

# Vereinfachung der markergestützten Selektion in der Pflanzenzüchtung durch die Datenbanken des Unternehmens IT-Breeding GmbH

## Implementation of the marker-assisted selection in the plant breeding programmes with the databases of IT-Breeding GmbH.

Lyudmyla V. Malysheva-Otto<sup>1\*</sup>, Fabian Gerstenberg und Lars-Gernot Otto

### Abstract

IT-Breeding GmbH, a start-up company from the Leibniz Institute of Plant Genetics and Crop Plant Research in Gatersleben, Germany, is a specialized service company which is active in the field of Research and Development with the focus on plant breeding. The company offers a combination of Product-Service. IT-Breeding develops an integrated resource of molecular and genetic information at the basis of plant traits. The resource is generated by (1) extensive and simultaneously highly specific acquisition of the relevant data from the Web, (2) analysis of the data based on company's internal know-how, (3) display of the evaluated data via uncomplicated user-friendly interface. The evaluated information IT-Breeding provides to the customers as Customer Databases for various plant species. The accompanying consulting service supports the breeders by the choice of molecular markers for specified traits, if necessary - by the choice of service company for genotyping, and by statistical evaluation and interpretation of the data after the molecular analysis. Implementation of the provided information, either in the form of the database or consulting, allows the breeders to create the programmes for the development of new varieties which will be significantly more time- and cost-effective. As a result the breeding company will gain the competitive advantage over competitors on the market.

### Keywords

Consulting services, databases, MAS, marker assisted selection, molecular data, plant traits, statistical data analysis

### Einleitung

Die Pflanzenzüchtung stellt eine der ältesten Wissenschaften der Menschheitsgeschichte dar. Jede Form der Pflanzenzüchtung, sowohl die klassische kreuzungsbasierte Züchtung als auch die Züchtung durch Einbeziehung verschiedener biotechnologischer Methoden, zielt auf die Veränderung des genetischen Materials der Pflanze, um neues und/oder verbessertes Pflanzenmaterial zu erzeugen.

Seit langer Zeit benutzen Züchter phänotypische Eigenschaften, um die besten Pflanzen unter den Nachkommen aus der Kreuzung von Eltern, die beide verschiedene wertvolle Eigenschaften besitzen, auszuwählen. Heutzutage sind die Züchter jedoch mit den drängenden globalen Herausforderungen konfrontiert, wie der rasant wachsenden Bevölkerung der Erde, begrenzten Boden- und Wasserressourcen und der Bedrohung durch den weltweiten Klimawandel. Von den Züchtern wird erwartet, dass Kulturpflanzen, die den neuen Anforderungen entsprechen und gleichzeitig eine hohe Ertragsleistung zeigen, in einem verkürzten Zeitrahmen entwickelt werden. Die einzige Möglichkeit dieser Herausforderung zu begegnen, ist die Integration der wissenschaftlichen Erkenntnisse über den genetischen Hintergrund der Merkmale, die in den letzten Jahren akkumuliert wurden.

### Datenmenge in der molekularen Pflanzenwissenschaft

In den letzten Jahrzehnten haben die Wissenschaftler große Datenmengen über pflanzliche Genome zusammengetragen, die die verschiedenen Aspekte des Pflanzenwachstums und der Pflanzenentwicklung betreffen. Diese Informationen werden in wissenschaftlichen Publikationen beschrieben und sind größtenteils über das Internet frei zugänglich zu erhalten. Wie das Beispiel der Literaturdatenbank PubMed zeigt (*Abbildung 1*), erhöhte sich die Menge an verfügbarer Information insbesondere in den letzten Jahren sehr stark.

Allerdings besteht das Problem darin, dass die vorhandenen Informationen zwar verfügbar sind, aber auf eine Vielzahl von Datenbanken verteilt sind (unterschiedliche Literaturdatenbanken und Webquellen wie z.B. GrainGenes, Gramene, MASWheat, etc.) und nicht systematisch erfasst werden. Hierbei werden die Daten auch nach unterschiedlichen Kriterien geordnet. Jeder Autor einer wissenschaftlichen Veröffentlichung präsentiert z. B. die Daten jeweils entsprechend seiner eigenen Herangehens- und Betrachtungsweise. Zurzeit gibt es Webquellen, die die molekularen Daten für bestimmte Pflanzenarten mit dem Ziel präsentieren, die wissenschaftliche Gemeinschaft über die verfügbaren Daten zu informieren (z.B. GrainGenes, Gramene, Komugi). Andererseits gibt es Webquellen,

