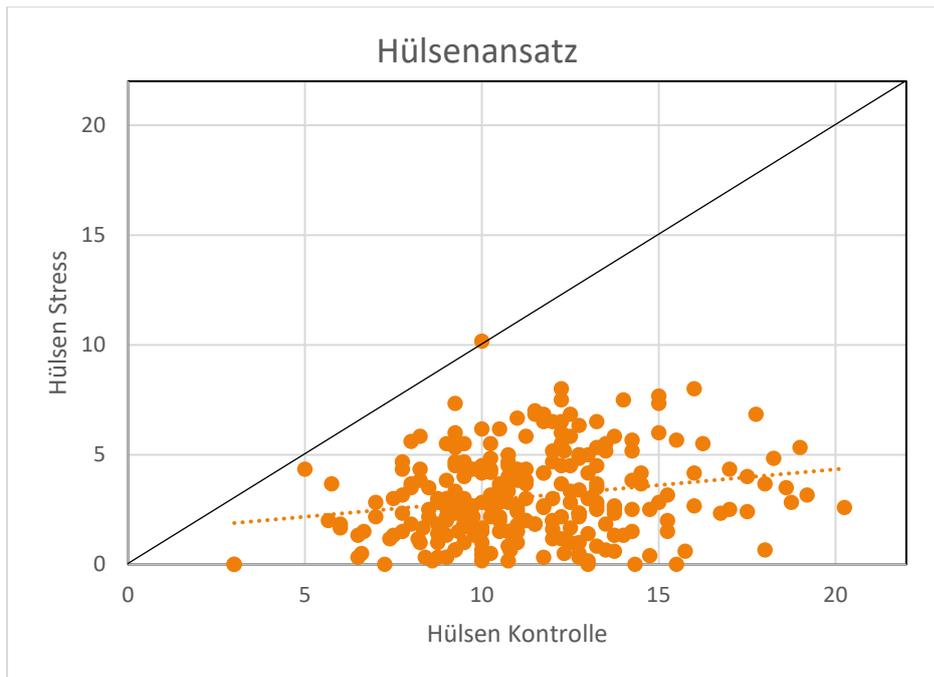


Abbildung 60: Hülsenansatz in der Kühlestress-Variante, aufgetragen gegen den Hülsenansatz der Kontrolle für die 250 in beiden Varianten geprüften Prüfglieder.

Im Vergleich der Ansätze Kreuzungsnachkommenschaft (1. Projektteil) und Sortiment (2. Projektteil) weist die Kontrolle der Kreuzungsnachkommenschaft einen etwas höheren mittleren Hülsenansatz bei vergleichsweise engerer Verteilung auf, was auf die Kreuzung zweier Leistungseltern zurückzuführen ist. Das Sortiment ist dagegen breiter aufgestellt mit einer leichten Verschiebung zu leistungsschwächeren Akzessionen. Unter Kühlestress wird der Unterschied noch deutlicher (Abbildung 61). Aus der Kreuzung einer kühletoleranten mit einer mittel-toleranten Sorte ging eine vergleichsweise tolerante Nachkommenschaft hervor. Das Sortiment zeigt unter Kühlestress eine deutliche Verschiebung zu einem geringeren Hülsenansatz und die Häufigkeitsverteilung wird auf Grund der Begrenzung durch den Ansatz 0 deutlich rechtsschief. Die Variabilität der Werte ist auch hier im Sortiment höher als in der Nachkommenschaft; es gibt aber nur wenige Akzessionen, die die der Nachkommenschaft noch übertreffen (Abbildung 49).

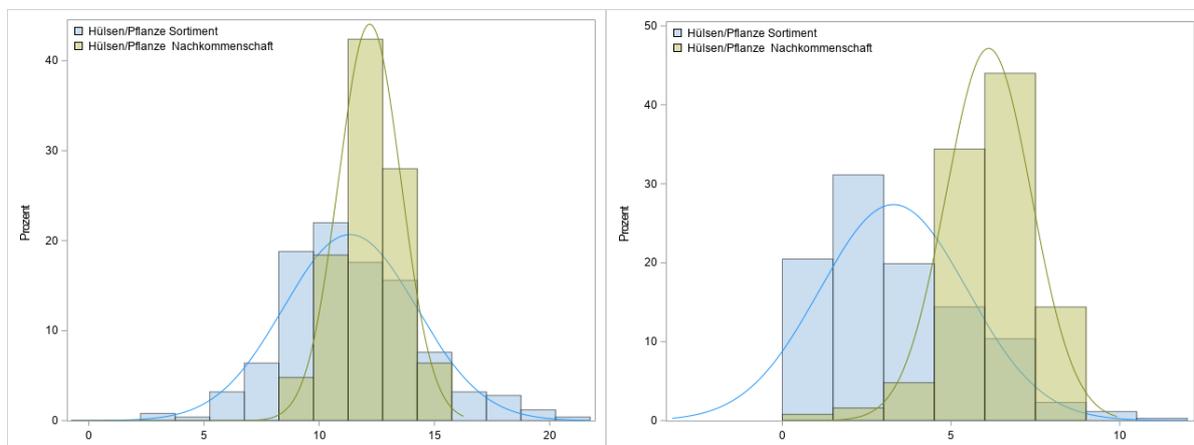


Abbildung 61: Häufigkeitsverteilungen des Hülsenansatzes für Nachkommenschaft und Sortiment in der Kontrolle (links) und in der Kühlestressvariante (rechts).

Im ersten Teil des Projektes wurde auch untersucht, in welchem Umfang sich die Heritabilitäten des Merkmals unter Stress- und Kontrollbedingungen verändern, wenn nur eine Kombination von 2, 3 oder 4 der 5 verfügbaren Superblöcke (SB, komplette Wiederholung) in die Analyse mit einfließen (Abbildung 50). Die Daten lieferten sehr robuste Werte, sodass selbst bei nur 2 Wiederholungen noch eine Heritabilität von 0,5 - 0,6 errechnet wurde. Daraufhin wurde entschieden, für die Phänotypisierung des Merkmals mit Ziel einer GWAS im 2. Projektteil die genotypische Variabilität zu Ungunsten der Reproduzierbarkeit zu erhöhen.

In diesem Ansatz des zweiten Projektteils mit nur 2 Wiederholungen lagen die Heritabilitäten mit 0,54 für die Kontrolle und 0,5 für die Stressvariante im unteren Bereich dieser Schätzung (Tabelle 20), waren für ein komplexes Merkmal aber immer noch gut.

Tabelle 20: Heritabilität des Hülsenansatzes unter Kontroll- und Kühlestressbedingungen im Sortiment.

Heritabilität h^2	Kontrolle	Kühlestress
Hülsenansatz	0,54	0,50

Genomweite Assoziationsstudie (GWAS) zum Hülsenansatz unter Kontroll- und Kühlestressbedingungen (HOH)

Von den untersuchten Sorten und Genbank-Akzessionen wurde eine Verwandtschaftsmatrix basierend auf den SNP-Daten erstellt (Abbildung 62).

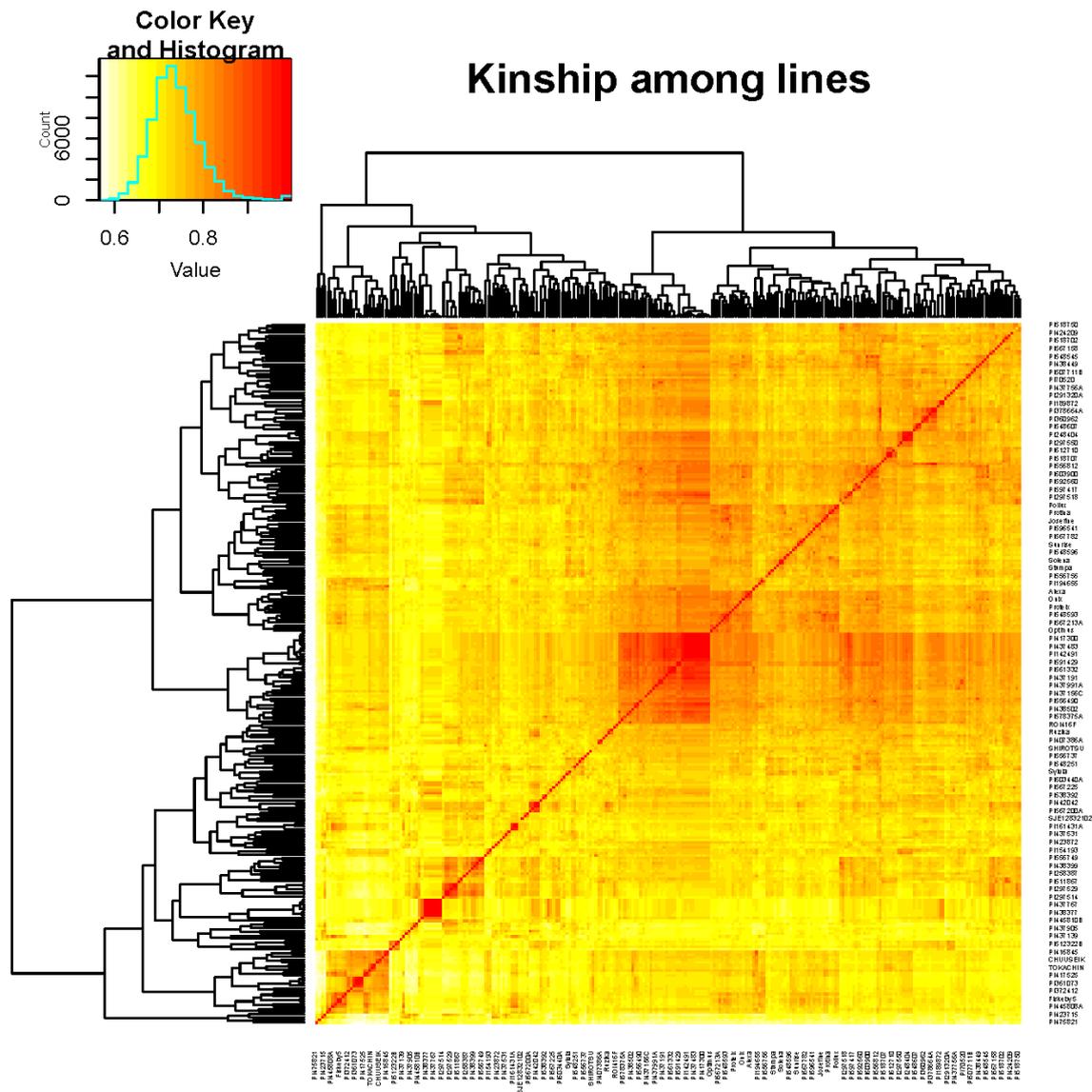


Abbildung 62: Darstellung der Verwandtschaftsmatrix, um die Beziehung zwischen den Akzessionen anzuzeigen. Die Namen auf der X-Achse und der rechten Y-Achse geben die einzelnen Akzessionen an, während die linke Y-Achse ein Dendrogramm darstellt, das die phylogenetische Beziehung auf der Grundlage der SNP-Daten zeigt.

Basierend auf den SNP-Daten wurden für alle Merkmale GWAS-Studien durchgeführt. Allerdings konnten keine signifikanten SNP-Merkmal-Assoziationen gefunden werden. Auch die in vorangegangenen Studien gefundenen QTL auf den Chromosomen 7 und 13 konnten nicht bestätigt werden.

5. DISKUSSION DER ERGEBNISSE

5.1 IPSP

Mit den beiden Kernkollektionen steht der mitteleuropäischen Sojazüchtung nun Genbankmaterial zur Verfügung, welches in der Zukunft gezielt und effektiv evaluiert werden kann. Saatgut für die ausgewählten Akzessionen kann bei der USDA Genbank zur freien Nutzung von Züchtern und Wissenschaftlern angefordert werden. Die Zusammenstellung dieser Kernkollektionen ist in zwei Schritten erfolgt: der erste Schritt umfasste die Selektion von Material, welches aufgrund seiner Herkunft grundlegend an mitteleuropäische Umweltbedingungen angepasst sein sollte. Die Verwendung von Umwelt-/Klimadaten zur Ableitung abiotischer Anpassung kompensiert dabei das Fehlen entsprechender Evaluierungsdaten und ermöglicht eine vielversprechende Vorauswahl von Genbankmaterial. Dieser Ansatz ist in der Literatur als *Focused Identification of Germplasm Strategy* beschrieben (Sanders et al. 2013; Street et al. 2008). Im zweiten Schritt wurden aus dieser Vorauswahl gemäß Kernkollektionsmethodik Akzessionen ausgewählt, die diese Gruppe ohne Redundanz in ihrer Diversität repräsentieren. Dies wird im Wesentlichen durch die Vermeidung genetischer „Doppelgänger“ in den Kernkollektionen erreicht. Beide Schritte stellen die Effektivität zukünftiger Phänotypisierungsbestrebungen sicher, indem (1) die richtige Materialgruppe evaluiert werden kann und (2) Kernakzessionen eine vergleichsweise große genetische Distanz zueinander aufweisen. So repräsentieren die Kernkollektionen in der Verteilung der Herkunftsumwelten das äußerste Ende kühler Herkunftsregionen hoher geografischer Breite und stellen damit einen guten Startpunkt für die Identifizierung vielversprechender Genbankakzessionen zur Verbesserung der abiotischen Anpassung an Mitteleuropa, aber auch für andere Merkmale, dar. Eine gründliche Phänotypisierung dieses Materials ist allerdings unumgänglich, um die vielversprechendsten Akzessionen für z.B. Testkreuzungen mit modernem Material zu identifizieren. Genbankmaterial wird moderne Sorten nur in Nebenaspekten komplementieren und um diese Überlegenheit vor dem genetischen Landrassenhintergrund erkennen zu können, sind Stressversuche und moderne Phänotypisierungsmethoden bspw. mit multispektral und RGB Kamera- und Drohnentechnik wertvolle Hilfsmittel. Vor dem Hintergrund der voranschreitenden Klimaveränderung mit zunehmend instabilen Wetterverhältnissen stellt Genbankmaterial hier eine zunehmend wertvoller werdende Ressource dar, um abiotische Anpassung zu untersuchen und ggf. in modernes Zuchtmaterial zu übertragen.

Die erfolgte Genotypisierung moderner europäischer Sojasorten ermöglichte es erstmals, das europäische Soja mit der globalen Sojapopulation, wie sie in der USDA Genbank repräsentiert ist, zu kontextualisieren. Unsere Ergebnisse hinsichtlich der grundlegenden genetischen Struktur, die von geografischer Herkunft und/oder Reifegruppenzugehörigkeit geprägt ist, stimmen im Wesentlichen mit früheren Untersuchungen, die dieses Material zum Gegenstand hatten, überein (Bandillo et al. 2015; Jarquin et al. 2016). Unsere Ergebnisse bestätigen außerdem die enge genetische Verwandtschaft des europäischen mit früh- bis mittelreifem nordamerikanischem Material (Hahn und Würschum 2014). Besondere Bedeutung kommt unserer Verknüpfung der ermittelten genetischen Ursprungspopulationen mit Herkunfts- und Umweltdaten der entsprechenden Akzessionen zu, die die Zuordnung zu chinesischen Ökotypen der Sojabohne (Hahn und Würschum 2014) erlaubte. Der Abgleich mit den Abstammungsverhältnissen in mitteleuropäischen Sojasorten zeigte, dass diese zu großen Anteilen eine Mischung aus japanischem, nord- und nordostchinesischem Material darstellen. Damit sind alle Materialgruppen, denen aufgrund ihrer Herkunft eine Anpassung an mitteleuropäische Umweltbedingungen unterstellt werden kann, prinzipiell vertreten. Material aus Nordostchina stellt hier insbesondere unter den frühesten europäischen Sorten die größten Anteile. Entsprechend scheint diese Gruppe eine besondere Rolle bei der bislang

erfolgten Umwelanpassung gespielt zu haben, was insbesondere auch auf photoperiodische Anpassung zurückzuführen sein könnte (Wolfgang und An 2017). Aufgrund des historischen Flaschenhalses der modernen Sojazüchtung (Gizlice et al. 1994) ist die Wahrscheinlichkeit hoch, dass diese Ursprungspopulationen weiterhin nützliche Variation beinhaltet die in modernem Zuchtmaterial nicht vertreten ist. Diese kann durch weiterführende, ggf. molekulare Arbeiten identifiziert und nutzbar gemacht werden (siehe Kapitel 7.1). Unsere Ergebnisse zur Populationsstruktur eröffnen insbesondere in Hinblick auf Fragestellungen zur Aufklärung der genetischen Architektur abiotischer Umwelanpassung vielversprechende Möglichkeiten für Modellstudien.

Die Untersuchungen von genetischen Ressourcen mit der Methode der Umweltassoziationskartierung (EAM) zeigte, dass sowohl bekannte Reifegene als auch bisher unbekannte Gene mit einer hohen statistischen Güte identifiziert werden können. Das Fehlen genauer geographischer Koordinaten in den Passportdaten konnte durch die Bildung von genetischen und regional verankerten Gruppen entlang eines latitudinalen Gradienten in China kompensiert werden. Neben der Assoziation einzelner Gene mit Umweltparametern zeigten sich auch umfangreiche Hinweise auf epistatische Interaktionen zwischen Genen, die aufgrund einer entsprechenden statistischen Kontrolle als robust angesehen werden können. In einem zweiten Schritt wurde untersucht, ob vorteilhafte Allele, die für eine Anpassung an kühle Regionen in Nordchina verantwortlich sind, auch im US-amerikanischen und europäischen Zuchtmaterial vorhanden sind. Diese Analyse ergab, dass einige bekannte Reifegene bereits im europäischen Material fixiert sind, die Mehrzahl der Gene aber nicht oder nur in wenigen Sorten vorhanden sind. Weil für diese Gene einfach Marker entwickelt werden können, ist es möglich durch gezielte Kreuzung von Sorten mit den entsprechenden Reifegenen.

5.2 LSA

5.2a QTL-Analysen für Ertrag, Reifezeit, Tausendkornmasse, Protein- und Ölgehalt

Agronomische Merkmale wie Ertrag, Pflanzenhöhe, Standfestigkeit und Saatgutgewicht stehen seit vielen Jahren im Mittelpunkt der Züchter. Als Hauptfaktoren für den Marktpreis von Soja zieht die Saatgutqualität zunehmend die Aufmerksamkeit der Züchter auf sich. Der negative Zusammenhang zwischen Ertrag und Qualität der Pflanzen macht es jedoch sehr schwierig, auf diese Merkmale zu selektieren. Daher ist die gleichzeitige Verbesserung von Ertrag und Qualität zu einem großen Problem für Sojabohnenzüchter geworden (Zhang et al. 2018).

In unseren Experimenten erhielten wir hohe Heritabilitätsschätzungen für alle Merkmale, was die gute Qualität unserer Versuchsdaten, die große genetische Variation unter den Nachkommen und die eher geringe Genotyp-Umwelt-Interaktion in unseren Untersuchungen verdeutlicht. Im Allgemeinen sind die Heritabilitäten in diesem mitteleuropäischen Sojamaterial vergleichbar mit den in früheren Studien für Soja aus verschiedenen Reifegruppen (000-VIII) publizierten Heritabilitätsschätzungen (siehe Kurasch et al. 2017). Selbst bei den Merkmalen mit geringer Variation zwischen den meisten der fünf Elternlinien wurden eine hochsignifikante genotypische Varianz sowie Merkmalsverteilungen beobachtet, bei denen die Nachkommen ihren jeweiligen Eltern weit unter- oder überlegen waren. Dies zeigte sich insbesondere bei den Familien, bei denen die Elternsorten in ihren Merkmalswerten sehr ähnlich sind. Diese für alle Merkmale und Familien beobachtete transgressive Segregation deutet darauf hin, dass die Eltern QTL-Allele tragen, die zumindest teilweise komplementär sind und die in ihren Nachkommen neu kombiniert werden. Dies bestätigt die beträchtliche genetische Vielfalt der Eltern und steht im Einklang mit früheren

Studien über agronomische Merkmale in Soja (Li et al. 2008; Mansur et al. 1996; Zhang et al. 2004).

Da ein zentrales Ziel des Sojaanbaus in Europa die Verbesserung der Proteinautarkie ist, stellt sich die Frage, wie die Proteinausbeute pro Fläche maximiert werden kann. Unsere Ergebnisse zeigten eine hohe positive Korrelation zwischen Proteinertrag und Bohnenertrag, die jedoch eine Autokorrelation ist, da der Proteinertrag als Bohnenertrag multipliziert mit dem Proteingehalt berechnet wird. Dennoch deutet die geringe Korrelation zwischen Proteinertrag und Proteingehalt auf eine viel höhere Relevanz der Erhöhung des Bohnenertrags für eine hohe Proteinausbeute hin als die Verbesserung des Proteingehalts. Unsere Ergebnisse zeigen, dass es möglich ist, zwei Arten von Linien zu unterscheiden, die den Proteinertrag maximieren. Erstere liefern hohe Proteinerträge pro Hektar hauptsächlich durch hohe Bohnenerträge, während der Proteingehalt zwischen 39 und 42 % liegt. Der zweite Typ zeigt etwas niedrigere Proteinerträge als der erste Typ, zeichnet sich aber durch einen sehr hohen Proteingehalt von über 42 % aus. Die Wahl des einen oder anderen Typs in der Pflanzenzüchtung hängt vom Verwendungszweck der geernteten Sojabohnen ab. So scheint es beispielsweise für Fütterungszwecke vielversprechender zu sein, sich auf Linien der ersten Art zu konzentrieren, während es bei lebensmitteltauglichen Sorten, insbesondere bei der Tofuproduktion, vorteilhaft ist, Linien der zweiten Art auszuwählen.

Eine Reihe von QTL, die wichtigen agronomischen und Qualitätsmerkmalen zugrunde liegen, wurden in den letzten Jahrzehnten publiziert. Bisher wurden zahlreiche QTL identifiziert, die Pflanzenhöhe, Samengewicht, Ölgehalt und Proteingehalt kontrollieren (www.soybase.org), basierend auf den unterschiedlichen genetischen Hintergründen, Anbauorten und statistischen Methoden. Darüber hinaus erschweren große Konfidenzintervalle um QTL die ursächliche Genidentifikation. In unseren Untersuchungen fanden wir je zwei QTL für den Ertrag, den Protein- und Ölgehalt, drei QTL für das Tausendkorngewicht und vier QTL für die Wuchshöhe. Diese wurden von verschiedenen Arbeitsgruppen gefunden und in soybase.org beschrieben.

