

Veröffentlichung der Ergebnisse von Forschungsvorhaben im BMBF-Programm

B I O L O G I E

Forschungsvorhaben: GABI-FUTURE - Verbundvorhaben: 'Nutzung genomweiter natürlicher Diversität in der Arabidopsis Population (GABI-GNADE)' (Teilvorhaben F)

Förderkennzeichen: 0315060F

Zuwendungsempfänger: Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstr. 3, 06466 Gatersleben

Ausführende Stelle: Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstr. 3, 06466 Gatersleben

Projektleitung: Herr Prof. Dr. Altmann

Laufzeit: 01.04.2008 bis 31.03.2011

"Das diesem Bericht zugrundeliegende BMBF-Forschungsvorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 0315060F gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor".

Forschungsvorhaben Teilprojekt: 0315060F

Thema: GABI-FUTURE-Verbundvorhaben: `Nutzung genomweiter natürlicher Diversität in der Arabidopsis Population (GABI-GNADE)` (Teilvorhaben F)

Schlussbericht zu Nr. 3.2

I.1. Aufgabenstellung

Im Verbundvorhaben GABI-GNADE sollten die Vorzüge von *Arabidopsis thaliana* genutzt werden, um die rasanten technischen Weiterentwicklungen und die sich damit eröffnenden neuen Strategien für die Untersuchung und die Nutzung natürlicher genetischer Diversität zu testen. Für dieses molekulargenetische Referenzsystem der Samenpflanzen ist gegenüber allen anderen Pflanzenarten der stärkste Informationszuwachs zu verzeichnen und es bietet durch die kurze Generationszeit und den geringsten logistischen Aufwand für experimentelle Untersuchungen beste Voraussetzungen, um im Vorlauf von Studien an Kulturpflanzen neue technische Verfahren zu prüfen. Natürliche Diversität ist die zentrale Ressource für die Pflanzenzüchtung. Sie stellt gleichzeitig eine sehr wichtige Basis für die Grundlagenforschung dar, da sie es ermöglicht, die an der Merkmalsvariation beteiligte genetische Kontrolle und die involvierten Interaktionen in komplexen Netzwerken zu detektieren. Die zu erheblicher Verbesserung der Nutzungs- und Untersuchungsmöglichkeiten führenden Fortschritte sind (i) stark wachsende Informationen zur Genomsequenzvariation, (ii) qualitative Sprünge in den Technologien für die Genotypisierung segregierender Populationen und zur Erzeugung reinerbiger Inzuchtlinien, (iii) existierende und entstehende Plattformen für die Hochdurchsatzphänotypisierung molekularer, biochemischer, physiologischer, morphologischer und agronomischer Merkmale und (iv) eine massive Steigerung in der Leistungsfähigkeit bioinformatischer und statistischer Analyseverfahren. Diese methodischen Entwicklungen werden in Kürze starke Veränderungen in der Pflanzenzüchtung ermöglichen.

Die Gesamtziele des Verbundprojektes waren:

- Die Resequenzierung und Assemblierung der Genome von zwei hoch diversen Arabidopsis-Linien unter Anwendung neuester Sequenzieretechnologien und Bioinformatikverfahren. Damit sollten wesentliche Fortschritte im Verständnis natürlicher Variation auf der Genom- und Arten-weiten Ebene erzielt werden.
- Die Entwicklung neuer Strategien für die Kombination von Datensätzen aus der Assoziationskartierung und Inzuchtkartierungspopulationen mit tiefen Sequenzinformationen und Profilanalysetechniken zur Unterstützung von Feinkartierungs- und Genklonierungsverfahren.
- Die Klonierung einiger Haupt-QTL (Quantitative Trait Loci) für Wachstums-bezogene Merkmale als Machbarkeitsstudie.

Mit diesem Projekt sollten wesentliche Fortschritte in der Erstellung wichtiger Ressourcen und grundlegender Kenntnisse im Referenzsystem Arabidopsis erzielt werden, eine schnellere

und effizientere Nutzung tiefer Sequenzinformationen ermöglichen, sobald diese für Kulturpflanzen verfügbar werden.

Die Aufgaben des Teilvorhabens F bestanden darin,

- an einer Meta-QTL-Analyse und QTL-Kreuzvalidierung von Wachstums- / Biomasse-QTL mitzuwirken, die zuvor in diversen Populationen der Partner detektiert worden waren,

- die Analyse von QTL-Regionen weiterzuführen, die mittels Assoziationskartierung bzw. Koppelungskartierung bestimmt worden waren, und funktionell relevante Polymorphismen zu identifizieren, sowie deren Segregation in heterogenen Inzuchtfamilien (HIFs) oder nahisogenen Linien (NIL) zu untersuchen,

und

- funktionelle Analysen von Kandidatengenen für Wachstums-/Biomasse oder metabolische QTL auszuführen.

