GABI – The German Plant Genome Research Program

Progress Report 2004 – 2007

Federal Ministry
of Education
and Research

Imprint
GABI – The German Plant Genome Research Program · Progress Report 2004 – 2007
Published by: GABI managing office · Dr. Jens Freitag · Dr. Matthias Arlt · Dr. Saskia Dombrowski, Dr. Dirk Büssis
c/o Max Planck Institute of Molecular Plant Physiology
Am Mühlenberg 1 · 14476 Potsdam-Golm · Germany
Telephone +49.331.567 8301 · Telefax +49.331.56789 8301
eMail freitag@mpimp-golm.mpg.de · www.gabi.de

ISBN-13: 978-3-00-019972-1

Federal Ministry
of Education
and Research

Gestaltung und Layout: Dirk Biermann, biermann@potsdam.de
Photos: Dr. Matthias Arlt, Dirk Biermann
Druck: Druckhaus Mitte, Berlin

GABI – The German Plant Genome Research Program

Progress Report 2004 – 2007

Foreword

Plants have recently attracted considerable attention in the media. Plants are in. Background for this new popularity is the quest for answers to global challenges. The ever-increasing human population and its demands are putting pressure on economies, societies and the environment. A new challenge is to maintain and increase prosperity in a sustainable manner. Since climate change has entered public awareness the importance of plants has been recognized even stronger. The basis of human well being is threatened by the climate change, as its effects are global. This new challenge has to be tackled.

How is all this connected with the increased focus on plant genome research? Plants represent a renewable resource. Plants are not only crops for food but are also becoming a commodity for industry. Plants can conserve energy efficiently. Plant carbohydrates and lipids are the basis for biofuels. Plants are an essential part of our environment. Plants are both the basis of all human and animal life and are instrumental in the future development of our society.

Already the first GABI progress report (1999-2004) identified this unique importance using the phrase „Plants, the basis of life". This still holds true. However, what has changed is the public perception of plant science. Plant science has become a more positive focus in public discussions, more so than plant genetic engineering. However, controversial discussions about plant science have restarted in the public. It will be important to further improve public awareness of plant science and to react to controversial discussion proactively.

The publication of the second GABI progress report concludes another funding period. We look back and reflect on what the GABI community has achieved over the past four years. With this progress report we are assessing our current situation. Furthermore the publication of the progress report represents the kick-off for the third funding period of the GABI program.

GABI 2 has enormously influenced and shaped the German and European research and development environment. During this funding period GABI has been established as a widely visible trademark. GABI has become a vibrant network consisting of more than 300 scientists in science and industry, a true public private partnership network. This progress report presents the scientific results of GABI 2 funding period, which could only be achieved through close interaction and co-operation of the partners. GABI has brought together and integrated scientists from many different areas, ranging from agricultural scientists, plant breeders and population geneticists to molecular geneticists, biochemists, bioinformaticians, plant physiologists, etc. GABI represents the most significant investment in international co-operations in life science in Europe. The co-operation with the French partner program Génoplante is almost legendary. The integration of Spanish researchers via the ERA Net Plant Genomics is another milestone of the second funding phase of GABI. This network created not only new scientific excellence but also a positive development of the participating scientists by cooperation, sharing resources and forming synergies in general. These international co-operations of GABI have become an integral part of the program. We will reflect on this in a second article.

Foreword

The challenge ahead lies in expanding and exploiting this basis in GABI-FUTURE, the third GABI funding period. It will be important to efficiently use the increased multidiciplinarity to improve research and development activities.

The GABI-FUTURE program ensures the continuation and even expansion of stable funding for plant genome research. GABI represents the most significant funding program for plant research not only in Germany but also in Europe. This concentration of funding enables and sets free synergies between industry and academia. This will make managing the GABI program even more challenging.

By almost doubling the annual funding the BMBF has demonstrated its satisfaction with the established basis of GABI and the ongoing research and its recognition of the relevance of advancing this scientific area for our society. With the start of GABI-FUTURE the focus of the research and development program will also change and broaden. New topics will be addressed, for example the use of plant biomass as biofuels or commodity as well as increased interaction with other research areas like process or conversion technology and engineering science. This new level of interaction will become an added feature of the third funding period GABI-FUTURE.

An important factor of this ambitious development is the partnership with industry. GABI will continue to represent a strong public private partnership during the third funding period. Industry co-finances the German plant genome research program by contributing about 30%. This will be 15 Mio. Euro during the third funding phase. These funds are in addition to the investments of the industry in their own research and development. This contribution ensures that strategic areas of our social and scientific development are tackled and further advanced.

The GABI community sees these impending global changes both as a challenge and an opportunity for future development. The road ahead is clear. Aims are to reduce the dependency on fossil resources using plant products and to increase the agricultural added value. This is the road map towards a "Knowledge Based Bio Economy", a vision shared between politics, industry and science. Moving more and more towards the production and use of renewable resources will ensure sustainability of our economy and thereby of our basis of life

The staff of the GABI managing office would like to thank all groups involved in the GABI research program for their excellent co-operation and look forward to continuing this interaction with the same spirit through the coming GABI phases.

Best regards from Potsdam
Jens Freitag & Dirk Büssis

Content

Content

Content

Content

Sitemap

Sitemap

Transnational Public Private Partnership – Reflection on a process going on

International cooperation as a key factor of plant genomics. The transnational funding initiative is unique within the European Union.

The cooperation between different national plant genome research programs is a model for the targeted bundling of research capacities, the know-how and resources in the life sciences sector. A strong network of academic and industrial co-operations has been formed since the turn of the millennium, driven by national programs respective research ministries and funding agencies. Scientific cooperation between basic researchers in France and Germany laid the groundwork for the initiative, with Spanish research groups following later. A network of research in multiple European Countries, the European Research Area Plant Genomics (ERA PG) was formed. The latest development of these multinational initiatives is that two European networks will link closely. The PLANT-KBBE, as spin off of the ERA-PG is concentrating on international cooperation based on Public-Private Partnerships. The ERA-PG will continue to coordinate the international cooperation of national funding agencies in Europe concentrating more on basic research activities. The German GABI-program, the French program Génoplante and Spain's PlantGen program form the core of the PLANT-KBBE initiative.

Plant Genomics refers to the analysis of plant genomes, which comprises physical mapping of the genome, genome sequencing and annotation (prediction and identification of genes), and the elucidation of the (biochemical) functions of the identified genes and gene products and their roles in determining the characteristics of the organisms.

Beginning in the Nineties

In the mid 1990s, it became clear that substantial progress in the breeding of optimized plants for use in agriculture and industry could not be achieved without comprehensive analysis of the functions of plant genes or even of all gene function of an organism at once. Throughout the world, considerable amounts of government and private funding were invested in relevant research. Such efforts were also made in some European countries, including France and Germany. Plant genome research programs were launched in both countries in 1998. In 1999, the German and French Ministries of Research agreed to cooperate in their respective national plant genome research programs, GABI and Génoplante. The intention was to pool the resources in the European Research Area and to achieve a critical mass of know-how and research capacity, with the belief of being successful facing fierce international competition.

Both research programs were characterized by their corresponding target objectives and basic structures, a result of the close relationship between science, industry and research policies. Both national programs where organized as public-private-partnerships. The cooperation agreed between the two ministries had two objective targets:

- On one hand, to bundle research capacities from both countries aimed at the functional analysis of genomes of important crop plants and to make important resources for research and service infrastructures in both countries mutually available.
- On the other hand, cooperation between the steering committees of both programs would be initiated and the transfer of research results into an application at the companies involved in both countries would be brought about through an efficient and demand-oriented technology transfer.

To begin to tackle these major tasks, experts in the field put their heads together in a series of meetings, which began with a Franco-German workshop in Bonn in May 2000, where the general thematic frame of the planned cooperation was defined. In October 2000, experts from both sides met to discuss the issues of proprietary protection of research results and the rights of research facilities and companies involved in the research and development projects. Finally, a scientific workshop was held in February 2001 in Montpellier, where definite research projects and the rules for their evaluation by a joint Franco-German panel of experts were established. The Montpellier workshop was a bottom up process of researchers working on the model organism Arabidopsis thaliana in both programs. The model organism became also a model how to organize research and developments activities internationally.

Initial projects

Due to the difficult legal issues that would have arisen on both sides through the cooperation of companies involved in bilateral projects, only four projects from basic research were initially brought to funding in 2001 with a three-year duration

Transnational Public Private Partnership – A reflection of a process going on

and financing totaling 4 million euros, . These projects aimed at achieving cooperation in Arabidopsis research and opening up, as well as utilizing resources in both countries.

As a pilot project for the cooperation between the companies involved in GABI and Génoplante, a large joint project – also funded within the framework of EUREKA – was applied for in parallel. The focus of this project was the development and application of biomolecular tools for the improvement of silage quality in maize and quality wheat.

At the same time, structures were established which were appropriate to guarantee the proper steering of the cooperation between GABI and Génoplante. The joint circle of experts and the immediate contacts between the managing offices of both programs were of crucial importance, as well as the establishment of a permanent working group to solve the legal issues within the cooperative undertaking.

A new stimulus from politics

At the Franco-German Science Forum in Paris on 12 February 2002, this development stage was deemed a good time to announce the two countries' cooperation in the field of plant genome research – a first in the life sciences. This impulse accelerated both participants' efforts to prepare additional projects with the involvement of industry and to thus strengthen the intended application-oriented character of the cooperation. A prerequisite for this was to overcome the difficulties resulting from the different legal administrations of the two programs and – deriving from these legal aspects – the various rights to be granted to the participating companies with respect to the research goals accomplished. The permanent working group that was set up to deal with the legal issues has now achieved this prerequisite. The group set up overriding rules for the utilization of joint resources and for the handling of proprietary rights resulting from the work. The group also developed an agreement for the transfer of material as well as a secrecy agreement and stipulated regulations for a case-by-case process to resolve any project-specific problems.

In this way, another five Franco-German joint projects were brought to funding in 2004. Four of these joint ventures were carried out with the participation of French and German businesses. The funding in this case again amounted to 4 million euros. Industry supplemented approximately another 1.5 million euros from their own resources. The goals of these projects were scientific issues of interest to the participating plant breeding businesses.

Spain – a third partner came on board

The dynamic development of the Spanish plant genome research programs during these years provided the stimulus for an investigation into possible synergies and areas for cooperation between Germany, France and Spain in this research area. In May 2003, a workshop organized by the Spanish Ministry of Education and Science in Madrid, with over 250 participants came to the conclusion that there are a multitude of shared viewpoints between the goals and contents of the three countries' programs. Among others, this applied in particular to the research areas of improved biotic and abiotic stress tolerance, the improvement of physiological parameters of crops, and research on questions relevant to nutrition as well as on sustainable agricultural production.

Under the title 'Functional and comparative genomics approaches for the investigation and use of the natural diversity', the participants at the workshop developed a total of 44 expressions of interest. The end result of the international assessment procedure was that nine joint projects were brought to funding by the participating research ministries in 2004. The funding sum totaled roughly 7.5 million euros. Experiences and procedures from the Franco-German cooperation formed the basis for the entire process of identifying projects, evaluation and funding. To accompany the project implementation and to resolve issues quickly, a new steering committee was established. Making up the committee, which meets regularly, are scientists from academia and industry from the three countries, representatives from the three managing offices of the national programs and representatives from the permanent working group dealing with the legal aspects of the cooperation.

As with the French and German programs, a further aim of the Spanish program was to improve the position in the area of application-oriented research in order to strengthen the competitiveness of the industry. A first and crucial step in rising to this challenge was the formal constitution of INVEGEN, a Spanish association of private companies with interests in diverse sectors (agrofood, forestry, bioenergy, etc.), by the end of 2005.

Transnational Public Private Partnership – A reflection of a process going on

Industry platforms

To put the Franco-German-Spanish cooperation in plant genome research on a stable foundation, representatives from industry, science and the various research ministries met in Weimar in March 2005 to discuss future topics and strategy-finding processes. As a result, the representatives of the ministries signed a memorandum of understanding about the scientific cooperation. Within this memorandum, the trilateral cooperation was recognized as the core activity in application-oriented plant genome research in Europe, milestones for a further cooperation were defined and a strategy-finding process was initiated. The integration of the trilateral cooperation into the ERA-Net on plant genome research, which has been in place since 2004 and in which the three countries are also involved, was of special significance here.

A milestone in the development of the trilateral cooperation was the decision by the industry platforms involved in the three respective national programs to develop a strategy paper to define the overriding goal of the co-operation. It was also intended to describe the future development of the cooperation with respect to the application of research results in industrial practice, the strengthening of the competitiveness of the industry partners involved and the reinforcement of the cooperation between science and economy.

The strategy paper, titled 'Genomic approaches for the investigation of genetic diversity in crop plants and its use for innovation', was discussed and adopted at a trilateral meeting in October 2005 in Cordoba. Since then, it has provided the essential basis for other future joint activities between the three countries. The paper emphasizes the contribution of plant genome research to solving central issues such as climate change, limitation of fossil resources, development of a sustainable agriculture and regional adaptation of crop plants. Simultaneously, it defines important fields of research and activities, which are of equal interest to the economy in the three countries involved. They are the following subjects in particular: improved quality in food and feed, and sustainable agriculture with improved productivity and yield stability. At the same time, the importance of plant research for the development of a knowledge-based bio-industry in Europe is emphasized.

The driving force behind the ERA-Net Plant Genomics

In 2004, eleven different European countries established the ERA-Net Plant Genomics, which serves to coordinate research and funding policies in participating countries in the field of plant genome research. Spain, France and Germany are among those involved in this ERA-Net.

In 2006, the first funding call within the framework of this ERA-Net was announced. Spain, France and Germany participated with their own call. The call, titled 'From genome analysis to product innovation', is based on a strategy paper from 2005 written by the industry platforms and is aimed at application-oriented subjects: improvement of food and feed quality, improvement of productivity and yield stability, and development of plants for innovative applications such as energy providers or green factories, improvement and stabilization of surrounding natural landscapes. The project management is to be executed preferably by the individual businesses.

After an international assessment, a total of fourteen project proposals were recommended for funding. Alongside businesses and scientists from the three countries, individual partners from Portugal, the United Kingdom, Italy, Switzerland and the Netherlands also took part in a variety of projects. The total funding sum is estimated to be approximately 15 million euros. This will be supplemented by substantial funds from industry, amounting to roughly 5 million euros.

These 20 million euros, which will go to specially designed cooperation projects, are made available under the potent national funding programs of the participating countries. In the framework of the national strategic development of plant research, these funding programs, which focus on the establishment of the knowledge-based bio-economy, are moving closer together and formulate common goals (see above). Transnational cooperation substantially aids this process. France, Spain and Germany are mobilizing together more than 70 million euros annually for plant genome research alone under their national programs.

This call resulted in the opening-up of this trilateral cooperation for further countries interested in application-oriented research within the framework of a public-private-partnership.

Transnational Public Private Partnership – A reflection of a process going on

The goal of these projects is to prepare the ground for the development of a knowledge-based bio-industry in Europe. The central concern is to secure the competitiveness of agriculture and industry, and thus jobs. In such a way, the three countries hope to continue to create visible momentum across Europe.

Plant-KBBE

In September 2007 a new quality of international cooperation was reached. On initiative of the "Ministère de l'Enseignement Supérieur et de la Recherche" (MESR, France), the "Agence Nationale de la Recherche" (ANR, France), the "Ministerio de Educación y Ciencia" (MEC, Spain) and the "Bundesministerium für Bildung und Forschung" (BMBF, Germany) a new research program was launched in Paris combining the national activities in plant genomics, plant systems biology as well as biotechnology and to support the establishment of a knowledge-based bio-economy in Europe (KBBE). The "Transnational Plant Alliance for Novel Technologies - toward implementing the Knowledge Based Bio-Economy' (PLANT-KBBE) in Europe" will become the trademark for international co-operation in life sciences in Europe- based on "Public-Private-Partnerships". The program will start in 2008 and will run until 2013. The purpose of the alliance is to further stimulate interdisciplinary research and development efforts in various sectors of our economies. Therefore, scientists from the public and the private sector, from universities, research institutions and multiple private companies representing plant breeders, chemical as well as food processing industry, chemical industry and bio-energy producers, are invited to support the program.

The initiated transnational research and development program "PLANT-KBBE" represents a consequent further development of positive experiences of France, Spain and Germany, with cooperation's on the project level existing over several years already. A cost efficient use of resources was practiced and significant synergisms were achieved based on the cooperation of the three national programs on plant genomics: i) Génoplante in France, ii) GABI (Genomanalyse im biologischen System Pflanze) in Germany and iii) Acción Estratégica de Genómica y Proteómica in Spain.

High-ranking representatives from the three ministries and the ANR signed a "Memorandum of Understanding" specifying the terms of the future transnational cooperation in Paris in September 2007. In the field of life sciences, disciplines such as plant genomics, plant systems biology and biotechnology will significantly contribute to the implementation of the knowledge-based bio-economy. "Bio-energy", "biomaterials" and "innovations for healthier and safer food" will be the topical priorities for the first program phase (2008-2013). Once established, the newly launched transnational program "PLANT-KBBE" will become open for more partners in Europe and beyond with an interest on interdisciplinary research and development projects based on public-private-partnerships.

GABI was, is and will be an essential part of the international cooperation. Mentioned as a highlight in the GABI 1 Progress Report the international cooperation became a necessary and wanted normality in GABI 2. In GABI-Future these qualitative and quantitative developments will continue and become enlarged.

The Plant GEM – a life sciences jewel

The Plant-Genomics European Meeting (Plant-GEM) has become a symbol of the international cooperation. The first conference was organized by GABI, Génoplante and GarNet in Berlin in 2001. Since then Plant-GEM has developed into THE hub of the European plant genomics community. After Berlin, York (U.K.), Lyon (France), Amsterdam (The Netherlands) and Tenerife (Spain) the 7th P-GEM will be organized in Albena (Bulgaria). After having so far mainly concentrated on Western European countries, in 2008 a new quality will be created symbolizing that the enlargement of the European community is real. Almost 500 scientists from all over Europe and beyond will attend the conference, presenting their work and discussing new research ideas. GABI as the initiator of the Plant-GEM continuously supports the conference financially and logistically.

Jens Freitag, Dirk Büssis, Dominique Job, Pablo Vera, Daniel Richard-Molard, Frank Laplace

WPG e.V.

At an early stage, German companies with interests in plant breeding, plant protection, and the plant processing industry recognized the potential of plant genome research for dynamic social and economic development. Therefore, economic enterprises have had a defining role since the planning stages of GABI.

Guidelines have been established that make GABI attractive for companies acting on different levels of the value-producing chain:

- Bringing together Germany's existing research capabilities and coordinately transforming its research infrastructure into an integrated network
- Connecting the interests of science and industry
- Establishing the promotion of quality research and creating favorable conditions for direct participation by industry
- Creating technology transfer strategies promoting the efficient transformation of research results into innovative products
- Guaranteeing comprehensive patent coverage
- Preserving the autonomy of participating companies: Every individual company may decide freely the way and extent to which it participates in GABI projects.

Once these guidelines were established, the registered association Wirtschaftsverbund Pflanzengenomforschung GABI (Economic Network of Plant Genome Research, WPG) was founded simultaneously with the public advertisement of the Genome Analysis of the Plant Biological System (GABI) promotion focus by the BMBF on September 18, 1998.

The WPG is a registered association seated in Bonn. The purpose of the association is the promotion of research – especially GABI. Initiated by industry, this organization has succeeded in achieving two essential aims:

1. Establishment of a central contact partner for industry on matters of policy, science, and economy as well as for all other GABI bodies.
2. Coordination of the activities of the companies participating in GABI and joint definition of binding proceedings and rules for members within the framework of the promotion regulations of the BMBF.

Thus, members of the WPG jointly assume the responsibility of enforcing the above guidelines and are therefore a supporting pillar of the whole GABI concept. In return, members have the opportunity to carry through their own projects and gain access to results generated by GABI.

12 companies founded the WPG but within the first year the number of members had increased to 24 and is currently 28. The five-member Managing Committee is responsible for all association matters. All positions are assigned on an honorary basis. The management is located at the Haus der Pflanzenzüchtung (House of Plant Breeding) in Bonn. The WPG Managing Committee acts as the central communication partner for industry. The Managing Committee selects delegates for the GABI control bodies from among its members. Two WPG members are represented in the Steering Committee. Furthermore, nearly half of the members of the Scientific Advisory Board come from the WPG. Thus, the WPG ensures a balanced influence on research decisions, both politically and scientifically, within GABI. The Managing Committee elaborates proposals for a contribution scheme and submits these proposals to the WPG members for discussion. The contribution scheme takes into account the very heterogeneous struc-

ture of the WPG members consisting of small and medium as well as large enterprises. The membership fees are primarily spent to finance the activities of the WPG and of the Patent and Licensing Agency (PLA for GABI). Furthermore, the Managing Committee decides upon the admission of new WPG members. An informal WPG membership application may be submitted to the Managing Office at any time. All enterprises in the field of the plant breeding, plant protection, and plant processing industries as well as in biotechnology may become members. These members (and their parental companies) must have their headquarters in Germany and carry out the major part of their research and development activities in Germany. WPG membership opens numerous possibilities. Together with experts from research and industry, an enterprise may establish research networks for special problem fields. Thus, the enterprise gets access to extensive expertise.

Furthermore, the activities of the PLA for GABI, guarantee an extensive information flow especially from the projects of the Research Area 1 to all WPG members. WPG members are informed on a regular basis about research projects aims as well as about the results. They receive publications in advance and have a time-limited exclusive reading right for primary data. Thus, each company may check the data in terms of usefulness for its own entrepreneurial activities. Furthermore, each member is counseled on the application for and granting of patents coming from the Research Area 1 and has the possibility to take out a license on a most-favored-nation basis. The complicated and cost-intensive analysis of the whole range of data can normally not be managed by individual members. By means of the network structure and the establishment of the PLA for GABI, the patent protection results from Research Area 1 are guaranteed - WPG members can access these results and extend their use for their own activities within or outside of GABI.

The general concept of the WPG does not aim solely at offering individual members ready solutions: It requires active co-operation in addition to financial contribution. This also applies to the international interlinking of GABI with other genome projects such as the French initiative GÉNOPLANTE. It is industry's responsibility to launch such pilot projects. National borders cannot be allowed to limit efficient research.

Managing Committee of the WPG
(state of April 2003)

Chairman
Andreas J. Büchting, KWS SAAT AG

Deputy
Dieter Berg, Bayer Crop Science GmbH
Gisbert Kley, Deutsche Saatveredelung
Lippstadt- Bremen GmbH

Members
Wolf v. Rhade, Nordsaat Saatzuchtgesellschaft mbH
Hans Kast, BASF AG

Managing Director
Carl Bulich, Wirtschaftsverbund Pflanzengenomforschung
GABi e.V., Kaufmannstrasse 71, D-53115 Bonn

Patent and Licensing Agency for GABI (PLA)

One of the major strategic objectives within GABI is to ensure an efficient transfer of knowledge and technology between all cooperating parties and to secure comprehensive protection of the results by intellectual property rights. In order to meet these criteria the Patent and Licensing Agency for GABI (PLA) has been established by the industrial platform of GABI named Wirtschaftsverbund Pflanzengenomforschung GABI e.V. (WPG) whose member companies carry the financial costs of PLA. Thus, PLA is one major contribution of industry to the GABI infrastructure.

In 1998 the German Plant Breeders Association (Bundesverband Deutscher Pflanzenzüchter BDP e.V.) has founded a subsidiary, named GVS (Gesellschaft für Erwerb und Verwertung von Schutzrechten GVS mbH) for the acquisition and exploitation of property rights within the field of green biotechnology. Taking advantage of this facility PLA was established at the GVS at the House of Plant Breeding in Bonn.

One major task of PLA is to check manuscripts before publication and to give advice for securing intellectual property rights. This holds true not only for manuscripts of scientific articles but also all kinds of posters and internet presentations, abstracts and lectures. Within six weeks PLA has to decide to either pass the paper or to secure the IPR. In addition, PLA supports scientists in modifying their manuscripts in order not to endanger subsequent patent potentials. If the decision is made to file a patent for GABI results, PLA supports the applicants by engaging a competent patent attorney and by absorbing the complete costs for the basic patent application. Concerning licensing offers for these patents, PLA acts as a mediator between the patent assignee and the members of WPG.

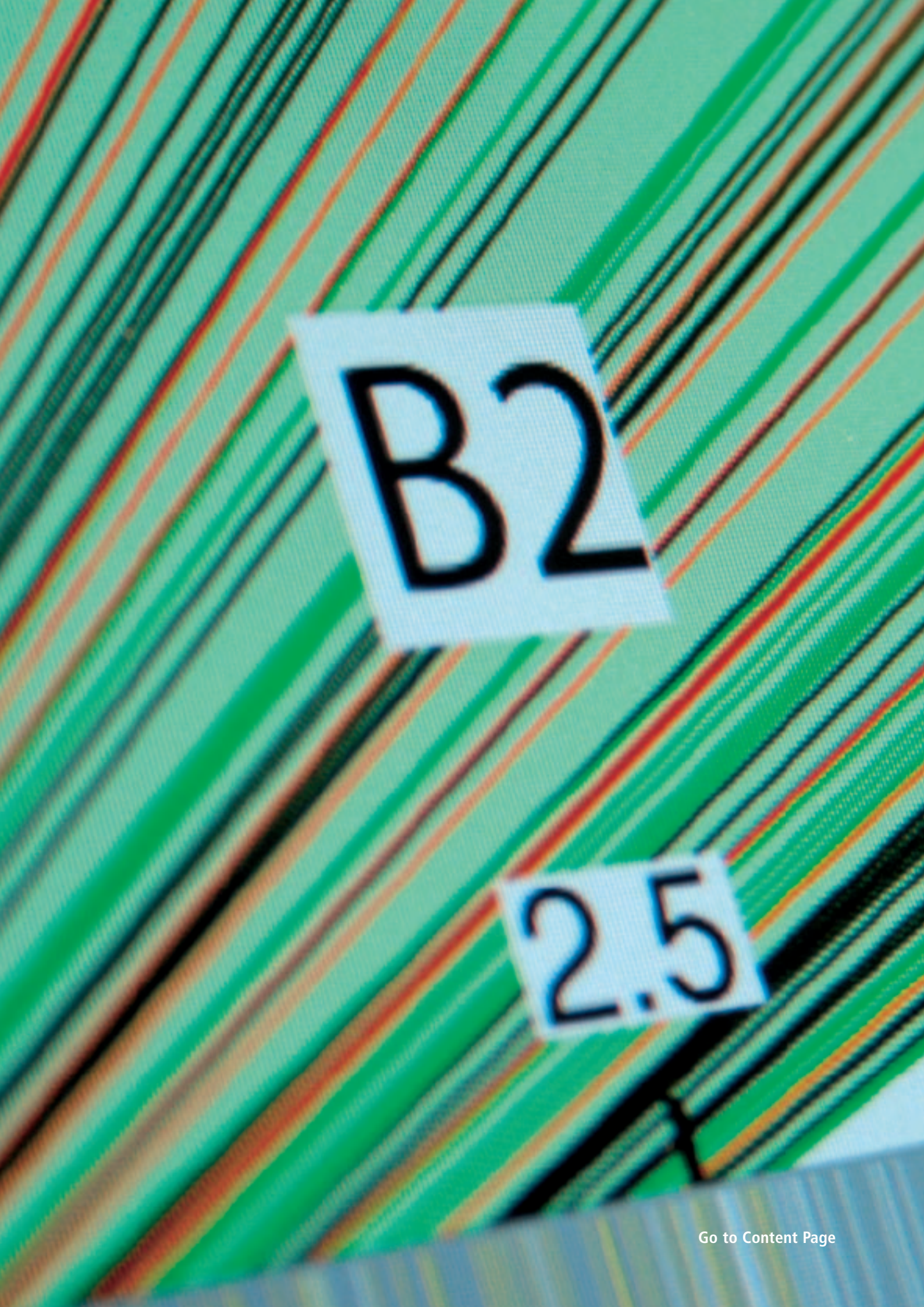
As patenting is only one route for enabling technology transfer and securing intellectual property rights, PLA also is engaged in setting up and stimulating alternative ways for efficient utilization of results, first of all by direct participation of member companies within GABI projects, including individual project agreements.

Being a link between academia and industry PLA also takes an active part within the boards and panels of GABI, participating in the meetings of the GABI Steering Council, the Scientific Advisory Board and the Scientific Coordination Committee. Furthermore PLA keeps close contact to the GABI secretary on the one hand and the GABI funding agency PtJ (Projektträger Jülich) on the other.

In order to guarantee a productive and trustful cooperation between the different partners, a comprehensive legal framework for GABI has to be built up and continuously updated. By support of a legal task group established by the WPG, named JurAG, PLA keeps a central role in designing and negotiating cooperation contracts for all joint projects. This is of particular importance for the international cooperation of GABI, first of all with Génoplante in France, followed by the extension to a trilateral cooperation including Spanish partners and also for the Canadian German Agrobiotech Teamworking (CGAT). In all cases, PLA takes a leading part on the German side for setting up the necessary legal agreements.

The next step with respect to these international activities is covered by a broader European dimension. PLA is actively contributing to the European Technology Platform Plants for the Future being a member of its Steering Council. On the other hand PLA also supports the initiative European Research Area Plant Genomics (ERA-PG) with special emphasis on contributions for the IP working group of ERA-PG.

In summary, during the past years PLA has become one essential part of the infrastructural fundament on which a success story like GABI could grow.



Go to Content Page

Bioinformatics

The increasing amount of biological data made the development of new ways of data mining and management imperative. Today's bioinformatics are an invaluable tool for the handling, interpretation and presentation of the huge datasets emerging from cutting-edge technologies applied not only in GABI projects. The range of implements includes innovative databases, statistical toolsets as well as automated analysis algorithms specifically designed for the needs of biological research.

Go to Content Page

GABI-PD: Integration of plant genome data

GABI-PD
Bioinformatics

Birgit Kersten*, Axel Nagel, Rico Basekow, Svenja Meyer
RZPD German Resource Center for Genome Research, Heubnerweg 6, 14059 Berlin
* corresponding author: kersten@rzpd.de, phone:+49-30-32639-261

Introduction

GabiPD, GABI's primary database is a widely-appreciated platform to centrally store and retrieve plant specific data. User-friendly graphical interfaces allow effective data searches via the internet (http://gabi.rzpd.de). The web-statistics show an constantly increasing user access which is being further promoted by GabiPD's participation in an international effort on Arabidopsis WebServices (http://gabi.rzpd.de/projects/webservices/).

Genomic, transcriptomic and proteomic data from different GABI groups are integrated in GabiPD as well as linked with each other and with public data. Comprehensive analysis of the stored data, such as sequence BLAST, clustering and ORF prediction is performed (Fig. 1). The analysis results are integrated, visualized and linked to the original data. Links to providers of GABI plant material are made available via the corresponding data (Fig. 1).

GabiPD supports GABI participants in bioinformatic analysis of their data and in data submissions to public repositories. Furthermore, specific database applications, such as BreedCAM (http://gabi.rzpd.de/projects/breed/) have been developed by the GabiPD team in close cooperation with GABI partners. All these activities are promoting the publication of GABI data and the extension of GabiPD as a reference system for publications.

Data integration and GreenCards

GabiPD acts as a flexible platform to the GABI/WPG community, providing access-controlled sharing of data thus allowing the exploitation of GABI-data in context of public data. The high flexibility of the software structure facilitates easy adaptation to new data types. Program modules, which allow a semi-automated data handling, ensure that data can be integrated fast and efficiently. Different graphical user interfaces can be used for data visualization and searches.

GreenCards is the main search- and visualization-interface for text-based data. GabiPD can be queried for clones, sequences, insertion mutants, SNPs, proteins, mapping information or BLAST results by entering, for example, sequence accessions, clone names, gene, marker, or genotype names, gene functions or simply keywords. Query hits are shown in an overview and can be selected for a detailed view to get numerous information on each search result. By a statistical function, all search hits are shown in a hierarchical order grouped by cultivar. Furthermore, download of search results in different formats is provided via GreenCards.

Genomic data in GabiPD

Genomic sequence information, e.g. from potato is accessible via GabiPD. If available, SNP/InDel information is provided and visualized with the sequences. SNPs are also displayed in the original traces. PoMaMo (Potato, Maps and More), a database specialized for potato genome data, was established within GabiPD in close cooperation with GABI Conquest (PD Christiane Gebhardt) (Meyer *et al.* 2005) (https://gabi.rzpd.de/projects/pomamo.shtml). In PoMaMo, genetic maps of potato chromosomes I to XII as well as Solanaceen function maps for pathogen resistance are visualized (Fig. 2). Recently, barley genetic maps have been integrated in GabiPD in cooperation with Dr. Nils Stein from the IPK, Gatersleben.

Mapping data are visualised via YAMB (Yet Another Map Browser), specifically developed by GabiPD (Meyer *et al.* 2005). Maps of different strains and consensus maps can be displayed in parallel, with same elements on homologous chromosomes linked with each other. The maps are interactive, that is all elements displayed on the maps are click able to provide more detailed information (Fig. 2).

EST sequence information and analysis

Comprehensive EST sequence information is displayed in GabiPD. The sequences are provided with EST clone descriptions in a text format (Fig. 1). In addition trace file views of the original sequences are accessible via GreenCards. Linked to the sequence data, results on different analysis are visualized. Beside BLAST analysis, which is being updated at least once a year, ORF prediction in all six frames and often results on cluster analysis are provided (Fig. 1). This complex data integration in a context allows for new interpretations of the data.

O. Thimm *et al.* **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes.** Plant Journal (2004) 37, 914-39. O. S. Meyer, A. Nagel and C. Gebhardt **PoMaMo-a comprehensive database for potato genome data.** Nucleic Acids Research (2005) 33 (Database issue), D666-70. O. A. Nagel *et al.* **Extension of the Visualisation Tool MapMan to Allow Statistical Analysis of Arrays, Display of Coresponding Genes and Comparison with Known Responses.** Plant Physiology (2005) 138, 1195-1204. O. T. Feilner *et al.* **High throughput identification of potential Arabidopsis mitogen-activated protein kinases substrates.** Molecular and Cellular Proteomics (2005) 4, 1558-68. O. E. Urbanczyk-Wochniak *et al.* Conversion of MapMan to Allow the Analysis of Transcript Data from Solanaceous Species: Effects of Genetic and Environmental Alterations in Energy Metabolism in the Leaf. Plant Mol Biol. (2006) 60, 773-92.

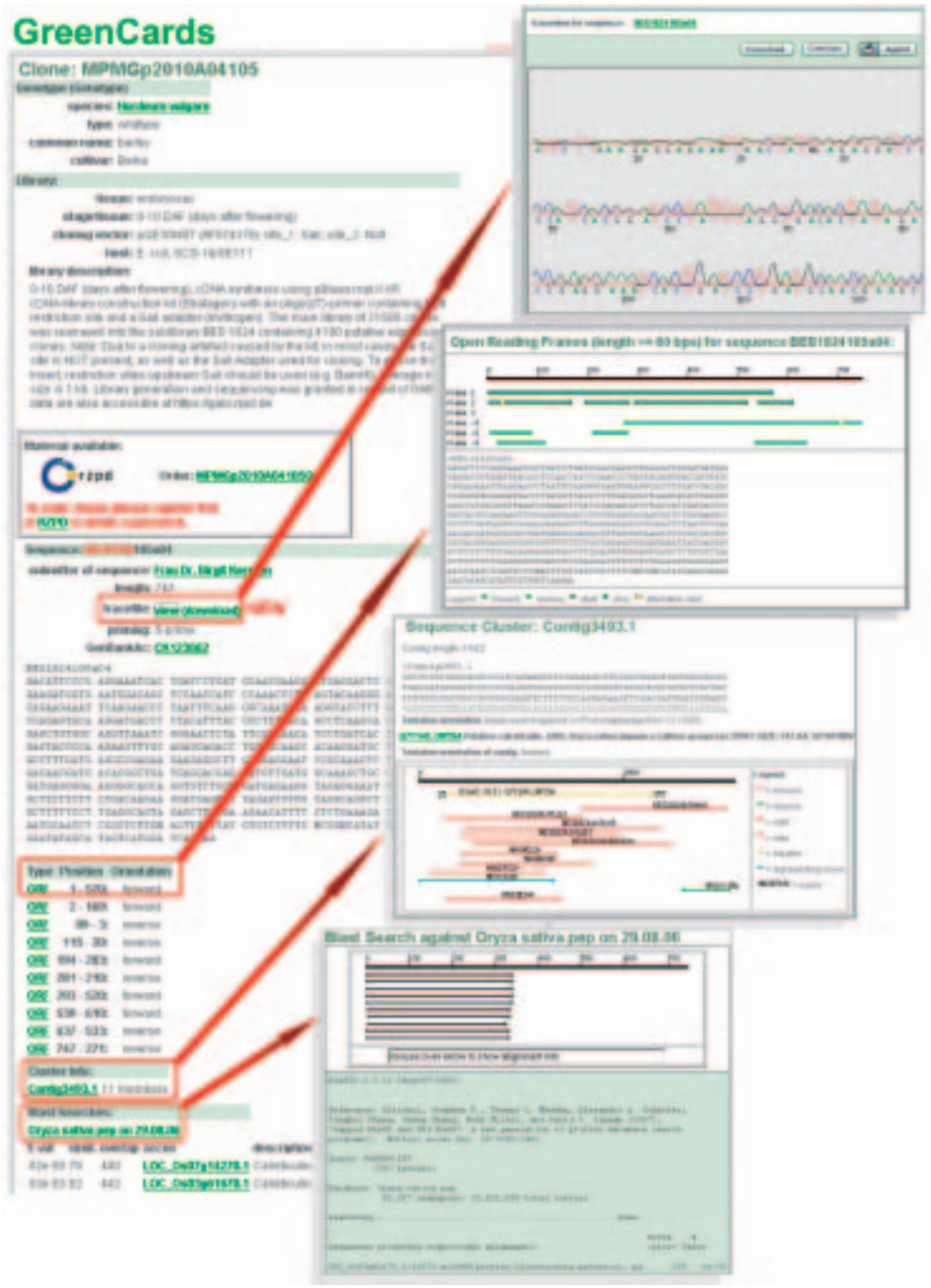


Fig. 1: *Data integration via GreenCards. Description of a barley EST clone in GabiPD including sequence (left). Access to the original trace file is provided (top right). Results on sequence analysis (right): ORF (open reading frame) predictions; cluster results including the consensus sequence of the clustercontig as well as display of the alignment of all clustercontig member sequences on the consensus; BLAST results. The Library MPMGp 2010 has been generated within GABI-Seed in cooperation with GABI-Lapp.*

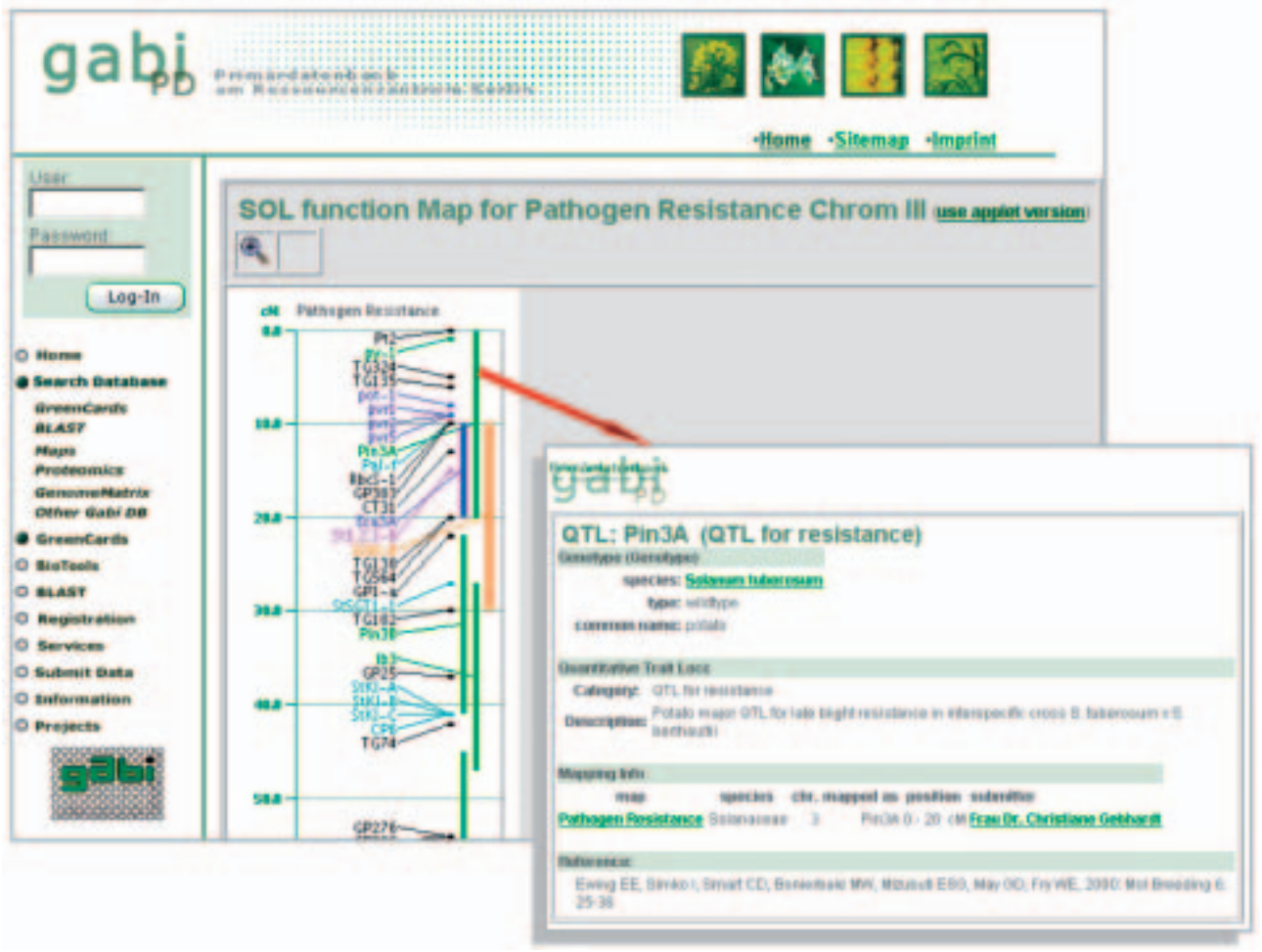


Fig. 2: *Solanaceae function maps for pathogene resistance (chromosome 3). Different type of resistance markers are depicted on the chromosome in different colors, QTLs as bars. Interactive access to detailed marker descriptions. The maps have been integrated by GABI PD in cooperation with GABI Conquest.*

Moreover, GabiPD performs project-specific sequence data analysis in cooperation with GABI-partners. Arabidopsis sequence analysis, e.g., led to the generation of an Arabidopsis Unigen set for protein expression (Feilner *et al.* 2005). This set has been used to develop the first plant protein microarrays for phosphoproteomic studies. Furthermore, barley EST sequences from the IPK Gatersleben have been clustered using the TIGR software TGICL and CAP3 in order to create an UNIgene set from barley.

MapMan – Mapping of expression profiling data onto selected pathways

Gene expression data, such as Arabidopsis Affymetrix data are integrated in GabiPD and displayed via GabiPD's web page by means of the MapMan tool (Web version). The Java program MapMan, which has been developed in the scope of the GABI-MapMan project (Prof. Mark Stitt), allows for mapping large data sets, such as expression profiling data, onto known pathways and cellular processes (Timm *et al.* 2004,

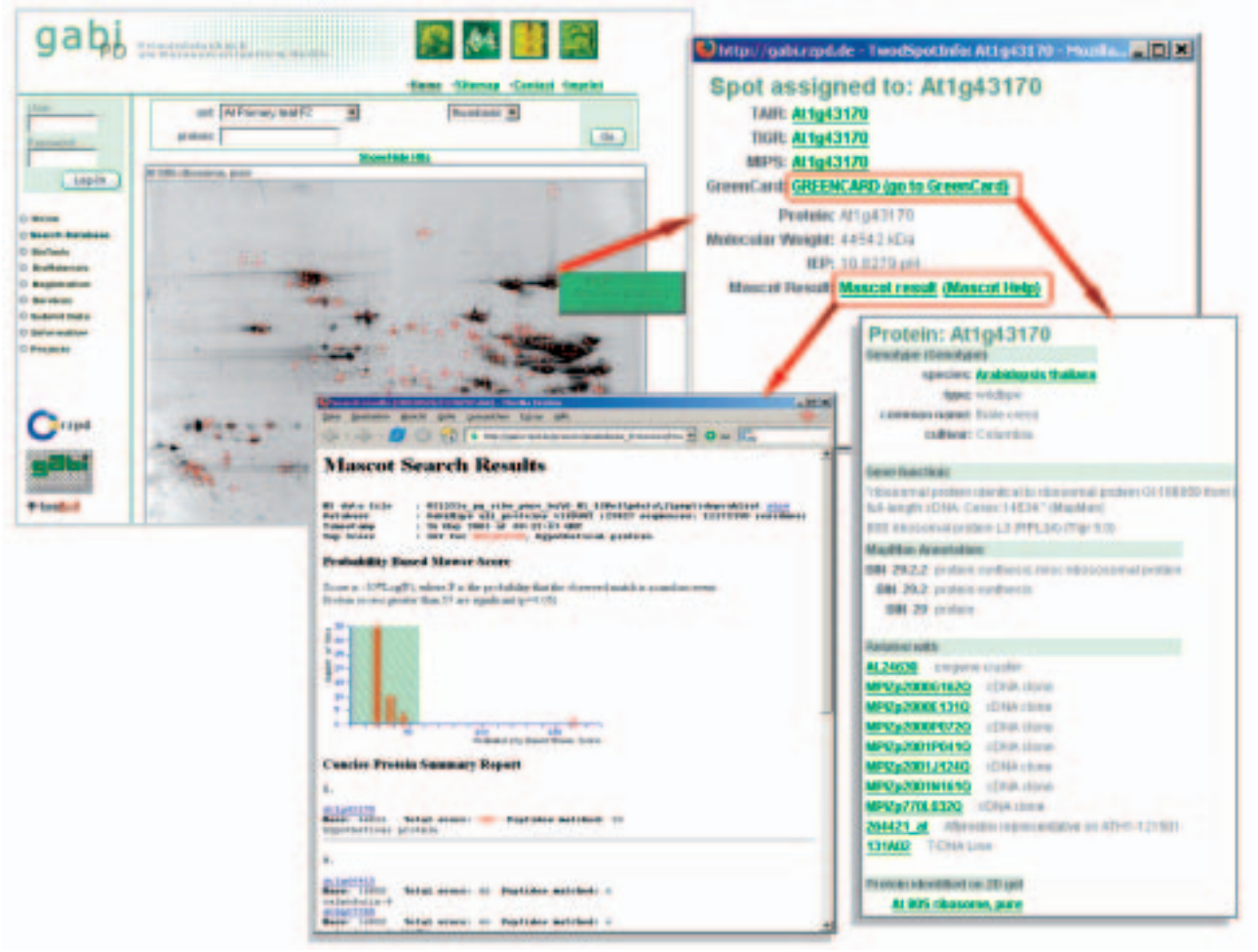


Fig. 3: *2DE image from Arabidopsis with spot annotations. Mouse-over function gives putative protein functions for the identified spots. On click, detailed protein information is provided, e.g. mascot search results of the MS data.*

Nagel *et al.* 2005, Urbanczyk-Wochniak *et al.* 2006). The user can select a pathway, e.g. glycolysis, and an experiment from a menu and can view up- and down-regulated genes in the context of the selected pathway. The desktop application of MapMan is provided via GabiPD for download (http://gabi.rzpd.de/projects/MapMan/).

Plant proteomics in GabiPD

Proteomic data from different Arabidopsis and rape tissues have been integrated in GabiPD in close cooperation with Dr. Patrick Giavalisco (GABI-Lapp) and Dr. Julia Kehr. Provision of innovative interfaces allow to view 2D-gel images in an interactive way (Fig. 3). Mouse-over function reveals putative protein functions for the identified spots in the 2-DE gel images.

Fig. 4: *GabiPD-BioTools for online sequence analysis. 11 different tools are provided, e.g. "EMMA" for the alignment of multiple DNA or protein sequences.*

On click, detailed information is accessible for the respective spot, such as MS analysis data, molecular weight and isoelectric point of the respective protein. Furthermore, information on Arabidopsis cDNA expression clones (Feilner *et al.* 2005) is provided at GabiPD's proteomics project page (http://gabi.rzpd.de/projects/Arabidopsis_Proteomics/). These clones, which are available via RZPD are useful to express Arabidopsis proteins of interest in *E. coli*, e.g. for the generation of specific antibodies or for further functional protein studies with the recombinant Arabidopsis proteins.

BioTools – Online tools for sequence analysis

Numerous tools from the EMBOSS package have been integrated in a web application, thus allowing online analysis of private sequence data via GabiPD (BioTools, http://gabi.rzpd.de/biotools). 11 different BioTools are currently available (Fig. 4), such as a tool for determination of restriction enzyme cleavage sites. DNA and protein sequences can be aligned (Fig. 4), translated etc. The BioTools can be further extended according to requests and requirements of the GABI-community.

A User-Driven Tool to Organise and Map Large Data Sets like Transcript and Metabolite Profiles onto Known Pathways and Cellular Processes: Tuning in Arabidopsis and Application to Support Genomics Research in Crops

MAPMAN
Bioinformatics

Mark Stitt 1; Björn Usadel 1; Jan Hanneman 1; Yves Gibon 1; Joachim Selbig 3; Dirk Walther 1; Daniel Weicht 1; Axel Nagel 2 and Birgit Kersten 2
1 Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam-Golm,
2 RZPD German Resource Center for Genome Research, Heubnerweg 6, 14059 Berlin
3 University of Potsdam, 14476 Potsdam-Golm

Introduction

Functional genomics opens many opportunities, but is limited by our ability to handle huge amounts of data, relate it to existing knowledge, and extract novel information. GABI-MAPMAN has established tools to support the visualization and interpretation of high throughout functional genomics experiments, including expression, protein, enzyme activitiy and metabolite profiling, in the context of pathways and other existing knowledge (Thimm *et al.*, 2004; Usadel *et al.*, 2005; Urbanczyk-Wochniak *et al.*, 2006). The goal was to provide a simple modular system which users can modify to meet their own needs.

MAPMAN: Visualization of the results form individual experiments

The core of the MAPMAN application is a set of SCAVENGER modules and the IMAGEANNOTATOR module. SCAVENGERs organize genes, enzymes, proteins or metabolites into hierarchical functional categories ('BINs'). The resulting ontologies ('mapping files') are imported into IMAGEANNOTATOR, and used to organize profiling data and visualize it via a false-color code on diagrams of biological processes ('maps'). The first release occurred in 2004, involving a publication (Thimm *et al.*, 2004) and the establishment of a web site (http://gabi.rzpd.de/projects/MapMan/) at which a web-based version can be used, and the application can also be downloaded to be installed and used locally. An extensive description of the application and help-pages are provided, as well as contacts for support. There have been over 2000 downloads of the application, and 5-6 requests for support are received each week.

SCAVENGER modules: SCAVENGER modules classify measured parameters into functional categories (BINS), which are structured in a hierarchical tree, termed a 'mapping file'. Each BIN and sub.BIN is identified by a unique numerical identifier, and these reflect the conceptual structure of the hierarchy (Fig. 1). A main aim is to minimize redundancy; this aids statistical analysis and compact visualization, and distinguishes them from GO. The 'mapping files' are distributed to the user as an xml, xls or txt file.

TranscriptSCAVENGERS classify genes into functional ontologies, and are used to support visualization and analysis of expression profiling data. The first TranscriptSCAVENGER was created for Arabidopsis, the first completely sequenced and best annotated, plant genome. A skeleton was developed using the TIGR2 annotation (Thimm *et al.*, 2004). Continual updating using literature, expert inputs and TIGR5 (Usadel *et al.*, 2005) resulted in re-assignment of 1000s of genes, and expansion of the ontology from 197 to >1000 categories.

TranscriptSCAVENGERS were subsequently created for tomato and potato (Urbanczyk-Wochniak *et al.*, 2006), maize (F. Poree and B. Usadel, unpublished, cooperation with GABI-COOL), barley (N. Sreenivasulu and B. Usadel unpublished cooperation with GABI-SEEDS). Collaborators are adapting MAPMAN to poplar (UPSC, Sweden), Medicago (Ardmore/Uni Bielefeld), soybean (Illinois State USA/CSIRO Australia) and potato and grape vine (Slovenia).

To transfer the BIN structure to crop and non-crop plants we developed a semi-automatic pipeline: a) automatic transfer of the gene identifiers to the new array, b) blasting sequences for the new species against the reference (Arabidopsis) c) blasting against the non-redundant NCBI protein database and evaluating the annotation pattern with string matching, d) domain scanning and GO term comparison and e) manual curation. New BINs can be created (e.g., in barley, inhibitory proteins are important in seed development). A more advanced tool is being developed to support storing, entering, deleting and modifying entries. It allows changes to be tracked, linked to an individual annotator, and reversed without loosing data.

MetaboliteSCAVENGER covers metabolites routinely detected by GC-TOF and LC-MS platforms (Thimm *et al.*, 2004, Kopka *et al.* 2005, and Gibon unpublished). ProteinSCAVENGER assigns ~60,000 proteins from various plants. EnzymeSCAVENGER covers 45 enzymes analyzed by a robotics-based platform (Gibon *et al.*, 2004 and Gibon unpublished).

○ Bläsing, O. Gibon, Y., Günther, M., Höhne, M., Osuna, D., Thimm, T., Scheible, W.-R., Morcuende, R. and Mark Stitt, M. (2005) **Sugars and Circadian Regulation make Major Contributions to the Global Regulation of Diurnal Gene Expression in Arabidopsis.** Plant Cell 17, 3257-3281 ○ Gibon, Y., Bläsing, O, Hannemann, J., Carillo, P., Höhne, M., Cross, J., Selbig, J., Stitt, M. (2004) **A robot-based platform to measure multiple enzyme activities using a set of cycling assays: comparison of changes of enzyme activities and transcript levels in Arabidopsis during diurnal cycles and in prolonged darkness.** Plant Cell: 16, 3304-3325 ○ Kopka. J., Fernie, A.F., Weckwerth, W., Gibon, Y., and Stitt, M. (2004) **Metabolite profiling in Plant Biology: Platforms and Destinations.** Genome Biology 5, 109 ○ Osuna, D., Usadel, B., Morcuende, R., Scheible, W:-R., Gibon, Y., Bläsing, O-E., Thimm, O., Höhne, M., Günter, M., Udvardi, M.K., Kamlage, B., Trethewey, R. and Stitt, M. (…) **Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings.** Plant J. In press ○ Thimm O., Bläsing O., Gibon Y. et al (2004) **MapMan: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes.** The Plant Journal 37, 914-939 ○ Usadel, B. ,Nagel, A., Thimm, O. *et al.* (2005) **Extension of the Visualisation Tool MapMan to Allow Statistical Analysis of Arrays, Display of Co-responding Genes and Comparison with Known Responses.** Plant Physiology 138, 1195-1204 ○ Urbanczyk-Wochniak, E., Usadel, B., Thimm, O. *et al.* (2006) **Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: effects of genetic and environmental alterations in energy metabolism in the leaf**. Plant Molecular Biology. 60, 773-792

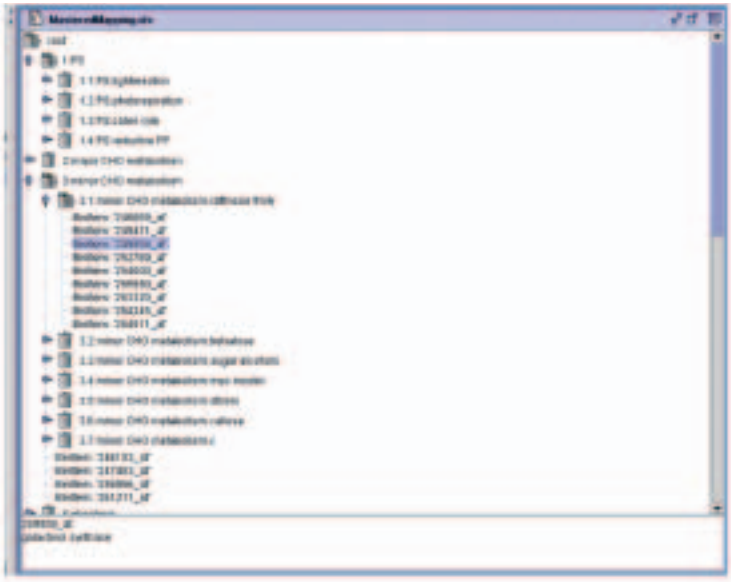


Fig. 1: *The panel to the left shows an excerpt of the BIN / sub-BIN structure. The excerpt shows BIN 1 (Photosynthesis), BIN 2 (Major carbohydrate metabolism), and BIN3 (Minor carbohydrate metabolism). Sub-BIN 3.1 (raffinose metabolism) is expanded to the level of single genes. The mouse over action to check assignment is illustrated for 249804_at, which encodes a galactinol synthase (see display at bottom of the screen): The lower panel shows a screen shot of the mapping file, with gene annotation.*

4.0.9	glycolysis.phosphoglycera	257699_at	phosphoglycerate kinase, putative similar to phosphoglycerate kinase, chloroplas
4.0.10	glycolysis.phosphoglycera	245689_at	phosphoglycerate mutase - like protein
4.0.10	glycolysis.phosphoglycera	262180_at	phosphoglycerate mutase, putative similar to phosphoglycerate mutase GI:89792
4.0.10	glycolysis.phosphoglycera	252169_at	putative protein several bacterial phosphoglycerate mutases; supported by cDNA:
4.0.11	glycolysis.enolase	266286_at	putative enolase (2-phospho-D-glycerate hydrolyase) ; supported by cDNA: gi_14
4.0.11	glycolysis.enolase	260392_at	putative enolase similar to ENOLASE (2-PHOSPHOGLYCERATE DEHYDRATAS
4.0.11	glycolysis.enolase	263924_at	enolase (2-phospho-D-glycerate hydrolyase) ; supported by cDNA: gi_15809969_
4.0.12	glycolysis.PK	263922_s_at	putative pyruvate kinase ;supported by full-length cDNA: Ceres:120685.
4.0.12	glycolysis.PK	258808_at	putative pyruvate kinase similar to pyruvate kinase, cytosolic isozyme GB:Q4295
4.0.12	glycolysis.PK	258065_at	putative pyruvate kinase similar to PYRUVATE KINASE, CYTOSOLIC ISOZYME
4.0.12	glycolysis.PK	247338_at	pyruvate kinase
4.0.12	glycolysis.PK	250526_at	pyruvate kinase
4.0.12	glycolysis.PK	248283_at	pyruvate kinase ;supported by full-length cDNA: Ceres:109919.
4.0.12	glycolysis.PK	247989_at	pyruvate kinase ;supported by full-length cDNA: Ceres:31580.
4.0.12	glycolysis.PK	251777_s_at	pyruvate kinase - like protein pyruvate kinase, Nicotiana tabacum, PIR:S41379
4.0.12	glycolysis.PK	254009_at	pyruvate kinase like protein pyruvate kinase, Nicotiana tabacum, PIR2:S41379
4.0.12	glycolysis.PK	252300_at	pyruvate kinase -like protein various pyruvate kinases from procaryotes
4.0.12	glycolysis.PK	260653_at	pyruvate kinase, plastid isozyme, putative similar to GB:Z28374 from [Nicotiana t
4.0.12	glycolysis.PK	256836_at	pyruvate kinase, putative similar to pyruvate kinase isozyme A, chloroplast precu

IMAGEANNOTATOR: In the IMAGEANNOTATOR module, the ontologies generated by the SCAVENGER modules are used to organize and visualize data on diagrams ('maps') of biological pathways or processes (Fig. 2). It is distributed to the user, and is also available as a web-resource. It is written in JAVA to make it compatible with multiple operating systems.