5.2b QTL-Analysen für Quellvermögen, Tofuausbeute, Tofufestigkeit, Tofuwert

Intrinsische Faktoren, die die Tofueigenschaften beeinflussen

Die kontinuierliche Verteilung aller untersuchten Merkmale sowie die transgressive Segregation zeigen eine quantitative Vererbung für alle bewerteten tofubezogenen Merkmale. Alle Merkmale wurden stark vom Genotyp beeinflusst, was im Einklang mit früheren Studien steht, die auch einen signifikanten genotypischen Beitrag zur Variation von Tofueigenschaften berichteten (Aziadekey et al. 2002; Poysa und Woodrow 2002); . Auch wenn Erblichkeitsschätzungen ein sehr wichtiger Indikator für den möglichen Erfolg des Züchters bei der genetischen Verbesserung des jeweiligen Merkmals sind, haben nur wenige Heritabilitäten für Tofumerkmale evaluiert und publiziert. Aziadekey et al. (2002) berichteten beispielsweise über Vererbungsmöglichkeiten für Sojamilchfeststoffe, Tofuausbeute und Festigkeit, die je nach bewertetem Material zwischen 0,59 und 0,82, 0,31 und 0,69 bzw. 0,37 und 0,61 lagen. Die in unserer Studie beobachteten Heritabilitätsschätzungen liegen in einem ähnlichen Bereich und bestätigen die quantitative Natur dieser Merkmale. Dennoch zeigen die signifikanten genotypischen Varianzkomponenten und die Heritabilitätsschätzungen von über 0,60 in unserer Studie das große Potenzial einer erfolgreichen Selektion für die bewerteten tofubezogenen Merkmale in Sojabohnenzüchtungsprogrammen. Nur die Einweichmerkmale zeigten niedrigere Vererbungsschätzungen von 0,4 für 1h bis 0,57 für insgesamt 25h Einweichen, was jedoch auf eine Fehlervarianz des Messverfahrens zurückzuführen sein könnte.

Der Anbauort beeinflusst die Tofueigenschaften erheblich.

Auch wenn die genotypische Varianz für alle untersuchten Merkmale signifikant zur phänotypischen Variation beitrug, waren sie auch stark vom Standort beeinflusst. Die Tofuausbeute und der Tofuwert werden aus Tofuertrag, Sojamilchertrag und Tofufestigkeit abgeleitet, so dass die Standortwirkung auf diese Merkmale wahrscheinlich durch die hohe Standortwirkung der ursprünglichen Merkmale, hauptsächlich Sojamilchertrag und Tofufestigkeit, erklärt wird. Die einzige Eigenschaft, die vom Anbauort nicht betroffen war, war Tofuertrag. Die physikalischen Eigenschaften von Tofuquark hängen höchstwahrscheinlich von den Eigenschaften der Sojamilch ab, da daraus Tofu hergestellt wird (Kim und Wicker 2005). Sojamilch und Sojamilchfeststoffe werden durch den Standort beeinflusst, aber für unsere Analyse haben wir gleiche Mengen an gleichmäßiger dichter Sojamilch verwendet, um Tofuproben herzustellen. Damit stehen die Tofuproben nicht mehr in direktem Zusammenhang mit Sojamilchfeststoffen. Dies könnte erklären, warum es keinen signifikanten Ortswirkungseffekt auf den Ertrag des Tofus gab, sondern auf die Tofuausbeute, die aus Sojamilchertrag und Tofuertrag berechnet wird. Der beobachtete Einfluss des Ortes auf die Tofueigenschaften entspricht den bisherigen Erkenntnissen. Signifikante Standortauswirkungen wurden auch von Poysa und Woodrow (2002), die Sojamilchfeststoffe betreffend, von Poysa et al. (2006) für die Tofufestigkeit, von Min et al. (2005) für die Tofuausbeute und die Tofufestigkeit und von Bhardwaj et al. (1999) für Sojamilchfeststoffe, die Tofufestigkeit und die Wasseraufnahme nach 1h und 16h berichtet. Wie von uns beobachtet, berichteten Bhardwaj et al. (1999) sogar über einen noch größeren Ortseffekt als den genotypischen Effekt für Tofustärke und Wasseraufnahme, während die anderen Studien den Genotyp als primäre Variationsquelle bezeichneten.

Der Standorteffekt ergibt sich höchstwahrscheinlich aus Veränderungen in der Bohnenzusammensetzung, wahrscheinlich einschließlich der Proteinmenge und -qualität, die sich auf die Tofueigenschaften Sojamilchertrag, Tofuertrag und Tofufestigkeit auswirken. Eine Meta-Analyse von Rotundo und Westgate (2009) zeigte, dass der Proteingehalt von mehreren Umweltparametern beeinflusst wird, wobei Temperatur- und Wasserstress den Proteingehalt erhöhen bzw. verringern. Neben dem eigentlichen Proteingehalt wurde auch festgestellt, dass die Proteinzusammensetzung und die Zusammensetzung der Untereinheiten durch den Standort und das Jahr beeinflusst werden (Fehr et al. 2003, 2003; Jaureguy et al. 2011; Poysa et al. 2006; Yang und James 2014). Jaureguy et al. (2011) kamen zu dem Schluss, dass Wachstumsbedingungen wie Temperatur, Niederschlag und Mikroumgebung sowie Pflanzenernährung den Proteingehalt und die Zusammensetzung der Bohnen beeinflussen können. Ob dies jedoch auf bestimmte Mikronährstoffe oder eine andere spezifische Bodenzusammensetzung zurückzuführen ist, ist noch unbekannt und bedarf weiterer Forschung. Der erhebliche Einfluss der Lage auf die hier und früher beobachteten Tofueigenschaften kann für Tofuproduzenten problematisch sein, die darauf abzielen, dem Verbraucher eine konstant hohe Qualität zu bieten. Daher ist weiteres Wissen über die besonderen Faktoren, die diesem Standortwirkungseffekt zugrunde liegen, von großem wirtschaftlichem Interesse, da es den Tofuproduzenten ermöglichen würde, ihren Landwirten detaillierte Anweisungen zu geben, wie sie die Ernte behandeln sollen, um eine gleichbleibende Qualität zu erhalten oder die Tofuqualität im Allgemeinen zu verbessern.

Tofumerkmale variieren innerhalb und zwischen den Familien.

Die beiden untersuchten Familien unterschieden sich signifikant für alle Merkmale mit Ausnahme der Tofufestigkeit, bei der sich auch die Eltern nicht unterschieden. Optimale Werte für die Tofufestigkeit für den europäischen Tofumarkt liegen zwischen 85 und 120 N. Tofu mit Werten unter 85 N ist zu wässrig und über 120 N wird der Tofu zu trocken und bröckelig. Die Werte für die Tofufestigkeit der bewerteten RILs lagen daher alle im optimalen Bereich, mit einem Gesamtmittelwert von 99,3 N. Aufgrund unterschiedlicher Bewertungsmethoden und

verschiedener Arten von produziertem Tofu sind die Tofufestigkeitswerte unserer Studie und die Werte anderer Studien, deren Werte von 1 N bis 37 N reichen, nicht vergleichbar (Benassi et al. 2011). Die Sorte Primus ist eine etablierte Tofusorte der Taifun Tofu GmbH und aufgrund ihres hohen Proteingehalts wird die Sorte Protina auch für die Lebensmittelproduktion empfohlen, während Sigalia eher eine Futtermittelsorte mit hohem Ertrag ist. In beiden Familien beobachteten wir transgressive Segregation, d.h. Nachkommen, die ihre Eltern in Bezug auf die tofubezogenen Eigenschaften übertroffen haben. Unter den Elternsorten zeigte Primus, wie erwartet, immer den höheren Wert für alle Merkmale, mit Ausnahme des Einweichfaktors, für den Protina den höchsten Wert hatte. Überraschenderweise schnitten die Nachkommen aus der Kreuzung Primus×Sigalia bei den meisten Tofueigenschaften mit Ausnahme des Einweichfaktors besser ab als die Nachkommen von Primus×Protina. Insbesondere für die Tofuausbeute wird die Differenz wirtschaftlich bedeutsam, da eine Erhöhung der Tofuausbeute um 0,1 % bei gleichem Bohnenertrag von 3 t/ha zu ca. 300 kg mehr Tofu pro ha führt. Die Sorte Protina liefert im Vergleich zu Primus und Sigalia einen höheren Proteingehalt. Es wurde festgestellt, dass der Proteingehalt mit der Tofutextur und -Qualität zusammenhängt, aber wir konnten auch Nachkommen mit guten Tofueigenschaften identifizieren, die von der Kreuzung Primus×Sigalia abgeleitet wurden, die im Durchschnitt einen niedrigeren Proteingehalt aufweist (Kurasch et al. 2017). Interessanterweise ist der Proteingehalt bei den einzelnen Familien in der Netzwerkanalyse eng mit den Tofuqualitätsmerkmalen Tofufestigkeit und Tofuwert verbunden, während er bei der familienübergreifenden Analyse nicht so ausgeprägt ist. Zusammenfassend lässt sich sagen, dass der tatsächliche Wert des Proteingehalts möglicherweise nicht der Hauptfaktor für die Bestimmung der Tofufestigkeit ist, aber dennoch bieten die Nachkommen mit erhöhtem Proteingehalt in Familien eine bessere Tofutextur.

Einweichmerkmale und ihre Beziehung mit Tofueigenschaften

Tofuhersteller interessieren sich nicht nur für Genotypen, die in 25 Stunden das meiste Wasser aufnehmen, sondern auch dafür, welcher Genotyp am besten zu einem bestehenden Tofuproduktionssystem passt. Es wurde gezeigt, dass ein Einweichfaktor von 2,2 ein empfohlener Mindestwert für die Weiterverarbeitung der Bohnen zu Sojamilch und Tofu ist (Pan und Tangratanavalee 2003). Unsere Familien hatten beide Durchschnittswerte für den Einweichfaktor größer als 2,2, mit einem etwas höheren Durchschnitt für die Familie Primus×Protina. Wir konnten auch zeigen, dass sich die drei Elternsorten und die Familien in ihrem Einweichverhalten deutlich unterscheiden, wobei beispielsweise Primus das geringste Wasser in der ersten Stunde und das meiste danach aufnimmt. Das umgekehrte gilt für Protina. Es gibt Tofuproduktionssysteme, die mit sehr kurzen Einweichzeiten arbeiten, wobei folglich Genotypen bevorzugt werden, die in den ersten Stunden am meisten einweichen, während bei anderen Systemen, die mit langen Einweichzeiten arbeiten, es am wichtigsten ist, dass die Bohnen am Ende mindestens das 2,2-fache ihres Ausgangsgewichts zugenommen haben. Der Einfluss des Einweichens auf Sojamilch und Tofu war familienbedingt. Für die Familie Primus×Sigalia war die Gesamtwasseraufnahme nach 25 Stunden positiv mit dem Sojamilchertrag, Tofuertrag und Tofuausbeute verbunden, während für die Familie Primus×Protina kein solcher Zusammenhang beobachtet wurde. Daher kann daraus kein allgemeiner Rückschluss auf den Zusammenhang zwischen Einweichfähigkeit und Tofueigenschaften gezogen werden. Allerdings können Einweichzeit und -temperatur die Löslichkeit von Sojaproteinen beeinflussen (Lima et al. 2014). Wahrscheinlich ist dieser Unterschied in der Löslichkeit von Sojaproteinen auch genetisch bedingt, was die Korrelation des Einweichens mit Sojamilchertrag, Tofuertrag und Tofuausbeute für die Familie Primus×Sigalia, nicht aber für die Familie Primus×Protina, erklären könnte.

Beziehungen zwischen Tofueigenschaften

Im Rahmen der Selektion in Zuchtprogrammen sollten die tofubezogenen Merkmale in Originalmerkmale wie Einweichfaktor, Sojamilchertrag, Tofuertrag und Tofufestigkeit unterteilt werden, während Tofuertrag und Tofuwert besser als abgeleitete Merkmale beschrieben werden, da sie aus den Originalmerkmalen berechnet werden. Es wird daher erwartet, dass die abgeleiteten Merkmale stark mit den ursprünglichen Merkmalen korrelieren, was durch unsere Ergebnisse bestätigt wurde. Tofuwert und Tofufestigkeit waren stark positiv korreliert, aber die Tofuausbeute war fast nicht mit dem Tofuwert korreliert, was auf einen viel höheren Einfluss der Tofufestigkeit auf den Tofuwert hinweist als die Tofuausbeute. Es ist daher fraglich, ob der berechnete Tofuwert ein guter Indikator für die allgemeine Tofueignung eines Genotyps ist oder ob mehr Wert auf die Tofuausbeute gelegt werden sollte. Ebenso behindert die Tofufestigkeit, die negativ mit der Tofuausbeute und verwandten Merkmalen wie Tofuertrag und Sojamilchertrag korreliert ist, die erfolgreiche gleichzeitige Selektion sowohl auf Tofuausbeute als auch auf Qualitätsmerkmale.

Die wohl wichtigsten Eigenschaften für die Tofuherstellung sind die Tofuausbeute und die Tofufestigkeit. Die beiden Globuline Glycinin (11S) und β -Conglycinin (7S) sollen verschiedene Tofuqualitätsmerkmale wie Tofuausbeute, Sojamilchausbeute und Tofufestigkeit beeinflussen (Cai et al. 1997; Cai und Chang 1999; James und Yang 2014; Poysa et al. 2006; Stanojevic et al. 2011; Yang und James 2013). Daher sollten die künftigen Arbeiten darauf abzielen, diese Proteinfractionen zu untersuchen, um ihren Beitrag zur Tofuqualität in der europäischen Sojabohne zu bewerten.

Agronomische Merkmale versus Tofueigenschaften

Die Netzwerkanalyse ermöglicht einen schnellen Überblick über die Zuordnung von Merkmalen, was bei einer großen Anzahl von Merkmalen sehr nützlich ist. Das Netzwerk der gemeinsamen Analyse aller Individuen hebt zwei wichtige Cluster mit den eng verwandten agronomischen Merkmalen Bohnenertrag, Proteinertrag und Pflanzenhöhe sowie der sehr bekannten engen negativen Verbindung zwischen Proteingehalt und Ölgehalt hervor. Der Bohnenertrag zeigte keine starke Korrelation zu einem der Tofu- und Einweichmerkmale, so dass die Selektion auf Tofueigenschaften den Bohnenertrag nicht beeinträchtigt, was natürlich ein wichtiges Merkmal für Sojabohnenzüchter und -anbauer ist.

Die Bohnengröße, die in unserer Studie als Tausend-Korn-Gewicht gemessen wurde, war für beide Familien eng mit der Tofufestigkeit und dem Tofuwert verbunden, jedoch nicht mit der Tofuausbeute. Die beobachtete positive Korrelation bedeutet, dass größere Samen zu härterem Tofu führten. Die Zuordnung von Saatgutgröße zu Tofueigenschaften ist in der Literatur nicht eindeutig und uneinheitlich. Einige berichteten von einer positiven Korrelation zwischen Saatgutgröße und Tofutextur sowie zwischen Saatgutgröße und Tofuausbeute Aziadekey et al. (2002), was unsere Erkenntnisse zur Tofufestigkeit bestätigt. Andere fanden jedoch keinen Zusammenhang zwischen Saatgutgröße und Tofutextur oder Tofuausbeute (Cai et al. 1997; LIM et al. 1990; Wang und Chang 1995), ähnlich wie unsere Ergebnisse bezüglich des Tofuertrags. Huang et al. (2014) untersuchten den Zusammenhang zwischen Saatgutgröße und Tofutextureigenschaften genauer, indem sie Saatgutmenge und -form zur Messung der Saatgutgröße verwendeten, fanden aber keine Korrelation zwischen diesen Merkmalen der Saatgutgröße und der Tofutextur. So bleibt der Zusammenhang zwischen Saatgutgröße und Tofuqualität schwer fassbar und kann von dem untersuchten Sojamaterial abhängen. Im Allgemeinen bevorzugen die Tofuproduzenten jedoch größeres Saatgut, obwohl es keinen nachgewiesenen Zusammenhang zwischen Saatgutgröße und Tofuqualität gibt, was darauf hindeutet, dass diese Präferenz hauptsächlich auf der Tradition oder Erfahrung mit bestimmten Tofusorten beruhen kann.

Die Netzwerkanalyse aller Individuen muss mit Bedacht interpretiert werden, da das allgemeine Netzwerk Artefakte produzieren kann. So war beispielsweise der negative

Zusammenhang zwischen Saatgutgröße und Einweichfaktor nur für die Familie Primus×Sigalia vorhanden, so dass die allgemeine Schlussfolgerung, dass Saatgutgröße und Einweichfaktor negativ korreliert sind, nicht gezogen werden kann, sondern von der spezifischen Familie, d.h. vom genetischen Hintergrund, abhängt.

Haupt-QTL identifiziert für die Saatgutqualitätsmerkmale

Die meisten der bisher in Soja durchgeführten QTL-Studien konzentrierten sich auf Qualitätsmerkmale wie Protein- und Ölkonzentration sowie Ertragsmerkmale wie das Gewicht des Saatguts. Diese Eigenschaften sind auch für Tofuhersteller interessant und wichtig, da sie proteinreiche Sorten mit großem Saatgut bevorzugen, obwohl wir in keinen positiven Effekt dieser Eigenschaften auf die Tofuproduktion gefunden haben. Bisher wurden in Soybase (<http://www.soybase.org/>) von 240 QTL bezogen auf die Sojaproteinkonzentration berichtet (Grant et al. 2010), die sich auf jedem Chromosom des Sojagenoms befinden. Aufgrund ihrer physikalischen Lokalisation können die hier auf den Chromosomen 15 und 20 identifizierte Haupt-QTL für Proteinkonzentration identisch sein mit den zuvor beschriebenen QTL 'Seed Protein 1-3' und 'Seed Protein 1-4' von Diers et al. (1992), "Seed Protein 3-12" von Brummer et al. (1997), "Seed Protein 30-1" und "Seed Protein 30-3" von Tajuddin et al. (2003), "cqSeed Protein-003" von Nichols et al. (2006) und "Seed Protein 37-8" von Wang et al. (2014). Das günstige Allel für die Proteinkonzentration wurde immer von den Eltern mit dem höheren Proteingehalt bereitgestellt. Die Population "Primus × Protina" hat einen durchschnittlich höheren Proteingehalt mit einem Mittelwert von 40,8 %, da beide Elternteile einen relativ hohen Proteingehalt haben. Der auf dem Chromosom 20 in den beiden anderen Populationen identifizierte QTL könnte daher auch zur Proteinkonzentration in 'Primus × Protina' beitragen, kann dort aber für das günstige Allel fixiert sein.

Eine große Anzahl von 318 QTL im Zusammenhang mit dem Tausendkorngewicht ist bei SoyBase aufgeführt, von denen 10 auf dem Chromosom 15 zu finden sind (Han et al. 2012; Kato et al. 2014, 2014; Lee et al. 2001; Li et al. 2008; Liu et al. 2011; Moongkanna et al. 2011; Orf et al. 1999; Teng et al. 2009) und 14 auf Chromosom 17 (Chapman et al. 2003; Gai et al. 2007; Hoeck et al. 2003; Kato et al. 2014; Kulkarni et al. 2017; Kuroda et al. 2013; Li und Durbin 2010; Mian et al. 1996; Zhang et al. 2004), die beiden Chromosomen mit dem größten Tausendkorngewicht-QTL, die hier identifiziert wurden. Im Falle des Tausendkorngewichts trug der Elternteil mit dem höheren TKG immer das QTL-Allel bei, das das Gewicht erhöht. Die QTL für Proteinkonzentration und das Tausendkorngewicht erscheinen als vielversprechende Ziele für die markergestützte Selektion, müssen aber in weiteren Populationen validiert werden.

Die genetische Architektur von Tofueigenschaften

Unsere Studie zu Tofueigenschaften umfasste zwei Populationen mit insgesamt 171 RILs und ist damit eine der größten Studien dieser Art und die erste in Europa. Obwohl dies insgesamt eine große Zahl ist, gab es nur 78 und 93 RILs in den beiden Populationen, was die QTL-Detektionsleistung einschränkt, insbesondere bei QTL mit kleinen Effekten. Wir fanden jeweils einzelne QTL für Sojamilchertrag, Tofufestigkeit und Tofuwert in der größeren Population "Primus × Protina", die 18,84 %, 16,12 % bzw. 15,07 % der genotypischen Varianz erklärten. Über die genetische Kontrolle und die molekularphysiologischen Mechanismen, die zu Tofuattributen wie Sojamilchertrag, Tofuertrag und Tofufestigkeit beitragen, ist bisher wenig bekannt. Wahrscheinlich sind wasserlösliche Proteine ein kritischer Faktor für die verschiedenen Eigenschaften der Tofueigenschaften, da der erste Schritt der Tofuproduktion die Auflösung von eingeweichten und gemahlenden Sojabohnen in Wasser ist, was zu Sojamilch führt. Tofu wird dann durch die Koagulation der löslichen Proteine in der erhaltenen Sojamilch gewonnen. (Lu et al. 2013) postulierten, dass eine wünschenswerte hohe Produktionsmenge an Tofu und Sojamilch aus Sojabohnen mit einem höheren Gehalt an

wasserlöslichen Proteinen gewonnen werden kann. Sie berichteten über mehrere QTL für den Gehalt an wasserlöslichem Protein auf zehn Chromosomen, von denen qsp5-1' (berichtet als 'Samenprotein 34-1' bei SoyBase) auf Chromosom 5 zwischen 92,45 cM - 95,95 cM liegt (basierend auf der öffentlichen Sojabohnen-Karte GmComposite_2003_A1). Die beiden QTL 'qsp8-4' und 'qsp8-5', die bei SoyBase als 'Samenprotein 34-4' und 'Samenprotein 34-5' aufgeführt sind, befinden sich beide auf Chromosom 8, im gleichen physikalischen Bereich wie unser QTL für Tofufestigkeit. Zhang (2008) berichtet über zwei QTL, die mit der Produktion von Tofu und Sojamilch verbunden sind, die sich auf dem Chromosom 8 zwischen Satt424 und Sat_162 befinden, das sich in der Nähe der genomischen Region der QTL 'qsp8-5' von Lu et al. (2013) befindet. Diese Ergebnisse deuten auf einen möglichen Zusammenhang zwischen der wasserlöslichen Proteinkonzentration und den Tofueigenschaften hin. Interessanterweise waren nicht alle QTL, die mit dem Gehalt an löslichem Protein verbunden sind, identisch mit denen, die für den Gehalt an Samenproteinen gefunden wurden, was die fehlende Korrelation zwischen Tofueigenschaften und dem Proteingehalt erklären würde (Lu et al. 2013). Es wird oft behauptet, dass ein hoher Proteingehalt eine Voraussetzung für Tofusorten ist, aber eine positive Korrelation zwischen Proteinkonzentration und Tofueigenschaften konnte nicht eindeutig nachgewiesen werden.