<sup>1</sup> IT-Breeding GmbH, Am Schwabeplan 1b, D-06466 GATERSLEBEN

\* Ansprechpartner: Dr. Lyudmyla V. MALYSHEVA-OTTO, malysheva-otto@it-breeding.de

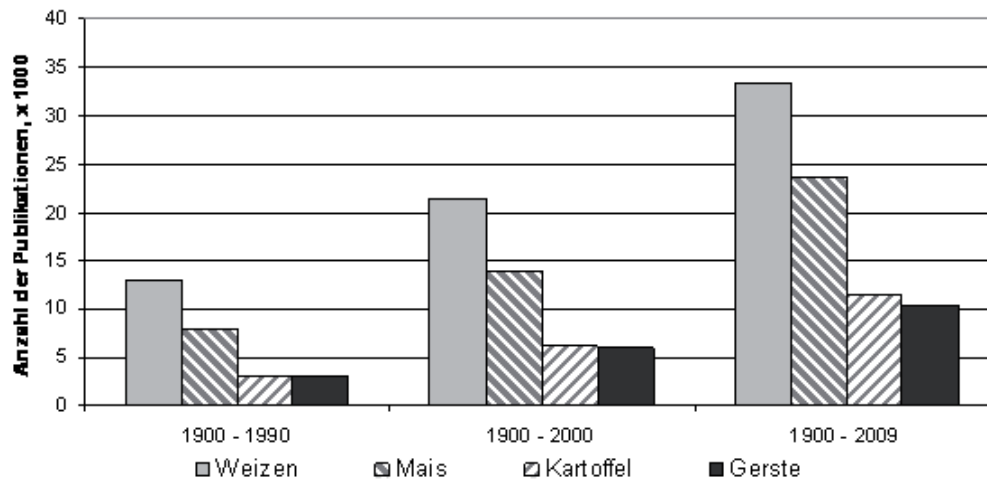


Abbildung 1: Anzahl der wissenschaftlichen Publikationen je nach Kulturart und Zeitraum

Figure 1: Number of publications ( $\times 1000$ ) on wheat, maize, potato and barley in the respective time periods

die von Interesse für Pflanzenzüchter und Saatguthändler sind, aber die molekularen Daten nicht betrachten (z.B. SeedQuest). Internationale Expertenanalysen gehen davon aus, dass die Nutzung biotechnologischer Verfahren künftig weiter voranschreiten wird und die Basis für so genannte Bioökonomie legt (BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2009).

IT-Breeding entwickelt eine Plattform für die Darstellung von molekularen und genetischen Daten, die speziell auf die Bedürfnisse des einzelnen Züchters abgestimmt sowie analysiert werden und über eine anwenderfreundliche Oberfläche verfügbar sind. Die Plattform dient als Basis, die es jedem erlaubt, auf die nötigen Informationen schnell und einfach zuzugreifen.

## Entwicklung und Struktur der Datenbanken

Um den Zugriff auf das vorhandene Wissen zu unterstützen, werden die Daten über die molekularen Hintergründe der Pflanzenmerkmale aus den verschiedenen Webquellen

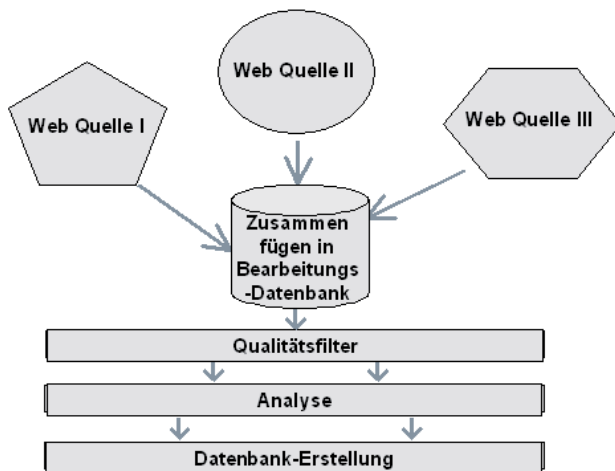


Abbildung 2: Ein Abriss der Entwicklung der Datenbank des Unternehmens

Figure 2: An outline of the development of the stock database of the company

integriert sowie ständig aktualisiert und ausgewertet. Die von IT-Breeding entwickelte Plattform präsentiert bei den molekularen Hintergründen der Pflanzenmerkmale Informationen über die Gene, QTLs und molekulare Marker, welche mit den diversen Pflanzenmerkmalen verbunden sind. Das Schema für die Entwicklung der Stammdatenbank, die die Grundlage für die spezifischen Leistungen des Unternehmens bildet, ist in *Abbildung 2* veranschaulicht.