The experiment file that is to be viewed is opened via a browser, together with an appropriate map and mapping file. In the display, each individual gene is displayed as a single feature, and all the genes in a particular BIN or subBIN are grouped together, like a floodlight. Responses are shown via a color scale. Depending on the level of resolution of the mapping file and display, large numbers of genes that are involved

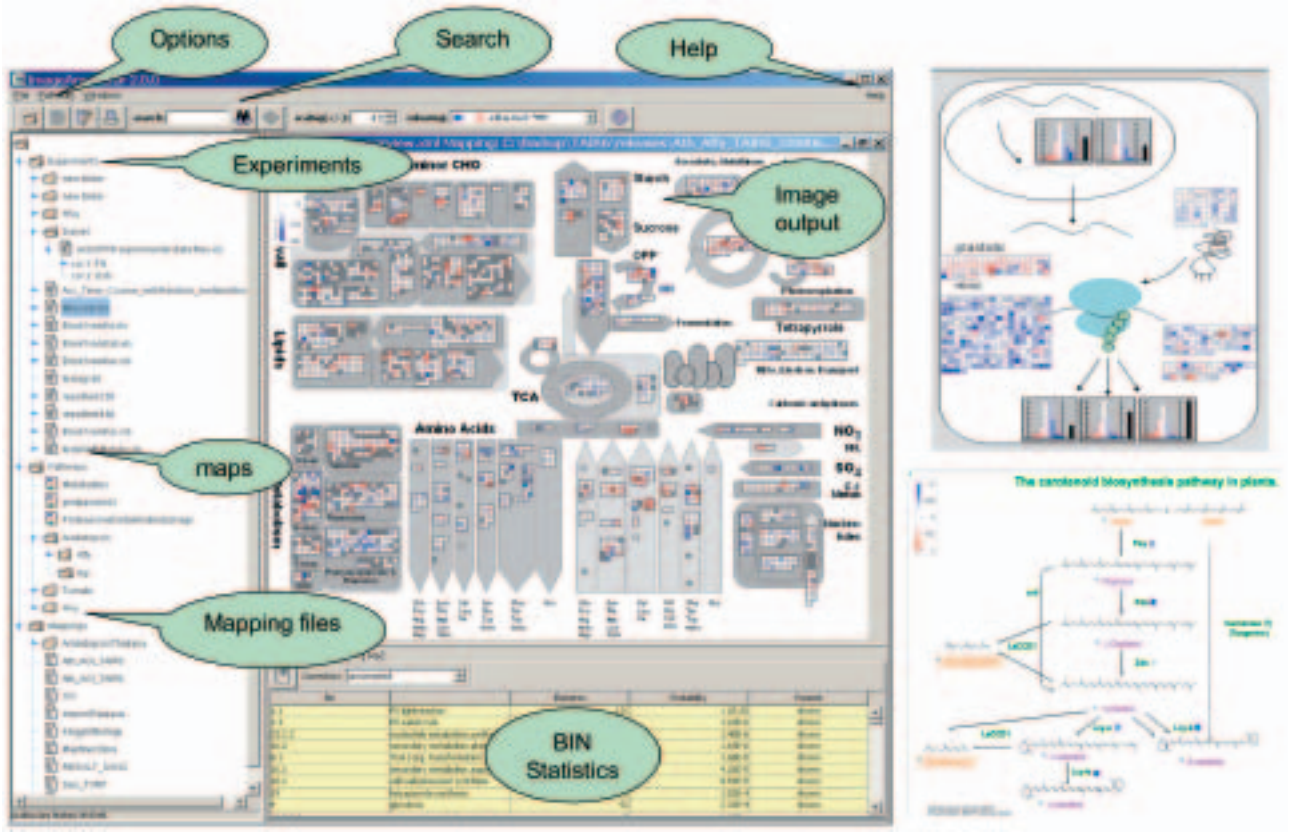


Fig. 2: *Screenshot of the IMAGEANNOTATOR interface, showing transcript profiling data.*

in a process may be grouped together, or genes can be subdivided down to partial processes or even to the level of the individual enzyme or protein. Individual features can be queried by mouse-click; information pops up (gene, BIN; annotation, response).

When a new map is made, it must first be customized. This is done via the IMAGEANNOTATOR interface. For each BIN or subBIN that is to be visualized, the user clicks at the position where the BIN or subBIN should be shown to activate a pop-up box, into which the numerical identifier is entered. This generates an xml file, which can be saved and is automatically opened when a map and the corresponding mapping file are selected in the browser. These files can be modified by the same procedure. For maps provided at the website, the xml files have already been generated, providing the user with a 'ready-to-go' tool. At present ~50 maps can be downloaded; this is continually expanding and users can add their own 'maps'

Following the first release of IMAGEANNOTATOR, it has been continually improved to provide more options and enhance the user interface. The browser is organized to group mapping files and 'maps' by organism and/or array-types to facilitate species- or array-specific customization. This is important for users who work with several species/ types of array. To allow rapid inspection of the growing number of maps, a mouse-over option opens up a 'thumbs-nail' miniature of a map (Fig. 3). The color scheme and scale can be chosen; and is displayed in a legend. The current build accepts datasets that have been processed in many different ways: as ratios, statistical analyses like p-values (Usadel *et al.*, 2005) or more sophisticated evaluations (see Bläsing *et al.* 2005). The modular structure accepts transcript, metabolite, protein and enzyme data; it is also possible to visualize multi-level data, with each data sort being distinguished by a different symbol. On-line calculation of Wilcoxons p-values provides a global overview; it shows if the response of genes in a particular BIN

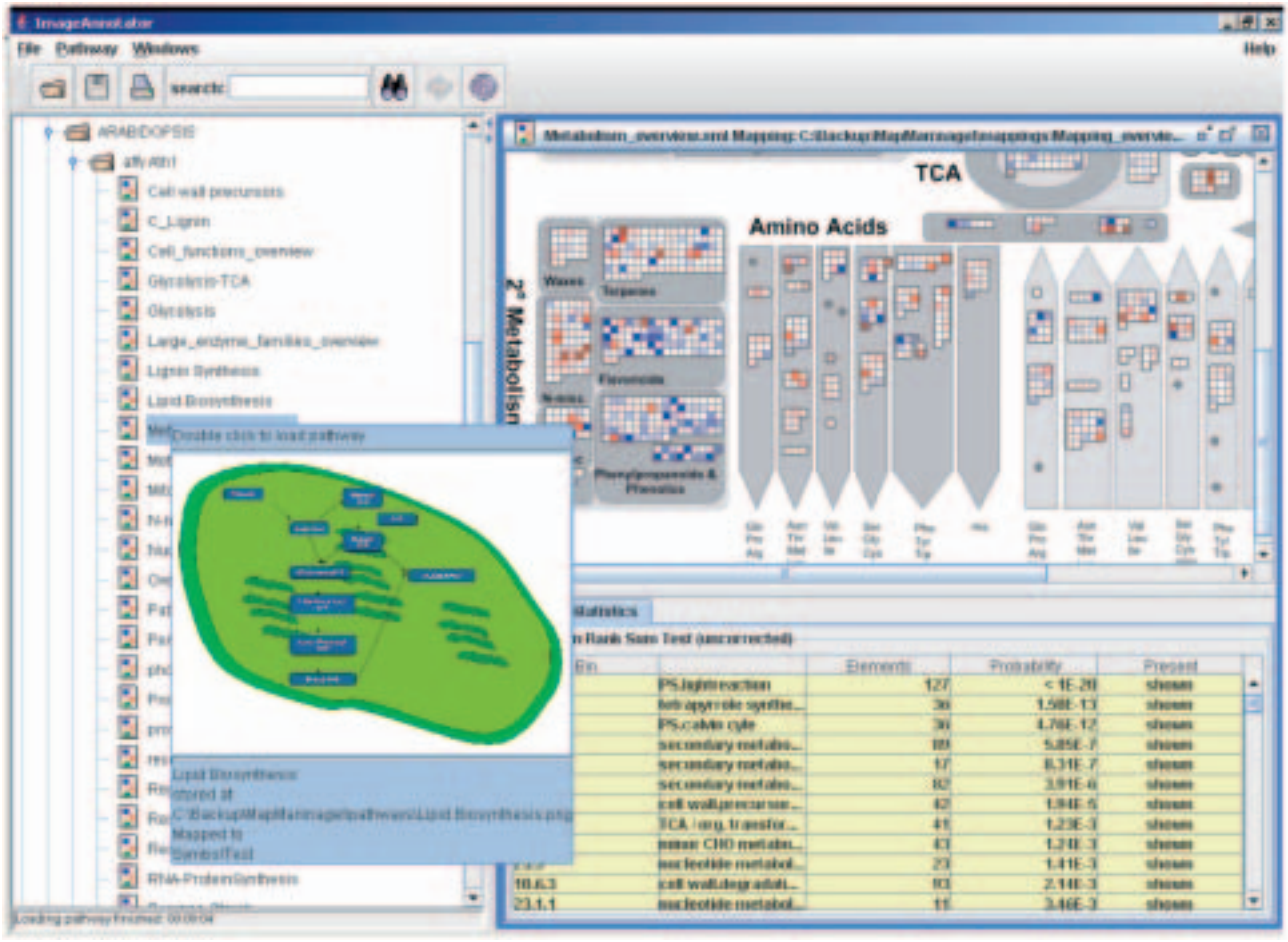


Fig. 3: *Screenshot of the IMAGEANNOTATOR interface, illustrating the pop-up thumbnail sketch of a map, and the option for viewing of the ranking of categories by the Wilcoxons test.*

is significantly different to all other genes on the array (Fig. 3). On-line filtering can be performed based on the quality of the data; for example, all genes can be grayed out whose expression is low, or whose expression is not significantly changed.

A major future aim is to increase the amount of functional information that users can access online, and combine with their data. To this end, a series of activities have been initiated: For example, an *on-line filtering option can be used to highlight genes that have a particular attribute*. A first build queries a background spreadsheet with digitalized values for various attributes of the gene or gene product. This will allow, for example, all genes to be highlighted that are thought to be located in the plastid, or all genes to be highlighted that are expressed in a particular organ. We are also adding ***BioMoby connectors***, and building links to large databases that provide more information about the attributes of genes (Fig. 4).

PAGEMAN: The IMMAGEANNOTATOR application visualizes the responses of 1000s of individual genes to a particular treatment. However, it cannot be used to easily compare responses between many different treatments. Tools like Genevestigator or AtGenExpress display the response of single genes-of-interest across large numbers of treatments, but do not allow a global comparison of the responses across all the treatments. One important future aim is to develop tools that compare global responses across large numbers of arrays. This will allow visualization of time course and treatment intensity experiments, as well as comparison of large numbers of pro-

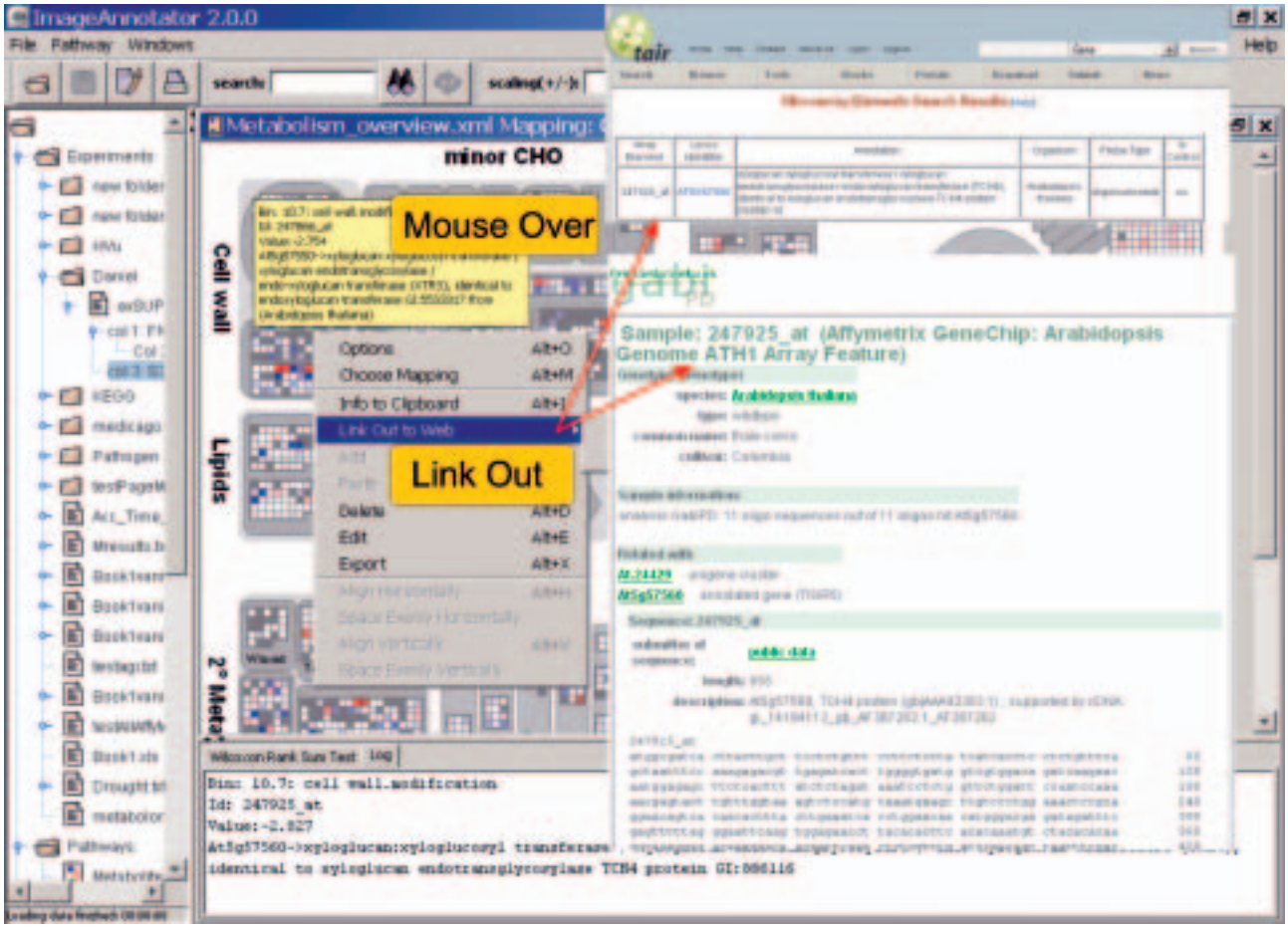


Fig. 4: *Screenshot of a visualization of transcript data illustrating the link-out to external databases.*

files to detect shared general responses. It may also be useful for comparing species responses, because it will not always be possible to establish 'true' homologs for comparisons at the single gene level, and because cross-species comparison will benefit form the ability to abstract fro several independent experiments for each species.

Our first approach (**PAGEMAN**) uses a combination of statistics and ontology to condense array datasets down to a small number of features, which provide information about biological response. It statistically evaluates the response of the groups of genes in each of the ~1000 hierarchical categories of the MAPMAN ontology, using Wilcoxons test or Fisher enrichment. This compresses the >20,000 data points from a microarray to a relatively small number of values. These can be exported as a table, or visualized in one dimension using false color (see Fig. 5). Sections of the ontology that do not show significant changes or are not of interest can be collapsed. Interesting features can be mouse-clicked to annotate them with the name of the biological categories and the p-value/enrichment. Multiple arrays can be lined up (like a series of PAGE gels) and visually inspected to identify similarities/differences in the global response (see e.g., Osuna *et al.*, in press). We showed that common features of diurnal changes between Arabidopsis and tomato can be identified by condensing to ontologies, followed by Principal Component Analysis (Urbanczyk-Wochniak *et al.*, 2006). PAGEMAN will greatly facilitate such comparisons. The first build is available at http://mapman.mpimp-golm.mpg.de/pageman.

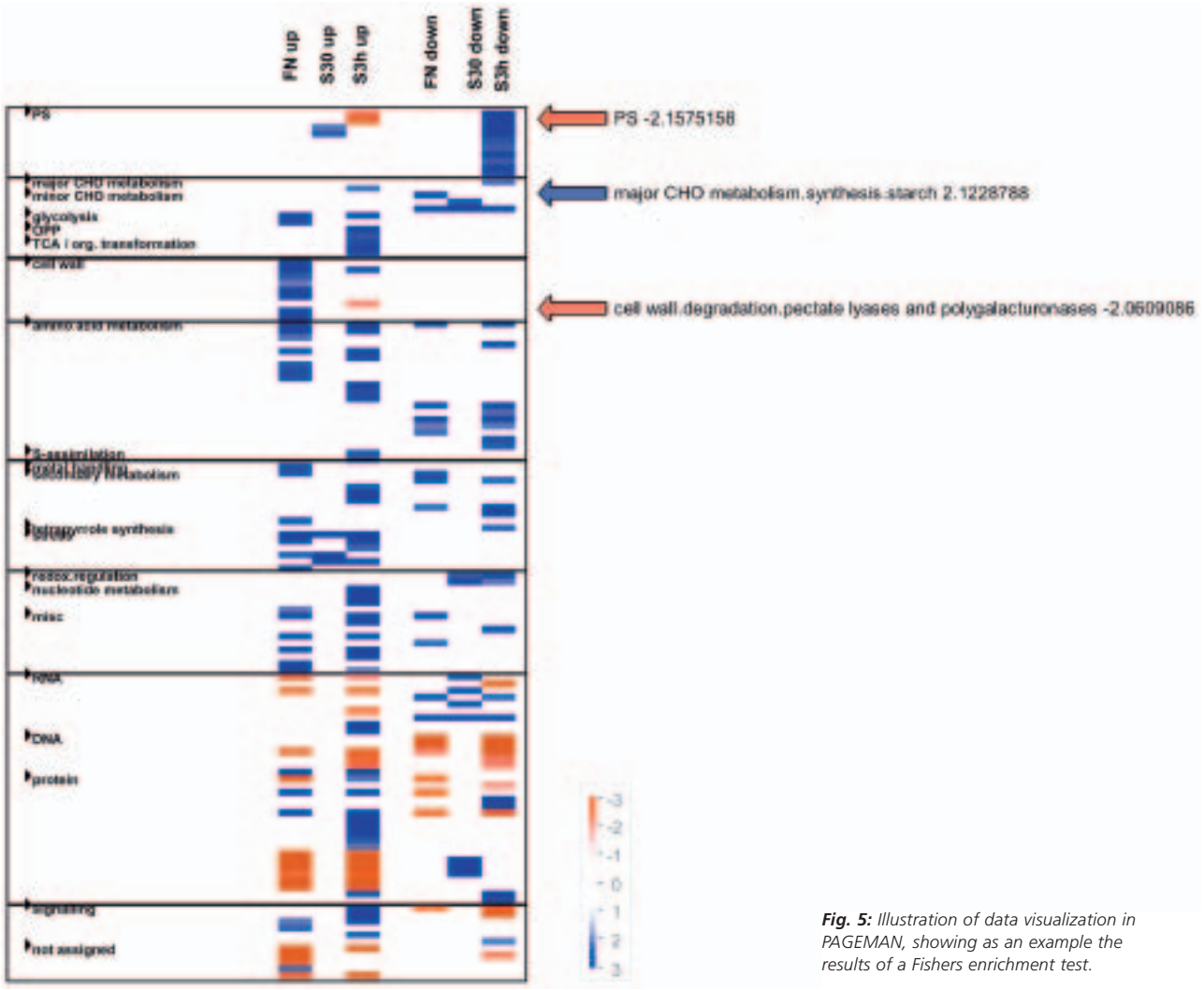


Fig. 5: *Illustration of data visualization in PAGEMAN, showing as an example the results of a Fishers enrichment test.*

GABI Matrix: Bioinformatic Resource for Genome Driven Plant Sciences

MATRIX
Bioinformatics

Georg Haberer, Remy Bruggmann, Heidrun Gundlach, Manuel Spannagl, Octave Noubibou, Hans Werner Mewes, Klaus Mayer
MIPS, Institut für Bioinformatik
GSF – Forschungszentrum für Umwelt und Gesundheit 85764 Neuherberg

Introduction

After the genome sequences of both *Arabidopsis* and rice are already available numerous plant genome sequencing are rapidly progressing. Thus in the near future a bouquet of plant genomes will be fully available. The ongoing projects circumvent both model genomes as well as crop genomes. Beside the need for comprehensive and structured information, genome and knowledge resources for the individual species and their genomes the availability of a range of plant genomes that represent a wide spectra and evolutionary range, bears the promise to undertake previously infeasible detailed and in depth comparative analysis among different species. Without doubt these analysis will give new insights into similarities and dissimilarities as well as specific characteristics of individual plant genomes. In addition these information resources will help to elucidate genetic elements that haven't been discovered thus far or have been difficult to detect.

Within GABI Matrix our work focuses on the different components and pillars necessary in such an "added-value pipeline". We developed a generic plant genome database system, PlantsDB, built on innovative technology platforms and populated with data from finished genome projects and constantly imported data from ongoing genome projects. Constantly updated genome databases and the well defined technical interfaces are fundamental for comparative and combinatorial analysis and the application of new tools and approaches for the analysis of plant genomes.

In the context of this article we will on one hand introduce the PlantsDB genome database concept and on the other use selected examples to illustrate our work in the field of genome analysis and comparative analysis (maize versus rice) as well as on the use of comparative sequence analysis in combination with expression data for the detection of cis elements in plants.

PlantsDB – MIPS Plant Genome Database Resources

Consistent and detailed data resources are a prerequisite for detailed and in depth cross-species comparisons and comparative phylogenetic analysis. Within GABI Matrix we developed PlantsDB. It aims to address this task by applying a generic though highly flexible modular database infrastructure for a wide range of plant genomic data. The respective species databases are updated and new data are continuously integrated either through adjustment against external resources or via the groups participation in a range of plant genome sequencing projects. The rapid cycle of data analysis and inclusion of analytical results into the respective databases thereby warrants rapid availability of the latest analytical results and data.

While individual organism databases provide an important pillar of PlantsDB, the focus of PlantsDB is extending beyond individual genomes. PlantsDB also aims to make available resources that are species spanning and address and support specific questions in comparative and integrative plant genomics. Topics and resources circumvent integrated resources for the detection and analysis of conserved orthologous sequence markers (COS markers), repeat catalogs and classification systems for all plant species, comparative views and search opportunities and a cis element database based on comparative sequence analysis. The PlantsDB resources are completed by the provided BioMOBY based web-service opportunities that support seamless navigation and combination of services provided by PlantsDB and partner databases worldwide. PlantsDB can be accessed at http://mips.gsf.de/projects/plants.

PlantsDB: Data access and retrieval

The MIPS plant genome resources provide access to all genomes included in common formats and similar interfaces. The main entry point for all databases included in PlantsDB can be found at http://mips.gsf.de/projects/plants. The web content is managed using the JBOSS 4.0.1 application server and Java Server Faces. For each genetic element a report page communicates detailed information on the respective entries and links to a list of the respective annotated genetic elements as well as to a graphical viewer. Specific genetic elements can be downloaded using the Genetic Element Retrieval System (GenERSys) tool as well as via web service. Cross-references in the report enable access to associated entries in external databases. Connection to the SIMAP database enables retrieval of sequence homologs from *Viridiplantae*, Fungi and Mammalia. For all genes, unspliced, spliced, coding DNA sequences as well as protein sequences are available and can be retrieved in HTML, XML or in FASTA format.

Together with other user selected tracks, genetic elements

○ 1. Joachim Messing, Arvind K. Bharti, Wojciech M. Karlowski, Heidrun Gundlach, Hye Ran Kim, Yeisoo Yu, Fusheng Wei, Galina Fuks, Cari Soderlund, Klaus F.X. Mayer, Rod A. **Wing, Map-based Analysis of Sequence Composition and Genome Organization of Maize.** Proc. Natl. Acad. Sci. USA, 101, 14349-14359 ○ 2. Georg Haberer, Sarah Young, Arvind Bharti, Heidrun Gundlach, Christina Raymond, Galina Fuks, Ed Butler, Rod Wing, Steve Rounsley, Chad Nussbaum, Bruce Birren Klaus Mayer and Joachim Messing, **Structure and Architecture of the Maize Genome**, Plant Physiology, 139, 1612-1624 (2005) ○ 3. Rémy Bruggmann, Arvind K. Bharti, Heidrun Gundlach, Jinsheng Lai, Ana C. Pontaroli, Georg Haberer, Galina Fuks, Charles Du, Fusheng Wei, Sarah Young, Christina Raymond, Matt C. Estep, Renyi Liu, Jeffrey L. Bennetzen, Agnes P. Chan, Pablo D. Rabinowicz, John Quackenbush, W. Brad Barbazuk, Rod A. Wing, Steve Rounsley, Bruce Birren, Chad Nusbaum, Klaus F.X. Mayer, and Joachim Messing, **Uneven Chromosome Contraction and Expansion in the Maize Genome**, Genome Research, 16 (10), 1241-1251 (2006) ○ 4. Georg Haberer, Michael Mader, Peter Kosarev Manuel Spannagl, Li Yang and Klaus F.X. Mayer, **Genome scale cis element detection through analysis of correlative expression modules and sequence conservation between *Arabidopsis thaliana* and *Brassica oleracea***, Plant Physiology, in press ○ 5. Manuel Spannagl, Octave Noubibou, Dirk Haase, Li Yang, Heidrun Gundlach, Kathrin Klee, Georg Haberer, Heiko Schoof and Klaus F. X. Mayer, **MIPSPlantsDB – Plant database resource for integrative and comparative plant genome research**, Nucleic Acids Research, in press

on a selected contig can be searched and viewed through Gbrowse. Gbrowse is a Generic Genome Browser combining a relational database and interactive web interfaces for displaying and manipulating annotation on genomes.

PlantsDB organism components

Under the umbrella of PlantsDB several individual species databases are contained. While the individual databases are physically separate, database structures are identical and user interfaces and services provided are similar. This is a prerequisite for easy and intuitive navigation as well as for comparative studies. At the time of writing the PlantsDB system comprises genome databases for *Arabidopsis*, *Medicago*, *Lotus*, Maize, rice and tomato. However due to the generic and modular architecture new and upcoming plant genome databases can be rapidly installed.

All of the individual plant genome database instances within the PlantsDB framework share common modules for data display, retrieval and services.

Analysis of the maize genome on completely sequenced genomic sequences

Maize, *Zea mays* L., is one of the most productive crops on earth. It serves as a model grass species to study many basic biological processes such as recombination, transposition, meiosis, paramutation, imprinting, gene expression, and plant development.Sequence analysis of the ends of BAC (BES) and random BAC clones has provided important clues on the increase in genome sizes. Compared to only 1/3rd in rice, the maize genome contains >2/3rd non-genic repetitive DNA elements. Moreover, in maize the majority (2/3rd) of non-genic repetitive DNA elements consists of LTR retrotransposon (64%), while in rice the proportion of DNA transposable is much higher (1.3% versus 10.4%). As a consequence, maize chromosomes have expanded more drastically because retrotransposons in contrast to DNA transposable elements cannot excise from their inserted position and amplify by a replicative mode. Based on alignments of maize LTR sequences, recent calculations have shown that the majority of retrotranspositions have occurred during the last 5 million years subsequent to the hybridization event of the two progenito. Therefore, in addition to genome duplication by whole genome duplication (WGD) subsequent retrotransposition led to further expansion of the maize genome. Besides its increase in size chromosomes underwent breakage and fusion cycles, leading to the formation of a mosaic of different homoeologous segments.

Comparative analysis gave important insights into the preservation of synteny. Although alignments of genetic markers indicated extensive synteny between grass genomes, sequence analysis of orthologous chromosomal regions showed extensive gene movement during speciation. The analysis of the duplicated regions of maize indicates that nearly in half of the cases the second copy of the duplicated genes was lost after tetraploidization. Therefore, maize does not have simply twice as many genes as rice. Based on the size of the transcriptome and an average gene size within 100 random BAC clones, the number of genes has been estimated to range between 42,000 and 56,000 consistent with the observed frequent loss of the second copy of the duplicated genes.

However, the comparison of orthologous regions from rice and maize described so far are based on relatively small chromosomal intervals with very few maize genes present. The maize clones were shotgun sequenced in highly parallel fashion and assembled into contiguous sequences. These sequenced maize regions provided us with a new source of transposable element (TE) structures and 479 non-TE related gene models. We also could determine the overall maize chromosome architecture in respect to the distribution of repeat sequences, genes, and chromatin landscape based on methyl-filtrated sequences. Comparison of these larger duplicated regions in maize showed that maize chromosomes not only expanded, but also contracted, showing the unevenness of insertion and deletions of genes and transposable elements in closely related species.

Cis regulatory element analysis by a combined expression correlation and phylogenetic conservation approach

Brassica enjoys a close evolutionary relationship to *Arabidopsis thaliana*. The two genera separated approximately 12-24 Mya ago. While *Arabidopsis thaliana* serves as a model for many research topics in plant genomics and plant biology *Brassica oleracea* with many subspecies covers a wide variety of commercially important vegetable crops such as broccoli, cauliflower and cabbage. The availability of a whole genome sequence for *Arabidopsis* as well as the large

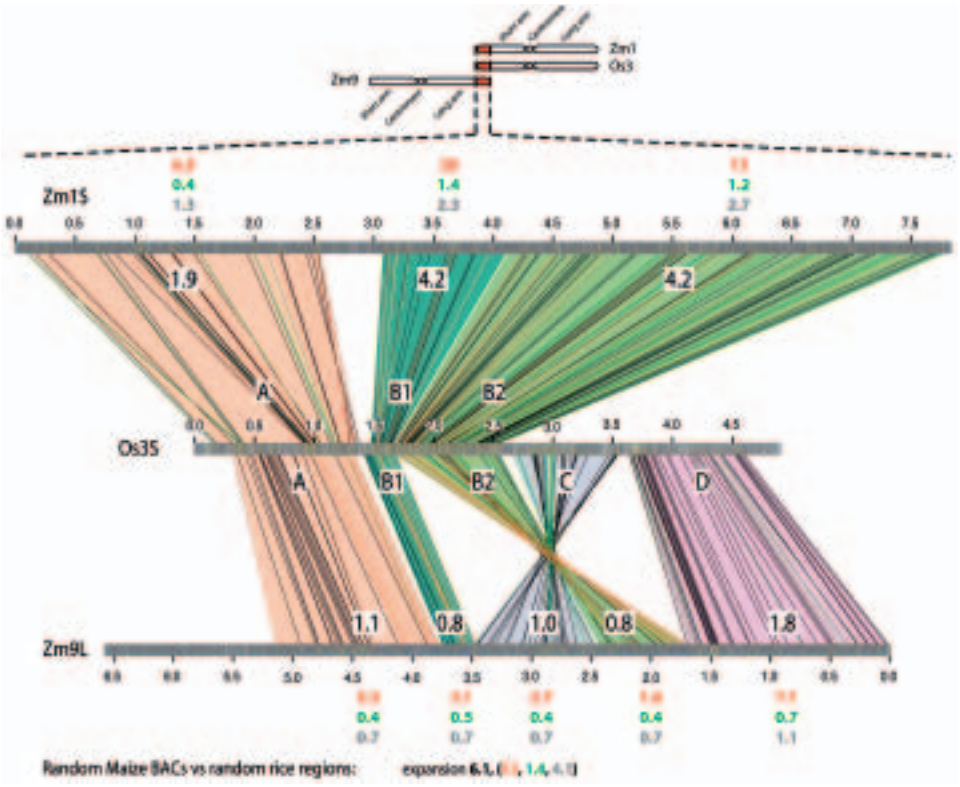


Fig. 1: *Alignments of chromosomal regions from Zm1S, Os3S and Zm9L. The three horizontal lines with scales in mega base pairs represent the chromosomal regions from Zm1S, Os3S and Zm9L. Above those a schematic diagram shows the alignment of these region in respect to their chromosomal locations. Vertical lines show conserved sequences between the chromosomal regions. Syntenic blocks are interrupted by regions without corresponding genes. Syntenic blocks are labeled with A, B1, B2, C, and D. The top line and the bottom line representing the two maize regions show expansion factors (A) for repeat elements (red), genes (green) and unassigned sequences (gray), and (B) for different types of LTR retrotransposons. The summary is given in red, Ty3/gypsy-like elements in green and Ty1/copia-like elements in gray. An insert provides the numbers for the 100 random BAC clones previously analyzed compared against randomly selected rice regions*

amounts of genomic survey sequences for *Brassica oleracea* and their close evolutionary relationship render them a remunerative model for comparative plant genomic studies.

Comparative genomics has been proven to be a powerful tool for the discovery of a large variety of functional elements by their conservation between related species. The utility and potential of *Brassica oleracea* GSS sequences for the improvement of genome annotation in *Arabidopsis thaliana* as well as for comparative studies of the repeat contents of both genomes has been reported.

Comparative genomic approaches have been shown to provide powerful means for the detection of genetic elements that are often difficult to discover by intrinsic measures due to their shortness and/or limited information content. Examples circumvent prominent and important genetic elements like micro-RNAs and *cis-regulatory* elements.

Fundamental for functional and contextual understanding of genes and their embedding and role within regulatory networks is the knowledge of the regulation of individual genes. Powerful technologies to monitor transcriptional states and dynamics are well established and widely applied. For *Arabidopsis*, a large high-quality expression dataset comprising 779 microarrays has been recently made available. These data thoroughly monitor the transcriptional state of the genome under various environmental conditions, for different organs and tissues as well as during distinct developmental phases.

The analysis of co-expressed genes under different conditions and states has been shown to be highly valuable for the analysis of shared *cis-regulatory* elements. A variety of algorithms like expectation-maximization or Gibbs-sampler has been developed to detect motifs which are overrepresented within sets of functionally related sequences.

Complementary, phylogenetic footprinting – and a related approach – phylogenetic shadowing detects *cis-elements* by their conservation between two or more evolutionary related sequences from orthologous genes. It has been successfully applied over a wide range of genera ranging from bacteria to yeasts and mammals. For plants, first pioneering studies have already been undertaken and encouraging results have been reported.

The majority of studies have used either conservation or overrepresentation to discover *cis-regulatory* elements. However, there are some reports which applied a combination of two evidences to evaluate and/or screen detected motifs. So far for plants neither large scale nor combinatorial studies for *cis-element* detection have been conducted.

GO/KEGG description	Detected Motif	Genes in category	Cat. genes with motif	p-value
ribosome biogenesis		94	68	1.14e-07
polarity specification of adaxial/abaxial axis		6	3	4.92e-02
dolichol biosynthesis		5	3	3.27e-04
response to osmotic stress		12	9	4.94e-02
regulation of transcription		483	282	3.45e-04
response to auxin stimulus		137	80	1.81e-02
response to red or far red light		19	11	1.83e-02
response to abscisic acid stimulus		49	10	3.18e-02
gravitropism		6	3	4.78e-02
sugar mediated signaling		14	9	1.07e-02
cytidine metabolism		9	4	6.69e-04
mitochondrial inner membrane transport		11	10	2.52e-02
giberellic acid mediated signaling		18	6	2.91e-03
carpel development		5	3	1.81e-02
flower development		24	8	2.94e-02
microtubule-based movement		66	27	3.58e-02
phenylpropanoid biosynthesis		7	4	3.94e-03
trehalose biosynthesis		22	8	6.58e-04
cellulose biosynthesis		42	9	3.68e-02
gluconeogenesis		56	50	6.48e-03
giberellin biosynthesis		13	8	2.58e-02

Fig. 2: *Enrichment of PhyloCon detected cis elements in GO and KEGG categories. Second column lists sequence logo representations of a profile discovered that is statistically overrepresented in a functional category/KEGG pathway shown in first column. Next columns show (i) the number of genes within the respective category which contain the motif in their promoters, (ii) the number of genes in the genome annotated for the category and (iii) the Bonferroni-corrected probability that the enrichments is obtained by chance. Example results are shown for GOSlim annotations in the upper part, for KEGG pathways in the part below the dashed line.*

We applied an algorithm, PhyloCon, which combines the detection of phylogenetic conservation within promoter regions with the co-expression of genes within a species. However, for the time being PhyloCon has not been applied on a genome scale and for *cis-element* discovery within higher eukaryotes.

The high-quality *Arabidopsis* genome sequence and the large amount of *Brassica oleracea* genomic survey sequences allows for a genome scale phylogenetic footprinting for conserved elements between the promoters of corresponding *Arabidopsis thaliana* and *Brassica oleracea* genes. Within this study we undertook a comprehensive analysis on thousands of *Brassica-Arabidopsis* orthologous promoter pairs. To generate co-expression information we analyzed a set of 81 microarray experiments totaling 779 chips. The resulting promoter sets were analyzed and a large number of candidate sites have been discovered. These sites are, both, derived from profiles conserved between orthologous promoters and associated with co-expression. Evaluation of our analysis results by means of experimentally validated *cis-regulatory* elements confirms the significance of the analysis results. This study provides a foundation for future *cis-regulatory* module analysis and analysis of regulatory circuits not restricted to *Brassicaceae* and plants in general. It further demonstrates the benefits of fragmented genome sequences to address complex problems in comparative genomics.

Biodiversity and functional genomics of endogenous small RNAs induced upon biotic and abiotic stresses in plants

PLASMAR
Bioinformatics

Yu Wang, Octave Noubibou, Klaus F. X. Mayer
MIPS, Institut für Bioinformatik
GSF – Forschungszentrum für Umwelt und Gesundheit 85764 Neuherberg

Introduction

One of the most astonishing scientific discoveries in the last decade is small RNA related gene regulation at transcription and post-transcription level[1, 2]. Small RNAs fall into three categories, microRNAs (miRNA), small interfering RNAs (siRNAs) and piwi-associated RNAs (piRNA). miRNAs and siRNAs are distinguished and classified based on their biogenesis rather than by their function. miRNAs derive from long, single-stranded miRNA precursor (pre-miRNAs) that fold imperfectly into double-stranded hairpin structures. The pre-miRNAs are processed sequentially by RNase III proteins of the Drosha/Dicer family1–4. In contrast, siRNAs are processed by members of the Dicer family from long, perfectly base-paired dsRNAs that derive from transcription of inverted repeat sequences, convergent transcription of sense-antisense gene pairs or synthesis by RNA-dependent RNA polymerases (RDRs)1–4. piRNAs have recently been discovered in mammalian testis and their functions are currently unknown. Plant miRNAs usually contain near perfect complementarities to their targets and introduce target cleavage. Animal miRNAs match to the 3'UTR of their target mRNAs. Most often a seed region, that is position 2-8 within the (animal) miRNA introduces translational repression. Functions of siRNAs during both transcriptional and post-transcriptional gene repression have been firmly established.

We use integrated methods to study plant small RNAs. Our strategy includes combination of both experimental and bioinformatics approaches to address the functional genomics of endogenous small interfering RNAs and microRNAs in plants. We are particularly interested in how these non-coding RNA molecules interact with genes in response to abiotic and biotic stress conditions or viral infections. Experiments were conducted to set up normal growth, stress-specific and viral infected small RNA libraries in *Arabidopsis*. A data mining pipeline has been established to automatically process sequence libraries containing a large set of plant small RNA. Complementary we have developed a machine learning ensemble algorithm to detect *cis* regulatory elements in promoters of plant microRNAs. Our results demonstrate that sequence conservation within promoters of miRNA families as well as between miRNA and its (potential) evolutionary progenitor gene can be exploited for understanding the regulation of microRNA genes.

Data Mining Pipeline for Automated Processing of Plant Small RNA Libraries

Recent advances of deep sequencing methods makes profiling of small RNA transcriptomes and collections a feasible experiment routine for investigating function of endogenous small interfering RNAs and microRNAs in plants[3]. Necessarily, the data avalanches trigged by these experiments [4-8] require fully automated processing of small RNA data. This includes, 1) implementation of a suitable database containing small RNAs; 2) collection and integration of all publically available (plant) small RNA data; 3) development and implementation of a data mining pipeline to automatically query the database and analyze small RNA libraries sequenced by our project partners. A detailed data processing schema is given in Figure 1. The pipeline input are raw sequences data which are cloned and sequenced by PLASMAR project partners. For instance, our French collaborator Gabriel Morilla has built two libraries, one made from RNA extracted from seedling of *Arabidopsis* treated with the biologically active peptide FLG22 and another one from seedling treated with a non-biologically active peptide. Small RNAs were purified and isolated from a PAGE gel. Two RNA adaptors have been ligated to the 5' and 3' end of the small RNAs. RT-PCR was used to obtain corresponding cDNAs. The computational task includes automatically removing the adapter sequences from 5' and 3' of small RNAs. The next step is to map all small RNA sequences to the *Arabidopsis* genome. Only those small RNAs which can be perfectly mapped to the genome were kept for further analysis. We are expecting to find new siRNAs and new miRNAs in the project. In addition the experimental set-up will give insights into the repertoire of small RNAs are active and activated in stressed plants. Thus, the next step in the data mining pipeline is to query our small RNA database to distinguish the known small RNAs and novel small RNAs. For novel small RNAs, we map them to the intergenic region of the *Arabidopsis* genome and look for inverted duplicates of

1. Couzin, J., **Nobel Prize in Physiology or Medicine. Method to silence genes earns loud praise.** Science, 2006. 314(5796): p. 34. 2. Bots, M., S. Maughan, and J. Nieuwland, **RNAi Nobel ignores vital groundwork on plants.** Nature, 2006. 443(7114): p. 906. 3. Miranda, K.C., *et al.*, **A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes.** Cell, 2006. 126(6): p. 1203-17. 4. Berezikov, E., *et al.*, **Diversity of microRNAs in human and chimpanzee brain.** Nat Genet, 2006. 5. Lu, C., *et al.*, **Elucidation of the small RNA component of the transcriptome.** Science, 2005. 309(5740): p. 1567-9. 6. Henderson, I.R., *et al.*, **Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning.** Nat Genet, 2006. 7. Lu, C., *et al.*, **MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant.** Genome Res, 2006. 16(10): p. 1276-88. 8. Qi, Y., *et al.*, **Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation.** Nature, 2006. 443(7114): p. 1008-12. 9. Wang, Y., T. Hindemitt, and K.F. Mayer, **Significant sequence similarities in promoters and precursors of *Arabidopsis thaliana* non-conserved microRNAs.** Bioinformatics, 2006. 22(21): p. 2585-9.

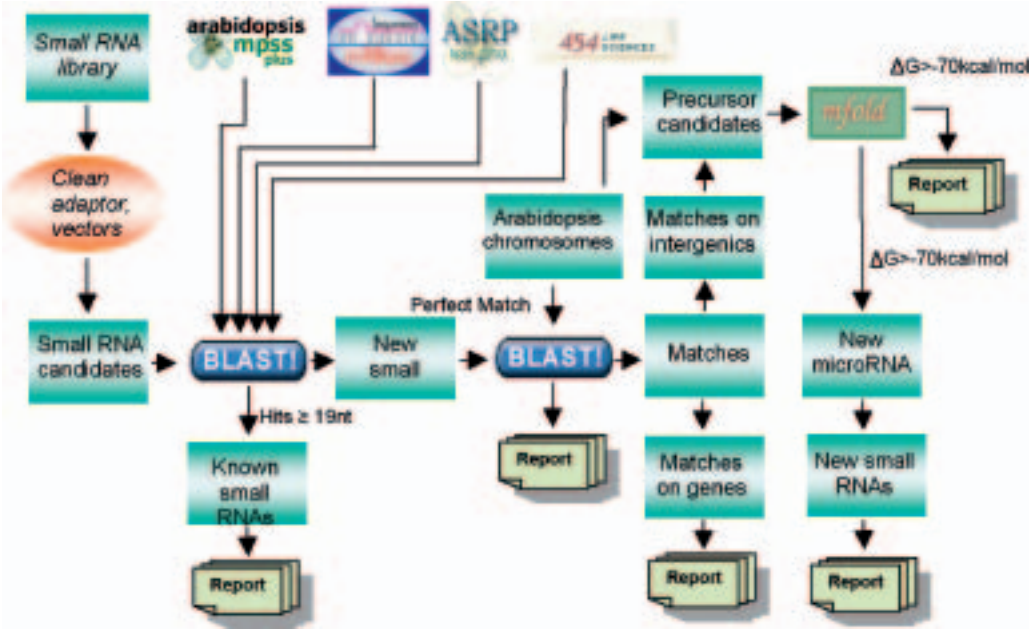


Fig. 1: *Data mining pipeline to process small RNA sequences.*

these sequences in the surrounding area. We use the sequence which contains both the small RNA sequence and its inverted repeat and apply MFOLD to analyse the secondary structures of these sequences. The ones that can form proper hairpin structures are identified as novel miRNAs. To date, we have processed eight sequencing libraries from the French partnering lab (O. Voinnet's research group) and two sequencing library from the Spanish collaborating labs.

As a crucial step in this project, we have collected half a million recently published small RNA sequences of *Arabidopsis thaliana*. These small RNAs are systematically reanalyzed and deposited into our PLASMAR database. A GBrowser has been set up to make it possible to access the database in a user-friendly manner (Figure 2). Small RNAs from various mutant backgrounds, namely dcl1-7, dcl234, rdr6, rdr2, and wild type Col-0 can be compared against each other. This information is very important for plant small RNA researchers since different small RNAs are usually involved in different cellular pathways. Therefore, small RNAs which are derived from different mutant backgrounds can form distinct sub-populations. This info is communicated in a user-friendly fashion using the Gbrowse genome browser.

Detection of *cis*- Regulatory Elements for Plant MicroRNAs

MicroRNAs (miRNAs) are about 21 nucleotide long, endogenous noncoding RNAs that regulate a large number of genes at the post-transcriptional level. In plants, miRNAs usually have nearly perfect matches to coding regions of their target messenger RNAs. Binding of miRNAs to these regions lead to transcript cleavage and degradation through the RISC complex.

Extensive conservation of miRNAs between different species and over long evolutionary distances such as between grasses and flowering plants has been reported. Plant miRNA genes are transcribed by RNA polymerase II. The primary miRNA transcripts (pri-miRNAs) are cleaved by RNase III like enzymes to generate miRNA precursors (pre-miRNAs) and subsequently to mature miRNAs.

The role of microRNAs as a part of gene regulatory machinery has been demonstrated by recent intensive research. However, the transcriptional regulation of microRNA genes them-

Fig. 2: *A GBrowser image depicts the genomic location of Arabidopsis thaliana microRNA ath-miR156e. This miRNA has been sequenced in wild type plants by MPSS technology and also in the mutant background of rdr2 and rdr6.*

selves is less well investigated. We have detected similarities within the precursor as well as the conserved regions within the promoter sequences of microRNAs (Figure 3,4)[9]. This provides additional evidence to support the inverted duplication model for plant microRNA evolution. Our finding suggests that the initial inverted duplication event, which potentially created miRNA from its targets (inverted duplication theory), has included the promoter region of target genes. For expanded miRNA families regulatory diversification through modulation of *cis* regulatory elements might be a mechanism leading to fixation of sequence redundant miRNA genes. In analogy to duplicated protein-coding genes promoter elements seem to be shared within paralogous MIR genes. This permits to sketch a powerful in silico approach which might be considered for the experimental and in silico analysis of microRNA promoters

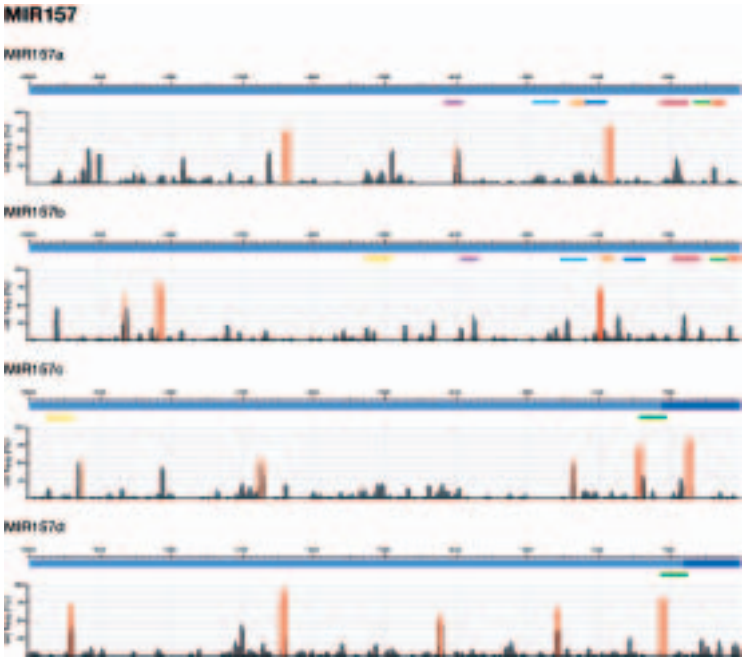


Fig. 3: *Sequence similarities within promoter regions of MIR157. The light blue horizontal bars represent upstream sequences. Colored bars represent conserved regions detected by DiAlign. The peaks represent the number of hits per nucleotide normalised by the number of total hits per sequence, which were calculated 50 times by MotifSampler. A peak is plotted in red when the height of the peak is larger than 50.*

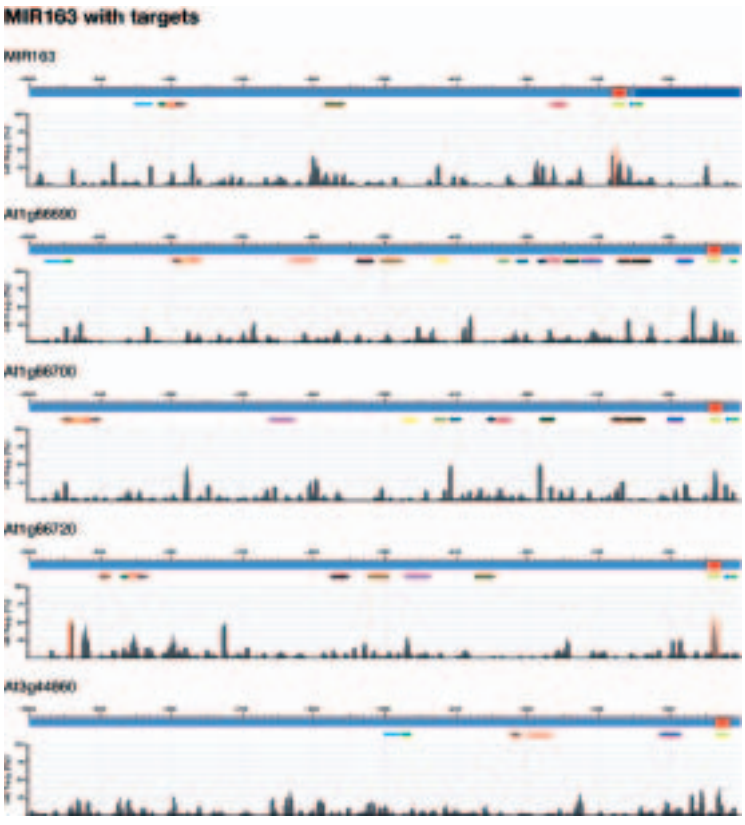


Fig. 4: *Sequence similarities within promoter regions of MIR163 and its target genes. The light blue horizontal bars represent upstream sequences. The dark blue bars at the right end of light blue bars indicates the 5′ UTR of target genes or 5′ ends of pri-miR163 which is not part of the microRNA stem loop. The red horizontal bars represent the identified TATA box. Colored bars represent conserved regions detected. The peaks represent the number of hits per nucleotide normalised by the number of total hits per sequence, which were calculated 50 times by MotifSampler. A peak is plotted in red when the height of the peak is higher than 50 percent.*

GABI-BRAIN: Development and implementation of innovative statistical concepts and computational tools for integrating genomics research and applied plant breeding programs

BRAIN
Bioinformatics

Martin Heckenberger 1; Andreas Büchse 1; Matthias Frisch 1; Hans Peter Maurer 1; Jens Möhring 1; Jasmina Muminović 1; Hans-Peter Piepho 1; Jochen C. Reif 1; Benjamin Stich 1; H. Friedrich Utz 1; Albrecht E. Melchinger 1; Frank Lichert 14; Josef Breun 11; Felix Dreyer 6; Erhard Ebmeyer 9; Erich Knopf 10; Jens Lübeck 13; Axel Schechert 3; Dieter Stelling 2; Stefan Streng 12; Eckhard Tacke 8; Arndt Zacharias 5; Heinrich Wortmann 4

1 Universität Hohenheim
2 Deutsche Saatveredelung Lippstadt-Bremen GmbH
3 Fr. Strube Saatzucht KG
4 Hybro Saatzucht GmbH & Co. KG
5 KWS Saat AG
6 Norddeutsche Pflanzenzucht Hans-Georg-Lembke KG
7 A. Dieckmann-Heimburg Saatzucht Sülbeck
8 Böhm Nordkartoffel Agrarproduktion oHG
9 Lochow-Petkus GmbH
10 Pflanzenzucht Dr. h.c. Carsten – Inh. Erhardt Eger KG
11 Saatzucht Josef Breun GdbR
12 Saatzucht Streng GmbH & Co. KG
13 Saka-Ragis Pflanzenzucht GbR
14 Gemeinschaft zur Förderung der privaten deutschen Pflanzenzüchtung e.V.

Introduction

Main objective of GABI is to conduct comprehensive research on structural and functional genomics, and to exploit this information in plant breeding and related areas. Numerous GABI-projects have contributed substantially to enormous improvements of available genomics tools. These tools are indispensable for molecular breeding of most important crops. However, the application of the results of many research experiments is still limited in classical plant breeding due to the following reasons:

- No system is currently available that efficiently merges genomic and molecular with phenotypic and plant breeding data, thus enabling a directed integration of genomic data into applied plant breeding.
- No or only incomplete concepts are available on how to use the information about known genes in applied plant breeding. For example, it is still completely unknown how to combine 10 (or 100) known genes in a single genotype as quickly and efficiently as possible.
- Favorable genes are generally identified and mapped in specially developed populations. In the majority of cases, these populations are recombinant inbred lines (RILs), near-isogenic lines (NILs) or doubled-haploid (DH) lines, which were developed from bi-parental crosses. Most suitable parental lines for such crosses are those that differ considerably regarding the trait of interest. However, such genotypes are mostly not present in actual breeding materials/pools. Therefore, the applicability of results from experimental populations is almost negligible in practical plant breeding. Furthermore, only two alleles of a monogenic trait can be studied in a bi-parental cross. In reality, number of alleles in breeding material is much higher.
- Routine evaluation of phenotypic data from field- or greenhouse-experiments generally ignores other available data such as pedigree or molecular marker data.

Successful application of resources and research results available from model species and crops for solving problems in applied plant breeding depends on:

- an efficient storage and handling of the rapidly growing sets of genomic and phenotypic data and their integrated analysis across experiments, years, and generations;
- innovative approaches for detecting genes, QTL, and alleles in elite germplasm, which make use of the extensive data generated in breeding programs;
- novel selection strategies facilitating an efficient use of the rapidly growing number of cloned and annotated genes for variety development in breeding programs;
- new statistical evaluation methods that consider all available information.

Data Management and Quality Assurance

Success of modern plant breeding programs depends on a comprehensive utilization of information about each genotype and its relatives (Figure 1). This information consists in classical breeding programs basically of phenotypic and pedigree data. With the advent of molecular markers, the amount of generated and evaluated data in plant breeding programs has considerably increased. Moreover, the length of a breeding cycle has been remarkably decreased by growing greenhouse- and winter nursery generations, and by applying marker-assisted selection in the early stages of development. An efficient data management is therefore indispensable for success of a commercial breeding program.

Data management and analysis systems, which are currently used in plant breeding, were developed with regard to specific problems. Some software packages offer tools for maintenance and biometrical analysis of molecular marker data, but do not allow any connections with phenotypic or pedigree data. Other software packages meet specific needs of plant breeding, such as maintenance of pedigree information and management of field books, however with no relation to genomic data. Numerous online databases maintain genomic data of a huge number of species. Nevertheless, no databank system is nowadays available, which integrates all functions required by modern plant breeding for an efficient data management.

○ Heckenberger, M., A.E. Melchinger, H.P. Maurer, and M. Frisch. **A comprehensive database management system for integrating phenotypic and genomic data in scientific and applied plant breeding programs.** Euphytica (2006) submitted. ○ Maurer, H.P, A.E. Melchinger, and M. Frisch. 2006. **Population genetical simulation and data analysis with Plabsoft.** Euphytica (2006) submitted. ○ Stich, B., A.E. Melchinger, H.-P. Piepho, M. Heckenberger, H.P. Maurer, and J.C. Reif. **A new test for family-based association mapping with inbred lines from plant breeding programs.** Theor. Appl. Genet. (2006a) 113, 1121-1130. ○ Stich, B., J. Yu., A.E. Melchinger, H.-P. Piepho, H.F. Utz, H.P. Maurer, and E.S. Buckler. **2006 Power to detect higher-order epistatic interactions in a metabolic pathway using a new mapping strategy.** Genetics (2006b) submitted. ○ Piepho, H.P., Büchse, A., and B. Truberg. **On the use of multiple lattice designs and a-designs in plant breeding trials.** Plant Breeding (2006) 125, 523-528.

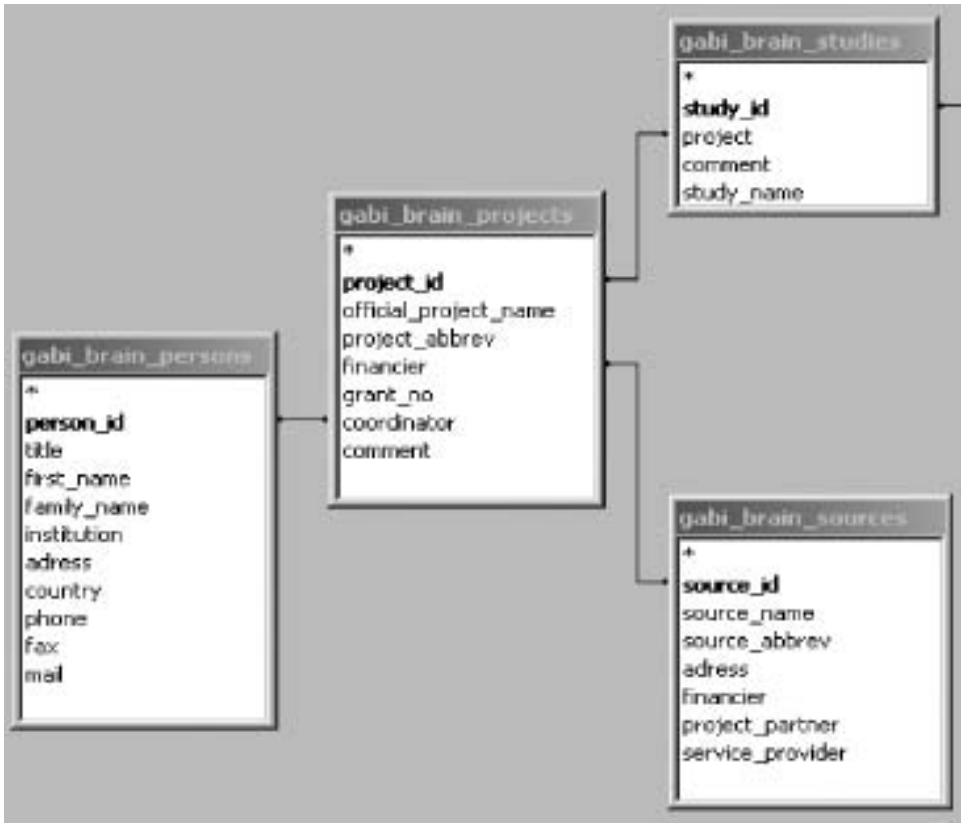


Fig. 1: *Integrated data management system is a basic prerequisite for successful integration of genomics into applied plant breeding*

Therefore, for making selection decisions, breeders generally use just a part of the available data. The full potential of the majority of the genomics approaches can be deployed only if the analysis is based on greater number of genotypes originating from more crosses as well as on actual and ancestral genotypes. These integrated analyses rely on highly complicated data structures and demand ideal data integrity and quality. We therefore developed a database system (Heckenberger *et al.* 2006) that enables directed integration of information from genomics into applied plant breeding.

Simulation Tools for Optimization of Breeding Programs in Genomics Era

The final goal of genome-projects such as GABI is to describe and characterize the structure and function of all (or at least the most essential) genes and alleles in model species and important crops. The application of this knowledge in plant breeding will most probably result in a fundamental change in paradigm, from phenotype-based to genotype-based breeding. In the latter case, selection is completely (or complementarily) based on the presence or absence of desired genes, which can be identified and characterized with the aid of genomic tools.