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. Aus Vorarbeiten bestand Zugang zu fortgeschrittenen Populationen für die funktionelle Analyse natürlicher Diversität bei Arabidopsis, darunter fünf Populationen rekombinanter Inzuchtlinien (RILs) reziproke Sätze von 63 bzw. 80 NILs der Akzessionen Col-0 und C24, sowie einer Assoziationskartierungspopulation von 178 diversen Arabidopsis-Akzessionen, die partiell genotypisiert war und für die in Zusammenarbeit mit dem Partner MPI-MP phänotypische Daten gesammelt worden waren. Weiterhin wurde eine umfangreiche Kollektion von 355 Arabidopsis Akzessionen (J. Borevitz, M. Nordborg and M. Quint), genotypisiert mit dem Affymetrix 250k SNP Array, für die Arbeiten verwendet. Ferner waren vorläufige Daten zu QTL für Wachstum/Biomasse bzw. Metabolitgehalte aus Analysen der RIL- und Assoziationspopulationen verfügbar, aus denen zudem Hinweise auf Zusammenhänge zwischen Metabolitgehalten und Wachstumseigenschaften vorlagen.

I.3. Planung und Ablauf des Vorhabens

Der Ablauf des Vorhabens entsprach weitgehend dem ursprünglichen Arbeitsplan. Im Folgenden sind die Arbeiten des Verbundprojektes dargestellt, zu denen im Rahmen des Teilprojektes F beigetragen wurde.

1. Zusammenstellung der umfangreichen Kollektion von 355 und einer weiteren Kollektion von 178 verschiedenen, von natürlichen Akzessionen abgeleiteten und genotypisierten Arabidopsis-Linien für Assoziationsstudien.
2. Phänotypische Analyse (Biomasse, Blattfläche) der selektierten Akzessionen und Ausführung von Assoziationstest zur Bestätigung der in der Col-0/C24 RIL-Population bestimmten Biomasse QTL
3. Analyse und Feinkartierung der durch Assoziationsstudien bestimmten QTL-Regionen
4. Identifizierung funktionell relevanter Polymorphismen zwischen den Genotypen Col-0 und C24 basierend auf den re-sequenzierten neuen C24-Daten

5. Kultivierung und Analyse von Nachkommen heterogener RILs und NILs (HIFs) selektierter Biomasse QTL Bereiche
6. Phänotypisierung und Genotypisierung der ausgewählten HIFs mit dem MASC-Markersatz (Törjek et al., 2003) zur Feinkartierung bestimmter QTL-Bereiche
7. QTL-Analyse mit Hilfe der genetischen und phänotypischen Daten der HIF-Populationen
8. Funktionelle Analyse von selektierten Kandidatengen der Biomasse und Metabolit QTL Bereiche

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

Zu Beginn des Projektes waren vergleichende Genom-Resequenzierungsarbeiten an 20 *A. thaliana* – Akzessionen bei Partner MPID im Gange, für die ein Array-basiertes Verfahren (1) herangezogen wurde und durch die eine hohe Anzahl Einzelnukleotidpolymorphismen (SNPs) identifiziert werden konnten. Die aufkommenden neuen Technologien der Sequenzierung durch Synthese (SBS), als erste die 454-Technik der Hochdurchsatz-Pyrosequenzierung (2) deuteten die Möglichkeit der Erhebung weit umfangreicher Sequenzdaten zu weiter sinkenden Kosten an.

Als Pflanzengenetische Ressourcen standen umfangreiche Populationen natürlicher *Arabidopsis*-Akkzessionen zur Verfügung, die aus vielen verschiedenen Habitaten stammten und entsprechende Adaptationen aufwiesen (3). Ferner waren RILs und NILs intensiv für die Identifizierung von QTL genutzt worden (z.B. 4-6) und es war in einigen Fällen die Feinkartierung und Klonierung der zugrundeliegenden Gene gelungen (z.B. 7-10), wobei es sich hier um einzelne erfolgreiche Beispiele handelte und mit den damaligen Verfahren keine systematische Abdeckung der Spezies-weiten genetischen Diversität möglich war.

Es lagen erste Ergebnisse zur Anwendung der in der Humangenetik zur Identifizierung von Krankheitsgenen genutzten Assoziationskartierung (11-14) in Pflanzen vor (15-29) und es gab Hinweise auf notwendige Informationen (30, 31) sowie Zugang zu diversen für diese Zwecke anwendbare Verfahrensweisen (32-34). Außerdem bestanden erste Kenntnisse über das Ausmaß des Koppelungsungleichgewichtes in Pflanzen und der Konsequenzen für die erforderliche Markerdichte vor (35-40). Die aufkommenden Techniken für Genom-weite Hochdichte-Markeranalysen ließen Assoziationsstudien in Pflanzengenetik und –züchtung machbar werden (41-47). Allerdings wiesen Daten zu transgressiver Segregation (48-66) darauf hin, dass in Inzuchtlinien vielfältige kompensatorische Effekte möglicherweise die Variationsbreite einschränken.

Neben hauptsächlich im privaten Sektor eingerichteten Plattformen für die Ganzpflanzenphänotypisierung (z.B. http://www.cropdesign.com/tech_traitmill.php) waren erste molekulare und biochemische Profilanalysetechniken (67, 68) für die Identifizierung und Klonierung von Metabolit-QTL (7, 69, 70) eingesetzt worden.

Zu dem im Vordergrund der Untersuchungen stehenden physiologischen Merkmal, dem vegetativen Wachstum, bestanden einerseits Informationen zur Bedeutung der Regulation von Zellteilung (71), Zellexpansion (72) und Meristemgröße (73, 74), andererseits Hinweise auf den Einfluss metabolischer Prozesse und der vorteilhaften Allokation von Ressourcen für die Ausprägung günstiger Pflanzenarchitekturen (75-78) und erste Hinweise auf integrierende, wachstumsregulierende Faktoren (79, 80).

