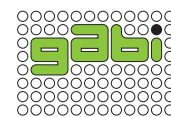


# Freier Zugang zu genetischen Pflanzenressourcen



**Eine der weltweit größten Sammlungen pflanzlicher "Knock-out"-Linien wird frei zugänglich.**

**Bernd Weisshaar und Heinz Saedler**

Die im Rahmen der deutschen Pflanzengenom-Initiative GABI (Genomanalyse im biologischen System Pflanze) am Kölner Max-Planck-Institut für Züchtungsforschung erstellten GABI-Kat-Linien werden an ein internationales Zentrum für pflanzliche Ressourcen übergeben und damit weltweit frei zugänglich gemacht. Bei den GABI-Kat-Linien handelt es sich um eine Sammlung genau definierter und katalogisierter Gendefekt-Mutanten der Referenzpflanze *Arabidopsis thaliana*. Das Studium der GABI-Kat-Linien führt Wissenschaftler zur Funktion der jeweils betroffenen Gene und deren Beitrag zu wichtigen pflanzlichen Eigenschaften. In Kulturpflanzen kann damit genau und schnell nach besseren Varianten wichtiger



Definition von "Gen-Treffern" (gene hits) bei GABI-Kat

**Gene gesucht werden, um so neue Pflanzensorten zu züchten, die z. B. Pilzattacken abwehren, Dürre und Frost widerstehen oder höhere Erträge bei geringerer Düngung liefern.**

In höheren Pflanzen ist die gezielte Veränderung von bestimmten Genen durch homologe Rekombination, wie sie in Hefe- oder Mauszellen möglich ist, nicht durchführbar.

Deshalb sind große Sammlungen von Insertionsmutanten, in denen die betroffenen Gene durch Sequenzierung der Insertionsstellen identifiziert wurden, von zentraler Bedeutung für den Fortschritt der Pflanzengenomforschung. Wir haben mit Ackerschmalwand (*Arabidopsis thaliana*) gearbeitet, einer weltweit verwendeten Referenzpflanze, deren Erbinformation vollständig entschlüsselt ist. Laut der aktuellen Version der *A. thaliana* Genomannotation (TIGR Ath genome annotation v5, siehe <http://www.tigr.org/tdb/e2k1/aht1.shtml>), beläuft sich die Zahl der proteinkodierenden Gene einschließlich der Pseudogene auf ungefähr 30.700. Die Gesamtzahl der verschiedenen Gene, für die zumindest ein Insertionsallel in den verschiedenen öffentlich verfügbaren Ressourcen aufgeführt ist, beläuft sich auf ungefähr 27.760.

Mit einem aus dem Bodenbakterium *Agrobacterium tumefaciens* stammenden, natürlichen System zur Genübertragung konnten wir für 16.900 der insgesamt 26.000 normalerweise intakten *A. thaliana*-Gene zum Teil mehrere unterschiedliche "Knock-out"-Mutanten erzeugen. T-DNA-Insertionsmutanten können in *A. thaliana* leicht hergestellt werden. T-DNA erzeugt eine geringe Anzahl an mutierten Loci pro Pflanze und die T-DNA-Markierung kann leicht und direkt mittels PCR nachgewiesen werden. Die für den Umgang mit T-DNA-Insertionsmutanten notwendige Genetik ist einfach, insbesondere im Vergleich zur Verwendung von autonomen Transposons oder Zwei-Komponenten-Transposon-Systemen als Insertionsmutagen. Von der internationalen Forschungsgemeinschaft wurden mehrere T-DNA



Blick in das Samenlager von GABI-Kat

mutagenisierte Populationen generiert und die Insertionsstellen der T-DNA relativ zu den betroffenen Genen in FST-Datenbanken (flankierender Sequenzbereich: Flanking Sequence Tags – FST) katalogisiert. Die so erzeugten Insertionsmutanten sind die Basis für die Bestimmung der Funktionen der ca. 26.000 Gene des *A. thaliana*-Genoms, die nicht als Pseudogene betrachtet werden.