Wir haben keinen QTL für Tofuertrag und Tofuausbeute gefunden. Zusammenfassend deuten unsere Ergebnisse auf einen hoch quantitativen Charakter aller Tofumerkmale hin, der durch zahlreiche kleine QTL-Effekte kontrolliert wird, die im QTL-Mapping nicht nachgewiesen werden können.

QTL für Quelleigenschaften

Das Einweichen der Sojabohnen ist der erste Schritt in der Tofuproduktion mit dem Ziel, die Samen weich zu machen und für die weitere Verarbeitung vorzubereiten. Dieser erste Schritt kann sehr zeitaufwendig sein, bis die Sojasamen einen ausreichenden Feuchtigkeitsgehalt von 120 % erreichen. Die Mindestzeit bei 25 °C, um diesen Richtwert zu erreichen, wurde mit 7 Stunden angegeben (Lima et al. 2014; Pan und Tangratanaavalee 2003). Um die Einwirkzeit zu verkürzen, kann die Temperatur geändert werden. Wenn eine höhere Temperatur die Einwirkzeit verkürzt, z.B. bei 55 °C und 70 °C wurde der empfohlene Feuchtigkeitsgehalt bereits nach 3 bzw. 2 Stunden erreicht (Lima et al. 2014). Die andere Seite der Medaille ist jedoch der von Lima et al. (2014) gemeldete Verlust des löslichen Proteingehalts bei Temperaturen über 25 °C, der die Produktion von Sojamilch und Tofu beeinträchtigen könnte. Dies zeigt die Bedeutung der Auswahl von Genotypen mit verbesserten Einweicheigenschaften, die an eine bestimmte Produktionsmethode angepasst sind, mit einem schnellen Einweichen bei niedrigen Temperaturen, um die löslichen Proteine zu erhalten. Möglicherweise könnte dies durch markerbasierte Ansätze unterstützt werden.

Bisher hat noch keine Studie die Genetik analysiert, die dem Quellverhalten von Sojasamen für die Verarbeitung zugrunde liegt. Die QTL, den wir für die Gesamtgewichtszunahme nach 25 Stunden gefunden haben, erklärten 10,9 %, 27,61 und 30,32 % der genotypischen Variation und lagen auf den Chromosomen 4, 20 und 11. Allerdings war ihre Wiederfindungshäufigkeit nach der Kreuzvalidierung eher gering, was auf ein geringes Vertrauen in diese QTL hindeutet. Die mit der Gewichtszunahme nach einer Stunde verbundenen Quell-QTL scheinen die zuverlässigsten Einweich-QTL zu sein. Sie wurden auf den Chromosomen 1, 14 und 15 identifiziert und erklären zwischen 42,40 % und 74,22 % der genotypischen Variation. Diese QTL könnten somit als Haupt-QTL klassifiziert werden und sind daher für die zukünftige Forschung interessant. Sie waren jedoch in jeder Population unterschiedlich, was ihre Verwendung in der markergestützten Selektion über Populationen hinweg in einem Zuchtprogramm einschränken kann. Obwohl die Heritabilitätsschätzungen eher niedrig waren, deuten diese Ergebnisse darauf hin, dass die Einweichmerkmale eine

weniger komplexe genetische Architektur aufweisen könnten als die Tofueigenschaften, wobei einige wichtige QTL die meiste Variation erklärten.

5.2c Genomische Selektion

In einer jüngst erschienenen Publikation zur genomischen Selektion bei Sojabohnen lautete die Key Message: „Im Vergleich zur phänotypischen Selektion wurde durch genomische Selektion eine höhere Effizienz erreicht“ (Matei et al. 2018). Ein hauptsächlichlicher Effekt ist, dass mit genomischer Selektion die Zeit, die für einen Selektionszyklus benötigt wird, verringert werden kann. In der zitierten Studie lag die Vorhersagegenauigkeit für den Bohnenertrag zwischen 0,37 bei der Untersuchung von Sorten und 0,74 wenn rekombinante Inzuchtlinien untersucht wurden. Andere Untersuchungen (Ma et al. 2016) ermittelten eine Vorhersagegenauigkeit von 0,47 – 0,49 für den Ertrag und bis zu 0,78 für die Pflanzenlänge. In unseren Untersuchungen fanden wir eine Vorhersagegenauigkeit von 0,65 für den Ertrag und 0,82 für die Pflanzenlänge lagen damit in einem ähnlichen Bereich.

Potenzial der genomischen Selektion für die Tofu-Sojabohnenzüchtung

Wir fanden keinen oder nur einen QTL pro Tofumerkmal, der meist nicht einen ausreichend hohen Anteil der genotypischen Varianz erklärte, um eine markergestützte Selektion zu rechtfertigen. Wir haben daher das Potenzial der genomischen Selektion evaluiert, da diese es ermöglicht, auch die Auswirkungen von unidentifizierten kleinen QTL zu erfassen. Die erhaltenen Schätzungen der Vorhersagegenauigkeit der Qualitätsmerkmale Proteingehalt und Samengewicht können als Referenz dienen und das im Allgemeinen hohe Potenzial der genomischen Selektion in der Sojabohnenzüchtung unterstreichen.

Die Einweichmerkmale zeigten vergleichsweise hohe Vorhersagegenauigkeiten, wenn auch mit einer sehr großen Variation innerhalb der einzelnen Populationen. Dennoch könnte die genomische Selektion auch für sie vielversprechend sein. Bezüglich der Tofumerkmale waren die Schätzungen der Vorhersagegenauigkeit jedoch eher gering mit großen Spannweiten in beiden Populationen. Unter den Tofumerkmalen hatten das Sojamilchgewicht und der Tofuwert die höchsten Schätzungen der Vorhersagegenauigkeit in der Population Primus × Sigalia mit 0,40 bzw. 0,44, gefolgt von der Tofuausbeute (0,33) und der Tofuhärte (0,23). Während in der Population "Primus × Sigalia" keine QTL identifiziert wurden, waren die Vorhersagegenauigkeiten im Durchschnitt etwas höher. Der Tofuertrag zeigte in beiden Populationen eine gleich geringe Vorhersagegenauigkeit.

Die geringe Vorhersagegenauigkeit der Tofueigenschaften im Vergleich zu den anderen untersuchten Merkmalen weist auf die Komplexität der Tofueigenschaften in zweierlei Hinsicht hin. Erstens kann ihre vermutlich komplexe genetische Architektur, wie bereits erwähnt, zu niedrigeren Vorhersagegenauigkeiten führen, da die Wirkungsschätzungen weniger zuverlässig werden, ein Trend, der auch für verschiedene Rapsmerkmale beobachtet wurde (Würschum et al. 2014). Andererseits ist es fraglich, ob die genetische Architektur der Tofueigenschaften wirklich so viel komplexer ist, als beispielsweise die der Proteinkonzentration. Der Unterschied kann also auch auf die weniger genauen phänotypischen Daten zurückzuführen sein, was sich in den niedrigeren Heritabilitätsschätzungen widerspiegelt. Hochwertige phänotypische Daten sind eine entscheidende Voraussetzung für jeden genomischen Ansatz und es müssen Anstrengungen unternommen werden, um die Phänotypisierung von tofubezogenen Merkmalen zu verbessern, wie von uns vorgeschlagen (Kurasch et al. 2018a). Trotz der hier erzielten eher geringen Vorhersagegenauigkeiten erscheint die genomische Selektion aufgrund des arbeits- und zeitaufwendigen Verfahrens zur Phänotypisierung der tofubezogenen Merkmale weiterhin vielversprechend. Wenn die Phänotypisierung nur einiger Individuen pro Population eine moderate Vorhersagegenauigkeit bei der Vorhersage der verbleibenden Individuen aus dieser

Population erreicht, würde dies die Züchtung von Tofueigenschaften erheblich unterstützen und somit weitere Forschungen rechtfertigen.

5.2d,e Kreuzung und Weiterführung Diallel1_neu und Diallel2_neu

Die Kreuzungen wurden wie geplant durchgeführt und das Material wurde mittels Single-Seed Descent (ssd) weitergeführt. Allerdings ist es sehr schwierig, Kreuzungen zwischen Sojapflanzen durchzuführen was einen systematischen Populationsaufbau stark erschwert. Die Anzahl Samen, die je 100 Kreuzungen erhalten werden konnten, lag zwischen 0 und 35.

5.2f QTL-Kartierung der Kühletoleranz

Siehe Kapitel 5.4.

5.2g Erweiterung des Genpools durch Kreuzungen mit selektiertem Material aus AP1

Siehe Kapitel 7.1 letzter Absatz.

5.2h Sortenprüfungen

Die Sortenprüfungen an den Standorten Hohenheim und Eckartsweier zeigten, dass die Erträge sehr stark vom jeweiligen Jahr abhängig sind. Das Jahr 2017 brachte an beiden Standorten die höchsten Spitzenerträge mit mehr als 47 dt/ha an beiden Standorten. Der Höchstertrag wurde jeweils erzielt von der Sorte ESMentor. Diese Sorte zeigte auch in Landessortenversuchen in Baden-Württemberg, Rheinland-Pfalz und Hessen sehr gute Erträge. So erzielte sie beispielsweise in Orschweier einen Ertrag von 54,5 dt/ha (Bechtold et al. 2017). 2015 und 2016 wurden am Standort Eckartsweier deutlich höhere Spitzenerträge erzielt als in Hohenheim, 2017 waren die Spitzenerträge an beiden Orten ähnlich hoch, 2018 waren die Erträge in Hohenheim höher, wahrscheinlich bedingt durch die Trockenheit. Betrachtet man alle Jahre, so ist auffällig, dass keine Sorte in allen Jahren und Orten in den Top5 für den Ertrag zu finden ist. Dies bedeutet, dass die Sorten insgesamt eine starke Umweltwechselwirkung aufweisen. Ob neuere Sorten über mehrere Standorte und Jahre stabiler sind, werden die zukünftigen Ertragsversuche zeigen. Erfreulich ist, dass für die doch noch eher geringe Anbaufläche in Deutschland sehr viele Sorten zur Verfügung stehen und dass auch neue Sorten gezüchtet werden.

5.2i Speed-Breeding System

Ein Speed-Breeding System für Sojabohnen schafft die Möglichkeit, bis zu 5 Generationen pro Jahr zu erzeugen. Besonders interessant wird es, wenn diese raschen Generationen mit genomischer oder phänomischer Selektion unterstützt werden. In unserem System wurden Kreuzungsnachkommenschaften angebaut, mit monogenen Markern auf das Vorhandensein des erwünschten Merkmals untersucht und die selektierten weitergeführt. Wenn nur ein Container zur Verfügung steht, hat man das Problem, dass man nicht so viele Generationen durchführen kann, da die Temperatur in der Kammer zur Durchführung von Kreuzungen niedriger sein muss, um ausreichend große Blüten zur Kastration und genügend lebensfähigen Pollen zu erhalten.

5.2k Ionomics

Elemente sind zusammen mit Nukleinsäuren, Proteinen und Stoffwechselprodukten ein wesentlicher Baustein der lebenden Zelle und an fast allen Prozessen in einem Organismus beteiligt. Das Verständnis der Funktionen und der Dynamik der Elemente ist daher entscheidend für das Verständnis der Funktionsweise des Lebens. Für Pflanzen, die alle Elemente mit Ausnahme von Kohlenstoff und Sauerstoff aus der Bodenumgebung aufnehmen müssen, ist die Kontrolle der Aufnahme und Verteilung von Elementen aus der lokalen Umgebung überlebenswichtig. Das Ionom ist definiert als "die Zusammensetzung der Mineralstoffe und Spurenelemente eines Organismus, die die anorganische Komponente der

zellulären und organismischen Systeme darstellt" (Salt et al. 2008). Es handelt sich um ein dynamisches Netzwerk von Elementen, die durch die Physiologie und Biochemie der Pflanze, die letztlich durch das Genom gesteuert werden, als Reaktion auf die Umwelt kontrolliert werden (Baxter 2010). Die verfügbare Literatur zu ionomischen Ansätzen in Kulturarten ist bisher begrenzt, jedoch wird das Ionen-Profilung neue Türen zum Verständnis der Regulation wichtiger agronomischer Merkmale öffnen (Shah et al. 2018). Der Gehalt an Mineralstoffen bestimmt direkt den ernährungsphysiologischen Wert von Sojaprodukten in der menschlichen Ernährung, aber ebenso, wenn Soja über den Umweg als Futtermittel zum Menschen gelangt. Im Sinne einer Biofortifikation muss es das Ziel sein, positive Mineralstoffe anzureichern, um diese über die Ernährung dem Menschen zur Verfügung zu stellen (Singh et al. 2022). Beispiele hierfür sind die Mineralstoffe Eisen, Zink und Selen, die für verschiedenste molekulare und physiologische Prozesse unentbehrlich sind. In unseren Untersuchungen konnten wir nachweisen, dass für den Eisen- und Zinkgehalt in Sojabohnen signifikante genotypische Unterschiede bestehen und es somit möglich ist, auf höhere Gehalte zu züchten. Die negative Korrelation, die vor allem bei Zink zum Ertrag besteht, macht es jedoch schwierig, gleichzeitig den Bohnenertrag zu erhöhen und einen hohen Zinkgehalt zu erzielen. Daher müssten die Studien ausgeweitet werden um zum Beispiel Genbankmaterial zu untersuchen um herauszufinden, ob es nicht Sojabohnen gibt, die einen deutlich höheren Zink und Eisengehalt aufweisen. Zudem ist noch unklar, wie stark Düngung und Bodeneinflüsse eine Rolle spielen. Interessant ist auch, dass recht hohe negative Beziehungen von Zink- und Kalziumgehalt zum Molybdängehalt auftraten. In unseren Untersuchungen wurden keine Proben mit einem nachweisbaren Cadmiumgehalt gefunden. Um hier weitere Untersuchungen zur Genetik der Cadmiumaufnahme durchzuführen, sollten Sojabohnen an Orten angebaut werden, die nachweislich eine hohe Cadmiumkonzentration bei Sojabohnen hervorrufen. Solche Orte wurden auch für Deutschland bereits publiziert (Franzaring et al. 2019).

5.2I Genotypisierung

Im Labor wurden publizierte Marker etabliert und teilweise in KASP-Marker umgewandelt. Somit können Kreuzungsnachkommen von versehentlichen Selbstungen unterschieden werden. Für verschiedene Merkmale stehen Marker zur Verfügung, die im Zuchtprogramm eingesetzt werden. Die Genotypisierung wurde weiter verbessert, es besteht jedoch noch Potential für weitere Verbesserungen, da der benötigte zweite Read noch nicht die optimale Qualität aufweist

5.2m regionale genomische Selektion

Die genomische Selektion wurde für eine Reihe von Kulturpflanzen bewertet. Für Sojabohnen liegen vergleichsweise wenige Studien vor, obwohl der Ansatz das Potenzial hat, die Züchtung auf Krankheitsresistenz abiotische Stressresistenz und ertrags- oder ertragsbezogene agronomische Merkmale zu verbessern (Zhu et al. 2021b).

Im Allgemeinen bleibt die ideale Zusammensetzung des Trainingsatzes bei der genomischen Selektion eine Herausforderung, aber die Befolgung der Richtlinien, die aus der vorliegenden Studie entstanden und publiziert sind (Zhu et al. 2021b), sollte die Chance erhöhen, eine zufriedenstellende Vorhersagegenauigkeit und damit den erwarteten Selektionsgewinn zu erzielen. Welche Art von Material als Trainingsatz zur Verfügung steht, hängt auch von der Kultur und der Organisation des Züchtungsprogramms ab. Wenn einige Individuen einer Familie bereits phänotypisiert wurden, lassen sich die verbleibenden Individuen am besten durch diese Vollgeschwisterstämme vorhersagen, sofern eine Mindestanzahl davon verfügbar ist (Jose J. Marulanda et al. 2015). Wenn die Anzahl der Vollgeschwister zu gering ist, können sie mit Halbgeschwisterfamilien kombiniert werden (Brauner et al. 2020). Sollen dagegen alle Individuen einer Familie vorhergesagt werden, um diejenigen auszuwählen, die in eine

intensive Feldprüfung gehen sollen, ist es am besten, auf Halbgeschwister oder ähnlich verwandtes Material zurückzugreifen. In Kenntnis der Vorhersagekraft von Halbgeschwisterfamilien könnte ein Züchtungsprogramm, das die genomische Selektion einbezieht, versuchen sicherzustellen, dass für jede neue Kreuzung, die initiiert wird und für die eine genomische Selektion auf die Nachkommenschaft angestrebt wird, genügend Halbgeschwisterfamilien zur Verfügung stehen, um einen zuverlässigen und effektiven zusammengesetzten Trainingsset zu bilden. Die Verwendung von nicht verwandten Familien scheint im Allgemeinen riskant zu sein, und ihre Einbeziehung in ein zusammengesetztes Trainingsset kann sich nur dann positiv auswirken, wenn das Trainingsset ansonsten zu klein ist und wenn sie eine ähnliche genetische Architektur aufweisen. Zusammenfassend lässt sich sagen, dass unsere Ergebnisse das Potenzial der genomischen Selektion für die Sojabohnenzüchtung verdeutlichen und in einem breiteren Kontext einige allgemeine Regeln für die Optimierung der genomischen Selektion in der Pflanzenzüchtung untermauern. Damit ist auch klar, dass eine regionale genomische Selektion nicht so einfach durchzuführen ist, da das dafür notwendige Trainingsset nicht die erforderliche Struktur haben wird. Inzwischen wurde jedoch auch das Konzept der phänotypischen Selektion, basierend auf Nahinfrarotspektren, entwickelt. Die Nahinfrarotspektroskopie wird in der Regel zur Vorhersage von Zielmerkmalen wie dem Proteingehalt verwendet, die darauf basieren, dass einige Wellenlängen Informationen über das Merkmal erfassen und somit zur Merkmalsvorhersage verwendet werden. (Rincenc et al. 2018) haben kürzlich die phänotypische Selektion als kostengünstige und durchsatzstarke Methode vorgeschlagen, die ebenso einfach wie brilliant ist. So wie molekulare Marker in der genomischen Vorhersage verwendet werden, können verschiedene Wellenlängen als Prädiktoren im statistischen Modell für die phänotypische Vorhersage verwendet werden. Unter Verwendung von NIRS-Daten aus einem Weizen- und einem Pappel-Datensatz konnten sie phänotypische Vorhersagefähigkeiten erreichen, die genauso hoch oder sogar besser waren als die mit molekularen Markerdaten. Anschließend wurde gezeigt, dass die phänotypische Vorhersage bei Sojabohnen, Weizen, Mais und Roggen gut funktioniert (Cuevas et al. 2019; Galán et al. 2020; Krause et al. 2019; Lane et al. 2020; Parmley et al. 2019).

Unsere Untersuchungen zeigten, dass die phänotypische Selektion genauso gut und teilweise besser anwendbar ist als die genomische Selektion (Zhu et al. 2021a). Für Vorhersagen an verschiedenen Standorten können NIRS-Daten von einem einzigen Standort verwendet werden. Allerdings sollten NIRS-Daten aus verschiedenen Umgebungen, z. B. aus verschiedenen Jahren, durch gemeinsame Genotypen in den Trainings- und Vorhersagesätzen verbunden werden. Die phänotypische Vorhersage scheint weniger anfällig für die Verwandtschaft zwischen den Individuen in den Trainings- und Vorhersagesätzen zu sein als die genomische Vorhersage, da im Allgemeinen Halbgeschwisterfamilien, aber auch nicht verwandte Familien hohe Vorhersagefähigkeiten erreichten. Außerdem führte die phänotypische Vorhersage bei gleichem Trainingsumfang zu einer höheren Vorhersagekraft als die genomische Vorhersage. Die phänotypische Vorhersage kann in verschiedenen Phasen eines Züchtungsprogramms angewandt werden, und insgesamt unterstreichen unsere Ergebnisse das Potenzial dieses Ansatzes zur Steigerung des genetischen Gewinns in der Pflanzenzüchtung. Die ersten NIR-Spektren von jungem Zuchtmaterial wurden erhoben und das Verfahren wird in das Zuchtprogramm überführt. Da die Verwandtschaft hierbei keine so große Rolle spielt, können jetzt auch regional unterschiedliche Selektionen durchgeführt werden.

5.3 IPM

Der Einzelnachweis für vier Hauptpathogene (*S. sclerotiorum*, *P. ultimum*, *R. solani* und *P. syringae* pv. *glycinea*) wurde etabliert. Die optimale Annealingtemperatur der Primerpaare für die DNA-Amplifikation aller o. g. Pathogenen lag zwischen 61,5°C und 62,5°C. Dies wird die

Etablierung des Multiplexverfahrens für diesen Pathogenen erleichtern. Die Einzelnachweise wurden an Blatt-, Stängel- und Saatgutproben angepasst. Des Weiteren wurden Duplexnachweise für Ss, Fs und Psg etabliert. Die Einzel- und Duplexnachweise wurden hierbei als qualitative Nachweise etabliert.

Allerdings muss für diese Reaktionen noch die Nachweisgrenze bestimmt werden. Dazu werden absteigende Konzentrationen von Pilzsporen bzw. Myzelmengen vorbereitet. Die DNA wird isoliert und mit Hilfe des oben entwickelten qPCR-Verfahrens analysiert. Ziel des Ansatzes ist es, die niedrigste noch nachweisbare Sporen- bzw. Myzelmenge zu ermitteln.