Literatur: (1) Hinds et al, 2005 Science 307, 1072; (2) Margulies et al, 2005 Nature 437, 376 ; (3) Koorneef et al, 2004 Annu Rev Plant Biol 55, 141; (4) Eshed and Zamir, 1995 Genetics 1147; (5) Loudet et al, 2005 Theor Appl Genet 110: 742; (6) Reymond et al, 2006 Plant, Cell Environ 29, 115; (7) Fridman et al, 2004 Science 305, 1786; (8) Kroymann and Mitchell-Olds, 2005 Nature 435, 95; (9) Werner et al, 2005 PNAS 102, 2460; (10) Alonso-Blanco et al., 2005 Plant Physiol 1304; (11) Buckler and Thornsberry, 2002 Curr Op Plant Biol 5, 107; (12) Risch, 2000 Nature 405, 847; (13) Jorde, 2000 Genet Res 10, 1435; (14) McCarthy et al, 2001 Genomics 78, 135; (15) Nordborg and Tavaré, 2002 Trends Genet 18, 83; (16) Gupta et al, 2005 Plant Mol Biol 57, 461; (17) Yu and Buckler, 2006 Curr Op Biotechnol 17, 155; (18) Thornsberry et al, 2001 Nat Genet 28,286; (19) Hansen et al, 2001 Genet Res 77, 61; (20) Wilson et al, 2004 Plant Cell 16, 2719-2733; (21) Szalma et al, 2005 Theor Appl Genet 110, 1324; (22) Li et al, 2005 Genetics 170, 813; (23) Ivandic et al, 2003 Plant Breeding 122, 300; (24) Gebhardt et al, 2004 Molecular Breeding 13, 93; (25) Simko et al, 2004 Mol Gen Genom 271, 522; (26) Kraakman et al, 2004 Genetics 168, 435; (27) Skot et al, 2005 Molecular Breeding 15, 233; (28) Breseghello et al, 2006 Genetics 172, 1165; (29) Aranzana et al, 2005 PloS Genetics 1, 531; (30) Pritchard 2001 Nat Genet 28, 203; (31) Flint-Garcia et al, 2003 Annu Rev Plant Biol 54, 357; (32) Pritchard et al, 2000 Genetics 255, 945; (33) Reich and Goldstein, 2001 Genetic Epidemiol 20, 4; (34) Yu et al, 2006 Nat Genet 38, 203; (35) Nordborg et al, 2002 Nat Genet 30, 190; (36) Nordborg et al, 2005 PloS Biology 3, 1289; (37) Remington et al, 2002 Proc Natl Acad Sci USA 98, 11479; (38) Tenaillon et al, 2001 Proc Natl Acad Sci USA 98, 9161; (39) Rafalski and Morgante, 2004 Trends Genet. 20, 103; (40) Kruglyak 1999 Nat Genet 22, 139; (41) Ponce et al, 1999 Mol Gen Genet 261, 408; (42) Cho et al, 1999 Nat Genet 23, 203; (43) Spiegelman et al, 2000 Plant Cell 12, 2485; (44) Swan et al, 2002 Genome Res 12, 1100; (45) Borevitz et al, 2003 Genome Research 13, 513; (46) Jansen et al, 2003 Crop Sci 43, 829; (47) Kim et al, 2006 Genetics 173,1125; (48) Alonso-Blanco et al, 1998 Genetics 149, 749; (49) Singh and Kaur 2001 Crop Improvement 28, 249; (50) Swarup et al, 1999 Plant J 20, 76; (51) Lexer et al, 2003 Molecular Ecology 12, 1225; (52) Monforte et al, 1997 Theor Appl Genet 95, 284; (53) Wang and Goldman 1997 Euphytica 96, 317; (54) Guzman-Maldonado et al, 2003 CropSci 43, 1029; (55) Yadav et al, 1998a Annals of Applied Biology 133, 227; (56) Yadav et al, 1998b Annals of Biology 14, 1; (57) Ming et al, 2002 Theor Appl Genet 105, 332; (58) Cherif and Harrabi, 1993 Plant Pathology 42, 617; (59) Stoddard and Herath, 2001 Austr J Agric Res 52, 73; (60) Terry et al, 1999 Entomologia Experimentalis et Applicanta 91, 465; (61) Landi et al, 2002 Maydica 47, 233; (62) Campbell and White, 1995 Phytopathology 85, 886; (63) Singh et al, 1993 Plant Disease 77, 460; (64) Knox et al, 1999 Can J Plant Pathol 21, 174; (65) Zhang et al, 2001 Phytopathology 9, 680-686; (66) Poysa 1993 Can J Plant Pathol 15, 301; (67) Gibon et al, 2004 Plant Cell 16: 3304; (68) Gibonet al, 2006 Genome Biology 7, R76; (69) Schauer et al, Nature Biotechnology, 2006; (70) Keurentjes et al, (2006) Nature Genetics 38, 842; (71) Vandepoele et al, 2002 Plant Cell 14, 903; (72) Coen et al, (2004) Proc Natl Acad Sci USA 101, 4728; (73) Beemster et al, 1996 J Exp Bot 47, 1663; (74) West et al, 2004 Plant Physiol 135, 1050; (75) Poorter et al., 2005 Oecologia 142, 360; (76) Poorter and Remkes, 1990 Oecologia 83, 553; (77) Sakamoto et al, 2006, Nature Biotech 24:105; (78) Poorter and Bergkotte 1992 Plant, Cell Environ 15, 221; (79) Achard et al, 2006 Science 311, 91; (80) Osuna et al, 2007 Plant J 49, 463.