### Europas größte Ressource

Nach der Erzeugung wurden die einzelnen Mutanten zu Pflanzenlinien vermehrt und das jeweils betroffene Gen mittels PCR-Methoden und DNA-Sequenzierung identifiziert. Aus den erhaltenen FST-Daten wurde ein "Online-Katalog" erstellt, der es erlaubt, für jedes gewünschte Gen die zugehörigen "Knock-out"-Pflanzenlinien zu bestellen. Bisher wurden auf diese Weise 63.000 Linien entwickelt und katalogisiert. Die GABI-Kat-Sammlung stellt damit die weltweit zweitgrößte Ressource für derartige Pflanzenlinien dar (siehe Tabelle "GABI-Kat in Zahlen"). Die größte derartige Sammlung ist die SALK T-DNA-Insertionsmutanten-Population, die 145.589 lokalisierte Insertionen enthält und damit 21.784 verschiedene Gene abdeckt (siehe <http://signal.salk.edu/cgi-bin/tdnaexpress>). Die GABI-Kat FST-Daten sind in EMBL-Bank/GenBank eingestellt und werden über das Internet mittels einer Webseite mit Suchmaske (<http://www.gabi-kat.de/>) verfügbar gemacht. Außerdem steht der vollständige Datensatz als Download-Datei für die Integration in andere Datenbanken wie SIGNAL (<http://signal.salk.edu/>) oder FLAGdb++ (<http://urgv.evry.inra.fr/projects/FLAGdb++/HTML/index.shtml>) zur Verfügung.

### Eine Insertionsmutante für jedes Gen?

Auf den ersten Blick scheint die Abdeckung aller Arabidopsisgene durch die weltweit in Datenbanken gelisteten FSTs bzw. T-DNA-Insertionslinien recht gut. Bei der Beurteilung der Zahlen muss jedoch beachtet werden, dass die tatsächliche Abdeckung des *A. thaliana*-Genoms mit Insertionen sehr von der genauen Definition eines "zählbaren" Insertionsallels abhängig ist. Die GABI-Kat Webseite (<http://www.gabi-kat.de/>) führt in der Datenbank-Ausgabe Nr. 17 (Februar 2005) zum Beispiel nur 11.987 Gene auf, in denen die Insertionsstelle zwischen den jeweiligen Start- und Stopkodons vermutet wird. Wir bezeichnen diese Insertionen als CDSi-Insertionen. Der

Begriff "CDSi" weist darauf hin, dass die entsprechende Insertion in der kodierenden Sequenz oder in den Introns zwischen dem ersten und dem letzten proteinkodierenden Exon lokalisiert ist (siehe Abb. S. 7). Weil bei diesen Insertionen das kodierte Protein verändert ist, und/oder eine Störung der Transkriptreifung vorliegen sollte, führen solche Insertionen zu drastischeren Effekten als Insertionen außerhalb dieser Region.

Die als "selling number" verwendbare Zahl der durch Insertionen abgedeckten Gene wird durch diese Definition, die sich auf wahrscheinlich brauchbare Insertionen beschränkt, deutlich von 16.664 auf 11.987 reduziert. Es kommt hinzu, dass nicht alle Insertionen, die auf Basis der FST-Daten in der EMBL-Bank/GenBank vorausgesagt wurden, in den folgenden Generationen der Pflanzenlinien wieder zu finden sind. Ein hierzu beitragendes Problem ist, dass einige Linien für die T1-Blattmaterial geerntet wurde und FST-Daten produziert wurden, keine Samen oder Samen ohne Keimfähigkeit bilden. Informationen über solche "verlorenen" Linien wurden bereits in die öffentlich zugängliche GABI-Kat Datenbank integriert. Unsere Erfahrungen zeigen auch, dass in der GABI-Kat-Population nur etwa 80% der aufgrund der aus T1-DNA generierten FSTs vorhergesagten Insertionen bestätigt werden können. Kollegen, die intensiv mit Insertionsmutantenslinien aus anderen Quellen arbeiten, beobachten ähnliche Erfolgsquoten. Die Quote kann geringfügig erhöht werden, wenn logistische Fehler, welche die Verknüpfung zwischen FST-Datensatz und Pflanzenlinie beeinflussen, korrigiert werden oder, wenn technische Probleme während des PCR-basierten Bestätigungsprozesses gelöst werden. Trotzdem bleibt die Tatsache bestehen, dass nicht alle vorhergesagten Insertionen in Form eines mutierten Allels des gewünschten Gens "materialisiert" werden können. Zusammengefasst können wir folgern, dass die Abdeckung der Gene im *A. thaliana* Genom mit wirklich brauchbaren Insertionsallelen deutlich geringer ist als auf der Basis der oben zitierten Zahlen erwartet werden könnte.