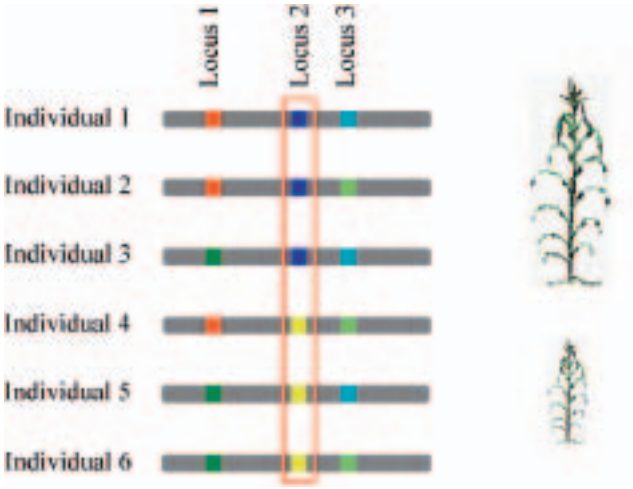


Fig. 2: *Illustration of the principal of association mapping: whereas no correlation exists on Locus 1 between marker genotype (= color of allele) and phenotype (large or small plant), on Locus 2 small plant always appears with the yellow marker allele, and large plant always with the blue marker allele.*

Principles of classical phenotypic selection are not directly applicable for genotype construction, which are commonly referred to as marker-assisted selection (MAS). Furthermore, MAS is difficult to implement and turns extremely complex with an increasing number of target genes. Hence, analytical solutions for the optimization of single breeding steps or a complete breeding program are not available.

Computer simulations became a standard tool for solving complex problems in many areas of research and technology. The University of Hohenheim conducted a pioneer study to develop software Plabsim, specifically designed to simulate plant breeding programs. Plabsim has already been used in many studies on optimal design of marker-assisted backcrossing programs for a single dominant gene, one recessive gene and a simultaneous introgression of two genes. These investigations indicated the complexity of optimization of MAS, even in cases with a single or two target genes. However, the optimization of MAS is meaningful and promising, owing to considerable savings in resources (e.g., more than 50% of marker data points).

With increasing number of target genes and alleles, which are detected in genome projects, the complexity of MAS rises exponentially. We developed software Plabsoft (Maurer *et al.* 2006) for the optimization and comparison of alternative breeding strategies integrating genomics and phenotypic data from practical breeding programs.

Statistical-genetical Analysis Tools:

1.) Pedigree-based QTL mapping with data from commercial breeding programs

Many QTL mapping studies have been conducted in plants, mostly on populations from bi-parental cross of inbred lines, developed particularly for these purposes. This approach possesses the above-mentioned disadvantages that QTLs capture only a part of allelic variation and that those alleles are usually not representative of the elite breeding material.

On the other hand, commercial breeding programs generate thousands of offspring each year, derived from various crossings between relatives. Furthermore, many agronomically important traits are tested in different environments. For the analyses of these materials with genomics tools, the constrictions of classical QTL-mapping studies must be overcome by novel, pedigree-based and/or haplotype-based QTL-mapping approaches. Essential idea of these novel approaches is QTL-mapping with simultaneous utilization of pedigree- and phenotypic data that are routinely collected in applied breeding programs. Thus, the complete QTL-variation present in breeding material can be estimated. Moreover, the breeder is enabled to search for the most favorable allele (allele-mining) both in elite breeding materials and in genetic resources.

This approach is very attractive for plant breeders because it allows the utilization of data readily available from the complex breeding populations and, subsequently, a direct application of the gained information in breeding programs. However, the application of these approaches requires suitable and robust tools for statistical analysis. Therefore, we developed a new test for family-based association mapping with inbred lines from plant breeding programs (Stich *et al.* 2006).

2.) Best linear unbiased prediction

Integrated analysis of genomics and phenotypic data requires a very high quality of field data. Only with the highest possible accuracy of the field data analysis, a relation between field performance and genomics data can be successfully established. We therefore compared alternative experimental designs with respect to their efficiency to assess the differences between genotypes (Piepho *et al.* 2005).

3.) Association mapping in plant breeding context

Association mapping (also designated as linkage disequilibrium (LD) mapping) utilizes linkage disequilibrium (LD) between a gene of interest and tightly linked marker for high-resolution mapping (Figure 2). The method was suggested in human genetics to detect new genes (including QTL) and estimate their effects. In contrast to pedigree- or haplotype-based QTL-mapping approaches, association mapping does not depend on controlled crossings, but requires (i) saturated linkage map and (ii) population-wide LD between markers and genes of interest. Therefore, we adapted statistical tests developed for human genetics for applications in plant breeding populations (Stich *et al.* 2006b).

Go to Content Page

Technology and Ressources

Many recent research projects took benefit in new technologies being available recently. The development of novel technologies like the emerging "omics"-technologies and the generation of far-ranging resources like whole-genome sequences, physical maps or collections of germplasm were a crucial prerequisite for modern biological science. Technologies and Resources developed in GABI consortia provide the basis for future research in GABI projects and beyond.

Go to Content Page

GABI-Kat: Sequencing of T-DNA flankings for »*in silico*« detection of mutant alleles

GABI-KAT
Arabidopsis

Bernd Weisshaar 1,2; Mario Rosso 1,2; Yong Li 1,2; Prisca Viehoever 2; Heinz Saedler 1
1) Max-Planck-Institut für Züchtungsforschung Köln (MPIZ), ADIS (DNA core facility), Carl-von-Linné-Weg 10, D-50829 Köln
2) Universität Bielefeld, Fakultät für Biologie, Lehrstuhl für Genomforschung, Universitätsstraße 25, D-33594 Bielefeld

Introduction

To build up resources for efficient progress in plant genomics, a large set of T-DNA mutagenised lines with insertion sites identified by DNA sequencing is being generated. The output is a database that describes which gene(s) have been disrupted in the lines, so that users can select specific mutants for their reverse genetics needs. Finally, mutant seeds are delivered to scientists around the world for physiological and gene-function relationship studies. Alleles from the GABI-Kat population make a significant contribution to saturate the *Arabidopsis thaliana* genome, which contains about 30,000 protein-coding genes, with NULL mutations. The ultimate goal is to use the new knowledge obtained from the model system *A. thaliana* for the further improvement of today's food crops.

Results

GABI-Kat was initiated at the Max-Planck-Institute for Plant Breeding Research (Cologne, Germany) in the year 1999 and the work started in June 2000. By building on experience from the ZIGIA project which centered on transposons for insertional mutagenesis, selection, growth and leaf harvest for DNA extraction from about 90,000 single lines was completed around February 2006. Within GABI-2, the project was continued to moderately increase the number of lines to be analysed (within GABI-1 it was 70,000), to further improve the quality of the mutant population, and to continue to serve the arabidopsis community with providing confirmed insertion lines. In addition, the transfer of GABI-Kat lines to the Nottingham Arabidopsis Stock Centre (NASC; http://www.arabidopsis.info/) was initiated. This was made possible by an agreement between the funding partners of GABI and the Max-Planck-Society (MPG) who announced in June 2005 that GABI-Kat lines will become freely available to the international research community. As a result, a new "Disclaimer MTA" was released which (i) clarifies the experimental nature of the material, (ii) states that the provision of the material occurs without any warranty, and (iii) excludes any reliability of the MPG for any use of the material. This MTA was made available in June 2005 and can be downloaded from the GABI-Kat website. Obviously, all other agreements which had been executed before remain effective as singed.

The transfer of the GABI-Kat lines to NASC makes sure that the availability of GABI-Kat lines is secured for the future. Lines with confirmed insertion sites are transferred as sets of

View of the GABI-Kat seed store. The colour-code of the paper clips for the seed bags is used to identify rows in the 96-well format (8 x 12) in which the plants are grown.

○ Li Y, Rosso MG, Strizhov N, Viehoever P, Weisshaar B. **GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in *Arabidopsis thaliana*.** Bioinformatics, 2003 19(11):1441-1442 ○ Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B. **An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics.** Plant Mol Biol. 2003 53(1-2):247-259 ○ Strizhov N, Li Y, Rosso MG, Viehoever P, Dekker KA, Weisshaar B. **High-throughput generation of sequence indexes from T-DNA mutagenized *Arabidopsis thaliana* lines.** BioTechniques. 2003 35(6):1164-1168 ○ Li Y, Rosso MG, Ulker B, Weisshaar B. **Analysis of T-DNA insertion site distribution patterns in *Arabidopsis thaliana* reveals special features of genes without insertions.** Genomics. 2006 87(5):645-652 ○ Li Y, Rosso MG, Viehoever P, Weisshaar B. **GABI-Kat SimpleSearch: an *Arabidopsis thaliana* T-DNA mutant database with detailed information for confirmed insertions.** Nucleic Acids Res., 2007 35: D874-D878 ○ The GABI-Kat home page (http://www.GABI-Kat.de/) lists 15 papers with contributions from GABI-Kat co-workers, and 91 additional publications that use GABI-Kat material (as of January 2007).

T3 seed harvested from a segregating family of single T2 plants. This will allow direct user access to potentially homozygous material for a given insertion line, because at least one homozygous plant should be among the 12 to 18 T2 plants which were grown. It should be noted that all T2 plants need to be resistant because they are grown on selective media. This also means that there should be no T2 plants homozygous for the wildtype allele in question unless there is a second locus which provides resistance. The transfer of lines from GABI-Kat to NASC has started in June 2005 and will continue until the end of the GABI-Kat project. At the time of this writing (January 2007) more than 58,000 seed stocks representing about 4,000 original lines have already been donated. As soon as NASC has processed these lines for uptake into their stock catalogue, they become available via the NASC website for ordering (http://www.arabidopsis.info/CollectionInfo?id=69). GABI-Kat no longer distributes these lines, and SimpleSearch presents the NASC order code that links to the NASC seed stock in these cases.

In addition to accessing lines that have been transferred to NASC from the stock centre, users can still order insertion lines directly from GABI-Kat. This access option will obviously concentrate on lines that are not yet confirmed and brought to T3. Usually, the user will get an aliquot of the T2 seeds of the respective line. In some cases, e.g. when all T2 seed of the line in question have been used for growing T2 plants, seed delivery is delayed until T3 seeds are harvested. Furthermore, we have examined our FST data set for CDSi insertion alleles of genes that are only covered by GABI-Kat. We found about 2,100 genes for which CDSi insertions are available from GABI-Kat only. These lines will be gradually entered into the confirmation pipeline and processed when possible without delaying direct user requests. If confirmed and after harvest of T3 seeds, and with an unavoidable delay that is caused by the time required for preparing donations at GABI-Kat, these lines will also be donated to NASC as described above. The plan now is to confirm all valuable lines so that finally all relevant GABI-Kat lines become available from NASC.

The authors wish to thank all members of the GABI-Kat team for their excellent work.

GABI-Kat in Numbers:

A) Timetable (dates):

01.06.2000:	Start of the project
01.01.2001:	All positions filled
01.05.2001:	Start of the production phase of FST sequence generation
09.07.2001:	Test-opening of the GABI-Kat website
15.08.2001:	Approval of the GABI-Kat MTA by PTJ
01.04.2002:	SimpleSearch updated to allow BLAST searches on FST sequences
01.06.2002:	Opening of the public GABI-Kat website (1st public MTA)
01.08.2002:	SimpleSearch updated to include graphic display of insertion sites
01.09.2002:	Start of the Génoplante-FST/GABI-Kat exchange program
15.02.2003:	New "gene hit" definition (major update of SimpleSearch)
15.12.2003:	More than 10425 individual gene internal (CDSi) hits
15.09.2004:	Re-annotation of all FSTs according to TIGR v5 genome release
13.06.2005:	Seed donation to NASC initiated
24.06.2005:	Release of the second public MTA (Disclaimer)
01.02.2006:	Last seeds of newly selected T1 plants harvested
25.06.2006:	SimpleSearch updated to include allele-specific confirmation sequences

B) Numbers (as of January 2007, release 21):

Resistant lines germinated and selected:	~92,600
Lines of which seeds are harvested:	~89,900
Lines with "genome hit" detected:	64,860
Number of genes with at least one FST hit (non-redundant):	17,089
Percentage of gene coverage:	64.6 %
Predicted insertion between ATG and STOP (non-redundant):	12,370
Percentage of gene coverage (CDSi-hit):	47.2 %
Lines requested in total:	6,025
Lines delivered to users:	4,255
Lines donated to NASC (as sets of several individual T3 seed bags)	4,043

GABI-Genoplante Natural Diversity: Evaluation of natural diversity in Arabidopsis accessions for traits of agronomic or basic importance

NATURAL-DIVERSITY
Arabidopsis

Mark Stitt 1; Thomas Altmann 2; Bernd Müller-Röber 2; Jürgen Kroymann 3; Stefan Clemens 4 and Mylène Durand-Tardif5
1 Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam-Golm
2 University of Potsdam, 14476 Potsdam-Golm
3 Max Planck Instiute of Chemical Ecology, Jena
4 Institute of Plant Biochemistry, Halle and 5INRA, Versailles

Introduction

Natural diversity in *Arabidopsis* represents an important but hitherto largely unexploited resource for basic and applied research. *Arabidopsis* probably originated in central Asia, but has spread to almost all continents. It is likely that selection acting in different environments on the *Arabidopsis* genome has led to diversity in genes affecting a large number of traits that are important for adaptation to growth and survival. If natural populations are adapted to their environment, they have an optimum growth in a range of physical, chemical and biotic surroundings. Moreover natural populations can respond differently to the modification of the environment, calling on their plasticity. The faculty to respond to the variation of the environment is conditioning the survival of populations in extreme episodes as well as their evolution.

Pioneered by work of Koornneef and coworkers, important functional diversity has been uncovered for a wide range of traits including resistance to biotic stresses (resistance to insects, fungi, bacteria, and viruses), tolerance to abiotic stress (such as high/low temperature, freezing, drought, metals, ozone), developmental characteristics (flowering time, plant size, seed size, venation pattern, trichome number), physiological aspects (such as seed dormancy, phosphate uptake, water-use efficiency), or biochemical traits (glucosinolate, seed oligosaccharide, epicuticular wax composition, several enzymatic activities) (reviewed in Alonso-Blanco and Koornneef, 2000).

Description of the Joint Project

The German and the French national plant genomics networks, GABI and Génoplante, have merge in a program that aims to inventory the natural phenotypic variation in collections of *Arabidopsis thaliana*. The goal was to analyze a broad range of traits across a large set of accessions. Before we started, such analyses had been generally limited to the examination of relatively few accessions/crosses/recombinant inbred lines. Against the background of the *Arabidopsis* genome programs, a systematic analysis of diversity in *Arabidopsis* will provide a model system for experimental analysis of genome evolution in higher plants. The discovery of the genes and of the alleles responsible for population adaptation and plasticity will also permit innovative strategies in plant breeding.

Arabidopsis thaliana was selected, because it is the model species for higher plants. It is to the plant kingdom what the fly D. melanogaster is to the animal one or mouse to the mammals: easy to care, rapid to grow, it gives a large offspring for each cross. Further, *A. thaliana* can be self-fertilized which is convenient for geneticists. Its genome is entirely sequenced which is of major importance for geneticists and physiologists. Any application for crops breeding will have to be extrapolated and tested.

A first crucial point was to agree for the biological material to be used. Both Germany and France have gathered natural populations from different geographic origin, the so called 'ecotypes'. To perform our studies, we need to work with characterized and homogenous lines. So, we picked one plant from each natural population and retained a "Single Seed Descent" for two or three generations in order to obtain homozygoty (pure lines) almost all the loci. Thus, we are not working with populations but with extracted lines that we call "accessions".

Prior to the program, a French group had set up a "core collection" of *A. thaliana* accessions (Fig. 1). The principle of a core collection is to maximize the number of alleles of genes "captured", the diverse forms of genes in a species, with a minimal number of accessions. Studying a given number of accessions of the core collection permits the survey of more genetic variability than the same number of accessions from a random sampling. Moreover, the fifteen laboratories in France have studied the same core collection of genetic material, allowing the accumulation of data on this common material. German groups studied the same core, plus some of their favorite accessions.

The information gained through the cooperative action of several expert laboratories has been collated in a central database Vnat (http://dbsgap.versailles.inra.fr/vnat/). This database was elaborated in France and subsequent work was supported by employment of personal form Vnat in Germany, It

○ Barrière, V., Denue, D, Briand, M., Simon, M., Jouanin, L. and Durand-Tardif, M. (2006) **Genetic variations of cell wall digestibility related traits in floral stems of Arabidopsis thaliana accessions as a basis for the improvement of the feeding value in maize and forage plants.** Theoretical and Applied genetics 113, 163-175. ○ Chevalier F., Pata M., Nacry P., Doumas P. and Rossignol R. (2003) **Effects of phosphate availability on the root system architecture : large-scale analysis of the natural variation between Arabidopsis accessions.** Plant Cell and Environment 26, 1839-1850. ○ Cross, J.M.., von Korff, M., Altmann, T., Bartzenko, L., Sulpice, R., Gibon, Y., Palacios, N. and Stitt, M. (2ß06) **Natural variation in carbon-nitrogen interactions: changes of metabolite levels and enzyme activities across 24 *Arabidopsis thaliana* accessions.** Plant Physiology 142, 1574-1588 ○ Granier C, Chenu K, Aguirrezabal L, Cookson S, Dauzat M, Hamard H, Thioux J, Rolland G, Bouchier-Combaud S, Lebaudy A, Muller B, Simonneau T and Tardieu F (2005) **PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to water deficit in *Arabidopsis thaliana* allows identification an accession with low sensitivity to water deficit.** New Phytologist 169, 623-635 ○ Meyer RC, Törjek O, Becher M, Altmann T (2004) **Heterosis of biomass production in Arabidopsis. Establishment during early development.** Plant Physiol 134:1813-1823 ○ Meyer RC, Steinfath M, Lisec J, Becher M, Witucka-Wall H, Törjek O, Fiehn O, Eckardt Ä, Willmitzer L, Selbig J, Altmann T (2006) **The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*.** Proc Natl Acad Sci USA (provisionally accepted) ○ Schmid KJ, Törjek O, Meyer R, Schmuths H, Hoffmann M, Altmann T (2006) **Evidence for a large-scale population structure of *Arabidopsis thaliana* from genome-wide single nucleotide polymorphism markers.** Theor Appl Genet 112:1104-1114 ○ Törjek O, Witucka-Wall H, Meyer RC, von Korff M, Kusterer B, Rautengarten C, Altmann T (2006) **Segregation distortion in Arabidopsis C24/Col-0 and Col-0/C24 recombinant inbred line populations is due to reduced fertility caused by epistatic interaction of two loci.** Theor Appl Genet 113:1551-1561

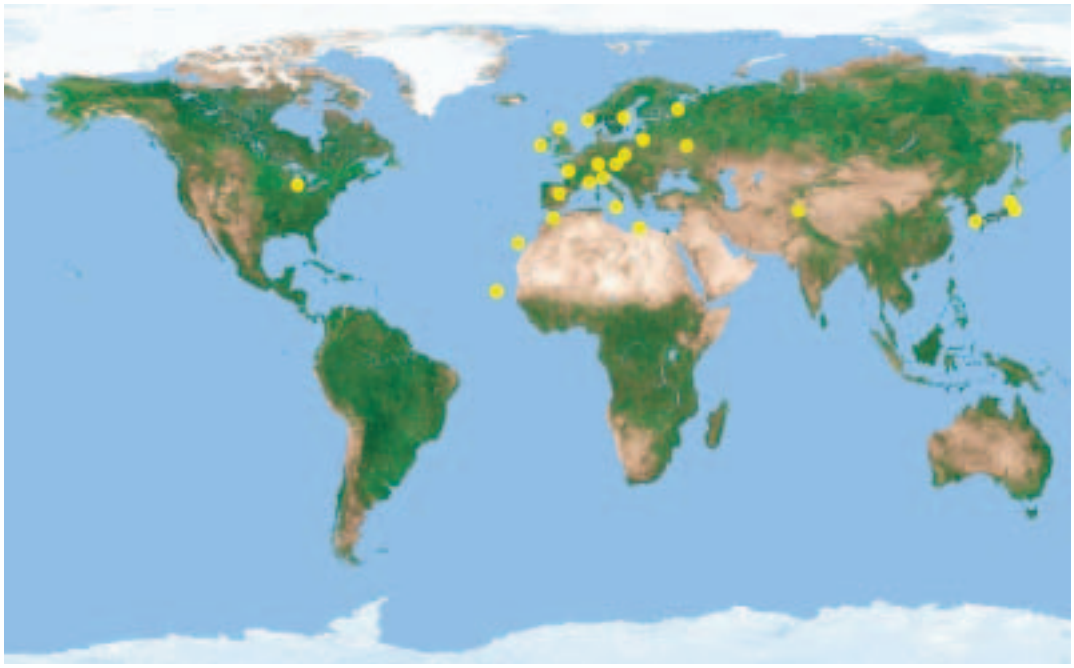


Fig. 1: *Geographic origin of the 24 accessions included in the Versailles core collection of Arabidopsis thaliana, which was used in this joint project.*

includes passport data on original ecotypes, current information about the French collection and results acquired in this program (Fig. 2). We precisely describe the experimental conditions of each study in the database. All the data are quantitative and presented as numeral tables that can be transferred on personal computers to realize statistical analysis. It constitutes a massive sum of data to be processed. The database was open to all partners during the project, and is progressively released to the public.

The program was organized according to the work-packages (Fig. 3). The main objective has been reached for most of the German and French teams: they know growth capacity of the accessions of the core collection, they know the extend of the phenotypic variation of the character they are interested in and have learnt which parameters to measure, and they are starting to understand the response of accessions to deviations from the standard experimental conditions. Those results are paving the way to the research and the isolation of genes implicated in the expression of complex characters, fundamental traits as well as agronomic ones.

Outlook

Natural diversity is a major tool for basic science, because it allows the analysis of complex networks of trait variation, and the central resource for plant breeding. Rapid advances are enabling its utilization are (i) increasing information on genome sequence variation, (ii) qualitative jumps in technologies for genotyping segregating populations and creating true-breeding inbred lines, (iii) existing and emerging platforms for high-throughput phenotyping of molecular, biochemical, physiological, morphological and agronomic traits, and (iv) a massive increase in the power of bioinformatics and statistics. These methodological developments will soon enable major changes in plant breeding.

A rate-limiting step in these advances is the characterization of phenotypic diversity. This joint GABI-Genoplante proj-

Fig. 2: *Home page of the web site VNat, where the results of this joint project are collected and will progressively be released into the public domain.*

ect has provided broad insights into phenotypic diversity for a woide range of traits across a large number of accessions. The next step is to identify and clone the genes involved, in order to allow their function and their role of natural diversity in evolution of that trait to be studies in *Arabidopsis*, and to highlight candidate genes for breeding in crop plants. Here, current projects by the German and French teams are pursuing two different and complementary approaches. The first involves mapping and subsequently fine-mapping of Loci involved in Quantitative Traits (QTLs) in inbred lines. Interesting accessions are crossed and the segregation of the characters studied by each team is analysed in the descendents of each cross. The second is association mapping. This involves analysis of a large number of individuals for (i) their genetic constitution, and (ii) a quantifiable trait of interest. The genetic constitution can be characterized by alleles of candidate genes [`candidate gene-based association studies´] or by anonymous genetic markers [`linkage disequilibrium (LD)-based association studies´]. Significant associations are sought between the trait of interest and the appearance of alleles/haplotypes of the genes, or alleles of markers.

GABI-Genoplante Natural Diversity:
Evaluation of natural diversity in Arabidopsis
accessions for traits of agronomic or basic importance

Fig. 3: *Web page of VNAT summarising the diverse work-packages and characters studied in the program*

GABI-TILL: Establishment of a central platform for testing lead gene function in crops based on TILLING

GABI-TILL
Cereals

Thomas Altmann 1; Sven Gottwald 2; Uwe Hohmann 3; Daniela Holtgräwe 4; Christian Jung 3; John Lunn 5; Nils Stein 2; Mark Stitt 5; Georg Strompen 1; Bernd Weisshaar 4
1 University of Potsdam, 14476 Potsdam-Golm
2 Institute of Plant Genetics and Crop Research, 06466 Gatersleben
3 University of Kiel, 24098 Kiel
4 University of Bielefeld, 33615 Bielefeld
5 Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam-Golm

Plant genome research is currently characterized by rapid advances in the identification of gene functions in plant reference species such as *Arabidopsis thaliana*. The result is a steep increase in the accumulation of lead gene information, which is awaiting utilization in crops to support breeding programs. An important first step towards the use of this massively increasing knowledge in reference species is a test of functional conservation of the homologous genes in the relevant crops. The most immediate approach to determine the role of a gene and its importance is the creation/identification of mutants with decreased (or increased) activity of the gene(s) of interest. While such approaches are straight-forward in model species, thanks to highly efficient resources such as (almost) saturating insertion mutant populations with identified insertion sites, they are not feasible in most crops. The only exceptions currently are maize (due to the presence of well characterized, active endogenous transposons) and rice (due to the availability of highly efficient T-DNA transformation procedures). In order to fill this gap, the GABI-TILL consortium partners have joined forces to establish TILLING platforms for two important crop species (barley and sugar beet) and the reference species Arabidopsis. The term TILLING (Targeting Induced Local Lesions IN Genomes; McCallum *et al.* 2000) refers to a high-throughput technique developed for the detection of mutations in genes. It is based on mismatch-specific endonuclease cleavage of heteroduplex DNA fragments formed upon PCR amplification of target gene sequences of individuals from a mutant population. Using this technology, a very broad spectrum of chemically induced or naturally occurring alleles can be identified in any organism including silent, partial or complete loss-of-function, and gain- or change-of-function mutations. In addition to the functional characterisation of crop species genes, mutant alleles with beneficial effects can be directly integrated into traditional breeding programs without involvement of transgenic approaches.

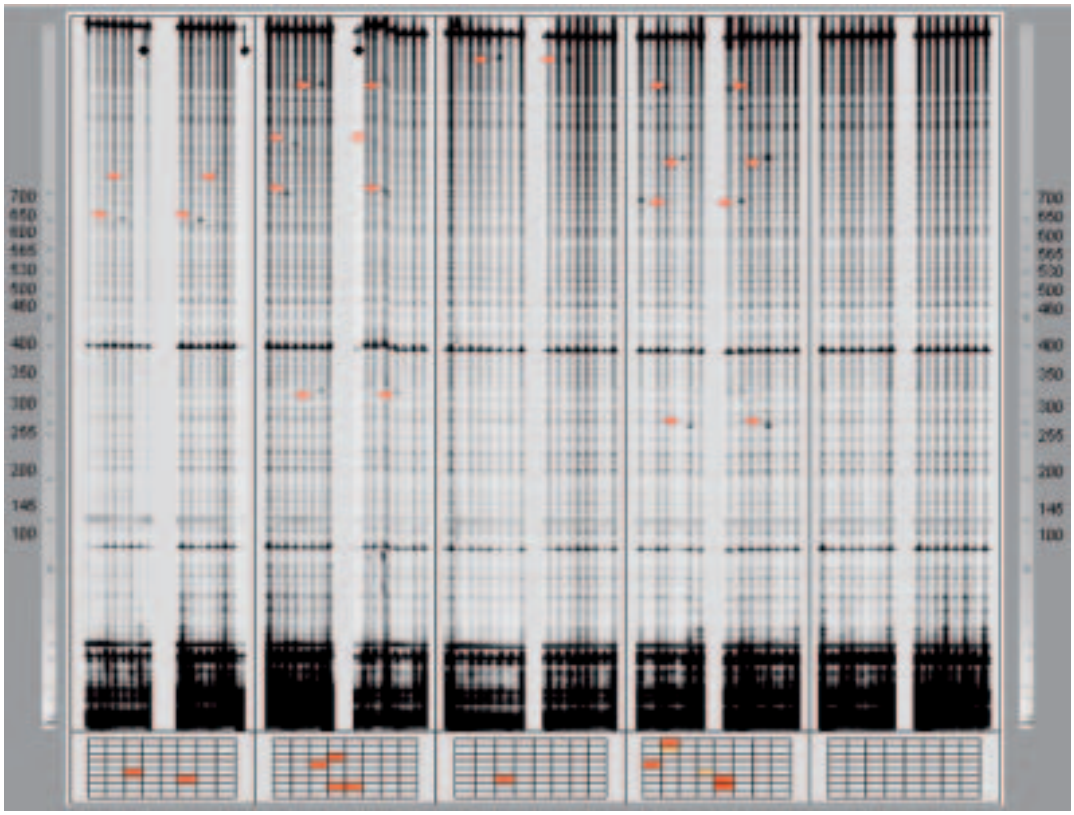


Fig. 1: *Example for a TILLING screen using a 2D-pooling strategy. This gel represents a TILLING screen on 320 individual lines using the 2D pooling strategy. The screen is subdivided into 5 partial screens on 64 individual lines each. A partial screen contains 8 horizontal and 8 vertical pools that give the coordinates for a positive mutant line. In this screen we were able to identify 11 individual mutations, indicated by red boxes in the 8x8 diagram on the bottom of the figure.*

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000) Nat. Biotech. 18: 455-457. Hohmann, U., Jacobs, G., and Jung, C. (2005), **An EMS mutagenesis protocol for sugar beet and isolation of non-bolting mutants.** Plant Breed. 124: 317-321.

In the GABI-TILL project the consortium partners have set up appropriately mutagenised plant populations, stored seed stocks, extracted and pooled DNA, implemented and optimised the TILLING procedures for the different crop species, performed TILLING screens for genes of (potential) commercial value, and set up a central database for the storage and access to the results. The GABI-TILL consortium is part of an international network of TILLING facilities and participates in the European Crop-TIL initiative. According to the advancement in the establishment of the TILLING resources for the different plant species, the platform has been opened to users.

Specific contributions

Supported by international co-operation, the University of Potsdam (partner 1, co-ordinator) has implemented and improved the TILLING technology and applies it now on a routine basis. This partner has transferred this knowledge to the other GABI-TILL partners in workshops and meetings. An Arabidopsis EMS mutant collection (35,000 M1 plants) composed of approx. 6,500 M2 plants has been established. DNA from approx. 6,000 M2 plants has been extracted and a novel 2D pooling strategy was implemented that improved the mutant detection efficiency and reduced the number of necessary reactions (see Fig 1). TILLING screens were performed for several genes of interest in "proof of concept studies" and hitherto 134 mutations were found in 11 genes, including *SPS5B* (encoding the major isoform of sucrose phosphate synthase in leaves) and *APS1* (encoding the catalytic subunit of ADP glucose pyrophosphorylase) genes. The EcoTILLING approach for analysis of natural diversity has been applied to 2 genes across 36 accessions and 25 polymorphisms were detected that were assigned to 6 and 7 different haplotypes, respectively.

To demonstrate the power of TILLING in complementing other mutant populations like GABI-KAT, a further 25 genes, for which no loss of function T-DNA insertion mutations are available, and 10 small genes (< 100 amino acids) have been selected and TILLING screens have been initiated. Within the current GABI-TILL project phase, the mutant collections and DNA pools will be increased to approx. 10,000 mutagenised M2 plants serving as the basis for TILLING screens carried out for interested users also outside the consortium.

The aforementioned *sps5B* and *aps1* mutants are analysed in detail by partner 5 (Max-Planck Institute of Molecular Plant Physiology). The two enzymes SPS and AGPase exert major control of fluxes through the pathways of sucrose and starch synthesis and are regulated allosterically and by post-translational mechanisms. The aim of the mutant analysis is to demonstrate the potential of TILLING to discover change-of-function mutants in addition to loss-of-function mutants. To date, 45 mutations have been identified in the *SPS5B* gene, of which 24 cause substitutions of conserved amino acid residues. Homozygous mutant plants are currently being screened by robotized assays of SPS activity under substrate-limiting and saturating conditions. Thirty-five mutations have been found in the *APS1* gene, with 18 of these leading to changes in conserved residues within the catalytic domain of the enzyme (see Fig.2). So far none of the homozygous mutant plants showed total loss of starch synthesis in leaves, indicating that all mutant forms have retained at least some activity. AGPase activity assays, and immunoblotting of proteins separated by non-reducing polyacrylamide gel electrophoresis, are being used to identify mutants with altered catalytic properties or redox modulation. Plants with novel forms of SPS or AGPase will be analyzed further to investigate the effects of the mutations on photosynthesis, carbon partitioning, growth and biomass.

Fig. 3: *Setup of a TILLING platform in barley. A) Greenhouse Cultivation of M1 plants B) Greenhouse Cultivation of M2 plants C) Field Cultivation of M2 plants D) Field Cultivation of M3 plants*

A TILLING platform for the malting barley cultivar "Barke", an elite central European 2-rowed malting barley with significant established genomics resources (~ 150,000 ESTs, DH mapping population), is being set up by the Institute of Plant Genetics and Crop Research Gatersleben (partner 2). About 88,600 barley seeds were mutagenized with different EMS doses. Treatments with up to 35mM EMS yielded 74% mean M1 germination and 17% completely sterile plants while higher concentrations resulted in only 37% germination and 100% sterility. More than 12,000 M1 plants have been cultivated to maturity at high-density in 299-well potting plates on flooding tables (Fig.3A) and leaf samples for DNA extraction and M3 seeds were harvested from over 9,000 M2 plants. Further 22,800 M1 plants are currently cultivated in the greenhouse. M2 mutants are grown either in the greenhouse (Fig.3B) or under field conditions (Fig.3C, D). Hetero-duplex analysis for mutation detection has been adapted to the barley requirements and pilot mutant screens for four genes (*Hv-eIF4E, MLO, Vrs1, HvCO1*) in about 1,800 plants indicated average mutation frequencies of 1 mutation/ ~0.6Mb. In eleven of twelve sequenced mutant alleles EMS-typical G-C/A-T transitions were identified. A slightly increased mutation frequency was observed in 4,608 mutants arranged in 8-fold 2D-DNA-pools and screened for *Hv-eIF4E* mutations.

A structured sugar beet mutant collection has been generated at the Plant Breeding Institute of the University of Kiel (partner 3) after treatments of seeds with different EMS concentrations. The M2 population encompasses 11,219 M2 plants derived from 1514 M1s. In total, 6.7% of the plants and 36% of the M2 families showed altered phenotypes with respect to growth, leaf and hypocotyl color (see Fig. 4), leaf shape, and bolting time (Hohmann *et al.* 2005). After establishment of sugar beet TILLING parameters normalized DNA isolated from 3000 M2 plants of 760 different M1s was pooled (4x to 8x pools). For mutant screens, sugar beet genes involved in floral transition were identified by BAC-library screening with sugar beet ESTs homologous to Arabidopsis flowering genes (*CO, FLD, LD, FVE, FLK* and *FCA*) and are presently sequenced. The full length *BvCO* gene was cloned in collaboration with the sugar beet research institute in Brooms Barn (England). TILLING screening of 384 DNA pools of 1520 M2 plants with 4 primer combinations covering the entire *BvCO* gene identified 12 polymorphic pools. Sequence analysis of individual plants from six pools revealed 3 transitions and 1 transversion. The data demonstrate that the beet population is useful for the identification of new sequence variants.

In addition to contributions to the set-up of large Arabidopsis mutant populations, the University of Bielefeld (partner 4) provides bioinformatics support to the GABI-TILL project and established the GABI-TILL central data base for integration of the contributions by the other partners and adminis-

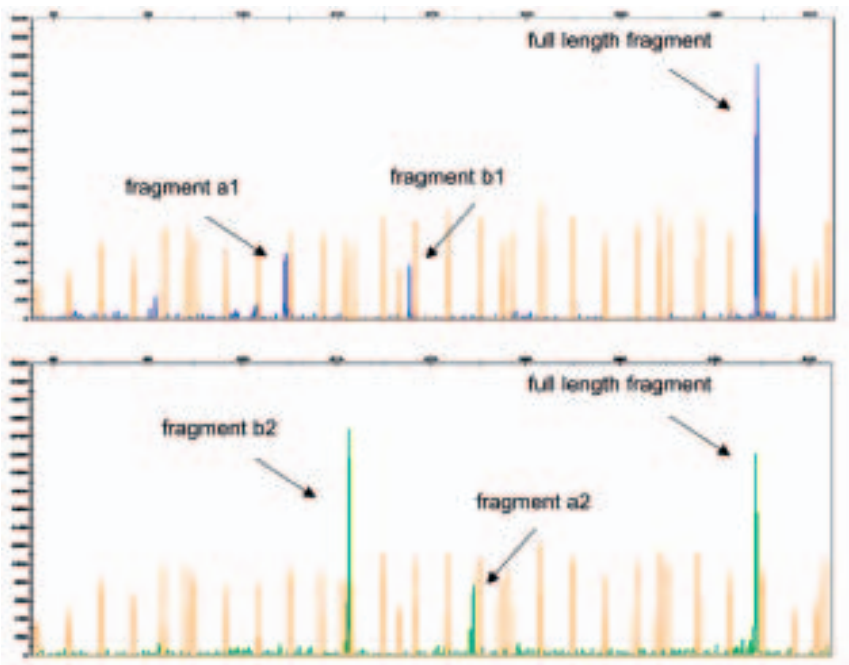


Fig. 5: *Analysis of mismatch cleavage products on ABI Genetic analyzer 3130 platform. Mismatch detection by capillary electrophoresis using an ABI Genetic analyzer 3130 followed by fragment analysis using LIZ standard and GenMapper software. One full length fragment (479 bp) was labeled with two fluorescent dyes (HEX and FAM). Two mismatch positions resulted in the formation of shorter and single labelled fragments after endonuclease treatment.*

trates the GABI-TILL website (http://www.gabi-till.de/). The implemented system for data management and presentation covers the public, the consortium-internal and the workgroup-internal areas. The public GABI-TILL website provides information about the project, the involved partners and the TILLING resources generated (populations, DNA pools). An access-restricted consortium-internal download centre provides access to relevant documents and datasets. To support the greenhouse and wet-lab TILLING workflow, a TILLING-specific LIMS (Laboratory Information Management System) is under development; first modules are already in use. The system provides import functions for existing datasets in standardised but adaptable tab-delimited text formats.

Experimental work was used to test and optimise modules of the TILLING LIMS and besides this serves as an additional DNA source for technology development, e.g. adaptation and optimisation of automated mutation detection using capillary electrophoresis on Applied Biosystems 3130xl series analyser (see Fig. 5). An automatic pipeline for confirmation of EMS-induced mutations on a sequence level including quality trimming of basecalled sequences, sequence annotation, multiple sequence alignment and evaluation of polymorphisms is about to be implemented. Quality control data is tracked through the entire analysis. A first version of a database and web-frontend for a "TILLING service" has been created and is now in the initial internal test phase.

GABI Beet Physical Map: A physical map of the sugar beet genome to integrate genetics and genomics

BEET-PHYSICAL-MAP
Sugar Beet

Britta Schulz 1; Bernd Weisshaar 2; Daniela Holtgräwe 2; Thomas Rosleff Sörensen 2; Georg Koch 3; Heinz Himmelbauer 4; Cornelia Lange 4; Juliane Dohm 4; Thomas Schmidt 5; Katharina Schneider 6
1 KWS SAAT AG, Einbeck
2 Institute for Genome Research, Bielefeld University, Bielefeld
3 Strube-Dieckmann, Nienstädt
4 Max-Planck-Institute for Molecular Genetics, Berlin
5 Institute of Botany, Technical University of Dresden, Dresden
6 GSF-National Research Center for Environment and Health, Neuherberg

Introduction

Specific features of the root enlargement and sugar deposition make sugar beet (*Beta vulgaris*) an interesting model system. The scientific goal of this consortium is the generation of a BAC-based map for sugar beet, strongly linked to the genetic map. A comprehensive physical map of the sugar beet genome is of central relevance for straight-forward positional cloning of agronomically important genes and for the integration of the existing molecular and genetic sugar beet resources. Map construction is based on the hybridisation of specific oligonucleotide probes to high-density BAC filter arrays. Existing EST as well as newly generated BAC end sequence data and increasing information about repeat sequences are resources used for probe design.

Project aims

(1) production of a data set required for physical map construction, based on hybridisation of oligonucleotide probes against BAC macroarrays,
(2) generation of BAC end sequences,
(3) high throughput detection of sequence polymorphisms (SNPs) and development of genetic markers,
(4) segregation analysis and construction of a high resolution genetic map,
(5) identification and characterization of repeat sequences,
(6) establishment and maintenance of the GABI-BPM database as a platform for presentation and exchange of data among partners and with the public.

Results

Production of a data set for BAC map construction

The physical map is based on clones from two different sugar beet BAC libraries. Library ZR/KIEL has been generated previously in the context of GABI-BEET from the double haploid sugar beet line KWS2320 [1]. Library SBI (owned by KWS and constructed by Keygene) was generated from heterozygous plant material. About 27,000 clones from each library are utilised, corresponding to approximately 9-fold coverage of the ~758 Mbp sugar beet genome. The selected set of BAC clones was transferred onto macroarrays, and filters were subsequently used for hybridisation with pooled oligonucleotide probes. The probes used are based on genetically mapped markers, sugar beet EST sequences [2] and sequenced BAC ends from ZR/Kiel clones. The majority of the presently available probes are EST-derived. The probe design is supported by a dedicated bioinformatics pipeline that combines interspecific sequence comparison via spliced alignments and the detection of sugar beet repetitive sequences. The alignment of a non-redundant set of 14,000 public sugar beet cDNA sequences against the genomes of presently sequenced plants (Arabidopsis, poplar, and rice) has resulted in the design of more than 10,000 probes. The probes are hybridised in pools, reducing workload more than 14-fold. Each probe is present in three different pools and by comparing their hybridisation results, clones can be assigned to individual probes. As of Nov. 1st 2006, probe/clone results have been collected for 8,700 probes and further 1,700 probes are being processed. Thus, a data set to build a representative first-generation map will be available soon. At the end of 2006, the second phase of the BAC map project will be initiated, with the aim to connect adjacent contigs, increase long-range connectivity, and cross-link genetic and physical maps.

Generation of BAC end sequences

From the ZR BAC-library generated in the former GABI-BEET project about 29,000 clones were selected for end sequencing providing approximately a 4.2-fold coverage – in terms of clones- of the sugar beet genome. Up to now more than 42,000 high quality BAC end sequences with an average length of 650 bp were generated on an Applied Biosystems Abi Prism 3730 DNA analyser and processed by a bioinformatic pipeline. The developed and implemented data analysis includes the detection of novel, probably species-specific repeats as well as sequence annotation, which in turn is useful for later synteny studies.

○ [1] Hohmann U, Jacobs G, Telgmann A, Gaafer RM, Alam, Jung, C (2003). **A bacterial artificial chromosome (BAC) library of sugar beet and a physical map of the region encompassing the boltinhg gene B.** Mol Genet. Genomics; 269: 126-136. ○ [2] Herwig R, Schulz B, Weisshaar B, Hennig S, Steinfath M, Drungowski M, Stahl D, Wruck W, Menze A, O'Brien J, Lehrach H, Radelof U (2002). **Construction of a 'unigene' cDNA clone set by oligonucleotide fingerprinting allows access to 25 000 potential sugar beet genes.** Plant J.; 32: 845-857.

○ Publications: Menzel G, Dechyeva D, Keller H, Lange C, Himmelbauer H, Schmidt T. ***Mariner*-based identification of miniature inverted-repeat transposable elements in *Beta vulgaris* L. Chromosome Res.** (in press)

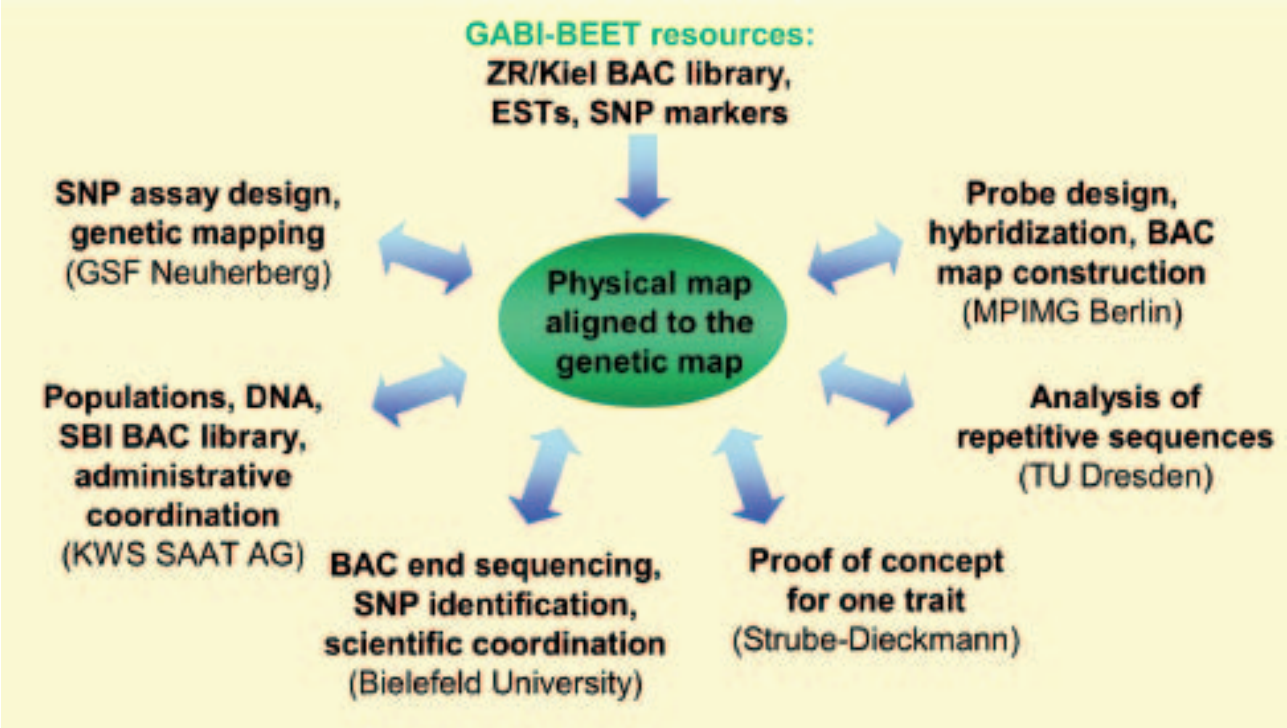


Fig. 1: *Graph summarizing the structure of the joint project GABI Beet Physical Map.*

Automated SNP detection and marker development

As a valuable tool for marker-assisted breeding and other genetic applications a dense set of DNA markers is under development. For the PCR-based generation of SNP-markers, DNA polymorphisms (SNPs/INDELs) between the parental lines of the mapping population are identified via direct sequencing of amplicons. Filtering out repetitive and duplicated sequences within the BAC end sequences followed by automatic primer design, multiple sequence alignment and evaluation of the alignment leads to the detection of a high number of polymorphic sites in the sugar beet genome. On the basis of processed BAC end sequences as targets for SNP marker development, about 4,000 amplimers were designed and are currently used for automated SNP detection. Setting up the whole approach for high throughput (96 well format) led to 1,434 high quality alignments. While more than 721 of these alignments lacked sequence variations, the remaining 713 alignments displayed different degrees of polymorphisms and are currently used for marker development. SNP genotyping is performed in 185 individual genotypes of the segregating F2 population provided by KWS. The aim of this work is the establishment of a high density genetic map of sugar beet which serves to anchor the BAC contigs of the physical map to the genetic map. The same target sequences that are applied in oligo-fingerprinting shall be used for genetic mapping. A high throughput genotyping pipeline was established for sugar beet based on the two amplification schemes 'Homogenous MassEXTEND' (hME) and i-plex which vary in their multiplexing efficiencies. Initial PCR amplification, primer

Fig. 2: *Picture of an explanted sugar beet plant.*

extension, and washing steps were carried out in the multiplex format allowing the parallel analysis of up to 28 loci. SNP alleles are detected by MALDI-TOF-mass spectrometry. Both amplification strategies proved successful, about 80% of designed marker assays yielded good results. Segregation analyses were started with 31 SNP markers developed from RFLP sequences and 187 SNP markers developed from EST sequences in a previous GABI project as a core framework. Based on BAC end sequences, assays for 244 new SNP marker have so far been developed of which 122 produced data for the final map calculation. There is evidence that in many cases paralogous low and medium repetitive sequences in the sugar beet genome may mimic SNPs and mislead the interpretation of the alignments.

Maps were calculated with the Joinmap Software using standard criteria. The most recent map comprises 340 markers including the 122 BAC end-derived SNP markers. The map covers 612.8 cM of the ~758 Mbp sugar beet genome statistically providing a marker density of one marker per 1.8 cM or 2.2 Mbp. BAC end-derived markers were distributed throughout the nine linkage groups of sugar beet (Table 1). Currently 168 more SNP markers are being analyzed. It is the aim to generate up to 2,000 SNP markers to crosslink the physical with the genetic map.

Identification and characterization of repeated sequences

The observed great proportion of repeat sequences within the sugar beet genome demands a special focus on identification and characterisation of species-specific low, medium and highly repeated sequences. Besides bioinformatic approaches to detection of repeat sequences within the generated genomic sequence data, a 500 bp shotgun library (SGZR00841) has been tested by hybridisation of high-density filters with all known repeat families comprising 84% of total repetitive DNA of the *B. vulgaris* genome. Satellite DNA included 79% from all detected repeats. A broad study of reverse transcriptase genes from Ty1-copia-retrotransposons has been performed. Sequence analysis revealed 22-99 % similarity between 105 clones which have been assigned into four distinct subfamilies. The reverse transcriptase gene fragments were further characterised by Southern hybridization and FISH demonstrating a dispersed genome organization.

Linkage group	SNP markers	Map length (cM)	BAC end markers
I	27	65.8	12
II	35	65.1	10
III	30	74	10
IV	28	73.7	9
V	56	78.5	24
VI	53	66.8	16
VII	47	69	17
VIII	23	62.5	7
IX	41	57.4	17
sum	**340**	**612.8**	**122**

Table 1: SNP marker distribution in the current genetic map of sugar beet

Four novel repeat families have been characterized. The satellite family pRv1 is 111 – 174 bp long and dispersed repeats pDvul1, pDvul2 and pDvul3 had sizes of 656 bp, 225 bp and 251 bp, respectively. FISH analysis showed that pRv1 is, along with pBv1, a centromeric satellite, pDvul2 is a chromosome-specific probe while pDvul1 and pDvul3 are dispersed along *B. vulgaris* chromosomes. Newly obtained information on repeated sequences is simultaneously incorporated into the probe-design and SNP-detection pipeline.

GABI-BPM database

Contributions from breeding companies and academic researchers are integrated in the GABI-BPM project. Therefore, the exchange of data is of central importance. Project partners may retrieve data via a secure website. The sequence data provided are used for development of SNP-based genetic markers and genotyping, designing 35mer oligonucleotides for oligo-fingerprinting of BACs, and the detection and analyses of (new) repeat sequences. Different datasets useful for various applications are continuously generated by an bioinformatic data processing pipeline, including vector clipping, quality trimming, repeat filtering and sequence annotation. In addition, quality control data are held in the database.

AB-QTL analysis in wheat: Detection of favorable genes for quality and yield traits from wheat wild species and their introgression into the elite germplasm of cultivated wheat by means of DNA markers

AB-QTL
Cereals

Antje Kunert; Anne Schubert; Ali Naz 1; Jens Léon; Klaus Pillen 1
Dept. of Crop Science & Resource Conservation, University of Bonn, Bonn
1 Current address: Max-Planck-Institute for Plant Breeding Research, Köln

Co-operating breeding companies
1 Deutsche Saatveredelung Lippstadt-Bremen GmbH, Lippstadt
2 Fr. Strube Saatzucht KG, Söllingen
3 Limagrain-Nickerson GmbH, Edemissen
4 Lochow-Petkus GmbH, Bergen
5 Saatzucht Josef Breun GdbR, Herzogenaurach
6 Saatzucht Schweiger OHG, Moosburg
7 W. von Borries Eckendorf GmbH & Co., Leopoldshöhe

Introduction

The AB-QTL (Advanced Backcross Quantitative Trait Locus) strategy was originally developed and applied by Tanksley & Nelson (1996) in order to improve modern elite tomatoes by incorporation of exotic tomato germplasm through natural backcross breeding. After the successful application of the AB-QTL strategy in small grain species like rice and barley (Xiao *et al.* 1996, Pillen *et al.*, 2003), the present AB-QTL project uses hexaploid wheat (*Triticum aestivum*) as the recipient of exotic donor germplasm. For this, synthetic wheat accessions serve as the donor of wild species alleles for the wheat A, B and D genomes. The synthetic wheats have been generated by hybridizations of *T. turgidum* ssp. *dicoccoides* and *T. tauschii*.

Aims

The particular aims of the wheat AB-QTL project are: (1) to localize QTLs in backcrosses between elite wheat varieties and exotic wheat accessions, (2) to detect favorable, trait improving exotic alleles at these QTLs, (3) to increase the genetic diversity present in our modern wheat varieties through the introgression of favorable exotic QTL alleles.

Material and Methods

The AB-QTL study is conducted with four advanced backcross populations (Fig 1).Two spring wheat populations have been backcrossed and selfed until the BC2F4 generation and two winter wheat populations were likewise generated up to the BC2F3 level. The wheat project started in August 2002. It was scheduled for a period of four years and finished in July 2006. During the first project phase (August 2002 through July 2003), the AB lines were propagated and, simultaneously, genotyped with SSR (simple sequence repeat) markers. During the second project phase (August 2003 through July 2006), the AB lines were phenotyped and finally used for QTL detection. During the seasons 2004 and 2005, the AB lines were field evaluated for agronomic and quality related traits as well as for pathogen resistances at eight locations in Germany. These locations include our own experimental station Dikopshof as well as the breeding stations of the seven co-operating breeding companies which are listed above. The evaluation of the AB lines cover the agronomic traits flowering time, plant height, lodging, maturation, spikes per square meter, grains per spike, thousand grain weight, total yield and harvest index. For quality analysis, the indirect parameters protein content, sedimentation and falling number and direct baking quality parameters including milling and baking tests were assessed at the "Bundesanstalt für Getreideforschung", Detmold. In addition, the AB lines were cultivated in extra plots without any fungicide treatment. Here, the possible field resistances of the AB lines against wheat diseases were evaluated. The disease reactions investigated include powdery mildew, *Septoria* leaf spot, *Drechslera* leaf blotch, leaf rust, yellow stripe rust and *Fusarium* head blight. Until the end of 2006, the phenotype and genotype data were statistically analyzed, resulting in the localization of significant QTLs for each parameter evaluated in the field.

Results and Discussion

Altogether 400 AB lines in spring wheat and 400 AB lines in winter wheat were propagated, genotyped with SSR markers and evaluated in field trials. These lines were developed from the spring wheat populations Triso*Synthetic84L (T84) and Devon*Synthetic84L (D84) as well as from the winter wheat populations Batis*Synthetic22L (B22) and Zentos*Synthetic86L (Z86). The AB lines were genotyped with 149 SSRs originating from public resources. A representation of the resulting graphical genotypes for population B22 is given in Fig. 2. After field evaluation of the above mentioned

○ A. Kunert *et al.* **AB-QTL analysis in winter wheat: I. Detection of favorable exotic alleles for baking quality traits introgressed from synthetic hexaploid wheat (*T. turgidum* ssp. *dicoccoides* _ *T. tauschii*).** (submitted). ○ A. Naz *et al.* **AB-QTL analysis in winter wheat: I. Genetic analysis of leaf rust seedling and adult plant resistance in the wheat advanced backcross population B22.** (in preparation). ○ K. Pillen *et al.* **Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.)** Theor. Appl. Genet. (2003) 107: 340-352. ○ S. D. Tanksley & J. C. Nelson **Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines.** Theor. Appl. Genet. (1996) 92, 191-203. ○ J. Xiao *et al.* **Genes from wild rice improve yield.** Nature (1996) 384, 223-224.

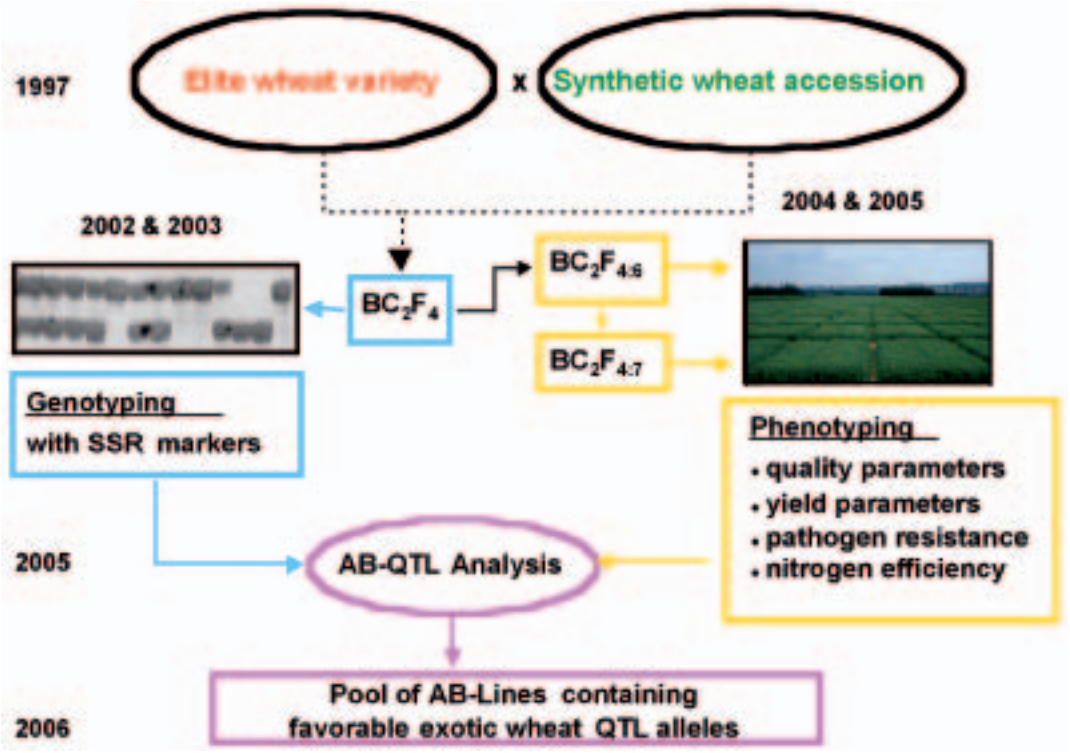


Fig. 1: *The Strategy of the AB-QTL analysis in wheat. Synthetic wheat accessions are used as donors in order to detect QTLs and to improve elite wheat varieties through the introgression of favorable, trait improving exotic QTL alleles.*

traits, a 3–factorial mixed model ANOVA was applied in order to identify QTLs as marker main effects or marker*environment interaction effects which contribute to the genetic regulation of the quantitative trait under investigation.

QTLs were located for all traits in all four populations. Depending on the trait, a varying fraction of the QTLs exhibited favorable effects associated with the exotic alleles. As an example for QTL results, the examination of baking quality in population B22 is given (Kunert *et al.*, submitted). Overall 38 QTLs for baking quality parameters were detected in the winter wheat population B22. The exotic allele improved trait performance at 14 QTLs (36.8 %). The favorable exotic alleles were mainly associated with grain protein content. However, the utmost improvement of trait performance due to the presence of exotic alleles was achieved for the traits falling number and sedimentation volume. At the QTL on chromosome 4B, the exotic allele increased falling number by 19.6 % and at the QTL on chromosome 6D the exotic allele led to an enlargement of the sedimentation volume by 21.7 %.

A second example of QTL results is given for the trait resist-

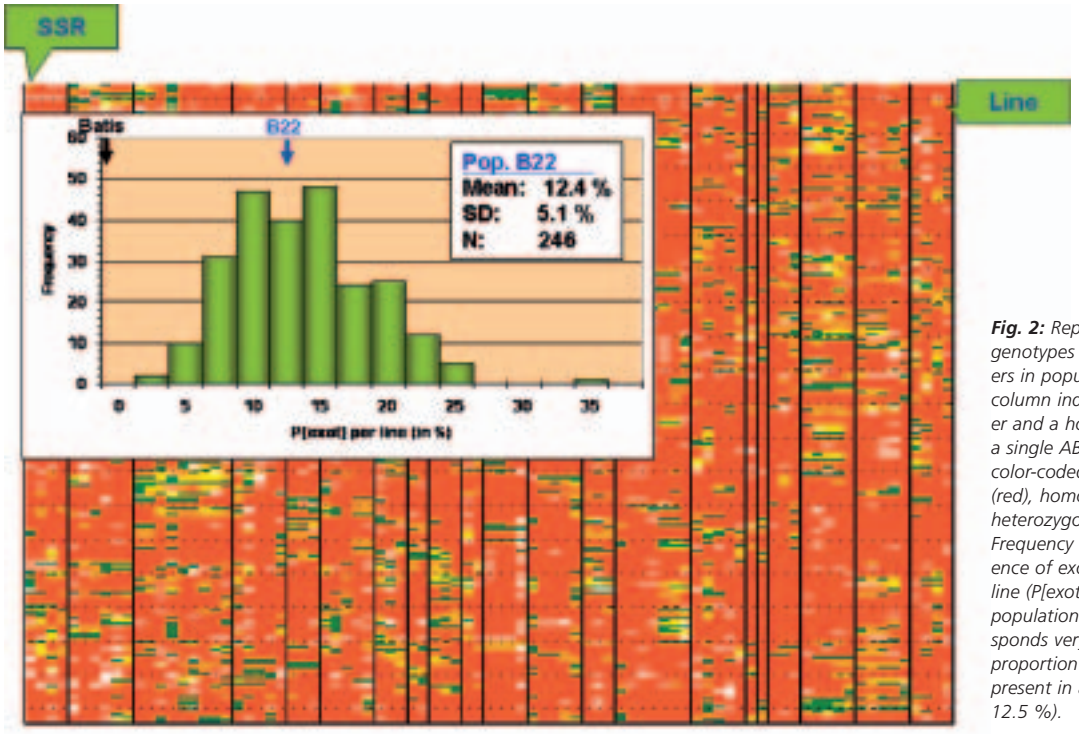


Fig. 2: *Representation of graphical genotypes assessed with SSR markers in population B22. A vertical column indicates a single SSR marker and a horizontal row represents a single AB line. Genotypes are color-coded as homozygous elite (red), homozygous exot (green) or heterozygous (yellow). Inset: Frequency distribution of the presence of exotic germplasm per AB line (P[exot] in population B22. The population mean (12.4 %) corresponds very well to the expected proportion of exotic germplasm present in a BC2F3 generation (i.e. 12.5 %).*

ance against leaf rust, a foliage disease caused by the pathogen *Puccinia tritici* (Naz *et. al.* in prep.). In population B22, 13 QTLs for seedling and adult plant resistance against *P. tritici* were identified. The strongest genetic effect was associated with a marker on chromosome 1D. At this locus, both, the seedling and the adult plant resistance were increased in AB lines containing the exotic allele by 46 and 43%, respectively, (Fig. 3).