In einem nächsten Schritt wurde untersucht, ob die verwendeten Primerpaare genug spezifisch und tatsächlich für Einzelnachweise der Erreger geeignet sind. Hierbei ist es essentiell auf eine ausreichende Sensitivität zu achten. Hierzu wurden die Nachweise auch auf nahe verwandten Arten durchgeführt, um zu zeigen, dass hier kein (falsch) positiver Nachweis erfolgt. Beispielweise ergab die Überprüfung der Spezifität der Primer für *R. solani*, dass diese nur ein Produkt mit der DNA des Isolates aus der Anastomose Gruppe 3 (AG-3) amplifizieren. Bei Verwendung der DNA eines Isolates aus der AG-5 wurde kein Produkt amplifiziert. Der Nachweis für *R. solani* ist somit bisher für die AG-3 realisiert. In ähnlicher Weise wurde die Spezifität der Primer für *P. ultimum* mit anderen *Pythium*-Arten (*P. debaryanum*, *P. arrhenomanes* und *P. phragmitis*) überprüft. Ausschließlich wies das Primerpaar qPyu92f/qPyu166r nur *P. ultimum* nach und kann somit als spezifisch angesehen werden. Gleichzeitig soll in diesem Ansatz die Reproduzierbarkeit des Verfahrens bestimmt werden. Der Pathogenitätstest verschiedener Krankheitserreger wurde auch optimiert und kann in den weiteren Versuchen verwendet werden.

Bei Versuchen zur Quantifizierung von *Diaporthe* spp. in einem parallelen Projekt zeigte sich, dass eine Quantifizierung aufgrund von Pilzsporen oder Myzelmengen unrealistisch ist. Stattdessen wurden dort Verdünnungsreihen genomischer DNA aus Reinkulturen der verschiedenen Pathogenen erstellt. Daraus ergeben sich dann entsprechende Standardgeraden. Mit Hilfe dieser Standardgeraden kann die Menge Pilz-DNA in den untersuchten Proben ermittelt werden. Für Pflanzenproben bietet sich dann eine „relative“ Quantifizierung an: Es wird parallel ebenfalls per Real-Time-PCR auch die Menge der pflanzlichen DNA bestimmt und dann die Schwere der Infektion in ng Pilz DNA / ng Pflanzen DNA angegeben. Aus diesen Standardgeraden lässt sich dann auch die Nachweisgrenze abschätzen; zur genauen Bestimmung der Nachweisgrenze müssen aber statistisch zahlreiche Proben nahe der Nachweisgrenze ausgewertet werden.

Theoretisch könnte man auch Sporen oder Myzel mit Pflanzenmaterial mischen und dann daraus DNA präparieren. (Das Problem bei der oben beschriebenen Vorgehensweise ist, dass bei geringen Mengen an Sporen oder Myzel die Effizienz der DNA-Präparation schlechter wird, was dann auch die Standardkurve beeinflusst. Dies spiegelt nicht korrekt wieder, was passiert, wenn man aus Pflanzenproben DNA präpariert.)

Für Bodenproben sollte die Nachweisgrenze separat bestimmt werden. Der Plan, den wir bis jetzt verfolgt haben, ist Bodenproben mit verschiedenen Sporenmengen zu inokulieren und so die Nachweisgrenze auszuloten.

5.4 JKI-RS

Die Anzahl der **Hülsen pro Pflanze** ist eines der wichtigsten Ertragsmerkmale der Sojabohne und eng mit dem Ertrag korreliert (Machado et al. 2017). Dabei scheint der Hülsenansatz auch die Ertragskomponente zu sein, die bei Kühlestress in der Blühphase am deutlichsten beeinflusst wird (Kurosaki et al. 2003). Über Ausfälle im Hülsenansatz verursacht durch Kühlestress in der Blüte berichten auch Hume und Jackson (1981) sowie Michailov et al. (1989). Bei inzwischen älteren Untersuchungen in Japan (Hokkaido) im Freiland wurde eine

stark negative Korrelation zwischen den Temperaturen im Juli zur Blüte der Sojabohnen und dem Kornertrag nachgewiesen (Horiguchi 1980).

Die im Rahmen des Projektes ‚Ausweitung des Sojaanbaus in Deutschland durch züchterische Anpassung sowie pflanzenbauliche und verarbeitungstechnische Optimierung‘ (FKZ 2811 NA 007) in 2011-2013 Jahren durchgeführten Gefäßversuche unter kontrollierten Bedingungen zeigten eine deutliche Variabilität des Hülsenansatzes von Sorten und Zuchtlinien in der Reaktion auf den Kühlestress. Der dabei entwickelte Test wurde jetzt genutzt, um nach einer Kreuzungsnachkommenschaft Merlin (kühletolerant im Hinblick auf den Hülsenansatz) x Sigalia (mittlere Kühlesensitivität) weitere 350 Akzessionen für eine GWAS zu charakterisieren mit dem Ziel, QTL für dieses Merkmal zu identifizieren.

Ein ähnlicher Test in der Klimakammer wird sowohl bei Kurosaki et al. (2003) als auch bei Cober et al. (2013) beschrieben, allerdings wurden bei letzterem die Nachttemperaturen in der Stressvariante sogar auf 5 °C abgesenkt. Analog zu Kurosaki et al. (2003) wird auch hier eine Stressdauer von mindestens 3 Wochen empfohlen, um die Genotypen zu differenzieren, allerdings wurden die Pflanzen erst nach einer Erholungsphase bonitiert.

Gefunden wurden in den hier vorgestellten Untersuchungen der Nachkommen der biparentalen Population ‚Merlin x Sigalia‘ in den Klimakammern ein Ranking der Eltern, welches dem der Vortests entsprach und eine breite Variabilität mit Transgression in der Nachkommenschaft für den Hülsenansatz. Heritabilitäten von 0,74 (Kontrolle) und 0,75 (Stress) zeigen, dass der Test für ein komplexes Merkmal vergleichsweise reproduzierbare Ergebnisse erbracht hat. Aus den Ergebnissen, dass sich die Heritabilität bei verringerter Anzahl von Wiederholungen nur im geringen Maße verschlechtert, lässt sich schließen, dass der Versuchsaufbau für ein mögliches Folgeprojekt modifiziert werden könnte. Es könnten dementsprechend bei geringerer Wiederholung mehr Populationen in der gleichen Zeit und bei gleichem Arbeitsaufwand getestet werden.

Es gibt inzwischen eine Reihe von Genomstudien, die bis zu 15 QTL für den Hülsenansatz pro Pflanze identifiziert haben (Sun et al. 2006; Vieira et al. 2006), im Folgenden wurden auch QTL für die Hülsenanzahl mit 1,2,3 und 4 enthaltenen Samen gefunden (Zhou et al. 2009). Aktuell ist ein Bericht veröffentlicht worden, in dem darüber hinaus auch QTL für die vertikale Verteilung der Hülsen am Stängel wie auch die Anzahl der Samen pro Hülse identifiziert wurden (Ning et al. 2018). Es ist aber (noch) keine Studie bekannt, in der QTL für den Hülsenansatz unter Kühlestress identifiziert wurden.

Mit Hilfe einer ‚bulk segregant analysis‘ wurden mit SSR-Markern aber 6 Chromosomenregionen identifiziert, die den Hülsenansatz unter Kühlestress potentiell beeinflussen (Cober et al. 2013). Die in diesem Projekt gefundenen QTL unterscheiden sich von den von (Cober et al. 2013) publizierten Regionen. Weiterhin legt die Überschneidung der signifikanten Chromosomenregionen unter beiden Bedingungen (auf Chromosom 7 und 13) nahe, dass es sich bei den gefundenen Genomabschnitten nicht allein um QTL für den Hülsenansatz bei Kühle handelt, sondern dass diese den Hülsenansatz unabhängig von den Temperaturbedingungen beeinflussen. Dies kann durch die Analyse der Varianzkomponenten in gewisser Weise bestätigt werden, da eine geringere ‚Genotyp x Treatment‘-Interaktion im Vergleich zur genotypischen Variation vorliegt. Aus züchterischer Sicht würde es also keinen großen Unterschied machen, im Umfang dieser Population unter Kontroll- oder Stressbedingungen zu selektieren. Um zu klären, ob diese Erkenntnisse sich auf weitere Populationen anwenden lassen, bedarf es weiterer Untersuchungen.

Wird die Sojabohne Kühlestress ausgesetzt, kommt es (auch oder auch primär) zu Veränderungen einer Reihe physiologischer Prozesse. Davon sind auch Merkmale wie der **Chlorophyllgehalt** betroffen, der unter Kühlestress sinkt (Balestrasse et al. 2010; Balko et al. 2014).

Erstere fanden bei 4 °C sogar einen Abfall um 57 %. Nach den Untersuchungen im o.g. Vorgängerprojekt, in dem der Chlorophyllgehalt indirekt als SPAD Werte erfasst wurde, wiesen diese Werte eine durchgängig positive, wenn auch nicht enge Korrelation zum Hülsenansatz auf und waren als Versuch der Nutzung indirekter Selektionskriterien in das Projekt einbezogen worden. Leider zeigten sie in der Nachkommenschaft zwar den erwarteten Abfall unter Stress, aber keine spezifische Reaktion der Linien, die in irgendeiner Form mit dem Hülsenansatz korreliert war. Eine Nutzung als indirektes Selektionskriterium ist damit ausgeschlossen. Dazu ist zu sagen, dass SPAD Messungen eine starke Beeinflussung durch Genotyp und Blattentwicklung aufweisen, über die schon Fritschi und Ray (2007) berichteten. Entsprechend konnten auch Vollmann et al. (2011) keinen direkten Zusammenhang zum Hülsenansatz herstellen.

Gefunden wurden in den hier vorgestellten Untersuchungen eine breite Variabilität im untersuchten Material - sowohl beim Hülsenansatz unter Kontroll-, als auch Kühlestressbedingungen - sowie signifikante Unterschiede in der Reaktion der Akzessionen auf den Stress. Heritabilitäten von 0,54 (Kontrolle) und 0,50 (Stress) zeigen, dass der Test für ein komplexes Merkmal selbst bei nur 2 Wiederholungen vergleichsweise reproduzierbare Ergebnisse erbracht hat.

Während für die Population Merlin x Sigalia QTL für die Anzahl an Hülsen und ein kühle-toleranter QTL gefunden wurde (Jähne et al. 2019) konnten dies in der GWAS-Analyse nicht bestätigt werden.

6. ANGABEN ZUM VORAUSSICHTLICHEN NUTZEN UND ZUR VERWERTBARKEIT DER ERGEBNISSE

(Wurden im Projekt praxisrelevante Ergebnisse erzielt? Falls nein, bitte begründen, falls ja: erläutern inwiefern diese Ergebnisse direkt praktisch anwendbar sind. Sofern praxisrelevante Ergebnisse gewonnen wurden, Erstellung eines Merkblatts zwecks Transfer dieser Ergebnisse in die Praxis)

6.1 IPSP

Es wurden keine Ergebnisse mit unmittelbarer Praxisrelevanz erzielt. Die in diesem Projektteil geleisteten Arbeiten sind der Züchtung neuer, angepassterer Sorten vorgelagert und besitzen entsprechenden „Grundlagenforschungscharakter“. Die erzielten Ergebnisse werden mittel- bis langfristig in Form von nützlicher genetischer Variation in der praktischen Pflanzenzüchtung ankommen und in der Folge auch durch die Verfügbarkeit neuer Sorten für die Landwirtschaft Praxisrelevanz entfalten.

Im AP1a wurde eine NoSQL-Datenbank entwickelt, die für die Speicherung der Projektdaten verwendet wurde. Diese Datenbank umfasst ein Datenbankschema, sowie Skripte zum Export und Import von Daten und ein Web-Interface. Die Struktur der Datenbank ist mit dem BreederAPI-Standard kompatibel. Die Projektdaten können über das Interface abgerufen und verwendet werden. Ein solches Datenbanksystem kann direkt für die praktische Züchtung verwendet werden, weil z. B. Stammbäume von Zuchtmaterial, welches unter der Verwendung von genetischen Ressourcen erzeugt wurden, speichert und für Abfragen verwendet. Dadurch ist es möglich, genetische Ressourcen, wie in diesem Arbeitspaket untersucht, schneller in Züchtungsprogramme einzubringen und sie mit BrAPI-kompatiblen Tools auszuwerten, so dass auch die im Projekt erhobenen Daten für die Züchtung zur Verfügung stehen.

Die in Arbeitspaketen AP1g-AP1i identifizierten Allele, Haplotypen und Marker-Merkmal-Assoziationen werden verwendet, um in der CC4EU und Genbanken geeignete Akzessionen zu identifizieren, die in Zusammenarbeit mit der LSA in ein Kreuzungsprogramm zur markergestützten Introgression nützlicher Allele mit Hilfe des Speed Breeding Systems eingebracht werden. Zusätzlich wurde eine Kooperation mit einem Züchtungsunternehmen vereinbart, um die Core Collection CC4EU an weiteren Standorten zu testen und vorteilhafte genetische Variation zu identifizieren.

6.2 LSA

Im Projekt wurden Sojastämme selektiert, die in manchen Merkmalen besser waren als die vorhandenen Sorten. Diese Stämme wurden und werden weiterhin an private Zuchtfirmen abgegeben, die die Stämme weiter prüfen und bei erfolgreichem Abschneiden als Sortenkandidaten in die Wertprüfung stellen, so dass sie anschließend in der landwirtschaftlichen Praxis ankommen.

Die QTL-Ergebnisse und die Ergebnisse zur genomischen Selektion wurden publiziert. Damit erhalten die privaten Pflanzenzüchter die Möglichkeit, die Ergebnisse zu nutzen und in ihre praktische Züchtungsarbeit zu integrieren.

6.3 IPM

Die Entwicklung neuartiger Strategien zum Schutz der Kulturpflanzen vor Krankheiten ist ein wesentlicher Beitrag zur Umsetzung neuer innovativer und übertragbarer Lösungen im Pflanzenschutzbereich. Aus den während des Vorhabens erzielten Ergebnissen werden schnelle Einzelnachweise wichtiger Sojapathogenen erarbeitet, die auf internationaler oder

nationaler Ebene in die entsprechenden Standardisierungsprozesse überführt werden können. Primäres Ziel dabei ist es, den ertragsrelevanten Erreger frühzeitig zu erfassen, so dass die Bekämpfungsmaßnahmen rechtzeitig angepasst werden können. Als zweiter praxisrelevanter Zweck der Arbeitspakete dient die Validierung der gefundenen Ergebnisse zur Erstellung der Empfehlungen und Kriterien für die Zertifizierung der entsprechenden Produkte (z. B. Saatgut). In einem an dieses Projekt angegliederten Promotionsverfahren wurde beispielweise ein qRT-PCR basierter quantitativer Nachweis für die Arten des *Diaporthe/Phomopsis*- Komplexes etabliert. Hierbei handelt es sich um verschiedene Saatgut-assoziierte Erreger, die zu erheblichen Verlusten während des Auflaufens der Kultur führen können. Es ist nicht geplant diesen Nachweis in das Multiplexverfahren zu integrieren. Vielmehr soll dieser Test unabhängig von den anderen Nachweisen weiter verfeinert werden und ultimativ in eine Methode zur Überprüfung der Saatgutqualität entwickelt werden.

6.4 JKI-RS

Die gefundenen QTL aus beiden Eltern zeigen, dass das Merkmal sehr komplex vererbt wird. Sie können im Zuchtprozess verwendet werden, um Eltern gezielt auszuwählen und positive Allele mit Bezug zur Kühletoleranz im Zuchtmaterial anzureichern.

Da die QTL in unabhängigem Material nicht bestätigt werden konnten, müssen weitere Untersuchungen zur Kühletoleranz durchgeführt werden. Ein praktischer Nutzen ist daher, auch unter Berücksichtigung der Dauer des Zuchtprozesses, erst in einem mittel- bis langfristigen Rahmen zu erwarten.

7 GEGENÜBERSTELLUNG DER URSPRÜNGLICH GEPLANTEN ZU DEN TATSÄCHLICH ERREICHTEN ZIELEN, HINWEISE AUF WEITERFÜHRENDE FRAGESTELLUNGEN

7.1 IPSP

Die Sammlung öffentlicher Daten zu verfügbaren genetischen Ressourcen der Sojabohne in einer Datenbank ist zum gegenwärtigen Zeitpunkt abgeschlossen. Mit *Soybase* (<https://www.soybase.org/>) steht außerdem ein öffentliches Verzeichnis zu den genetischen Grundlagen für publizierte Ertrags-, Reife-, Resistenz- und Qualitätsgene bzw. QTL und derer physikalischer Position im Genom zur Verfügung. Die Ziele der Arbeitspakete AP1a und AP1b sind damit erreicht worden.

Hinsichtlich der Ziele des Arbeitspakets AP1c wurde mit der Auswahl einer Kandidatengruppe (3.663 Akzessionen) mit potentieller Anpassung an mitteleuropäische Umweltverhältnisse und der darauf aufbauenden Zusammenstellung zweier Kernkollektionen von jeweils fünf bzw. zehn Prozent (d.h. 183 bzw. 366 Akzessionen) ein Grundstein für weitere Arbeiten zur Identifizierung züchterisch interessanter Akzessionen gelegt. Beide Kollektionen sind sowohl für die effiziente Suche nach abiotischer Anpassung als auch für andere Merkmale geeignet. Der mittelfristigen Auswahl der vielversprechendsten Kreuzungspartner für Kreuzungen mit modernem Zuchtmaterial muss nun jedoch eine ausführliche Phänotypisierung vorausgehen. Die genotypischen Differenzierungsanalysen und die Konsultation der *Soybase* Datensätze haben ergeben, dass Genombereiche die für abiotische Anpassung an mitteleuropäische Umweltbedingungen ausschlaggebend sind, sowohl in der Kandidatengruppe als auch in den Kernkollektionen angereichert sind. Bislang standen hier in der mitteleuropäischen Sojazüchtung die Reifegene E1 – E4 im Fokus (Kurasch et al. 2017). Neben diesen kartierten wir in den genetischen Ressourcen zahlreiche weitere interessante Genombereiche. Für das Feinmapping dieser Signale ist nun die Generierung von Sequenzdaten für das Genbankmaterial notwendig. Eine zusätzliche Resequenzierung von modernem Zuchtmaterial würde in der Folge das Allele-mining relevanter Genorte (d.h. Genorte für die modernes Zuchtmaterial keine positive Variation aufweist) und die Entwicklung funktionaler Marker begünstigen.

Im Bereich des Arbeitspakets AP1d wurden anstelle von genetischen Ressourcen moderne europäische Sorten mit dem SoySNP50K Array (Song et al. 2013) genotypisiert. Dies lag darin begründet, dass für die genetischen Ressourcen zu Projektbeginn schon hinreichend genotypische Daten vorlagen (Song et al. 2015) die für modernes Material fehlten. Die Verwendung des Arrays stellte hier das kostengünstigste Verfahren dar und ermöglichte es erstmals, moderne europäische Sojasorten mit der globalen Sojapopulation systematisch zu kontextualisieren. Analysen zur genetischen Struktur zeigten u.a., dass es per se keine asiatischen Sojapopulationen mit potentieller Umwelteignung für Mitteleuropa gibt, die nicht ohnehin schon in modernem Zuchtmaterial repräsentiert sind. Für die praktische Züchtung ist das eine gute Nachricht, macht nun aber die Untersuchung der genetischen Architektur von Umwelтанpassung nötig, um hier weitere Selektionsgewinne zu erzielen. Aufbauend auf unserer Arbeit wären dazu insbesondere genetische Differenzierungsstudien basierend auf den vorliegenden genotypischen als auch auf zu generierenden Resequenzierungsdaten zu nennen. Hier bietet sich jeweils der Vergleich der beiden nordchinesischen Ursprungspopulationen mit der südchinesischen an. Ein Vergleich der nordchinesischen Materialgruppen miteinander könnte komplementäre Anpassungsstrategien offenbaren. Ein vielversprechendes Modell stellt auch die japanische Ursprungspopulation dar, die mit einer vergleichsweise geringen genetischen Differenzierung das Reifegruppenportfolio und eine große Bandbreite diverser Anbaumumwelten abdeckt. Hier wären potentiell sogar

Assoziationsstudien mit Umweltdaten möglich, da sich aus den Passportdaten eine hinreichend kleinteilig, provinzielle Herkunft ableiten lässt. Beide Ansätze laufen darauf hinaus, die genetischen Grundlagen von bisher nicht in modernem Zuchtmaterial verankerter Umwelthanpassung zu identifizieren. Hier werden die nun vorhandenen genotypischen Daten (und ggf. zu erzeugende Resequenzierungsdaten) den direkten Abgleich mit Zuchtmaterial erlauben.

Arbeitspaket AP1e ist das einzige Arbeitspaket, in dem die Projektziele nicht erreicht wurden, da es von unserer Seite bislang noch zu keiner Auflage eines Prebreeding-Programmes gekommen ist. Dies ist jedoch der erforderlichen Sorgfalt geschuldet, die bei der Auswahl potenter Kreuzungspartner in Genbankmaterial notwendig ist. Mit den oben beschriebenen Arbeiten und insbesondere mit den Kernkollektionen sind hier jedoch Weichen gestellt worden und die gezielte Integrierung nützlicher Variation in existierende Züchtungspopulationen ist mittelfristig absehbar.

Das wichtigste Arbeitsziel von AP1f war die Sequenzierung und -analyse der CC4EU Core Collection. Dieses Ziel ist wegen der Corona-Pandemie mit einiger Verzögerung erreicht worden. Die Core Collection ist sequenziert, aber noch nicht vollumfänglich ausgewertet. Aufgrund der zwischenzeitlich publizierten GmHapMap (Torkamaneh et al. 2020) konnten die 50kSNP Marker der core collection für die Imputierung und weitere Analysen verwendet werden. Die Genomsequenzierungsdaten werden jedoch in Folgeanwendungen verwendet, um z.B. die Rolle von anderen genetischen Varianten als SNPs bei Marker-Merkmal-Assoziationen zu verwenden und die Sequenzierungsdaten für die Vorhersage des Phänotyps zu verwenden.

Die Feinkartierung von und Markerentwicklung von Reifegenen (AP1g) wurde teilweise erreicht. Die GmHapMap ermöglichte die Lokalisation von signifikanten Marker-Merkmal-Assoziationen mit einer hohen Präzision, so dass die Zahl der Kandidatengene, für die Marker entwickelt werden muss, sehr stark reduziert werden konnte. Auf die Entwicklung KASP-Markern wurde verzichtet, weil noch nicht feststand, welche Kreuzungen durchgeführt werden. Die Allele Mining Pipeline wurde etabliert (AP1h), so dass mit Hilfe der GmHapMap Kandidatengene für QTL identifiziert und weiter charakterisiert beziehungsweise unmittelbar für die Züchtung verwendet können.