### Tausende von Anfragen

nach GABI-Kat-Linien aus dem In- und Ausland liegen vor. Die große Nachfrage ist ein Beweis für die hohe Qualität und Zuverlässigkeit der in GABI erstellten Ressource. Insgesamt 2.750 Linien wurden bislang an Wissenschaftler in aller Welt versandt. Die Anzahl der Publi-

kationen, in denen über Ergebnisse berichtet wird, die mit Hilfe von GABI-Kat Linien erarbeitet wurden, nimmt ständig zu. Um diese umfangreiche Sammlung dauerhaft für die Wissenschaft zu erhalten, wurde jetzt mit dem bis 2007 dauernden Transfer der GABI-Kat-Linien an das Nottingham Arabidopsis Stock-Center (NASC) begonnen. Das NASC ist die zentrale europäische Sicherungs- und Verteilungseinheit für genetische Ressourcen und befindet sich in Großbritannien (<http://www.arabidopsis.info/>). Auf der Basis einer Vereinbarung zwischen den Förderpartnern von GABI und der Max-Planck-Gesellschaft (MPG), die im Juni 2005 bekannt gegeben wurde, können GABI-Kat-Linien der internationalen Forschungsgemeinde nun auf diesem Weg zur Verfügung gestellt werden. Die neue Vereinbarung zum Materialtransfer (MTA – material transfer agreement) wurde im Juni 2005 bereitgestellt und ist über die GABI-Kat-Webseite erhältlich. Selbstverständlich bleiben alle bisher unterzeichneten Vereinbarungen in der unterzeichneten Form gültig.

### Der Transfer ist gestartet

Im Folgenden wird der Transfer der Linien an das NASC mit einigen relevanten Details beschrieben. Die GABI-Kat-Linien, die bereits bestätigt wurden, werden als Gruppen von T3-Samen transferiert. Ein "T3-Set" besteht dabei aus den T3-Samen einer segregierenden Familie, wobei die Samen von einzelnen T2-Pflanzen getrennt geerntet wurden. Dies ermöglicht den Nutzern den direkten Zugriff auf potentiell homozygotes Material für eine bestimmte Insertionslinie. Wenigstens eine homozygote Pflanze befindet sich mit hoher Wahrscheinlichkeit unter den 12 bis 18 herangezogenen T2-Pflanzen einer Linie. Idealerweise würde die Genotypisierungsinformation über die jeweiligen T2-Pflanzen von den Nutzern der Ressource abgefragt und am NASC gesammelt werden, so dass bekannt wird, welche T2-Pflanze homozygot und welche heterozygot für die jeweilige Insertion war. Alle T2-Pflanzen müssen resistent gewesen sein, weil sie auf selektiven Medien herangezogen werden. Dies bedeutet auch, dass es keine überlebenden T2-Pflanzen geben sollte, die homozygot für das in Frage stehende Wildtyp-Allel sind, es sei denn, es existiert ein zweiter Locus, der Resistenz liefert.