These results as well as similar findings for other traits in all four advanced backcross populations indicate that the synthetic wheat donors derived from wild emmer x *T. tauschii* carry favorable QTL alleles for baking quality as well as for agronomic and pathogen resistance traits. Currently, AB lines containing favorable wild species QTL alleles are selected and made available to (i) co-operating wheat breeders as improved breeding material and to (ii) plant molecular biologists as a genetic resource tool for future studies on the molecular regulation of quantitative traits. After purification of the QTL alleles into QTL-NILs, these lines can be used as a starting point for succeeding studies of the differential expression of QTL-related genes which might ultimately lead to the map-based cloning of the underlying genes for quantitative traits.

AB-QTL analysis in wheat: Detection of favorable genes for quality and yield traits from wheat wild species and their introgression into the elite germplasm of cultivated wheat by means of DNA markers

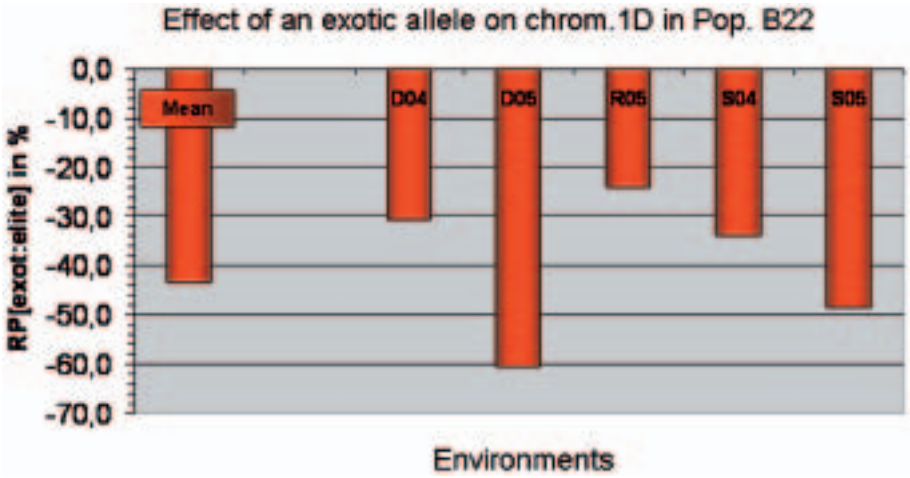


Fig. 3: *The effect of an exotic QTL allele on chromosome 1D on leaf rust symptoms. The relative performance (RP[exot:elite] indicates the strength of reduction of disease symptoms in AB lines of Population B22 due to the presence of the exotic allele on chromosome 1D. The columns represent the resistance effect averaged across all environments (mean) and in individual environments (D04 through S05), respectively.*

Metabolomics-Platform: Transferring functional genomics technology from Arabidopsis to crop plants: metabolic profiling in aid of rapeseed biotechnology

METABOLOMICS
Rapeseed

Edda von Roepenack-Lahaye, Christoph Böttcher, Dierk Scheel and Stephan Clemens

Introduction

Plants have to cope with a variety of constantly occurring and potentially stressful changes in their environment, such as pathogen attack, cold, shade, drought, high levels of light or high salinity. Most of the plant responses to such changes involve the perception of signals and the synthesis of proteins, peptides or low molecular weight compounds (primary and secondary metabolites). A characteristic of plant life is the production of an enormous number of secondary metabolites. These are known to have crucial roles in plant development, as well as in the interaction of plants with their biotic and abiotic environment. They can function as signals in symbiotic interactions, help defending against phytopathogenic bacteria and fungi or deterring herbivores. Also they provide protection against a variety of abiotic stresses such as UV-light, drought or high salt concentrations. In order to get a better understanding of how plants are using their arsenal of natural products in the daily battle of survival in an ever changing environment, methods are required, which allow sensitive detection, quantification and identification of secondary compounds. Metabolomic approaches are one possible answer to these requirements. The information resulting from such non-targeted analyses is of tremendous value for gene function analyses. It will provide molecular insights into physiological states and biological networks, will allow to identify disease markers or to study the impact of environmental perturbations on a biological system.

Because of the chemical diversity of the metabolome, which for any given multi-cellular species comprises a mixture of thousands of compounds differing in size, polarity etc., and varying in abundance by several orders of magnitude, the need for multi-parallel analytical techniques is obvious and well-accepted. LC-MS is a versatile analytical tool. It covers a wide mass range and allows to target many compound classes not detectable by GC-MS. In addition modern LC-MS setups offer superior options to structurally elucidate unknown metabolites using accurate mass determination and fragmentation options.

Results and Discussion

Within the Gabi1 project we established a reliable LC-MS profiling platform using capillary liquid chromatography coupled to electrospray ionization quadrupole time of flight mass spectrometry (von Roepenack-Lahaye *et al.*, 2004). Within the frame of Gabi2 the metabolite profiling was developed further, constantly reevaluated and extended towards the analysis of crop plants or more specific various rapeseed lines. Doubts about the feasibility and reliability of LC-MS-based metabolite profiling have been raised repeatedly, because LC-MS - especially with electrospray ionization - can be subject to matrix effects. The term matrix effects refers to alterations of ionization efficiency of analytes by the presence of co-eluting substances. This could severely compromise quantification. We evaluated matrix effects for our metabolomics platform with several methodological approaches. Our data demonstrated that there are indeed significant absolute matrix effects when comparing highly divergent samples. However, relative matrix effects are negligible - unless extremely divergent matrices are compared - and do not compromise the relative quantification that is aimed for in non-targeted metabolomics studies (Böttcher *et al.*, 2007).

Data-analysis was and is still a major bottleneck of metabolomics. In cooperation with the bioinformatics group at the IPB (AG Massenspektrometrie & Bioinformatik) methods for raw data deconvolution have been developed and improved, respectively. The former use of the data deconvolution software MetaboliteID was shifted to an open source software called XCMS and tailored to the needs of the LC-MS profiling platform (Klie and Neumann, 2006). In order to facilitate the identification and structural elucidation of metabolic compounds the commercial software package "ACD" was acquired. The realms of this software were extended to allow - next to managing LC-MS and MS-MS spectra - sophisticated search algorithms within the steadily increasing database. In addition, an LC-MS-profile database was setup in the institute providing additional data-mining tools.

○ Edda von Roepenack-Lahaye,Thomas Degenkolb, Michael Zerjeski, Mathias Franz, Udo Roth, Ludger Wessjohann, Jürgen Schmidt, Dierk Scheel, Stephan Clemens. **Profiling of Arabidopsis Secondary Metabolites by Capillary Liquid Chromatography Coupled to Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry1.** Plant Physiology, (2004), 134, 548–559 ○ Christoph Böttcher, Edda v. Roepenack-Lahaye, Edith Willscher, Dierk Scheel, Stephan Clemens. **An evaluation of matrix effects in metabolite profiling based on capillary liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry** Analytical Chemistry, (2007) 79 (4), 1507-1513 ○ Sebastian Klie and Steffen Neumann. **Storage and Processing of Mass Spectrometry Data.** Proc. of 17th Int. Conference on Databases and Expert Systems (2006). ○ Christoph Böttcher, Edda von Roepenack-Lahaye, Dierk Scheel and Stephan Clemens **LC-MS profiling of the *tt4/tt5* mutants in *Arabidopsis*: A metabolomics approach** in preparation ○ Christof Stoll, Wilfried Lühs, M. Karim Zarhloul, Wolfgang Friedt . **Genetic modification of saturated fatty acids in oilseed rape (*Brassica napus*).** European Journal of Lipid Science and Technology, (2005) 107, 4, 244-248

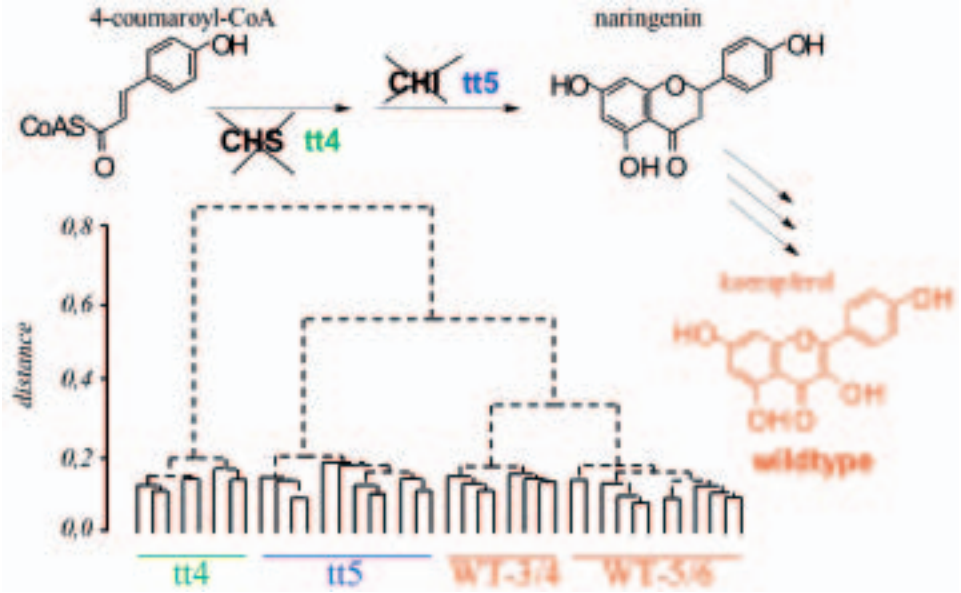


Fig.1: *In the tt4 and tt5 mutants the flavonoid pathway leading from 4-coumaroyl-CoA to kaempferol is blocked at the stage of the chalcone synthase and isomerase. The HCA showed a clear distinction in the degree of relationship between the two mutant and the wildtype seed material as well as the two WT seed batches (due to their different age).*

In the frame of Gabi1 the metabolite profiling was originally based on using *Arabidopsis thaliana* leaves and roots as the source of plant material. Within Gabi2 the focus shifted to crop plants, with a special interest in rapeseed. Extraction and profiling protocols were optimized using as a first step *Arabidopsis* seed material. Seeds seem to contain in general lower levels of those classes of secondary metabolites which could be examined with our profiling platform. This was indicated by a reduced number of mass signals (roughly about 800 signals in contrast to 1500 signals in leaves) that could be reliably detected by LC-MS. In addition, seeds in contrast to roots and leaves contain high levels of choline esters. To structurally characterize this important compound class extensive MS-MS-analyses were performed resulting in the identification of 25 of these nitrogen containing esters.

We performed as initial optimization experiments the analysis of seed material from *Arabidopsis* lines harboring mutations in the well known flavonoid pathway. The transparent testa mutants, *tt4* and *tt5,* are possessing a non functional chalcone synthase and chalcone isomerase, respectively (Fig. 1). The metabolic phenotype of these plants results in the more or less complete lack of kaempferol and its derived compounds, hence the light yellow coloring of the seed testa. The non-targeted profiling approach showed not only the expected absence of specific members of the flavonoid pathway. Several flavonoids, which have not been described before, accumulated to high levels in the mutant lines indicating a redirection of the metabolic fluxes (Böttcher *et al.*, in preparation). The hierarchical cluster analysis (HCA) of this experiment showed not only a clear separation of the different mutant and wildtype materials, but also a clear distinction between seed batches of different age (seed batch WT-3/4 was one year older than WT-5/6, Fig. 1). Additional examinations pointed to a changing metabolome during seed storage, emphasizing the need for truly comparable starting materials for profiling experiments.

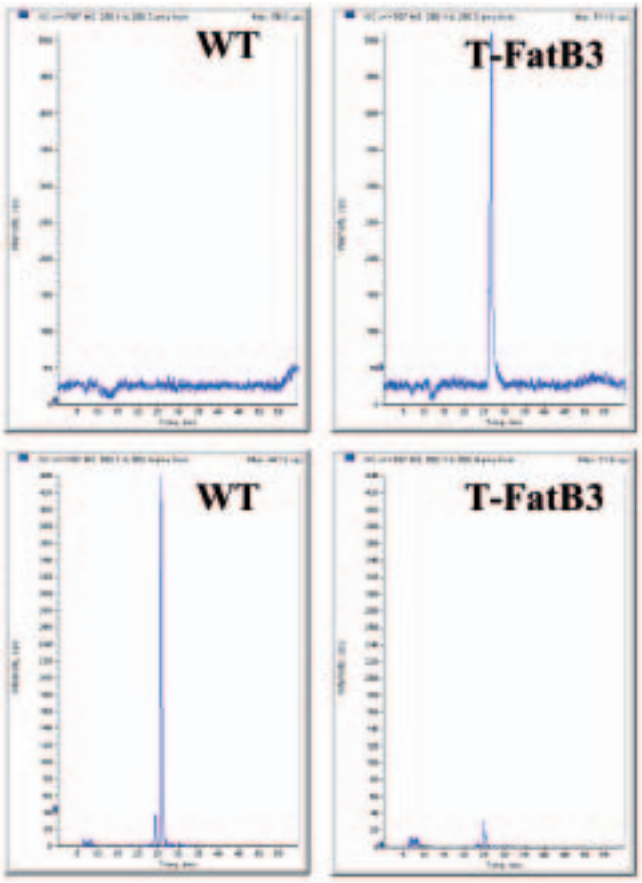


Fig. 2: *Extracted ion chromatograms of the mass signals m/z 256,23/RT27,11 (upper panel, induced in the transformants) and m/z 560,28/RT25,33 (lower panel,reduced in the transformants) in seeds of a transformant line carrying the ClFatB3 thioesterase, and the wildtype rapeseed.*

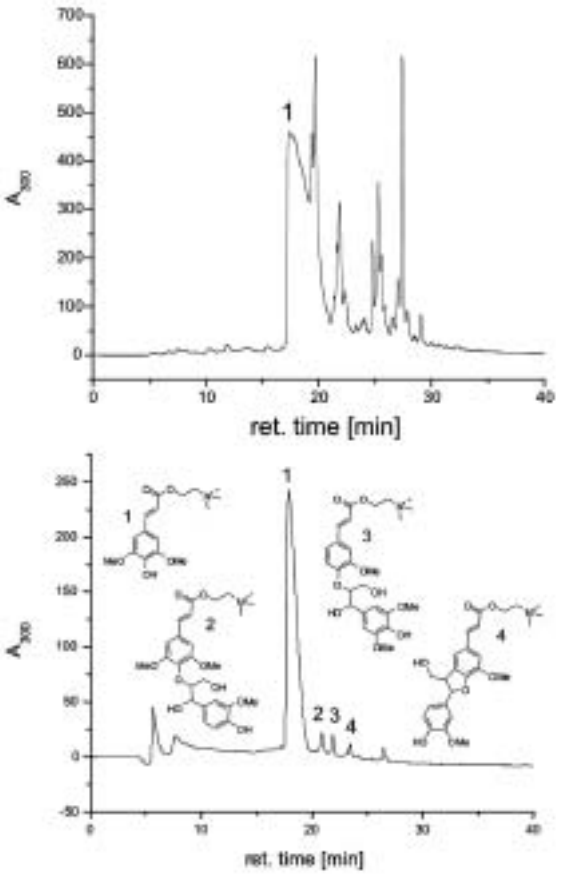


Fig. 3: *UV-traces of a rapeseed extract before (upper panel) and after (lower panel) fractionating for choline esters. The main choline esters in rape seed material are presented 1) Sinapoylcholine, 2) 3-{4-[2-Hydroxy-2-(4-hydroxy-3-methoxyphenyl)-1-hydroxymethylethoxy]-3,5-dimethoxy-phenyl}acryloylcholine, 3) 3-{4-[2-Hydroxy-2-(4-hydroxy-3,5-dimethoxyphenyl)-1-hydroxymethylethoxy]-3-methoxy-phenyl}acryloyl-choline, 4) 3-[2-(4-Hydroxy-3-methoxyphenyl)-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-yl]acryloylcholine.*

Having optimized the metabolomic approach for *Arabidopsis thaliana* the system was applied to rapeseed analysis. For proof of concept experiments we compared the wildtype rape line Drakkar with transformants carrying the ClFatB3 gene, encoding a thioesterase of *Cuphea lanceolata* (Stoll *et al.*, 2005). Five transformant lines were choosen, which showed a reduced long chain fatty acid content (for example oleic acid) and increased shorter chain fatty acid derivatives (for example palmitin). The search for consistent metabolic alterations revealed that about 10% of the detected mass signals (the total sum of reliable integrated *m/z* were up to 300) were either increased or reduced in the transformants (Fig.2). As expected quite a number of the corresponding metabolites appear to be involved in the fatty acid pathway.

The fact that only about 300 mass signals could be reliably detected indicates either that these particular rapeseed lines contained lower amounts of the compounds in question or that a massive matrix effect (due to very high choline ester

contents in rape) masked quite a portion of the natural products during the LC-MS analysis. Weak cation exchange chromatography was employed to deplete the rape seed extracts of the choline ester content, which was then analysed separately. In Fig.3 the results of the fractioning scheme are shown highlighting the four main choline esters in rapeseed.

Material and Methods

The Arabidopsis seed material was extracted with 80% MeOH. The extracts were reduced to dryness and resolved in 10% MeOH. For rapeseed (Drakkar, T98-45/29/1-14 E4; T98-45/21/1-10 C1; T98-45/29/1-16 E7; T98-45/29/1-14 D5; T98-45/29/1-14 E6; Stoll *et al.*, 2005) a pre-extraction step with hexan was included. The dried extracts were resolved in 50% MeOH. For cation exchange chromatography solid phase cartridges were used according to the manufacturers description (strata-X-CW, Phenomenex). The LC-MS profiling was performed as described previously (von Roepenack-Lahaye *et al.*, 2004), using XCMS (http://metlin.scripps.edu) for data deconvolution. The developed data mining tools for the ACD and XCMS software, as well as the ACD / ESI-MS database will be publicly available at http://msbi.ipb-halle.de/MetWare.

Go to Content Page

Development

Plants undergo dramatic changes in habitus and metabolism as they grow and develop. From the resting seed to the mature plant a tremendous amount of changes yet highly regulated occur. Understanding spatial and temporal activation and cut off of regulatory switches affecting development and differentiation provides the basis for improvement and selection of plants.

Go to Content Page

ARABIDO-SEED: Establishing the network of seed gene expression and analysis of its biodiversity

ARABIDO-SEED
Arabidopsis

Gudrun Mönke, Tran My Linh, Udo Conrad, Urs Hähnel, Lothar Altschmied, Bernd Weisshaar 1, Prisca Viehöver 1, Michaela Mohr, Ivo Grosse, Astrid Vorwieger and Helmut Bäumlein
IPK Gatersleben and
1 Bielefeld University

Worldwide, seeds are the main source of human nutrition and animal feed as well as an important resource for the chemical industry. For that reason the understanding of seed formation is essential for plant breeding and genetic modification of plants to improve the properties of seeds with respect to yield, nutritional value and technical use. Since seeds are complex organs composed of several tissues with different genetic constitution, their formation requires many coordinated and highly regulated steps, including cell and tissue differentiation, regulation of growth, accumulation of storage compounds, acquisition of desiccation tolerance as well as reversible maintenance of dormancy.

Transcription factors (TF) are key components for the control of development in response to internal and external signals. Therefore, understanding their action, namely the regulation of their target genes, will direct future efforts in plant breeding and genetic modification. Although major regulators of seed development like ABI3, LEC1, LEC2 and FUS3 have been identified and functionally characterised to some extent, a genome-wide analysis of TF-target gene networks underlying seed development has not been attempted so far.

The trilateral project called ARABIDO-SEED between France, Spain and Germany aims a) at the characterisation of KO mutants defective in relevant TF genes by transcriptome analysis, b) at the identification of TF target promoters by chromatin immunoprecipitation, c) at QTL analysis of transcript levels on a set of RILs to identify regulatory loci, and d) at the development of bioinformatics resources for data handling, integration, analysis and visualisation.

Here, we report the current progress of the German contribution focussing mainly on the identification of TF target genes *via* transgenic lines with regulated TF expression in combination with chromatin immunoprecipitation as well as the development of bioinformatic tools to define TF binding sites in target promoters.

Regulated ectopic expression of selected TF genes

Two experimental approaches for TF gene regulation have been applied: a) TF fusion to the glucocorticoid receptor (GR) domain and b) the estradiol regulated XVE-system. Homozygous *Arabidopsis* lines of the following constructs have been generated: 35S::FUS3::GR; 35S::LEC1::GR; 35S::MYB44::GR; 35S::MYB77::GR; 35S::ET2::GR; 35S::ABI3::GR. The functionality of the 35S::ABI3::GR construct has been proven by transient co-expression of ABI3 controlled gene promoters when induced with dexamethasone, a ligand for the GR domain. Permanent induction of the 35::LEC1::GR construct during germination leads to impaired growth of transgenic seedlings, ectopic embryo formation between root and hypocotyl and embryo like structures on root tips similar to the *pkl* mutant. The latter structures accumulate seed storage proteins as it is shown in figure 1.

The TF mentioned above have also been expressed under the control of the estradiol inducible XVE element and homozygous lines are under selection. A short term induction protocol has been established using the GFP::XVE and the FUS3::XVE lines. Permanently induced LEC1::XVE seedling exhibit partially similar phenotypes as described above for the corresponding GR fusion construct.

Chromatin immunoprecipitation

An affinity-purified antibody against the TF ABI3 was used for chromatin immunoprecipitation. The approach could successfully enrich the seed specific napin 3 gene promoter. The specificity was controlled by precipitation with a rabbit serum IgG fraction, purified by proteinA affinity chromatography. The enrichment was detected by comparing PCR signals using primers for a non-target GAPDH (gap)- gene promoter *versus* the seed specific napin (nap) gene promoter as well as chromatin precipitation with either unspecific (IgG) or ABI3-specific antibodies. The results could be reproduced by semi-quati-

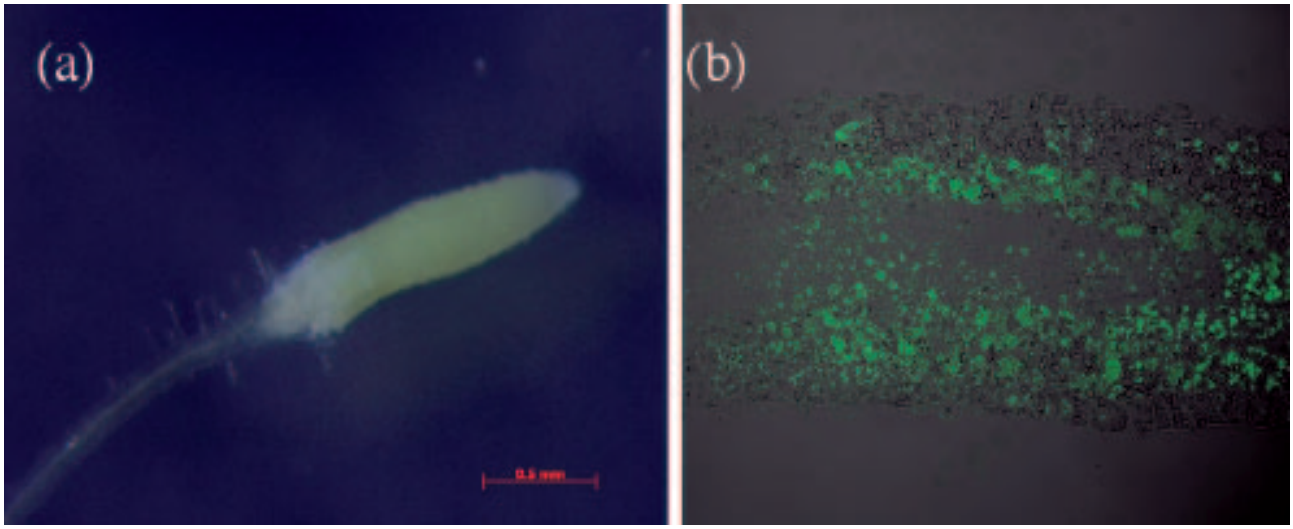


Fig. 1: *Embryogenic structures on root tips of dexamethasone induced 35S::LEC1::GR lines (a) and immunological localisation of cruciferin in these structures (b).*

tative PCR and further confirmed by real-time PCR. The next figure shows the results of a semi-quantitaive PCR experiment.

Promoter array hybridisation

Chromatin probes precipitated either with specific (ABI3) or unspecific (IgG) antibodies have been used for first promoter array hybridizations. For this purpose a collection of promoter fragments for *Arabidopsis thaliana* was obtained from the European SAP initiative (systematic analysis of *Arabidopsis* promoters; www.psb.ugent.be/SAP/). 11904 amplicons from the genomic DNA of *Arabidopsis* were purified and concentrated *via* reverse osmosis. All purified fragments were run on 2.5% agarose gels to determine fragment size and approximate concentration. Fragments were spotted in duplicate on two 12 x 8 cm Nylon membranes using a MicoGrid III robot. In addition to promoter fragments, guide spots containing non-plant plasmid DNA were printed onto these promoter arrays for evaluation purposes.

A labeling protocol for fragmented chromatin DNA obtained from *Arabidopsis* seeds with and without immunoprecipitation was established. This labelling protocol follows the principle of adapter ligation-mediated PCR (Lee *et al.* Nature Protocols 1: 729, 2006). In brief, ends of sonicated chromatin were filled with Klenow polymerase, a non-phosphorylated adapter was ligated, and a polymerase chain reaction (PCR) was performed using the adapter sequence as primer. Amplification products were purified using reverse osmosis and labeled with [_-33P]dCTP *via* random priming.

Due to low signal intensity, an initial hybridization experiment with the promoter arrays yielded only 8 promoter fragments with a significant difference of signal intensities

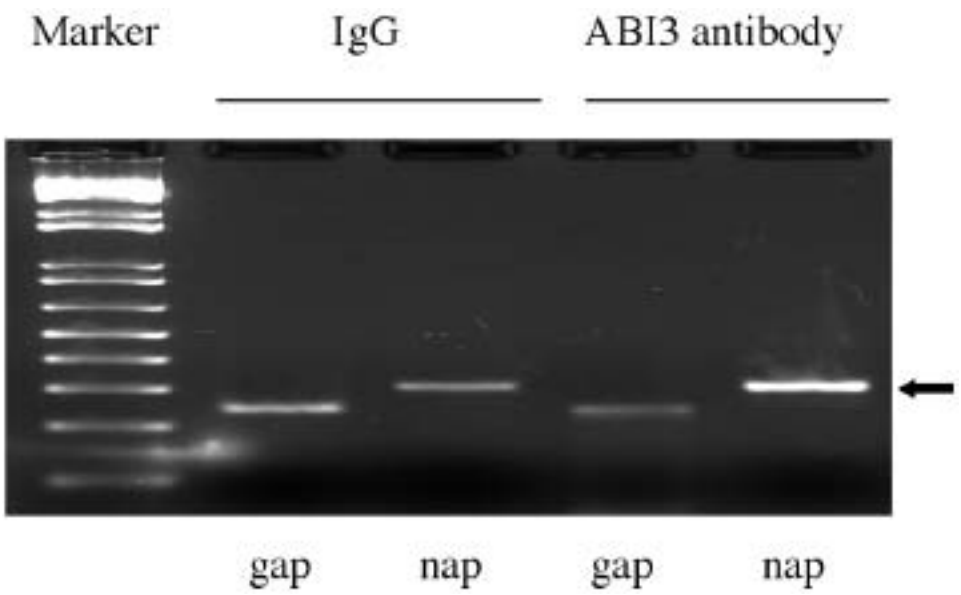


Fig. 2: *Enrichment of the napin 3 (nap) gene promoter by chromatin immunoprecipitation using ABI3-specific antibodies. Chromatin was isolated from developing seed 12 days after pollination and controls were performed using IgG and amplification of the GAPDH (gap) gene promoter.*

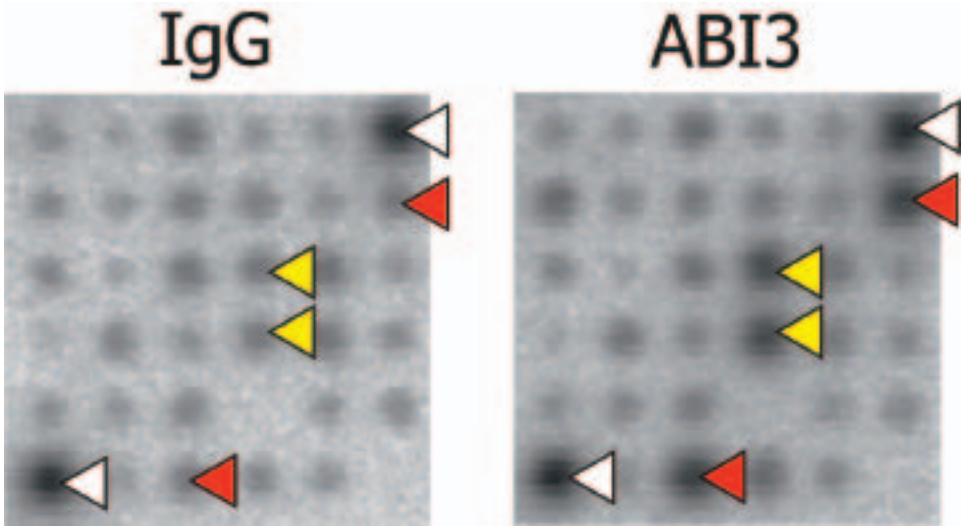


Fig. 3: *Promoter array hybridised with labelled chromatin precipitated with ABI3 antibodies and IgG as control. Duplicated spots of differentially precipitated Arabidopsis promoters are marked with red and yellow triangle while guide spots are marked with open triangles.*

between chromatin precipitated with IgG as control and chromatin precipitated with an antibody for the transcription factor ABI3. Two of these potential ABI3 target promoters contain the expected RY binding motif and the respective genes show increasing mRNA levels on Affymetrix chips (www.genevestigator.ethz.ch/at/) during seed development. Meanwhile the signal intensity in hybridization experiments has been improved significantly (figure 3). Currently, tests for technical and biological reproducibility are under way as well as first attempts to identify the most appropriate control (total chromatin or IgG precipitated-chromatin), to correlate potential target promoters with mRNA levels of the respective genes, and to identify potential transcription factor binding sites in unaligned promoter sequences using novel algorithms developed and implemented at the IPK in Gatersleben.

Prediction of target genes of known TF

The described molecular experiments have been complemented by the development and application of bioinformatic tools to identify and predict putative TF binding sites in target promoters.

Thus, the newly developed program CoMoFinder allows the prediction of target genes that contain a given composite motif in their promoter. Previously, CoMoFinder required the input data as Genbank files. Now CoMoFinder was extended to read the input data directly from the Sequence and Annotation Marts of the Plant Data Warehouse at IPK Gatersleben. CoMoFinder is available at http://www.bic-gh.de/comofinder/.

A novel algorithm for de-novo motif discovery in unaligned sequences was implemented to aid the identification of potential TF binding sites in target promoters identified *via* chromatin immunoprecipitation. Since de-novo motif discovery is computationally demanding, in the past the de-novo motif discovery program EMMA could be used by only one user at a given time, and the runtime of EMMA had to be restricted to 10 minutes per job. This allowed only simple analyses on few and short DNA sequences. To remove these limitations, EMMA has been connected to the Cluster Execution Framework of the Linux Cluster at IPK Gatersleben. Up to 180 jobs can now run in parallel, and the execution time limit could be extended to 24 hours per job. This allows multiple parallel analyses, which may involve several hundred promoter regions of up to 1 kb. Computation on the Linux Cluster can be started *via* a web-interface available at http://www.bic-gh.de/emma/.

Barley as a model and a crop: Gene expression networks determining seed traits (GABI-SEED II)

SEED 2
Cereals

Ulrich Wobus, Hans-Peter Mock, Christoph Pietsch, Volodymyr Radchuk, Marion Röder, Falk Schreiber, Udo Seiffert, Nese Sreenivasulu, Marc Strickert, Katja Witzel and Winfriede Weschke
Leibniz Institute of Plant Genetics and Crop Plant Research, D-06466 Gatersleben

Introduction

Crop plant seeds provide the most important foodstuff for humans and their domesticated animals. Therefore, understanding seed development provides the necessary insight into processes regulating quantity and quality of seed storage products, the major compounds used for food, feed, and industrial processing. Although we do already have a reasonable knowledge on genes being responsible for the synthesis and storage of, for instance, sugars, starch, storage proteins, fatty acids etc., the gene regulatory networks determining parameters such as protein and oil composition, percentage of storage products in a seed, seed size, weight and number as well as other quantitative parameters are mostly unknown. Molecular genetic analysis of such traits is usually based on observable phenotypes, without knowledge of the trait-determining genetic architecture, and on anonymous polymorphisms linked to quantitative trait loci (QTL). However, genomic approaches have massively increased our ability to study genes and regulatory networks involved in complex trait expression. Recently, several genes have been identified and QTL effects were even tracked down to the level of single nucleotides, the quantitative trait nucleotide (QTN). Generally, QTL analyses have illuminated the formal genetic basis of many traits important for seed production and plant breeding. An important step from QTL to genes is the separation of an individual QTL from other segregating loci, i.e. the construction of near-isogenic lines (NIL) or introgression lines. A set of such lines has been used in our investigation. These lines offer specific advantages for studies dealing with the role of the "genomic background" for the expression of defined genes (gene-gene interactions) and for the use of alien genetic material for crop improvement.

The described project aims at using the tools developed in GABI phase I to explore the full potential of wild barley (*Hordeum spontaneum*) germ plasma introgressed into a modern cultivar line (*H. vulgare* cv. Brenda). By using a range of genetic, biochemical, molecular and bioinformatics tools a unique data set is being created allowing the development of a detailed molecular understanding of agronomic relevant seed traits.

The project is centered around Genetical Genomics (a technology to identify trait-related genes by treating gene expression patterns as inherited **Q**uantitative **T**raits and the respective genomic **L**oci as QTLs) with the following specific goals: **(1)** Extensive molecular characterization (transcript and protein profiling) of barley introgression lines for marker-assisted breeding; **(2)** Calculation of QTLs for molecular parameters and comparison to known trait QTLs; **(3)** Extensive genetic mapping of key regulators/e(xpression)QTL-forming ESTs; **(4)** Bioinformatics-based definition of gene networks involved in trait determination; **(5)** Identification of regulatory genes with a significant role in trait expression; **(6)** Specific applied output: definition of developmental- and trait-specific key regulators; trait-specific molecular markers; **(7)** Prospect: targeted attempts of seed trait engineering.

Results

Analysis of barley introgression lines [C. Pietsch, M. Röder] (Scientists in charge)

An advanced doubled haploid backcross population consisting of introgression lines between the spring barley cultivar 'Brenda' and the wild barley *Hordeum spontaneum* accession HS213 has been used for expression profiling, protein profiling and as a mapping population for e(xpression)QTL analysis. Advanced backcross populations represent a library of introgressed chromosomal segments of a wild donor line in the genetic background of an elite variety (Figure 1).

Expression profiling of grain development of the introgression lines at four developmental stages [V. Radchuk, N.Sreenivasulu, M. Strickert,]

In order to assess the genetic architecture of regulatory networks underlying grain traits, a custom made 12K seed array of barley has been used to systematically explore variance in mRNA abundance of developing seeds (4, 8, 16 and 25 DAF) of 22 of the above described doubled haploid introgression lines (BC3-DH). Measuring the transcript level variation between 'Brenda' versus all 22 introgression lines across the four developmental time points covering pre-storage phase (4 DAF), intermediate phase (8 DAF), early storage phase with a highrate of starch synthesis (16 DAF) and later storage phase with ongoing starch, storage protein and lipid accumulation (25DAF) allowed us to discern putative genetic

○ Borisjuk, L.; Hajirezaei, M.; Klukas, C.; Rolletschek, H.; Schreiber,F.: **Integrating data from biological experiments into metabolic networks with the DBE information system.** In Silico Biology 5 (2005) 93-102. ○ Junker, B. H.; Klukas, C.; Schreiber, F.: **VANTED: a system for advanced data analysis and visualization in the context of biological networks.** BMC Bioinformatics 7 (2006) e109. ○ Radchuk, V.; Sreenivasulu, N.; Radchuk, R.; Wobus, U.; Weschke, W.: **The methylation cycle and its possible function in barley endosperm development.** Plant Mol. Biol. 59 (2005) 289-307. ○ Radchuk, V.; Borisjuk, L.; Radchuk, R.; Steinbiss, H.-H.; Rolletschek, H.; Broeders, S.; Wobus, U.: **Jekyll encodes a novel protein involved in the sexual reproduction of barley.** The Plant Cell 18 (2006) 1652-1666. ○ Sreenivasulu, N.; Radchuk, V.; Strickert, M.; Miersch, O.; Weschke, W.; Wobus, U.: **Gene expression patterns reveal tissue-specific signaling networks controlling programmed cell death and ABA-regulated maturation in developing barley seeds.** Plant J. 47 (2006) 310-327. ○ Strickert, M.; Seiffert, U.; Sreenivasulu, N.; Weschke, W.; Villmann, T.; Hammer, B.: **Generalized Relevance LVQ (GRLVQ) with Correlation Measures for Gene Expression Data.** Neurocomputing 69 (2006) 651-659. ○ Strickert, M; Sreenivasulu, N.; Peterek, S.; Weschke, W.; Mock, H.-P.; Seiffert, U.: **Unsupervised feature selection for biomarker identification in chromatography and gene expression data.** Lecture Notes in Artificial Intelligence 4087 (2006) 274-285. ○ Strickert, M.; Teichmann, S.; Srennivasulu, N.; Seiffert, U.: **High-troughput multidimensional scaling (HiT-MDS) for cDNA-array expression data.** Lecture Notes in Computer Science 3696 (2005) 625-633. ○ Weise, S.; Grosse, I.; Klukas, C.; Koschützki, D.; Scholz, U.; Schreiber, F.; Junker; B. H.: **Meta-All: a system for managing metabolic pathway information.** BMC Bioinformatics 7 (2006) e465.

Fig. 1: *Graphical genotypes of the advanced backcross population between Hordeum spontaneum accession HS213 and the spring barley cultivar 'Brenda' with HS213 introgressed segments in red.*

networks that differ between lines and to identify important regulators controlling main storage events during barley seed development. To verify the data two sets of seed material from plants independently grown in a phytochamber have been used. Presently we are comparing two normalization protocols, median centering and quantile normalization methods, for estimating the optimal variance of transcript abundance between the elite parent 'Brenda' and the introgression line set. To place the observed variance of transcript abundance within a broader physiological context we plan to implement correlation analysis of 518,584 expression points to define pathway networks influencing seed traits. The results will allow us to identify groups of introgression lines showing altered changes in storage pathways. These correlation networks will be compared with additionally measured data of metabolites, enzyme activity, and protein levels.

2D-analysis of protein composition of ripe grains (protein profiling) [K. Witzel, H.-P. Mock]

In order to analyze the quantitative and qualitative differences in the proteome of mature seeds of the parent line Brenda and the 42 introgression lines, 2D gel electrophoresis was used. Protein spots differing in expression level were identified by mass spectrometry (MALDI-TOF MS, LC-ESI-Q-TOF MS) after computer-based image analysis.

In a first experiment soluble proteins were analyzed. 20-30 spots per line were found to be differentially expressed and subsequently identified. They represent mostly proteins of primary and secondary metabolism, disease/defence proteins and chaperones involved in protein destination and storage. Most of the identified proteins appeared in several spots, implicating different isoforms (protein variants including post-translational modifications). For instance, Z-type serpin was

Fig. 2: *Expression analysis of protein spots 55 and 155, identified as protein Z-type serpin, in 15 selected introgression lines and the parent line Brenda. Left: estimated expression levels; right: part of the 2-D gel with the respective protein spots. Shown are only lines with expression levels at least 2 times different from Brenda.*

found in 6 different spots. The abundance of these isoforms can vary in different lines as shown in Figure 2.

The identification of 1560 selected spots led to the following result. 50% of the proteins could be identified by using barley data (which are still scarce) in open-access data bases. The remaining spots were analyzed by LC-ESI-Q-TOF MS and *de novo* sequencing. Subsequent data base searches identified another 40% of the spots leading to a total of 90% of identified proteins.

After compilation of all identifiers of the different data bases (NCBI, TIGR, UniProt, KEGG, GO) for each protein, bioinformatics analyses (in co-operation with Christian Klukas of IPK) were started with the aim to visualize the data with the help of the VANTED software (see below) to ease interpretation, to map the data onto biological networks (for instance KEGG) and to identify common patterns within the introgression lines. This enabled us to visualize single protein level differences between lines (see Figure 2), the differences of a single spot between all lines and the integration of orthologous genes classified according to KEGG ontology into metabolic pathways or in gene/protein categories (KEGG BRITE) (Figure 3).

Estimation of QTLs for mRNA expression level variability (eQTLs) and protein quantity and quality variation (pQTLs) [C. Pietsch, M. Röder]

As described above, a set of introgression lines (Figure 1) has been used for eQTL analysis. Gene expression profiles were sampled at different developmental stages (see above) based on the 12K array. Gene-expression signals were quantile-normalised and log2 transformed.

The QTL analysis of approximately 47 000 expression profiles resulted in 1373 valid eQTLs representing 1264 ESTs surpassing a $p \leq 0.05$ genome wide threshold derived by 10.000 permutations. An accumulation of eQTL signals has been observed especially on chromosome 2H and 7H belonging predominantly to eQTLs from 25 DAF.

Due to the expense and large experimental efforts, replication of eQTL experiments described in the literature with independently grown material remains scarce. Since the recovery of QTL signals in independent experiments depends on the heritability of the gene-expression profile and the magnitude of the allelic effects, results can vary considerably in different experiments even under controlled growth conditions. Therefore we conducted eQTL analyses of two aligned data sets from 22 introgression lines of independently grown plant material. The comparative analysis revealed a total of 2720 eQTLs corresponding to 2376 ESTs with a LOD cut-off of 3 in at least one experiment (Figure 4). A set of 698 eQTL in common between both experiments corresponding to 622 ESTs will be subjected to further validation experiments (Figure 5).

Confirmed eQTL signals can be further evaluated in two different ways (Figure 5). Firstly, the development of molecular markers for eQTL and the co-localization of eQTL signal and marker location in a subsequent QTL analysis discriminates *cis*-regulatory against trans-regulatory polymorphisms influencing gene-expression. Furthermore, advanced backcross populations allow the selection of lines harbouring the donor allele at the eQTL position in order to conduct small- to medium-sized follow-up experiments to confirm QTL signals especially for trans-acting factors e.g. 'eQTL hotspots'. Secondly, evaluated eQTL signals can be subjected to a more detailed characterization, for instance by initiating secondary mapping populations from selected introgression lines for fine-mapping or map based cloning strategies.

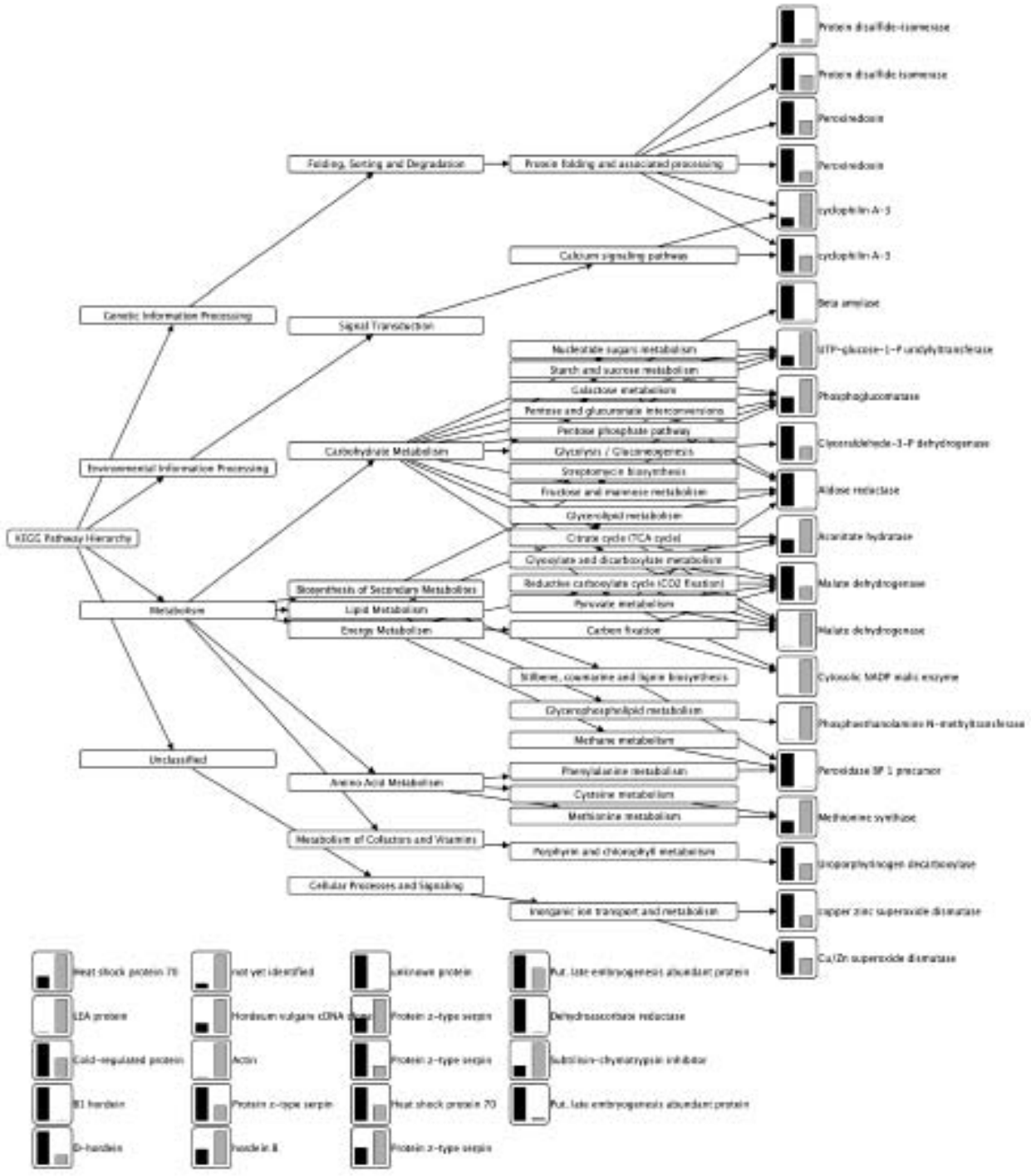


Fig. 3: *Representation of proteome data of the parent line Brenda (black columns) and the introgression line 6-6-3 (gray column) based on KEGG ontology (KO)-identifiers. Whereas KEGG PATHWAY charts are only based on molecular interactions and reactions, KEGG BRITE considers several kinds of relationships at the genetic and molecular level and extends KEGG PATHWAY by defining gene-protein categories. However, only proteins with a KO-identifier can be included. Proteins without identifiers (for instance, B1 hordein, protein Z-type serpin, putative Late embryogenesis abundant protein) cannot be included in the networks. [Data representation programmes were developed by B. Junker, C. Klukas and F. Schreiber]*

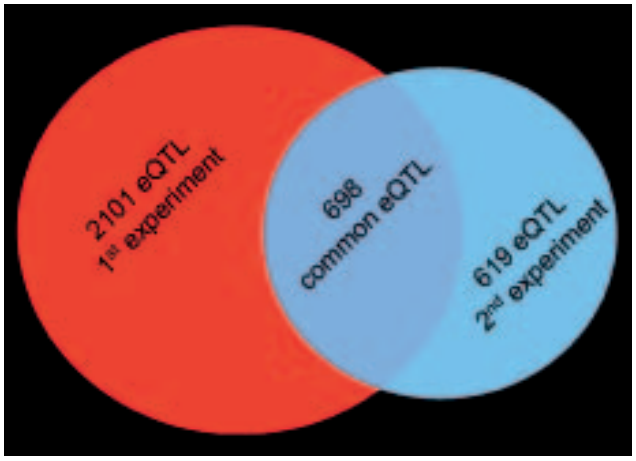


Fig. 4: *Number of eQTLs in two independent experiments applying a LOD score of ≥ 3. Common eQTL are defined by a LOD cut-off of 3 in one experiment and at least of LOD 2 in the other experiment.*

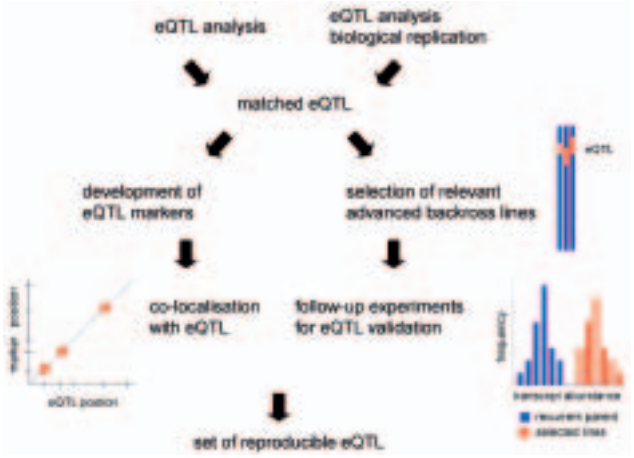


Fig. 5: *Workflow for eQTL validation*

Identification of key regulators of barley grain development

Precise annotation of gene function and tissue-specific distribution of main regulators [N. Sreenivasulu]

For identification of transcription factors (TFs) from the 12K seed array, tBLASTX searches were performed using predicted open reading frame sequences of TFs identified in the *Arabidopsis* and rice genome sequences as well as from TRANSFAC. Furthermore, functional domains were scanned and classified to at least 20 families. In order to identify putative kinases we have used sequence information from Plants P database.

Among the differentially expressed gene set, 508 (21%) encode putative regulators such as TFs as well as hormone and signaling components. Their preferential expression in different tissues during different times of seed development was thoroughly investigated (Sreenivasulu *et al.*, 2006).

Detailed expression analysis of seed development in the parent line 'Brenda' with advanced statistical and artificial intelligence methods [M. Strickert, N. Sreenivasulu, U. Seiffert, V. Radchuk, W. Weschke]

The comprehensive transcriptional profiling studies described above revealed all together 365,397 hybridization signals from 31 duplicated experiments covering four developmental stages. Upon filtering, approximately 20% (2,384 clones) of the total unigene set present on the macro array were retained as representing differentially expressed genes with respect to developmental time and/or tissue.

These high-dimensional data with up to 11786 genes per experiment required utilization and development of robust methods for clustering, classification and factor analysis. The three main requirements were:

1. Availability of different types of similarity for gene expression patterns,
2. Tolerance with respect to background signals and noisy conditions, and
3. Mitigation of the 'curse of dimensionality' which is a common problem in machine learning.

Particularly, prototype-based methods turned out to be suitable for addressing these demands: neural gas (NG) and self-organizing maps (SOM) have been used for data grouping and ordering, respectively; generalized relevance learning vector quantization (GRLVQ) for both classification of labelled gene expression data and simultaneously finding functionally relevant genes [http://srng.webhop.net/]. Thereby, we found that, besides Euclidean distances and Pearson correlation, also rank-based transformations, including Spearman rank correlation, are very useful in combination with the above methods for processing gene expression data. Also the concept of adaptive, data-driven similarity measures proved to be very powerful.

In addition, new visualization techniques (see Fig. 6) have been developed for the display of high-dimensional data, the most important of which being a further improvement of the recently proposed high-throughput multidimensional scaling (HiT-MDS-2) [http://hitmds.webhop.net/]. This method clearly

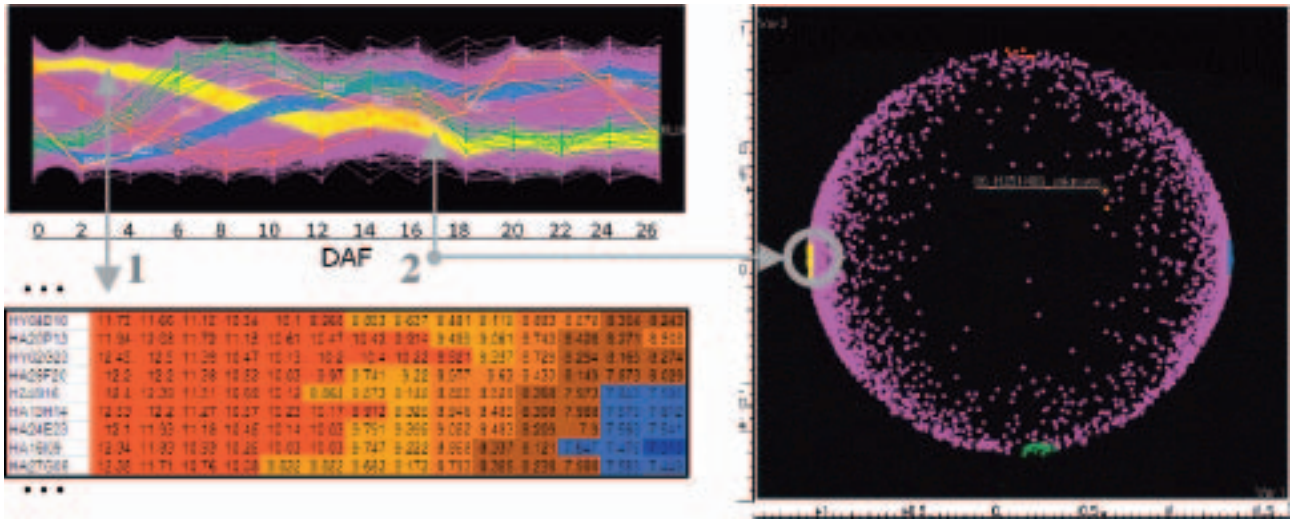


Fig. 6: *Gene space exploration in developing barley endosperm at days 0-26 after flowering. Link 1 connects normalized temporal gene expression data (yellow lines, upper left panel) with data ordered by a self-organizing map (SOM) and displayed in a spreadsheet application (some genes shown in lower left panel). Link 2 shows how specifically down-regulated genes (yellow lines) are associated with points in a correlation-based 2-D space of genes (yellow points in right panel), obtained by HiT-MDS-2. Other differential characteristics are displayed in red, blue, and green. The spatial specificity of the 2-D space and the similarity-preserving mapping allow data screening tasks like gene browsing, visual clustering of regulatory patterns, and outlier detection.*
Finally, a new method has been developed for factor analysis in unlabeled data: in contrast to principal component analysis (PCA), not the factor loadings (eigenvalues) of rotated Euclidean experiment spaces are calculated, but gene ratings are directly derived from original inter-experiment relationships for arbitrary differentiable similarity measures. This approach takes – better than PCA – into account the nature of expression intensity comparisons for the identification of differentially regulated genes.

outperforms other methods like principal component projections in terms of display accuracy and versatility. HiT-MDS-2 allows a systematic screening of inter-experiment relationships, such as developmental stages of developing barley endosperm, pericarp, and embryo tissue. More importantly, two-dimensional displays can be created for subspaces of the 11786 genes, allowing visual association of genes expression patterns with their functional categories.

Visualization of the transcriptional reprogramming of biochemical pathways [F. Schreiber, N. Srennivasulu, B. Junker]

This part of the project deals with the development of a database for metabolic pathways including compartments, transport processes and several other information and the further development of analysis and visualization methods for experimental data in the context of biochemical pathways and functional hierarchies. It has been financed by IPK.

Together with the IPK groups Bioinformatics and Plant Data Warehouse we designed and implemented Meta-All [http://bic-gh.de/meta-all], a plant specific metabolic pathway database. Meta-All stores hand-curated information about metabolic pathways together with the organism and its taxonomy, the location in the plant at which they are occurring, transport processes, detailed kinetic information about the enzymes, and references to publications. Meta-All is used within GABI-SEED II to store structural and kinetic models derived from information obtained in other work-packages and from literature.

New visualization methods have been developed and implemented in the VANTED tool (Visualisation and Analysis of NeTworks containing Experimental Data, http://vanted.ipk-gatersleben.de) for an advanced visual exploration and statistical analysis of complex biochemical data sets obtained in other work-packages of GABI-SEED II. The emphasis is on the linkage of experimental data, for example containing metabolite concentrations, expression profiles or protein levels, with metabolic networks and functional hierarchies such as the gene ontology (GO) or KEGG ontology (KO). It offers a number of statistical functions to analyze the data, is connected to several in-house and external databases (DBE, KEGG, FLAREX DB, Meta-All) and supports different formats for data exchange with other tools. Figure 7 provides one example for visualizations computed with VANTED.

Furthermore, gene sequences present on the 12K seed array were annotated using various reference databases and assigned to MapMan ontology by homology relationships. We developed a MapMan tool for barley and implemented a

Fig. 7: *Visualization by VANTED of transcriptome data taken from endosperm during DAF 0 to 26 in two day intervals. MapMan functional hierarchies of lipid metabolism are loaded onto VANTED graph elements and time series clustered data obtained by self organizing maps (SOM) are visualized based on an automatic graph layout (red – up-regulated, blue - down-regulated, grey – other patterns). The inlet shows the expression profile during DAF 0 to 26 of clone cn8068 representing a GDSL lipase involved in lipid degradation.*

MapMan visualization tool to identify the most relevant pathways involved during the onset of storage in developing barley grains (N. Sreenivasulu in cooperation with B. Usadel).

Cluster analysis of gene expression to identify transcription factors and kinases specifically expressed during reprogramming of biochemical pathways

[N. Sreenivasulu, V. Radchuk, W. Weschke]

For identification of key regulators specifically expressed during reprogramming of biochemical pathways during barley seed development, we first inspected cluster members encoding regulators and biochemical pathway genes. To gain further support for the identified candidates, the promoter regions of the rice orthologs of respective cluster member genes were searched for enriched cis elements. Based on both approaches we have linked regulators to ongoing tissue-specific cellular physiological events during barley seed development between fertilization and the late maturation stage. A process such as Programmed Cell Death (PCD) is executed by different sets of proteases in different seed tissues (pericarp and endosperm). PCD appears to be under the control of ethylene and jasmonic acid signal transduction pathways in the pericarp, and of ethylene and ABA in the endosperm (Figure 7). It is of special interest that several steps in the suggested ethylene pathway are mediated by the expression of tissue-specific gene family members. A further significant outcome of this study is the identification of various key regulators of maturation pathways in endosperm and embryo. Our data also suggests that ABA exerts a control over starch biosynthetic genes in the endosperm *via* SNF1 kinase and several TFs, whereas in embryo ABA influences a different gene network including oleosin and storage globulin genes (Figure 8). Desiccation-related genes in the embryo appear to be controlled in two different but probably interacting ways: ABA dependent *via* ABA responsive element binding factors including DREB 2B, and ABA independent *via* TF DREB 2A.

