

Forschungsvorhaben Teilprojekt: 0315958A

Thema: Verbundprojekt OPTIMAL: Auf genetischen Markern und Biomarkern basierte prädikative Züchtung von Mais-Kulturvarietäten

Schlussbericht zu Nr. 3.2

I.1. Aufgabenstellung

Ziel des Verbundvorhabens OPTIMAL war die Entwicklung und Optimierung eines neuartigen Hybridzuchtansatzes, der DNA Polymorphismen als genetische Marker mit Metabolitprofilen (MPs) und Ganzpflanzenmerkmale als Biomarker kombiniert, um Hybridleistung vorherzusagen. Insbesondere sollten umfangreiche Daten erhoben werden, auf deren Basis neue Vorhersagemodelle entwickelt und die Effizienz der neuen Verfahrensweise im Vergleich zu (a) Standardmethoden der Hybridvorhersage basierend auf ‚general combining ability‘ (Kombinationsfähigkeit; GCA) und (b) neuen, auf genomischer Selektion beruhenden Methoden der Hybridvorhersage mit hochdichten SNP Markern verglichen werden können. Die Untersuchungen sollten mit adaptierten Elitezuchtlinien erfolgen, um die Verfahren möglichst praxis-nah evaluieren zu können und unmittelbar mit der Etablierung und Anwendung der neuen Züchtungsstrategie beginnen zu können. Dazu sollten in einer kollaborativen Anstrengung im OPTIMAL-Verbundvorhaben mehrere Aspekte untersucht werden: (i) Reproduzierbarkeit und Wert (Informationsgehalt und Stabilität) von Metabolitprofilen von unterschiedlichen Pflanzenorganen (Spross, Blätter, Wurzel) aus unterschiedlichen Umweltbedingungen (Feld, Gewächshaus, Phytokammer) sollten verglichen und bewertet werden. (ii) Informationen über Präsenz/Absenz Variation (PAV) und Kopienzahlvariation (CNV) in europäischen Zuchtlinien sollten mittels vergleichender genomischer Hybridisierung (comparative genomic hybridization; CGH) gewonnen werden. (iii) Verfahren zur automatisierten und nicht-invasiven Erhebung phänomischer Daten sollten verfeinert werden. (iv) Das Verhältnis von genetischen Markerdaten (SNPs, PAVs, CNVs), Biomarkerdaten (MPs, phänomische Daten) und agronomisch wichtigen Merkmalen sollten für eine ‚optimale Prediktorauswahl‘ (‚optimal feature (predictor) selection‘) erkundet werden, und die vorhandenen Methoden der Hybridvorhersage mit SNPs auch für die anderen Marker angepasst und weiterentwickelt werden. (v) Die Präzision der Hybridvorhersage mit den unterschiedlichen Markertypen sollte bestimmt werden. (vi) MPs und phänomische Daten sollten erhoben und analysiert werden, um ein besseres Verständnis der molekularen Mechanismen von Heterosis zu erreichen. (vii) Letztendlich sollten verbesserte Methoden der Hybridvorhersage identifiziert, einschließlich neuer statistischer Verfahren, und verbesserte Selektionsschemata für Hybridzuchtungsprogramme entworfen werden.

Neben der Koordination des OPTIMAL-Gesamtverbundes bestand die Aufgabenstellung des Teilprojektes A in der Erhebung entscheidender umfangreicher Datensätze an den vom Projektpartner 4 (UH) zur Verfügung gestellten Inzuchtlinien und Hybriden und die Bereitstellung dieser Daten für Partner 7 (UP) für die kombinierte Analyse zusammen mit den von den Projektpartnern 5 (MPI-MP-LW) und 6 (MPI-MP-MS) an den selben Linien gesammelten Daten. Dazu waren nach der Optimierung der Kultivierungsbedingungen tiefen

Phänotypisierungsdaten für ca. 200 Inzuchtlinien und ca. 400 Hybriden mit Hilfe der automatisierten Phänotypisierungsanlage (LemnaTec) des IPK zu erheben, mittels GC-MS-Analysen Metabolitprofildaten für die kultivierten Pflanzen (primär für die Inzuchtlinien) aufzunehmen und gemeinsam mit dem Unterauftragnehmer (Prof. P. Schnable, Iowa State University, USA) Informationen bzgl. An-/Abwesenheitsvariation (PAV) bzw. Kopiezahlvariation (CNV) von Sequenzen in den Genome der ca. 200 Inzuchtlinien zu erheben. Die gesammelten Rohdatensätze waren dabei einer intensiven Vorverarbeitung und individuellen Primäranalyse zu unterziehen, sodass für die jeweiligen Analyseverfahren und für jeden untersuchten Genotyp jeweils sehr umfangreiche Messwertsätze zusammengestellt und für die integrierte Analyse zur Verfügung gestellt werden konnten.

I.2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Für die Durchführung des Vorhabens standen der ausführenden Stelle durch das Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung die nötigen Labor- und Pflanzenkulturräume und –einrichtungen zur Verfügung. Vom Projektpartner 4 (Universität Hohenheim) wurde eine Kollektion von 200 Elite Maisinzuchtlinien (100 Dent-Linien, 100 Flint-Linien), sowie 396 dazugehörige faktorielle Testkreuzungen zur Verfügung gestellt.

I.3. Planung und Ablauf des Vorhabens

Die geplanten Projekttreffen wurden dem Arbeitsfortschritt angepasst. Das erste Meeting fand im Frühling 2012 statt (April 2012, Potsdam-Golm), um zuvor ausreichend Zeit für die Personalrekrutierung zu erlauben. Das zweite Meeting wurde im Frühling 2013 abgehalten (März 2013, Hohenheim), und ein drittes im Herbst 2013 (November 2013, Einbeck). Daneben fanden vielfältige Treffen bzw. Telefonate oder Telefonkonferenzen individueller Partner statt, besonders auch im Kontakt mit dem Unterauftragnehmer zur Abstimmung der Verfahrensweise zur Identifizierung von CNV/PAV, zur Beobachtung des Projektfortschritts und zur Diskussion der erhaltenen Ergebnisse. Ferner fand im April 2015 ein/e Treffen/Telefonkonferenz der an der Datenanalyse zentral beteiligten akademischen Partner (P1, P4, P5, P6, P7) zur Abstimmung der abschließenden Datenanalyse statt. Nachdem diese weitestgehend erfolgt ist, ist ein weiteres Treffen aller Partner für einen abschließenden Informations- und Erfahrungsaustausch vorgesehen. Veränderungen im Ablaufplan (siehe unten) machten es nötig, eine zuwendungsneutrale Verlängerung zu beantragen und den Arbeitsplan entsprechend anzupassen. Die Berichtstermine wurden gemäß der Bestimmungen angeglichen.

Der Ablauf des Vorhabens wurde dahingehend geändert, dass in Teilprojekt A die Experimente zur Optimierung der Gewächshausbedingungen auf Grund der späten Verfügbarkeit des Saatguts erst nach sechs Monaten Verzögerung durchgeführt werden konnten. Diese Verzögerung wirkte sich auch auf Phänotypisierung und Ernte des Pflanzenmaterials für Metabolitanalysen aus. Dies konnte nur teilweise durch verkürzte Kultivierungszeiten ausgeglichen werden und war einer der wesentlichen Gründe für die Notwendigkeit einer Projektlaufzeitverlängerung.