I.5. Zusammenarbeit mit anderen Stellen

Das Projektvorhaben wurde in enger Zusammenarbeit mit den anderen beteiligten Partnern im Verbundprojekt GABI-GNADE durchgeführt. In regelmäßigen Intervallen fanden Projekttreffen statt, mindestens einmal pro Jahr am Standort eines Partners sowie am Rande des GABI-Statusseminars. 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. Spezifische Kooperationen und zusätzliche Treffen erfolgten in Bezug auf den Austausch von Zwischenergebnissen und von vorläufigen Daten wie der Bereitstellung der C24-Resequenzierten Information durch den Partner MPIZ und mit den anderen Projektpartnern vom MPIMP.

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

Im GABI-GNADE Projekt erfolgte die Bestätigung der in der Col-0 und C24 RIL-Population bestimmten Biomasse QTL in zwei Kartierungspopulationen mit jeweils 355 und 178 Arabidopsis Akzessionen. Beide Kartierungspopulationen wurden mit 250.000 SNP Markern (Affymetrix 250K SNP Array) genotypisiert, wobei die Genotypdaten für einen Großteil der Akzessionen von den Arbeitsgruppen M. Nordborg und J. Borevitz zur Verfügung gestellt wurden und 87 Linien mit derselben Methode vom IPK genotypisiert wurden. Die SNP-Daten dieser zusätzlichen Linien wurden in Zusammenarbeit mit der Nordborg Gruppe in deren Analyse-Pipeline analysiert, um so die Kompatibilität mit den bereits vorhandenen Daten zu gewährleisten. Durch begrenzte Assoziationstests mit Kandidatenregionen auf den Chromosomen 1 unten, 3 oben und 4 oben konnten die identifizierten Biomasse QTL bestätigt werden. Die Biomasse QTL auf den Chromosomen 1 unten und 3 oben wurden dadurch von jeweils 3,65 MB auf 420 kb bzw. 3.6 Mb auf 500 kb eingegrenzt. Weiterhin konnten die Biomasse QTL auf Chromosom 1 oben, 1 unten und 3 unten auch unter Stickstoffmangel Bedingungen identifiziert werden. Als Ergänzung der auf dem 250K Affymetrix Array repräsentierten SNPs zwischen C24 und Col-0 wurden C24 ‚re-sequencing‘ Daten (Cosmid-Daten und Daten vom Partner MPIDB) herangezogen, um weitere SNPs in den QTL Regionen zu identifizieren. Aus diesen SNPs wurden 384 für einen neuen Illumina Golden Gate OPA ausgewählt, mit dem die 447 Akzessionen genotypisiert wurden. Zur Bestätigung und weiteren Feinkartierung der Biomasse QTL Bereiche und somit zur Identifizierung möglicher Kandidatengene mit Wachstumseinfluss wurden für die QTL Regionen auf Chromosome 1 unten, 3 oben und 3 unten Nachkommen segregierender RIL und IL Familien (HIFs) auf Blattfläche hin untersucht. Die Genotypisierung der untersuchten Individuen erfolgte mit MASC-Markern auf der MegaBace-Plattform.

Die Kandidatenregion des Biomasse QTL auf Chromosom 4 konnte erfolgreich auf 50 kb und 14 Gene feinkartiert werden. Für diese 14 Kandidatengene wurden KO-Linien von SALK und GABI-Kat geordert, geselbtest und mittels PCR analysiert. Von 12 Genen stehen mindestens 2 homozygote, von 2 Genen nur heterozygote KO-Linien zur Verfügung. Erste Analysen während der Vermehrung gaben Hinweise auf Wachstumsunterschiede in 2 Linien.

Die weitere Charakterisierung der Kandidatengene erfolgte durch Kreuzungsexperimente. Dafür wurden KO-Linien aller 14 Gene mit einer IL gekreuzt, die eine C24-Introgression in