Summary

In the GABI-SEED II joint project an extended genetical genomics approach is used to delineate the structure and function of regulatory networks underlying grain traits. Molecular and phenotypic characters from transcript abundance to grain weight are mapped on a population of introgression lines [ILs] between the barley spring cultivar 'Brenda' and a wild barley accession. Correlation analyses are still ongoing. As a prelim-

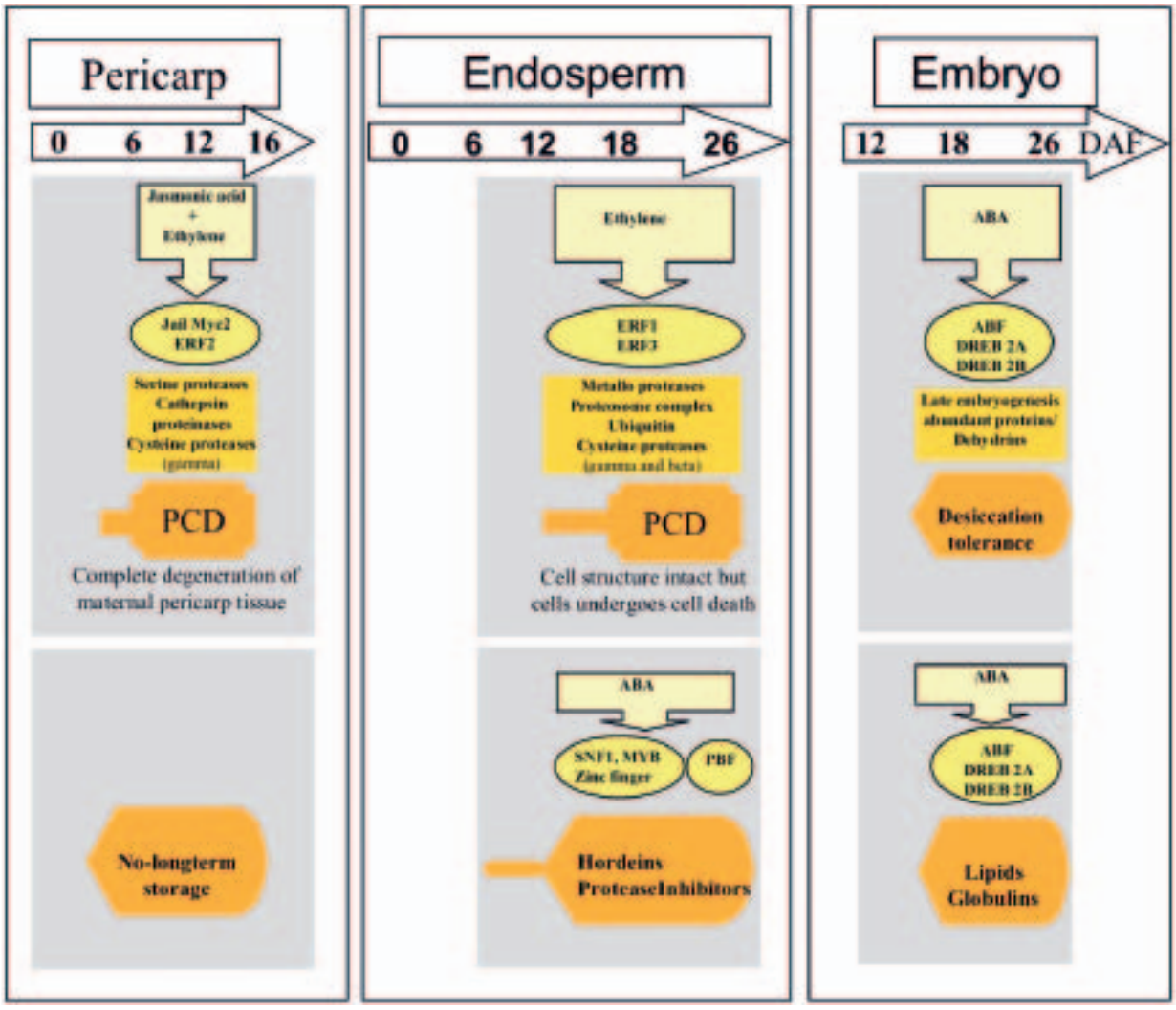


Fig. 8: *Model scheme representing putative key regulators controlling PCD and storage processes in distinct tissues of developing barley seed. Time series transcriptome analysis, selected metabolite profiling and analyzing homeologous rice gene promoters for known cis elements reveal key regulatory components putatively controlling PCD and storage pathways. Only a fraction of the candidate regulatory genes identified in this study are shown here. (Taken from Sreenivasulu et al., 2006)*

inary summery we conclude: (1) expression profiling during 'Brenda' grain development including nearly 12,000 genes revealed putatively co-regulated gene clusters and new insights into, for instance, transcription factor regulated and hormone-dependent pathways; (2) between 22 ILs changes in functionally defined pathways (for instance devoted to storage) are evident; (3) protein profiling of soluble proteins (storage protein analysis is ongoing) of mature grains revealed >1500 protein spots of which ~90% could be identified. As a mean, 20 – 30 spots varied per line; (4) QTL analysis of ~47,000 expression profiles yielded 1373 e(xpression)QTLs representing 1264 ESTs. In an independent experimental repetition approx. 700 eQTLs were confirmed but have to be further validated; (5) new bioinformatics data storage, analysis and visualization tools have been developed to interpret the experimental results. In order to reach more firm conclusions and applicable results a complete compilation and bioinformatics analysis of all data is necessary together with refined experiments in the future.

GENOSOME-Comparative Genomics of Solanaceae Meristems Regulation of potato tuber dormancy

GENOSOME
Potato

Melanie Senning 1; Burkhard Steuernagel 2; Uwe Sonnewald 1; Uwe Scholz 2; Sophia Biemelt 1
1 Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
2 Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany

Introduction

The aim of this part of the GENOSOME project is to identify key regulators of potato tuber dormancy. Potato tuber dormancy characterizes the period during which tuber bud growth does not occur and ends with the appearance of visible sprouting. Naturally, the length of the dormancy period is controlled by exogenous factors such as temperature, water and nutrient supply as well as by endogenous changes. Thus, during the dormancy period tubers develop from a sink into a source organ supporting the outgrowth of the newly developing sprout. This process is accompanied by structural and metabolic changes as well as by an altered level of gene expression. In addition, phytohormons are supposed to play a pivotal role. While ABA and ethylene are believed to act as negative regulators of tuber sprouting, gibberellins (GA) and cytokinins are generally considered as growth promoters.

Although our knowledge concerning changes occurring during tuber dormancy has considerably increased within the last years, the molecular mechanisms triggering tuber sprouting are still unclear. Within the project we are aiming at (i) identifying genes involved in the regulation of meristematic activity, since it is assumed that re-activation of the bud meristem coincides with breakage of dormancy and (ii) unravelling the role of GA in tuber sprouting. In addition, (iii) the expression pattern of meristem-specific promoters from *Arabidopsis thaliana* are investigated in potato plants to analyse whether meristem organisation is conserved between these species and to provide specific promoters for functional analysis of target genes.

Approaches and current achievements

Identification of regulators controlling potato tuber dormancy

In order to identify genes controlling tuber sprouting, changes in gene expression were detected by transcript profiling. Therefore, labelled cDNA probes prepared from dormant and sprouting tuber buds were used to hybridise custom-made macro-arrays and TIGR-chips. These data were stored in an expression database (FLAREX) (http://pgrc.ipk-gatersleben.de/flarex), which has been established at the IPK.

The comparative analysis revealed a number of differentially expressed genes, e.g. using the TIGR-chips 353 genes were found to be up-regulated more than 2-fold, whereas 84 genes were down-regulated at the 2-fold level during tuber sprouting. Among several transcription factors and metabolic enzymes, members of the GA-regulated *GAST1* family were strongly induced during bud breakage. One member of the

Tab. 1: *Characteristics of potato tuber cDNA libraries.*

Name	Tissue, Treatment	Phage-Vector	Plasmid-Vector	No. of ESTs	No. of Singletons	No significant hit in TIGR (%)
STDB	Solanum tuberosum dormant buds	λZAPII,	pBlue SK-	1455	805	8,38
SDBT	Solanum tuberosum dormant buds (TripleEx)	λTriple Ex	Tripl Ex	561	211	0,36
SDBN	Solanum tuberosum dormant buds, **normalized library**		PCR blunt	2028	1622	6,95
SSBT	Solanum tuberosum sprouting buds (TripleEx)	λTriple Ex	Tripl Ex	1661	511	0,66
SSBN	Solanum tuberosum sprouting buds, **normalized library**		PCR blunt	1128	892	4,17
Total				**6833**	**4042**	**15**

Hajirezaei M., Sonnewald U. (1999) **Inhibition of potato tuber sprouting: Low levels of cytosolic pyrophosphate lead to non-sprouting tubers from transgenic potato plants.** Potato Research 42, 353-372. Künne C., Lange M., Funke T.,Miehe H. Grosse I., Scholz U. (2005) **CR-EST: A resource for crop ESTs.** Nucleic Acids Research 33, D619–D621.

Tab. 2: *Overview of currently available expression data sets obtained within the GENOSME project.*

Project partner	Array used	No. of hybridisations	status
CSIC	TIGR Micro-array (potato)	6	hybridized
FAU / IPK	Macro-arrays (potato)	13	integrated
MPIZ	Micro-array (tomato)	10	hybridized
FAU	TIGR Micro-array (potato)	4	hybridized
MPIZ	Affymetrix (tomato)	12	integrated
MPIZ	Affymetrix (tomato)	6	intended
FAU	POCI Agilent (potato)	10	intended

gene family, termed GARP (GA regulated protein), was also identified in previous experiments and therefore chosen as a first candidate for further analysis. The different expression during tuber dormancy was verified by northern blot analysis. However, transgenic plants constitutively silencing the gene by means of RNAi did not show visible phenotypic changes and tubers had no significant altered sprouting behaviour. Besides, the function of other GAST-family members and of selected transcription factors during the onset of sprouting is being investigated.

Although the TIGR-chip was already used for our experiments, genes involved in regulation of meristematic activity might be under-represented. Therefore, the intention was to produce meristem- specific cDNA-arrays. To this end, normalized cDNA libraries were prepared from dormant and active tuber buds; referred to as SDBN (Solanum tuberosum dormant buds normalized) and SSBN (Solanum tuberosum sprouting buds normalized), respectively. 2304 ESTs from SDBN and 1152 ESTs from SSBN were sequenced and BLASTX2 searches were performed against the NRPEP protein database at the German Resource Centre for Genome Research (http://gabi.rzpd.de/). Together with ESTs from cDNA libraries generated previously a cluster analysis was performed indicating that the normalized libraries have a higher percentage (80%) of singletons compared to conventionally prepared cDNA libraries (Table 1). Moreover, BLASTN comparison

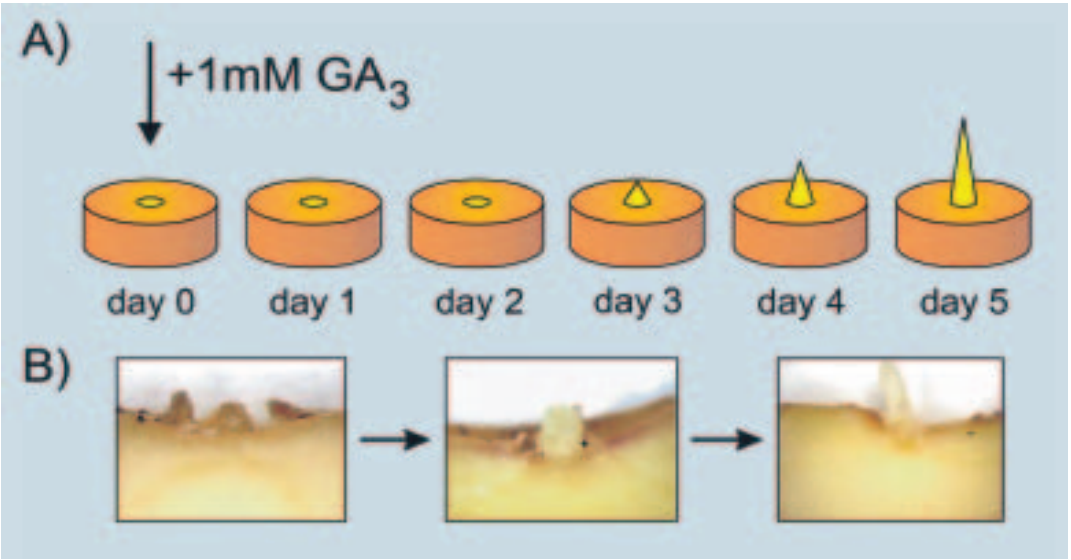


Fig. 1: *GA_3 induced tuber sprouting. Application of a 1mM GA_3 solution to excised discs of potato tubers containing a bud triggers outgrowth of the sprout. A) Schematic illustration. B) Cross sections through tuber discs before and 3 and 5 days after GA_3-treatment.*

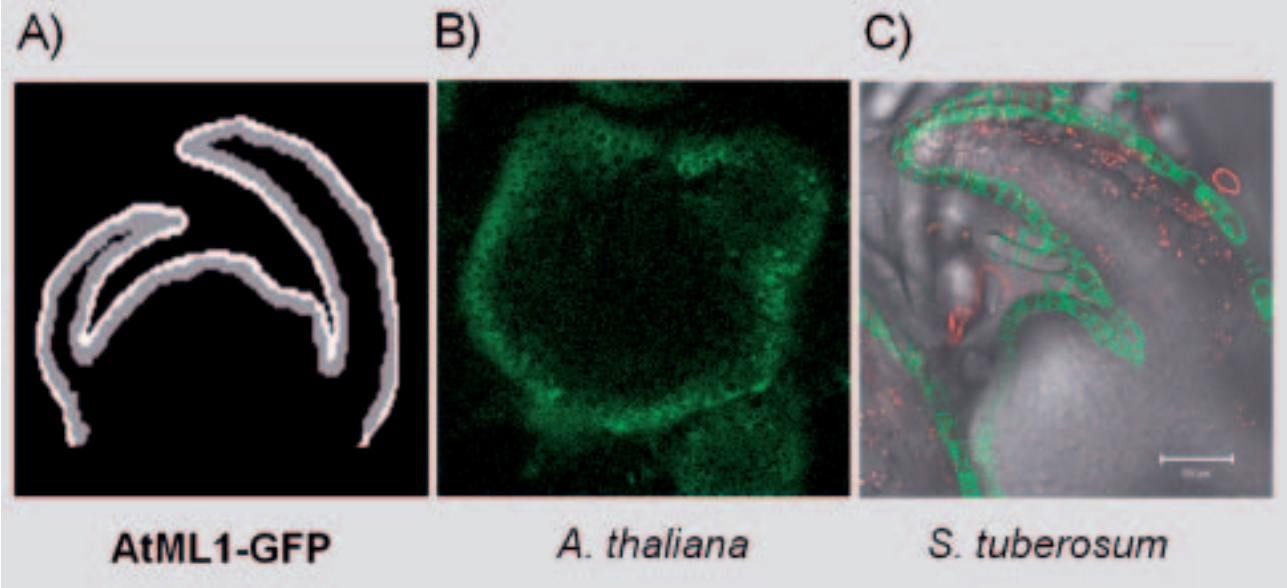


Fig. 2: *Comparative analysis of the AtMERISTEM LAYER1 (ATML1) promoter in Arabidopsis thaliana and Solanum tuberosum. AtML1 promoter is active in the outermost epidermal layer of shoot apical meristem as indicated schematically in (A). B) shows a cross section through the shoot apical meristem of a transgenic A. thaliana plants expressing AtML1:GFP construct (picture is a courtesy of P. Laufs (INRA)). C) shows a longitudinal section through the shoot apical meristem of a transgenic potato plants expressing the same construct indicating that the AtML1 promoter has a similar expression in potato as in Arabidopsis.*

against the available TIGR data (gene index version 10; 38.000 unigenes) revealed that our libraries provide new sequence information (Table 1). EST sequences, BLASTX2 annotation as well as clustering results were integrated in the CR-EST database (Künne *et al.*, 2005; http://pgrc.ipk-gatersleben.de/cr-est/). All EST sequences have recently been released to the GABI- and public community.

All ESTs obtained were included into a cluster analysis to derive a specific set of oligos for the production of a potato oligo chip by "Agilent". This 40K oligo chip covers all public and non-public available potato ESTs, and will commonly be used by a consortium ("Potato Oligo Chip Initiative" (POCI)). Since all available potato ESTs are represented on the POCI-Chip, it is best suited for further experiments.

The role of GA in tuber sprouting

GA is known as an inductor of tuber sprouting, and exogenously applied GA can release tuber dormancy. To unravel the role of GA, transgenic plants have been created expressing the GA biosynthetic genes GA20-oxidase or GA2-oxidase to increase or decrease the amounts of bioactive GAs, respectively. The analysis of transgenic plants expressing these genes under control of the CaMV 35S promoter revealed no impact on sprouting behaviour, although plants show characteristic phenotypic changes and modified levels of bioactive GAs. One reason for the unchanged sprouting behaviour might be the low activity of the CaMV 35S promoter during tuber storage. To achieve higher expression during tuber storage, transgenic plants were generated expressing of both genes under control of the chimeric STLS1/ CaMV 35S promoter (Hajirezaei and

Sonnewald, 1999). The transgenic plants show elongated or stunted shoot growth being indicative for changed GA-levels. Interestingly, the analysis of a first set of tubers revealed slightly altered onset of tuber sprouting. A more detailed analysis with a higher number of tubers is ongoing to verify this result.

In order to determine transcriptional changes associated with GA-induced sprouting, an assay was established allowing the rapid induction of tuber sprouting (Figure 1). To this end, tuber buds are excised from still dormant tubers and incubated in petri dishes on filter paper soaked with 1mM GA_3 solution which triggers tuber sprouting within 4-7 days. This assay also facilitates the analysis of temporally changes in gene expression occurring during the onset of tuber sprouting.

Analysis of meristem-specific promoters from Arabidopsis in potato

The aim of this part is to compare the meristem organisation and function between *Arabidopsis* and members of *Solanaceae* species and to identify meristem-specific promoters suitable to drive the expression of target genes. Promoters which were shown to be active in different parts of the meristem or at different developmental stages in *Arabidopsis* were selected and fused to GUS and/ or GFP as reporter genes (P. Laufs, K. Nikovics, INRA Versailles). These include the proliferation markers CYCLIN B and CYCLIN-DEPENDENT KINASE 2 and a set of five promoters allowing visualization of meristem organization, namely the promoters driving expression of AINTEGUMENTA, WUSCHEL (WUS), SHOOT MERISTEM LESS (STM), UNUSUAL FLORAL ORGANS and AtMERISTEM LAYER1 (AtML1). Because homologues of these genes are not known or are only poorly characterized in *Solanaceae*, regulatory sequences from *Arabidopsis* which had been already described were used. Transgenic potato plants expressing each of the promoter-reporter gene fusions have been generated. The primary screening was performed in the apical meristem. So far, for all promoters activity could be detected in potato except for STM and WUS. The expression pattern of the different promoters investigated was similar in *Arabidopsis* and potato, as exemplified in figure 2 for the AtML1 promoter. The expression is being investigated in other tissues, e.g. flower, roots, stolons, with special emphasis on tuberization and tuber sprouting.

Comparative data analysis

The comparative analysis of expression profiles derived from different tissues and different types of arrays will be done in this part of the GENOSOME project. In cooperation with the BIC-GH group a data warehouse is being developed to integrate heterogeneous data and to facilitate data comparison. Besides expression data from the GENOSOME partners public available expression data from SGED, TAIR and PLEXdb will be imported. Table 2 gives an overview about the current state of this process.

GENOSOME-Comparative Genomics of Solanaceae Meristems Transcriptional control of lateral meristem initiation

GENOSOME
Potato

Gregor Schmitz, Klaus Theres
Max Planck Institute für Pflanzenzüchtungsforschung, Köln

Introduction

During plant development, lateral meristems are formed in the axils of leaves. Later these newly initiated meristems develop into side-shoots which contribute to a large extent to the above ground plant body. The number and size of side shoots influence yield and quality of many plant derived products.

In tomato, side shoot development is undesired. In the two mutants, *lateral suppressor* (*ls*) and *blind* (*bl*), the process of lateral meristem initiation is affected and no sign of lateral meristem differentiation can be detected. Due to side effects on inflorescence and flower development these mutants can not be used for breeding.

Early steps of side shoot development are rather poorly understood, and only a limited number of genes regulating this process have been identified. To find more genes that influence lateral meristem initiation and subsequent steps of lateral bud development the transcriptome of leaf axils will be compared between wild-type and the two tomato branching mutants *ls* and *bl*. Data will be compared to results obtained in the other parts of the GENOSOME project.

Results and Discussion

Transcript profiling of genotypes with different branching capacities

Stages of lateral meristem development during the vegetative growth phase in wild-type tomato have been defined in comparison to the development of the subtending leaf. In axils of young leaves of about 5 mm length, no sign of lateral meristem development was detected, but cells smaller than the surrounding cells were visible along the border between leaf and stem. In axils of leaves of about 15 mm length, newly formed lateral meristems can be detected in the middle of this area as small mounts of tissue just above the site of leaf inception on the developing stem (Fig. 1 A,D). In axils of leaves of about 35 mm length, lateral meristems have developed the first leaf primordia (Fig. 1 B,C,E,F).

lateral suppressor mutants show no sign of lateral meristem formation in vegetative leaf axils. In *blind* plants, most of the leaf axils were like in *ls*, but in the first two or three leaf axils side-shoots developed like in wild-type plants.

Two stages of leaf axil development have been selected to isolate material for RNA preparations from three different plant lines:

Leaf axils of wild-type, and *ls* and *bl* mutants without lateral meristems have been excised from axils of leaves between 3 and 7 mm in size.

Leaf axils in which a small mount of cells can be detected were isolated from axils of leaves between 20 and 35 mm in size from wild-type plants. Axils of the same developmental stages not showing lateral bud development were cut from the mutants.

To isolate axillary tissue, leaves were removed starting with the oldest leaf and from each plant a single axil of the chosen developmental stage was cut with a razor blade and immediately frozen in liquid nitrogen. The isolated tissue sections were about 0.1 to 0.2 mm in all dimensions. For RNA preparations 30 to 40 leaf axils were sampled. Each of the 6 different sample isolations was repeated at least 5 times.

Since only very small amounts of tissue were harvested, RNA was isolated with the RNeasy Micro Kit from Qiagen. Amounts of RNA were determined with a Nanodrop-Spectrophotometer. The obtained amounts were in the range of 1 to 2 µg. Integrity of the RNA samples was confirmed using an Agilent Bioanalyzer.

RNA was amplified using the SuperScript RNA amplification system from Invitrogen. Yield of amplified RNA (aRNA) was between 20 and 30 µg starting from 1 µg of total RNA. Quality of aRNA samples was checked by random primed synthesis of cDNA on the RNA, followed by cloning and sequencing of a limited number of cDNA fragments. Of 15 cDNA sequences, 12 corresponded to single copy protein coding tomato genes and two sequences contained repetitive elements present in many other tomato ESTs. One cDNA

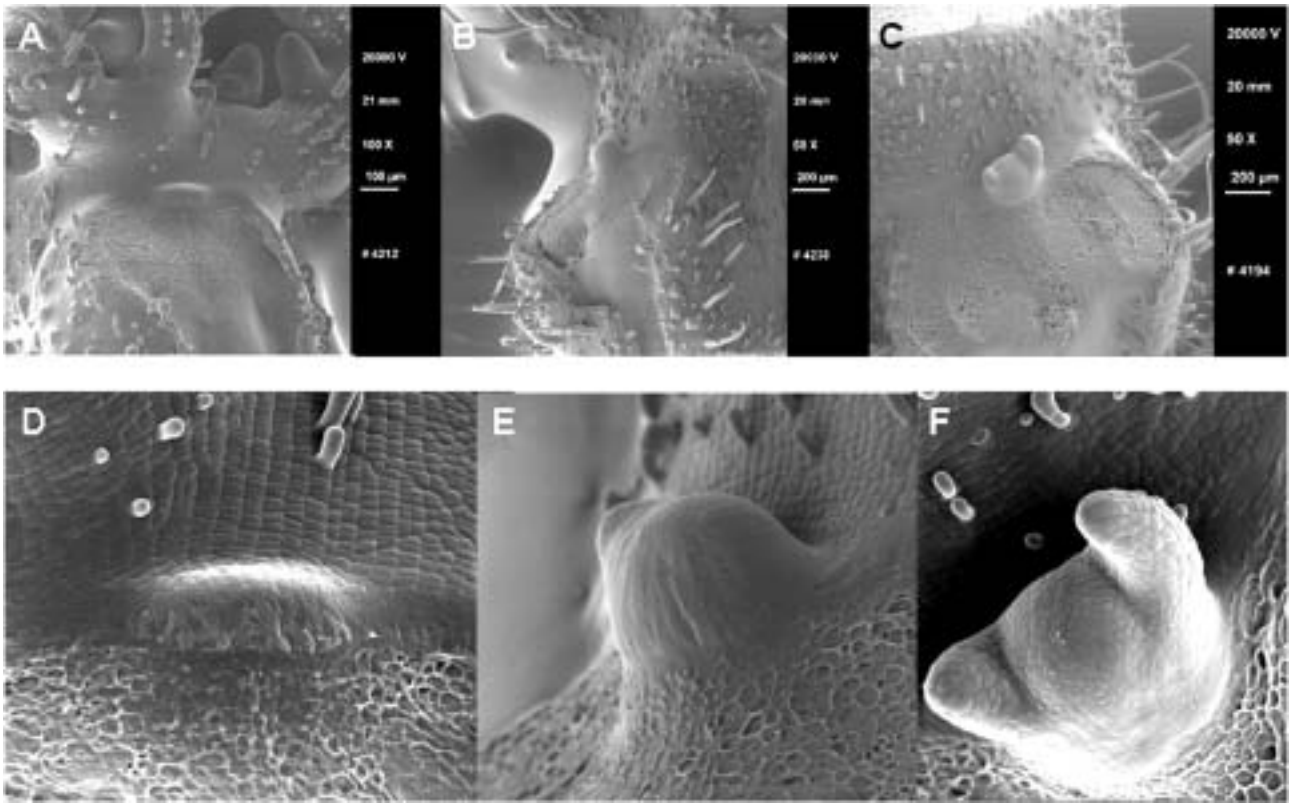


Fig. 1: *Early steps of lateral bud development in tomato. Leaves were removed to visualize the leaf axil region. A,D: Stage of bulge formation. B,F: Lateral meristem with first leaf primordium. C,F: Lateral meristem with two leaf primordia and elongated first internode.*

sequence had no match in the SolGenes database, but the open reading frame displayed high similarity to an Arabidopsis protein sequence. This analysis demonstrated that most of the amplified RNA corresponded to the mRNA fraction of the isolated RNA and can be used for micro-array hybridizations.

Amino allyl-UTP labelled aRNA was synthesized from RNA isolated from the two developmental stages of axil development of wild-type, *ls*, and *bl*. Samples of the RNA were coupled to Cy3 or Cy5 and used for micro-array hybridizations.

These hybridizations were performed at ENSAT in Toulouse using the newly developed tomato oligo chip. The tomato genome oligo set V1.0 contains 12,160 oligos including 300 controls and represents 11,862 unigenes of tomato. Oligos in this set are mostly designed from gene sequences from the Lycopersicon Combined Build #3 unigene database of Cornell University and a number of additional oligos were designed from sequences from GenBank.

Hybridizations will be performed with 4 biological replicates including a dye swap between replicate 1 and 2 versus replicate 3 and 4 (Table 1). Comparisons are done between wild-type and each of the two mutants at both developmental steps resulting in a total of 24 hybridizations. Hybridizations of biological replicates 1 to 3 are finished and replicate 4 is in preparation.

In a second set of experiments the Affymetrix Tomato Chip is used for hybridizations. The 8300 genes spotted on this chip have a partial overlap with the genes on the tomato oligo chip. At least three independent hybridizations will be carried out with the six different biological materials. Results of two sets of hybridizations are available and will be analyzed.

Identifying direct targets of the meristem initiation regulators *Ls* and *Bl*

Constructs containing Glucocorticoid-Receptor-Lateral *suppressor* fusions were introduced into *ls* mutants. Transformant lines that displayed the *ls* mutant phenotype under non-inducing conditions, but a wild-type branching in the presence of dexamethasone were propagated to obtain high numbers of transgenic seeds. Seedlings will be induced with dexamethasone and RNA will be isolated at different time points during the induction process to identify genes regulated directly by *Ls*.

Tab. 1: *Experimental setup of microarray hybridization experiments. WT – wild-type control, ls – lateral suppressor[1] mutant plants, bl – blind[to2] mutant plants. Stages of leaf axil development are defined by the length of the subtending leaf, 5 mm – leaves were between 3 and 7 mm in size, 30 mm – leaves were between 20 and 35 mm in size.*

	Biological replicate			
Dye	**1**	**2**	**3**	**4**
green	**WT 5 mm**	**WT 5 mm**	**ls 5 mm**	**ls 5 mm**
red	**ls 5 mm**	**ls 5 mm**	**WT 5 mm**	**WT 5 mm**
green	**WT 30 mm**	**WT 30 mm**	**ls 30 mm**	**ls 30 mm**
red	**ls 30 mm**	**ls 30 mm**	**WT 30 mm**	**WT 30 mm**
green	**WT 5 mm**	**WT 5 mm**	**bl 5 mm**	**bl 5 mm**
red	**bl 5 mm**	**bl 5 mm**	**WT 5 mm**	**WT 5 mm**
green	**WT 30 mm**	**WT 30 mm**	**bl 30 mm**	**bl 30 mm**
red	**bl 30 mm**	**bl 30 mm**	**WT 30 mm**	**WT 30 mm**

Functional genetic map for branching in tomato

To correlate the activities of specific candidate genes to functions in side-shoot development, populations segregating for different branching behaviour were developed by crossing the weakly branching *Solanum lycopersicum* variety "Moneymaker" to strongly branching wild tomato relatives *S. pennellii*, *S. pimpinellifolium*, and *S. habrochaites*. These populations will be used for QTL mapping of branching traits in addition to the published *S. pennellii* QTLs.

We developed a set of 60 CAPS markers, which are evenly distributed over the 12 tomato chromosomes for mapping in *S. lycopersicum* x *L. pennellii* populations. In a pilot experiment 9 genes were integrated into the genetic map. We mapped six R2R3-repeat MYB genes related to *Blind*, that may be candidates for branching regulators and 3 GAI related genes. A recently published set of CAPS markers (www.sgn.cornell.edu/documents/markers/cosii.xls) can be used for mapping in the same population as well as in *S. lycopersicum* x *S. pimpinellifolium* populations.

Seed material of a first set of 80 mutant lines with altered branching patterns derived from an EMS and from a fast neutron mutagenesis experiment (http://zamir.sgn.cornell.edu/mutants/) has been amplified and mutants with reduced axillary meristem initiation or outgrowth were selected. Additional lines will be analyzed in next spring/summer. DNA of plants with reduced or enhanced branching will be isolated and can be used to identify alterations in candidate genes.

Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance

REGULATORS
Arabidopsis

George Coupland 1; Detlef Weigel 2; Javier Paz Ares 3; Pilar Carbonero 4; Vincent Colot 5; Alain Lecharny 5
1 Max Planck Institute for Plant Breeding, Köln.
2 Max Planck Institute for Developmental Biology, Tübingen
3 Centro Nacional de Biotecnologia, Madrid
4 Ciudad Universitaria, Madrid
5 Unité de Recherches en Génomique Végétale, Evry

Summary of general goals of programme

This programme exploits inter-species conservation in promoter sequences to identify regulators of agronomic performance. Promoters of Arabidopsis genes implicated in reproductive development, phosphate starvation and germination will be compared with those of orthologous genes from 6 other species within and close to the *Brassicaceae*. Once identified, conserved motifs will be used in systematic yeast one-hybrid assays using an ORF library of most known Arabidopsis transcription factors (TFs). This library will be based on a partial library already available from the EC-funded REGIA programme, but will be extended in this work. Candidate TFs will be validated *in planta* by chromatin immunoprecipitation followed by array hybridization. The role of identified TFs in the biological processes of interest will then be examined by identifying T-DNA insertions in the genes encoding them, and assessing the effect on the phenotype of the mutant plants.

Development of communal resources that underpin REGULATORS

A major tool to be developed in this programme was an extensive library of cloned cDNAs of Arabidopsis transcription factors in GATEWAY cassettes that could be used for high-throughput analyses. The transcription factor (TF) library previously made in the EC funded REGIA programme was to be thoroughly characterized and extended by cloning a further 280 TF ORFs into GATEWAY entry clones. All 280 new TF ORFs were cloned in the proposed vector (pENTRY 221) and sequenced. Priority was given to clones known to be incorrect in the REGIA collection based on bioinformatics analysis (see Bioinformatics report below) and to clones suggested by participants. The 280 ORFs cloned so far belong to the MYBR1, GRAS, bZIP, ERF, NAC, Trihelix, ZZ, Chromo- and Bromodomain, and HSF families.

All sequences of the REGIA-TF collection and additional clones from GABI-KAT (carried out previously by ADIS, Bernd Weisshaar, MPIZ) were compiled and evaluated. The analysis revealed that from the 1024 clones in the original library 935 represented non-redundant clones. In addition, approximately 200 of these clones may not be suited for certain analyses, because they contain intron sequences, frame-shifts or are not full-length. All 280 new clones were as expected. Therefore the sequence verified library consists of 1015 non-redundant clones, approximately 50% of the transcription factor complement of Arabidopsis. All of these clones are available in microtitre plates and are fully supported via the internet-based REGULATOR database developed specifically for this project.

The number of Arabidopsis genes that encode potential or known TFs was reevaluated. Different lists of Arabidopsis TF sequences were collected, and combined with an analysis of Pfam domains as well as of specific or known conserved structures in order to construct a complete and up-to-date list of TF genes in *Arabidopsis thaliana*. Multiple alignments, phylogenetic trees and human expertise further support the existing family relationships between TF proteins. These results have been made available to the REGULATORS participants through the web site. Gene and family names are linked to detailed information (such as sequence, gene structure) in the FLAGdb^{++} database (http://urgv.evry.inra.fr/projects/FLAGdb++/HTML/index.shtml). The REGULATOR web site will be publicly available at the end of the programme.

Within REGULATORS, the major application of the TF library was to construct a specialized yeast one hybrid library. To this end 895 of the REGIA TF ORFs and 192 of the new TF ORFs were successfully recombined to the yeast vector pDEST22 and transformed into the yeast strain YM4271. Currently a library of 1080 yeast strains is available (some of which contain redundant clones, see above) for yeast one hybrid screens and this will be extended with the missing 88 new ORFs.

Identification of promoter motifs

The programme is focused on identifying conserved promoter motifs in genes induced during the floral transition, during seed germination and in response to phosphate (Pi) starvation. Genes will be identified using expression arrays and conserved promoter motifs in these genes through promoter sequence analysis in Arabidopsis and relatives in the *Brassicaceae*. In this report, we use the work on Pi starvation as an example. To identify genes associated with Pi starvation,

expression profiling was performed in wild-type plants and in *phr1* mutants, which are impaired in the activity of a key transcription factor involved in response to phosphate starvation. 600 genes were identified as induced and 200 as repressed in wild-type plants in the conditions tested. More than half of the responsive genes show reduced responsiveness in the mutant, which demonstrates the importance of PHR1 in mounting the Pi starvation response. Moreover, the PHR1 binding motif is overrepresented in the promoters of induced genes (55% in Pi starvation induced versus 22% in all Arabidopsis genes). In contrast, in the promoters of the repressed genes the PHR1 binding motif is not over- (or under-) represented, suggesting that repression by PHR1 is indirect.

Phylogenomics and promoter analyses will be used *in planta* to confirm the key role of PHR1 binding sites. For the former approach, promoter sequences from genes involved in flowering, phosphate starvation and germination isolated from several species of the *Brassicaceae* were isolated and conserved motifs were identified through sequence comparisons. So far twelve phylogenomic analyses have been performed. In all cases the results are similar, with several conserved motifs being present in each promoter analysed. In the case of Pi starvation, three promoters that contained the PHR1 binding motif were characterized in between 4 and 6 species. The motif was found to be conserved in the promoters from all *Brassicaceae* species which were analysed. In contrast to simple expectations, only one of the other 15 conserved boxes that were identified was found to be overrepresented in Pi starvation responsive genes in Arabidopsis. This might suggest that PHR1 is quite promiscuous and establishes functional interactions with a vast array of transcriptional factors, and because these bind to many different motifs no additional binding site is particularly enriched in Pi starvation responsive genes.

Functional analyses of conserved motifs in the promoters of genes that are highly and specifically responsive to Pi starvation were carried out. Of the five conserved motifs identified (two corresponding to PHR1 binding motifs), mutation of two of these (corresponding to one of the PHR1 binding sites and an unrelated motif) impaired Pi starvation responses. One could assume that the additional conserved boxes not affecting Pi starvation responses could play a role in determining gene expression in unrelated conditions. However, so far analyses of transgenic plants harbouring the mutant reporters during all developmental stages and organs did not reveal any effect of mutation of these conserved boxes.

Establishment of ChIP to chip technology

To test whether the TF binding sites identified in the Y2H are relevant *in vivo*, chromatin immunoprecipitation will be used. Furthermore by hybridizing the chromatin immunoprecipitated DNA to a tiling array of chromosome 4 the full array of target sites for the TF on this chromosome will be identified. To develop this system for the analysis of flowering-time regulation transgenic plants were generated that express the Arabidopsis transcription factor FLC, which has a major role in repressing flowering, fused to an HA-epitope. ChIP experiments using the HA antibody were carried out and binding of FLC to sites within genes known to be regulated by this transcription factor was demonstrated by showing that fragments from these genes were enriched in the precipitated DNA. The DNA is being hybridized to the chromosome 4 tiling array. In addition, the binding sites of the chromatin associated protein TFL2, which represses the expression of several genes with key functions in flowering were analyzed. Transgenic plants expressing TFL2-HA were made, chromatin immunoprecipitations carried out and the resulting DNA hybridized to the chromosome 4 array. The TFL2 protein did not bind to heterochromatin but exclusively to gene rich regions, and the specific genes to which TFL2 binds were marked with a modified histone mark (trimethylated lysine 27 on histone 3) that was not previously predicted to be associated with TFL2. This work required the development of new statistical approaches for the analysis of the data that will be essential for the analysis of future hybridizations based on transcription factors characterized in REGULATORS. The ChIP to chip technology is currently being transferred to whole-genome tiling arrays employing both longer (Nimblegen) and shorter (Affymetrix) oligonucleotides.

Functional genomics of transcription factors (TFs) in maize seed development

MAIZE-TF
Maize

Wolfgang Werr 1; Alfons Gierl 2; Xavier Sarda 3; Pascual Perez 4 and Peter Rogowsky 5
1 Institut für Entwicklungsbiologie, Universität zu Köln, Gyrhofstr. 17, D-50923 Köln, Germany
2 Lehrstuhl für Genetik, Technische Universität München, Am Hochanger 8, D-85350 Freising, Germany
3 Biogemma SA, Génomique Végétale, 1 rue Pierre Fontaine, 91058 Evry, France
4 Biogemma SA, Biologie Cellulaire et Moléculaire, 24 ave des Landais, F-63177 Aubière, France
5 Reproduction et Développement des Plantes, UMR879 INRA-CNRS-ENSL-UCBL, 46 allée d'Italie, F-69364 Lyon, France

Introduction

The goal of this project is to determine the function of transcription factors (TFs) involved in maize seed development. 13 TFs have been chosen for functional analysis via ectopic expression and/or inhibition of expression (RNAi). The phenotype of the corresponding transgenic maize plants is characterised by morphological observations and cytological sections. Their transcriptome is analysed in order to identify the developmental or metabolic pathways in which the TFs are involved. Alterations in the agronomic performance are pinpointed by NIRS analysis. Carbohydrate flux analyses in selected transgenic plants and inbred lines complement the phenotypic analysis.

The project addresses maize seed development as a whole because early development, maturation and entrance into dormancy are strongly interconnected and all three influence both the yield and quality of the maize seed. The specific knowledge gathered in maize will be useful for future applications concerning these processes in maize as well as in other cereals.

Materials and Methods

Choice of TFs with preferential expression in the maize seed

In an exhaustive approach to identify transcription factors expressed in the maize kernel, the maize EST database of Genoplante was searched by tblastn with 1463 protein sequences of TFs present in the Arabidopsis genome. 693 putative TFs were identified. Their expression profiles were extracted from the data available for the Genoplante cDNA microarray. 69 of them showed higher expression in kernels than in roots and leaves and 33 of the 69 preferential expression in a single seed compartment (embryo, endosperm, pericarp or pedicel). In this list of 33 newly identified TFs, semi-quantitative RT-PCR analysis allowed to give priority to TFs with maxima at certain developmental stages during seed development.

In an alternative approach the maize members of the WOX, NAC and HD-ZIP IV families of TFs, members of which are essential for embryo development in Arabidopsis, were identified in EST assemblies and their expression pattern surveyed by *in situ* hybridisation. For functional analysis priority was given to family members with specific expression in well defined domains of the kernel.

Constructs for maize transformation

Novel integrative vectors geared for Agrobacterium mediated maize transformation were produced that allow GATEWAY based cloning either for over-expression or RNAi. Each vector exists in 3 versions: with the constitutive actin promoter, the constitutive CsVMV promoter and the seed-specific HMWG promoter. All vectors contain a Basta resistance cassette for selection and a GFP cassette for the detection of transgenic seed.

Transcriptome analysis

The Genoplante micro-array consisting of 58752 long oligonucleotides (70 mer) was hybridised with RNA extracted from appropriate tissues from transgenic and wildtype plants. RNA was amplified by *in vitro* transcription. Experiments were performed in biological triplicate as well as with a dye swap.

Retrobiosynthetic NMR analysis

The carbohydrate metabolism in the developing maize kernel was monitored by *in vitro* culture of kernels, labeling with [U-$^{13}C_6$] glucose at 15-22 DAP and analysis of ^{13}C-labeling patterns by NMR or GC-MS. Glucose isotopologue patterns from starch hydrolysates were determined as described by Ettenhuber *et al.* (2005).

NIRS analysis

Near infrared spectroscopy (NIRS) was carried out at the Limagrain facility in Riom (France) on individual kernels stemming from segregating ears. After the analysis the kernels were sown and the plantlets genotyped for the presence or absence of the transgene.

C. Ettenhuber *et al.* **Changes in flux pattern of the carbohydrate metabolism during kernel development in maize.** Phytochemistry (2005)66, 2632-2642 ○ A.-S. Khaled *et al.* **Engrailed-ZmOCL1 fusions cause a transient reduction of kernel size in maize.** Plant Mol Biol (2005)58, 123-139. ○ J. Nardmann and W. Werr **The shoot stem cell niche in Angiosperms: expression patterns of WUS orthologues in rice and maize imply major modifications in the course of mono- and dicot evolution.** Mol Biol Evol. (2006) in press. ○ G. Spielbauer *et al.* **Robustness of central carbohydrate metabolism in developing maize kernels.** Phytochemistry (2006) 67, 1460-1475 ○ R. Zimmermann and W. Werr **Pattern formation in the monocot embryo as revealed by NAM and CUC3 orthologues from *Zea mays* L.** Plant Mol Biol (2005) 58, 669-85.

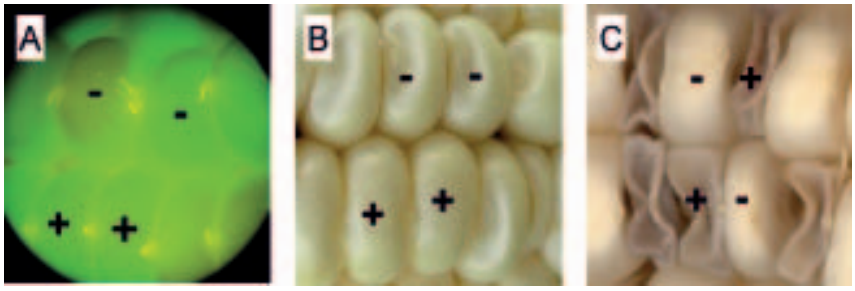


Fig. 1: *Maize ear segregating for Shrunken2 RNAi. Transgenic seed (+) expressing Sh2 RNAi (HMWG promoter) and GFP (CsVMV promoter) are roundish and fluorescent at 22 DAP (A, B) and collapsed at maturity (C). Non-transformed seed (-) is not fluorescent and shows a small indentation at both stages.*

Results

Non-destructive identification of transgenic kernels on segregating ears

Primary maize transformants are hemizygous and the identification of transgenic and non-transgenic seeds on segregating ears has been a painstaking task. In order to overcome this limitation, new transformation vectors allowing a non-destructive and direct identification of transformed seeds were generated. As a common feature all these transformation vectors contain a constitutively expressed GFP.

To demonstrate the effectiveness of these vectors, maize was transformed with an RNAi construct aiming at the repression of *Shrunken2*. The *sh2* mutant is characterised by a severe collapse of the endosperm. Ears of hemizygous plants pollinated by wildtype pollen showed a 1:1 segregation of collapsed, fluorescent kernels and plump, non-fluorescent kernels, demonstrating both the efficiency of the RNAi construct in phenocopying a shrunken phenotype and the usefulness of the GFP marker (Fig. 1).

Spatial expression pattern of selected TFs

In situ hybridisation was not only the final step in the selection process of the 13 TFs for functional analysis but also allowed us to confront the spatial expression profiles of maize genes with the ones present in the literature for their orthologues in Arabidopsis. To this end comprehensive phylogenies were established for the HD-ZIP IV (Khaled et al 2005), NAC (Zimmermann and Werr 2005) and WOX (Nardmann and Werr 2006) families of TFs. Examples of rather specific expression profiles in particular domains of the seed are shown in Fig. 2. For several genes only part of the expression pattern was conserved between maize and Arabidopsis, leading to the conclusion that even in fundamental developmental processes, such as the maintenance of meristems, important differences exist between the two species.

Phenotypic analysis of transgenic plants

Since maize transformation is a long process and since detailed morphological characterisations cannot be done with T0 plants, the bulk of the results will arrive after the end of the project. Nevertheless some results are available for constructs transformed at the very beginning of the project. An example is transgenic plants over-expressing *OCL1*, which showed a pleiotrophic phenotype including an important

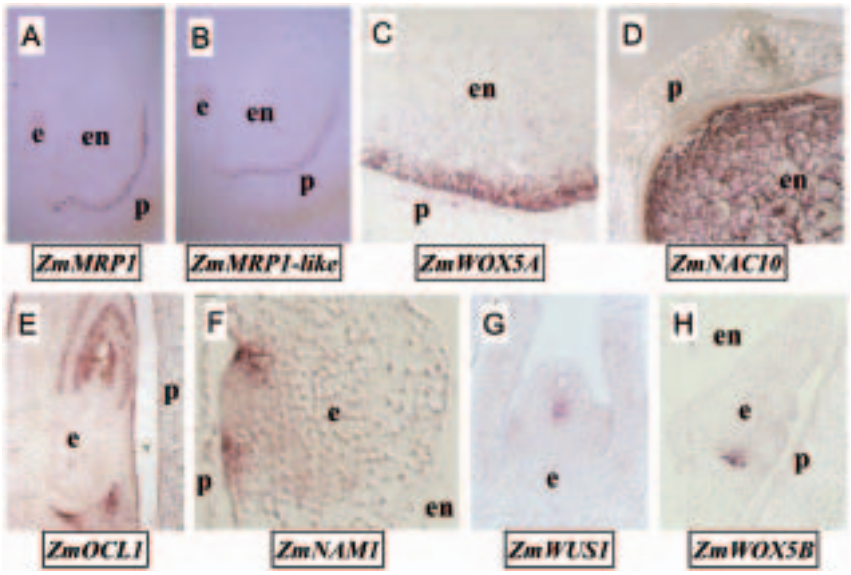


Fig. 2: *In situ hybridisation with selected TFs. Specific expression in the BETL (A, B, C), in the starchy endosperm (D), in the epidermis (E), at the SAM boundary (F), in the SAM (G) or the RAM (H) can be seen. e, embryo; en, endosperm; p, pericarp.*

delay in flowering time. While no morphological aberrations could be detected in cytological sections of the kernels, NIRS analyses predicted an increase in soluble sugars and a decrease in starch (Table 1). Biochemical quantifications will have to be carried out to confirm this result.

Transcriptome analysis of transgenic maize plants

The comparative analysis of the transcriptome is one way to identify developmental or metabolic pathways impaired in transgenic plants over-expressing or repressing target genes. An additional interest in the case of TFs is that at least some of the genes up- or down-regulated likely are direct target genes.

An example was the comparison between transgenic kernels harbouring the *OCL4*-RNAi construct and wildtype kernels from the same 13 DAP ear. A gene list of 23 up-regulated and 32 down-regulated genes was established. A pilot study confirmed more than half of the candidates by semi-quantitative RT-PCR, while the annotations were frequently not informative and did not allow to identify common functions.

Conservation of central carbohydrate fluxes in mutants of starch synthesis

Analysis of metabolic fluxes is an emerging component of functional genomics that is complementary to established "omics" approaches. In this study, the effect of down-regulation of the *Sh2* gene by RNAi was compared to the effect of a null mutation in this gene. *Sh2* encodes the large subunit of AGPase, a key regulator of starch biosynthesis. In agreement with our previous finding that carbohydrate metabolism in maize kernels is robust against genetic and environmental perturbations (Spielbauer *et al.*, 2006), both the *sh2* mutant and *sh2* RNAi kernels showed no massive reprogramming of carbohydrate fluxes compared to the wildtype (Fig. 3). However, both materials showed a significant reduction in the amount of {111111} glucose, which points to an increased activity of processes involved in hexose cycling relative to net starch synthesis.

Tab. 1: *NIRS analysis of OCL1 over-expressing kernels*

Trait	Transgenic (Average)	Wildtype (Average)	Transgenic (StdDev)	Wildtype (StdDev)	t-test* Anova
% All soluble sugars / Dry weight	2,08	1,73	0,42	0,35	0,006
% Fructose / Dry weight	0,11	0,03	0,08	0,09	0,007
% Starch / Dry weight	62,99	65,27	2,87	2,77	0,012
% Fructose / Sugars	5,48	3,89	2,11	2,43	0,026
Methionine	0,26	0,22	0,05	0,06	0,031
% Sucrose / Dry weight	1,65	1,34	0,49	0,39	0,032
% Oleic acid / Lipids	28,09	26,66	1,95	2,41	0,035
% All sugars / Dry weight	85,03	86,29	2,04	1,70	0,040

** traits significant at the 99% level are highlighted in red, those at the 95% level in orange.*

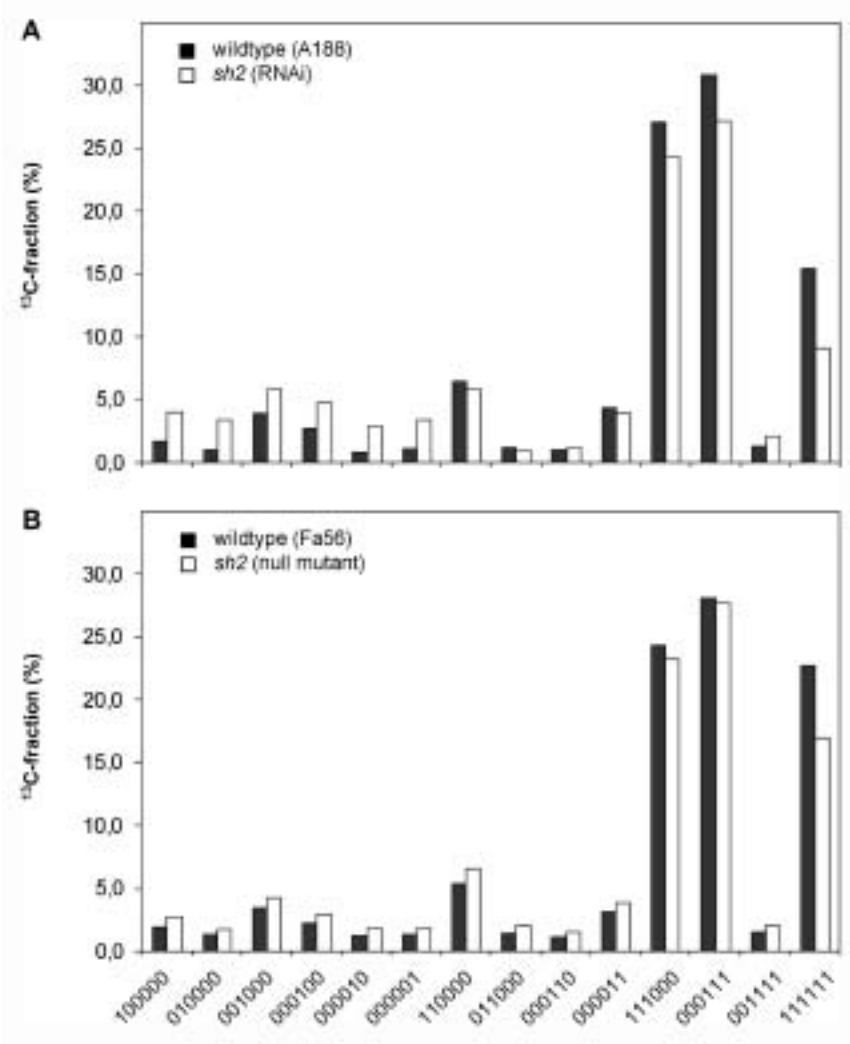


Fig. 3: *Analysis of isotope labeling patterns in starch-glucose from maize kernels. Down-regulation of the Sh2 gene by RNAi (A) affords a phenocopy of the sh2 null mutant (B) with respect to the redistribution of 13C. This demonstrates the ability to modify carbon metabolism using kernel-targeted RNAi.*

Bridging Genomics and Genetic Diversity: Association between Gene Polymorphism and Trait Variation in Cereals

RYE-BARLEY-DIVERSITY
Cereals

Silke Stracke 1, Grit Haseneyer 1,3, Andreas Graner 1, Hartwig H. Geiger 3, Sascha Sauer 2, Viktor Korzun 4, Claus Einfeldt 5, Milena Ouzunova 6, Thomas Presterl 6
1 Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany
2 University of Hohenheim, Institute of Plant Breeding, Seed Science and Population Genetics, Stuttgart, Germany
3 Max-Planck-Institute for Molecular Genetics, Berlin, Germany
4 Lochow-Petkus GmbH, Bergen, Germany
5 Dr. J. Ackermann & Co., Irlbach, Germany
6 KWS SAAT AG, Einbeck, Germany

Génoplante Collaborators: A. Charcosset 7, D. Brunel 8, C. Ravel 9, G. Charmet 9, J.-B. Veyrieras 7, B. Courtois 10, M. Deu 10, M. Rousset 7, A. Murigneux 11, L. Camus-Kulandaivelu 7, C. Dupuits 8, P. Dubreuil 11, D. Manicacci 7, I. Bonnin 7, J. Cornouiller 7, I. Goldringer 7
7 INRA - UPS - INAPG – CNRS, Station de Génétique Végétale, Ferme du Moulon, Gif sur Yvette, France,
8 INRA-CNG, Evry cedex , France
9 INRA, Clermont-Ferrand, France
10 CIRAD-Biotrop, Montpellier, France
11 Biogemma, Aubière, France

During the last decade, linkage disequilibrium-based association studies have received increased attention of plant geneticists and breeders. The success of this approach is highly dependent on the reproduction mode of the species, the genetic population structure of the investigated germplasm, and the evolutionary and/or breeding history. Considering this range of influencing factors, a GABI-Génoplante collaboration with focus on six important cereal species (barley, wheat, rye, maize, sorghum, rice) was initiated to investigate the organisation of nucleotide diversity, linkage disequilibrium, and associations between gene polymorphism and trait variation, among large collections of genetically diverse germplasms. For the comparative approach we selected two model traits in cereals: heading date/flowering time and grain quality covering the content of starch and of major seed storage proteins. The choice of the traits was guided by (i) the agronomic relevance across species, (ii) the possibility to perform accurate and repeatable phenotypic evaluation of large collections, (iii) the existing genetic knowledge and molecular tools, and (iv) preliminary information available from previous Génoplante and GABI programs. The main objectives of this project are:
(1) to gain further understanding of genes involved in controlling the traits of interest and to identify alleles of potential value for practical breeding programs and
(2) to gather knowledge on the potential and applicability of marker/trait associations to species with contrasting reproduction modes and breeding histories.

The working strategy comprises the following steps (Figure 1):

To guarantee sufficient power of the statistical tests, we used collection sizes of 200-400 accessions representing of diverse genetic origins, different selection cycles and/or differences in end use. Phenotypic evaluations of the collection were conducted for each species in independent field trials spanning more than one year and location. Additionally to the field scores, the grain quality analysis was performed by Near Infra Red Spectroscopy (NIRS).

The genetic background diversity within each collection was assessed by using neutral molecular markers (SSR, RFLP or SNP) and the effect of population structure on linkage disequilibrium is considered in the statistical models to test associations. A new statistical framework for analysing the structure of a population was developed by the INRA group

	Heading date/flowering time					Grain quality							
	Hd1	*Hd3a*	*Hd6*	*D8*	*Phy B*	*O2*	*OHP*	*DOF*	*GaMYB*	*CWINV*	*SUT*	*Bt2*	*Sh2*
Maize	■	■		■		■						■	■
Sorghum	■	■	■		■	■						■	■
Rice	■	■	■		■	■							
Wheat	■	■				■	■	■	■				
Barley	■	■		■		■	■	■	■	■	■		
Rye	■	■				■	■		■	■	■		

Tab. 1: *Summary of the selected candidate genes in the different species: heading date 1, 3a, 6 (Hd1, Hd3a, Hd6), dwarf 8 (D8), phytochrome B (Phy B), opaque 2 (O2), O2 heterodimerizing protein (OHP), endosperm-specific DOF-transcription factor (DOF), gibberellin acid regulated MYB-transcription factor (GaMYB), cell wall-bound invertase (CWINV), sucrose transporter (SUT), Shrunken-2 (Bt2), Brittle-2 (Sh2).*

○ L. Camus-Kulandaivelu, J.-B. Veyrieras, D. Madur, V. Combes, M. Fourmann, S. Barraud, P. Dubreuil, B. Gouesnard, D. Manicacci, A. Charcosset (2006) **Maize adaptation to temperate climate: relationship between population structure and polymorphism of *Dwarf8* gene.** Genetics 172: 2449-2463 ○ N. Stein, M. Prasad, U. Scholz, T. Thiel, H. Zhang, M. Wolf, R. Kota, R. Varshney, D. Perovic, I. Grosse, A. Graner (2007) Theor. Appl. Genet. (in press) **A 1000 loci transcript map of the barley genome – new anchoring points for integrative grass genomics.** ○ J.-B. Veyrieras, L. Camus-Kulandaivelu, A. Charcosset (submitted) **Mining population structure using PCA.**

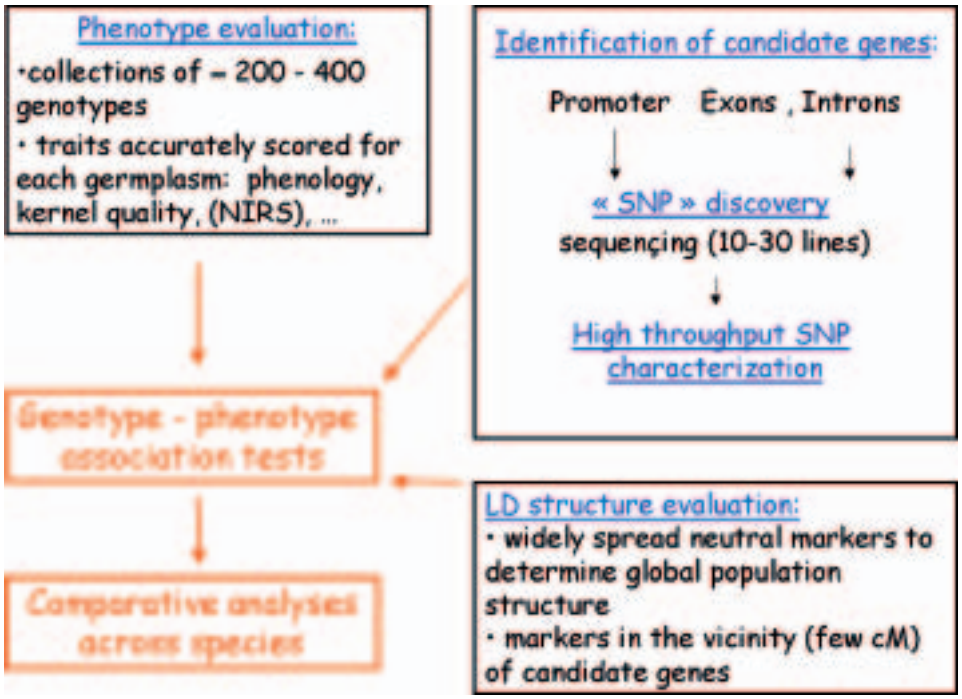


Fig. 1: *Flow diagram of the working plan*

(Veyrieras *et al.*, submitted). This approach, which explicitly aims at controlling the effect of population structure on LD is being used in parallel to the STRUCTURE software for a common data analysis of population structure.