Fortschritte in Next-Generation-Sequencing (NGS) während der Projektlaufzeit führten zu einem zusätzlichen Vorversuch, um verschiedene Methoden zur Identifizierung von CNV/PAV zu vergleichen: die ursprünglich geplante Array-basierte genomische Hybridisierung (aCGH) und das neue sequence-capture basierte NGS (scNGS) -Verfahren. Dieser Vorversuch wurde im Schnable Labor an der Iowa State University, USA, durchgeführt und ergab, dass scNGS ermöglicht, mehr Gene mit höherer Auflösung zu analysieren, und dadurch die Detektion von PAV/CNV verbessert werden kann. Zusätzlich, und ohne weitere Kosten, können die Millionen von scNGS reads auch zur Identifizierung von möglichen SNPs verwendet werden. Daher wurde nach ausführlichen Beratungen beschlossen, scNGS statt aCGH anzuwenden, und dies vom Schnable Labor als Subunternehmer durchführen zu lassen.

Im Folgenden sind die Arbeiten des Verbundprojektes dargestellt, zu denen im Rahmen des Teilprojektes A beigetragen wurde:

1. Etablierung optimierter Kulturbedingungen im klimatisierten Gewächshaus
2. Erhebung manueller phänotypischer Daten (Frisch- und Trockengewicht) und mittels Bildanalyse aus Hochdurchsatz-Bilddaten extrahierte phänotypische Daten für 198 Inzuchtlinien und 396 ihrer Hybride (bereitgestellt durch Partner 4, UH).
3. Metabolitprofile (469 polare Metabolite, davon 134 mit bekannter chemischer Struktur) von 198 öffentlichen Inzuchtlinien.
4. Etablierung von Zeanome Sequence Capture in Kooperation mit ISU.
5. PAV/CNV Profile von 201 UH-Inzuchtlinien.
6. 252 Millionen SNPs in 198 UH-Inzuchtlinien.

I.4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Hybridzüchtung nutzt das biologische Phänomen 'Heterosis', das die erhöhte Leistungsfähigkeit von Hybriden gegenüber ihren (Inzucht) Eltern bezeichnet [1]. Heterosis ist besonders auffällig in Mais und wird seit über einem Jahrhundert wissenschaftlich untersucht [2]. Nutzung von Heterosis in der Hybridzüchtung beinhaltet die Entwicklung von Inzuchtlinien und die Selektion der besten Kombinationen zwischen ihnen [3]. Die Anwendung des Doppelt-Haploiden-Verfahrens in Mais erleichterte und beschleunigte die Produktion neuer Zuchtlinien [4-7], so dass die Anzahl der in jedem Zuchtzyklus produzierten Linien, und damit die mögliche Anzahl von Kombinationen zur Hybridbildung, erheblich angestiegen ist [8,9]. Dadurch bedeutet(e) die Beurteilung neuer Inzuchtlinien und die Vorhersage der Hybridleistung der großen Anzahl möglicher Hybride eine wesentliche Herausforderung für die moderne Maiszüchtung [10-15]. Zu Beginn des Projektes selektierten die meisten Züchter Hybridkombinationen auf Grund der Elternleistungen oder Schätzungen ihrer Kombinationsfähigkeit (GCA) [16]. Zur besseren Vorhersage der Hybridleistung wurde das BLUP Verfahren ((BLUP=Best Linear Unbiased Prediction) aus der Tierzüchtung auch für die Maiszüchtung vorgeschlagen [17-19]. Alle genannten Methoden erforderten umfangreiche Feldversuche, und waren daher sehr zeitaufwendig und teuer. Versuche, Hybridleistung in interpool Hybriden mittels Marker-basierten genetischen Distanzen vorherzusagen, schlugen fehl [20]. Ein ähnliches Verfahren mit Transkriptions-basierten Distanzen brachte jedoch vielversprechende Ergebnisse [21]. Weitere Prediktoren beruhten

auf QTL Schätzwert, die durch multiple lineare Regression ermittelt wurden [22-26]. Diese Methoden erreichten eine mit dem klassischen GCA Verfahren vergleichbare Präzision, aber zu sehr viel niedrigeren Kosten. Mit der hoch-dichten Markerabdeckung des Genoms konnte auch genomische Selektion [27] zur Vorhersage der Hybridleistung herangezogen werden. Zu Beginn dieses Projektes wurde dieses Verfahren von P4 innerhalb des Synbreed Kompetenzclusters [28] evaluiert – mit vielversprechenden Erkenntnissen während der Laufzeit von OPTIMAL [29, 30]

In OPTIMAL sollte ein neues Konzept der Leistungsvorhersage basierend auf Metabolitprofilen als Biomarkern auf sein Anwendungspotential in der Maiszüchtung untersucht werden. Signifikante und spezifische Zusammenhänge zwischen Biomasse und Metabolitzusammensetzung von Pflanzen wurden ursprünglich in großen *Arabidopsis thaliana* Populationen und entsprechenden Testkreuzungen unter Verwendung multivariater statistischer Verfahren entdeckt [31-33]. Weitere Analysen zeigten, dass auch Hybridleistung vorhergesagt werden konnte, und dass die Präzision durch Einbeziehung genetischer Markerdaten verbessert werden konnte [34,35]. Vorläufige Ergebnisse des GABI-ENERGY Verbundprojekts mit einer möglichst diversen Maispopulation deuteten zu Beginn von OPTIMAL einen ähnlichen Zusammenhang auch in der Kulturpflanze Mais an. Allerdings sollten in OPTIMAL Elite Zuchtlinien aus aktuellen Zuchtprogrammen verwendet werden, und neben Metabolitprofilen auch genetische Markerdaten in die Modelle inkorporiert werden. Zusätzlich zu den Analysen in Inzuchtlinien sollten auch Metabolitprofile und phänomische Daten in Hybriden erhoben werden, mit dem Ziel, genetische und molekulare Mechanismen, die zu erhöhter Hybridleistung führen, analysieren zu können. Es wurde auch vermutet, dass durch die Hybriddaten die Vorhersagemodelle verbessert werden könnten.

Strukturelle genomische Variation (PAV, CNV) wurde zu Beginn von OPTIMAL, zusammen mit SNPs, gerade als wichtige Faktoren für die Ausprägung agronomischer Merkmale und Hybridleistung vorgeschlagen. Es war gezeigt worden, dass CNVs und PAVs in Mais sehr häufig vorkommen [36,37,38], und es wurde vermutet, dass diese strukturelle Variation zwischen Maisinzuchtlinien entscheidend zu der hohen phänotypischen Diversität und Plastizität, und auch zu Heterosis beiträgt [36,35]. So könnten Tandem-Duplikationen für die Evolution von Krankheitsresistenzen wichtig sein. CNV könnte die Evolution neuer Expressionsmuster erleichtern, und die große Zahl von PAV Kombinationen in Hybriden könnte zu viele neuartigen Genzusammensetzungen in Hybriden führen. Daher sollten mehrere hundert Mais Zuchtlinien im Rahmen von OPTIMAL auf ihre strukturelle Genomvariation hin analysiert werden.