dem entsprechenden Bereich auf Chromosom 4 oben enthält. Die phänotypische Analyse der Kreuzungsnachkommen wies auf einen Verlust der Heterosis bei Mutationen in mehreren Kandidatengen hin. Diese Gene sind auch durch vermehrte Interaktionen gekennzeichnet. Weiterhin wird die Kandidatenregion auf Chromosom 4 von fünf C24-Cosmiden abgedeckt (C24 genomische DNA im binären Cosmid pCLD04541). Für diese Cosmide wurde eine Shotgun-Sequenzierung durchgeführt. Die Sequenz konnte daraufhin mit Hilfe der von D. Weigel zur Verfügung gestellten Resequenzierungsdaten vollständig assembliert und mit der Col-0 Sequenz verglichen werden. Gegenüber Col-0 enthält das Genom der Akzession C24 in diesem Bereich eine ca. 2 kb große Insertion mit einem ORF und eine ca. 500 bp Deletion. Zusätzliche Screens lieferten 4 sich überschneidende Col-0-Cosmide, die diese QTL Region abdecken. Für Komplementierungsanalysen wurden die 9 Cosmide zunächst mittels ‚triparental mating‘ in *Agrobacterium tumefaciens* transformiert. Weiterhin konnten alle 4 Col-0 Cosmide in Col-0 Pflanzen und 1 Col-0 Cosmid in C24-Pflanzen transformiert werden. Komplementär dazu wurden 4 C24-Cosmide sowohl in Col-0 als auch in C24 Pflanzen transformiert. Die erfolgreiche Transformation wurde durch PCR Analysen bestätigt. Von den von Lise et al. 2007 vorgeschlagenen Kandidatengen für metabolische QTL wurden fünf Gene identifiziert, deren KO zu einer signifikanten Veränderung desjenigen Metaboliten führt, der vom Genprodukt (putativ) umgesetzt wird (Fumarat, Tyrosin, Harnstoff, Nikotinat und Phosphat). Für drei Kandidatengene waren keine KO-Linien verfügbar und es wurden RNAi und TILLING Linien hergestellt bzw. geordert. Eine RNAi Linie zeigt einen deutlichen Phänotyp bei reduzierter Expression des Kandidatengens, und in zwei rückgekreuzten TILLING Linien deutet der Screen auf eine Stop-Mutation hin. Der Fumarat mQTL auf Chromosom 5 wurde genauer untersucht. Das Col-0 QTL-Allel führt zu einer vermehrten Ansammlung von Fumarat gegenüber dem C24 QTL-Allel. Innerhalb des QTL Konfidenzintervalls liegt als direktes Kandidatengen *AT5G50950*, das auf Grund von Sequenzhomologie als putative Fumarase annotiert wurde. GC-MS Metabolitprofile von KO-Linien und Col-0 bestätigten einen verminderten Fumaratgehalt in den Mutanten. Zudem war der Gehalt von Malat signifikant verändert, in Übereinstimmung mit der reversiblen Reaktion, die von einer Fumarase katalysiert wird. Ein Vergleich der *AT5G50950* Sequenzen von Col-0 und C24 identifizierte 3 SNPs, die einen Aminosäureaustausch in der Proteinsequenz verursachen. Der Fumarase-mRNA Gehalt der beiden Akzessionen C24 und Col-0 wurde mittels quantitativer RT-PCR verglichen. Es konnte nachgewiesen werden, dass das Col-0 Allel der Fumarase ca. 16 mal höher exprimiert ist als das C24 Allel. Diese Ergebnisse bestätigen den Effekt des Col-QTL Allels und sind ein deutlicher Hinweis, dass die unterschiedliche Expression des nun bestätigten Fumarasegens auf Chromosom 5 die Ursache für die natürliche Variation des Fumaratgehalts an diesem Locus ist (Brotman et al. 2011, J Plant Physiol 168:1387). Ferner wurde innerhalb eines Tyrosin mQTL auf Chromosom 5 ein Kandidatengen für eine vermutete Tyrosinaminotransferase (*AT5G53970*) identifiziert und funktionell charakterisiert. Die Komplementation einer *E. coli* – Mutante und die Analyse zweier unabhängiger Arabidopsis Insertionsmutanten mittels GC-MS lieferten den Nachweis der physiologischen Funktion des codierten Enzyms, das nicht in die Tyrosin-Synthese sondern in dessen Nutzung für die Erzeugung anderer Stoffwechselprodukte involviert ist. Die Beobachtung einer drastischen Reduktion der Tocopherolgehalte in den Insertionsmutanten führte zu der

Schlussfolgerung, dass *AT5G53970* eine Tyrosinaminotransferase codiert, die eine entscheidende Funktion in der Tocopherolsynthese hat (Riewe et al., zur Veröffentlichung eingereicht).