Comparative sequencing of the candidate genes was achieved in one representative set of genotypes for each species in order to identify polymorphisms and to assess the nucleotide diversity of the analysed fragments. This step was followed by high throughput genotyping of the entire collections at the French and the German genotyping platforms.

A set of candidate genes for both traits was selected taking into account biological knowledge on model species, information on co-localisation with QTL, and synteny (Tab. 1). The type of homology of the investigated genes of the different species (orthologs or paralogs) was defined according to mapping position and sequence comparison. Three out of the selected genes (*Hd1*, *Hd3a* and *O2*) are studied for all six species (*O2* alignment, Figure 2) and will therefore provide comprehensive data sets that allow to investigate the likely consequences of reproduction mode and selection history on the organisation of nucleotide diversity and consequently on the choice of statistical models. All other genes considered are studied in at least two species.

GABI Project Part

The German contribution is mainly focused on spring barley. In addition, rye was integrated to identify orthologous candidate genes based on the sequence information from barley and to compare barley and rye for their nucleotide and haplotype diversity and the extent of linkage disequilibrium.

For maize KWS SAAT AG collaborated closely together with the French partners Biogemma and INRA to evaluate a common set of 370 genotypes (Camus-Kulandaivelu *et al.*, 2006) for flowering time and to investigate the selected candidate

genes. The concerted project also benefited from complementary bioinformatics efforts conducted in Germany (contribution of MIPS to the physical map and sequence organisation) and France (development of methodologies for LD analysis).

Barley Collection

The barley collection of 375 cultivated spring accessions comprises 150 German breeding lines provided by the breeding companies Lochow Petkus GmbH and Dr. J. Ackermann & Co. and 225 diverse genebank accessions of the IPK Gatersleben originating from Europe, West and East Asia, and America. The genebank accessions were mainly selected from the Barley Core Collection (BCC) that represents the genetic diversity of cultivated barley.

Phenotyping of the barley collection

The phenotypic variation of the collection was evaluated in 2004 and 2005 at three different locations (University of Hohenheim, Dr. J Ackermann & Co Irlbach, and Lochow-Petkus GmbH Bergen/Wohlde) with three replications each. As targets for association studies in barley, heading date, plant height, and thousand-grain weight were assessed. In addition, raw protein and starch content were measured by NIRS at the FAL in Braunschweig in collaboration with C. Paul. The field trials in both years revealed significantly genetic variations for all traits. The statistical analysis revealed in a high reproducibility of the field data at each location and in both years. As expected, the genebank material showed a higher phenotypic diversity than the breeding material for all traits.

Genetic background structure of the barley collection

The genetic background structure of the collection was assessed by applying 46 EST-based SSR markers developed in GABI Phase I (STEIN *et al.* submitted). According to the frequencies of the marker alleles, the germplasm grouped into two subclusters reflecting two-rowed and six-rowed barley, respectively. This grouping will be considered in the subsequent association studies in order to prevent spurious associations.

Nucleotide diversity of the barley candidate genes

HvCO1 (*Hd1*), *HvFT* (*Hd3a*) and *HvCKA2* (*Hd6*) are investigated as candidate genes for heading date. Additionally, SLN1, counterpart to the maize gene *dwarf8* is analysed for the traits plant height and heading date. With regard to grain storage protein content the transcription factors *HvGaMYB*, *HvBLZ1* (*WOHP*), *HvBLZ2* (*O2*), *BPBF* (*DOF*) are studied, while the diversity in *HvSUT1* and *HvCWINV1* is analysed in respect of the grain starch content and thousand-grain weight.

Due to a high SNP density in the gene fragments, genotyping was mostly performed by direct sequencing covering 600 – 3300 bp of the candidate genes. In few gene fragments showing a low polymorphic rate, site-specific GOOD Assays were developed. SNP density ranged from 1 SNP/18 bp to 1 SNP/192 bp among candidate genes. Comparative analyses between the breeding and genbank materials revealed marked differences in nucleotide and haplotype diversity: Only 1 - 2 haplotypes occurred in the first group but 5 – 19 in the second.

The analyses of association between the identified polymorphic sites and the phenotypic variations will be finalized in collaboration with H.-P. Piepho (University of Hehenheim) in the near future. In case of a significant association, an extended LD study in the physical and genetic vicinity of a candidate gene is indented.

Rye collection

The rye collection comprises 75 genotypes randomly sampled from each of two old open-pollinated varieties (Halo and Carokurz, respectively), 160 modern inbred lines, and 66 genebank accessions (including 5 wild ryes). In view of the planned haplotype analysis, all heterozygous genotypes were crossed to a homozygous inbred line (L301) in 2004, and DNA was extracted from leaf samples of the F1s.

The amplification of rye genes being homologous to the barley genes *HvCO1, HvBLZ1, HvBLZ2, HvGamyb* and *HvSUT1* was successful. The genes will be mapped in order to verify the homology.

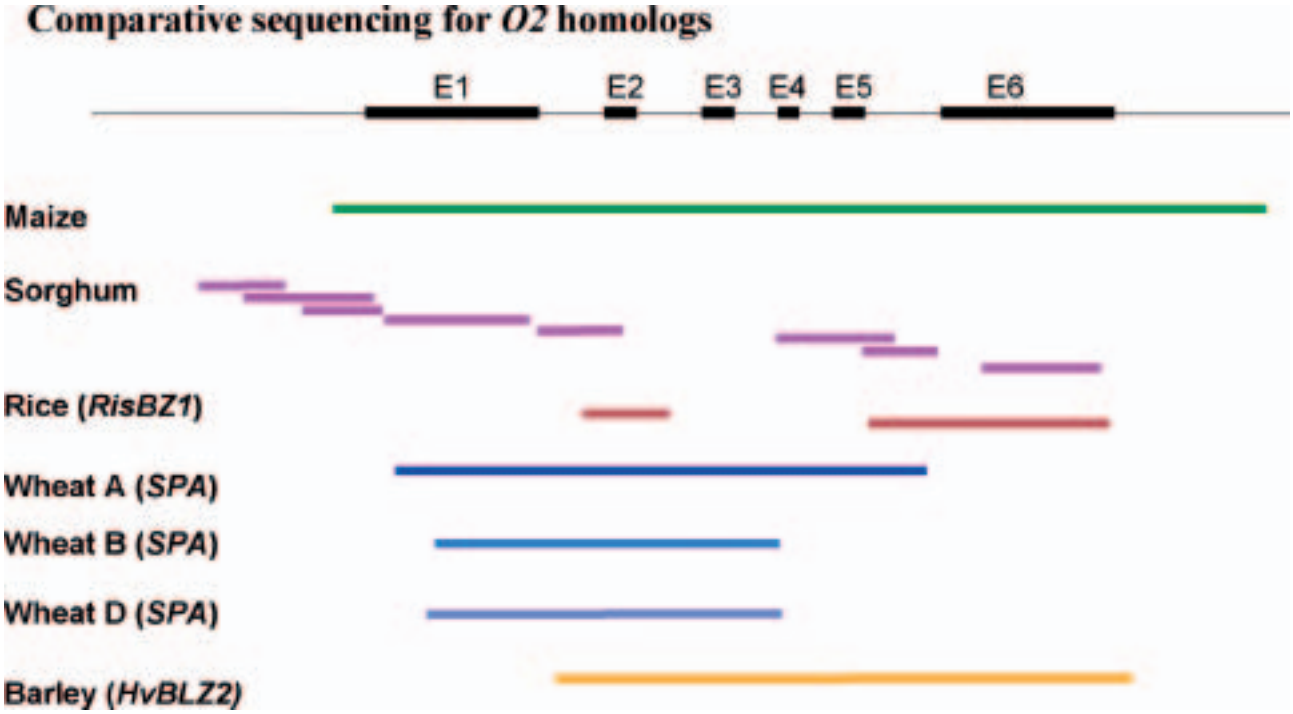


Fig. 2: *Sketch of the interspecific fragment alignment for the Opaque 2 (O2) homologs. Gene structure based on maize with exons E1-E6 is given on the top.*

Stress Adaption

Plants are permanently exposed to environmental deviance continuously interacting with their surroundings. As sessile organisms a change of location is mostly impossible, therefore plants developed multiple ways of response to biotic or abiotic challenge. Understanding these reactions provides knowledge for breeding new varieties being able to cope with predators, pathogens or adverse environmental conditions.

Go to Content Page

Go to Content Page

Impact of modifications in the lignin metabolic pathway on fungal and bacterial pathogen resistance using Arabidopsis mutants and transgenic lines

LIGNIN-METABOLIC PATHWAY
Arabidopsis

Lise Jouanin 1; Dominique Buffard 1; Cao-Trung Do 1; Nikolaus Schlaich 2; Clarissa Mazur 2; Harald Keller 3; Catherine Etienne 3; Anthony Pégard 3; Caroline Levis 4; Pascal Lepècheur 4; Dominique Roby 5; Maurine Tronchet 5; Thomas Kroj 5

1 Biologie cellulaire, INRA, 78026 Versailles, France
2 RWTH-BioIII, Pflanzenphysiologie, Worringerweg 1, 52074 Aachen, Germany
3 UMR Interactions Plantes-Microorganismes et Santé végétale, BP167, 06903 Sophia Antipolis
4 PMDV, INRA 78026 Versailles, France
5 Interactions Plantes-Microorganismes, BP52627, 31326 Castanet-Tolosan cedex, France

Introduction

Phenylpropanoids products or derivatives have been proposed to play a major role in plant defence against pests and pathogens. Lignification constitutes a protective barrier which may hinder or slow the pathogen spread from the infection site. Many programs are underway to modify lignin content and/or structure in order to improve plants for use in paper pulp making or in animal feeding. Until now very few studies have focused on the potential impact of these modifications on pathogen resistance. It is therefore primordial to determine if these modifications which may be beneficial for agro-industrial uses of plants, would not have detrimental consequences on their pathogen resistance.

In the frame of a GABI-GENOPLANTE program, we have proposed to determine if Arabidopsis lines deregulated in the lignin biosynthetic pathway display modified response to fungal and bacterial pathogens known to infect the plant. The lignin characteristics of Arabidopsis knockout lines for genes encoding the different enzymes of the monolignol pathway were determined. The consequences of the mutations on bacterial (*Pseudomonas syringae*, *Xanthomonas campestris*), oomycete (*Hyaloperonospora parasitica*) and fungal (*Alternaria brassicicola*, *Botrytis cinerea*) pathogens were studied. This work will allow to get information on the possibility to improve lignin parameters without negative impact on pathogen resistance.

Material and methods

Arabidopsis mutant lines (ecotypes Wassilewskija (Ws-0) or Columbia (Col-0) were identified in T-DNA insertion collections of Versailles, SALK or GABI. They were characterized with standard molecular techniques. The lignin content and composition of floral stems was determined using Klason and thioacidolysis methods, respectively.

The bioassay tests with the different pathogens were performed according to standard protocols used in the different laboratories.

Results

The main characteristics of the mutant lines were determined and summarized in Table 1. Most of the mutants do not have a specific phenotype when grown in greenhouse conditions. However, the size of the *ccr1* mutants is highly reduced and senescence is delayed and the stems of the *cad C x cad D* mutant are red and limp. In addition, modifications of the soluble phenolic pool (mainly at the level of sinapate esters) are observed in lines such as *comt1* and *ccr1*.

Bioassays with *Hyaloperonospora parasitica* (Keller team)

The oomycete *Hyaloperonospora parasitica* is an obligate leaf pathogen that provokes Downy Mildew of crucifers. Its biotrophic lifestyle is characterised by the formation of haustoria as intracellular feeding structures within the host tissues. The Arabidopsis ecotypes, Columbia (Col-0) and Wassilewskija (Ws-0) are resistant to the *H. parasitica* isolates, Emwa1 and Noco2, respectively, due to resistance gene-mediated genetic resistance, but susceptible to isolates Noco2 and Emwa1, respectively. During the compatible interaction, germinating spores enter the intercellular spaces by the means of appressoria. Intercellular growing hyphae form haustoria inside the host cells (2- to 5 days post-inoculation), and generate conidiophores for asexual reproduction using natural leaf openings (5- to 7 days post inoculation).

In the frame of the present project, we analyzed the influence of mutations in genes coding for the lignin biosynthetic enzymes on compatible and incompatible interactions with the oomycete pathogen. At different developmental stages of wild-type and mutant lines, we compared the parameters of interaction, including intercellular growth, haustoria formation, and the quality and quantity of reproduction.

None of the analyzed mutations had an influence on the outcome of the resistance gene-mediated incompatible interaction between *H. parasitica* and *A. thaliana*. All analyses were, therefore, focalized on the compatible interaction between Col-0 and Noco2, as well as between Ws-0 and Emwa1. All analyzed mutants were infected successfully by the compatible *H. parasitica* isolates. Neither the extent of intercellular hyphal growth, nor the capacity of hyphae to branch was altered in the lignin mutants when compared to wild-type Arabidopsis. The emergence of conidiophores from Arabidopsis leaf tissues is an indicator for the success of the invasion cycle, which completes asexual reproduction. The determination of the number of conidia produced during this

○ Eudes *et al.* **Evidence for a role of AtCAD1 in lignification of elongating stems of *Arabidopsis thaliana*.** Planta (2006) in press. ○ Goujon *et al.* **A new *Arabidopsis thaliana* mutant deficient in the expression of O-methyltransferase impacts lignins and sinapoyl esters.** Plant Molecular Biology (2003) 51: 973-989. ○ Lorrain *et al.* **VASCULAR ASSOCIATED DEATH 1, a novel GRAM domain-containing protein, is a regulator of cell death and defense responses in vascular tissues.** Plant Cell (2004) 16: 2217-2232. ○ Meyer *et al.* **Optimization of pathogenicity assays to study the *Arabidopsis thaliana Xanthomonas campestris* pv. *campestris* pathosystem.** Mol Plant Pathol (2005) 6: 327-334. ○ Siewers *et al.* **Functional analysis of the cytochrome P450 monooxygenase gene bcbot1 of Botrytis cinerea: role in botrydial biosynthesis and impact on virulence in different wild strains.** Mol. Plant-Microbe Inter. (2005) 18:602-12 ○ Sibout *et al.* **Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenase in Arabidopsis. Isolation and characterization of the corresponding mutants.** Plant Physiology (2003) 132: 848-860. ○ Sibout *et al.* **Cinnamyl alcohol dehydrogenase C and D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis.** Plant Cell (2005) 2059-2076. ○ Viaud *et al.* **Identification of calcineurin and cyclophilin A-dependant genes by cDNA arrays analysis in the phytopathogenic fungus *Botrytis cinerea*.** Mol. Microbiol.(2003) 50: 1451-1465

Tab. 1: *Main characteristics of the Arabidopsis mutant lines*

Gene	N° AGI	Mutant	Ecotype	Lignin content	Lignin composition	Reference
CCoAOMT1	At4g34050	*ccomt1*	Col-0	-30%	S/G	
*CCR*1	At1g15960	*ccr1s* *ccr1g*	Col-0	-50%	More condensed	
CCR 2	At1g80820	*ccr2*	Ws-0	=	=	
OMT 1	At5g54150	*comt1*	Ws-0	=	G +5-OH-G	Goujon *et al.*, 2003
F5H 1	At4g36220	*f5h1*	Ws-0	=	G	
CAD C	At3g19450	*cad c*	Ws-0	=	=	Sibout *et al.*, 2003
CAD D	At4g34230	*cad d*	Ws-0	- 10%	coniferaldehyde	Sibout *et al.*, 2003
CADC *CAD D*		*cad C x cad d*	Ws-0	-30%	Coniferaldehyde +sinapaldehyde	Sibout *et al.*, 2005
CADB1	At4g37980	*cad B1*	Col-0	=	=	Eudes *et al.*, 2006
CADB2	At4g37990	*cadB2*	Ws-0	=	=	Eudes *et al.*, 2006

cycle is thus the general mean to analyze differences in susceptibility and resistance between plant lines. The principal finding with respect to this criterion was that none of the mutants was significantly more susceptible to *H. parasitica* infection.

In contrast, we repeatedly observed an increased resistance phenotype of two independent lines that were mutated in the *COMT-1* gene. On these mutants, the production of conidia was significantly reduced on all layers leaves tested from plants at different developmental stages. During the interaction with theses mutants, oomycete hyphae initiated the development of apparently normal feeding structures, but mature haustoria appeared to be deformed, were unstable, and frequently underwent lysis within the host cell. A knock-out of the *COMT-1* gene thus creates an oomycete-unfriendly environment, either through the absence of a compound required for pathogen growth, or through the accumulation of a side product, which is toxic to the microorganism.

An indicator of suboptimal conditions for oomycete growth in plant tissues is the induction of sexual reproduction, which leads to the formation of oospores. Under the defined environmental conditions used in our growth chambers, sexual reproduction of *H. parasitica* is rarely observed, and asexual reproduction is the rule. With the exception of the *COMT-1* mutants, we never detected the formation of oospores on the analyzed wild-type and mutant *A. thaliana* lines after inoculation with *H. parasitica*. However, on both independent lines mutated in the *COMT-1* gene, sexual reproduction of the oomycete was frequently observed, and correlated with reduced conidia production. Moreover, sexual, and asexual reproduction occurred on the same individuals, e.g. presenting the oomycete sporulating asexually on one leaf, but forming oospores on another. The observed reduction in conidia appears thus to be due to a repression of asexual reproduction to the profit of oospore formation, rather than to an enhanced resistance of the mutant lines.

In **conclusion**, modifications in the lignin biosynthetic pathway did not alter the general outcome of compatible and incompatible interactions between *A. thaliana* and *H. parasitica*. Among the about 65 quantitative resistance tests that were performed, we did not observe significantly increased susceptibility of the mutants, but reduced rates of asexual sporulation were frequently observed on the *COMT-1* mutants. However, our findings showed that an analysis of resistance and susceptibility to *H. parasitica* requires more aspects of the oomycete lifestyle to be analyzed, than a simple counting of asexual spores. Using this criterion only, *COMT-1* mutant were more resistant to *H. parasitica*, thus indicating that modifications in lignin synthesis would not have negative impacts on the management of Downy Mildew. Nevertheless,

COMT-1 mutants underwent sexual reproduction, and oospores are the most robust and persisting means of oomycete proliferation. Considering additionally the fact that the formation of oospores allows the pathogen to adapt to new environmental conditions, an alteration of lignin content through *COMT-1* knockouts would bear a clear risk of negatively influence Downy Mildew control in agriculture.

Bioassays with *Xanthomonas campestris* (Roby team)

Xanthomonas campestris pv. *campestris* (*Xcc*), a natural pathogen of crucifers, has been tested on the different mutant lines, using a wound inoculation procedure (Meyer *et al.*, 2005). This inoculation procedure in which bacteria are introduced in the central vein by wounding, closely reflects aspects of a natural infection process, *Xcc* being a vascular pathogen. In addition, this is a semi-quantitative assay, based on the scoring of the disease symptoms. Among the Arabidopsis mutant lines which were tested, only that corresponding to the COMT1 gene (*comt1*), showed a more susceptible phenotype, intermediary between those of the two wild-type ecotypes Ws-0 (resistant) and Sf-2 (susceptible). This line being also, but more clearly, affected to infection by *Pseudomonas syringae* pv. *tomato* (described below), this result has not been further studied.

Bioassays with *Pseudomonas syringae* (Roby and Schlaich teams)

Partner 4 (D. Roby's team). Two strains of *Pseudomonas syringae* pv. *tomato* (Pst) were routinely used by Partner 4: a virulent one (DC3000) and an avirulent one (DC3000/*avrPphB*) on the ecotype Ws-0, the genetic background of most of the lines. In this case, the inoculation procedure consists in the introduction of a bacterial suspension in the leaf mesophyll cells by infiltration. Evaluation of *in planta* bacterial growth was also used to quantitatively assess the susceptibility/resistance of each line (Lorrain *et al.*, 2004). Among the different lines which were found to be affected in their resistance to Pst, the highest difference was observed for the double mutant *cad C x cad D*. The work was thus mainly focused on this mutant. Its susceptibility was clearly established by both tests (bacterial growth 10 fold higher in the double mutant as compared to the wild-type), and shown to be additive of the responses of the two single mutants. This mutant is also affected in the response to several strains of Pst harbouring different avr genes.

The expression of the *CAD* genes was determined during incompatible and compatible interactions. *CAD* D, *CAD* B2 and *CAD* G are induced whereas *CAD* C and *CAD* 1 are repressed. Two hypotheses can be proposed to explain the phenotype of the *cad c cad d* mutant: i) absence or alteration of the lignification process as a defence mechanism, inducing a decrease in basal resistance, or ii) perturbation of defence signalling due to absence or modification in monolignol pools. To test the first hypothesis, expression of genes of the lignin pathway (*C4H*, *OMT*1, *CCoAOMT*1, *CCR*1) was analyzed. The level of expression of these genes was not modified except for *CCR*2 which was induced. For the second hypothesis, genes involved in defence response were tested. *PR2*, *PR5* and *PR3* were not affected whereas *PR1* expression was repressed.

Considering these data, and although the lignification process has not been detectable under our conditions in response to pathogen infection, neither in the mutant line nor in the wild-type, these results suggest that modifications of the lignin content of the cell wall might affect defence responses, and consequently resistance to infection. A manuscript reporting these data, is currently in preparation.

The analyses of N. L. Schlaich's team using the hemi-biotrophic bacterium Pst DC3000 and Pst DC3000 expressing the avirulence gene *avrRpt2* yielded very similar results as described by partner 4.

Bioassays with *Alternaria brassicicola* (Schlaich team)

To test the response of the monolignol biosynthesis mutant lines to a necrotrophic fungus, we used *Alternaria brassicicola* infections. Attempted infections by this pathogen can only be counteracted through so called basal defences, i.e. they are

not strain specific. Moreover, basal defence often is not very effective and thus *Arabidopsis* is usually more or less susceptible to this pathogen. Spore suspensions were applied to the adaxial side of leaves of four week old plants. After three days at high humidity and another day at 70% relative humidity symptoms were scored. Several lines were under these conditions found to be distinctly more susceptible to this pathogen than the wildtype control: *comt1a*, *comt1b*, *f5h1*, *cad C x cad D* and *ccomt*.

It was interesting to see that also for defence against a necrotrophic pathogen the *cad C x cad D* line was more susceptible. Also the fact that mutations in several methyl-transferases render the lines more susceptible to infection by *Alternaria* raises the possibility that methylated monolignols and the polymers derived from them are particularly effective against this fungus.

Bioassays with *Botrytis cinerea* (Levis team)

Botrytis cinerea is a necrotrophic fungus capable of infecting a wide range of plants including *A. thaliana*. To test phenotypic consequences of mutated lignin biosynthetic genes inactivation in the resistance against *B. cinerea*, we used two fungal strains with differences in aggressivity. The T4 strain is less agressive than the B05.10 strain (Siewers *et al.*, 2005). Pathogenicity assays were performed as described by Viaud *et al.* (2003) except that the final concentration was of 10^6 conidia ml^{-1}. Arabidopsis excised leaves were inoculated with 10 µl droplets of conidial suspensions. Disease development was checked at 3, 4, 5 and 7 dpi by measuring the radial spread from the inoculation point to the lesion margin. Pathogenicity assays were repeated ten times and each repetition comprised ten independent measures of lesion size per strain.

Among the nine tested lines, three were found to be significantly affected in their resistance to *B. cinerea*. The *cadB2* mutant was more susceptible to T4 strain infection (+46%) but a little less susceptible to BO5.10 strain infection (-12%) at 7 dpi (i.e. at the colonization stage). The *cadB2* gene is probably involved in stress response and may confer a late partial resistance against the weakly aggressive T4 strain. This hypothesis is in agreement with the expression profile (www.genevestigator.ethz.ch) observed for this gene that shows a 28-fold over-expression during the interaction with *Botrytis*. For the *comt1* mutant, the sensitivity was also increased with the T4 strain (+47%), but no significant difference was observed with B05.10.

On the other hand, the *ccr1g* mutant showed a complete different phenotype. It was more resistant to both strains, but difference in the lesion expansion was observed. The B05.10 strain infection showed a reduction of the colonization at 7 dpi (31%) whereas the T4 strain infection lesion was decreased (-32%) and completely stopped at 3 dpi. It has been shown that the lignin composition of the *ccr1g* mutant is condensed and this modification could be responsible for the increasing resistance against *B. cinerea*.

Discussion

Lignin has a negative impact on forage digestibility and on paper pulp making. Modification of lignin content and composition can improve crops (alfalfa, maize…) and trees (poplars, eucalyptus…). Reduction of lignin content is beneficial for agro-industrial use of plants but need to be moderate in order to have no impact on plant size. In this work, *Arabidopsis* knockout mutants were used to determine the impact of absence of one specific gene on plant morphology, lignin parameters and resistance to pathogens. The results obtained with each mutant suggest that each case is particular. Therefore, if a transgenic line of interest for its lignification pattern is commercialized, pathogenic bioassays with the main pathogens of the transformed target plant need to be performed.

DILEMA: Diversity of Disease Resistance Mechanisms in Arabidopsis

DILEMA
Arabidopsis

Jane Parker 1; Carmen Castresana 2; Dominique Roby 3
1 Max-Planck Insititute for Plant Breeding Research, Cologne
2 National Centre of Biotechnology, Madrid, Spain
3 CNRS Laboratory for Plant-Microbe Interactions, Toulouse, France

Introduction

In response to pathogen attack plants activate complex cellular reprogramming that leads to the induction of a broad spectrum of defences. While much progress has been made in understanding the mechanisms by which plants detect and defend themselves against particular microbes, there has been scant investigation of the potentially rich diversity of resistance traits generated through intra-specific recombination of genomes. These traits are likely to constitute the predominant form of natural disease resistance in the field in which a balance has to be maintained between activation of energy costly defences and the demands of growth, reproduction and effective competition with other species. In DILEMA we aimed to investigate the phenotypic variation of selected geographical accessions, recombinant inbred (RI) populations and defence signalling mutants of the model species, Arabidopsis, to define regulatory steps in resistance to multiple pathogen types. The project framework was designed to integrate information on plant defence responses at the levels of global gene transcription, mRNA translation and key post-translational events and/or regulatory protein activities. Knowledge derived from these studies should provide new biotechnological tools to control plant diseases that remain the most costly and growth limiting factors in agricultural production and food supply.

Goals and Methods

A long term goal of DILEMA is to isolate genes or gene combinations that enhance resistance to both necrotrophic and biotrophic pathogens. The project was divided into four work packages (WPs). In WP1 "Natural Variation" we scored individual lines from a range of Arabidopsis accessions and selected RI populations for response phenotypes to bacterial and fungal pathogens that encompass infection modes from necrotrophy, hemi-biotrophy and biotrophy. Results from these studies have been collated and individual lines that exhibit reproducible differences in resistance and susceptibility to one or more pathogen are being taken on for more detailed analysis. In WP2 "Global mRNA Analysis" the global transcription profiles of wild type and mutants affected in several defence signalling pathways were examined and gene expression signatures identified. Also, different methods to extract polysomal RNA containing actively transcribed mRNAs for profiling have been explored. Our bioinformatics partner, Fabrice Legeai, has been involved in WPs 1 and 2 by establishing a database collating information on pathology phenotypes and accession natural variation. In WP3 "Posttranslational Modifications and Interactions", analysis of key translational and post-translational events that may distinguish resistance-susceptibility phenotypes was undertaken. In particular, we focused on protein ubiquitination and sumoylation, as potentially pivotal controls of plant responses to pathogens and environmental stresses. A further aim was to examine the activities, post translational modifications and inter-molecular associations of defence regulators that were previously isolated by participants of the consortium. In WP4 "Analysis of Gene Function", we have extracted data from gene expression microarrays to identify new components of plant defence that can then be integrated into the existing signalling networks of Arabidopsis.

Results

Natural Variation in Arabidopsis Responses to Pathogens

Screens of Arabidopsis parental accessions for different RI populations were done using the bacterial pathogens, *Pseudomonas syringae* and *Xanthomonas campestris*, an oomycete necrotroph, *Pythium irregulare* and an oomycete biotrophic pathogen, *Hyaloperonospora parasitica*. Analysis of low level (basal) resistance traits to *P. syringae* revealed two major quantitative trait loci (QTLs), *PRP-Ps1* and *PRP-Ps2*, in the RIL population of Bay-0 x Sha (Beyreuth x Shahdara) as shown in Figure 1 [1]. The QTLs were validated in Heterogeneous Inbred Lines (HIFs) selected from Bay-0 x Sha RILs. Analysis of marker gene expression using the HIFs showed that

○ 1. L. Perchepied *et al.* **Natural variation in partial resistance to Pseudomonas syringae is controlled by two major QTLs in *Arabidopsis thaliana*.** PloS One (in press). ○ 2. B. Adie *et al.* **Characterization of Pythium irregulare-*Arabidopsis thaliana* interaction uncovers ABA as an essential signal for pathogen resistance.** Submitted to The Plant Cell. ○ 3. B. Feys *et al.* **Arabidopsis SENESCENCE ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity.** The Plant Cell (2005)17, 2601-2613. ○ 4. M. Bartsch *et al.* **Regulators of salicylic acid-independent EDS1 signaling in Arabidopsis immunity and cell death.** The Plant Cell (2006)18, 1038-1051. ○ 5. T. Colby *et al.* **SUMO-conjugating and SUMO-deconjugating enzymes from Arabidopsis.** Plant Physiology (2006) 142, 318-332. ○ 6. Vellosilo *et al.* **Oxylipins produced by the 9-lipoxygenase pathway regulate lateral root development and defense responses through a novel signaling cascade.** Submitted.

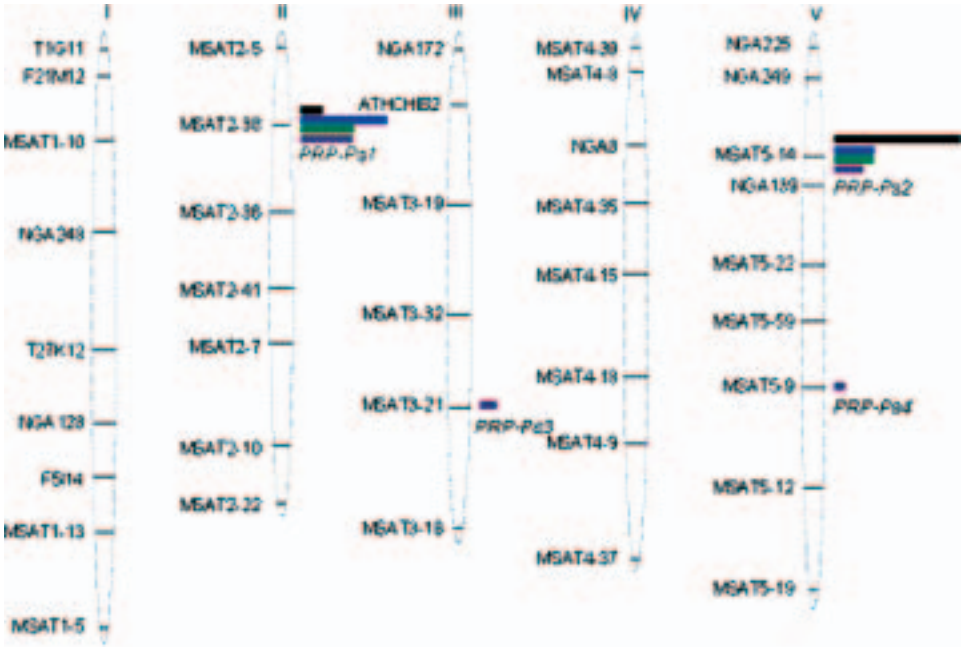


Fig. 1: *Arabidopsis QTLs controlling partial resistance to Pseudomonas syringae pv. tomato DC3000 in the Bay-0 X Shahdara recombinant inbred line population. The detected QTLs are represented by bars located at the closest marker position (black, experiment 1; blue, experiment 2; green, experiment 3; purple, experiment 4) on the Bay-0 x Shahdara genetic map. The length of the bar is proportional to the QTL effect (R2).*

PRP-Ps2 influences the expression of the salicylic acid (SA) dependent defence genes, *PR1*, *ICS* and *PR51*. Further QTLs governing resistance to *X. campestris* were identified in Col-*gl* x Kas-1 RILs, and the fine mapping of the major QTL, representing more than 50% of the variation, is underway. Selected RIL and HIF lines are being tested with different pathogens to ascertain the extent and range of resistance phenotypes. Analysis of natural variation in resistance to *P. irregulare* or *H. parasitica* revealed extensive genetic variability in host responses. Upon *P. irregulare* infection, a strong inverse correlation was found between Arabidopsis accession size and disease susceptibility, suggesting that morphology contributes a large genetic component to the host response. Infections with *H. parasitica* revealed that dominant loci governing acute (cell death-associated) resistance predominate in this highly co-evolved host-pathogen interaction.

Global mRNA Analysis of Arabidopsis Defence

Global expression profiles were examined in Arabidopsis wild type and null *dox1* mutants that have lost α-dioxygenase activity catalyzing primary oxygenation of fatty acids. DOX1 increases in tissues undergoing hypersensitive cell death giving rise to accumulation of new lipid derivatives of unknown function. RNA from leaves responding to inoculation with avirulent *P. syringae* pv *tomato* DC3000 expressing *avrRpm1* was extracted at 0, 8, 24 and 48 h after bacterial inoculation and used to profile transcriptional changes in replicated microarray analyses. Fifteen down-regulated and 26 up-regulated genes in wild type that have significantly different expression in *dox1* mutants have been selected for further study. Further, transcriptome-wide analyses of Arabidopsis seedlings responding to the 9-lipoxygenase (9-LOX) derivative, 9-hydroxylinolenic acid, suggested that the 9-LOX pathway participates in activation of defense and developmental responses that serve to limit pathogen entry and modulate plant growth [6].

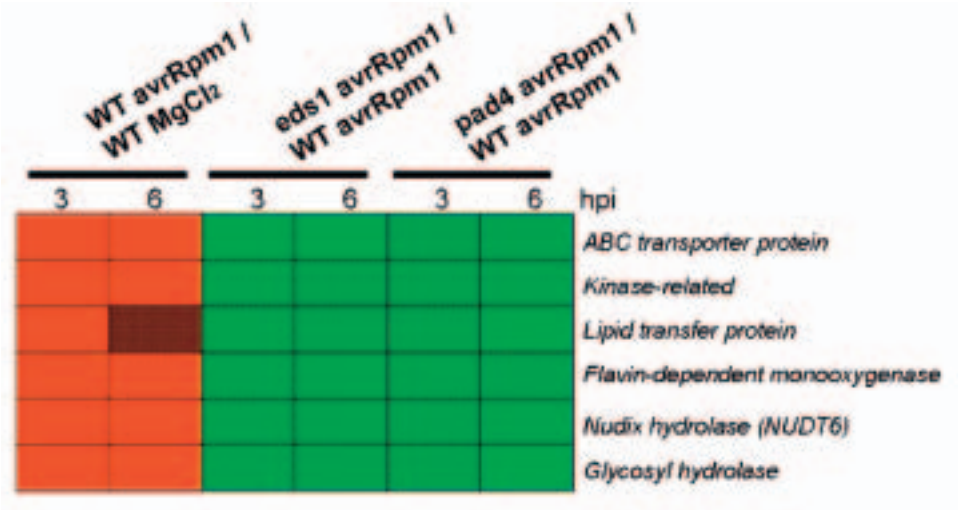


Fig. 2: *Analysis on Affymetrix (ATH1) GeneChips of Arabidopsis genes that are responsive to infiltration of avirulent P.syringae strain DC3000 expressing avrRpm1, reveals a group of six genes, including FMO1 and a highly sequence-related homologue of NUDT7 (NUDT6), whose up-regulation depends on both EDS1 and PAD4. The pathogen inducibility of these genes and its dependence on functional EDS1 and PAD4 was verified by semi-quantitative RT-PCR (not shown).*

Whole genome transcript profiles were also obtained for *P. irregulare*-infected versus mock-inoculated wild type plants and mutants affected in hormonal signalling pathways that together mediate defence to this pathogen (*coi1*: JA; *ein2*: ethylene; *sid2*: SA). This analysis highlights the predominant role of JA and has uncovered an unexpected contribution of abscisic acid (ABA) in activation of defences against *P. irregulare* [2].

Postranslational Protein Modifications and Interactions

One aspect of this project was to define interactions between the central defence regulator EDS1 and its signalling partners, PAD4 and SAG101. Nucleo-cytoplasmic EDS1 complexes control the extent of low level (basal) resistance by regulating production of SA and other as yet unidentified signals [3,4]. Protein interaction analyses in yeast and after purification of *E.coli* expressed proteins indicate that cooperativity rather than competition between EDS1 binding of PAD4 and SAG101 is important for effective defence signal relay. Also, combined over expression of EDS1 and PAD4 in transgenic plants was found to enhance basal resistance to virulent pathogens whereas over expression of EDS1 or PAD4 individually does not alter resistance. The influence of EDS1/PAD4 over expression on responses to infection by other pathogens and on plant growth and development is now being tested within DILEMA. A second aspect of this analysis was to investigate ubiquitination and sumoylation as potentially important control steps in plant defence. An *in vitro* assay was established that has enabled characterisation of the activities of SUMO-conjugating and SUMO-deconjugating enzymes in Arabidopsis [5].

Discovery of Defence Regulators

Gene expression microarray (Affymetrix) data derived compared from a previous analysis of early Arabidopsis wild type responses with those of the *eds1* and *pad4* mutants to *P. syringae* inoculation have provided a wealth of information on further potential basal defence components. Analysis of targeted gene knock out lines for candidate genes has led to the identification of one positive regulator, a flavin monooxygenase (FMO1), and a negative regulator, a cytosolic nudix hydrolase (NUDT7) of the plant immune response (Figure 2) [4]. These new components are now being positioned in the plant defence network and their molecular and biochemical activities characterized.

Discussion and Perspective

The complementary approaches used by partners of the DILEMA programme to identify components of broad spectrum disease resistance have produced important new information. Identification of QTLs for low level resistance to bacteria shows that novel traits can be mapped and potentially cloned. New components and regulatory steps of plant defence have been identified. These activities now need to be integrated, together with the natural resistance traits, into the plant stress response network of Arabidopsis to provide a fuller understanding of the balance of plant defence activation, growth and yield.

Trilateral project: Germany, France, Spain

Exploitation of the natural diversity of grape through functional genomics for improved resistance and quality

CORE-GRAPE-GENE
Grape

Gisela Neuhaus, Rudolf Eibach, Erika Maul, Reinhard Töpfer and Eva Zyprian
Federal Centre for Breeding Research on Cultivated Plants
Institute for Grapevine Breeding Geilweilerhof, 76833 Siebeldingen, Germany

Introduction

European viticulture is endangered by two fungal diseases affecting cultivated grapevine varieties. These are caused by the obligate biotrophic pathogens *Plasmopara viticola* and *Erysiphe necator* (downy and powdery mildew) that invaded European vineyards during the 19th century. Both cause environmental and economical damage as they necessitate an enormous pesticide use. Resistance to these pathogens is absent in European *Vitis vinifera* and exclusively found in North-American and Asian *Vitis* wild species. Such wild grapevines thus are valuable genetic resources that need to be characterized and exploited for resistance breeding in combination with best fruit quality. The genetic diversity in relation to resistance and quality of a representative collection of grapevine accessions including cultured grapevine *Vitis vinifera*, European wild *Vitis silvestris* as well as resistant *Vitis* sp. genotypes and interspecific hybrids is being studied in this trilateral cooperation (CoreGrapeGene, France/Génoplante; Spain/MCyT, Germany/GABI) to finally identify genetic sequence variants responsible of these traits.

The first aim of this joint project is to select a representative grapevine core set comprising a maximum of phenotypic and general genetic diversity with mimimal redundancy from the large *Vitis* germplasm collections present at Vassal (F), El Encin (E) and Siebeldingen (D). This core collection will serve in the second step to analyse sequence diversity of trait related candidate genes. In Germany, the impact is mainly placed on pathogen resistance, while French and Spanish partners focus on fertility and berry quality traits.

Materials and Methods

Phenotypic evaluation of disease resistance

At Geilweilerhof, phenotypic data for fungal disease resistance have been collected under natural infection pressure in field plantations for many years. Data were recorded according to internationally recognized OIV (International Organisation of Vine and Wine) descriptors of quantitative resistance to downy mildew (*Plasmopara viticola*, descriptors OIV 452, 453) and powdery mildew (*Erysiphe necator*, descriptors OIV 455, 456) on leaves and fruit clusters. Resistance is scored in levels from 1 to 9 (1 = fully susceptible, 9 = resistant).

Selection of plant material for genotyping

Based on these phenotypic data, 357 *Vitis* accessions from the *Vitis* germplasm repository at Geilweilerhof comprising fungus-resistant grapevines including interspecific hybrids were chosen. Their putative origin of resistance traits has been followed (as far as known) by studying the pedigrees of the resistant accessions utilizing an internal database. Ten *Vitis vinifera* cultivars were added to the set to facilitate harmonization of the data within the consortium.

Genotyping

Genetic variability within the selected grapevines was characterized by analyzing allelic variation at 20 Simple Sequence Repeat (SSR) loci. Evenly genome distributed markers were chosen based on previous genetic mapping (Adam-Blondon *et al.*, 2004) and used in accordance with the French and Spanish groups. DNA extraction of leaf tissue was performed using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) with addition of PVP40 (1% w/v) to the extraction buffer. PCR amplifications were optimized for annealing temperatures using a Mastergradient Cycler™ (Eppendorf AG, Hamburg) and were performed in a final volume of 10µl using 8 ng template DNA, 1x NH_4-buffer pH 8.8, 0.15 mM dNTPs (each), 1.5 mM $MgCl_2$, 0.25 µM primer and 0.3 U *Taq* DNA polymerase. Primer-specific allelic ladders spanning the range of the expected allele sizes were created for optimal length determinations. Fluorescently labeled primers (6-FAM, HEX and NED) were used for analysis employing an automated capillary sequencer (ABI Prism3100, GeneticAnalyzer, Applied Biosystems, Weiterstadt) in multiplex assays. Amplification products were analysed by GeneScan™ version 3.7 and scored by Genotyper™ version 3.7 (Applied Biosystems, Weiterstadt).

Within the genotype set studied in Siebeldingen a pool of about 35 accessions in common with the French and Spanish sample sets has been selected to harmonize determination of allele sizes and their coding to compare data among the joint groups.

○ Adam-Blondon *et al.* 2004. **Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics.** Theor Appl Genet 109: 1017-1027 ○ Gouesnard *et al.* 2001. **MSTRAT: An Algorithm for Building Germplasm CoreCollections by Maximizing Allelic or Phenotypic Richness.** J Heredity 92: 93-94 ○ Lamboy, W.F. and Alpha, C.G. 1989. **Using Simple Sequence Repeats (SSRs) for DNA Fingerprinting Germplasm Accessions of Grape (*Vitis L.*) Species.** J Amer Soc Hort Sci 123 (2): 182-188 ○ Neuhaus *et al.* **Exploitation of the Genetic Diversity: an approach to elucidate resistance. *Acta Horticulturae*** (submitted) ○ Whitt S.R and Buckler E.S, 2003. **Using natural allelic diversity to evaluate gene function.** Methods in Mol Biol 236: Plant Functional Genomics: Methods and protocols in : E. Grotewold (ed.) Humana Press, Inc. Totowa, NJ

Construction of a core collection

Based on the phenotypic data and the allelic diversity a *Vitis* core set has been created employing SSR data of 357 *Vitis* genotypes by maximizing genetic diversity against individual number. The core set was constructed by using statistical analysis employing the software MSTRAT Vers. 4.0 (http://www.montpellier.inra.fr/gap/MSTRAT/mstratno.htm; Gousnard *et al.*, 2001).

Results and Discussion

The aim of this project is to exploit the genetic diversity existing in cultivated grapevines and wild *Vitis* species for breeding improved resistant grapevines. Most of the genotypes investigated in the German part of the study arose from interspecific crosses derived from resistant grapevines of American and Asian origin. To minimize redundancy of resistance sources in the final core set, the contribution of resistance donors has been determined by pedigree studies for each accession. As evident from these ancestry studies, the selected collection mainly includes source material from the American genepool, e.g. *V. labrusca*, *V. rupestris* or *V. lincecumii*. Only one Asian source – *V. amurensis* – was found as ancestor among the 357 accessions investigated. In total 15 different putative resistance sources in 38 combinations were identified (Tab.1).

Genotyping at 20 SSR loci showed strongly elevated allelic variability in resistant germplasm derived from wildtypes as compared to *Vitis vinifera* sub-collections. The SSR loci in the set of 357 genotypes exhibited about 740 different alleles, on average 37 alleles per locus. The maximum number of different alleles was found at locus SSR13 with 69 variant allele sizes, whereas the minimum number was detected for locus SSR9 exhibiting 13 different allelic length variants (Fig. 1). At most of the loci the selected genotypes showed more than 30 different allele sizes. This high diversity required the use of additional reference alleles for correct size determinations and extended analyses as compared to the *V. vinifera* subsets analysed in France and Spain.

Plants showing only one PCR amplification product at an SSR locus may be either homozygous or heterozygous lacking one allele. This can one allele. This can not be distinguished without segregation data. Also, apparent tri- and four-allelic genotypes still need some verification. Respective genotypes were preliminarily included in the analysis of the core set. In a few cases no amplification was possible. This may be due to the fact that some loci may not be conserved across all the accessions (Lamboy and Alpha, 1998). We expected an extended number of different allele sizes for most of the interspecific hybrids, which holds true. In the study of Lamboy and Alpha (1998) a maximum of 38 different allele lengths was found at the most polymorphic locus in *Vitis*.

For construction of a *Vitis* core set, the redundancy within different core set sizes based on 357 resistant *Vitis* genotypes was estimated and the number of variables in relation to the size of the core set was optimized. These calculations indicated that a subset of 100 genotypes comprises 94 % allelic diversity, whereas 135 genotypes could be defined covering 99 % of the allelic diversity. 150-170 individuals will be sufficient to represent 100 % allelic diversity of the whole set

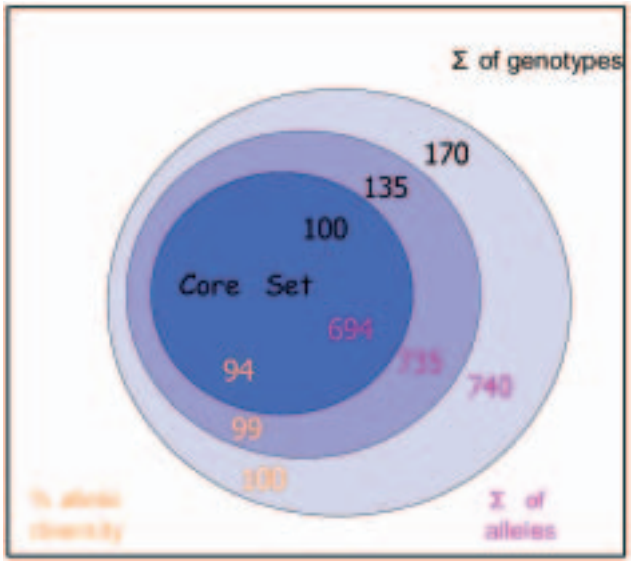


Fig. 2: *Construction of an optimized core set based on the allelic variability of 357 grapevine accessions with resistance towards mildew diseases. Numbers of genotypes are shown in relation to corresponding allelic diversity coverage.*

Tab.1: *Summary of resistance source combinations identified within Vitis genotypes with known pedigree out of the 357 accessions analyzed. Putative resistance donors are indicated by green typing in the parentages.*

Genotype	Location	Col.[1]	Use[2]	SP[3]	Parentage	Resistance Source[4]	F(%)[5]
Concord	53-16-012	N	W, T	V. lab.	not known	*V. lab.*	15,6
Villard Blanc	53-23-012	B	W	I.C.	Seibel 6468 x Subereux	*V. aest., berl., cin., linc., rup.*	14,8
Roi des Noirs	67-43-023	N	W	I.C.	S. 29 x Danugue	*V. linc., rup*	10,9
Siegfriedrebe	53-13-011	B	W	I.C.	Oberlin Noir S.P. x Riesling	*V. rip.*	9,0
Cascade	53-23-006	N	W	I.C.	Seibel 7042 x Gloire de Seibel	*V. aest., cin., lab., linc., rip., rup*	5,8
Leon Millot	53-17-007	N	W	I.C.	Mill. et Grass. 101-14 S.P. x Goldriesling	*V. rip., rup.*	5,5
De Chaunac	53-16-005	N	W	I.C.	Seibel 5163 x Seibel 793	*V. lab., linc., rip., rup.*	4,7
Chambourcin	53-16-004	N	T	I.C.	S. V. 12-417 x Chancellor	*V. aest., berl., cin., lab., linc., rip., rup.*	3,9
Couderc 161-49	53-19-008	N	R	I.C.	*V. berlandieri x V. riparia*	*V. berl., rip.*	2,7
Humbert-Chapon 3	67-41-036	N	W	I.C.	? x Gaillard 2	*V. lab., rip., rup.*	2,7
Rubilande	53-23-004	R	W, T	I.C.	Bienvenu x Roi des Noirs	*V. aest., cin., linc., rup.*	2,3
Othello	67-43-113	N	W	I.C.	Clinton x Black Hamburg	*V. lab., rip.*	2,3
V. rupestris du Lot	76-50-011	-	R	V. rup.	*V. rupestris*	*V. rup.*	2,0
Lednice C 62	53-14-019	B	-	I.C.	(*V.amur.* x Malingre Precoce O.P.) x Kocsis Irma	*V. amur.*	1,6
Sori	53-13-013	N	R	I.C.	Solonis x V. riparia	*V. longii, rip.*	1,6
Cook	53-16-014	N	W	I.C.	(King x Delaware) x Niagara	*V. aest., lab.*	1,2
Malegue 57	67-40-084	N	-	I.C.	*V. berlandieri* x Blanquette	*V. berl.*	1,2
Freiburg 52-62	53-24-011	B		I.C.	Traminer x (S.V. 12-481 x (P. gris x Chass. bl.))	*V. berl., rup.*	1,2
Remaily 66-54-4	53-14-026	N	-	I.C.	*V. doaniana* O.P.	*V. doaniana*	1,2

(Fig.2). This number is much higher than the 48 respective 60 cultivars identified in France and Spain to represent the *Vitis vinifera* core.

In the next step, re-sequencing of selected candidate genes will be done in the core set to determine patterns of allelic diversity, haplotype structures and linkage disequilibrium (Whitt and Buckler, 2003). Furthermore, all data will be integrated with those obtained in the Spanish and French partners groups through a common secured web page constructed within the project. Finally, we aim to publish genotyping results in a database available for application in marker assisted selection and fingerprinting of cultivars of grape species broadening the genetic basis for future grapevine breeding.

Acknowledgement

We would like to thank the Bundesministerium für Bildung und Forschung (BMBF) for the financial support through GABI. Special thanks is expressed to Patrice This (INRA, UMR Diversité et Génomes des Plantes Cultivées) for help in statistical analysis.

Tab.1: *continued*

Muscat Bailey A	53-17-012	N	W,T	I.C.	Bailey x Muscat Hamburg	*V. lab., linc.*	1,2
Olmo 3-54	53-17-020	-	-	V. rot	*V. rotundifolia* x *V. vinifera*	*V. rotund.*	1,2
Joannes Seyve 11369	40-03-019	N	W	I.C.	S. 6468 x Chancellor	*V. aest., berl., cin., linc., rip., rup.*	0,8
Seibel 7324	67-41-095	N	W	I.C.	S. 5163 x S. 880	*V. aest., cin., lab., linc., rup.*	0,8
Misket Kailachki	53-14-021	B	W, T	I.C.	M. Hamburg x Villard Blanc	*V. berl., linc., rup.*	0,8
Jacquez	53-17-002	N	W, T	I.C.	*V. aestivalis* x *V.vinifera*	*V. aest.*	0,3
Droujba	53-14-016	B	W	I.C.	Misket Kailachki x (Zarya Severa x M. Hamburg)	*V. aest., amur., berl., cin., linc., rup.*	0,3
Vincent	53-18-018	N	W	I.C.	(Lomanto x Seneca) x Chelois	*V. aest., champ., cin., lab., linc., rup.*	0,3
Excelsior	53-13-012	B	W	I.C.	(*V. rupestris* x Othello) x Herbemont	*V. aest., cin., lab., rip., rup.*	0,3
Seibel 5450	67-42-026	N	W	I.C.	S. 867 x S.4182	*V. berl., lab., linc., rip., rup.*	0,3
Eger 7	53-14-018	N	W	I.C.	Roucaneuf x Szelekiv	*V. berl., linc., rip., rup.*	0,3
Naumburg 2-76-24	53-22-011	N	W, R	I.C.	Silvaner x Couderc 16-16	*V. berl., longii, rip.*	0,3
Dog Ridge	53-16-018	N	R	I.C.	*V. rupestris* x *V. candicans*	*V. cand., rup.*	0,3
Champanel	39-38-008	N	R	I.C.	*V. champinii* x Worden	*V. champinii, lab.*	0,3
V. coignetiae Pulliat	53-14-002	N	R	ttt	ttt	*V. coig.*	0,3
Freiburg 423-51	53-22-016	N	W	I.C.	Triumph x *V. cinerea*	*V. lab., cin.*	0,3
Cayuga White	53-16-002	B	W	I.C.	Seyval x Schuyler	*V. lab., linc., rup.*	0,3
Orlando Seedless	67-42-094	B	T	I.C.	Florida D 4-167 x Florida F9-68	*V. lab., linc., simpsonii*	0,3
Couderc 2	39-32-010	N	W	I.C.	(*V. lincecumii* x *V. riparia*) x Pinot	*V. linc., rip.*	0,3

[1] N = Black, B = White, [2] R = rootstock, T= table grape, W = wine grape, [3] I.C. = Interspecific Cross, ttt = true to type, [4] different resistance sources: V. aestivalis (V. aest) , V. amurensis (V. amur.), V. berlandieri (V. berl.), V. candicans (V. cand.), V. cinerea (V. cin.), V. champinii (V. champ.), V. coignetiae (V. coig.), V. doaniana, V. labrusca (V. lab.), V. lincecumii (V. linc.), V. longii, V. riparia (V. rip.), V. rotundifolia (V. rotund.), V. rupestris (V. rup.), V. simpsonii, [5] F = frequency within the set of 257 genotypes of known parentage. For 100 genotypes of the complete set the ancestry could not be determined.

Naturally occurring nucleotide diversity in candidate genes for forest tree adaptation: magnitude, distribution and association with quantitative trait variation

DIGENFOR Trees

Barbara Vornam 1; Oliver Gailing 1; Reiner Finkeldey 1; Carmen Collada 2; M. Ángeles Guevera 2; Álvaro Soto 2; Nuria de María 2; Santiago González-Martínez 2; Luis Díaz 2; Ricardo Alia 2; Ismael Aranda 2; José Climent 2; M.Teresa Cervera 2; Pablo Goicoechea 3; Valerie Léger 4; Emmanuelle Eveno 4; Jeremy Derory 4; Pauline Garnier-Géré 4; Antoine Kremer 4; Christophe Plomion 4

1 Institute of Forest Genetics and Forest Tree Breeding, University of Goettingen, Germany
2 Genetica y Ecofisiologia Forestal, Dept. Sistemas y recursos Forestales (CIFOR-INIA) Madrid, Spain, Dept.de Biotechnologia, ETSIM (UPM), Madrid Spain, Unidad Mixta de Genómica y Ecosisiologica INIA-UPM, Madrid, Spain
3 NEIKER, Vitoria-Gazteiz, Spain
4 Equipe Génétique Forestière, UMR BIOGECO, INRA, Cestas, France

Introduction

It is difficult to overstate the importance of trees. Trees are essential components of the natural landscape and play a crucial role in global carbon budgeting, the response to global climate change and the maintenance of biodiversity. Trees also form the foundation of multibillion Euro forest products industries, including the conversion of biomass to energy. Despite their importance from both environmental and economic perspectives, little is known about the cellular mechanisms that underpin the growth and survival of trees. This is surprising, given that an understanding of these mechanisms will guide efforts aimed at ensuring the long-term maintenance of forest health, and the enhancement of forest productivity.

In this project, we concentrate on two forest tree species: *Pinus pinaster*. and *Quercus petraea*, representative of the two main genera of the European forest. *Pinus pinaster* belongs to the Pinus genus, one of the world's most important crops and one of the most important genera of trees in the world's forest. As gymnosperms, they represent one of the oldest and most successful groups of higher plants. *Quercus petraea* belongs to the genus Quercus, probably the widest distributed tree genus worldwide. Sessile oak belongs to the white oaks and has a continental distribution from Spain to the Ural mountains. The species is characterized by an extremely high genetic diversity and differentiation over its natural range especially for adaptive traits as bud burst.

Objectives

Growth, development and productivity of long-lived organisms such as forest trees are continuously challenged by abiotic stresses, and may also be greatly affected by rapid climatic changes in the near future. For forest trees the predicted increasing rate of global change bears the risk of reduced potential for long-term adaptation compared to annual plants with much higher reproduction rates. The objectives of this trinational (GÉNOPLANTE, France and MCyt, Spain) population genomics project are to identify mutations in candidate genes of adaptive significance based on nucleotide diversity pattern analysis and their validation in natural populations of the non-model forest tree species sessile oak (*Quercus petraea*) and maritime pine (*Pinus pinaster*).

Thirtyfive candidate genes were chosen for each species in order to validate functionally important SNPs (single nucleotide polymorphisms) in associating nucleotide diversity with the phenotypic variation of bud phenology in oak, drought stress response and bud phenology in pine.

The results will be germane to conservation programs, where information on the genetic control of the physiological functions of trees will aid to preserve critically important genotypes. Another important area of application for functional genomics in particular is the adaptation of trees to environmental changes of all sorts, and physiological responses to abiotic stresses.

Plant material and candidate genes

Maritime pine: Five natural populations were sampled along a rainfall gradient in central Spain. In the case of drought stress the candidate genes were chosen from EST sequences that have been generated from cDNA libraries (http://cbi.labri.fr/outils/SAM/COMPLETE/index.php) and in the case of bud phenology orthologous sequences derived from *Arabidopsis thaliana* were used. So far 27 genes or gene fragments were amplified and sequenced using the DNA from megagametophytes which represent haploid (maternal) tissue.

Sessile oak: Five natural populations were sampled along an altitudinal gradient in the French Pyrenees. Candidate genes related to bud phenology were chosen from EST sequences that have been generated from cDNA libraries (Derory *et al.* 2006). So far 12 genes or gene fragments have been amplified using the DNA from leaves. The gene fragments were directly sequenced and sequenced after cloning (three to six clones per transformation).

In both species the absence of population structures was verified by neutral SSR-markers.