Die ausgewerteten Informationen werden Kunden in Form von kundenspezifisch erstellten Datenbanken (KSDBen) für bestimmte Kulturpflanzenarten (z.B. Weizen, Gerste, Raps) angeboten. Die Oberfläche der KSDBen hat eine sehr klare Struktur mit Fenstern für die Auswahl der Pflanzenmerkmale auf der einen Bildschirmseite und deren genetischen Parametern (Marker, Gene und QTLs) auf der anderen Seite. Dies ermöglicht eine schnelle und unkomplizierte Suche nach den gewünschten Informationen.

## Anwendung des Fachwissens bei der Entwicklung neuer Sorten

Das erworbene Wissen über den genetischen Hintergrund von Pflanzenmerkmalen dient als eine effiziente Basis für moderne Pflanzenzüchtung. Es wird von Wissenschaftlern fortdauernd weiterentwickelt und vergrößert. Auf dieser Basis können neue Sorten schneller und zielgerichteter entwickelt werden. Das Wissen kann hierbei während zweier Schritte bei der Entwicklung neuer Sorten angewandt werden: (1) Auswahl der Eltern für das Kreuzungsprogramm (siehe z.B. GUR and ZAMIR 2004), (2) Evaluierung und Auswahl der Nachkommen (siehe z.B. KUCHEL et al. 2005, BRENNAN and MARTIN 2007).

Falls die Assoziation zwischen einem molekularen Marker und einem bestimmten Merkmal bekannt ist, kann der Marker benutzt werden, um die besten Elternpflanzen für die Kreuzung zu finden. Das Germplasma wird mit Hilfe des Markers durchsucht, um die Sorte oder Akzession zu finden, die das erforderliche Markerallel aufweist, was als Beleg für die Existenz des gewünschten Merkmals angesehen wird. Während der Evaluierung der Kreuzungsnachkommen

können Protokolle für die markergestützte Selektion (entwickelt auf der Grundlage der QTL-Detektion oder Marker-Merkmal Assoziation) angewandt werden, um den Züchtungsprozess zu beschleunigen.

### Der Umfang der kundenspezifisch erstellten Datenbanken und der Beratung

Die Kundendatenbank von IT-Breeding gibt den Züchtern die Informationen über die Gene, von denen bekannt ist, dass diese mit spezifischen Merkmalen verbunden sind (inkl. Referenzen zu den Sorten, die in die Untersuchungen einbezogen wurden) und ob die Marker innerhalb oder in Assoziation mit diesen Genen entwickelt wurden. Bei den Informationen über die molekularen Marker werden die Marker genannt, die mit spezifischen Merkmalen verbunden sind (QTLs) und es wird aufgeführt, ob die Marker erfolgreich für die markergestützte Selektion genutzt wurden (mit Referenz zu den untersuchenden Populationen). Die Datenbanken werden individuell für jeden Kunden erstellt. Sie enthalten umfangreiche Informationen oder können in Form von Modulen erstellt werden (siehe *Abbildung 3*), um spezielle Fragestellungen leichter zu beantworten.

Die Kundendatenbanken versetzen den Züchter in die Lage, schnell und ohne zusätzlichen Zeitaufwand die weltweit veröffentlichten Daten über die Gene und über die Marker für die Pflanzenmerkmale, welche ihn interessieren, zielgerichtet und zeitnah zu durchsuchen. Daraus ergibt sich eine bestmögliche Informations- und Entscheidungsgrundlage, um ein Züchtungsprogramm für die oder das gewünschte(n) Merkmal(e) zu erstellen.

Das flankierende Beratungsangebot der IT-Breeding GmbH unterstützt die Züchter bei der Auswahl der Marker, bei Bedarf auch bei der Suche nach dem kosteneffektivsten Dienstleister für die Genotypisierung, und bei der statistischen Auswertung und Interpretation der Daten nach der molekularen Analyse (mit Hilfe der Programme Diversity, PopGene, Arlequin, NTSYS, TASSEL).