II.2. Die wichtigsten Positionen des zahlenmäßigen Nachweises

- Die identifizierten Biomasse QTL in der Col-0 und C24 RIL-Population wurden in zwei mit dem Affymetrix 250K SNP Array genotypisierten Kartierungspopulationen mit jeweils 355 und 178 Arabidopsis Akzessionen bestätigt.
- Bestätigung und Feinkartierung der Kandidatenregionen auf Chromosome 1 unten, 3 oben und 4 durch begrenzte Assoziationstests. Die Biomasse QTL auf Chromosom 1 unten und 3 oben konnten dadurch von jeweils 3,65 MB auf 420kb und 3,6 Mb auf 500 kb eingegrenzt werden.
- Genotypisierung von 447 Akzessionen mit einem neu erstellten 384plex Illumina Golden Gate Array, als Ergänzung der auf dem 250K Affymetrix Chip repräsentierten SNPs zwischen C24 und Col-0.
- Feinkartierung der Biomasse QTL Regionen auf Chromosom 1 unten, 3 oben und 3 unten mittels Genotypisierung mit SNP-Markern auf der MegaBace-Plattform und Phänotypisierung auf Blattfläche von Nachkommen segregierender RIL und IL Familien (HIFs) in diesen QTL-Bereichen.
- Erfolgreiche Feinkartierung der Kandidatenregion des Biomasse QTL auf Chromosom 4 auf 50kb und 14 Gene, sowie deren funktionelle Untersuchung mittels T-DNA KO Mutanten und Kreuzungsanalysen.
- Sequenzierung von fünf C24-Cosmiden der Kandidatenregion auf Chromosom 4. Vollständige Assemblierung der Sequenz mit Hilfe der von D. Weigel zur Verfügung gestellten Resequenzierungsdaten. Im Vergleich zu Col-0 enthält das Genom der Akzession C24 in diesem Bereich eine ca. 2 kb große Insertion mit einem ORF und eine ca. 500 bp Deletion.
- Zusätzliche Screens der Region auf Chromosom 4 lieferten 4 sich überschneidende Col-0-Cosmide, die diese QTL Region abdecken. Für Komplementierungsanalysen wurden die 9 Cosmide zunächst mittels ‚triparental mating‘ in *Agrobacterium tumefaciens* transformiert und sodann zur Transformation von *Arabidopsis thaliana* Col-0 und C24 eingesetzt.
- Identifizierung von 5 Kandidatengen für metabolische QTL, deren KO zu einer signifikanten Veränderung desjenigen Metaboliten führt, der vom Genprodukt (putativ) umgesetzt wird (Fumarat, Tyrosin, Harnstoff, Nikotinat und Phosphat). Für drei Kandidatengene waren keine KO-Linien verfügbar und es wurden RNAi und TILLING Linien hergestellt bzw. geordert. Eine RNAi Linie zeigt einen deutlichen Phänotyp bei reduzierter Expression des Kandidatengens, und in zwei rückgekreuzten TILLING Linien deutet der Screen auf eine Stop-Mutation hin.
- Detaillierte Untersuchung eines Fumarat mQTL auf Chromosom 5. Das Col-0 QTL-Allel führt zu einer vermehrten Ansammlung von Fumarat gegenüber dem C24 QTL-Allel. Innerhalb des QTL Konfidenzintervalls liegt als direktes Kandidatengen *AT5G50950*, das auf Grund von Sequenzhomologie als putative Fumarase annotiert wurde. GC-MS Metabolitprofile von KO-Linien und Col-0 bestätigten einen

verminderten Fumaratgehalt in den Mutanten. Zudem war der Gehalt von Malat signifikant verändert, in Übereinstimmung mit der reversiblen Reaktion, die von einer Fumarase katalysiert wird.

- Ein Vergleich der AT5G50950 Sequenzen von Col-0 und C24 identifizierte 3 SNPs, die einen Aminosäureaustausch in der Proteinsequenz verursachen. Der Fumarase-mRNA Gehalt der beiden Akzessionen C24 und Col-0 wurde mittels quantitativer RT-PCR verglichen. Es konnte nachgewiesen werden, dass das Col-0 Allel der Fumarase ca. 16 mal höher exprimiert ist als das C24 Allel. Diese Ergebnisse bestätigen den Effekt des Col- QTL Allels und sind ein deutlicher Hinweis, dass die unterschiedliche Expression des nun bestätigten Fumarasegens auf Chromosom 5 die Ursache für die natürliche Variation des Fumaratgehalts an diesem Locus ist (Brotman et al. 2011, J Plant Physiol 168:1387).
- Detaillierte Analyse eines Tyrosin mQTL auf Chromosom 5 mit der Identifizierung und funktionellen Charakterisierung einer vermuteten Tyrosinaminotransferase (*AT5G53970*) mit dem Nachweis seiner Funktion in der Nutzung von Tyrosin für die Erzeugung anderer Stoffwechselprodukte insbesondere Tocopherolen (Vitamin E) (Riewe et al., zur Veröffentlichung eingereicht).

II.3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

In diesem Projekt wurden grundlegende Kenntnisse über die Anwendbarkeit neuartiger Technologien im Bereich Genotypisierung und Sequenzierung für das Referenzsystem *Arabidopsis* gewonnen und neue Strategien für die Kombination von Datensätzen aus der Assoziationskartierung und Koppelungskartierung zur Unterstützung von Feinkartierungs- und Genklonierungsverfahren entwickelt. Die in diesem Projekt geleisteten Arbeiten und bereitgestellten Ressourcen ermöglichen eine effizientere Nutzung der natürlichen genetischen Diversität in *Arabidopsis*. Insbesondere die Kombination von QTL-Kartierung, Resequenzierung und *de novo* Sequenzierung hat zu neuen Erkenntnissen über (mögliche) genetische Ursachen von Wachstumsunterschieden und Metabolitveränderungen geführt. Die erhaltenen Ergebnisse sind von entscheidender Bedeutung für die Abschätzung der Anwendbarkeit der getesteten Verfahren in Pflanzenzüchtungsprogrammen und sie geben klare Hinweise für die Voraussetzungen, unter denen sie erfolgreich eingesetzt werden können. Damit entsprechen die erhaltenen Ergebnisse und somit die geleistete Arbeit den Zielen des Förderprogramms.

