

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*-Akzessionen 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](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.