Literaturangaben

- [1] Lippman, Z.B. and Zamir, D. (2007) Heterosis: revisiting the magic. Trends Genet. 23: 60-66.
- [2] Springer, N.M. and Stupar, R.M. (2007) Allelic variation and heterosis in maize: How do two halves make more than a whole? Genome Res. 17: 264-275.
- [3] Becker, H.C. (1993) Pflanzenzüchtung. Ulmer Verlag.
- [4] Geiger, H.H. (2009) Double haploids. In: Bennetzen J., Hake, S. (eds.) Handbook of maize - genetics and genomics, Springer Science + Business Media, LLC, New York, pp. 641-657.

- [5] Melchinger, A.E., Longin, C.F.H., Utz, H.F., and Reif, J.C. (2005) Hybrid maize breeding with doubled haploid lines: quantitative genetic and selection theory for optimum allocation of resources. *In: 41th Illinois Corn Breeders' School, University of Illinois at Urbana-Champaign, 2005*, pp. 8-21.
- [6] Prigge, V., Sánchez, C., Dhillon, B.S., Schipprack, W., Araus, J.L., Bänziger, M., and Melchinger, A.E. (2011) Doubled haploids in tropical maize: I. Effects of inducers and source germplasm on *in vivo* haploid induction rates. *Crop Sci.* 51: 1498-1506.
- [7] Röber, F.K., Gordillo, G.A., and Geiger, H.H. (2005) *In vivo* haploid induction in maize – performance of new inducers and significance of double haploid lines in hybrid breeding. *Maydica* 50: 275-283.
- [8] Seitz, G. (2005) The use of doubled haploids in corn breeding. *In: Proceedings of the forty first annual Illinois corn breeders' school 2005, Urbana-Champaign, Illinois, USA*, pp. 1-8.
- [9] Smith J.S.C., Hussain, T., Jones, E.S., Graham, G., Podlich, D., Wall, S., and Williams, M. (2008) Use of double haploids in maize breeding: implications for intellectual property protection and genetic diversity in hybrid crops. *Mol. Breeding* 22: 51-59.
- [10] Longin, C.F.H., Utz, H.F., Reif, J.C., Schipprack, W., and Melchinger, A.E. (2006) Hybrid maize breeding with doubled haploids: I. One-stage versus two-stage selection for testcross performance. *Theor. Appl. Genet.* 112: 903-912.
- [11] Longin, C.F.H., Utz, H.F., Melchinger, A.E., and Reif, J.C. (2007) Hybrid maize breeding with doubled haploids: II. Optimum type and number of testers in two-stage selection for general combining ability. *Theor. Appl. Genet.* 114: 393-402.
- [12] Longin, C.F.H., Utz, H.F., Reif, J.C., Wegenast, T., Schipprack, W., and Melchinger, A.E.M. (2007) Hybrid maize breeding with doubled haploids: III. Efficiency of early testing prior to doubled haploid production in two-stage selection for testcross performance. *Theor. Appl. Genet.* 115: 519-527.
- [13] Wegenast, T., Longin, C.F.H., Utz, H.F., Melchinger, A.E., Maurer, H.P., and Reif, J.C. (2008) Hybrid maize breeding with doubled haploids. IV. Number versus size of crosses and importance of parental selection in two-stage selection for testcross performance. *Theor. Appl. Genet.* 117: 251-260.
- [14] Wegenast, T., Utz, H.F., Longin, C.F.H., Maurer, H.P., Dhillon, B.S., and Melchinger, A.E. (2010) Hybrid maize breeding with doubled haploids: V. Selection strategies for testcross performance with variable sizes of crosses and S1 families. *Theor. Appl. Genet.* 120: 699-708.
- [15] Mi, X., Wegenast T., Utz, H.F., Dhillon, B.S., and Melchinger, A.E. (2011) Best linear unbiased prediction (BLUP) and optimum allocation of test resources in maize breeding with doubled haploids. *Theor. Appl. Genet.* 123: 1-10.
- [16] Hallauer, A.R. and Miranda Filho, J.B. (1988). *Quantitative Genetics in Maize Breeding*. Iowa State University Press, Ames, IA.
- [17] Bernardo, R. (1994). Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci* 34: 20-25.
- [18] Bernardo, R. (1995). Genetic models for predicting maize single-cross performance in unbalanced yield trial data. *Crop Sci* 35: 141-147.
- [19] Piepho, H.P. (2009). Ridge regression and extensions for genomewide selection in maize. *Crop Sci* 49:1165-1176.

- [20] Melchinger, A.E. (1999). Genetic diversity and heterosis. p. 99-118. In: J.G. Coors, and S. Pandey (eds.). *The Genetics and Exploitation of Heterosis in Crops*. ASA - CSSA, Madison, WI.
- [21] Frisch, M., Thiemann, A., Fu, J., Schrag, T.A., Scholten, S., and Melchinger, A.E. (2010). Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. *Theor Appl Genet* 120: 441-450.
- [22] Vuylsteke, M., Kuiper, M., and Stam, P. (2000). Chromosomal regions involved in hybrid performance and heterosis: their AFLP (R)-based identification and practical use in prediction models. *Heredity* 85: 208-218.
- [23] Schrag, T.A., Melchinger, A.E., Sørensen, A.P., and Frisch, M. (2006). Prediction of single-cross hybrid performance for grain yield and grain dry matter content in maize using AFLP markers associated with QTL. *Theor Appl Genet* 113: 1037-1047.
- [24] Schrag, T.A., Maurer, H.P., Melchinger, A.E., Piepho, H.-P., Peleman, J., and Frisch, M. (2007). Prediction of single-cross hybrid performance in maize using haplotype blocks associated with QTL for grain yield. *Theor Appl Genet* 114: 1345-1355.
- [25] Schrag, T.A., Möhring, J., Maurer, H.P., Dhillon, B.S., Melchinger, A.E., Piepho, H.-P., Sørensen, A.P., and Frisch, M. (2009). Molecular marker-based prediction of hybrid performance in maize using unbalanced data from multiple experiments with factorial crosses. *Theor Appl Genet* 118: 741-751.
- [26] Schrag, T.A., Möhring, J., Melchinger, A.E., Kusterer, B., Dhillon, B.S., Piepho, H.-P., and Frisch, M. (2010). Prediction of hybrid performance in maize using molecular markers and joint analyses of hybrids and parental inbreds. *Theor Appl Genet* 120: 451-461.
- [27] Meuwissen, T.H.E., Hayes, B.J., and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819-1829.
- [28] www.synbreed.tum.de
- [29] Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer, L., Melchinger, A.E. (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics* 44: 217–220.
- [30] Technow, F., Schrag, T.A., Schipprack, W., Bauer, E., Simianer, H., Albrecht E. Melchinger, A.E. (2014) Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. *Genetics* 197: 1343-1355
- [31] Meyer, R.C., Steinfath, M., Lisec, J., Becher, M., Witucka-wall, H., Törjek, O., Fiehn, O., Eckardt, Ä., Willmitzer, L., Selbig, J., and Altmann, T. (2007). The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 104: 4759-4764.
- [32] Sulpice, R., Pyl, E.T., Ishihara, H., Trenkamp, S., Steinfath, M., Witucka-Wall, H., Gibon, Y., Usadel, B., Poree, F., Piques, M.C., von Korff, M., Steinhauser, M.C., Keurentjes, J.J.B., Guenther, M., Hoehne, M., Selbig, J., Fernie, A.R., Altmann, T., and Stitt, M. (2009). Starch as a major integrator in the regulation of plant growth. *Proc Natl Acad Sci USA* 106: 10348-10353.
- [33] Sulpice R., Trenkamp S., Steinfath M., Usadel B., Gibon Y. Witucka-Wall H., Pyl E.T., Tschoep H., Steinhauser M.C., Guenther M., Hoehne M., Rohwer J.M., Altmann T., Fernie A.R., and Stitt M. (2010). Network analysis of enzyme activities and metabolite levels and

their relationship to biomass in a large panel of Arabidopsis accessions. *Plant Cell* 22: 2872–2893.