Die Assoziationskartierung und Umweltassoziationsstudien mit europäischem (und asiatischem) Material wurden erfolgreich durchgeführt (AP1i), so dass neue Marker für die Selektion von Blühzeitpunktgenen und Reifegenen für zukünftige Kreuzungen und Züchtungsprogramme entwickelt werden können.

7.2 LSA

Die geplanten Ziele wurden weitgehend erreicht, die Grundlagen für eine genomische Selektion bei Sojabohnen wurde gelegt. Die genomische Selektion kann jedoch nur dann ihre volle Leistung ausprägen, wenn weitere Methoden in der Sortenentwicklung hinzugenommen werden. Im Zuchtprozess ist es notwendig, relativ homozygotes Material auf dem Feld zu prüfen, so dass die gewonnenen Ergebnisse auch auf die nächste Generation übertragbar sind. Um mehr Generationen pro Jahr und somit schneller homozygotes Material zu bekommen, nutzt die LSA einen Winterzuchtgarten in Costa Rica. Da jedoch extreme Wetterphänomene einige Male diesen Winterzuchtgarten zerstört oder beeinträchtigt haben und auch die Qualität des zurückgesendeten Saatguts nicht immer befriedigend ist, ist es sinnvoll, diesen um ein Indoor-Wachstumssystem mit LED Lichtern zu erweitern. Die Fortschritte in der LED-Technologie sind immens, so dass hier, ähnlich wie im Bereich der Genotypisierung, in der letzten Zeit enorme Kosteneinsparungen zu erzielen waren und daher die neuesten LED-Generationen genutzt werden kann. Unsere Arbeiten haben gezeigt, dass dadurch 4-6 Generationen in einem Jahr erreicht werden, für die Sojazüchtungsarbeiten der

LSA wurde ein Speed-Breeding Container entwickelt, der inzwischen voll in die Zucharbeiten integriert ist.

Unsere Untersuchungen zur Tofuausbeute und Tofuqualität haben gezeigt, dass signifikante genotypische Unterschiede im Zuchtmaterial vorhanden sind. Allerdings sind die Ortseffekte ebenfalls erheblich. Eventuell spielt die Aufnahme von verschiedensten Mineralstoff- und Spurenelementen der unterschiedlichen Böden, an denen die Sojabohnen für die Untersuchungen angebaut wurden, also das Ionom, eine entscheidende Rolle. Die Züchtung von Soja, insbesondere von Tofusorten, erfordert somit ein besseres Verständnis der genotypischen Variation des Ionoms der Sojabohnen, hin zu einer zielgerichteten Verbesserung. Die relativ neu entwickelte Röntgenfluoreszenzanalyse ist, da sie kostengünstig und fähig zum Hochdurchsatz ist, sehr gut dafür geeignet, weitergehende Erkenntnisse zum Ionom der Sojabohne zu liefern. Zahlreiche Untersuchungen wurden durchgeführt wobei klar wurde, dass vielfältige Beziehungen zwischen der Aufnahme der Mineralstoffe bestehen. Für die Elemente Magnesium, Phosphor, Schwefel, Kalium, Kalzium, Mangan, Eisen, Nickel, Kupfer, Zinn und Molybdän wurden Kalibrationen erstellt

Soja ist eine Kurztagespflanze und wird in unterschiedliche Reifegruppen eingeteilt. Während in warmen, südlichen Regionen in Deutschland, wie dem südlichen Rheintal, Reifegruppe I angebaut werden kann und zur Reife kommt, werden für weiter nördliche und kühlere Regionen Sorten der Reifegruppen 0 bis 0000 benötigt. Dies wird insbesondere deutlich, wenn die Sortenergebnisse unterschiedlicher Regionen, z.B. in Landessortenversuchen oder in Sortenversuchen, wie sie bei uns im Arbeitspaket 2h durchgeführt wurden, verglichen werden. So zeigte beispielsweise die Sorte Regina, die von der LSA gezüchtet wurde, im Mittel in Versuchen in Baden-Württemberg, Hessen und Rheinland-Pfalz im Jahr 2017 den Relativertrag 106 und war damit die ertragsstärkste Sorte. Die Relativerträge schwankten jedoch an den einzelnen Orten zwischen 99 und 116. In Bayern schwankten die Relativerträge zwischen 93 und 112, in Thüringen zwischen 95 und 127, in Sachsen-Anhalt zwischen 98 und 101. In Ostdeutschland wurde noch unterschieden zwischen Löss-Standorten und Diluvial-Standorten. Während Regina auf Löss-Standorten hohe Relativerträge mit 107 erzielte, zeigte sie auf Diluvial-Standorten nur Erträge von relativ 96.

Neben der Adaptation erfordert die Züchtung regional optimierter Sojasorten auch eine unterschiedliche Genetik bezüglich des Ertrags, auch wegen der substantiellen Genotyp-Umwelt-Interaktion. Dies bedeutet, dass im Unterschied zu anderen Kulturarten wie dem Weizen, bei Soja keine einzelne Sorte in ganz Deutschland mit höchstem Ertrag anbaubar ist. Um dem gerecht zu werden, sollte zukünftig eine regionale Selektion etabliert werden. Dies bedeutet, dass Merkmalsdaten aus regionalen Sortenversuchen genutzt werden, um Selektionsmodelle für die einzelnen Zielregionen zu etablieren. Unsere grundlegenden Arbeiten zur genomischen Selektion haben gezeigt, dass eine genomische Selektion einen gewissen Verwandtschaftsgrad der zu selektierenden Linien benötigt. Dies ist für eine phänomische Selektion nicht der Fall. Daher wird diese zukünftig in der Sojazüchtung eingesetzt, um eine regionale Selektion durchzuführen.

7.3 IPM

Die Analyse verschiedener Proben mit unterschiedlicher Herkunft und Zusammensetzung wird im Wesentlichen die frühzeitige Diagnose von bestehenden Problemen im Feld vereinfachen. Großes Interesse besteht bei Pflanzenschutzberatern und Landwirten auch an einer Beprobung des Saatgutes und des Bodens, möglichst noch vor der Aussaat. In diesem Arbeitspaket wurden diesbezüglich praxisrelevante Ergebnisse teilweise erzielt. Die Grundlagen des Nachweisverfahrens unterschiedlicher Pathogenen wurden etabliert. Einzelne gefährliche Krankheitserreger können mit Hilfe des etablierten Systems nachgewiesen werden. Das Verfahren soll allerdings noch als Multiplex-System und an

verschiedenen Probentypen validiert werden. Für das Multiplex-System muss nochmals die Sensitivität getestet werden. Diese Sensitivitätsprüfung soll mit dotierten Probenmatrices erfolgen. Dazu kann z.B. das Saatgut mit einer bestimmten Anzahl Pilzsporen bzw. Bakterienzellen versetzt werden, wobei sich die verwendeten Mengen in ihrer Größenordnung an den ermittelten Nachweisgrenzen orientieren. Es ist geplant auch den zu erarbeitenden Nematodennachweis zu integrieren, allerdings muss hier zunächst der Einzelnachweis etabliert werden. Eine Gegenüberstellung der ursprünglich geplanten und tatsächlich erreichten Ziele erfolgt in Tabelle 21: Gegenüberstellung der Vorhaben und deren Ergebnis.

Tabelle 21: Gegenüberstellung der Vorhaben und deren Ergebnis.

Aktivität	Ergebnis
Etablierung eines quantitativen Nachweisverfahrens für verschiedene Pathogene	Das Nachweisverfahren wurde erfolgreich für Ss, Fs und Psg etabliert.
Etablierung eines quantitativen Nachweisverfahrens für verschiedene Probentypen	Die Nachweisverfahren wurden erfolgreich für Blatt-, Stängel- und Saatgutproben etabliert.
Etablierung einer quantitativen Methodik zur Erfassung von Nematoden	Es erfolgte eine Probennahme und Auszählung von Nematoden in Bodenproben für konventionelle und ökologische Betriebe.
Etablierung des Multiplex-Systems auf verschiedene Proben	Duplex-Nachweise wurden erfolgreich für Ss und Fs, sowie Ss und Psg etabliert.
Testung verschiedener Proben	Das Duplex-System für Ss und Fs wurde zur Testung von Saatgutproben verwendet.
Publikationen	Es wurden Tagungsbeiträge veröffentlicht.

7.4 JKI-RS

Im Teilprojekt 4 wurden alle ursprünglich geplanten Ziele erreicht. Es wurden zusätzliche Arbeiten aufgenommen: (1) eine 5. Wiederholung des Gefäßversuches unter kontrollierten Bedingungen und (2) ein Anbau der Kreuzungsnachkommenschaft in 2 Wiederholungen im Feld 2016 zur phänotypischen Charakterisierung der biparentalen Kreuzungspopulation. Die Phänotypisierungsdaten wurden den Projektpartnern zur weiteren Verrechnung zur Verfügung gestellt und die QTL Analysen erfolgreich durchgeführt. In der Verlängerung des Projektes wurden die ursprünglich geplanten Ziele insofern nicht erreicht, als auf Grund der Saatgutqualität nicht alle Akzessionen keimten bzw. gesunde Jungpflanzen hervorbrachten. In der Konsequenz konnten nur 347 von 400 Akzessionen in die Analyse einfließen. Zusätzlich wurde in 2020 für TP 1 ein Feldversuch zur Charakterisierung einer core collection der amerikanischen Genbank unter norddeutschen Bedingungen durchgeführt. Die Daten wurden dem Projektpartner zur Verfügung gestellt.

8 ZUSAMMENFASSUNG

IPM

Zur Einleitung effizienter Pflanzenschutzmaßnahmen ist es essentiell, über die Art und Menge möglicherweise vorhandener Pathogene rechtzeitig Aufschluss zu gewinnen. Im Zuge der Forschungsarbeiten wurden Einzelnachweise basierend auf TaqMan Real-time-PCR entwickelt, die auch eine Erfassung bestimmter Pathogene ermöglichen sollen. Der Einzelnachweis zumindest für die drei Hauptpathogene (Ss, Fs und Psg) wurde etabliert. Die optimale Annealingtemperatur der Primerpaare für die DNA-Amplifikation aller o. g. Pathogenen lag bei 60,0 °C.

Zudem wurde die Pathogenität der angegebenen Krankheitserreger geprüft und das Inokulationsverfahren für die weiteren Untersuchungen etabliert. Hierzu wurden die Nachweise auch auf nahe verwandten Arten durchgeführt. Die Überprüfung der Spezifität der Primer- und Sonden-Kombinationen ergab, dass diese keine nicht-Ziel-DNA nachweisen.

Nach Etablierung der Einzelnachweise wurden Duplex-Nachweise für Ss und Fs, sowie Ss und Psg etabliert. Diese Nachweise sollen dann zur frühzeitigen Detektion von Erregern in Soja-Feldern verwendet werden.

IPSP

Genbanken enthalten Tausende von Sojaakzessionen, die für eine Erweiterung der genetischen Diversität bei der Züchtung von klimaangepassten Sojasorten verwendet werden können. Im Rahmen des Projekts konnten wir durch Kombination von Klimaparametern, imputierten Genomsequenzen und Phänotypisierungsdaten in Genbankmaterial, einer daraus abgeleiteten Kernkollektion sowie europäischen Sorten bekannte und neue Gene für den Blühzeitpunkt und die Reife identifizieren. Wir konnten auch zeigen, dass viele dieser Gene in epistatischen Interaktionen stehen. Zusammengefasst ermöglichen unsere Arbeiten das effiziente Design von molekularen Markern, um mit ihrer Hilfe vorteilhafte Allele für Anpassungsmerkmale effizient in aktuelles Zuchtmaterial einzukreuzen.

LSA

Die Pflanzenzüchtung spielt eine zentrale Rolle bei der Etablierung von Soja als Kulturpflanze für Mitteleuropa, da die Anpassung mit agronomischen Eigenschaften und hohem Ertrag kombiniert werden muss. Wir fanden heraus, dass selbst angepasste Sorten für Reifegene variieren und somit eine starke Segregation für die Reife in ihren Nachkommen aufweisen. Dies ist von Vorteil, da es die Züchtung von sehr frühen Linien ermöglicht. Eine bessere Kenntnis der Auswirkungen der Reifegene in mitteleuropäischem Sojazuchtmaterial wird diese Klassifizierung unterstützen und wird zukünftig auch eine markergestützte Vorauswahl von Linien ermöglichen, die in unterschiedlichen Reifezonen angebaut werden können. Dies könnte ein Jahr lang Feldversuche einsparen, die in unserem Fall erforderlich waren, um die Reife des damals nicht ausgewählten Materials zu bestimmen. Unsere Ergebnisse zeigen deutlich, dass hochertragreiche Linien gezüchtet werden können, die es ermöglichen, die Sojaanbauregion weiter nördlich zu erweitern. Darüber hinaus bieten die gewonnenen Varianzkomponenten, Erbanlagen und Merkmalskorrelationen eine solide Grundlage für die Gestaltung von Soja-Zuchtprogrammen. Die Untersuchungen zeigen, dass eine genomische Selektion als zusätzliches Instrument der Sojazüchtung in Zukunft eingesetzt werden kann. Die Tofu-Labormethode erwies sich als wertvolles Instrument, um vielversprechende Nachkommen in Sojabohnenzüchtungsprogrammen auf ihre Eignung für die Sojamilch- und

Tofuproduktion zu testen. Die bewerteten Tofumerkmale zeigten alle mittlere bis hohe Heritabilitäten und veranschaulichten das Potenzial für eine genetische Verbesserung dieser Merkmale. Es wurde keine signifikante negative Korrelation zwischen agronomischen und Tofueigenschaften beobachtet, so dass die damit einhergehende Verbesserung der agronomischen und Tofuqualitätsmerkmale möglich erscheint. Wir konnten auch ein unterschiedliches Einweichverhalten zwischen drei Familien beobachten, was helfen könnte, Linien für Tofuproduktionsmethoden zu klassifizieren und zu identifizieren, die unterschiedliche Einweichzeiten erfordern. Aus den Zusammenhängen zwischen Proteingehalt und Tofueigenschaften lässt sich kein klares Bild ziehen, da der Proteingehalt positiv mit den Tofuqualitätsmerkmalen in der einen Familie korreliert war, in der anderen aber nicht. Dies zeigt jedoch die Notwendigkeit, die Tofu-Eigenschaften durch die Herstellung von Tofu zu testen, anstatt den Proteingehalt als Indikator für die Eignung eines Genotyps als Lebensmittelsorte zu verwenden. Weitere Untersuchungen zum Genotyp-Umwelt-Effekt, der zu einer variablen Tofufestigkeit führt, würden es dem Landwirt und dem Tofuhersteller ermöglichen, dem Verbraucher eine konstant hohe oder sogar verbesserte Tofuqualität zur Verfügung zu stellen. Unsere Ergebnisse deuten auf eine komplexe genetische Architektur hin, die den Tofumerkmalen zugrunde liegt, mit Ausnahme von vielleicht den Einweichmerkmalen, die durch einige wichtige QTL kontrolliert werden können. Daher erscheint die markergestützte Selektion als keine praktikable Option zur Verbesserung der Tofueigenschaften. Unsere Ergebnisse zeigten hohe Genauigkeiten der genomweiten Vorhersage für die Proteinkonzentration und das Tausendkorngewicht sowie für die Einweichmerkmale, aber mittlere bis niedrige Vorhersagegenauigkeiten für die Tofu-Eigenschaften. Folglich liegt die Verbesserung des Phänotypisierungsprozesses dieser Merkmale im Mittelpunkt ihrer zukünftigen Verbesserung. Während der Durchsatz eines Laboratoriums im Labormaßstab begrenzt bleiben wird, werden verbesserte phänotypische Daten wahrscheinlich zu etwas höheren Vorhersagegenauigkeiten führen. Dann erscheint eine Kombination aus Phänotypisierung und genomischer Selektion vielversprechend, um die Identifizierung von Sojagenotypen mit verbesserter Tofu-Qualität zu unterstützen.

Der im Vorhaben entwickelte Speed-Breeding Containers stellt eine wesentliche Verbesserung im Zuchtprozess dar. Die Anzahl erfolgreicher Kreuzungen wurde deutlich erhöht und die Erstellung homozygoter Linien konnte beschleunigt werden. Die Untersuchungen zum Ionon der Sojabohne zeigten erste Zusammenhänge zwischen der Aufnahme unterschiedlicher Mineralstoffe auf, in diesem neuen Feld besteht jedoch noch ein erheblicher Forschungsbedarf.

Ganz neu wurden Arbeiten zur phänomischen Selektion in den Zuchtprozess mit aufgenommen. An grundlegenden Arbeiten konnten wir zeigen, dass die phänomische Selektion der genomischen Selektion teilweise überlegen ist. Da die phänomische Selektion einfacher und kostengünstiger durchzuführen ist, wird diese weiter in die Zuchtarbeiten integriert.

Die Zuchtaktivitäten insgesamt waren erfolgreich, es konnten bereits erste Linien an die private Pflanzenzüchtung abgegeben werden, die derzeit in weiteren Prüfungen ermitteln, ob die Linien als Sorten zugelassen werden können.

JKI-RS

Sojabohnen stellen hohe Anforderungen an Temperatur und Wasserversorgung. Geeignete Sorten mit hohem Proteinertrag sind bislang in Deutschland nur für klimatisch günstige Lagen verfügbar. Gerade in kühleren Lagen, wie z.B. in Norddeutschland, ist die Gefahr von Ertragseinbußen auf Grund von Kühlstress groß.

Ziel dieses Arbeitspakets war es, 120 Linien einer vorhandenen Population, die in die genomische Selektion eingeht (Projektteil LSA) hinsichtlich ihres Hülsenansatzes nach einer Kühlestressphase zu charakterisieren, um Marker für die Kühletoleranz zu identifizieren.

Die Heritabilität des Hülsenansatzes als Maß für die genetische Stabilität fiel mit 0,74 für die Kontrolle und 0,75 für den unter Kühlestress vergleichsweise hoch aus. Das ist indirekt ein Zeichen dafür, dass das Testsystem reproduzierbare Ergebnisse ermöglicht hat. Mittels eines ‚Genotype-by-sequencing‘ (GBS) -ansatzes und durch die anschließende QTL-Analyse der SNP-Marker auf Einzelpflanzenbasis wurden für die ‚Merlin x Sigalia‘-Population 3 QTL für den Hülsenansatz unter Kontroll- (Chromosom 4, 7, 13), sowie 2 QTL unter Stressbedingungen (Chromosom 7, 13) gefunden und durch Kreuzvalidierungen bestätigt. Dabei zeigte sich, dass Allele von beiden Eltern einen positiven Effekt auf den Hülsenansatz haben und nicht nur Allele des toleranten Merlin-Elters. Dies untermauert zusätzlich zu dem Fund, dass es nicht den einen großen QTL gibt, den quantitativen Charakter des Merkmals.

Ziel im zweiten Projektteil des Arbeitspakets war es, 400 Akzessionen hinsichtlich ihres Hülsenansatzes nach einer Kühlestressphase zu charakterisieren, um mittels einer GWAS Marker für die Kühletoleranz zu identifizieren und ggf. die bei der Charakterisierung der Nachkommenschaft aus der Kreuzung Merlin x Sigalia gefundenen Marker aus dem ersten Teil des Projektes zu validieren.

Wie erwartet ergab sich aus den Tests ein signifikanter Effekt der Akzession, der Variante mit einer deutlichen Reduktion des Hülsenansatzes unter Stress, sowie signifikante Unterschiede in der Reaktion der Akzessionen auf den Stress als Grundlage einer Selektion. Die Heritabilität des Hülsenansatzes als Maß für die genetische Stabilität fiel mit 0,50 für die Kontrolle und 0,53 für den Hülsenansatz unter Kühlestress nicht mehr so hoch aus, wie im ersten Teil des Projektes, basiert aber auch nur auf 2 statt 5 Wiederholungen. Das ist indirekt ein Zeichen dafür, dass das Testsystem auch mit 2 Wiederholungen noch reproduzierbare Ergebnisse ermöglicht hat. Die GWAS-Analyse ergab keine signifikanten QTL, auch die QTL der ‚Merlin x Sigalia‘-Population waren in der GWS-Analyse nicht mehr signifikant.