Der Transfer von bestätigten Linien von GABI-Kat zum NASC hat im Juni 2005 begonnen und wird bis zum voraussichtlichen Ende des GABI-Kat-Projektes im Jahr 2007 fortge-

## GABI-Kat in Zahlen

### A) Zeitplan

|             |  |
|-------------|--|
| 01.06.2000: | Beginn der Förderung des Projektes   |
| 01.05.2001: | Beginn der Produktionsphase der FST-Generierung  |
| 01.05.2002: | Veröffentlichung des ersten öffentlichen MTA   |
| 01.05.2002: | Erste Grobsequenzierung von FSTs an EMBL-Bank/GenBank  |
| 01.06.2002: | Eröffnung der SimpleSearch Suchmaschine über Internet  |
| 01.08.2002: | Aktualisierung von SimpleSearch mit graphischen Darstellungen der Insertionsstellen relativ zu Genen |
| 15.02.2003: | Neue Definition eines „Gen-Treffers“   |
| 15.09.2004: | Re-Annotation aller FSTs anhand des TIGR v5 Genom-Datensatzes (Genomannotation)                      |
| 13.06.2005: | Erste Samenübergabe an das NASC  |
| 24.06.2005: | Veröffentlichung des zweiten öffentlichen MTA (Disclaimer)   |

### B) Die Zahlen im August 2005

|  |         |
|--|---------|
| Resistente Linien gekeimt und selektiert:                  | 89.000  |
| Linien mit erfolgreicher DNA-Extraktion und T2-Samenernte: | 83.000  |
| Gesamtzahl der GK FSTs in der EMBL-Bank/GenBank:           | 106.000 |
| Linien mit erfassten „Genomtreffern“:                      | 63.000  |
| Zahl der Gene mit mindestens einem „Gen-Treffer“:          | 16.900  |
| Prozentuale Genabdeckung (+/- 300 bp Definition):          | 64,5%   |
| Zahl der Gene mit mindestens einem „CDSi-Treffer“:         | 12.200  |
| Prozentuale CDSi-Abdeckung:                                | 46,1%   |
| Insgesamt angefragte Linien:                               | 3.300   |
| Bestätigte und verschickte Linien:                         | 2.750   |

setzt werden. Im August 2005 waren bereits mehr als 5.000 Samenstocks an das NASC abgegeben. Dies entspricht etwa 450 Familien oder "T3-sets", dem Äquivalent von Originallinien. Sobald das NASC diese Linien zur Aufnahme in die Sammlung aufbereitet hat, werden sie für Bestellungen über die NASC Webseite zur Verfügung stehen – im August 2005 sind dies 175 "T3 sets" (2.365 Samenstocks). GABI-Kat wird diese Linien nicht weiter an User abgeben. Die Vorbereitung der hohen Zahl an T3-Samentütchen und die benötigte Annotation und Logistik sind sehr arbeitsintensiv. Wir gehen davon aus, dass wir pro Woche ungefähr 1.200 Samentütchen an das NASC schicken können, was einer Abgabe von ca. 100 bestätigten Linien pro Woche entspricht. Die GABI-Kat Webseite stellt die zusätzlichen Informationen zur Verfügbarkeit von Linien am NASC seit August 2005 bereit. Diese Informationen werden nach jeder neuen Linienabgabe aktualisiert.

Linien, die bereits ans NASC überführt wurden, können beim Stockcenter für eine Gebühr von 10 britischen Pfund pro "T3-Set" zzgl. einer Bearbeitungsgebühr von 8,50 britischen Pfund pro Bestellung abgerufen werden. Zusätzlich können weiterhin Insertionslinien direkt bei GABI-Kat bestellt werden. Diese

Zugangsmöglichkeit sollte sich selbstverständlich auf Linien konzentrieren, die noch nicht bestätigt sind und deswegen auch nicht als T3-Samen ans NASC abgegeben werden konnten. Wenn diese Linien bestätigt sind, werden sie gegen eine Gebühr von 100 Euro pro Linie und nach Unterzeichnung des o. g. MTAs an den Besteller abgegeben. Die Gebühr beinhaltet die Verpackungs- und Portokosten. Normalerweise erhält der Besteller unmittelbar nach der positiven Bestätigung ein Aliquot der T2-Samen der entsprechenden Linie. Wenn jedoch z. B. alle T2-Samen einer Linie zum Anziehen von T2-Pflanzen verwendet wurden, kommt es zu einer Verzögerung bei der Lieferung der Samen bis T3-Samen geerntet wurden.