[34] Gärtner, T., Steinfath, M., Andorf, S., Lisec, J., Meyer, R.C., Altmann, T., Willmitzer, L., and Selbig, J. (2009). Improved heterosis prediction by combining information on DNA- and metabolic markers. *PLoS ONE* 4: e5220.

[35] Steinfath, M., Gärtner, T., Lisec, J., Meyer, R.C., Altmann, T., Willmitzer, L., and Selbig, J. (2010). Prediction of hybrid biomass in *Arabidopsis thaliana* by selected parental SNP and metabolic markers. *Theor Appl Genet* 120: 239-247.

[36] Springer, N.M., Ying, K., Fu, Y., Ji, T., Yeh, C.-T., Jia, Y., Wu, W., Richmond, T., Kitzman, J., Rosenbaum, H., Iniguez, A.L., Barbazuk, W.B., Jeddloh, J.A., Nettleton, D., and Schnable, P.S. (2009). Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet* 5: e1000734.

[37] Beló, A., Beatty, M.K., Hondred, D., Fengler, K.A., Li, B., and Rafalski, A. (2010) Allelic genome structural variations in maize detected by array comparative genome hybridization. *Theor Appl Genet* 120: 355-367.

I.5. Zusammenarbeit mit anderen Stellen

Das Projektvorhaben wurde in enger Zusammenarbeit mit den anderen beteiligten Partnern im Verbundprojekt OPTIMAL durchgeführt. In regelmäßigen Intervallen fanden Projekttreffen statt, mindestens einmal pro Jahr am Standort eines Partners sowie am Rande der Plant2020 Statusseminare. Bei diesen Treffen wurde der Projektfortschritt besprochen, weiterführende Arbeiten geplant und Interaktionen mit anderen Partnern innerhalb und außerhalb des GABI-Programms diskutiert. Auf diese Weise konnten alle erforderlichen Informationen zwischen den Partnern ausgetauscht werden. Der bei P1 beschäftigte Postdoc absolvierte während der Projektlaufzeit einen Forschungsaufenthalt zur scNGS-Durchführung und -Datenanalyse an der Iowa State University .

II.1. Verwendung der Zuwendung und des erzielten Ergebnisses im Einzelnen

Von P4 selektierte und zur Verfügung gestellte Maisinzuchtlinien (25 Linien insgesamt) wurden in drei Vorversuchsserien zur Optimierung der kontrollierten Gewächshausbedingungen angezogen und vermessen. Dabei wurden ausgehend von den Standardanzuchtverfahren unterschiedliche Temperatur- und Beleuchtungsregime getestet und die Bedingungen gemäß der Beobachtungen folgendermaßen angepasst: 15/10°C Tag/Nachttemperatur während der ersten drei Wochen, dann 20/13°C Tag/Nachttemperatur während einer Woche und schließlich 25/18°C Tag/Nachttemperatur während zwei weiterer Wochen. In einer Versuchsreihe wurde zudem in einem Block mit Hilfe von Ventilatoren Wind simuliert, um die Auswirkungen der mechanischen Stimulation auf das Pflanzenwachstum zu testen. Das Frischgewicht der Pflanzen aus den angepassten Gewächshausbedingungen korrelierte signifikant ($r = 0.836$; $p > 0.01$) mit dem Frischgewicht von Pflanzen aus Feldversuchen (Tab. 1), und zwar deutlich höher als das Frischgewicht von Pflanzen, die unter Standardbedingungen angezogen wurden ($r = 0.641$; $p < 0.01$).

Windsimulation erhöhte die Korrelation zwischen angepassten Bedingungen und Feldversuchen noch leicht ($r = 0.900$; $p < 0.01$).

Die verbesserten Kultivierungsbedingungen wurden für die Phänotypisierungsexperimente noch etwas verändert, um vergleichbare Bedingungen für die Bestimmung der frühen Biomasse und Probennahme für Metabolitprofile (3-Blatt Stadium) zwischen Feld und Gewächshaus zu erreichen: 15/10°C Tag/Nacht Temperatur während vier Wochen, dann 20/13°C Tag/Nacht Temperatur während einer Woche und schließlich 25/18°C Tag/Nacht Temperatur für weitere zwei Wochen. Probennahmen für MP fanden 27-29 Tage nach Aussaat statt, also unmittelbar vor dem ersten Wechsel im Temperaturregime.

Ein Set von 198 Inzuchtlinien des Partners P4 wurde in zwei aufeinanderfolgenden Jahren unter den optimierten Kultivierungsbedingungen in der automatisierten IPK-Phänotypisierungsanlage für große Pflanzen (LemnaTec) angezogen (Abb. 1). Dabei wurden die Pflanzen täglich von oben und vier Seiten im sichtbaren Lichtbereich, Fluoreszenz und NIR aufgenommen. Die Hälfte der Pflanzen wurde nach 27-29 Tagen für MP beprobt und anschließend für (frühe) Frisch- und Trockengewichtsbestimmung geerntet, die andere Hälfte wurde am Ende des Experiments für (mittlere) Frisch- und Trockengewichtsbestimmung geerntet. Von diesen Linien hatte P4 auch Proben für MP von Feldversuchen genommen. Zusätzlich wurden auch 396 Hybride aus faktoriellen Testkreuzungen der Inzuchtlinien (von P4) unter den gleichen Bedingungen in zwei Jahren angezogen, beprobt, fotografiert und vermessen. Die Datenauswertung erfolgte mittels des im IPK entwickelten Bildanalyseprogramms IAP, das im Rahmen des OPTIMAL Projektes um einige Mais-spezifische Algorithmen erweitert wurde, z. B. die Erfassung des Blattrollens und die Analyse von Farbe und Fluoreszenz- und NIR-Intensität in unterschiedlichen Regionen der Pflanze.

Die 198 Inzuchtlinien zeigten eine hohe Variation in Biomasseproduktion innerhalb der Population (Abb. 2a), mit einer hohen Korrelation ($r = 0.899$, $p < 0,001$) zwischen gewogenem Frischgewicht und nach Bildanalyse geschätztem Frischgewicht (Abb. 2b). Mit ca. 100 hoch reproduzierbaren digitalen Bildanalyse-Merkmalen konnte zwischen Flint- und Dent-Linien unterschieden werden, mit nur wenigen Überschneidungen (Abb. 3). ‚Deep Phenotyping‘ bietet also die Möglichkeit, unser Verständnis der phänotypischen Variation zu steigern, da Phänotypen sehr viel detaillierter erfasst werden können, und dadurch eine präzisere Kopplung zwischen genomischen und phänotypischen Datensätzen erreicht werden kann. Die phänotypischen Daten der Inzuchtlinien und der Hybriden wurden P7 für die Entwicklung und Testung der Vorhersage-Modelle zur Verfügung gestellt.