9 REFERENCES

- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19:1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Aziadekey M, Schapaugh WT, Herald TJ (2002) Genotype by environment interaction for soymilk and tofu quality. *J Food Quality* 25:243–259. <https://doi.org/10.1111/j.1745-4557.2002.tb01022.x>
- Balestrasse KB, Tomaro ML, Battle A, Noriega GO (2010) The role of 5-aminolevulinic acid in the response to cold stress in soybean plants. *Phytochemistry* 71:2038–2045. <https://doi.org/10.1016/j.phytochem.2010.07.012>
- Balko C, Hahn V, Ordon F (2014) Kühletoleranz bei der Sojabohne (*Glycine max* (L.) Merr.)–Voraussetzung für die Ausweitung des Sojaanbaus in Deutschland. 1,27 MB / *Journal für Kulturpflanzen* 66(11) 2014. <https://doi.org/10.5073/JFK.2014.11.02>
- Bandillo N, Jarquin D, Song Q, Nelson R, Cregan P, Specht J, Lorenz A (2015) A Population Structure and Genome-Wide Association Analysis on the USDA Soybean Germplasm Collection. *Plant Genome* 8:0. <https://doi.org/10.3835/plantgenome2015.04.0024>
- Bari A, Street K, Mackay M, Endresen DTF, Pauw E de, Amri A (2012) Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. *Genet Resour Crop Evol* 59:1465–1481. <https://doi.org/10.1007/s10722-011-9775-5>
- Baxter I (2010) Ionomics: The functional genomics of elements. *Brief Funct Genomics* 9:149–156. <https://doi.org/10.1093/bfpg/elp055>
- Bechtold K, Hartung K, Müller-Belami K, Würfel T (2017) Kurzinformatio Sommerleguminosen, Sojabohnen Ernte 2017
- Becker H (2011) *Pflanzenzüchtung*, 2nd edn. utb-studi-e-book, vol 1744. UTB GmbH; Ulmer, Stuttgart
- Benassi VdT, Yamashita F, Prudencio SH (2011) A statistical approach to define some tofu processing conditions. *Ciênc. Tecnol. Aliment.* 31:897–904. <https://doi.org/10.1590/S0101-20612011000400011>
- Beukelaer H de, Smýkal P, Davenport GF, Fack V (2012) Core Hunter II: fast core subset selection based on multiple genetic diversity measures using Mixed Replica search. *BMC Bioinformatics* 13:312. <https://doi.org/10.1186/1471-2105-13-312>
- Beukelaer HD, Davenport GF, Fack V (2017) Multi-Purpose Core Subset Selection. www.corehunter.org
- Bhardwaj HL, Bhagsari AS, Joshi JM, Rangappa M, Sapra VT, Rao MSS (1999) Yield and Quality of Soymilk and Tofu Made from Soybean Genotypes Grown at Four Locations. *Crop Science* 39:401. <https://doi.org/10.2135/cropsci1999.0011183X0039000200017x>
- Bohra A (2013) Emerging paradigms in genomics-based crop improvement. *ScientificWorldJournal* 2013:585467. <https://doi.org/10.1155/2013/585467>
- Boyrie L, Moreau C, Frugier F, Jacquet C, Bonhomme M (2021) A linkage disequilibrium-based statistical test for Genome-Wide Epistatic Selection Scans in structured populations. *Heredity (Edinb)* 126:77–91. <https://doi.org/10.1038/s41437-020-0349-1>
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Brauner PC, Müller D, Molenaar WS, Melchinger AE (2020) Genomic prediction with multiple biparental families. *Theor Appl Genet* 133:133–147. <https://doi.org/10.1007/s00122-019-03445-7>
- Brown DM, Bootsma A (1993) *Crop Heat Units for Corn and Other Warm Season Crops in Ontario*:4pp

- Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker RC (1997) Mapping QTL for Seed Protein and Oil Content in Eight Soybean Populations. *Crop Science* 37:370. <https://doi.org/10.2135/cropsci1997.0011183X003700020011x>
- Cai T, Chang K, Shih M, Hou H, Ji M (1997) Comparison of bench and production scale methods for making soymilk and tofu from 13 soybean varieties. *Food Research International* 30:659–668. [https://doi.org/10.1016/S0963-9969\(98\)00032-5](https://doi.org/10.1016/S0963-9969(98)00032-5)
- Cai T, Chang K-C (1999) Processing Effect on Soybean Storage Proteins and Their Relationship with Tofu Quality. *J. Agric. Food Chem.* 47:720–727. <https://doi.org/10.1021/jf980571z>
- Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Müller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat Genet* 43:956–963. <https://doi.org/10.1038/ng.911>
- Chapman A, Pantalone VR, Ustun A, Allen FL, Landau-Ellis D, Trigiano RN, Gresshoff PM (2003) Quantitative trait loci for agronomic and seed quality traits in an F2 and F4:6 soybean population. *Euphytica* 129:387–393. <https://doi.org/10.1023/A:1022282726117>
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Braz. J. Plant Physiol.* 138:963–971
- Cober ER, Molnar SJ, Rai S, Soper JF, Voldeng HD (2013) Selection for Cold Tolerance during Flowering in Short-Season Soybean. *Crop Science* 53:1356. <https://doi.org/10.2135/cropsci2012.08.0487>
- Cuevas J, Montesinos-López O, Juliana P, Guzmán C, Pérez-Rodríguez P, González-Bucio J, Burgueño J, Montesinos-López A, Crossa J (2019) Deep Kernel for Genomic and Near Infrared Predictions in Multi-environment Breeding Trials. *G3 (Bethesda)* 9:2913–2924. <https://doi.org/10.1534/g3.119.400493>
- Cullen DW, Toth IK, Boonham N, Walsh K, Barker I, Lees AK (2007) Development and Validation of Conventional and Quantitative Polymerase Chain Reaction Assays for the Detection of Storage Rot Potato Pathogens, *Phytophthora erythroseptica*, *Pythium ultimum* and *Phoma foveata*. *J Phytopathol* 155:309–315. <https://doi.org/10.1111/j.1439-0434.2007.01233.x>
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R (2011) The variant call format and VCFtools. *Bioinformatics* 27:2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. *J Anim Breed Genet* 124:331–341. <https://doi.org/10.1111/j.1439-0388.2007.00701.x>
- Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. *Theor Appl Genet* 83:608–612. <https://doi.org/10.1007/BF00226905>
- Elhady A, Heuer H, Hallmann J (2018) Plant parasitic nematodes on soybean in expanding production areas of temperate regions. *J Plant Dis Prot* 125:567–576. <https://doi.org/10.1007/s41348-018-0188-y>
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Endelman JB (2011) Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome Journal* 4:250
- Fehr WR, Hoeck JA, Johnson SL, Murphy PA, Nott JD, Padilla GI, Welke GA (2003) Genotype and Environment Influence on Protein Components of Soybean. *Crop Science* 43:511. <https://doi.org/10.2135/cropsci2003.0511>
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol* 37:4302–4315. <https://doi.org/10.1002/joc.5086>

- Fliege CE, Ward RA, Vogel P, Nguyen H, Quach T, Guo M, Viana JPG, Santos LB, Specht JE, Clemente TE, Hudson ME, Diers BW (2022) Fine mapping and cloning of the major seed protein quantitative trait loci on soybean chromosome 20. *The Plant Journal*. <https://doi.org/10.1111/tpj.15658>
- Frankel O (1984) *Genetic Perspectives of Germplasm Conservation*:161–170
- Franzaring J, Fangmeier A, Schlosser S, Hahn V (2019) Cadmium concentrations in German soybeans are elevated in conurbations and in regions dominated by mining and the metal industry. *J. Sci. Food Agric.* 99:3711–3715. <https://doi.org/10.1002/jsfa.9548>
- Fritschi FB, Ray JD (2007) Soybean leaf nitrogen, chlorophyll content, and chlorophyll a/b ratio. *Photosynthetica* 45:92–98. <https://doi.org/10.1007/s11099-007-0014-4>
- Funatsuki H, Kurosaki H, Murakami T, Matsuba S, Kawaguchi K, Yumoto S, Sato Y (2003) Deficiency of a cytosolic ascorbate peroxidase associated with chilling tolerance in soybean. *Theor Appl Genet* 106:494–502. <https://doi.org/10.1007/s00122-002-1108-9>
- Gai J, Wang Y, Wu X, Chen S (2007) A comparative study on segregation analysis and QTL mapping of quantitative traits in plants—with a case in soybean. *Front. Agric. China* 1:1–7. <https://doi.org/10.1007/s11703-007-0001-3>
- Galán RJ, Bernal-Vasquez A-M, Jebsen C, Piepho H-P, Thorwarth P, Steffan P, Gordillo A, Miedaner T (2020) Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. *Theor Appl Genet* 133:3001–3015. <https://doi.org/10.1007/s00122-020-03651-8>
- Gass T, Schori A, Fossati A, Soldati A, Stamp P (1996) Cold tolerance of soybean (*Glycine max* (L.) Merr.) during the reproductive phase. *European Journal of Agronomy* 5:71–88. [https://doi.org/10.1016/S1161-0301\(96\)02011-4](https://doi.org/10.1016/S1161-0301(96)02011-4)
- Gautier M (2015) Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics* 201:1555–1579. <https://doi.org/10.1534/genetics.115.181453>
- Gawenda I, Thorwarth P, Günther T, Ordon F, Schmid KJ (2015) Genome-wide association studies in elite varieties of German winter barley using single-marker and haplotype-based methods. *Plant Breed* 134:28–39. <https://doi.org/10.1111/pbr.12237>
- Gilmour A, Cullis B, Gogel B, Thompson R (2009) ASReml user guide release 3.0. VSN International Ltd Hemel Hempstead
- Gizlice Z, Carter TE, Burton JW (1994) Genetic Base for North American Public Soybean Cultivars Released between 1947 and 1988. *Crop Science* 34:1143. <https://doi.org/10.2135/cropsci1994.0011183X003400050001x>
- Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346. <https://doi.org/10.1371/journal.pone.0090346>
- Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res* 38:D843-6. <https://doi.org/10.1093/nar/gkp798>
- Günther T, Coop G (2013) Robust identification of local adaptation from allele frequencies. *Genetics* 195:205–220. <https://doi.org/10.1534/genetics.113.152462>
- Günther T, Schmid KJ (2011) Improved haplotype-based detection of ongoing selective sweeps towards an application in *Arabidopsis thaliana*. *BMC Res Notes* 4:232. <https://doi.org/10.1186/1756-0500-4-232>
- Hahn V, Würschum T (2014) Molecular genetic characterization of Central European soybean breeding germplasm. *Plant Breed* 133:748–755. <https://doi.org/10.1111/pbr.12212>
- Hallauer AR, Miranda JB (1981) *Quantitative genetics in maize breeding*. Iowa State University Press, Ames

- Han Y, Li D, Zhu D, Li H, Li X, Teng W, Li W (2012) QTL analysis of soybean seed weight across multi-genetic backgrounds and environments. *Theor Appl Genet* 125:671–683. <https://doi.org/10.1007/s00122-012-1859-x>
- Hang Vu TT, Lawn RJ, Bielig LM, Molnar SJ, Xia L, Kilian A (2012) Development and initial evaluation of diversity array technology for soybean and mungbean. *Euphytica* 186:741–754. <https://doi.org/10.1007/s10681-011-0573-0>
- Haupt M, Schmid K (2020) Combining focused identification of germplasm and core collection strategies to identify genebank accessions for central European soybean breeding. *Plant Cell Environ* 43:1421–1436. <https://doi.org/10.1111/pce.13761>
- Haupt M, Schmid K (2022) Using landscape genomics to infer genomic regions involved in environmental adaptation of soybean genebank accessions. *bioRxiv:2022.02.18.480989*. <https://doi.org/10.1101/2022.02.18.480989>
- Herten K, Hestand MS, Vermeesch JR, van Houdt JKJ (2015) GBSX: a toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinformatics* 16. <https://doi.org/10.1186/s12859-015-0514-3>
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol* 25:1965–1978. <https://doi.org/10.1002/joc.1276>
- Hoeck JA, Fehr, W. R., Shoemaker RC, Welke. G. A., Johnson SL, Cianzio SR (2003) Molecular Marker Analysis of Seed Size in Soybean. *Crop Science* 43:68
- Horiguchi I (1980) The damage forecast of rice and soybean yield caused by cool summer temperature. *Memoirs of the Faculty of Agriculture Hokkaido University* 12:222–229
- Htwe NMPS, Luo Z-Q, Jin L-G, Nadon B, Wang K-J, Qiu L-J (2015) Functional marker development of miR1511-InDel and allelic diversity within the genus *Glycine*. *BMC Genomics* 16:467. <https://doi.org/10.1186/s12864-015-1665-3>
- Huang M, Kaletunc G, St. Martin S, Feller M, McHale LK (2014) Correlations of seed traits with tofu texture in 48 soybean cultivars and breeding lines. *Plant Breed* 133:67–73. <https://doi.org/10.1111/pbr.12122>
- Hui W, Gel YR, Gastwirth JL (2008) lawstat : An R Package for Law, Public Policy and Biostatistics. *J. Stat. Soft.* 28. <https://doi.org/10.18637/jss.v028.i03>
- Hume D, Jackson A (1981) Pod Formation in Soybeans at Low Temperatures. *Crop Science* 21:933–937
- Jähne F, Balko C, Hahn V, Würschum T, Leiser WL (2019) Cold stress tolerance of soybeans during flowering: QTL mapping and efficient selection strategies under controlled conditions. *Plant Breed* 138:708–720. <https://doi.org/10.1111/pbr.12734>
- Jähne F, Hahn V, Würschum T, Leiser WL (2020) Speed breeding short-day crops by LED-controlled light schemes. *Theor Appl Genet* 133:2335–2342. <https://doi.org/10.1007/s00122-020-03601-4>
- James AT, Yang A (2014) Influence of globulin subunit composition of soybean proteins on silken tofu quality. 2. Absence of 11SA4 improves the effect of protein content on tofu hardness. *Crop Pasture Sci.* 65:268. <https://doi.org/10.1071/CP13399>
- Jarquín 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 (Bethesda)* 6:2329–2341. <https://doi.org/10.1534/g3.116.031443>
- Jauregui LM, Chen P, Scaboo AM (2011) Heritability and correlations among food-grade traits in soybean. *Plant Breed* 130:647–652. <https://doi.org/10.1111/j.1439-0523.2011.01887.x>
- Jose J. Marulanda, Albrecht E. Melchinger, Tobias Würschum (2015) Genomic selection in biparental populations: assessment of parameters for optimum estimation set design. *Plant Breeding* 134:623–630. <https://doi.org/10.1111/pbr.12317>

- Kato S, Sayama T, Fujii K, Yumoto S, Kono Y, Hwang T-Y, Kikuchi A, Takada Y, Tanaka Y, Shiraiwa T, Ishimoto M (2014) A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. *Theor Appl Genet* 127:1365–1374. <https://doi.org/10.1007/s00122-014-2304-0>
- Kim Y, Wicker L (2005) Soybean cultivars impact quality and function of soymilk and tofu. *J. Sci. Food Agric.* 85:2514–2518. <https://doi.org/10.1002/jsfa.2287>
- Kim TG, Knudsen GR (2008) Quantitative real-time PCR effectively detects and quantifies colonization of sclerotia of *Sclerotinia sclerotiorum* by *Trichoderma* spp. *Applied Soil Ecology* 40:100–108. <https://doi.org/10.1016/j.apsoil.2008.03.013>
- Klasen JR, Barbez E, Meier L, Meinshausen N, Bühlmann P, Koornneef M, Busch W, Schneeberger K (2016) A multi-marker association method for genome-wide association studies without the need for population structure correction. *Nat Commun* 7:13299. <https://doi.org/10.1038/ncomms13299>
- Krause MR, González-Pérez L, Crossa J, Pérez-Rodríguez P, Montesinos-López O, Singh RP, Dreisigacker S, Poland J, Rutkoski J, Sorrells M, Gore MA, Mondal S (2019) Hyperspectral Reflectance-Derived Relationship Matrices for Genomic Prediction of Grain Yield in Wheat. *G3 (Bethesda)* 9:1231–1247. <https://doi.org/10.1534/g3.118.200856>
- Kulkarni KP, Asekova S, Lee D-H, Bilyeu K, Song JT, Lee J-D (2017) Mapping QTLs for 100-seed weight in an interspecific soybean cross of Williams 82 (*Glycine max*) and PI 366121 (*Glycine soja*). *Crop Pasture Sci.* 68:148. <https://doi.org/10.1071/CP16246>
- Kurasch AK, Hahn V, Leiser WL, Starck N, Würschum T (2017) Phenotypic Analysis of Major Agronomic Traits in 1008 RILs from a Diallel of Early European Soybean Varieties. *Crop Sci.* 57:726–738. <https://doi.org/10.2135/cropsci2016.05.0318>
- Kurasch AK, Hahn V, Miersch M, Bachteler K, Würschum T (2018a) Analysis of tofu-related traits by a bench-scale tofu production method and their relationship with agronomic traits in European soybean. *Plant Breed* 137:271–282. <https://doi.org/10.1111/pbr.12581>
- Kurasch AK, Leiser WL, Bachteler K, Miersch M, Hahn V, Würschum T (2018b) Evaluation of the genetic architecture of tofu traits in soybean towards genomics-assisted breeding. *Plant Breed* 137:873–882. <https://doi.org/10.1111/pbr.12651>
- Kuroda Y, Kaga A, Tomooka N, Yano H, Takada Y, Kato S, Vaughan D (2013) QTL affecting fitness of hybrids between wild and cultivated soybeans in experimental fields. *Ecol Evol* 3:2150–2168. <https://doi.org/10.1002/ece3.606>
- Kurosaki H, Yumoto S, Matsukawa I (2003) Pod Setting Pattern during and after Low Temperature and the Mechanism of Cold-Weather Tolerance at the Flowering Stage in Soybeans. *Plant Production Science* 6:247–254. <https://doi.org/10.1626/pps.6.247>
- Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. *Braz. J. Plant Physiol.* 124:743–756
- Lane HM, Murray SC, Montesinos-López OA, Montesinos-López A, Crossa J, Rooney DK, Barrero - Farfan ID, La Fuente GN de, Morgan CLS (2020) Phenomic selection and prediction of maize grain yield from near - infrared reflectance spectroscopy of kernels. *Plant phenome j.* 3. <https://doi.org/10.1002/ppj2.20002>
- Lawn RJ, Hume D (1985) Response of tropical and temperate soybean genotypes to temperature during early reproductive growth. *Crop Science* 25:137–142
- Lee SH, Park KY, Lee HS, Park EH, Boerma HR (2001) Genetic mapping of QTLs conditioning soybean sprout yield and quality. *Theor Appl Genet* 103:702–709. <https://doi.org/10.1007/s001220100595>
- Lees AK, Cullen DW, Sullivan L, Nicolson MJ (2002) Development of conventional and quantitative real-time PCR assays for the detection and identification of *Rhizoctonia*

- solani AG-3 in potato and soil. *Plant Pathology* 51:293–302. <https://doi.org/10.1046/j.1365-3059.2002.00712.x>
- Li W, Zheng D-H, Van K, Lee S-H (2008) QTL Mapping for Major Agronomic Traits across Two Years in Soybean (*Glycine max* L. Merr.). *Journal of Crop Science Biotechnology* 11:171–190
- Li H (2011) Improving SNP discovery by base alignment quality. *Bioinformatics* 27:1157–1158. <https://doi.org/10.1093/bioinformatics/btr076>
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- LIM BT, DeMAN JM, DeMAN L, BUZZELL RI (1990) Yield and Quality of Tofu as Affected by Soybean and Soymilk Characteristics. Calcium Sulfate Coagulant. *J Food Science* 55:1088–1092. <https://doi.org/10.1111/j.1365-2621.1990.tb01605.x>
- Lima FSd, Kurozawa LE, Ida EI (2014) The effects of soybean soaking on grain properties and isoflavones loss. *LWT - Food Science and Technology* 59:1274–1282. <https://doi.org/10.1016/j.lwt.2014.04.032>
- Liu W, Kim MY, Kang YJ, Van K, Lee Y-H, Srinives P, Yuan DL, Lee S-H (2011) QTL identification of flowering time at three different latitudes reveals homeologous genomic regions that control flowering in soybean. *Theor Appl Genet* 123:545–553. <https://doi.org/10.1007/s00122-011-1606-8>
- Lu W, Wen Z, Li H, Yuan D, Li J, Zhang H, Huang Z, Cui S, Du W (2013) Identification of the quantitative trait loci (QTL) underlying water soluble protein content in soybean. *Theor Appl Genet* 126:425–433. <https://doi.org/10.1007/s00122-012-1990-8>
- Ma Y, Reif JC, Jiang Y, Wen Z, Wang D, Liu Z, Guo Y, Wei S, Wang S, Yang C, Wang H, Yang C, Lu W, Xu R, Zhou R, Wang R, Sun Z, Chen H, Zhang W, Wu J, Hu G, Liu C, Luan X, Fu Y, Guo T, Han T, Zhang M, Sun B, Zhang L, Chen W, Wu C, Sun S, Yuan B, Zhou X, Han D, Yan H, Li W, Qiu L (2016) Potential of marker selection to increase prediction accuracy of genomic selection in soybean (*Glycine max* L.). *Mol Breeding* 36:113. <https://doi.org/10.1007/s11032-016-0504-9>
- Machado BQV, Nogueira APO, Hamawaki OT, Rezende GF, Jorge GL, Silveira IC, Medeiros LA, Hamawaki RL, Hamawaki CDL (2017) Phenotypic and genotypic correlations between soybean agronomic traits and path analysis. *Genet Mol Res* 16. <https://doi.org/10.4238/gmr16029696>
- Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic Mapping of Agronomic Traits Using Recombinant Inbred Lines of Soybean. *Crop Science* 36:1327. <https://doi.org/10.2135/cropsci1996.0011183X003600050042x>
- Mao T, Jiang Z, Han Y, Teng W, Zhao X, Li W (2013) Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. *Plant Breed* 132:630–641. <https://doi.org/10.1111/pbr.12091>
- Matei G, Woyann LG, Milioli AS, Bem Oliveira I de, Zdziarski AD, Zanella R, Coelho ASG, Finatto T, Benin G (2018) Genomic selection in soybean: accuracy and time gain in relation to phenotypic selection. *Mol Breeding* 38:32. <https://doi.org/10.1007/s11032-018-0872-4>
- Mian MA, Bailey MA, Tamulonis JP, Shipe ER, Carter TE, Parrott WA, Ashley DA, Hussey RS, Boerma HR (1996) Molecular markers associated with seed weight in two soybean populations. *Theor Appl Genet* 93:1011–1016. <https://doi.org/10.1007/BF00230118>
- Michailov VG, Marjuskin VF, Sichkar VI (1989) Cold tolerance during reproductive growth of soybean. *Soybean Genetic Newsletter* 16:211–213
- Min S, Yu Y, StMartin S (2005) Effect of Soybean Varieties and Growing Locations on the Physical and Chemical Properties of Soymilk and Tofu. *Journal of Food Science* 70:C8-C12