### Mehr als 2000 exklusive Gene in GABI-Kat

Wir haben unseren FST-Datensatz nach CDSi-Insertionsallelen der Gene untersucht, die soweit bekannt nur durch GABI-Kat abgedeckt sind. Wir haben ungefähr 2.100 Gene gefunden, für die CDSi Insertionen nur bei GABI-Kat erhältlich sind. Diese Linien werden nach und nach in die Bestätigungs-Pipeline aufgenommen und – sofern ohne Verzögerung bei direkten Nutzer-Anfragen möglich – abgearbeitet.

Nach positivem Ausgang der Bestätigung und im Anschluss an eine Ernte von T3-Samen, werden diese Linien ebenfalls wie oben beschrieben ans NASC abgegeben, genau wie Linien die auf Nutzeranfrage hin neu bestätigt werden. Letztlich werden alle bestätigten Linien ans NASC abgegeben, wobei es zwangsläufig zu einer gewissen Verzögerung kommt. Dies ist zurückzuführen auf die Dauer der Reifung der T3-Samen, die Vorbereitung der vielen T3-Samentütchen für die Abgabe bei GABI-Kat, und auch den Aufnahmeprozess der neuen Samenstocks am NASC.

Mit der Unterstützung bei Erstellung und Freigabe der GABI-Kat-Linien leisten das Bundesministerium für Bildung und Forschung, die im Wirtschaftsverbund Pflanzengenomforschung GABI organisierte deutsche Wirtschaft und die Max-Planck-Gesellschaft einen wichtigen, viel beachteten Beitrag für die gemeinsame weltweite Zusammenarbeit in der Wissenschaft. Die erleichterten Zugangsbedingungen, die durch die privaten Förderungspartner des GABI Programms möglich gemacht wurden, sollten in einer weiteren, intensiven Nutzung der GABI-Kat-Insertionsmutanten resultieren. Damit leisten sie einen wichtigen Betrag zur stetig zunehmenden Zahl an Pflanzengenen, von deren Funktionen wir wenigstens eine gewisse Vorstellung haben.

### Kontakt

Bernd Weisshaar

*MPI für Züchtungsforschung, Köln*

Bitte senden Sie Ihre Emails bzgl.

GABI-Kat an: [info@gabi-kat.de](mailto:info@gabi-kat.de)

Aktuelle Adresse:

Universität Bielefeld, Fakultät für Biologie, 33594 Bielefeld

Email: [bernd.weisshaar@uni-bielefeld.de](mailto:bernd.weisshaar@uni-bielefeld.de)

### Referenzen

*GABI-Kat:*

<http://www.gabi-kat.de/>

*SIGnAL T-DNA Express:*

<http://signal.salk.edu/cgi-bin/tdnaexpress>

*TIGR Ath genome annotation:*

[www.tigr.org/tdb/e2k1/ath1/ath1.shtml](http://www.tigr.org/tdb/e2k1/ath1/ath1.shtml)

*NASC:*

[www.arabidopsis.info/](http://www.arabidopsis.info/)

*GABI Programm:*

[www.gabi.de/](http://www.gabi.de/)

*MPI für Züchtungsforschung:*

<http://www.mpiz-koeln.mpg.de/>

*WPG:*

[http://www.gabi.de/21seiten/e\\_06\\_e.php](http://www.gabi.de/21seiten/e_06_e.php)