Das tiefgefrorene Blattmaterial aller Inzuchtlinien wurde in einem automatisierten Cryogrinder gemahlen und aliquotiert. Metabolitextrakte wurden manuell erzeugt und mittels GC-MS mit vorgeschalteter in-line Derivatisierung mit Autosampler vermessen. Von jeder Linie wurden vier Replikate analysiert. Insgesamt wurden 469 polare Metabolite quantifiziert, davon 134 Metabolite mit bekannter chemischer Struktur. Die Datensätze bestätigen die gute Reproduzierbarkeit zwischen Experimenten und Replikaten. Die Variation zwischen Linien mit extremen Werten für Metabolitgehalt bewegte sich innerhalb von 1-2 Größenordnungen (Abb. 4). Die Metabolitdaten wurden gemäß des Experimentaufbaus normalisiert und P7 (UP) für die Entwicklung und Testung der Vorhersage-Modelle übermittelt.

Während der Laufzeit des Projektes wurde scNGS als alternative Methode zur Entdeckung von CNV und vor allem PAV entwickelt. Eigene Vorversuche in Kollaboration mit Pat

Schnables Labor an der Iowa State University (ISU) ergaben, dass aCGH zwar robust CNV in gut charakterisierten Genomen erkannte, aber Deletionen im Referenzgenom (erwartungsgemäß) nicht detektieren konnte, und auch offensichtliche Schwächen bei der Erkennung von CNV in schlecht charakterisierten/assemblierten Regionen des Referenzgenoms zeigte. scNGS umgeht diese Schwierigkeiten und lieferte in den Vorversuchen präzisere und umfangreichere Identifizierungen von CNV und PAV. Daher wurde in Abstimmung mit allen beteiligten Partnern beschlossen dieses neue Verfahren anzuwenden und es wurden insgesamt vom OPTIMAL-Konsortium DNA-Proben von 396 Inzuchtlinien (IPK=201, LMG=95, KWS=100) isoliert und gemeinsam mit dem Unterauftragnehmer (ISU) scNGS-Analysen durchgeführt. Diese Linien waren auch mit dem SNP50 Array genotypisiert worden. Der bei P1 beschäftigte Postdoc absolvierte einen Forschungsaufenthalt an der ISU und begleitete die dortigen Arbeiten vor Ort. Für die Anreicherung exonischer DNA wurde ein „Zeanome Capture“ entwickelt, das 108 Mb Exon-Sequenzen abdeckte (probe design space), davon 60 Mb B73 Exom und 48 Mb andere Mais Exom-Sequenzen (z.B. US NAM Elternlinien und Teosinte) (Abb. 5). Jede Linie wurde zwei Capture-Verfahren unterzogen und zwei unabhängigen Sequenzier-Reaktionen unterzogen (Illumina HiSeq 2000; 2 x 100 bp paired-end). Im Durchschnitt wurden 2 x 35 Millionen Reads pro Inzuchtlinie erhalten. Nach dem Durchlaufen eines komplexen Sequenzanalyseverfahrens wurden mit diesen Datensätzen insgesamt 30.228 Konsensus-PAV Segmente einer Länge von mindestens 300 bp und 252,4 Millionen SNPs detektiert, pro Inzuchtlinie im Durchschnitt 1,25 Mio. SNPs (Abb. 5).

Wie sämtliche erhobenen phänotypischen und metabolischen Daten wurden auch die genomischen Daten nach entsprechender Voranalyse an Partner 7 (UP) weitergegeben und dort für das Optimieren, Trainieren und Evaluieren von Hybridleistungsvorhersagemodellen herangezogen. Die Daten stehen ferner für weiterführende Analysen, insbesondere für Korrelations- und Assoziationsstudien zur Verfügung, mit denen Bezüge zwischen den verschiedenen Ebenen der Merkmalsausprägung (Metabolite – Ganzpflanzenmerkmale) und zwischen Erbgut und Merkmalsausprägung (mittels GWAS) detektiert werden können.

Tabelle 1: Frischgewichtskorrelationen von 25 Maisinzuchtlinien unter verschiedenen Wachstumsbedingungen

	1	2	3	4	5	6
Frühe Biomasse GWH (1)						
Mittlere Biomasse GWH (2)	0,658**					
Frühe Biomasse GWH_WS (3)	0,802**	0,766**				
Mittlere Biomasse GWH_WS (4)	0,557**	0,848**	0,752**			
Frühe Biomasse_Feld (5)	0,836**	0,748**	0,900**	0,669**		
Finale Biomasse_Feld (6)	0,696**	0,668**	0,613**	0,538**	0,769**	
Frühe Biomasse_GWH_Normal (7)	0,513*	0,733**	0,684**	0,648**	0,641**	0,458*

**p < 0,01 , * P < 0,05 (Signifikanzniveau)

`GWH`: in OPTIMAL angepasste Bedingungen im Gewächshaus, `WS`: Windsimulation mit Ventilatoren, `Feld`: Frischgewicht unter Feldbedingungen (Daten von P4, UHOH), `GWH-Normal`: Standard Gewächshausbedingungen (25/20°C Tag/Nacht Temperatur während der gesamten Wachstumsphase).



Abbildung 1: Phänotypisierung der 198 Maiszuchtlinien in 2 Blöcken mit je 4 Replikaten in automatisierten Phänotypisierungsanlage für große Pflanzen. Die Töpfe werden über ein Förderbandsystem automatisch zu den drei Aufnahmestationen (Reflexion von sichtbaren und nahe-Infraroten Licht, Fluoreszenz), der Waage und der Bewässerungsstation gefahren.

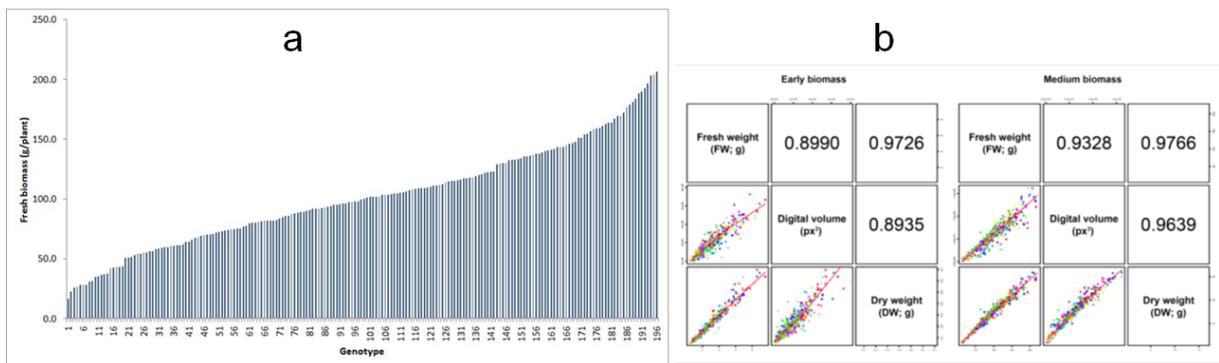


Abbildung 2: (a) Verteilung des Frischgewichts der 198 Maiszuchtlinien und (b) Pearson Korrelationen zwischen gemessenem und nach Bildanalyse geschätztem Frischgewicht ($p < 0,001$).