- Money D, Gardner K, Migicovsky Z, Schwaninger H, Zhong G-Y, Myles S (2015) LinkImpute: Fast and Accurate Genotype Imputation for Nonmodel Organisms. *G3* (Bethesda) 5:2383–2390. <https://doi.org/10.1534/g3.115.021667>
- Moongkanna J, Nakasathien S, Novitzky WP, Kwanyuen P, Sinchaisri P, Srinives P (2011) SSR Markers Linking to Seed Traits and Total Oil Content in Soybean. *Thai Journal of Agricultural Science* 44:233–241
- Nelson RL (2011) Managing self-pollinated germplasm collections to maximize utilization. *Plant Genet. Resour.* 9:123–133. <https://doi.org/10.1017/S147926211000047X>
- Nichols DM, Glover KD, Carlson SR, Specht JE, Diers BW (2006) Fine Mapping of a Seed Protein QTL on Soybean Linkage Group I and Its Correlated Effects on Agronomic Traits. *Crop Science* 46:834. <https://doi.org/10.2135/cropsci2005.05-0168>
- Ning H, Yuan J, Dong Q, Li W, Xue H, Wang Y, Tian Y, Li W-X (2018) Identification of QTLs related to the vertical distribution and seed-set of pod number in soybean *Glycine max* (L.) Merri. *PLoS One* 13:e0195830. <https://doi.org/10.1371/journal.pone.0195830>
- Orf JH, Chase K, Jarvik T, Mansur L, Cregan PB, Adler, F.R., Lark, K.G. (1999) Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. *Crop Science* 39:1642–1651
- Pan Z, Tangratanaalee W (2003) Characteristics of soybeans as affected by soaking conditions. *LWT - Food Science and Technology* 36:143–151. [https://doi.org/10.1016/S0023-6438\(02\)00202-5](https://doi.org/10.1016/S0023-6438(02)00202-5)
- Parmley K, Nagasubramanian K, Sarkar S, Ganapathysubramanian B, Singh AK (2019) Development of Optimized Phenomic Predictors for Efficient Plant Breeding Decisions Using Phenomic-Assisted Selection in Soybean. *Plant Phenomics* 2019:5809404. <https://doi.org/10.34133/2019/5809404>
- Piepho H-P, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. *Braz. J. Plant Physiol.* 177:1881–1888. <https://doi.org/10.1534/genetics.107.074229>
- Poysa V, Woodrow L (2002) Stability of soybean seed composition and its effect on soymilk and tofu yield and quality. *Food Research International* 35:337–345. [https://doi.org/10.1016/S0963-9969\(01\)00125-9](https://doi.org/10.1016/S0963-9969(01)00125-9)
- Poysa V, Woodrow L, Yu K (2006) Effect of soy protein subunit composition on tofu quality. *Food Research International* 39:309–317. <https://doi.org/10.1016/j.foodres.2005.08.003>
- R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Reif JC, Melchinger AE, Frisch M (2005) Genetical and Mathematical Properties of Similarity and Dissimilarity Coefficients Applied in Plant Breeding and Seed Bank Management. *Crop Science* 45:1–7
- Reuscher S, Furuta T (2016) ABHgenotypeR: Easy visualization of ABH Genotypes
- Rincen R, Charpentier J-P, Faivre-Rampant P, Paux E, Le Gouis J, Bastien C, Segura V (2018) Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar. *G3* (Bethesda) 8:3961–3972. <https://doi.org/10.1534/g3.118.200760>
- Rotundo JL, Westgate ME (2009) Meta-analysis of environmental effects on soybean seed composition. *Field Crops Research* 110:147–156. <https://doi.org/10.1016/j.fcr.2008.07.012>
- Salt DE, Baxter I, Lahner B (2008) Ionomics and the study of the plant ionome. *Annu Rev Plant Biol* 59:709–733. <https://doi.org/10.1146/annurev.arplant.59.032607.092942>
- Sanders R, Bari A, Street K, Devlin M (2013) A new approach to mining agricultural gene banks – to speed the pace of research innovation for food security. 'FIGS' - the Focused Identification of Germplasm Strategy

- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X-C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183. <https://doi.org/10.1038/nature08670>
- Shah T, Xu J, Zou X, Cheng Y, Nasir M, Zhang X (2018) Omics Approaches for Engineering Wheat Production under Abiotic Stresses. *Int J Mol Sci* 19:2390. <https://doi.org/10.3390/ijms19082390>
- Singh A, Jaiswal A, Singh A, Tomar RS, Kumar A (2022) Chapter 9 - Plant ionomics: toward high-throughput nutrient profiling. In: Kumar V (ed) *Plant Nutrition and Food Security in the Era of Climate Change*. Elsevier Science & Technology, San Diego, pp 227–254
- Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2013) Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS One* 8:e54985. <https://doi.org/10.1371/journal.pone.0054985>
- Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2015) Fingerprinting Soybean Germplasm and Its Utility in Genomic Research. *G3 (Bethesda)* 5:1999–2006. <https://doi.org/10.1534/g3.115.019000>
- Stanojevic SP, Barac MB, Pesic MB, Vucelic-Radovic BV (2011) Assessment of soy genotype and processing method on quality of soybean tofu. *J. Agric. Food Chem.* 59:7368–7376. <https://doi.org/10.1021/jf2006672>
- Steinhoff J, Liu W, Maurer HP, Würschum T, C. Friedrich HL, Ranc N, Reif JC (2011) Multiple-Line Cross Quantitative Trait Locus Mapping in European Elite Maize. *Crop Science* 51:2505. <https://doi.org/10.2135/cropsci2011.03.0181>
- Street K, Mackay M, Zuev E, Kaul N, El Bouhssini M, Konopka J, Mitrofanova O (2008) Diving into the genepool: a rational system to access specific traits from large germplasm collections:28–31
- Sun D, Li W, Zhang Z, Chen Q, Ning H, Qiu L, Sun G (2006) Quantitative trait loci analysis for the developmental behavior of Soybean (*Glycine max* L. Merr.). *Theor Appl Genet* 112:665–673. <https://doi.org/10.1007/s00122-005-0169-y>
- Tajuddin T, Watanabe S, Yamanaka N, Harada K (2003) Analysis of Quantitative Trait Loci for Protein and Lipid Contents in Soybean Seeds Using Recombinant Inbred Lines. *Breed. Sci.* 53:133–140. <https://doi.org/10.1270/jsbbs.53.133>
- Taylor J, Butler D (2017) R Package ASMap: Efficient Genetic Linkage Map Construction and Diagnosis. *J. Stat. Soft.* 79. <https://doi.org/10.18637/jss.v079.i06>
- Teng W, Han Y, Du Y, Sun D, Zhang Z, Qiu L, Sun G, Li W (2009) QTL analyses of seed weight during the development of soybean (*Glycine max* L. Merr.). *Heredity (Edinb)* 102:372–380. <https://doi.org/10.1038/hdy.2008.108>
- Thachuk C, Crossa J, Franco J, Dreisigacker S, Warburton M, Davenport GF (2009) Core Hunter: an algorithm for sampling genetic resources based on multiple genetic measures. *BMC Bioinformatics* 10:243. <https://doi.org/10.1186/1471-2105-10-243>
- Torkamaneh D, Laroche J, Valliyodan B, O'Donoghue L, Cober E, Rajcan I, Vilela Abdelnoor R, Sreedasyam A, Schmutz J, Nguyen HT, Belzile F (2020) Soybean (*Glycine max*) Haplotype Map (GmHapMap): a universal resource for soybean translational and functional genomics. *Plant Biotechnol J.* <https://doi.org/10.1111/pbi.13466>
- Tsubokura Y, Watanabe S, Xia Z, Kanamori H, Yamagata H, Kaga A, Katayose Y, Abe J, Ishimoto M, Harada K (2014) Natural variation in the genes responsible for maturity loci

- E1, E2, E3 and E4 in soybean. *Ann Bot* 113:429–441.
<https://doi.org/10.1093/aob/mct269>
- Utz HF (2011) *PlabMQTL*: Software for meta-QTL analysis with composite interval mapping. Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim
- Utz HF, Melchinger A, Schön C (2000) Bias and Sampling Error of the Estimated Proportion of Genotypic Variance Explained by Quantitative Trait Loci Determined From Experimental Data in Maize Using Cross Validation and Validation With Independent Samples. *Genetics* 154:1839–1894
- Utz, H.F., Melchinger, A.E. (1996) *PLABQTL*: a program for composite interval mapping of QTL. *Journal of Quantitative Trait Loci* 2:1–5
- van Heerden PDR, Kiddle G, Pellny TK, Mkwala PW, Jordaan A, Strauss AJ, Beer M de, Schlüter U, Kunert KJ, Foyer CH (2008) Regulation of respiration and the oxygen diffusion barrier in soybean protect symbiotic nitrogen fixation from chilling-induced inhibition and shoots from premature senescence. *Plant Physiol* 148:316–327.
<https://doi.org/10.1104/pp.108.123422>
- Vieira AJD, Oliveira DAd, Soares TCB, Schuster I, Piovesan ND, Martínez CA, Barros EGd, Moreira MA (2006) Use of the QTL approach to the study of soybean trait relationships in two populations of recombinant inbred lines at the F7 and F8 generations. *Braz. J. Plant Physiol.* 18:281–290. <https://doi.org/10.1590/S1677-04202006000200004>
- Vollmann J, Walter H, Sato T, Schweiger P (2011) Digital image analysis and chlorophyll metering for phenotyping the effects of nodulation in soybean. *Computers and Electronics in Agriculture* 75:190–195. <https://doi.org/10.1016/j.compag.2010.11.003>
- Wang L, Guan R, Zhangxiong L, Chang R, Qiu L (2006) Genetic Diversity of Chinese Cultivated Soybean Revealed by SSR Markers. *Crop Science* 46:1032–1038
- Wang X, Jiang G-L, Green M, Scott RA, Song Q, Hyten DL, Cregan PB (2014) Identification and validation of quantitative trait loci for seed yield, oil and protein contents in two recombinant inbred line populations of soybean. *Mol Genet Genomics* 289:935–949.
<https://doi.org/10.1007/s00438-014-0865-x>
- Wang C-CR, Chang SK-C (1995) Physicochemical Properties and Tofu Quality of Soybean Cultivar Proto. *J. Agric. Food Chem.* 43:3029–3034. <https://doi.org/10.1021/jf00060a008>
- Weir BS (1996) *Genetic Data Analysis II: Methods for discrete population genetic data*. Sinauer Associates, Sunderland
- Wickland DP, Battu G, Hudson KA, Diers BW, Hudson ME (2017) A comparison of genotyping-by-sequencing analysis methods on low-coverage crop datasets shows advantages of a new workflow, GB-eaSy. *BMC Bioinformatics* 18:586.
<https://doi.org/10.1186/s12859-017-2000-6>
- Wolfgang G, An Y-QC (2017) Genetic separation of southern and northern soybean breeding programs in North America and their associated allelic variation at four maturity loci. *Mol Breeding* 37:8. <https://doi.org/10.1007/s11032-016-0611-7>
- Würschum T (2012) Mapping QTL for agronomic traits in breeding populations. *Theor Appl Genet* 125:201–210. <https://doi.org/10.1007/s00122-012-1887-6>
- Würschum T, Abel S, Zhao Y (2014) Potential of genomic selection in rapeseed (*Brassica napus* L.) breeding. *Plant Breed* 133:45–51. <https://doi.org/10.1111/pbr.12137>
- Xu R, Tambong JT (2011) A TaqMan real-time PCR assay targeting the cytochrome o ubiquinol oxidase subunit II gene for detection of several pathovars of *Pseudomonas syringae*. *Canadian Journal of Plant Pathology* 33:318–331.
<https://doi.org/10.1080/07060661.2011.600335>
- Xu M, Xu Z, Liu B, Kong F, Tsubokura Y, Watanabe S, Xia Z, Harada K, Kanazawa A, Yamada T, Abe J (2013) Genetic variation in four maturity genes affects photoperiod

- insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biol* 13:91. <https://doi.org/10.1186/1471-2229-13-91>
- Yang A, James AT (2013) Effects of soybean protein composition and processing conditions on silken tofu properties. *J. Sci. Food Agric.* 93:3065–3071. <https://doi.org/10.1002/jsfa.6140>
- Yang A, James AT (2014) Influence of globulin subunit composition of soybean proteins on silken tofu quality. 1. Effect of growing location and 11SA4 and 7Sa' deficiency. *Crop Pasture Sci.* 65:259. <https://doi.org/10.1071/CP13398>
- Zhang W-K, Wang Y-J, Luo G-Z, Zhang J-S, He C-Y, Wu X-L, Gai J-Y, Chen S-Y (2004) QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. *Theor Appl Genet* 108:1131–1139. <https://doi.org/10.1007/s00122-003-1527-2>
- Zhang H-M (2008) QTL Mapping of Tofu and Soymilk Output in RIL Population NJRISX of Soybean. *A A S* 34. <https://doi.org/10.3724/SP.J.1006.2008.00067>
- Zhang Y, Li W, Lin Y, Zhang L, Wang C, Xu R (2018) Construction of a high-density genetic map and mapping of QTLs for soybean (*Glycine max*) agronomic and seed quality traits by specific length amplified fragment sequencing. *BMC Genomics* 19:641. <https://doi.org/10.1186/s12864-018-5035-9>
- Zhao Y, Gowda M, Liu W, Würschum T, Maurer HP, Longin FH, Ranc N, Reif JC (2012) Accuracy of genomic selection in European maize elite breeding populations. *Theor Appl Genet* 124:769–776. <https://doi.org/10.1007/s00122-011-1745-y>
- Zhou R, Chen H-F, Wang X-Z, Zhang X-J, Shan Z-H, Wu X-J, Cai S-P, Qiu D-Z, Zhou X-A, Wu J-S (2009) QTL Analysis of Yield, Yield Components, and Lodging in Soybean. *A A S* 35:821–830. <https://doi.org/10.3724/SP.J.1006.2009.00821>
- Zhu X, Willmar L, Leiser, Volker Hahn, Tobias Würschum (2021a) Phenomic selection is competitive with genomic selection for breeding of complex traits. *The Plant Phenome Journal*:e20027
- Zhu X, Leiser WL, Hahn V, Würschum T (2021b) Training set design in genomic prediction with multiple biparental families. *Plant Genome* 14:e20124. <https://doi.org/10.1002/tpg2.20124>

10 ÜBERBLICK ÜBER DIE IM BERICHTSZEITRAUM REALISIERTEN VERÖFFENTLICHUNGEN ZUM PROJEKT, BISHERIGE UND GEPLANTE AKTIVITÄTEN ZUM PROJEKT

PUBLIKATIONEN:

- Haupt, M, Schmid K (2020) Combining focused identification of germplasm and core collection strategies to identify genebank accessions for central European soybean breeding. *Plant, Cell & Environment*, 43(6), 1421-1436.
- Haupt, M, Schmid, K (2022) Using landscape genomics to infer genomic regions involved in environmental adaptation of soybean genebank accessions. *bioRxiv* DOI:2022.02.18.480989.
- Jähne, F, Balko C, Hahn V, Würschum T, Leiser WL (2019) Cold stress tolerance of soybeans during flowering: QTL mapping and efficient selection strategies under controlled conditions. *Plant Breeding*. DOI: 10.1111/pbr.12734.
- Jähne, F, Hahn V, Würschum T, Leiser WL 2020. Speed-Breeding short day crops by LED-controlled light schemes. *Theoretical and Applied Genetics*, 133: 2335-2342. DOI:10.1007/s00122-020-03601-4.
- Kurasch AK, Hahn V, Leiser WL, Starck N, Würschum T (2017) Phenotypic Analysis of Major Agronomic Traits in 1008 RILs from a Diallel of Early European Soybean Varieties. *Crop Science* 57:726. doi: 10.2135/cropsci2016.05.0318.
- Kurasch AK, Hahn V, Miersch M, Bachteler K, Würschum T (2018a) Analysis of tofu-related traits by a bench-scale tofu production method and their relationship with agronomic traits in European soybean. *Plant Breeding* 137:271–282. doi: 10.1111/pbr.12581.
- Kurasch AK, Leiser WL, Bachteler K, Miersch M, Hahn V, Würschum T (2018b) Evaluation of the genetic architecture of tofu traits in soybean towards genomics-assisted breeding. *Plant Breeding* 137:873–882. doi: 10.1111/pbr.12651 .
- Zhu X, Leiser WL, Hahn V, Würschum T (2021) Identification of seed protein and oil related QTL in 944 RILs from a diallel of early-maturing European soybean. *The Crop Journal* 9, 238-247. DOI:10.1016/j.cj.2020.06.006.
- Zhu X, Leiser WL, Hahn V, Würschum T (2021) Identification of QTL for seed yield and agronomic traits in 944 soybean (*Glycine max*) RILs from a diallel cross of early-maturing varieties. *Plant Breeding* 140, 254-266. DOI:10.1111/pbr.12900.
- Zhu X, Leiser WL, Hahn V, Würschum T (2021) Training set design in genomic prediction with multiple biparental families. *The Plant Genome* 14, e20124. doi.org/10.1002/tpg2.20124.
- Zhu X, Leiser WL, Hahn V, Würschum T (2021) Phenomic selection is competitive with genomic selection for breeding of complex traits. *The Plant Phenome Journal* 2021;4:e20027, DOI: 10.1002/ppj2.20027.

VORTRÄGE

- Hahn, V.: Die Züchtung von Tofu-Sojabohnen. Taifun Sojarunde, Freiburg, 5.9. 2015
- Hahn, V.: Herausforderung Sojazüchtung. Sojatagung 2015, Freising, 26.-27.11. 2015
- Hahn, V.: Aktuelles aus der Sojaforschung: Sojazüchtung, Feldtag Leguminosen im Ökologischen Landbau, Kleinhohenheim, 20.06. 2016
- Hahn, V.: Genomik-basierte Verbesserung des heimischen Sojazuchtmaterials und Etablierung eines molekularen Screeningsystems für Soja-Pathogene. Kongress

- "Hülsenfrüchte - Wegweiser für eine nachhaltigere Landwirtschaft", Berlin, 3. + 4. 11. 2016
- Hahn, V.: Sojazüchtung. Sitzung des Vorstandes des Landesbauernverbandes BW, Hohenheim, 17.11. 2016
- Hahn, V. 2017. Sojazüchtung für Deutschland – Wieso, Weshalb, Warum? JuLe-Tagung Karlsruhe 24.06.2017
- Hahn, V. 2017. Breeding Research on Soybean: Small Bean – Great Importance, Doktorandenseminar Pflanzenzüchtung, JLU Gießen, 20.11.2017
- Hahn, V. 2017. Sojazüchtung: Stand und Perspektiven – Wo geht es hin? Sojatagung 2017, Rastatt, 6.-7.12. 2017
- Hahn, V. 2018. Soja – warum jetzt auch in Deutschland? Vortrag im Rahmen einer Vortragsreihe beim veganen Brunch, Geislingen, 12.5.
- Hahn, V. 2018. Perspektiven der Sojazüchtung für Deutschland. GPZ/GFPi-Sommertagung Öl und Eiweißpflanzen, Hohenlieth, 4.-5. Juni 2018
- Hahn, V. 2019. Soja – warum jetzt auch in Deutschland? Vegan Street Days Stuttgart, 30. Mai 2019
- Hahn, V. 2019. European soya – possibilities and limitations. 22nd European Symposium on Poultry Nutrition, Danzig, Polen, 10-13. Juni 2019
- Hahn, V. 2020. Sojabohnen auf mitteldeutschen Äckern – was lässt die Pflanzenzüchtung erwarten? UFOP Perspektivforum 2020, Dresden, 23.9.2020
- Hahn, V. 2021. Soybean breeding. Im Rahmen der Veranstaltungsreihe ‚Breeding schemes and programs in plant and animal breeding‘, Universität Göttingen, Online, 20.05.2021
- Hahn, V. 2021. Soja – Warum jetzt auch in Deutschland? Vortrag im Rahmen einer Vortragsreihe beim veganen Brunch, Aalen, Online, 6.6.2021
- Hahn, V. 2021. From a wild plant to global importance: soybeans from a plant-breeding perspective. Tagung: Soy in the Anthropocene, Wien, hybrid 3.11.2021
- Hahn, V. 2021. Züchtung – Stand und Perspektiven. Vortrag im Rahmen der 52. Pflanzenbaulichen Vortragstagung mit dem Titel: Körnerleguminosen – eine tragende Säule eines zukunftsfähigen Ackerbaus? Online 18.11.2021
- Hirschburger, D., A. El-Hasan and R.T. Vögele 2016. Etablierung eines molekularen Screeningsystems für Soja-Pathogene. GPZ-Tagung, AG 14 Öl- und Eiweißpflanzen, Scharnhorst, 29.-30.6.2016

POSTER

- Hirschburger, D., A. El-Hasan and R.T. Vögele 2016. Establishment of a quantitative molecular detection system for soybean pathogens. DPG-Tagung, AK Wirt-Parasit-Interaktionen, Gießen, 17.-18.3.2016
- Hirschburger, D., A. El-Hasan and R.T. Vögele 2017. Establishment of a quantitative molecular detection system for *Pseudomonas syringae* on soybean. DPG-Tagung, AK Wirt-Parasit-Interaktionen, Rostock, 16.-17.3.2017
- Hirschburger, D., A. El-Hasan and R.T. Vögele 2017. Etablierung eines quantitativen molekularen Nachweissystems für Soja-Pathogene -ein SojaGenoPath-Projekt. Soja-Tagung 2017, Rastatt, 6.-7.12. 2017
- Hirschburger, D., A. El-Hasan and R.T. Vögele 2018. Development of a quantitative molecular detection system for the main pathogens causing damping-off on soybean. DPG-Tagung, AK Wirt-Parasit-Interaktionen, Köln, 15.-16.3.2018

ANHANG

Anhang 1: Liste des Pflanzenmaterials für die Kühleuntersuchungen am JKI

Laufende Nummer	Pedigree	WDH	Mit/Ohne Kontrolle	Ansatz	WDH	Mit/Ohne Kontrolle	Ansatz
1	PI445800	1	MK	A1	2	MK	A13
2	PI438444	1	MK	A1	2	MK	A13
3	PI549051A	1	MK	A1	2	MK	A13
4	PI549078	1	MK	nicht gekeimt	2		
5	PI556517	1	MK	A1	2	MK	A14
6	PI561332	1	MK	A1	2	MK	A14
7	PI449457	1	MK	A1	2	MK	A14
8	PI445824B	1	MK	A1	2	MK	A13
9	PI507711B	1	MK	A1	2	MK	A14
10	PI548607	1	MK	A1	2	MK	A13
11	PI561584	1	MK	A1	2	MK	A13
12	PI442042	1	OK	A1	2	OK	A13
13	PI514672	1	MK	A1	2	MK	A14
14	PI507704A	1	MK	A1	2	MK	A14
15	PI543932	1	MK	A1	2	MK	A15
16	PI507681A	1	MK	A1	2	MK	A15
17	PI555463	1	MK	A1	2	MK	A15
18	PI445791	1	MK	nicht gekeimt	2		
19	PI507685A	1	MK	A1	2	MK	A13
20	PI445804	1	MK	A1	2	MK	A13
21	PI445815	1	MK	nicht gekeimt	2		
22	PI438370	1	MK	nicht gekeimt	2		
23	PI561277	1	MK	A8	2	MK	A14
24	PI561278A	1	MK	A1	2	MK	A13
25	PI556756	1	MK	A1	2	MK	A14
26	PI556856	1	MK	nicht gekeimt	2		
27	PI437531	1	MK	A1	2	MK	A14
28	PI556812	1	MK	A1	2	MK	A15
29	PI584527	1	MK	nicht gekeimt	2		
30	PI556895	1	MK	nicht gekeimt	2		
31	PI567222	1	OK	A1	2	OK	A13
32	PI597404	1	OK	A7	2	OK	A18
33	PI445810B	1	MK	A1	2	MK	A15
34	PI556737	1	OK	A1	2	OK	A14
35	PI567156A	1	OK	A7	2	OK	A20
36	PI567217B	1	MK	nicht gekeimt	2		
37	PI556767	1	MK	A8	2	MK	A14
38	PI437823	1	MK	A1	2	MK	A14
39	PI438060	1	MK	A1	2	MK	A15
40	PI567223	1	OK	A7	2	OK	A20