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

In diesem Vorhaben wurden sehr wichtige Erkenntnisse über die effiziente Nutzung natürlicher genetischer Diversität unter Einbeziehung neuer Techniken zur Erfassung der Sequenzdiversität von Gesamtgenomen und der Integration von Ansätzen der Assoziations- und Koppelungskartierung gewonnen. Diese Informationen können nun herangezogen werden, um wertvolle genetische Variation in pflanzengenetischen Ressourcen von Kulturpflanzen aufzuspüren und diese optimal in der Pflanzenzüchtung einzusetzen. Sie helfen weiterhin die Auswahl der wirksamsten und kosteneffizientesten Strategien zu treffen und unterstützen die Planung und Gestaltung entsprechender Zuchtprogramme und

experimenteller Designs. Die erhaltenen Kenntnisse beziehen sich auf neue Markerplattformen mit extrem hoher Markerdichte bzw. mit spezifisch zusammengestellten Markersätzen, die in hohem Durchsatz angewendet werden können. Ferner wurden wichtige Erfahrungen in Bezug auf die Nutzung vielfältiger Quellen genetischer Variation und neuartiger Populationen gemacht und auf die Integration und Verwendung von QTL-Informationen aus komplementären Datensätzen. Diese Erkenntnisse wurden bereits vor ihrer Publikation den GABI-FUTURE-Partnern einschließlich den beteiligten Wirtschaftspartnern zur Kenntnis gebracht, die diese für die Optimierung ihrer Strategien unter Nutzung kosteneffizienter Verfahren für die Identifizierung vorteilhafter genetischer Diversität einsetzen können.

Die im Laufe der Arbeiten im Teilprojekt F erlangten Kenntnisse zu Kandidatengenen für Wachstums-modulierende Faktoren, insbesondere solchen, die Hybridwüchsigkeit (Heterosis) verursachen können, sind von hohem Wert für weiterführende vergleichende Untersuchungen an Kulturpflanzen und damit für neue Möglichkeiten der effektiveren und gezielteren Gestaltung der Hybridzüchtung. Eine Umsetzung dieser Strategie ist in einem weiterführenden Projekt in Zusammenarbeit mit einem Wirtschaftspartner geplant.

Die molekular identifizierten und in ihrer Funktion verifizierten Gene, deren Variation zu Änderungen in Metabolitgehalten führen (Fumarase, Tyrosinaminotransferase), eröffnen neue Zugänge zur Modulation wichtiger primärer und sekundärer Stoffwechselwege. Damit lassen sich einerseits die Fumarat-Gehalte in Pflanzen modulieren, die ggf. eine wichtige physiologische Bedeutung als alternative diurnale C-Speicher besitzen und damit potenziell zur Optimierung von wachstumsbezogenen Stoffwechselprozessen herangezogen werden können. Andererseits bietet das identifizierte Tyrosinaminotransferasegen eine neue Möglichkeit die Tocopherol- (Vitamin E-)Gehalte in Pflanzen zu steigern und damit die Produktion und die Verfügbarkeit dieses wichtigen Vitamins zu verbessern.

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

Eine sehr umfangreiche Assoziationsstudie unter Verwendung von Sätzen von 95 bzw. 96 Arabidopsis-Akzessionen, die partiell mit denen im GABI-GNADE-Projekt untersuchten Linien überlappen, und Nutzung von 250k SNP-Genotypdaten wurde von Atwell et al. (2010; Nature 465, 627-631), für 107 Phänotypen ausgeführt, darunter verschiedene Wachstumsparameter. Allerdings wurden dort keine Stoffwechselmerkmale berücksichtigt. Der Abgleich der durch Atwell et al. (2010) für wachstumsbezogene Parameter identifizierten Gene mit möglichen Kandidatengenen der identifizierten Biomasse QTL Bereiche aus dem GABI-GNADE-Projekt zeigte einzelne Übereinstimmungen in Genen für die Bereiche auf Chromosome 1 unten, sowie 3 oben und unten.

Die rasante Entwicklung hocheffizienter Sequenzieretechniken in den letzten Jahren hat vermehrt zu einer Anwendung dieser Methoden bei Kulturpflanzen geführt. Ähnliche Erkenntnisse, die die Verbindung dieser Techniken mit QTL Analysen und Profilanalysemethoden darstellen, wurden bisher an keiner anderen Stelle beschrieben.

II.6. Erfolgte oder geplante Veröffentlichungen des Ergebnisses

Die erzielten Ergebnisse wurden in sehr renommierten internationalen wissenschaftlichen Fachzeitschriften publiziert:

Brotman, Y., Riewe, D., Lisek, J., Meyer, R.C., Willmitzer, L., and Altmann, T. (2011) Identification of enzymatic and regulatory genes of plant metabolism through QTL analysis in *Arabidopsis*. *J. Plant Physiol.* **168**, 1387-1394.

Childs, L.H., Witucka-Wall, H., Günther, T., Sulpice, R., von Korff, M., Stitt, M., Walther, D., Schmid, K.J., and Altmann, T. (2010) Single feature polymorphism (SFP)-based selective sweep identification and association mapping of growth-related metabolic traits in *Arabidopsis thaliana*. *BMC Genomics* **11**, 188.

Lisek, J., Meyer, R.C., Steinfath, M., Redestig, H., Becher, M., Witucka-Wall, H., Fiehn, O., Törjék, O., Selbig, J., Altmann, T., and Willmitzer, L. (2008) Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *Plant J.* **53**, 960-972.

Lisek, J., Steinfath, M., Meyer, R.C., Selbig, J., Melchinger, A.E., Willmitzer, L., and Altmann T. (2009) Identification of heterotic metabolite QTL in *Arabidopsis thaliana* RIL and IL populations. *Plant J* **59**, 777-788.

Meyer, R.C., Kusterer, B., Lisek, J., Steinfath, M., Becher, M., Scharr, H., Melchinger, A.E., Selbig, J., Schurr, U., Willmitzer, L., and Altmann T. (2010) QTL analysis of early stage heterosis for biomass in *Arabidopsis*. *Theor. Appl. Genet.* **120**, 227-237.