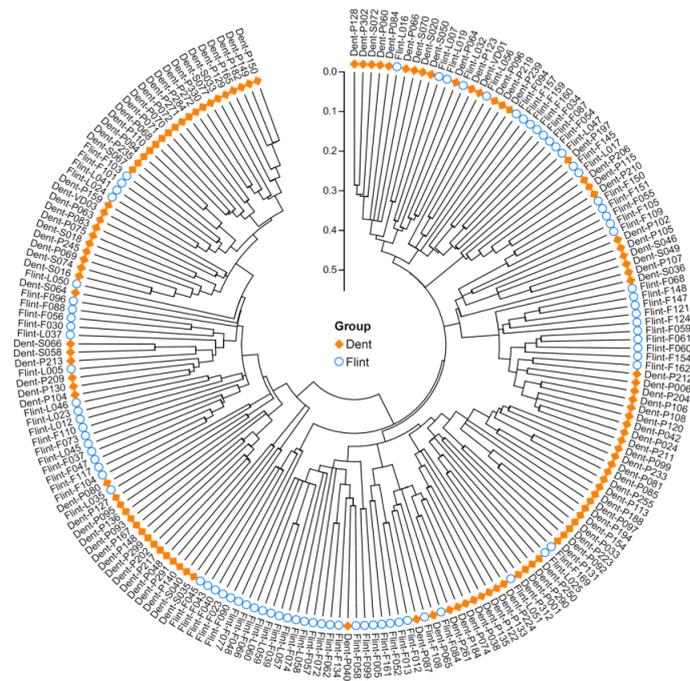


Abbildung 3: Phylophänomischer Baum konstruiert unter Verwendung von 100 hoch reproduzierbaren ($r > 0,7$; $p < 0,001$) digitalen Bildanalyse-Merkmale der drei Aufnahmetypen (VIS, NIR, Fluoreszenz).

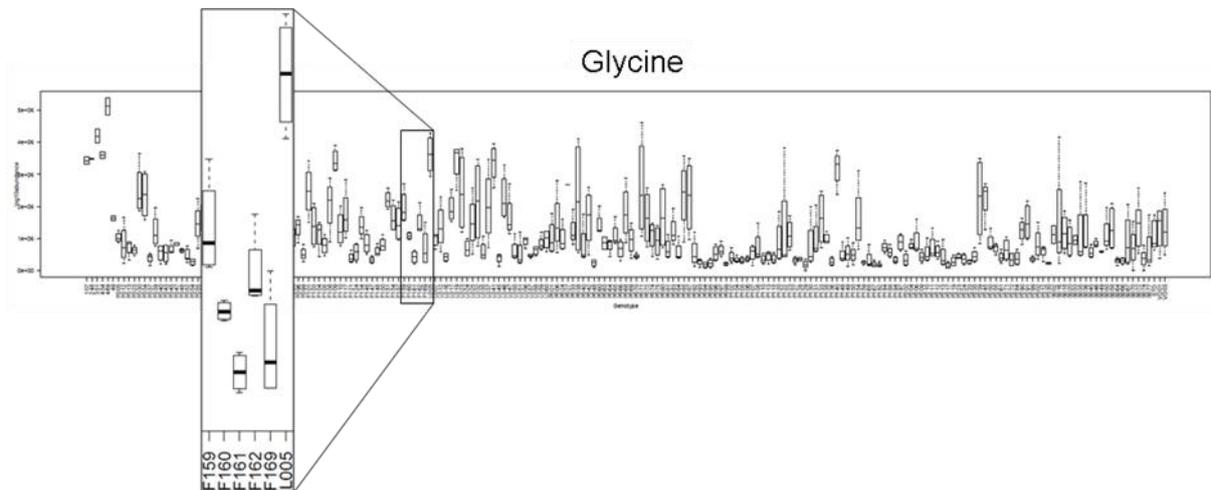


Abbildung 4. Variation des Glycingehalts in 198 Maisinzuchtlinien, stellvertretend für die 469 im Blatt gemessenen Metabolite ($n = 4$)

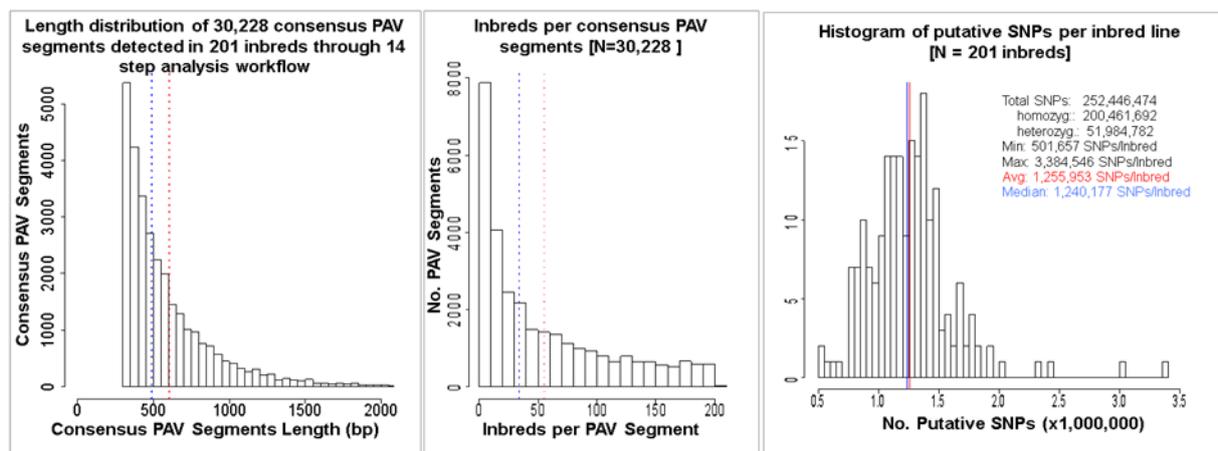


Abbildung 5: Ergebnisse von Zeanome Sequence Capture und Illumina Sequenzierung der genomischen DNA von 201 Maisinzuchtlinien. Links und Mitte: PAV; Rechts: SNP Detektierung. Die Analyse wurde zeitgleich auch mit 100 KWS und 95 LGE Linien durchgeführt.