41	PI603319	1	OK	A2	2	OK	A14
42	PI232999	1	MK	A2	2	MK	A13
43	PI556927	1	MK	nicht gekeimt	2		
44	PI458541	1	MK	A2	2	MK	A13
45	PI319537A	1	MK	A2	2	MK	A14
46	PI578433	1	MK	A2	2	MK	A14
47	PI248404	1	MK	A2	2	MK	A15
48	PI583365	1	MK	A2	2	MK	A15
49	PI548651	1	MK	A2	2	MK	A13
50	Fiskeby3	1	MK	nicht gekeimt	2		
51	PI438405	1	MK	nicht gekeimt	2		
52	PI548251	1	Ok	A2	2	OK	A15
53	PI437270A	1	MK	nicht gekeimt	2		
54	PI548345	1	MK	nicht gekeimt	2		
55	PI438456	1	MK	A2	2	MK	A13
56	PI361104	1	MK	A2	2	MK	A13
57	PI291300	1	MK	A2	2	MK	A14
58	PI297505	1	MK	A2	2	MK	A14
59	PI437551	1	MK	nicht gekeimt	2		
60	PI437175	1	MK	nicht gekeimt	2		
61	PI378655	1	MK	nicht gekeimt	2		
62	PI556599	1	MK	nicht gekeimt	2		
63	PI603169	1	MK	nicht gekeimt	2		
64	PI597390	1	MK	nicht gekeimt	2		
65	PI542056	1	MK	nicht gekeimt	2		
66	PI417295	1	MK	nicht gekeimt	2		
67	PI445795	1	MK	nicht gekeimt	2		
68	PI567212C	1	MK	nicht gekeimt	2		
69	PI417519B	1	MK	nicht gekeimt	2		
70	PI548508	1	MK	nicht gekeimt	2		
71	PI513382	1	MK	A3	2	MK	A11
72	PI548164	1	MK	A3	2	MK	A11
73	PI438160A	1	MK	A3	2	MK	A11
74	PI361085A	1	MK	A3	2	MK	A12
75	PI603440A	1	MK	A3	2	MK	A11
76	PI227330	1	MK	A3	2	MK	A11
77	PI43844i	1	MK	A3	2	MK	A18
78	PI153214	1	MK	A3	2	MK	A18
79	PI438377	1	MK	A3	2	MK	A12
80	PI438494	1	MK	A3	2	MK	A18
81	PI567220B	1	MK	A3	2	MK	A11
82	PI578391	1	MK	A3	2	MK	A11

83	FiskebyIV	1	MK	A3	2	MK	A12
84	PI407386A	1	MK	A3	2	MK	A12
85	PI437756A	1	MK	A3	2	MK	A18
86	PI548560	1	MK	A3	2	MK	A18
87	PI512322B	1	MK	A3	2	MK	A18
88	PI424209	1	MK	A3	2	MK	A19
89	PI437740B	1	MK	A3	2	MK	A19
90	PI437816	1	MK	A3	2	MK	A12
91	PI189872	1	MK	A3	2	MK	A11
92	PI561308	1	MK	A3	2	MK	A11
93	PI438404	1	MK	A3	2	MK	A12
94	PI561280	1	MK	A3	2	MK	A12
95	PI194643	1	MK	A3	2	MK	A18
96	PI437147	1	MK	A3	2	MK	A18
97	PI475821	1	MK	A3	2	MK	A19
98	PI437290	1	MK	A3	2	MK	A19
99	PI360952	1	MK	A3	2	MK	A18
100	PI417300	1	MK	A3	2	MK	A19
101	PI407713	1	MK	A4	2	MK	A11
102	PI437597B	1	MK	A4	2	MK	A11
103	PI161431B	1	MK	A4	2	MK	A12
104	PI567167	1	MK	A4	2	MK	A12
105	PI132207	1	MK	A4	2	MK	A18
106	PI438319B	1	MK	A4	2	MK	A18
107	PI567200A	1	MK	A4	2	MK	A19
108	PI597426	1	MK	A4	2	MK	A19
110	PI378674A	1	MK	A4	2	MK	A19
111	PI347544A	1	MK	A4	2	MK	A11
112	PI290116A	1	MK	A4	2	MK	A11
113	PI612740	1	MK	A4	2	MK	A12
114	PI437148	1	MK	A4	2	MK	A12
115	PI522192B	1	MK	A4	2	MK	A18
116	FC31122	1	MK	A4	2	MK	A18
117	PI319538	1	MK	A4	2	MK	A19
118	PI518706C	1	MK	A4	2	MK	A19
119	PI578415	1	MK	A4	2	MK	A19
120	PI407708B	1	MK	A4	2	MK	A19
121	PI295950	1	MK	A4	2	MK	A11
122	PI142491	1	MK	A4	2	MK	A11
123	PI423716	1	MK	A4	2	MK	A12
124	PI556667	1	MK	A4	2	MK	A12
125	PI556768	1	MK	A4	2	MK	A18
126	PI504506	1	MK	A4	2	MK	A18
127	PI290150	1	MK	A4	2	MK	A19
128	PI556490	1	MK	A4	2	MK	A19
129	PI424199	1	MK	A4	2	MK	A20
130	PI437204	1	MK	A4	2	MK	A20

131	PI291274A	1	MK	A4	2	MK	A19
132	PI438457	1	MK	A5	2	MK	A11
133	PI81775	1	MK	A5	2	MK	A12
134	PI592560	1	MK	A5	2	MK	A12
135	PI542065	1	MK	A5	2	MK	A18
136	PI548283	1	MK	A5	2	MK	A18
137	PI518702	1	MK	A5	2	MK	A19
138	PI297516	1	MK	A5	2	MK	A19
139	PI194656	1	MK	nicht gekeimt	2		
140	PI360964	1	MK	A5	2	MK	A12
141	PI578375A	1	MK	A5	2	MK	A20
142	PI319536C	1	MK	A5	2	MK	A11
143	PI438359	1	MK	A5	2	MK	A11
144	PI548378	1	MK	A5	2	MK	A12
145	PI290141	1	MK	A5	2	MK	A12
146	PI317335	1	MK	A5	2	MK	A18
147	PI92470	1	MK	A5	2	MK	A18
148	PI297518	1	MK	A5	2	MK	A19
149	PI597395	1	MK	A5	2	MK	A19
150	PI438502	1	MK	A5	2	MK	A20
151	PI561282C	1	MK	A5	2	MK	A20
152	PI592389	1	MK	A5	2	MK	A11
153	PI378664A	1	MK	A5	2	MK	A11
154	PI546487	1	MK	A5	2	MK	A12
156	FiskebyV	1	MK	A5	2	MK	A18
157	PI548615	1	MK	A5	2	MK	A18
158	PI603294	1	MK	A5	2	MK	A19
159	FC30684	1	MK	A5	2	MK	A19
160	PI290118	1	MK	A5	2	MK	A20
161	PI542769	1	MK	A6	2	MK	A20
162	PI297529	1	MK	A6	2	MK	A11
163	PI567209A	1	MK	A6	2	MK	A11
164	PI407708A	1	MK	A6	2	MK	A12
165	PI204652	1	MK	A6	2	MK	A12
166	PI438326	1	MK	A6	2	MK	A18
167	PI511867	1	MK	A6	2	MK	A18
168	PI238924	1	MK	A6	2	MK	A19
169	PI238925	1	MK	A6	2	MK	A19
170	PI437991A	1	MK	A6	2	MK	A12
171	PI458825A	1	MK	A6	2	MK	A12
172	PI291320A	1	MK	A6	2	MK	A11
173	PI508083	1	MK	A6	2	MK	A11
174	PI417511	1	MK	A6	2	MK	A12
175	PI437852	1	MK	A6	2	MK	A12
176	PI567783	1	MK	A6	2	MK	A18
177	PI297508	1	MK	A6	2	MK	A18
178	PI437557	1	MK	A6	2	MK	A19

179	PI518707	1	MK	A6	2	MK	A19
180	PI189955	1	MK	nicht gekeimt	2		
181	PI437767	1	MK	A6	2	MK	A20
182	PI548501	1	MK	A6	2	MK	A20
183	PI227331	1	MK	A6	2	MK	A11
184	PI567174B	1	MK	A6	2	MK	A11
185	PI437481	1	MK	A6	2	MK	A12
186	PI378664C	1	MK	A6	2	MK	A12
187	PI70520	1	MK	A6	2	MK	A18
188	PI153217	1	MK	A6	2	MK	A18
189	PI603297	1	MK	A6	2	MK	A19
190	PI153275	1	MK	nicht gekeimt	2		
191	PI153284	1	MK	nicht gekeimt	2		
192	PI154196	1	MK	nicht gekeimt	2		
193	PI603151B	1	MK	A7	2	MK	A19
194	PI567213A	1	MK	A7	2	MK	A20
195	PI153272	1	MK	nicht gekeimt	2		
196	PI297550	1	MK	A7	2	MK	A12
197	PI567225	1	MK	A7	2	MK	A11
198	PI597417	1	MK	A7	2	MK	A11
199	PI438349	1	MK	A7	2	MK	A12
200	PI567212D	1	MK	A7	2	MK	A12
201	PI506700	1	MK	A7	2	MK	A18
202	PI258387	1	MK	A7	2	MK	A18
203	PI438373	1	MK	A7	2	MK	A19
204	PI437566	1	MK	A7	2	MK	A19
205	PI445819	1	MK	A7	2	MK	A20
206	PI507285	1	MK	A7	2	MK	A20
207	PI378669C	1	MK	A7	2	MK	A11
208	PI153209	1	MK	nicht gekeimt	2		
209	PI290124	1	MK	nicht gekeimt	2		
210	PI567785	1	MK	A7	2	MK	A11
211	PI555464	1	MK	A7	2	MK	A12
212	PI518753	1	MK	A7	2	MK	A20
213	PI437316	1	MK	A7	2	MK	A18
214	PI437296	1	MK	A7	2	MK	A18
215	PI437139	1	MK	A7	2	MK	A19
216	PI507700	1	MK	A7	2	MK	A19
217	PI556749	1	MK	A7	2	MK	A20
218	PI567157B	1	MK	A7	2	MK	A20
219	PI438273	1	MK	A7	2	MK	A20
220	PI437877A	1	MK	A7	2	MK	A20
221	PI437556	1	MK	A8	2	MK	A15
222	PI437741	1	MK	A8	2	MK	A15
223	PI437645B	1	MK	A8	2	MK	A15

224	PI437706	1	MK	A8	2	MK	A13
225	PI591431	1	MK	A8	2	MK	A13
226	PI437469B	1	MK	A8	2	MK	A13
227	PI437905	1	MK	A8	2	MK	A14
228	PI437436A	1	MK	A8	2	MK	A14
229	PI437833	1	MK	A8	2	MK	A15
230	PI437579	1	MK	A8	2	MK	A15
231	PI153219	1	MK	A8	2	MK	A14
232	PI548642	1	MK	A8	2	MK	A15
233	FruhwirtsSchwarze Pedoja	1	MK	A8	2	MK	A13
234	PI507673	1	MK	A8	2	MK	A13
235	Primus	1	MK	nicht gekeimt	2		
236	PI132214	1	MK	nicht gekeimt	2		
237	PI578422	1	MK	A8	2	MK	A15
238	PI597427A	1	MK	A8	2	MK	A15
239	PI437928	1	MK	A8	2	MK	A15
240	PI442045	1	MK	A8	2	MK	A15
241	PI567209B	1	MK	A8	2	MK	A13
242	((PrimusxGallec)I5x (ProteixxProtina)I4)	1	MK	A8	2	MK	A13
243	PI437920A	1	MK	A8	2	MK	A14
244	PI548380	1	MK	A8	2	MK	A14
245	PI612710	1	MK	A8	2	MK	A15
246	PI548375	1	MK	A8	2	MK	A15
247	PI596541	1	MK	A8	2	MK	A15
248	PI437178	1	MK	A8	2	MK	A16
249	PI522192A	1	MK	A8	2	MK	A13
250	PI189880	1	MK	A8	2	MK	A13
251	PI347567	1	MK	A9	2	MK	A14
252	Regina	1	MK	nicht gekeimt	2		
253	PI445834	1	MK	nicht gekeimt	2		
254	PI317333	1	MK	nicht gekeimt	2		
255	Korus	1	MK	A9	2	MK	A14
256	PI153223	1	MK	A9	2	MK	A15
257	PI438449	1	MK	A9	2	MK	A15
258	PI567782	1	MK	A9	2	MK	A16
259	PI154191	1	MK	A9	2	MK	A16
260	PI417509	1	MK	A9	2	MK	A13
261	PI189924	1	MK	A9	2	MK	A13
262	PI153234	1	OK	A9	2	OK	A13
263	Taifun 3	1	MK	A9	2	MK	A14
264	(PI437179)I1	1	MK	A9	2	MK	A14
265	Bettina	1	MK	nicht gekeimt	2		
266	FC30689	1	MK	A9	2		
267	PI424206	1	MK	A9	2	MK	A15

268	PI538392	1	MK	A9	2	MK	A16
269	PI548593	1	MK	A9	2	MK	A16
270	PI567171	1	MK	A9	2	MK	A13
271	PI548539	1	OK	A9	2	OK	A13
272	Protina	1	MK	A9	2	MK	A13
273	PI319537B	1	MK	A9	2	MK	A14
274	PI438399	1	MK	A9	2	MK	A14
275	PI291318	1	MK	A9	2	MK	A15
276	PI437483	1	MK	A9	2	MK	A15
277	PI612617A	1	MK	A9	2	MK	A16
278	PI561353	1	OK	A9	2	OK	A14
279	PI549074	1	MK	nicht gekeimt	2		
280	PI154193	1	MK	A9	2	MK	A16
281	PI506609	1	OK	A10	2	OK	A15
282	PI70520	1	OK	A10	2	OK	A16
283	PI518750	1	OK	A10	2	OK	A16
284	PI404155B	1	MK	A10	2	MK	A13
285	PI437998	1	MK	A10	2	MK	A13
286	PI161431A	1	MK	A10	2	MK	A14
287	PI603900	1	MK	A10	2	MK	A14
288	PI593938	1	OK	A10	2	OK	A13
290	PI416904C	1	MK	A10	2	MK	A15
291	(ESMentorx(Primusx Sultana)I5)	1	MK	A10	2	MK	A15
292	PI548596	1	MK	A10	2	MK	A16
293	PI437474B	1	MK	A10	2	MK	A16
294	PI153303	1	MK	A10	2	MK	A13
295	PI437320	1	OK	A10	2	OK	A13
296	PI549079	1	MK	A10	2	MK	A13
297	PI445789	1	MK	A10	2	MK	A14
298	PI189886	1	MK	A10	2	MK	A14
299	PI591429	1	MK	A10	2	MK	A15
300	PI549073	1	MK	A10	2	MK	A15
301	PI 546047	1	OK	A7	2	OK	A11
302	PI 546050	1	OK	nicht gekeimt	2		
303	PI 445797	1	OK	nicht gekeimt	2		
304	PI 546042	1	OK	A7	2	OK	A12
305	PI 445808A	1	OK	A7	2	OK	A18
306	PI 438389	1	OK	A7	2	OK	A18
308	PI548545	1	MK	A7	2	MK	A19
309	PI567211B	1	OK	A7	2	OK	A12
310	PI 445802	1	OK	A2	2	OK	A15
311	PI 464910	1	OK	A2	2	OK	A13
312	PI507373	1	OK	A2	2	OK	A13
313	PI 507678	1	OK	A3	2	OK	A11
314	PI196501	1	OK	A3	2	OK	A20
315	PI361073	1	OK	A3	2	OK	A12

316	PI372411	1	OK	A3	2	OK	A12
317	PI372412	1	OK	A3	2	OK	A18
318	PI194655	1	OK	A3	2	OK	A18
319	PI417564	1	OK	A3	2	OK	A12
320	PI423715	1	OK	A3	2	OK	A19
321	PI430491	1	OK	A3	2	OK	A11
322	PI189857	1	OK	A3	2	OK	A11
323	PI290127	1	OK	A3	2	OK	A20
324	PI372413	1	OK	A3	2	OK	A12
325	PI416845	1	OK	A4	2	OK	A18
326	PI417541	1	OK	A4	2	OK	A18
327	PI437153A	1	OK	A4	2	OK	A19
328	PI437191	1	OK	A4	2	OK	A19
329	PI297514	1	OK	A4	2	OK	A11
330	PI36653	1	OK	A4	2	OK	A11
331	PI378677C	1	OK	A4	2	OK	A12
332	PI417525	1	OK	A4	2	OK	A11
333	PI423872	1	OK	A4	2	OK	A18
334	TOKACHINAGAHA	1	OK	A4	2	OK	A18
335	PI437081A	1	OK	A4	2	OK	A19
336	PI437156C	1	OK	A4	2	OK	A19
337	PI437162	1	OK	A5	2	OK	A11
338	PI437187	1	OK	A5	2	OK	A11
339	OOYACHI2	1	OK	nicht gekeimt	2		
340	WASEOUSODE (SHIKAOIITOH)	1	OK	nicht gekeimt	2		
341	KOGANEDAIZU	1	OK	nicht gekeimt	2		
342	HOSOBA1	1	OK	nicht gekeimt	2		
343	CHUUSEIKURODAIRYU	1	OK	A5	2		
344	RANKOSHI1	1	OK	A5	2	OK	A11
345	HADAKAKURO2	1	OK	A5	2	OK	A18
346	TOIKU101	1	OK	A5	2	OK	A18
347	TOKEI778	1	OK	nicht gekeimt	2		
348	TO3N	1	OK	A5	2		
349	(PI196501)I1	1	OK	A6	2	OK	A11
350	(PI361073)I1	1	OK	A6	2	OK	A20
351	(PI372411)I1	1	OK	A6	2	OK	A20
352	(PI372412)I1	1	OK	A6	2	OK	A18
353	(PI194655)I1	1	OK	A6	2	OK	A18
354	(PI417564)I1	1	OK	A6	2	OK	A19
356	(PI430491)I1	1	OK	A6	2	OK	A20
357	(PI189857)I1	1	OK	nicht gekeimt	2		
358	(PI290127)I1	1	OK	A6	2		
359	(PI372413)I1	1	OK	A6	2	OK	A20
360	(PI416845)I1	1	OK	nicht gekeimt	2		

383	1152	1	OK	nicht gekeimt	2		
384	((PrimusxProtina)I5x (ProteixxProtina)I4)I6	1	OK	A8	2	OK	A16
385	(ESMentorx(Primusx Sultana)I5)I5	1	OK	nicht gekeimt	2		
386	Adsoy	1	OK	nicht gekeimt	2		
387	Alexa	1	OK	A9	2	OK	A14
388	Aligator	1	OK	nicht gekeimt	2		
389	Amandine	1	OK	A9	2	OK	A16
390	Amarok	1	OK	nicht gekeimt	2		
391	Amphor	1	OK	nicht gekeimt	2		
392	Ana	1	OK	nicht gekeimt	2		
393	Ascasubi	1	OK	A9	2	OK	A16
394	Aveline	1	OK	A9	2	OK	A13
395	Bohemians	1	OK	nicht gekeimt	2		
396	Buenos	1	OK	nicht gekeimt	2		
397	Capnor	1	OK	A10	2	OK	A15
398	ESMentor	1	OK	nicht gekeimt	2		
399	ESSenator	1	OK	nicht gekeimt	2		
400	Eugen	1	OK	nicht gekeimt	2		
401	Fiskeby2	1	OK	A10	2	OK	A15
402	Gallec	1	MK	A10	2	MK	A16
403	Idefix	1	OK	A10	2	OK	A16
404	Josefine	1	OK	A10	2	OK	A14
405	KANAGAWAWASE	1	OK	A10	2	OK	A14
406	KITAHOMARE	1	OK	nicht gekeimt	2		
407	KITAMISHIRO56	1	OK	nicht gekeimt	2		
408	Lissabon	1	MK	A10	2	MK	A15
409	Lotus	1	OK	A5	2	OK	A12
410	Luna	1	OK	A5	2	OK	A11
411	Marquise	1	OK	nicht gekeimt	2		
412	Naya	1	OK	A6	2	OK	A12
413	Obelix	1	OK	A7	2	OK	A20
414	Onix	1	OK	A7	2	OK	A20
415	Optimus	1	OK	A8	2	OK	A16
416	Paradis	1	OK	nicht gekeimt	2		
417	Pollux	1	OK	A9	2	OK	A13
418	Proteix	1	OK	A9	2	OK	A15
419	ROI-416F	1	OK	A10	2	OK	A15
420	Ruzika	1	OK	A6	2	OK	A12
421	S/W	1	OK	nicht gekeimt	2		
422	SHINSEI60	1	OK	nicht gekeimt	2		

423	SHIROTSURUNOKO	1	MK	A10	2	MK	A16
424	Shouna	1	OK	A10	2	OK	A16
425	Sigalia	1	OK	nicht gekeimt	2		
426	Sirelia	1	OK	nicht gekeimt	2		
427	Solena	1	MK	A10	2	MK	A14
428	Stumpa	1	OK	A10	2	OK	A14
429	Sultana	1	OK	nicht gekeimt	2		
430	Sunrise	1	OK	A8	2		
431	SY Eliot	1	OK	A8	2	OK	A14
432	Sylvia	1	OK	A8	2	OK	A15
433	TENPOKUSHIROME	1	OK	A8	2	OK	A13
434	Tiguan	1	OK	A8	2	OK	A14
435	Tourmaline	1	OK	A8	2	OK	A15
436	Triumpf	1	OK	nicht gekeimt	2		
437	Valjevka	1	OK	A8	2		
438	Zlata	1	OK	A8	2	OK	A14
439	PI180499	1	MK	A5	2	MK	A20
440	PI578388A	1	MK	A5	2	MK	A11
441	PI567158	1	OK	A10	2	OK	A15
442	Lenka	1	MK	nicht gekeimt	2		
443	Bettina	1	OK	A10	2		
444	Merlin	1	OK	A10	2	OK	A16
445	PI561280	1	MK	A10	2	MK	A16
446	PI561282	1	MK	A10	2	MK	A16