### Vorteile für Züchter

Die Dienstleistungen der IT-Breeding GmbH helfen bei dem Transfer des angesammelten Wissens der molekularen Pflanzenwissenschaften in die Praxis. Die Kundendatenbanken von IT-Breeding kombinieren die Vorteile von Suchmaschinen mit der Informationsanalyse. Die Datenbanken stellen

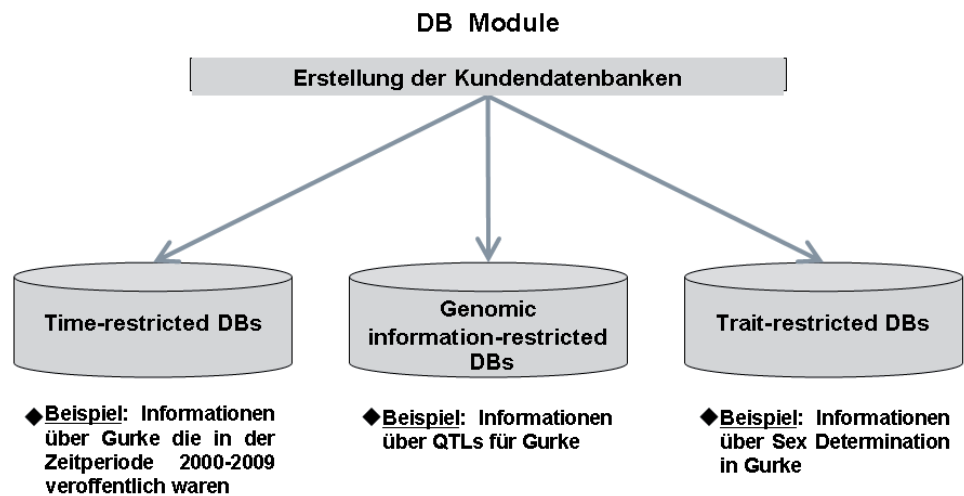


Abbildung 3: Ein Überblick über die eventuellen kundenspezifischen Datenbank-Module  
Figure 3: An overview of the possible customer-specific database modules

eine kundenspezifische Informationsplattform dar, deren sorgfältig überprüfte und analysierte Daten aus renommierten Quellen stammen und kontinuierlich gepflegt wie auch aktualisiert werden. Sie stellen eine Informationsanalyse durch die Extraktion der für das Züchtungsprogramm wichtigsten Daten bereit. Die Beratungsdienstleistung garantiert individuelle Kundenbetreuung, insbesondere in Hinblick auf spezielle Kundenwünsche und -bedürfnisse.

Durch die Auswahl der Dienstleistungen von IT-Breeding erhält der Kunde folgende Vorteile: (i) eine zuverlässige Quelle an integrierten Informationen über die genetischen Hintergründe der gewünschten Pflanzenmerkmale, kein Verlust wichtiger Informationen, (ii) konstante Integration der neuesten Publikationen, immer aufgearbeitet in der gleichen, anwenderorientierten Form, (iii) deutliche Zeit- und Geldersparnis gegenüber einer selbst durchgeführten Erfassung und Analyse von ähnlichen Informationen seitens des Züchters, (iv) die Informationen bzgl. der molekularen Marker erleichtern die Integration der markergestützten Selektion in die Pflanzenzüchtungsprogramme, (v) bei der Beratung Zugriff auf unsere Kompetenzen, welche 100% mit den Fragestellungen der Pflanzenzüchter verbunden sind.

### Literatur

- BRENNAN JP, MARTIN EP, 2007: Returns to investment in new breeding technologies. *Euphytica* 157, 337-349.
- GUR A, ZAMIR D, 2004: Unused natural variation can lift yield barriers in plant breeding. *PLoS Biology* 2, e245. DOI:10.1371/journal.pbio.0020245
- KUCHEL H, YE G, FOX R, JEFFERIES S, 2005: Genetic and economic analysis of a targeted marker-assisted wheat breeding strategy. *Mol Breed* 16, 67-78.
- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG, 2009: Runder Tisch zur Pflanzengenetik [Available online: <http://www.bmbf.de/de/13622.php>; accessed 20 Jan 2010].