II.2. Die wichtigsten Positionen des zahlenmäßigen Nachweises

- Mittels drei Versuchsserien an 25 Inzuchtlinien wurden Optimierungen der kontrollierten Wachstumsbedingungen vorgenommen, die zu einer sehr substantziellen Steigerung der Übereinstimmung mit entsprechenden Felddaten führte (von $r = 0,641$ auf $r = 0,900$ für frühe Biomasse)

- 198 Maisinzuchtlinien und 396 Maishybride wurden jeweils in zwei unabhängigen Experimenten unter den optimierten Kultivierungsbedingungen im automatisierten Phänotypisierungssystem angezogen und einer detaillierten Phänotypisierung unterzogen (für die Inzuchtlinien insgesamt 4 Wiederholungen à 4 Individuen; für die Hybriden insgesamt 2 Wiederholungen à 4 Individuen), bei denen jeweils Messwerte für ca. 100 hoch reproduzierbare aus Bilddaten extrahierte Merkmale erhoben wurden; ferner wurden jeweils die früh (nach 4 Wochen) und bis zur mittleren Entwicklungsphase (nach 7 Wochen) akkumulierte Pflanzenbiomasse (Frisch- und Trockengewichte) bestimmt
- Von den 198 genotypisierten und phänotypisierten Inzuchtlinien wurden von je 4 Individuen Blattproben genommen und GC-MS basierte Metabolitprofile erstellt (469 polare Substanzen davon 134 mit bekannten chemischer Struktur); analog wurden von den 396 phänotypisierten Hybriden von je 2 Individuen Blattproben genommen und für zukünftige Analysen gelagert.
- Von den 198 genotypisierten und phänotypisierten Inzuchtlinien (UH-Kollektion) wurden scNGS basierte PAV- und CNV Profile erstellt (2 Replikate); im Durchschnitt wurden 2 x 35 Millionen Reads pro Inzuchtlinie erhalten, mit denen insgesamt 30.228 Konsensus-PAV Segmente einer Länge von mindestens 300 bp abgeleitet wurden und 252,4 Millionen SNPs detektiert wurden (pro Inzuchtlinie im Durchschnitt 1,25 Mio. SNPs).

II.3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die in diesem Projekt ausgeführten Arbeiten entsprechen in vollem Umfang den Zielstellungen des Projektes und der der Zuwendung zugrunde liegenden Aufgabenstellung. Es wurden alle experimentellen Arbeiten in der vorgesehenen und für das Erreichen der Projektziele erforderlichen Art und Weise ausgeführt und die erhobenen Primärdaten (vor-)verarbeitet und den Projektpartnern für die geplante integrierte Analyse bereitgestellt. Auf diese Weise wurden grundlegend neuartige und in ihrer Tiefe massiv über bisher verfügbare Datensätze hinausgehende Informationen an Mais-Zuchtmaterial des öffentlichen Zuchtprogramms der Universität Hohenheim gewonnen. Diese Daten umfassen sowohl einzigartige phänotypische Daten, sehr umfangreiche Metabolitprofilaten und neuartige Genotypdaten (PAV-Informationen in Kombination mit wesentlich vertieften SNP-Daten). Diese Datenbasis stellte eine entscheidende und notwendige Grundlage für die durch Partner 7 (UP) erfolgreich ausgeführte Weiterentwicklung und Verbesserung von Vorhersagemodellen für die Hybridleistung dar. Somit konnten auf Basis dieser Daten und in Zusammenarbeit mit den anderen Projektpartnern entscheidende neue Informationen gewonnen werden und neue Verfahrensweisen entwickelt werden, die den Wirtschaftspartnern zur Kenntnis gebracht wurden. Diese können somit unmittelbar in das Repertoire der praktischen Hybridzuchtverfahren aufgenommen und angewendet werden. Über diesen für die Mais-Hybridzüchtung unmittelbar relevanten Erkenntnisgewinn hinaus wurde mit den erhobenen Daten eine sehr wertvolle Basis für weitergehende Untersuchungen der Zusammenhänge zwischen Erbgutvariation und der Ausprägung phänotypischer Ganzpflanzenmerkmale und biochemischer Merkmale (Metabolitgehalte) geschaffen. Die in

diesem Projekt geleisteten Arbeiten und die damit gewonnenen Daten und Erkenntnisse ermöglichen eine Steigerung und Stabilisierung der Ertragsbildung der Kulturpflanze Mais, indem eine effizientere züchterische Nutzung der natürlichen genetischen Diversität zur Erzeugung von Hochleistungshybriden erreicht werden kann. Die erreichten Projektergebnisse entsprechen damit in vollem Umfang den Zielen des Förderprogramms.

II.4. Voraussichtlicher Nutzen (Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplanes)

In diesem Vorhaben wurde entscheidend dazu beigetragen sehr wertvolle Erkenntnisse über die Möglichkeiten der Vorhersage der Leistung von Mais-Hybriden, insbesondere deren Ertragsbildung, auf der Basis von Daten der Elternlinien gewonnen. Die Kenntnisse betreffen den Wert und die Aussagekraft unterschiedlicher Datentypen (genetische Marker: SNPs, PAVs; Bio(chemische) Marker: Metabolitprofile; und phänotypische Marker/Ausprägungen). Diese Kenntnisse und darauf aufbauende neue Züchtungsverfahren können unmittelbar in praktische Züchtungsprogramme aufgenommen werden, da sie an aktuellem Zuchtmaterial (des Partners 4, UH) gewonnen wurden, den als Projektpartner beteiligten Pflanzenzüchtungsunternehmen unmittelbar zur Kenntnis gebracht wurden und diese Wirtschaftspartner im Rahmen des Projektes in die Lage versetzt wurden, vergleichende Untersuchungen an ihrem eigenen proprietären Zuchtmaterial durchzuführen. Sie helfen die Auswahl und Gestaltung der wirksamsten und kosteneffizientesten Strategien zu treffen und unterstützen die Planung und Umsetzung entsprechender Zuchtprogramme und experimenteller Designs mit denen aufwändige phänotypische Selektionen ggf. reduziert werden können. Damit ist eine unmittelbare Prüfung der Wirtschaftlichkeit (im Sinne einer Zeit- und/oder Kostenersparnis) einer Anwendung der neuartigen Verfahrensweisen ermöglicht und es kann ein möglicher Weiterentwicklungsbedarf abgeschätzt werden.

Neben der möglichen unmittelbaren wirtschaftlichen Nutzung der in diesem Projekt gewonnenen Ergebnisse stellen die sehr umfangreichen gesammelten Daten eine höchst wertvolle Basis für den weiterführenden wissenschaftlichen Erkenntnisgewinn dar: Mit Hilfe der genetischen Markerdaten (inklusive der neuartigen und weitreichenden PAV-Informationen) und den erhobenen biochemischen und phänotypischen Daten lassen sich vielfältige ursächliche Beziehungen zwischen Erbgutvariation und der Merkmalsausprägung herleiten, die zu einer weitergehenden Funktionszuweisung von Genen und der Bedeutung bestimmter Elemente im Maisgenom führen können. Die erstellte Datenbasis eröffnet damit weitreichende Möglichkeiten der wissenschaftlichen Verwertung und der Nutzung für weiterführende Untersuchungen.