Preliminary results and discussion

Maritime pine: Preliminary estimates of haplotype and nucleotide diversities were moderately high with Hd = 0,476 and π = 0,00351 on average (tab 1) and similar to the results obtained in previous functional and expressional candidate

Tab. 1: *Haplotype- and nucleotide-diversity of candidate genes related to drought stress tolerance and bud phenology in maritime pine.*

Candidate Gene	Total fragment length (bp)	Nb SNP	Nb Indels.	h Nb of haplotypes	H Haplotype diversity	Nucleotide diversity (π 10-3) per site total
AOX	483	12	0	20	0.850	5.43
APs II	300	4	0	4	0.187	0.96
GP130	285	9	0	7	0.274	1.90
GlP	170	9	0	8	0.417	3.65
SPS	427	9	0	12	0.770	3.79
ERD3	795	5	0	6	0.451	1.34
DHN2	456	9	2	7	0.664	5.32
LP3_1	411	8	2	10	0.742	7.39
LP3_3	406	8	0	7	0.813	5.21
PP2C	589	5	2	5	0.503	1.38
RD21	941	21	3	12	0.784	6.83
GS1a	714	5	0	6	0.562	2.85
GS1b	975	9	1	3	0.531	4.23
Gs1a-Prom	693	6	2	3	0.603	4.12
Gs1b-Prom	610	3	1	4	0.682	1.59
PHYO-I	575	0	0	1	0.000	0.00
PHYO-II+III	1038	0	0	1	0.000	0.00
PHYN-I+II	1111	14	0	4	0.643	4.11
PHYP	401	6	0	2	0.153	2.29
CRY1	891	3	0	4	0.432	0.53
CRY2	702	0	0	1	0.000	0.00
CCoAOMT-I+II	1730	39	5	15	0,728	3.64
GLUCAN-I+II	1404	31	4	12	0.789	9.30
PR-AGP4I+II	1550	75	5	8	0,651	7.96
C4H-I	750	7	0	4	0.737	4.99
C4H-II	650	3	0	2	0.248	1.29
GRP2-I+II	575	10	0	12	0.719	3.65
RINGP	867	13	1	8	0.795	4.62
Total or average	**20 499**	**323**	**28**	**188**	**0.476**	**3.51**

gene studies of pine (González-Martinez *et al.* 2006). We detected one SNP per 63 bp across coding and non-coding regions on average. In gene fragments of the genes *PHY* O and *CRY* 2 of 1613 bp and 702 bp length, respectively, no SNPs or indels were detected. Insertions/deletions (indels) were found in 6 genes, ranging from 2 to 18 bp. Some of the so far analysed genes have non-synonymous substitutions at a relatively high frequency (> 0.3) which may be indicative of natural selection. Based on neutrality tests these genes show also an excess of intermediate-frequency alleles, less haplo-

Tab. 2: *Haplotype- and nucleotide-diversity of candidate genes related to bud phenology in sessile oak.*

CG	Total fragment length (bp)	Nb SNP	Nb Indels.	h Nb of haplotypes	H Haplotype diversity	Nucleotide diversity (π 10-3) per site total
Cat2	336	7	1	13	0.606	3.16
PCNT 115	556	24	1	17	0.845	6.61
CHZFP	385	13	2	14	0.452	2.09
Dhn2	521	17	5	12	0.741	4.71
Dhn3	542	40	2	25	0.933	11.96
GST	477	22	0	11	0.850	8.27
Cry1	367	6	0	4	0.643	4.20
H3	373	4	1	7	0.872	3.93
DAG2	360	1	0	2	0.571	1.59
YSL1	315	2	0	5	0.750	3.02
AUX-REP	443	7	1	7	0.873	5.28
RASI	327	8	0	9	0.978	10.19
Total or average	**4981**	**151**	**13**	**126**	**0.759**	**5.42**

types as expected and an excess of haplotypes with derived mutations as typical signatures of selection. These genes are also used for association mapping based on linkage disequilibrium which offers an alternative method for mapping adaptive traits in forest trees using ancestral recombination events in natural populations in order to associate marker and phenotype. The extend of linkage disequilibrium regarding only parsimony informative sites within the *RD21*A like gene (a cysteine protease) is shown in Fig.1. *RD21*A had strong LD to the 3' end of the fragment, a result that contrasts with the moderate levels of LD found in this gene when sequences from the full geographic range of maritime pine have been analyzed. Higher LD is expected in recent populations due to founder effects during range expansion.

Sessile oak: In contrast to pine the estimates of haplotype and nucleotide diversities were relatively high in oaks with $H_d = 0.759$ and $\pi = 0.00542$ on average (tab 2) so far, and one SNP per 33 bp across coding and non-coding regions were detected on average. Insertions/deletions (indels) were found in 7 genes, ranging from 2 to 36 bp. The genes *Dhn* 2 and *Dhn* 3 had indels within the coding regions, including a 30 bp indel in *Dhn* 2 and a 36 bp indel in *Dhn* 3. The lengths of these indels were multiples of 3 bp, therefore the reading frames are not shifted. For *Dhn* 2 only a partial gene sequence was obtained, while the *Dhn* 3 gene is represented by a full length sequence of 542 bp including the 5' and 3'UTR regions. 14 SNPs were found in the coding region of the *Dhn* 3 gene, six of them led to non-synonymous substitutions. The indels found in the other analysed genes showed microsatellite-like motifs and were located within the non-coding regions.

Association studies and field evaluation of the adaptive traits

Maritime pine: Different phenotypic traits were measured in two provenance-progeny tests (5 populations represented by 20 to 30 trees per population) installed in France (good conditions) and Spain (stressful conditions). The traits under evaluation (e. g. height, diameter, ontogeny, etc.) showed significant population and family effects indicating a genetic control of this phenotypic variation.

Sessile oak: The *in situ* monitoring of temperature and bud burst was duplicated along two altitudinal transects in France (Luz and Ossau) to ensure for repeatability of the results. Bud bursting was recorded in the stands on 10 randomly sampled trees. First results showed a linear variation of bud burst as a function of altitude, and the inter- and intra-population variation is much larger than in other tree species studied so far.

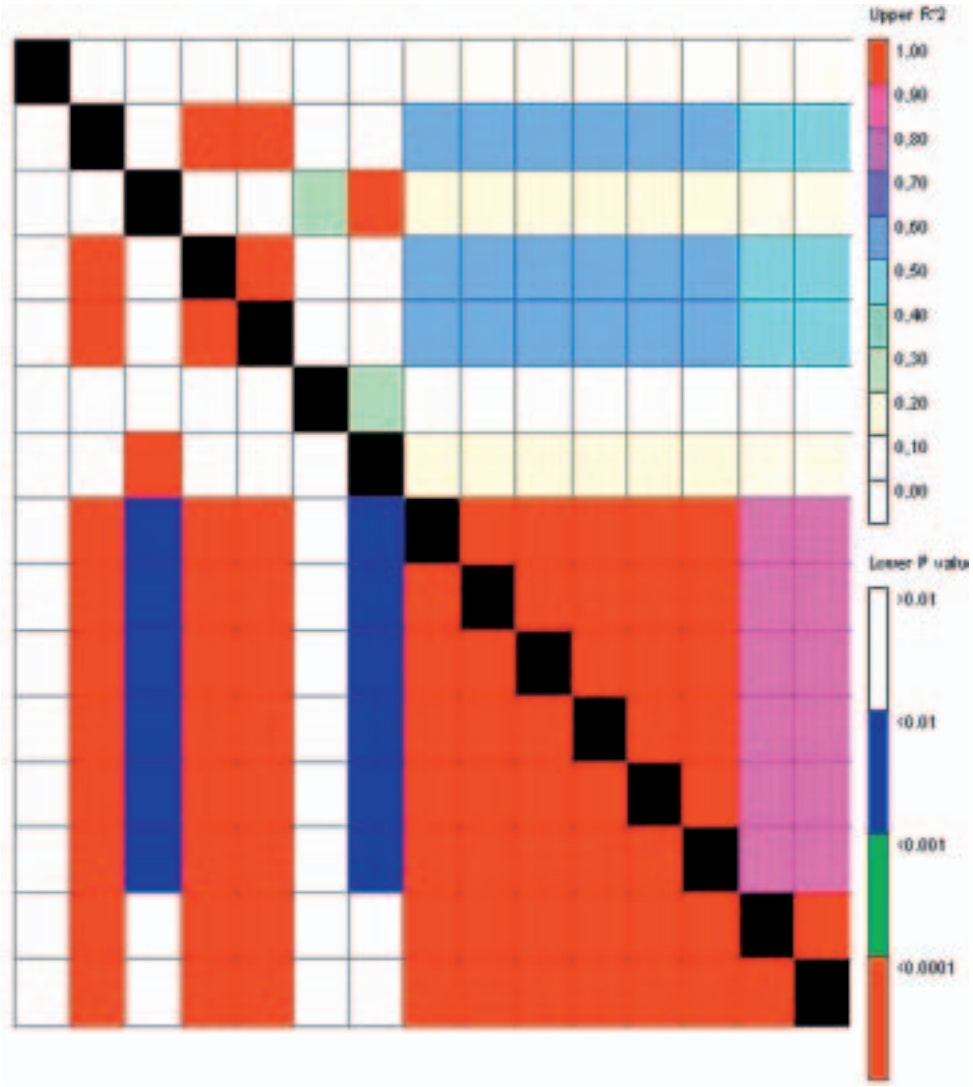


Fig. 1: *LD (linkage disequilibrium) plot within the RD21A like gene of maritime pine. Parsimony informative sites are plotted on both the X-axis and Y-axis. Pairwise calculation of LD (r^2) are displayed above the diagonal with the corresponding P-values for Fischer's exact test displayed below the diagonal. Coloration is indicative of the corresponding P-value or r^2 values from the bars on right.*

Identification and Characterization of Genes controlling quantitative Pathogen-Resistance in Potato (*Solanum tuberosum*) by a Candidate Gene Approach, PHASE 2 (CONQUEST-2)

CONQUEST
Potato

Christiane Gebhardt 1; Agim Ballvora 1; Ute Achenbach 1; Jens Lübeck 2; Josef Strahwald 2; Eckart Tacke 3; Hans-Reinhard Hofferbert 4; Birgit Kersten 5; Svenja Meyer 5; Rico Basekow 5
1 MPI für Züchtungsforschung, Köln
2 Saka-Ragis Pflanzenzucht GbR, Windeby
3 Bioplant, Ebstorf
4 BNA, Ebstorf
5 RZPD German Resource Center for Genome Research, Berlin

Introduction

The GABI-CONQUEST project has two major goals: Unraveling the molecular basis of natural variation of pathogen resistance in potato and developing DNA-based markers to be used for the selection of cultivars with superior resistance. The pathogens of primary interest are the oömycete *Phytophthora infestans* causing the late blight disease and the parasitic root cyst nematode *Globodera pallida*. During the 1st project phase (CONQUEST-1, 1999-2003), ca. 1400 SNP (single nucleotide polymorphism) and 127 InDel (Insertion-Deletion) markers were identified that are physically linked to candidate genes for pathogen resistance (genes functional in pathogen recognition, defense signaling or defense response). The markers tag most regions of the potato genome known to harbor QRL (Quantitative Resistance Locus) (Rickert *et al.* 2003). SNPs linked to a major QRL against *G. pallida* were identified and used to develop haplotype specific PCR-based marker assays. A physical map was constructed covering part of a resistance hot spot on potato chromosome V, where major QTL for resistance to both *P. infestans* and *G. pallida* are located. Ten BAC (Bacterial Artificial Chromosome) insertions were sequenced resulting in 600 kb annotated potato genomic sequence (Ballvora *et al.* 2007). The PoMaMo (Potato Maps and More) database (https://gabi.rzpd.de/projects/ Pomamo/) was initiated in collaboration with RZPD within the GABI 1 project "GABI Primary Database" (GabiPD) (Meyer *et al.* 2005). These results are the assets for the second phase of the project, which comprises (i) an association mapping experiment for resistance to *P. infestans* and *G. pallida* in two populations of 96 tetraploid breeding clones each, (ii) the functional characterization of candidate genes in the resistance hot spot on potato chromosome V and (iii) the development of database application tools for association mapping in tetraploid potato.

Materials and Methods

Association mapping of resistance to *P. infestans* and *G. pallida*.

SARA and BNA each assembled a population of 96 tetraploid breeding clones from their respective breeding programs and evaluated them in two years in replicated field trials for quantitative resistance to *P. infestans* (Figure 1) in combination with other agronomic characteristics such as maturity type. The breeding clones were also evaluated for resistance to *G. pallida* in a pot test. The same populations were genotyped with DNA-based markers known to be linked to resistance loci. SNPs were scored by sequence analysis of amplified DNA fragments. This method allowed the assessment of the allele dosage in tetraploid potato as basis for the deduction of haplotypes at a given locus. CAPS (Cleaved Amplified Polymorphic Sequence), SCAR (Sequence Characterized Amplified Region) and ASA (Allele Specific Amplification) markers linked to resistance loci (Bormann *et al.* 2004 and unpublished results) were also scored in the populations. SSR (Simple Sequence Repeat) markers was scored by allele separation on Spreadex gels (Elchrom Scientific, Cham, Switzerland).

Functional characterization of candidate genes.

Based on genomic sequence information from the resistance hot spot on potato chromosome V, (i) primers were designed for studying the expression of annotated genes by RT-PCR and (ii) four siRNA constructs were made for silencing members of the *R1* resistance gene family. Potato genotypes having QTL for late blight resistance on chromosome V were transformed with the siRNA constructs.

○ Ballvora *et al.* **Comparative sequence analysis of Solanum and Arabidopsis in a hot spot for pathogen resistance on potato chromosome V reveals a patchwork of conserved and rapidly evolving genome segments.** (2007) BMC Genomics, in press. ○ Bormann C. *et al.* **Tagging quantitative trait loci for maturity-corrected late blight resistance in tetraploid potato with PCR-based candidate gene markers.** Mol Plant Microbe Interactions (2004) 17, 1126-1138. ○ Meyer S, Nagel A, Gebhardt C. **PoMaMo – a comprehensive database for potato genome data.** Nucleic Acids Research (2005) 33, Database issue, D666-D670. ○ Rickert A. *et al.* **First-generation SNP/InDel markers tagging loci for pathogen resistance in the potato genome.** Plant Biotechnology Journal (2003) 1, 399-410. ○ Sattarzadeh A. *et al.* (2006) **Single nucleotide polymorphism (SNP) genotyping as basis for developing a PCR-based marker highly diagnostic for potato varieties with high resistance to Globodera pallida pathotype Pa2/3.** Mol Breed (2006) 18, 301-312.

Fig. 1: *Phenotypic variation in the field of resistance to P. infestans in the CONQUEST-2 populations of tetraploid potatoes.*

Results

Associations between DNA-based markers and quantitative resistance to *P. infestans* and *G. pallida*. The two populations of 96 tetraploid breeding clones each were genotyped for 223 SNP markers at 22 sequenced loci, for 76 polymorphic DNA fragments resulting from CAPS, SCAR or ASA markers at 24 loci and for 166 SSR alleles at 30 loci (Figure 2). Preliminary single marker association tests (Student's t-test, ANOVA) indicate significant ($p < 0.01$) associations between both types of resistance and all types of DNA-markers. The number of significant associations is higher than expected by chance alone. Markers diagnostic either for increased or decreased resistance were identified.

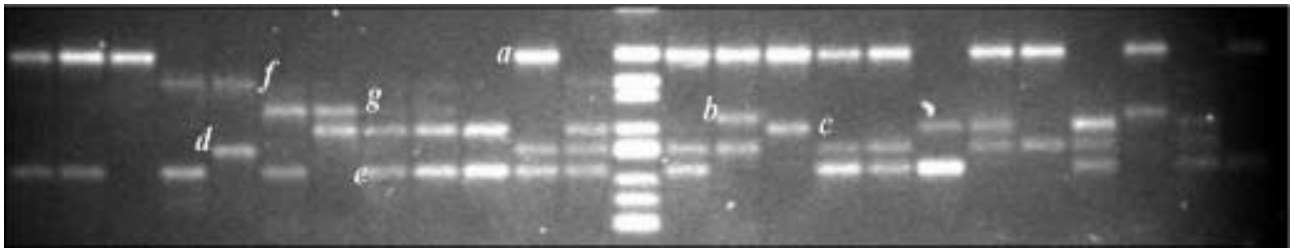


Fig. 2: *Genotypic variation in the CONQUEST-2 populations exemplified with the SSR marker STM0038. Seven SSR alleles (a to g) can be distinguished. The allele dosage in each tetraploid individual can be estimated by the band intensity.*

A highly diagnostic marker for resistance to G. pallida pathotype Pa2/3.

Thirty four potato varieties with known levels of resistance to *G. pallida* according to the variety descriptions and 21 susceptible varieties were screened by PCR with the allele specific 'HC' marker developed in CONQUEST-1. The HC marker was present in 21 of the 23 most resistant varieties and was absent in all susceptible varieties tested. The HC marker is therefore diagnostic for high resistance levels to *G. pallida* in tetraploid potato (Sattarzadeh *et al.* 2006).

Functional characterization of candidate genes.

Primers were designed for 23 open reading frames (ORFs) in the genomic sequence on chromosome V. Their expression was studied in leaves of two genotypes with different levels of quantitative resistance to late blight before and after infection with *P. infestans*. Fourteen of the ORFs were expressed at various levels in leave tissue. No clear response to infection with *P. infestans* was observed with these genes.

Ca. 200 transgenic siRNA lines were regenerated from Agrobacterium mediated transformation of one diploid and two tetraploid potato genotypes with four R1-siRNA constructs. Preliminary infection tests indicate that some siRNA lines are compromised for resistance to *P. infestans*.

The PoMaMo database.

A 'Solanaceae function map for resistance' was compiled from 177 publications on mapping and cloning resistance genes in potato, tomato, pepper and tobacco and made publicly available in the PoMaMo database in collaboration with GabiPD. The desktop-database application ConquestExplorer was developed as a structured repository for CONQUEST-2 project data. This tool facilitates the integration, visualization and analysis of phenotypic and molecular data generated in the project (Figure 3). Moreover, PotHap, the prototype of a haplotyping tool was implemented allowing for prediction of haplotypes in tetraploid species based on unphased SNP genotype data.

Discussion

The results obtained so far indicate that association mapping is a feasible and valuable approach in tetraploid potato, which can lead to DNA-based markers to be used in potato breeding. At present, the HC marker is the best marker available for diagnosis of high resistance to *G. pallida* in commercial varieties and breeding clones. The large number of marker/trait associations found with candidate genes but also with 'non-candidate' markers, suggests that linkage disequilibrium is extensive in the genetic material analysed. The functional analysis of candidate genes did not give conclusive results until now.

Fig. 3: *Integration, administration, visualization and analysis of CONQUEST-2 project data in the ConquestExplorer application. (A) Central user interface. (B) Molecular maps (C) Tracefile viewer with integrated SNP visualization. (D) Statistic tests for phenotype/genotype associations. (E) Result interface of the haplotyping tool PotHap.*

Genomics of Chilling Tolerance in Maize

COOL
Maize

Milena Ouzunova 1; Thomas Presterl 1; Walter Schmidt 1; Günter Strittmatter 1; Karin Ernst 2; Nino Baliashvilli 2; Peter Westhoff 2; Evelyn M. Möller 3; Katinka Wilde 3; Hartwig H. Geiger 3; Fabien Poree 4; Oliver Bläsing 4; Mark Stitt 4
1 KWS SAAT AG
2 Heinrich-Heine-Universität, Düsseldorf
3 Universität Hohenheim, Stuttgart
4 MPI für Molekulare Pflanzenphysiologie, Golm

Scientific background and objectives

Due to its tropical origin maize is a chilling-sensitive species. To date plant breeding has been successful in adapting maize to the cool, temperate regions of Europe by relying mostly on a chilling escape strategy. The present project used a chilling sensitive and a tolerant line (Figure 1) for generation of a doubled haploid mapping population and pursues a genomics-based genetical strategy combined with exploitation of physical mapping information to identify and map, as precise as possible, genomic regions that confer chilling tolerance to maize. The cloning of a quantitative trait locus (QTL) for chilling tolerance at the molecular level is envisaged and based on work commenced in GABI phase I.

Fine-mapping of chilling tolerance QTL using near-isogenic lines and marker development

The construction of a set of overlapping near-isogenic lines (NILs) has enabled precise fine-mapping of chilling tolerance QTL, which were identified in the segregating doubled-haploid line (DHL) population established and analysed in GABI phase I (Figure 2). First priority was given to a QTL on chromosome 4 explaining 34% of the phenotypic variation for the line performance in plant fresh weight under chilling conditions and to a QTL on chromosome 5 that determines 14% of this trait. Second priority was on a QTL on chromosome 6 which explains 8% of the grain yield and of the fresh weight performance of the hybrids. The molecular markers developed for chilling tolerance associated genes and for fine mapping of the QTL will be useful for breeding of chilling tolerant maize lines via marker-assisted selection. In comparison to marker development in GABI phase I the PCR success rate increased remarkably using genomic, coding, and repeat-masked sequences for primer design. Such sequences were provided by MIPS (Munich Information Center for Protein Sequences) based on BAC-end sequences anchored on the preliminary BAC-contig fingerprint map which allowed a targeted marker development for the regions of interest on and the establishment of an integrated genetic and physical map. The overall polymorphisms frequency based on approx. 900 loci evaluated was 45%. Totally, 61 gene-based markers were mapped across the genome, thereof 11, 12, and 10 located on chromosomes 4, 5, and 6 respectively. The average marker distances for these three chromsomes are 1.5 cM, 1.1 cM, and 1.6 cM respectively. The expected co-segregation of 20 differentially expressed genes (detected by SSH) and QTL does not exist. Of the remaining genes retrieved from literature and databases 6 resp. 5, and 1 are located within the QTL regions on chromosomes 4, 5, and 6.

Genetically oriented transcriptome analysis

Because the evaluation of NILs in the field is a time-consuming process several tests were performed to identify a simple phenotypic marker for chilling tolerance under growth-chamber conditions. For a QTL on chromosome 4, such a marker was found: The Cold-sensitive SL line and the chilling-tolerant TH line differ in the development of chlorotic spots during several days of recovery after a seven-day chilling stress. The presence of an apparent chlorotic phenotype in certain parts of younger leaves showed a clear correlation with chilling sensitivity (low fresh matter yield in the field) when different NILs of chromosome 4 were analysed. Hence this phenotype makes it possible to screen 30 lines for chilling tolerance associated with the QTL on chromosome 4 within four weeks. The region responsible for this phenotype has been restricted to 6 cM, and the development of additional markers will reduce this distance further.

During the first phase of GABI the SSH method had been used to identify genes that are induced or repressed by chilling treatment in the tolerant TH line. Subsequently the development of NILs enabled us to use this method for identifying genes which map to, or whose expression is regulated by a QTL. In this attempt material of the sensitive parent SL and certain NILs was used. By this approach the SSH products were enriched for sequences differentially expressed between tolerant and sensitive lines. SSH products were cloned, sequenced and arrayed on nylon membranes. By hybridisation with cDNAs made from chilling-treated tolerant and sensitive maize lines, genes were identified that were differentially expressed. Expression patterns were confirmed by northernblot hybridisation. Candidate genes whose differential expression could be confirmed were mapped.

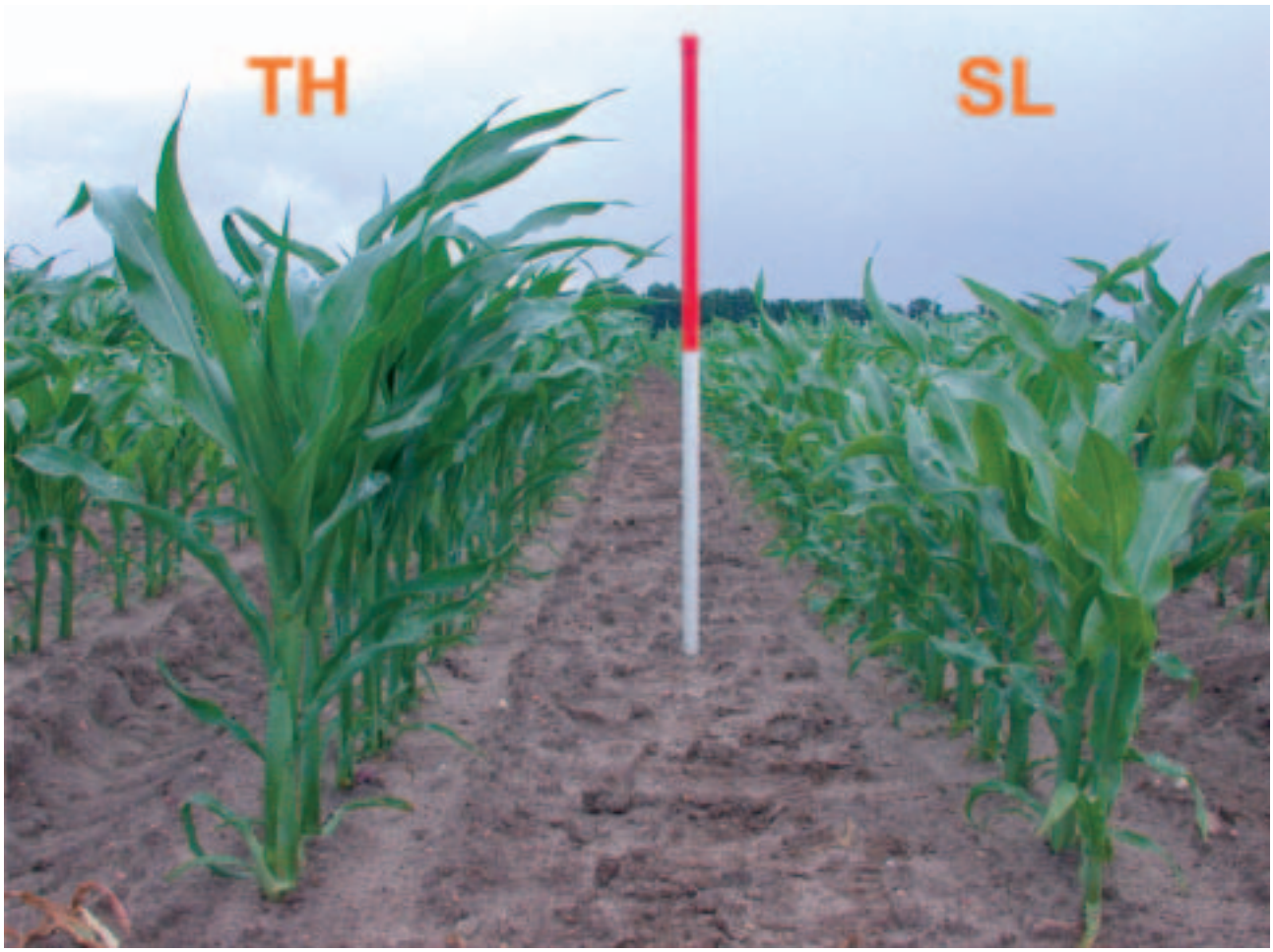


Fig. 1: *Clear growths differences height between the chilling tolerant parent TH (left) and sensitive parent SL (right).*

In the meantime a commercial BAC library became available, and BAC contigs that cover the QTL donor segments are now going to form the basis for a positional cloning approach. The QTL on chromosome 5 (about 10 cM) can be covered by a nearly complete BAC contig of 96 BACs. 160 BACs covering the QTL were selected for a cloning approach. At time different methods are tested to identify non-repetitive regions in these BACs for sequencing. The data obtained will be used to develop additional markers. In another approach the BACs are used to identify coding regions.

More than 1000 BACs arranged in several small contigs are necessary to cover the 6cM QTL of chromosome 4. This number is too large for a simple positional cloning approach. Hence only hundred BACs distributed over the QTL were selected for marker development.

Metabolic fingerprinting

Metabolic fingerprinting was carried with leaf, stem, and root tissues to identify the tissue type(s) and process(es) most affected by chilling. With a simplified experimental systems for

studying responses to low temperature in maize (analogous to 'gauntlets' already developed in *Arabidopsis*), and existing *Arabidopsis* full-genome transcript profiles that provide information about the response of this chilling-tolerant species to a fall in temperature were evaluated to support analysis of the response of maize to similar shifts, and prioritize candidate genes. Both SL and TH genotypes were grown under standard conditions (14/10h light, 25/22°C) for two weeks, and then transferred at the end of the night to chilling conditions (10/8°C). 4th and 5th leaves were sampled after 4hours and 1 week after the application of chilling treatment. After one week of chilling temperatures, an increase of fructose in the 4th leaf and of sucrose and starch in the 5th leaf of the chilling sensitive parent SL could be observed, when compared to the answer in the chilling resistant parent TH (Fig. 3). Other tissues and metabolites evaluated (such as nitrate, chlorophyll a, protein content, G6P, glucose, amino acids) did not display significant changes in response to chilling treatment between both parents.

Breeding maize lines with superior chilling tolerance

A second-cycle DHL population was developed from a selected fraction of the GABI mapping population and field tested in eight environments for chilling tolerance. An increase of plant fresh-matter yield with increase in number of positive QTL alleles was found, indicating the usefulness of the QTL for marker assisted selection for chilling tolerance. Furthermore, NILs carrying donor segments of tropical highland maize genotypes were analyzed with molecular markers and tested for chilling tolerance in field experiments. Donor chromosome segments identified in these NILs confirmed the most important QTL for chilling tolerance found in the GABI population, but also novel QTL alleles were detected in the exotic germplasm.

Summary and outlook

All milestones for NIL development and high resolution mapping of the QTL of interest in the DH line mapping population were achieved. Overall, the chosen strategies and the use of *in silico* generated and physically anchored coding and repeat-masked genomic sequences turned to be very beneficial for PCR success and hence marker development. As the co-localization of diffentially expressed genes and QTL for chilling tolerance does not exist the future main focus is set on map-based cloning of genes within the QTL region by using the publicly available BACs and information of the integrated genetic GABI map and the physical BAC-contig fingerprint map.

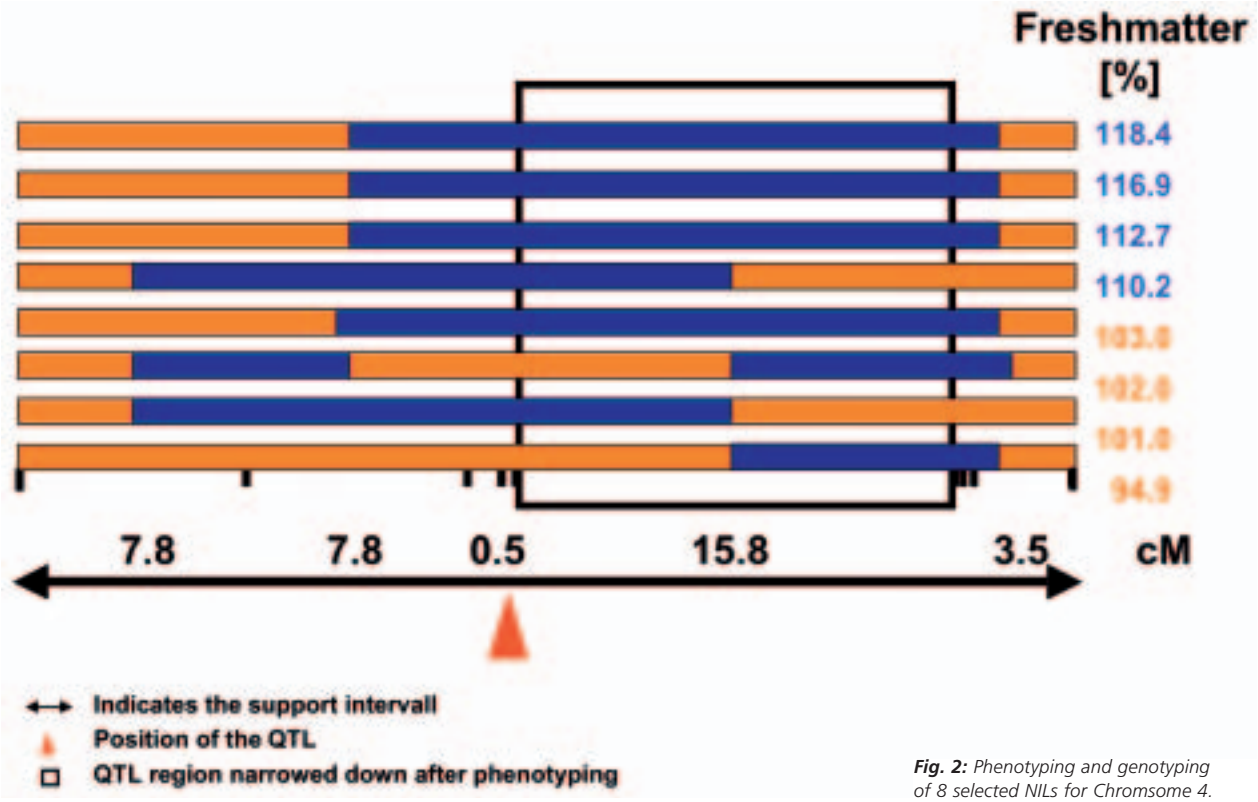


Fig. 2: *Phenotyping and genotyping of 8 selected NILs for Chromsome 4.*

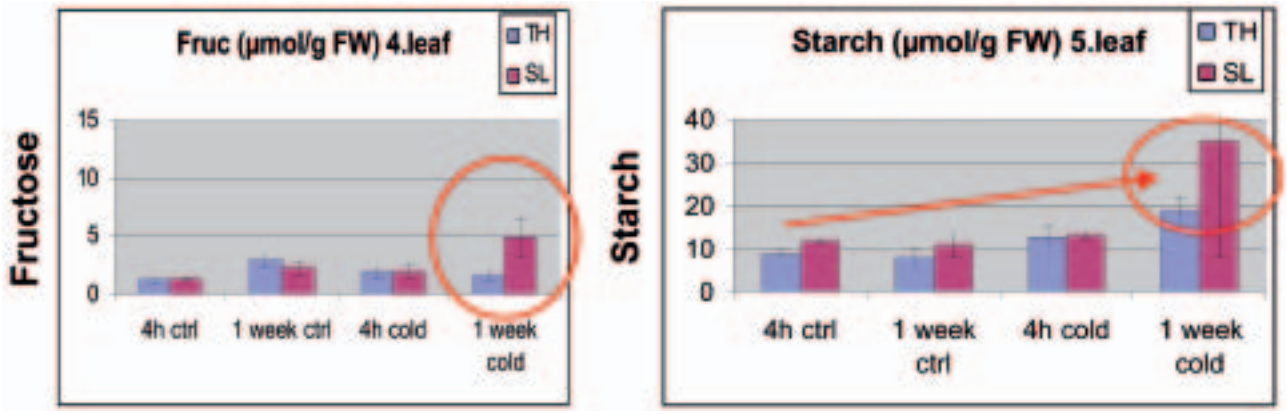


Fig. 3: *TH (blue) and SL (red) genotypes were grown for two weeks under standard conditions, 14/10 light, 25/22°C, and then transferred at the end of the night period to chilling conditions, 10/8°C. Fourth and fifth leaves were sampled 4 hours and 1 week after the transfer. Metabolites were extracted, fructose (left) and starch (right) contents measured (Geigenberger et al 1996, Hendrick et al, 2003).*

Quality and Yield

Improving crop value will be the major future task of applied plant research. Farmers will benefit from improved quantitative traits like thousand-grain weight or gross biomass production. On the other hand high-quality crops and vegetable with improved properties including optimised dietary attributes, better flavour and health benefits are the goals of modern applied plant genomics.

Go to Content Page

Go to Content Page

GABI-EVAST: EValuation of the power of ASsociation Tests versus QTL mapping

EVAST
Arabidopsis

Thomas Altmann 1; Britta Schulz 2; Mark Stitt 3; Ronan Sulpice 3; Maria v.Korff 1; Hanna Witucka-Wall 1
1 University of Potsdam, 14476 Potsdam-Golm
2 KWS Saat AG, 37574 Einbeck
3 Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam-Golm

Central carbohydrate and nitrogen metabolism is of fundamental importance for plant productivity, both, for vegetative and generative growth and for the accumulation of storage compounds in harvest organs. While the underlying physiology of sugar metabolism and starch accumulation, the genes required for the pathways, and the properties of the encoded proteins is one of the best understood areas of plant metabolism, the factors responsible for its higher level regulation especially in relation to the control of growth and the maturation of storage organs are largely unknown. C/N interactions are more complex and less well understood: while considerable background knowledge is available about the pathways, enzymes, and the genes that encode them, less is known about the regulation of individual enzymes and about globally acting regulation networks that will modulate these complex interactions.

Most previous approaches to understanding the regulation of metabolism have either analysed the response to changes in the environmental, or the response to changes in the expression of genes involved in the process. In GABI-EVAST, we use natural variation as a starting point. In the past, the use of natural variation was rather descriptive; differences might be found, but it was hardly possible to determine if they were causal. This is changing because (i) high throughput profiling methods allow far more phenotypic parameters to be analysed in (ii) a large number of genotypes and (iii) dramatic advances in genotyping technologies are making it possible for differences to be linked to discrete genomics regions and ultimately indivudal genes.

Carbohydrate and nitrogen metabolism are addressed in the GABI-EVAST project by a combined application of cutting edge genotyping tools with the use of a multilevel phenotyping platform, which allows parallel measurements of metabolic traits and the activities of a large number of enzymes. Using genotype and phenotype data on a collection of Arabidopsis natural accession-derived lines, recombinant inbred lines (RILs), and introgression lines (ILs), association tests are performed and QTL mapping is carried out using candidate cardinal genes and molecular markers. Thus, approx. 130 natural accession-derived lines (genotyped with 261 SNP-based markers), 100 lines of which are subjected to genome tiling chip hybridisation for identification of single feature polymorphisms (SFPs), 69 genetically characterised Col-0/C24 ILs, a population of 160 genotyped RILs, and a collection of 175 RILs derived from crosses among 22 different parental lines and genotyped with 149 SNP-based markers are analysed in this project for accumulation of vegetative biomass, for levels of sugar, starch, amino acids, organic acids, protein, and chlorophyll, and for up to 40 enzyme activities. Using the obtained results, the viability of genome-wide (linkage disequilibrium-based) association testing using molecular markers or candidate gene-based association tests is analysed and compared to the power of `classical´ QTL mapping using RILs or ILs in the reference species *Arabidopsis thaliana*. Furthermore, genomic regions are identified, which harbour genes with functions relevant to the mediation or the control of plant primary metabolism and growth. In parallel, the approaches and techniques developed in Arabidopsis for metabolic analyses are tested in a 'bridge' project in sugar beet breeding lines to study a set of metabolic traits, which are of central economic importance: the sucrose level, and the ratio of sucrose/reducing sugars and the C/N ratio in beets. Furthermore, candidate cardinal sugar beet genes are tested for sequence variation.

Specific results

An association mapping population of 130 diverse Arabidopsis accessions has been assembled and genotyped with a previously (GABI1) developed set of 112 equally spaced SNP-based markers. The population includes the French CC24 and the 20 accessions re-sequenced through the Weigel/Ecker/Perlegen project, it has been filled to include accessions from as diverse as possible geographic sites and climatic conditions at collection site, and has been maximized for genetic diversity and optimized for allele frequency using genotype data of c. 380 accessions, using the set of 112 equally-spaced SNP markers. Using the genotype data, the population has been shown to exhibit little structure (Fig. 1) and thus to be suitable for efficient detection of marker-trait associations. The population has recently been genotyped with another set of 149 SNP-markers and initial association tests have been performed using the collected phenotype data that yielded numerous significant marker-trait associations. In addition, the methodology of genomic DNA hybridization to genome tiling chips (according to Borevitz *et al.*, 2003) has

○ Borevitz, J.O., Liang, D., Plouffe, D., Chang, H.-S., Zhu, T., Weigel, D., Berry, C.C., Winzeler, E., Chory, J. (2003) **Large-scale identification of single-feature polymorphisms in complex genomes.** Genome Res. 13: 513-523. ○ Cross, J.M., von Korff, M., Altmann, T., Bartetzko, L., Sulpice, R., Gibon, Y., Palacios, N., Stitt, M. (2006) **Variation of enzyme activities and metabolite levels in 24 *Arabidopsis thaliana* accessions growing in carbon-limited conditions.** Plant Physiol., in press, DOI:10.1104/pp.106.086629.

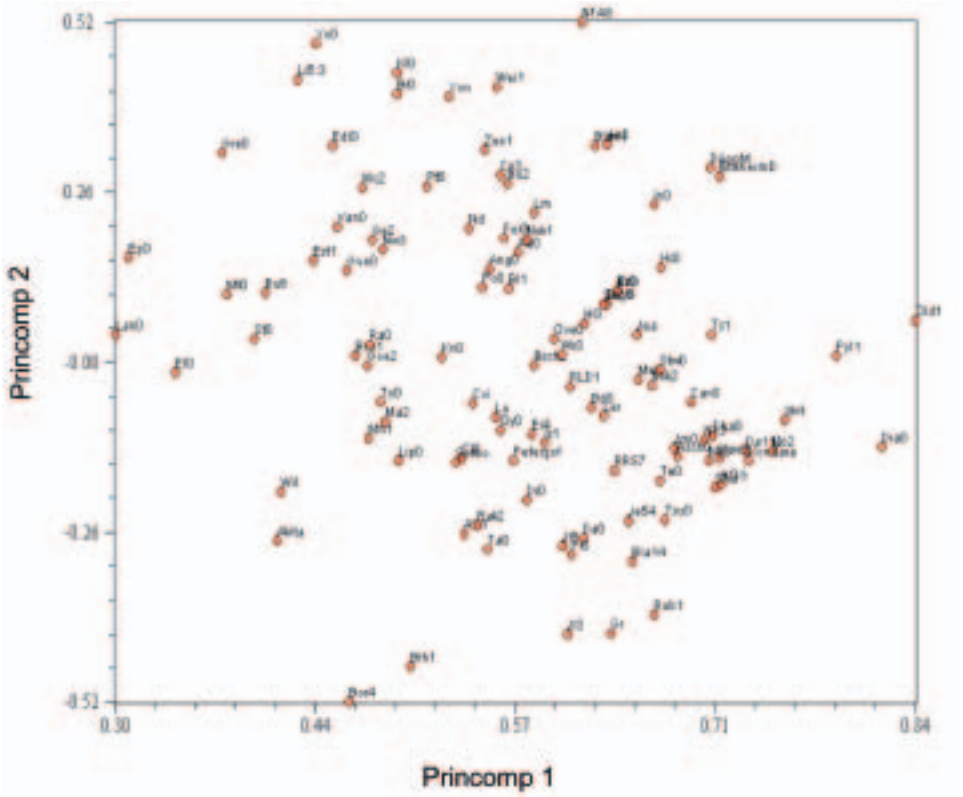


Fig. 1: *Analysis of the structure of the association mapping population. The collection of lines has been compiled from natural accessions of diverse origin and includes sets of lines that have been maximized for genetic diversity. Genetic relations were analysed using genotype data based on 112 markers subjected to principle component analysis. The distribution of the lines according to principal components 1 and 2 are shown.*

been optimised. 100 accessions have been selected to exhibit maximum phenotypic diversity and are in the process of being genotyped by genomic DNA hybridization to genome tiling chips covering 5.6 million 25mer oligonucleotides. The extracted hybridization data should provide at least 4-6000 polymorphisms for each accession in comparison to the Col-0 reference (Fig. 2).

The 130 lines of the association mapping population have been phenotyped for growth and a set of metabolic parameters in multiple replicated experiments. Plants were grown in short days and low light (C-limiting) growth conditions and analysed for growth (rosette weight [Fig. 3], leaf area, growth in the week before harvest), central metabolites that are major resources for growth (starch, sucrose, glucose, fructose, amino acids, malate, fumarate), major structural components (protein, chlorophyll) and >40 enzyme activities in central C and N metabolism. These analyses are currently being extended by carrying out GC-TOF metabolite profiling. Emerging results are: Most of the enzymes show a positive relation and some a significant correlation (R = + 0.3 to +0.5) and metabolites (especially starch and amino acids) a weak negative trend with growth (R = -0.2 to -0.5) (Fig. 4). In faster growing lines, less carbohydrate and amino acids are left at the end of each 24 h light/dark cycle. This indicates that one key component of the 'fast growth' character is increased activities of enzymes in some pathways of central metabolism, which drive faster fluxes (Cross *et al.*, 2006).

Initial association tests were performed using a simple ANOVA model and 56 traits (with genetic variance >25% of total variance) and 142 marker genotypes (with >25 lines per genotype class) yielded 110 marker effects that were significant at FDR < 0.01, and 40 that were significant even at Bonferoni corrected $p < 0.05$. In several cases, QTLs for growth, enzymes or metabolites co-associated, which may in-

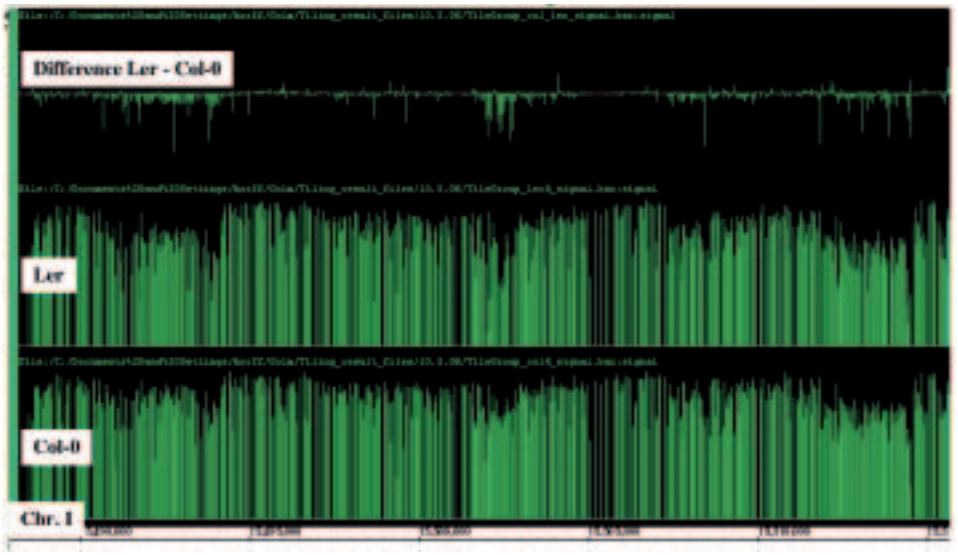


Fig. 2: *Single feature signals obtained by DNA hybridisation to oligo nucleotide GeneChips (genome tiling arrays) that cover the Arabidopsis genome at a spacing of 10 nucleotides. Hybridization signals of a representative 25 kb region of chromosome 1 are shown for DNA of the Arabidopsis accessions Col-0 (bottom) and Ler (center) and the signal differences (Ler – Col-0; top). Negative values indicate lower hybridization signals for the Ler DNA and indicate DNA-Sequence polymorphisms.*

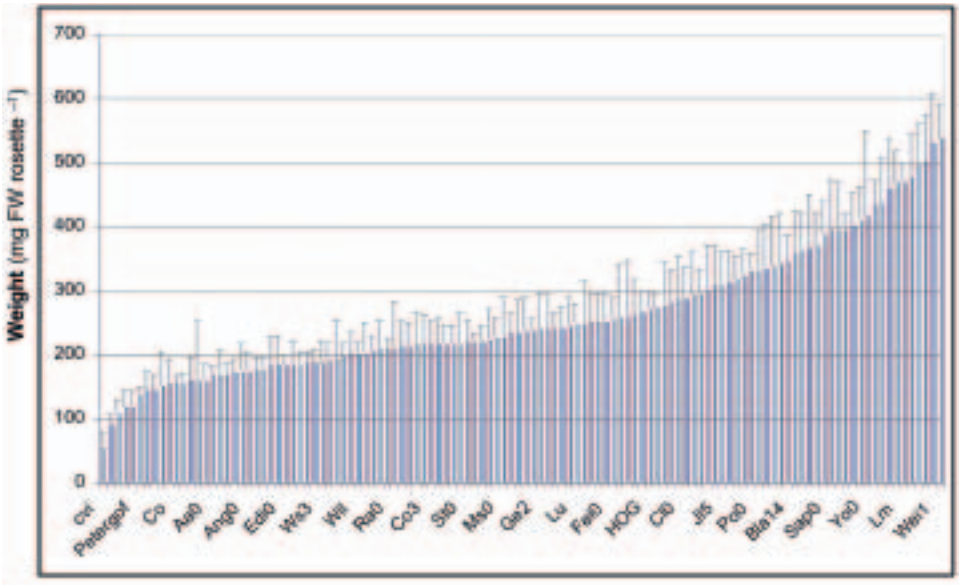


Fig. 3: *Variation in size (biomass) of 113 natural Arabidopsis thaliana accessions upon growth at controlled conditions. Mean fresh weight per plant is shown (rosette; +/- SE) after 5 week cultivation under short day conditions (8h light / 16h darkness).*

dicate that they identify genes with higher order regulatory roles. These data show that genome-wide association studies are feasible using this population and that highly informative results can be expected upon expansion to the >5000 marker data of the tiling chip hybridizations. The results (identified responsible genomic regions) will be used to explore the Weigel/Ecker/Perlegen data set from re-sequencing of 20 accessions to identify haplotypes in the gene families that encode the measured enzymes. The obtained information will be combined with marker distribution to predict haplotype distribution through the 100 accessions.

The aforementioned complex population of 175 RILs derived from crosses involving 22 different parental lines has been cultivated in two experiments each with 4 replica of 4 pooled plants and will be subjected to the same phenotypic characterization (growth, metabolic traits, and some selected

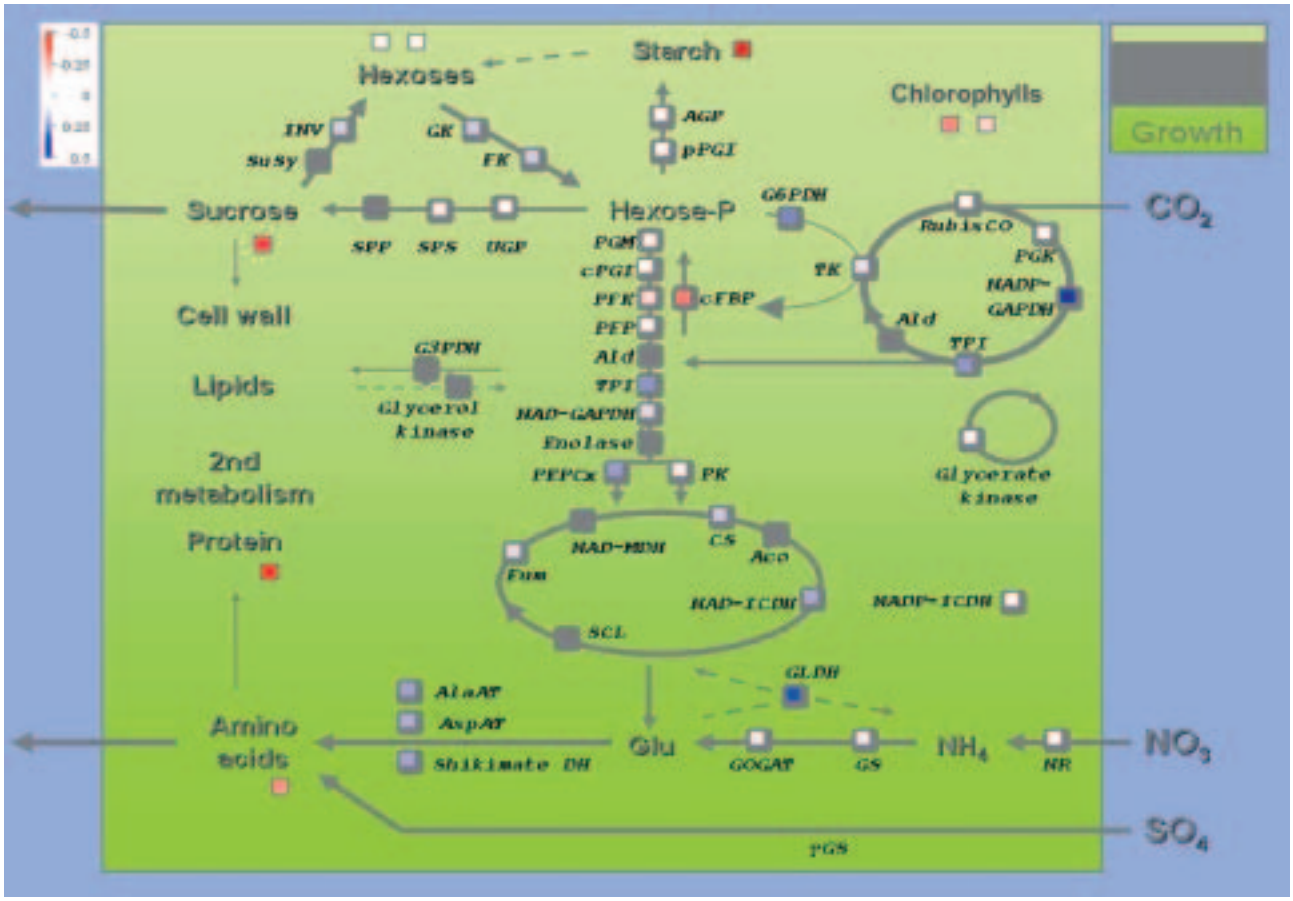


Fig. 4: *Pair wise correlations between growth (fresh weight biomass), several metabolites (sugars, amino acids), chlorophyll, protein, and 34 enzymes of primary carbon and nitrogen metabolism observed in 130 Arabidopsis accessions. Positive and negative correlations are indicated by increasingly deep blue and red shading, respectively.*

enzymes activities) as the association mapping population. Likewise, the set of 69 ILs with single or dual C24 genome segment introgressions into Col-0 has been grown and phenotypically characterized for biomass. In addition, pilot studies have been carried out to investigate the prospects of applying metabolite and enzyme profiling to growing sugar beet taproots. Assays were optimsed for use with this tissue, and two different sampling strategies tested for harvesting beets from the field. This made clear that large numbers of samples from individual beets are required to provide the necessary statistical power to distinguish between genotypes. The further application of this approach in sugar beet will depend on the development of techniques to allow rapid and facile sampling and processing of representative tissue from large numbers of individual beets.

ARAMEMNON – a database of plant membrane proteins

ARAMEMNON
Arabidopsis

Rainer Schwacke, Ulf-Ingo Flügge
University of Cologne, Botanical Institute, Cologne

Introduction

Membrane proteins play critical roles in all living cells. Although these proteins are estimated to constitute some 25% of all open reading frames in fully sequenced genomes, only a few of them have been characterized in more detail due to difficulties with experimental techniques. While predictions of properties like the number and position of transmembrane spans are feasible by using computational methods, often the results of different prediction methods are also different. The ARAMEMNON database has been developed to compile the many predicted data for plant membrane proteins from several sources and to make them comparable. The comparison should not be restricted to predictions of a single protein but also include predictions for related protein sequences.

Features of the database

The current ARAMEMNON database release 4.0 contains information about all putative membrane proteins of the nuclear and organellar genomes of *Arabidopsis thaliana* and *Oryza sativa*, and 4800 membrane proteins of more than 200 other seed plant species. The data included in the database could be grouped by four major topics (Fig.1).

Description

- manually curated protein descriptions for more than 3500 proteins
- bibliographic references for more than 2000 proteins
- transporter classification according to the Transport Classification (TC) system

Fig. 1: *Example result web page after searching for sulfate transporters. Result data are separately selectable and grouped by the four major topics description, sequence, topology and relationships (homologues).*

○ A.P. Weber, R. Schwacke, U.I. Flügge **Solute transporters of the plastid envelope membrane** Annu. Rev. Plant Biol. (2005) 56, 133-164. ○ R. Schwacke, U.I. Flügge, R. Kunze **Plant membrane proteome databases** Plant Physiol. Biochem. (2004) 42 (12), 1023-1034. ○ R. Schwacke *et al.* **ARAMEMNON, a novel database for Arabidopsis integral membrane proteins** Plant Physiol. (2003) 131(1), 16-26.

- ontology terms according to the Gene Ontology (GO) Consortium
- conserved protein sequence motifs as defined by the PFAM database

Sequence

- data about the protein, cDNA and DNA sequence and about the exon positions
- different cDNA / protein models including splice variants (*A. thaliana* and *O. sativa*)

Topology

- 16 individual predictions and two consensus predictions for the transmembrane alpha-helices (Fig.2)
- 13 individual predictions and a consensus prediction for the subcellular location

Relationships

- data about relationships to paralogous and orthologous plant membrane proteins
- simple phylogenetic tree view of paralogously related proteins
- multiple transmembrane alpha-helices alignments and sequence alignments of related proteins

Conclusion

Since its first publication in autumn 2002 ARAMEMNON has become an established and comprehensive resource for plant membrane protein data. The database is used by mostly non-commercial national and international research institutions. ARAMEMNON is publicly available at http://aramemnon.botanik.uni-koeln.de.

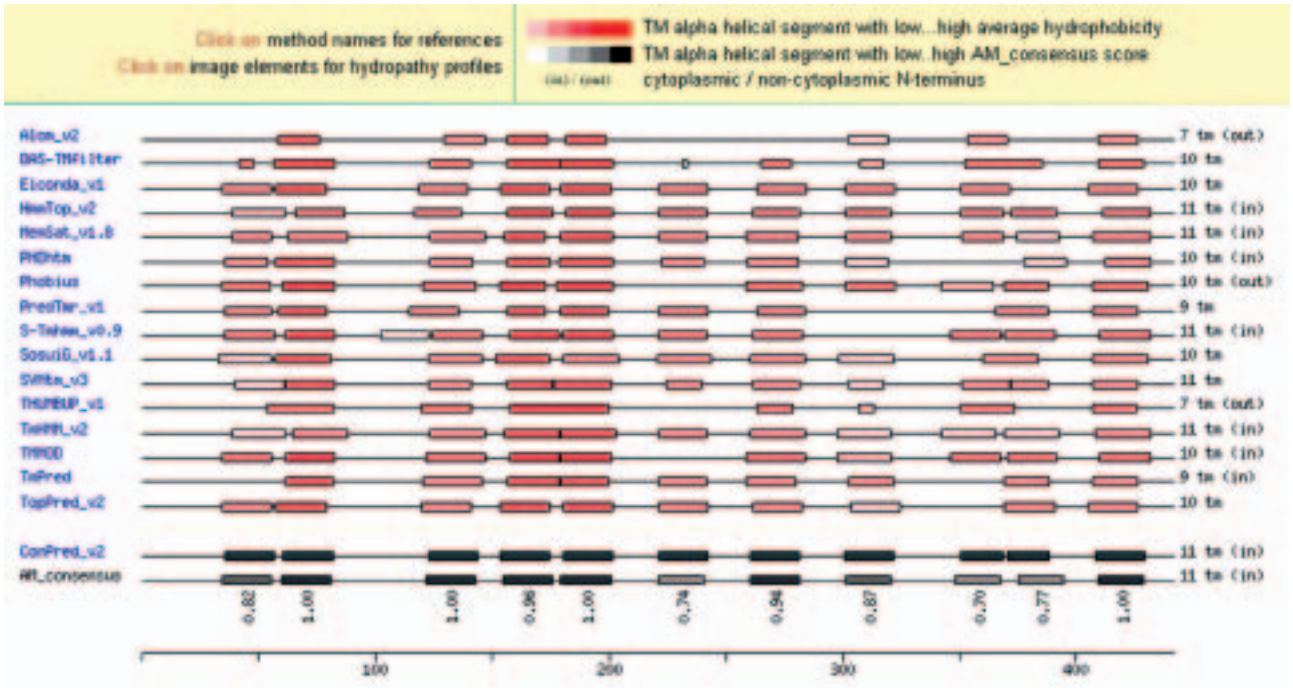


Fig. 2: *Comparison of 16 individual predictions for alpha helix transmembrane regions. In addition results from two consensus predictions are shown.*

Identification of inherent factors determining seed quality, germination efficiency and early seedling vigour of sugar beet by transcriptome and proteome profiling

SUGAR-BEET-SEED
Sugar Beet

Partner 1: Juliane Meinhard 1; Andreas Menze 1; Uwe Fischer 1*
Partner 2: Julie Catusse 2; Claudette Job 2; Dominique Job 2*
Partner 3: Elena Pestsova 3, Peter Westhoff 3
1 KWS SAAT AG, Grimsehlstr. 31, 37555 Einbeck, Germany
2 CNRS/Bayer CropScience joint laboratory (UMR 2847), Bayer CropScience, 14-20 rue Pierre Baizet, 69263 Lyon, France
3 Institut für Entwicklungs und Molekularbiologie der Pflanzen, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany
* Project coordinators

Introduction

A major concern of seed companies is the production of seeds of high quality, particularily in term of germinative quality. Losses of germination vigour can be encountered either at the time of seed production, their conditioning for the production and/or their storage. The availability of molecular and/or biochemical markers of seed vigour would make it possible to better characterize seed material in order to optimize protocols for seed production. To discover such markers, global approaches of gene expression profiling (transcriptomics, proteomics) are applied to seeds of sugar beet, a crop species of major agronomic interest. Up to the present the knowledge about gene expression profiles and protein patterns related to the sugar beet seed and the germination process is rather limited.

Within this joint GABI-Génoplante project partner 1 (KWS SAAT AG, Einbeck, Germany) provides seeds of selected batches of sugar beet, partner 2 (CNRS/Bayer CropScience joint laboratory, UMR 2847, Lyon, France) is in charge of the proteome and partner 3 (Heinrich-Heine-Universität Düsseldorf, Germany) of the corresponding transcriptome analysis.

Material and methods

Seed samples

Fruits of a triploid monogerm sugar beet (*Beta vulgaris* L.) seed lot were produced in Italy in 2002 and processed according to KWS commercial standards. Germination characteristics were evaluated and samples corresponding to early (T1, 1% germinated seeds), medium (T50, 50% germinated seeds) and late germination stages (T95, 95% germinated seeds) were used for further analyses. Similar samples were generated also for seeds that experienced a priming or ageing treatment (or a combination of both) in order to improve or reduce the corresponding germination vigour.