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., Günther, M., Höhne, 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.

Sulpice, R., Trenkamp, S., Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H., Pyl, E.-T., Tschoep, H., Steinhauser, M.C., Günther, M., Höhne, 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.

Törjék, O., Meyer, R.C., Zehnsdorf, M., Teltow, M., Strompen, G., Witucka-Wall, H., Blacha, A., and Altmann, T. (2008) Construction and analysis of 2 reciprocal *Arabidopsis* introgression line populations. *J. Hered.* **99**, 396-406.

Weitere Publikationen mit jüngsten Ergebnissen und den Ergebnissen weiterführender Untersuchungen sind zur Publikation eingereicht bzw. in Vorbereitung.

Darunter:

Riewe, D., Koohi, M., Lisek, J., Pfeiffer, M., Lippmann, R., Willmitzer, L., and Altmann, T. (2011) A tyrosine aminotransferase involved in tocopherol synthesis in *Arabidopsis* (eingereicht)

Berichtsblatt

1. ISBN oder ISSN	2. Berichtsart Schlussbericht	
3a. Titel des Berichts GABI-FUTURE-Verbundvorhaben: `Nutzung genomweiter natürlicher Diversität in der Arabidopsis Population (GABI-GNADE)` (Teilvorhaben F)		
3b. Titel der Publikation		
4a. Autoren des Berichts (Name, Vorname(n)) Altmann, Thomas; Meyer, Rhonda C.; Weigelt-Fischer, Kathleen	5. Abschlussdatum des Vorhabens März 2011	
	6. Veröffentlichungsdatum	
4b. Autoren der Publikation (Name, Vorname(n))	7. Form der Publikation	
	9. Ber.Nr. Durchführende Institution	
8. Durchführende Institution(en) (Name, Adresse) Leibniz Institut für Pflanzengenetik und Kulturpflanzenforschung Corrensstraße 3 06466 Gatersleben	10. Förderkennzeichen 0315060F	
	11a. Seitenzahl Bericht 13	
	11b. Seitenzahl Publikation	
	12. Literaturangaben 89	
13. Fördernde Institution (Name, Adresse) Bundesministerium für Bildung und Forschung (BMBF) 53107 Bonn	14. Tabellen 0	
	15. Abbildungen 0	
	16. Zusätzliche Angaben	
17. Vorgelegt bei (Titel, Ort, Datum)		
18. Kurzfassung Die Vorzüge der Modellpflanze <i>Arabidopsis thaliana</i> sollten genutzt werden, um die rasanten technischen Weiterentwicklungen und die sich damit eröffnenden neuen Strategien für die Untersuchung und die Nutzung natürlicher genetischer Diversität zu testen und so natürliche Polymorphismen in Genen zu identifizieren, die in Wachstumsprozesse involviert sind. Natürliche Diversität ist nicht nur eine zentrale Ressource für die Pflanzenzüchtung sondern stellt gleichzeitig eine sehr wichtige Basis für die Grundlagenforschung dar. Unter Verwendung von RIL und zusätzlichen IL Populationen der Arabidopsis Akzessionen Col-0 und C24 konnten insgesamt 7 QTL für Biomasse und 86 QTL für bekannte Metabolite identifiziert und bestätigt werden. Diese kombinierten Daten bieten eine hervorragende Basis, um funktionell relevante Variationen in Genen, die in Verbindung mit Wachstums- und Stoffwechselprozessen stehen, zu identifizieren. Zur Validierung und Feinkartierung der Biomasse QTL Bereiche wurden Assoziationskartierungen an zwei genotypisierten und auf Biomasse hin phänotypisierten Kartierungspopulationen mit jeweils 355 und 178 Arabidopsis Akzessionen und zusätzlich mit Nachkommen segregierender RIL und IL Familien (HIFs) durchgeführt. Dadurch wurden die Biomasse QTL auf den Chromosomen 1 unten und 3 oben von jeweils 3,65 MB auf 420 kb bzw. 3,6 Mb auf 500 kb eingegrenzt. Die Kandidatenregion des Biomasse QTL auf Chromosom 4 oben konnte erfolgreich auf 50 kb und 14 Gene feinkartiert werden. Große Fortschritte und Weiterentwicklungen der Methoden und Techniken erfolgten einerseits im Bereich der Populationskartierungen als auch auf dem Gebiet der Resequenzierungen ganzer Genome. So konnten mit Hilfe der Resequenzierungsdaten der Akzession C24 Sequenzen von C24-Cosmiden in der Kandidatenregion auf Chromosome 4 oben vollständig assembliert und mit der Col-0 Sequenz abgeglichen werden. Im Vergleich zu Col-0 weist das Genom der Akzession C24 neben vielen Einzelnukleotidpolymorphismen und kleinen Insertions-Deletions-Variationen in diesem Bereich eine ca. 2 kb große Insertion mit einem ORF und eine ca 500 bp Deletion auf. Im Rahmen dieses Projektes wurden ferner zwei Kandidatengene von Metabolit-QTL durch Analyse von T-DNA-Insertionsmutanten als ursächlich für die Variation in Metabolitgehalten identifiziert und ihre Funktion im pflanzlichen Stoffwechsel aufgeklärt.		
19. Schlagwörter		
20. Verlag	21. Preis	