II.5. Während der Durchführung des Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Während der Projektlaufzeit wurde nicht bekannt, dass an anderer Stelle ein in der Komplexität und Multidimensionalität (von der Detektion von PAV/CNV auf Genomebene, über Metabolit- und Phänotypvariation im klimatisierten Gewächshaus und im Feld) und der gegebenen Vielfalt der möglichen und geprüften Vorhersagemodelle vergleichbares Projekt

durchgeführt worden wäre. Für bestimmte Teilbereiche wurden jedoch konzeptionell und inhaltlich vergleichbare bzw. ähnliche Untersuchungen für verschiedene Pflanzenarten ausgeführt und publiziert. So wurde eine Hybridleistungsvorhersage bei Mais auf Basis von mRNA-Profilen der 21 Elternlinien von 98 Hybriden mit ähnlicher Präzision erreicht, wie mit genetischen (AFLP-) Markern (Zenke-Philippi et al. (2016) *BMC Genomics* 17:262); es wurden signifikante Hybridleistungsvorhersagen auf der Basis von Metabolitprofilen von Weizenlinien (135 Elternlinien von 1604 Weizenhybriden) erzielt, die allerdings niedrigere Präzision aufwiesen, als die auf Basis von genetischen Markerdaten (90k SNPs), und trotz ihrer Unabhängigkeit nicht zu einer Verbesserung der Genotyp-basierten Vorhersage führten (Zhao et al. (2015) *Proc. Natl. Acad. Sci. USA* 112:16624-15629); und es wurde für eine Kollektion von 306 Reishybriden eine Vorhersagegenauigkeit der Hybridleistung auf Basis von Metabolitprofilen der 18 elterlichen Inzuchtlinien erzielt (Dan et al. (2016) *Sci. Reports* 6:21732 doi: 10.1038/srep21732). Im Jahr 2015 wurde eine in den USA vorgenommene Patentanmeldung (Anmeldenummer: US 12/344,887, Veröffentlichungsnummer: US8996318 B2) erteilt und eine weitere Patentanmeldung (Anmeldenummer: US 14/664,590, Veröffentlichungsnummer: US20150197821 A1) veröffentlicht, die sich auf ein Hybridleistungsvorhersageverfahren bei Pflanzen unter Anwendung vergleichender genomischer Hybridisierung zur Detektion von Kopiezahlvariation bezieht.

II.6. Erfolgte oder geplante Veröffentlichungen des Ergebnisses

Ein Teil der erzielten Ergebnisse wurde bereits in internationalen wissenschaftlichen Fachzeitschriften publiziert:

Castellini, A., C. Edlich-Muth, M. Muraya, C. Klukas, T. Altmann & J. Selbig (2015) Towards a graph-theoretic approach to hybrid performance prediction from large-scale phenotypic data. In *M. Lones, A. Tyrrell, S. Smith, G. Fogel (eds) Information Processing in Cells and Tissues: 10th International Conference, IPCAT 2015, San Diego, CA, USA, September 14-16, 2015, Proceedings* (Vol. 9303, pp. 173-184). Springer.

Junker, A., M.M. Muraya, K. Weigelt-Fischer, F.A. Arana-Ceballos, C. Klukas, A.E. Melchinger, R.C. Meyer, D. Riewe & T. Altmann (2015) Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Front. Plant Sci.* 5: 770.

Edlich-Muth, C., Castellini, A., Westhues, M., Muraya, M., Riewe, D., Schlereth, A., Schrag, T., Cuadros Inostroza, A., Melchinger, A., Willmitzer, L., Stitt, M., Altmann, T., Salbig, J., and Nikoloski, Z. (2016) Metabolic profiles and genetic support lead to highly predictive and generalizable models of hybrid performance in maize. *Proc. Natl. Acad. Sci. USA*, eingereicht.

Weitere Veröffentlichungen zu den für die Mais-Inzuchtlinien erhaltenen PAV/CNV-Ergebnissen, zu Genom-weiten Marker-Merkmal-Assoziationen für biochemische Merkmale (Metabolit-Profilen) und Ganzpflanzenmerkmale (Phänotyp-Profilen), sowie zu genetischen Faktoren, die die dynamischen Reaktionen der Pflanzen auf wechselnde Umweltbedingungen bestimmen sind in Vorbereitung.

IV. Kurzfassung (Berichtsblatt; Document Control Sheet)

1. ISBN oder ISSN	2. Berichtsart (Schlussbericht oder Veröffentlichung) Schlussbericht
3. Titel Forschungsvorhaben Teilprojekt: 0315958A Thema: Verbundprojekt OPTIMAL: Auf genetischen Markern und Biomarkern basierte prädikative Züchtung von Mais-Kulturvarietäten	
4. Autor(en) [Name(n), Vorname(n)] Meyer, Rhonda C. Altmann, Thomas	5. Abschlussdatum des Vorhabens 30.06.2015
	6. Veröffentlichungsdatum
	7. Form der Publikation Bericht
8. Durchführende Institution(en) (Name, Adresse) Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) Gatersleben, Corrensstraße 3, 06466 Stadt Seeland OT Gatersleben	9. Ber. Nr. Durchführende Institution
	10. Förderkennzeichen 0315958A
	11. Seitenzahl 17
12. Fördernde Institution (Name, Adresse) Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	13. Literaturangaben 44
	14. Tabellen 1
	15. Abbildungen 5
16. Zusätzliche Angaben	
17. Vorgelegt bei (Titel, Ort, Datum)	
18. Kurzfassung Ziel des Verbundvorhabens OPTIMAL war die Entwicklung und Optimierung eines neuartigen Hybridzuchtansatzes, der DNA Polymorphismen als genetische Marker mit Metabolitprofilen (MPs) und Ganzpflanzenmerkmale als Biomarker kombiniert, um Hybridleistung vorherzusagen. Neben der Koordination des OPTIMAL-Gesamtverbundes bestand die Aufgabenstellung des Teilprojektes A in der Erhebung entscheidender umfangreicher Datensätze an den vom Projektpartner 4 (UH) zur Verfügung gestellten Inzuchtlinien und Hybriden und die Bereitstellung dieser Daten für Partner 7 (UP) für die kombinierte Analyse zusammen mit den von den Projektpartnern 5 (MPI-MP-LW) und 6 (MPI-MP-MS) an den selben Linien gesammelten Daten. Nach der Optimierung der kontrollierten Wachstumsbedingungen wurde für 198 genotypisierte Maisinzuchtlinien und für 396 Maishybride mit Hilfe eines automatisierten Kultivierungs- und Phänotypisierungssystem des IPK umfangreichste Phänotypdaten erhoben. Ferner wurden jeweils die früh (nach 4 Wochen) und bis zur mittleren Entwicklungsphase (nach 7 Wochen) akkumulierte die Pflanzenbiomasse (Frisch- und Trockengewichte) bestimmt. Für die 198 Inzuchtlinien wurden zudem GC-MS basierte Metabolitprofile erstellt und es wurden für diese Linien weiterhin scNGS basierte PAV- und CNV Profile erhoben, mit denen insgesamt 30.228 Konsensus-PAV Segmente einer Länge von mindestens 300 bp abgeleitet und 252,4 Millionen SNPs detektiert wurden. Nach entsprechender Voranalyse wurden sämtliche erhobenen phänotypischen, metabolischen und genomischen Daten durch Partner 7 (UP) erfolgreich für das Optimieren, Trainieren und Evaluieren von Hybridleistungsvorhersagemodellen herangezogen. Somit konnten in Zusammenarbeit mit den anderen Projektpartnern entscheidende neue Informationen gewonnen werden und neue Verfahrensweisen für die Vorhersage der Hybridleistung auf Basis der Daten von Elternlinien entwickelt werden, die den Wirtschaftspartnern zur Kenntnis gebracht wurden. Ferner stehen die in diesem Projekt erhobenen Daten zusammen mit den aus Vorarbeiten vorhandenen Daten für weiterführende Untersuchungen, insbesondere für Korrelations- und Assoziationsstudien, zur Verfügung.	

19. Schlagwörter

Mais, Hybridleistung, Vorhersage, Phänotypisierung, Metabolitprofile, Sequenz-An-/Abwesenheits-Polymorphismen

20. Verlag

21. Preis