Transcriptomics

Four cDNA libraries each consisting of approximately 20,000 clones were developed from dry (T0) and germinating seeds (T1, T50, T95). ESTs were obtained by non-random sequencing of 5'-ends of cDNA clones belonging to the different libraries. Macroarray filters containing PCR fragments of the 2,789 cDNA clones were developed commercially (RZPD, Berlin). Array hybridizations were performed at least twice per each developmental time point using different biological replicates of plant material. The results were evaluated by AIDA (Raytest), Excel and Genesis software.

Proteomics

The proteome approach is based on previous research work of the project partner on Arabidopsis seeds (Job *et al.*, 2005). A sequential extraction protocol was optimized for sugar beet seed to enable the separation of different protein classes according to their solubilization characteristics: albumins are soluble in H2O; globulins in salt solutions; prolamins in 90% EtOH; and glutelins in acidic/basic solutions. Protein quantifications as well as SDS-PAGE experiments were done according to standard procedures. The conditions for the 2D gel electrophoresis, the protocols for protein staining and gel analyses were described previously (Job *et al.* 2005). The quantita-

Cutoff of differential expression	Number of differentially expressed unigenes
≥ 2 fold	1,042 (46 %)
≥ 3 fold	656 (29 %)
≥ 5 fold	361 (16%)
≥ 10 fold	202 (9 %)
Total number of cDNAs/unigenes on array	2,251 (100%)

Tab. 1: *Number of differentially expressed genes under different cutoff value*

○ D. Bellin *et al.* **EST sequencing, annotation and macroarray transcriptome analysis identify preferentially root-expressed genes in sugar beet.** Plant Biol (2002) 4: 700-710. ○ B.G. De los Reyes, S.J. Myers, J.M. McGrath **Differential induction of glyoxylate cycle enzymes by stress as a marker for seedling vigor in sugar beet (*Beta vulgaris*).** MGG (2003) 269: 692-698. ○ C. Job *et al.* **Patterns of protein carbonylation in Arabidopsis seeds and during germination.** Plant Physiol (2005) 138: 790-802.

tive determination of spot volumes was performed from gels carried out in triplicate. Data were subjected to statistical analysis by one-way analysis of variance. Where F values indicated significance ($P<0.05$), individual means were compared using Student's t test (alpha = 5%).

Results and discussion

Expressed Sequence Tags (ESTs) from dry and germinating seeds

From four representative cDNA libraries a set of 2.789 seed specific ESTs was developed and subjected to assembly analysis with available sugar beet ESTs. 2013 ESTs were found to be members of contigs while 776 singletons represent newly described sugar beet genes. Overall a set of 2.251 unigenes was derived, annotated using public databases and grouped into functional categories. Most of the publically available sugar beet ESTs originated so far from leaf, root or inflorescence specific DNA libraries (Bellin *et al.*, 2002) while ESTs derived from seeds are seldom (de los Reyes *et al.*, 2003).

Gene expression profiling of germination

After evaluation of the array results 1.042 unigenes were considered to be differentially expressed during germination (Tab.1). By hierarchical clustering genes were divided into a first cluster comprising 399 down-regulated genes (subclusters A-E) and a second one including 363 up-regulated genes (subclusters H-J) as well as 280 genes that show an 'intermediate' expression pattern, i.e. being up-regulated at the early and down regulated at the later stages of germination or *vice versa* (subclusters F-G) (Fig. 1A). Separate functional analysis for up- and down-regulated genes was performed by using the gene ontology classification (Fig. 1B). The cluster of up-regulated genes reflects an overall activation of the metabolism during imbibition and germination of seeds. The cluster of down-regulated genes contains a lot of genes coding for seed storage proteins, LEAs and dehydrins. For several genes of interest (i.e. cys-peroxiredoxin, α-amylase and isocitrate lyase) the data obtained by macroarray analysis were verified by RNA dot-blot hybridization.

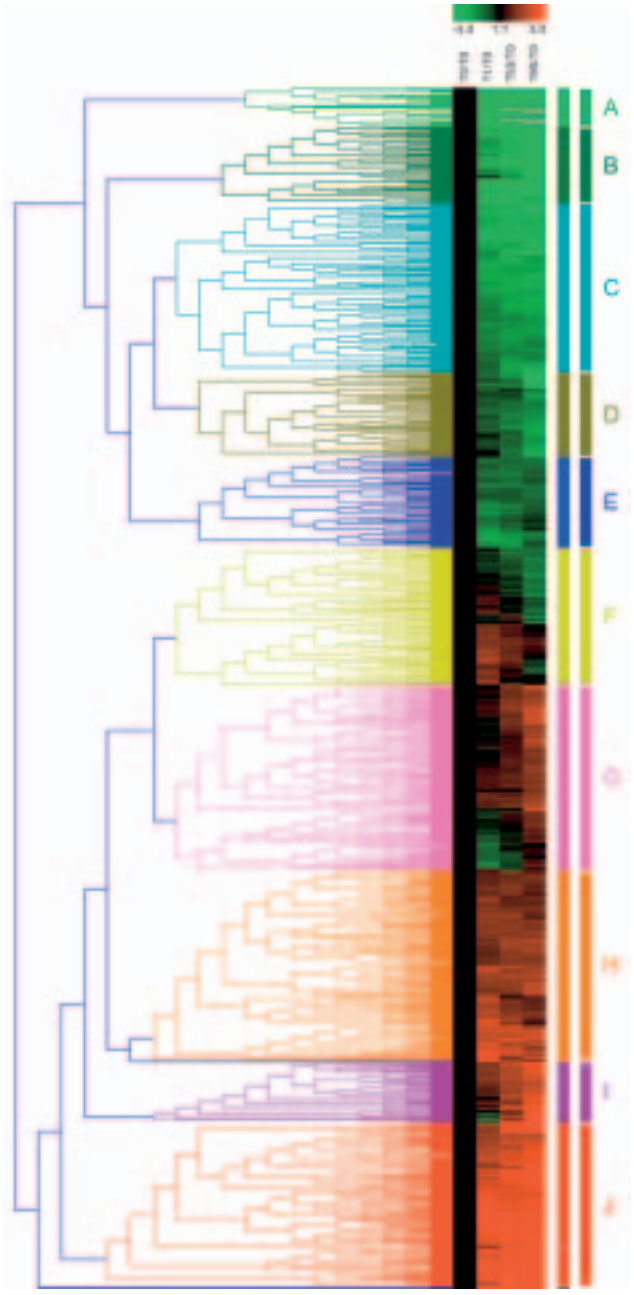


Fig. 1A: *Hierarchical clustering of 1042 genes differentially expressed during sugar beet seed germination. Every horizontal row represents an individual gene and developmental stages are represented in vertical columns. Dry seeds (T0) have been considered as reference time point and logarithmically scaled (base 2) signal intensity ratios (T1/T0, T50/T0 and T95/T0) are given. Red colour represents up-regulated genes and green colour marks down-regulated genes.*

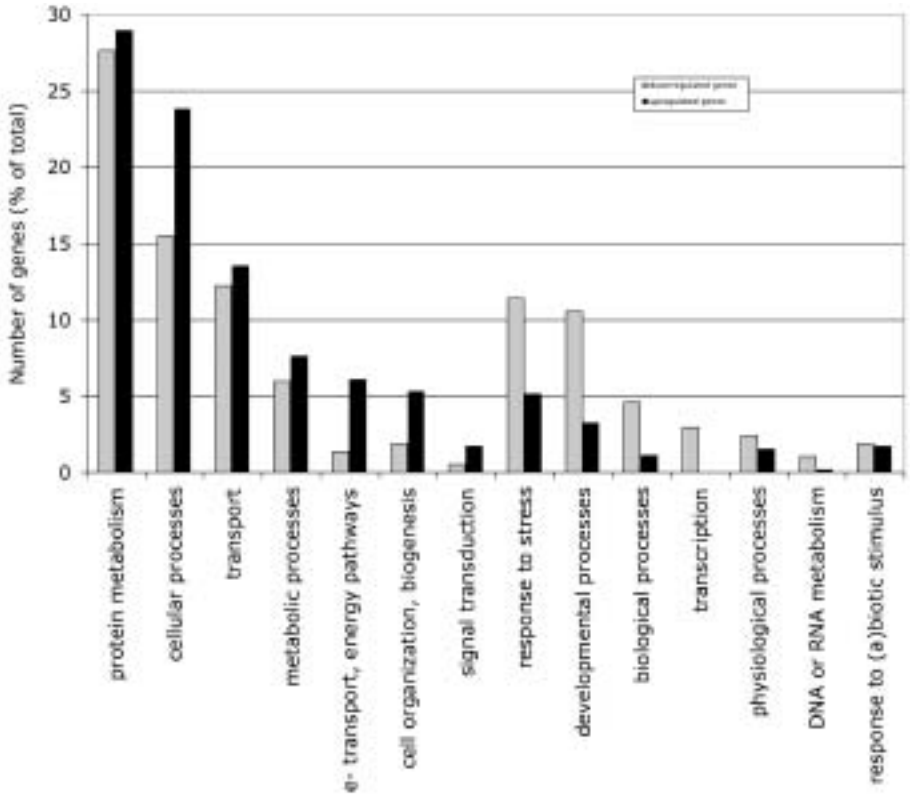


Fig. 1B: *Ontological classification of genes differentially expressed during germination of sugar beet seeds. Comparison of up- (subclusters H-J, Fig. 1A) and down-regulated (subclusters A-E) genes. The genes annotated with an e-value threshold of less than 1,00E-20 are considered.*

Proteomics of the dry mature seed

At present, 567 proteins [387 albumins (Fig.2A) + 142 globulins + 38 glutelins] were identified by LC/MS-MS mass spectrometry on the reference protein maps. On the average, the success rate of protein identification was in the range of 80%, thanks to the availability of public sugar beet ESTs (GABI-Beet). Albumins were classified according to gene ontology to reveal important mechanisms required for germination (Fig.2B). To characterize the tissue specificity of seed proteins, seeds were dissected to investigate the proteomes of roots, cotyledons, and perisperm. 89 proteins were identified whose accumulation level is more important in roots while for 128 proteins the accumulation in cotyledons is more pronounced. A major part of proteins that are differently accumulated between cotyledons and roots are metabolic enzymes. As an example part of the pathway related to methionine biosynthesis could be reconstructed in cotyledons suggesting the importance of methionine metabolism for seeds.

Proteomics of the germinating seed

The analysis of the 2D gels and the characterization of proteins expressing varying accumulation patterns during the process of germination is in progress (445 proteins identified). This will generate information about the metabolic processes that are set up in seeds during germination.

Proteomics and transcriptomics for seeds of varying germination vigour

An analysis of primed and aged seed samples was realized highlighting proteins/genes whose expression patterns are correlated to the treatment-related physiological modifications.

Conclusion and prospects

At present we have characterized more than 1000 sugar beet seed proteins, as well as about 2,800 ESTs from dry and germinated seed samples. This unique ressource allows for the first time a detailed comparative analysis of the proteome and transcriptome of dry mature as well as of germinating sugar beet seeds, generating new insights into the molecular mechanisms determining the development of a new seedling.

Acknowledgments

The PhD thesis of Julie Catusse is supported by the French Ministry of research. We thank Jean-Marc Strub and Alain Van Dorsselaer (CNRS, Strasbourg) for the characterization of sugar beet seed proteins.

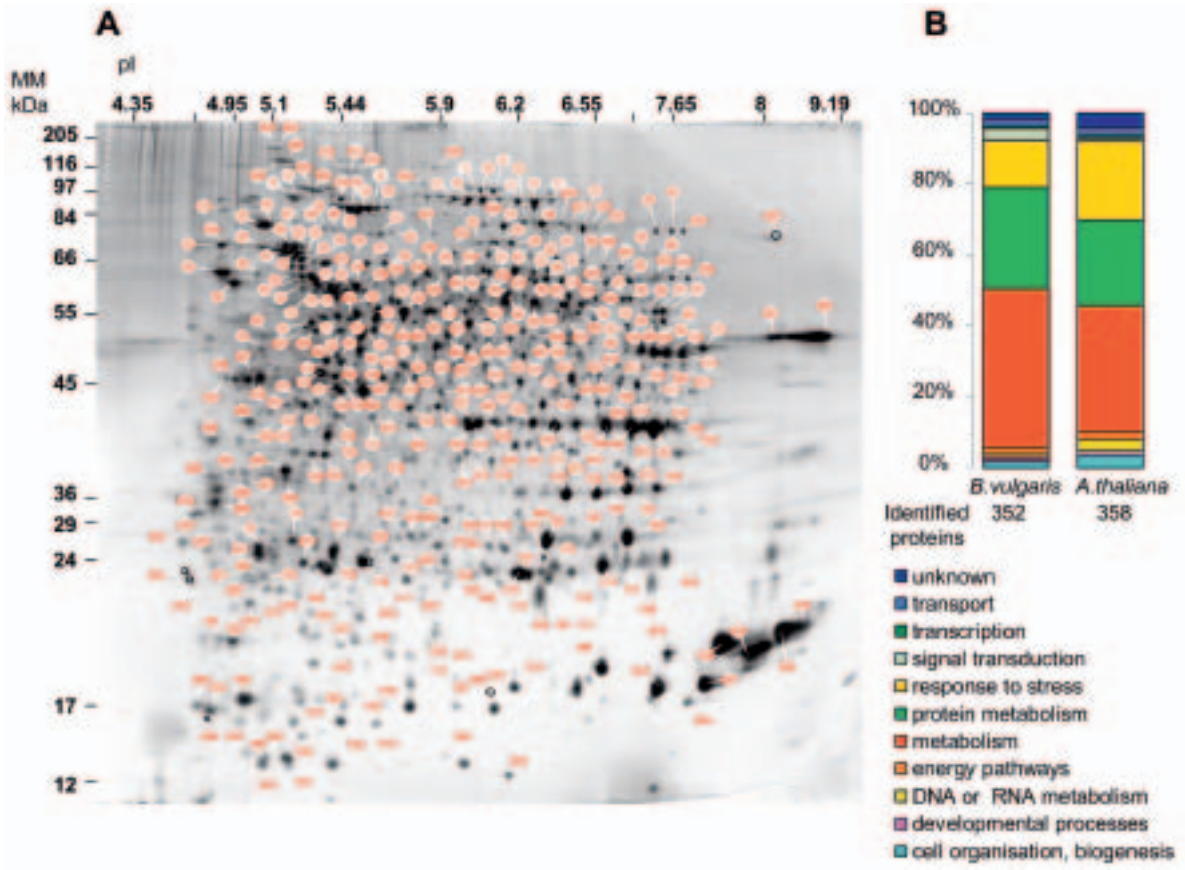


Fig. 2: *(A) Reference map by 2D gel electrophoresis of albumins extracted from dry mature seeds of sugar beet. (B) Classification according to gene ontology of proteins identified in sugar beet and Arabidopsis seed protein extracts.*

Comparative genomics between arabidopsis and rapeseed for genes directing seed-specific flavonoid biosynthesis

COMPARATIVE GENOMICS
Rapeseed

Bettina Kah 1; Boulos Chalhoub 2; Gunhild Leckband 3; Nathalie Nesi 4; Milena Ouzunova 5; Martin Sagasser 1; Rod Snowdon 6; Loïc Lepiniec 7; Bernd Weisshaar 1
1 Bielefeld University, Chair of Genome Research, Bielefeld
2 Unité Mixte de Recherche en Génomique Végétale, INRA-CNRS-Univ. Evry, Evry
3 Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Holtsee
4 UMR 118 INRA-Agrocampus Rennes, Amélioration des Plantes et Biotechnologies Végétales, le Rheu
5 KWS Saat AG, Einbeck
6 Justus-Liebig-University Giessen, Department of Crop Science and Plant Breeding
7 UMR 204 INRA-INAPG, Laboratoire de Biologie des Semences, Versailles

Introduction

Rapeseed is the most important oil crop in Europe. However, its utilisation as a source of high-quality protein is limited by the negative impact of high amounts of fibre and anti-nutritive flavonoids (i.e. condensed tannins) in rapeseed seeds. The flavonoid composition in the seed coat is controlled by a large number of genes and little is known about the inheritance of this trait in rapeseed. Therefore, breeding of yellow seeded rapeseed lines remains very difficult. Isolation and functional analysis of the genes involved in the regulation of this character is essential for breeding stable yellow seeded varieties of rapeseed.

A number of genes encoding flavonoid biosynthetic enzymes or regulatory factors have already been identified in *Arabidopsis thaliana*. Unravelling the diverse aspects of flavonoid biosynthesis in this model crucifer will accelerate the selection of interesting genes in *Brassica napus*. Indeed, the genomes of both species exhibit considerable similarities, which facilitates the recovery of flavonoid biosynthetic genes from rapeseed. In this project, we use the tools of functional and comparative plant genomics in order to specifically reduce the level of anti-nutritive condensed tannins in rapeseed.

The main topics of the project are the identification and functional analysis of *A. thaliana* genes involved in flavonoid metabolism and the identification of homologous genes in *Brassica napus* for crop improvement. Once genes of interest and/or molecular markers have been characterised, modulating the production of specific flavonoids is feasible by breeding, selection or molecular engineering.

Results

Arabidopsis thaliana

In order to identify and characterize potentially new *tt*-loci from *A. thaliana*, T-DNA insertion mutants of the GABI-Kat project and activation-tagged lines from NASC were screened for yellow seeded lines. These lines were characterized by pigment phenotyping and analysis of the flavonoid composition. Eight out of ten GABI-Kat lines analysed by the group at Bielefeld University were shown to be mutated in genes allelic to already known *TT*-genes. The metabolic profiling and phenotyping from the of new *A. thaliana tt*-mutants performed in the group of Loïc Lepiniec led to five new *tt*-mutants. For the unknown *tt*-lines from both groups crossing and genetic complementation experiments were performed that will give results soon. We hope to continue molecular analysis of any confirmed new locus.

The functional analysis of known *tt*-genes is focused on AtTT1, a putative WIP transcription factor, and AtTT8, a BHLH transcription factor. For *TT1* no real NULL-mutant is known. Therefore, one important aim is the identification of such a mutant. Currently, transposon lines are being screened for footprint alleles and TILLING populations are searched for new *tt1*-mutants. In addition, coexpression studies were performed in order to find putative targets for *TT1*. Some interesting targets were identified but need further analyses.

TT8 expression was shown by fusions to a GUS reporter gene in seeds and vegetative tissue where flavonoid accumulation takes place and in the outer integument of the seed. This is consistent with TT8's implication in mucilage biosynthesis. By introduction of this construct into different *tt*-mutants and/or quantification of *TT8* mRNA by RT-PCR it could be shown that the activity of the *TT8* promoter can be controlled by different MYB factors (TT2 or PAP1), TTG1, and different BHLHs (EGL3 as well as TT8 itself). By the use of transgenic plants expressing inducible TT2 and TTG1 proteins it could be shown that TT2 and TTG1 can directly activate TT8 expression *in planta* (BAUDRY *et al.*, 2006).

○ A. Baudry, M. Caboche and L. Lépiniec **TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*.** The Plant Journal 2006 46(5), 768-79.

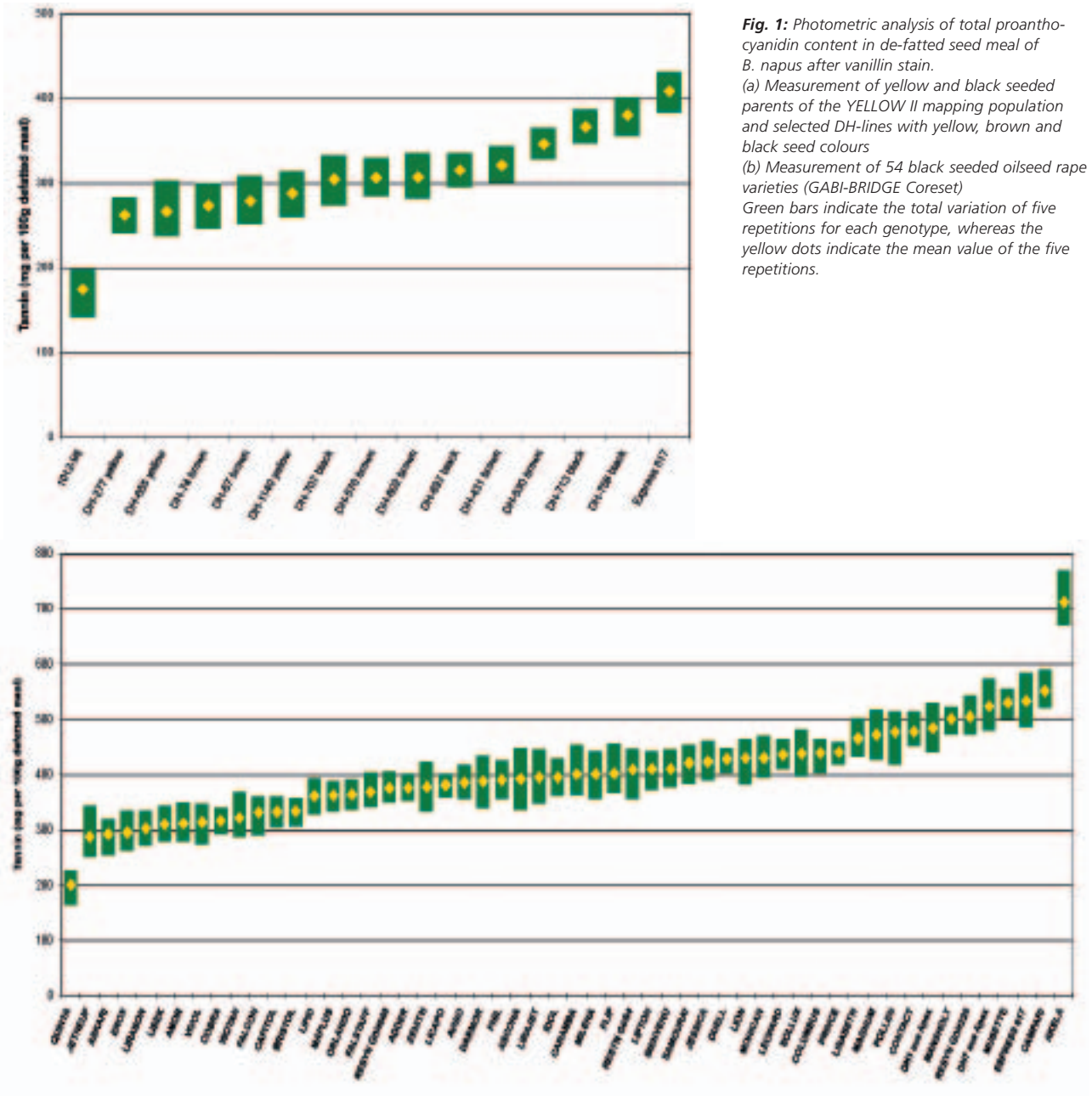


Fig. 1: *Photometric analysis of total proanthocyanidin content in de-fatted seed meal of B. napus after vanillin stain.*
(a) Measurement of yellow and black seeded parents of the YELLOW II mapping population and selected DH-lines with yellow, brown and black seed colours
(b) Measurement of 54 black seeded oilseed rape varieties (GABI-BRIDGE Coreset)
Green bars indicate the total variation of five repetitions for each genotype, whereas the yellow dots indicate the mean value of the five repetitions.

Brassica napus

For KWS and NPZ as breeding companies, the primary goal is to get greater variation for the "yellow seed" trait into the germplasm. This can later on be crossed to actual breeding material in order to change unwanted ingredients like i.e. tannins. Therefore conventional oilseed rape material is characterised for the traits tannin content, tannin composition and seed colour by the group of Rod Snowdon (contracted by KWS). The tested material consists of 50 diverse European oilseed rape genotypes of the GABI-BRIDGE core set and 30 further genotypes of the YELLOW II population, preferably yellow and brown seeded material. A clear reduction of tannin content was demonstrated in yellow seeded material (see figure 1a), but a great variation could also be shown in the black seeded material itself (see figure 1b). A closer look at the tannin composition of 80 genotypes was done by performing by HPLC analyses. It could be shown that certain compounds differ drastically between yellow and black seeded material.

In order to investigate the flavonoid biosynthesis pathway in rapeseed we want to make use of comparative genomics between *A. thaliana* and *B. napus*. One aim is the isolation and functional characterization of Bn*TT*-genes with special emphasis on Bn*TT1*, Bn*TT2*, Bn*TT7*, Bn*TT12*, Bn*TT16* and Bn*BAN*.

Boulos Chalhoub and co-workers analyzed the *B. napus* gene content for the candidate genes Bn*TT1*, Bn*TT2*, Bn*TT7* and Bn*BAN*. This was done by screening a BAC library by PCR techniques and subsequent partial sequencing. Complete genomic sequences of four Bn*TT1*, two Bn*TT2* and four Bn*BAN* gene copies were recovered including promoter and terminator sequences. As a proof of concept the sequences of Bn*TT1.1* and both Bn*TT2* genes as well as Bn*TT16.1* were used for complementation of the respective *A. thaliana* mutants. The complementation of Bn*BAN* is under way.

Three candidate genes were used to do comparative sequencing in a panel of ten mapping parents from KWS. Identified polymorphisms are being used for mapping. As expected from the results of other projects only low diversity was found in the regions of the candidate genes. So far two Bn*BAN* genes were mapped. The next step now will be the correlation of these loci with QTLs for seed colour and tannin content. For the other copies of Bn*BAN* and Bn*TT1* and Bn*TT2* useful polymorphisms still need to be detected in the german mapping populations. The group of Nathalie Nesi was able to localise Bn*TT2.2* on the Darmor x Yudal genetic map.

One of the identified Bn*TT1.1*-containing BAC clones was chosen for sequencing using a shot gun approach. After partial sequencing of a more than 200 kb insert sequence information of intergenic regions is available that will be used for further comparative sequencing. We expect to find more sequence diversity in these regions.

Stable transformants containing RNAi constructs were generated for Bn*TT2*, Bn*BAN* and Bn*TT1* in rapeseed. For Bn*TT2* and Bn*BAN*, transformation was performed by Nathalie Nesi and co-workers. No obvious phenotype could be detected. The transformation for Bn*TT1* has been performed recently and the T_1-seeds will be available in a few month. Simultaneously RNAi for *TT1* was performed in *A. thaliana* using the same construct that was generated for the Brassica transformation (consists of a region of Bn*TT1*) and the according AGRIKOLA-RNAi construct (consists of a part of At*TT1*), T_2-seeds for both transgenics do not show any obvious phenotype, but will be investigated further.

EMS-mutagenesis and subsequent TILLING in *B. napus* is done by INRA Rennes and INRA Evry in order to find knock out mutants for our candidate genes. After initial problems with the polyploid rapeseed genome have been solved, they developed a TILLING strategy. The strategy was validated with the

Bn*TT2.2* sequence, where 16 independent mutations could be identified and verified by sequencing. The first sub-population is highly mutagenized containing approximately one mutation per 20 kb.

Outlook

The functional analyses of At*TT* genes will continue, including the search for NULL allele of At*tt1* and the characterization of new tt mutants. The functional analyses of Bn*TT* genes (mainly *TT1, TT2, TT8* and *BAN*) are underway including gene expression analyses, complementation assay in *A. thaliana* and RNAi in both *B. napus* and *A. thaliana*. It has been already demonstrated that Bn*TT1*, Bn*TT2* and Bn*TT16* restore the wildtype phenotype when ectopically expressed in *A. thaliana* mutant. Experiments are in progress for the other *TT* genes. The mutant lines identified by TILLING and transgenic plants generated will be characterized at the phenotypic level. The candidate genes will be analyzed by genetic mapping, association studies and analysis of allelic diversity.

Trilateral project

GENMETFRUQUAL – Genomic and Metabolomic Exploration of Fruit Devlopment and Quality in Tomato

GENMETFRUQUAL
Tomato

Maria-Ines Zanor 1; Mondheur Boyazen 2; Mathilde Causse 3; Christophe Rothan 4; Rafael Fernandez 5; Antonio Granell 6; Joachim Kopka 1; Bjoern Usadel 1; Alisdair R. Fernie 1

1 Max-Planck-Institut für Molekulare Pflanzenphysiologie, Postdam-Golm, Germany
2 INRA-Toulouse, France
3 INRA-Avignon, France
4 INRA-Bordeaux, France
5 EELM-Almeria, Spain
6 IBMCP-Valencia, Spain

Foreword

Tomato fruit quality for the fresh market has been strongly improved during the last 30 years. Firmness and long shelf life, specially required for shipping to distant markets, is one of the most important characteristics for fresh market breeders. With the availability of tomato all year round, consumers began to complain about tomato flavour. Mainly, the increase in firmness observed in modern varieties is also responsible for consumer complaints. Taste-panel studies have shown that sweetness and sourness are the major determinants of tomato flavour preference together with the major components of organoleptic quality (appearance, colour, aroma and texture).The aim of the GENMETFRUQUAL consortium was to gain better understanding of molecular factors underlying fruit quality. For this purpose a range of genetically diverse tomatoes were characterised at the genomic level using a broad spectrum of transcriptomic, proteomic and metabolomic approaches. Genotypes used included recognised cultivars as well as transgenics exhibiting alterations in hormone and vitamin metabolism as well as introgression lines in which genome segments from either a different cultivated or a wild species tomato were substituted into elite cultivars. Once completed the data will be published in a relational databases in order to allow linkage of phenotyping data from the above mentioned technology platforms and where available consumer taste trials.

Objectives

GENMETFRUQUALs aim was to resolve the molecular basis of fruit quality within the tomato. For this purpose collaborative interaction was established between experimental groups of different expertise and database tools were established in order to allow integration and interrogation of various genomic and quantitative genetic datasets. Three complementary approaches were taken. First, the analysis of the broad genetic diversity supplied by introgression populations containing marker-defined substitutions of wild species chromosomal segments and already characterised from the consumer perspective. Secondly, parallel transcriptomic and metabolomic profiling was carried out in order to identify candidate genes for the alteration of fruit quality at the level of chemical composition of the fruit. Finally, transgenic lines that exhibit alterations in various quality attributes (namely seed set or vitamin content) were analysed at the level of chemical composition in order to get a more comprehensive picture of the mechanisms underlying the observed changes in fruit quality. Given that the tomato genome is currently being sequenced results from this study should provide important knowledge that can be transferred to other climacteric fleshy fruits.

Tools for data visualisation, integration and correlation

As a pre-requisite for this project the establishment of bioinformatic tools for data analysis, evaluation and warehousing including the integration of multi-factoral dataset was essential. This was achieved in a range of ways firstly the establishment of a database of mass spectral tags (Schauer *et al.*, 2005) was generated which greatly improves mass spectral annotation of tomato chromatograms. This database represents a valuable resource that is freely available to the plant community and promises to be a great aid in future identification of unknown metabolites. It is currently embedded in the comprehensive systems biology database (csbdb; http://csbdb.mpimp-golm.mpg.de) a forum that will facilitate the warehousing and subsequent data-mining of diverse data sets provided by transcriptomics, proteomics and metabolomics experiments. Furthermore the MapMan program for visualisation of Arabidopsis transcript profiles has been converted in order that it can be used to display transcript changes from Solanaceous species such as tomato (Urbanczyk-Wochniak *et al.*, 2006). Both correlation analysis and MapMan have been utilized in the analysis of volatiles and in changes of metabolites in hormone affected tomatoes with changes in the transcriptome which revealed mechanistic insights into the causality of these changes.

Alhagdow M, Mounet F, Nunes-Nesi A. *et al.* (2006). **Changes in ascorbate redox state through silencing of L-galactono-1,4-lactone dehydrogenase (GLDH) affects plant and fruit development in tomato** (in review). Lecomte L, Duffe P, Buret M, *et al.* (2004). **Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds.** Theor Appl Genetics 109: 658-668. Schauer N, Semel Y, Roessner U *et al.* (2006) **Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement.** Nat Biotechnol 24: 447-454. Schauer N, Steinhauser D, Strelkov S *et al.* (2005) **GC-MS libraries for the rapid identification of metabolites in complex biological samples.** FEBS Lett 579:1332-1337. Urbanczyk-Wochniak E, Usadel B, Thimm O *et al.* (2006) **Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: effects of genetic and environmental alterations in energy metabolism en the leaf.** Plant Mol Biol 60: 773-792.

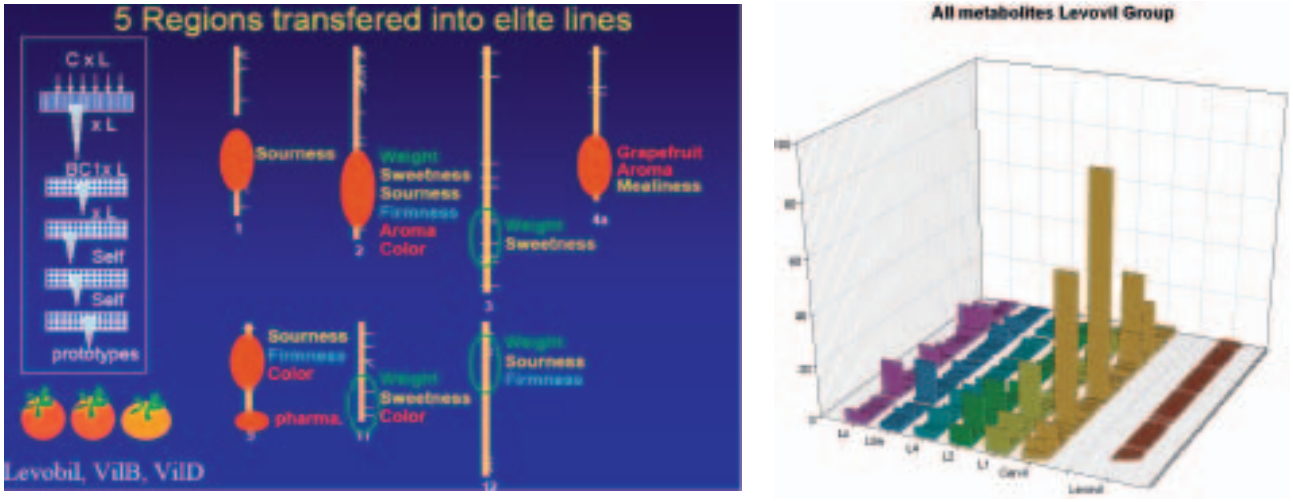


Fig. 1: *Upper panel: Schematic representation of the generation of introgression lines (described in detail in Lacomte et al. 2004) containing a QTL map of organoleptic properties. Red areas indicate regions intogressed into Levobil, VilB or VilD elite lines Lower panel: histogram illustrating the degree of metabolic change between a subset of the parental and introgression lines.: Interestingly a reasonable degree of transgressive behaviour is displayed.*

Analysis of chemical composition in introgression line populations of tomato

Naturally occurring variation in wild species can be used to increase the genetic diversity of cultivated crops and improve agronomic value. Populations of introgression lines carrying wild species alleles facilitate the identification of traits associated with the introgressed regions, and hence the characterization of the biochemistry and genetics underlying these phenotypes. This approach has been taken on two fronts. First, the identification of QTL for metabolites believed to be important determinants of quality was carried out in a *S.pennelli* vs *S. lycopersicum* introgression population (Schauer *et al.*, 2006). Secondly, exclusively within this project, introgression lines that have already been evaluated by the consumer for taste and organoleptic properties (Lecomte *et al.*, 2004) were comprehensively characterised at the metabolite level in order to evaluate the chemical composition recognised to confer these properties. Taste-panel studies have shown that sweetness and sourness are the major determinants of tomato flavour preference together with the major components of organoleptic quality (appearance, colour, aroma and texture). The quantitative trait loci (QTLs) detected for sourness and sweetness mapped within few regions on chromosomes 2, 3, 9 and 11. After the transfer of the favourable regions into two other elite lines (VilB and VilC) 36 selected introgressed lines (ILs) were subjected to metabolic profiling. The metabolite concentrations of fruit pericarp harvested were evaluated using gas chromatography/mass spectrometry (GC/MS) metabolite profiling. At least 66 metabolites of known structure were accurately quantified in every chromatogram. These compounds include most plant amino and organic acids, sugars, sugar alcohols, fatty acids and vitamin E (a- tocopherol). Interestingly many other metabolites were also highly variable across the parental lines. For example, the Cervil parent showed higher amount of the amino acids aspartate and phenylalanine (used as flavour intensifiers) than observed in the Levovil parent (whereas only phenylalanine was higher in respect to the VilB parent line). A representative dataset from

the Leovivl population is presented in Figure 1. The following conclusions could be made. Firstly, the ILs adopt both characteristics of the parental lines and on occasion exhibit transgressive behaviour. Secondly, some of the exhibited changes in metabolite content are very large however these tend to be stress-associated metabolites such as proline or poorly measured metabolites such as 2-OG, these need to be interpreted with extreme caution. With the exception of these compounds the range in metabolite content in these crosses is relatively similar to that seen in the aforementioned *S. pennelli* vs *S. lycopersicum* ILs. A further series of *S. pimpinelifolium* vs *S. lycopersicum* ILs was also worked on during this project, however, as yet not at the metabolomic level.

Combined transcript and metabolite profiling for identification of candidate genes

In addition transcript profiling across a developmental time course was carried out using the TOM1 chip in parallel with GC-MS based metabolite profiling for this purpose organic acids, amino acids, sugars cell wall monomers and pigments were profiled. Simple correlation analysis was carried out in order to determine transcripts that correlated with the content of biotechnologically important metabolites (Carrari *et al.*, in press). Interestingly, certain compound classes are highly correlated with the same transcripts suggesting that they are highly likely regulated at the level of transcription. A subset of the resultant data matrix is presented in Figure 2 below. In addition to its use for the identification of candidate genes this data also provided important fundamental insight into mechanisms of metabolic regulation within the fruit.

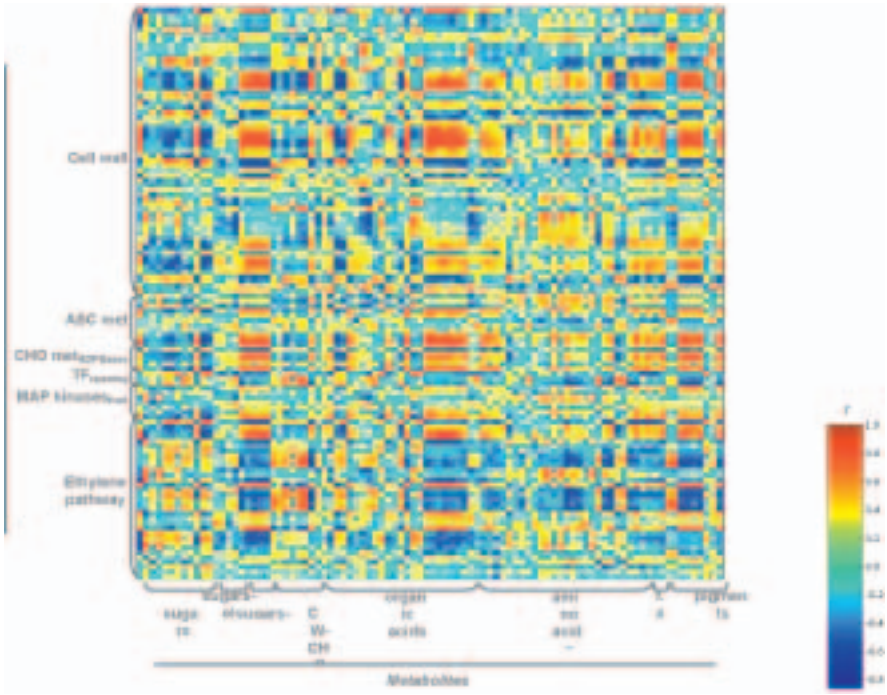


Fig. 2: *Heatmap of correlations between selected transcripts and metabolites. Each dot indicates a given r value resulting from a Spearman correlation analysis in a false colour scale. RI TFs TDR: ripening related transcription factors (TDR family); CHO-AGPses: carbohydrate metabolism ADPglucose pyrophosphorylases.*

Metabolite profiling as a tool to gain mechanistic insight into factors underlying diverse changes in fruit quality

More focussed studies including both transcriptomic and metabolomic analysis of fruits subjected to hormonal application or transgenics exhibiting alterations in giberellic acid, auxin and ethylene have additionally been carried out. Although detailed evaluation of these results is yet to be carried out. In addition a targeted approach was also taken in which tomato fruit altered in either the galacto-lactone dehydrogenase or phosphomannose epimerease reaction were created via antisense technology. These plants exhibited dramatic phenotypic and metabolic changes (Alhagdow *et al.*, 2006) including alteration in Krebs cycle metabolism which underlines previous work suggesting the importance of this pathway in fruit quality and furthermore provides important information relevant for future metabolic engineering strategies aimed at elevating vitamin C content in fruits

Outlook

Once completely gathered the data of this project will give a relatively comprehensive overview of molecular and metabolic features defining fruit quality and its development. The data will finally be warehoused in a bespoke database whose interrogation should enable the possibility to identify key regulatory genes involved in metabolic and developmental processes responsible for fruit quality in tomato and other crops for which it serves as a model organism.

GABI-MALT: An integrated approach to the genetic and functional dissection of malting quality in barley SNP-detection and haplotype analysis in candidate genes for malting

MALT
Cereals

Inge Matthies 1; Jutta Förster 2; Marion Röder 1
1 Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstr. 3, D-06466 Gatersleben
2 SAATEN-UNION Resistenzlabor GmbH (SURL), Hovedisser Strasse 92, D-33818 Leopoldshöhe

Background and Rationale

Apart from yield, a good malting and brewing quality is the most important breeding aim in any spring barley breeding program. Despite its salient importance, its quantitative inheritance prevented a detailed genetic analysis on the structural level. The project GABI-Malt is splitted into four subprojects and is carried out in close collaboration with all project partners. Subproject 4 represents a broad approach for harnessing the structural genetic diversity present in the barley germplasm by associating haplotypes of candidate genes to malting quality parameters with the goal of marker-development for those traits. This is performed in three main steps (1) identification of candidate genes using functional genomics approaches, (2) analysis of the allelic diversity of malting related candidate genes, and (3) association of haplotypes and single nucleotide polymorphism (SNP)-patterns with malting quality parameters. In a public private partnership we are cooperating with SURL in terms of marker development.

The obtained results will increase our understanding of the genetic diversity in genes of metabolic pathways underlying malting quality. The findings will contribute to the development of new cultivars with improved physiological performance by exploiting the allelic diversity present in barley.

Identification of candidate genes in barley with impact on malting quality

Our focus is the analysis of SNPs and haplotypes of candidate genes for malting and their association to phenotypic malting parameters. For this purpose candidate genes from two sources were chosen:

- A total of 48 genes were selected according to their relevance to malting and brewing in literature (Fox *et al.* 2003; Hayes *et al.* 2003). These are genes coding for enzymes known to be important in starch-degradation of the germinating barley grain, such as α-amylases, α-glucosidases, saccharases, as well as those responsible for cell wall degrading activities, such as xylanases, β-glucanases and genes coding for enzymes and proteins with more general functions, such as peptidases, phosphatases, kinases and lipases and genes coding for storage proteins like hordeins or certain inhibitors (e.g. BASI).
- As second source of candidates are presented by 16 ESTs, which were specifically expressed in malting cultivars were included in this research program (Potokina *et al.* 2004; provided by A. Graner, Subproject 1).

Assessment of haplotype structures and marker development

The assessment of haplotype structure of all selected candidate genes was performed in several steps (Fig. 1). By sequence analysis of PCR-products 444 primer combinations covering partial gene fragments of 500-800 bp size were screened on 8 highly diverse reference genotypes. Of all screened primer combinations, 54.2 % amplified single-copy-fragments and were subsequently tested on 64 barley cultivars, which were selected by LfL Freising (Subproject 3). Out of those, 30 % of all sequenced fragments possessed SNP- and 12.35 % INDEL-polymorphisms.

After analysis of the gene structure, those SNPs showing high polymorphisms in 64 cultivars and leading to amino-acid exchanges after transcription were preferentially used for development of suitable pyrosequencing assays in order to perform high-throughput-genotyping in a larger set of genotypes.

Association analysis

By combining genotypic and phenotypic information, it is possible to perform association studies to a large extent (Fig. 1). For this purpose, data of 120 malting and brewing parameters of 250 barley cultivars were collected either from public sources such as 'Bundessortenamt', 'Landessortenversuche' and 'Braugerstenjahrbücher' of the past 20 years or delivered by LfL Freising from actual field and micromalting experiments of 64 cultivars in Freising 2004 and 2005. All data are handled in a database called "MetaBrew" which is developed in collaboration with the Bioinformatics group at IPK. Up to now, approximately 80 000 datapoints are available.

Association studies were performed with the software TASSSEL, assuming the "General Linear Model" (GLM) and significant relationships between single SNPs or haplotypes and certain malting properties could be determined for some of the candidate genes. As example 1,4-β-xylanase (EC 3.2.18), which hydrolyzes cell wall arabinoxylans, is outlined here (Fig. 2). The analysis of a gene fragment coding for 1,4-ß-xylanase X-1, tested on 72 cultivars resulted in 5 haplotypes based on 4 SNPs (Fig. 2c). These were correlated with

○ G.P. Fox *et al.* **Molecular basis of barley quality.** Australian Journal of Agricultural Research (2003) 54, 1081–1101. ○ P.M. Hayes *et al.* **Genetic diversity for quantitatively inherited agronomic and malting quality traits.** In: R von Bothmer, Th van Hintum, H Knüpffer, Sato K (eds), Diversity in barley (*Hordeum vulgare*), (2003) pp. 201-226. Elsevier Science B.V., Amsterdam, The Netherlands. ○ E. Potokina *et al.* **Functional association between malting quality trait components and cDNA array based expression patterns in barley (*Hordeum vulgare* L.)** Mol. Breed. (2004) 14, 153-170.

13 different traits, derived from micromalting experiments by LfL in Freising (subproject 3) and significant assocations could be found for the parameters viscosity and friability for haplotype H4_GM111 at $p<0.05$ (Table 1).

Prospects

All results obtained in this study provide a complex dataset for performing association genetics in barley. We gained a lot of information on genetic diversity of malting related genes by identification of sequence polymorphism. Furthermore, the developed SNP-markers for malting quality are useful for the screening of unexploited germplasm and for marker assisted selection. Therefore, the pyrosequencing markers developed in this project part, provide a valuable and excellent tool for marker assisted selection in any malting barley breeding program.

Cooperating Project Partners

- Saaten-Union-Resistenzlabor (SURL), Leopoldshöhe
- Bayerische Landesanstalt für Landwirtschaft (LfL), Freising
- IPK Gatersleben, Group Genome Diversity (Prof. A. Graner)

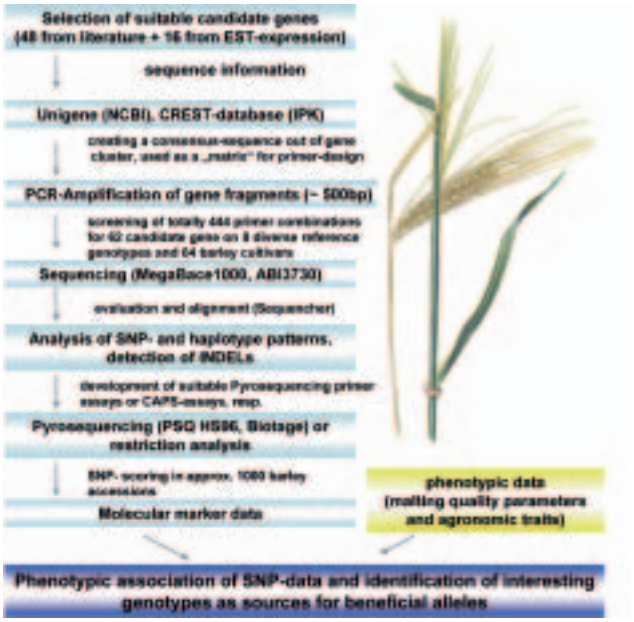
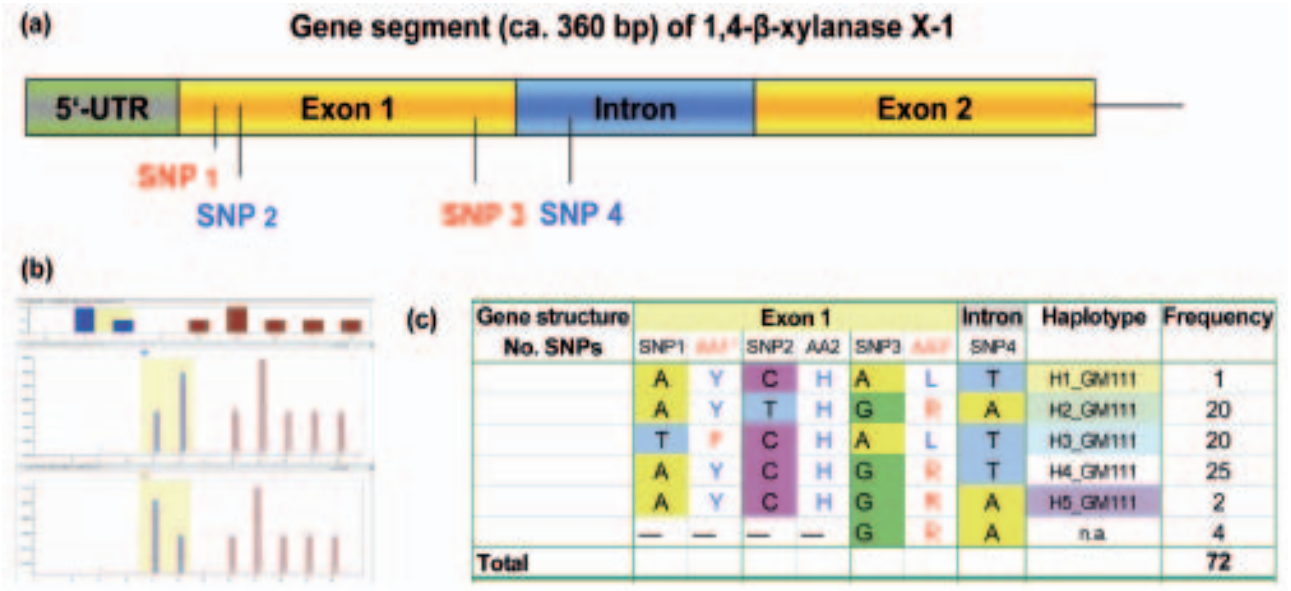


Fig. 1: *Procedure of SNP-marker development for malting quality in barley.*



Gene structure	Exon 1						Intron	Haplotype	Frequency
No. SNPs	SNP1	AA1	SNP2	AA2	SNP3	AA3	SNP4		
	A	Y	C	H	A	L	T	H1_GM111	1
	A	Y	T	H	G	R	A	H2_GM111	20
	T	F	C	H	A	L	T	H3_GM111	20
	A	Y	C	H	G	R	T	H4_GM111	25
	A	Y	C	H	G	R	A	H5_GM111	2
	—	—	—	—	G	R	A	n.a.	4
Total									72

Fig. 2: *(a) Partial gene structure of (1,4)-β-Xylan Endohydrolase Isoenzyme I. Non-synonymous SNPs which are responsible for amino acid (AA) changes are labeled red and marked with * in (c), (b) Pyrograms of SNP3 in a fragment of 1,4-β-xylanase X-1, derived by PCR with primer-combination GM111, (c) Detected SNPs and haplotype pattern in a 350 bp-fragment of this 1,4-β-xylanase X-1 gene, amplified with primer-combination GM111 in a set of 8 references and 64 barley cultivars.*

Selection and evaluation of a complete set of wild barley introgression lines (ILs)

MALT
Cereals

Inga Schmalenbach, Klaus Pillen
Max-Planck-Institute for Plant Breeding Research, Köln

Co-operations
1 Nordsaat Saatzucht GmbH, Böhnshausen
2 Saatzucht Josef Breun GdbR, Herzogenaurach
3 Lochow-Petkus GmbH, Bergen-Wohlde
4 Saatenunion Resistenzlabor, Leopoldshöhe
5 IPK, Gatersleben
6 LfL, Freising-Weihenstephan
7 University of Bonn, Dikopshof Research Station

Introduction

The development of introgression lines (ILs) provides an opportunity to use the genetic potential of wild species. As described by Zamir (2001), a complete set of ILs is supposed to represent the entirety of the wild genome, while each line contains only one single marker-defined chromosomal segment of the wild parent and the remaining genome is consistently derived from the elite parent. After the development of the ILs by repeated backcrossing, subsequent selfing and parallel marker-assisted selection, they are a reliable and stable resource with characterized genotypes. This resource can be used versatilely. A great advantage of the ILs is the small part of the wild genome, which each line carries. Therefore, the ILs facilitate the fine mapping of QTLs and enable the statistical detection of less phenotypic effects because of reduced linkage drag. Furthermore, the ILs can be used for the investigation of the phenotypic effects of QTL interactions by crossing those lines, which carry QTLs. Within the scope of the GABI-Malt project, it is intended to select a complete set of wild barley introgression lines for the first time. This set will be used both to verify QTL effects and to identify new QTLs for malting quality traits, agronomic traits and disease resistances.

Project description

The project started with 40 pre-ILs (candidate introgression lines) of the spring barley population S42, which were selected from 301 BC2DH lines in the previous project GABI-Diversity (Figure 1). The DH lines were generated from a primary cross between the spring barley cultivar Scarlett (*Hordeum vulgare* ssp. *vulgare*) and the wild barley accession ISR 42-8 from Israel (*Hordeum vulgare* ssp. *spontaneum*). These lines were used for several AB-QTL studies in GABI-Diversity (von Korff *et al.* 2005, 2006, and submitted). The 40 pre-ILs were selected on the basis of the following criteria: (1) The selected lines should possess a low percentage of the wild genome. (2) They ought to contain only a few donor segments additional to the favored introgression. (3) The introgressions in the selected lines should be overlapping to ensure that the entire exotic genome is covered within the complete set of lines (von Korff *et al.* 2004). As described in Figure 2, the introgression lines are developed by backcrossing of the pre-ILs with the recurrent parent once again (BC3) in order to further reduce the portion of the exotic genome and to minimize the favored introgression. The BC3 lines were subsequently selfed several times to receive complete homozygous lines. In BC3S2 useful plants were selected as introgression lines on the basis of genotype data of 98 SSR markers. These selected ILs were selfed twice again (BC3S4). In this generation the finished ILs, which only contain the favored introgression, were verified on the basis of single SSR loci. In addition, more finished ILs are selected from the pool of the ILs under construction, which at present still contain one or more heterozygous loci. All finished ILs were subsequently propagated (BC3S4:6) to receive seeds for field trials at three locations and for micro malting. After completion of phenotyping, these data and the genotype data will be analyzed in the form of a marker-phenotype association study both to verify QTL effects and to identify new QTLs.

The present state of the project

So far, there are altogether 92 selected ILs, which can be divided into 55 finished ILs and 37 ILs still under construction (Figure 3). All finished ILs were verified successfully and 39 of these lines are propagated in winter 2006/07 in New Zealand in cooperation with Lochow-Petkus to receive a sufficient number of seeds for the phenotype studies in 2007. Approximately one third of the finished ILs (32) are singular, this means that they are not completely represented by other ILs. These lines are absolutely necessary for the set of ILs, whereas the other lines are preferable, because they contain smaller exotic introgressions, but not essential for the aim of the project. At present 70.3 % of the wild genome are covered by finished ILs, whith a range from 56.5 % to 92.6 % of the covered genome per chromosome (chromosome 5H and 4H,

M. von Korff *et al.* **Development of candidate introgression lines using an exotic barley accession (*Hordeum vulgare* ssp. *spontaneum*) as donor.** Theor Appl Genet (2004) 109:1736-1745 M. von Korff *et al.* **AB-QTL analysis in spring barley: I. Detection of resistance genes against powdery mildew, leaf rust and scald introgressed from wild barley.** Theor Appl Genet (2005) 111: 583-590. M. von Korff *et al.* **AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (*H. vulgare* ssp. *spontaneum*).** Theor Appl Genet (2006) 112: 1221-1231. M. von Korff *et al.* **AB-QTL analysis in spring barley: III. Identification of exotic alleles for the improvement of malting quality in spring barley (*H. vulgare* ssp. *spontaneum*)** (submitted). D. Zamir **Improving plant breeding with exotic genetic libraries.** Nat Rev Genet (2001) 2:983-989

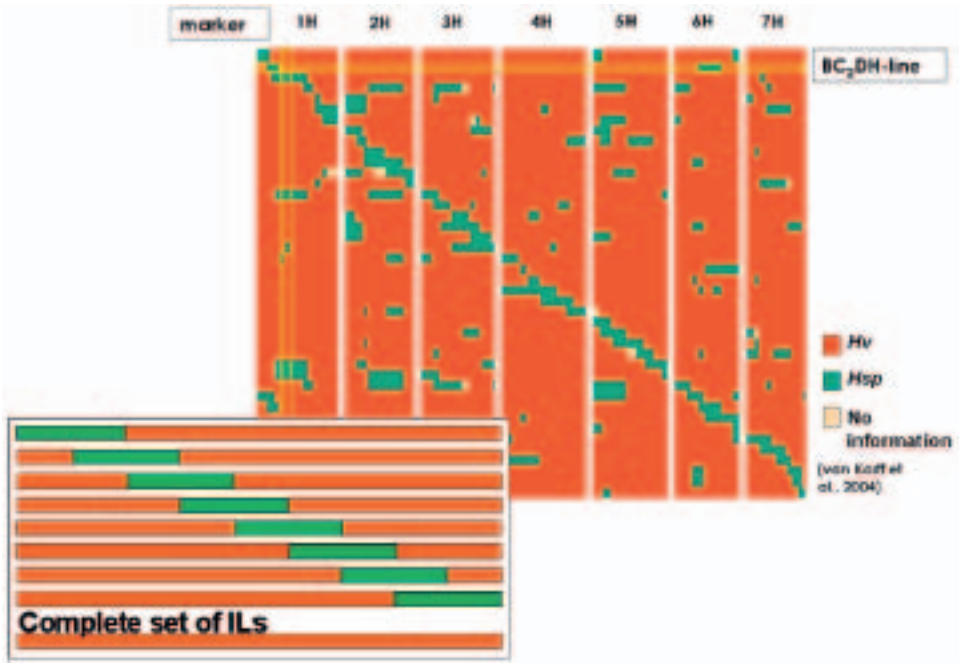


Fig. 1: *Graphical genotypes of 40 BC2DH lines of the population S42 and schematic illustration of a complete set of introgression lines (ILs) (von Korff et al. 2004, modified). The BC2DH lines are completely homozygous and contain several introgressions of the wild barley genome. Therefore, they provide a basis for the development of a complete set of ILs. These ILs only contain the favored introgression whereas all additional introgressions are eliminated. The introgressions have to overlap to ensure that the complete wild genome is covered by the entire set of ILs.*

respectively). The finished ILs contain an average exotic introgression of 31.7 cM (range: 7.0 - 100.5 cM). Eight lines include small introgressions of less than 15 cM. At present the ILs under construction are genotyped again in subsequent selfing generations. This means that all loci, which were yet heterozygous in BC3S2, are investigated again with SSR markers in BC3S4. The first aim is to select more finished ILs in order to finally cover the whole wild barley genome. Secondly, the selection of ILs with very small introgressions is intended, so that in future a set of ILs with introgressions smaller than 5 cM will be anticipated. Such a set could facilitate the fine mapping of QTLs, because it would make it possible to narrow down the position of a QTL. The development of ILs with very small exotic introgressions requires a larger density of markers (e.g., one marker/cM). This could be achieved by genotyping with a SNP chip (Illumina 1.5 k barley SNP chip). At present there are five chromosomal regions with neither finished ILs nor ILs under construction. Useful BC3 plants with more than one exotic introgression will be crossed with Scarlett once again to close these gaps. For one region on chromosome 7H there have already been no appropriate pre-ILs in the pool of BC2DH lines. It might be possible that the adjacent ILs in this region are already overlapping. This would be revealed by the genotyping of new markers. Otherwise, it would be necessary to cross the elite and the wild parent again to develop ILs, which would bridge the gap on chromosome 7H.

Future work

The complete set of ILs can be characterized on different levels. In the current project, the phenotype of the ILs will be evaluated during the year 2007 in field trials at three locations in Germany. These are the experimental station Dikopshof (University of Bonn, North Rhine-Westphalia) and the breeding stations Gudow (Nordsaat, Schleswig-Holstein) and Herzogenaurach (Saatzucht Breun, Bavaria). At all three

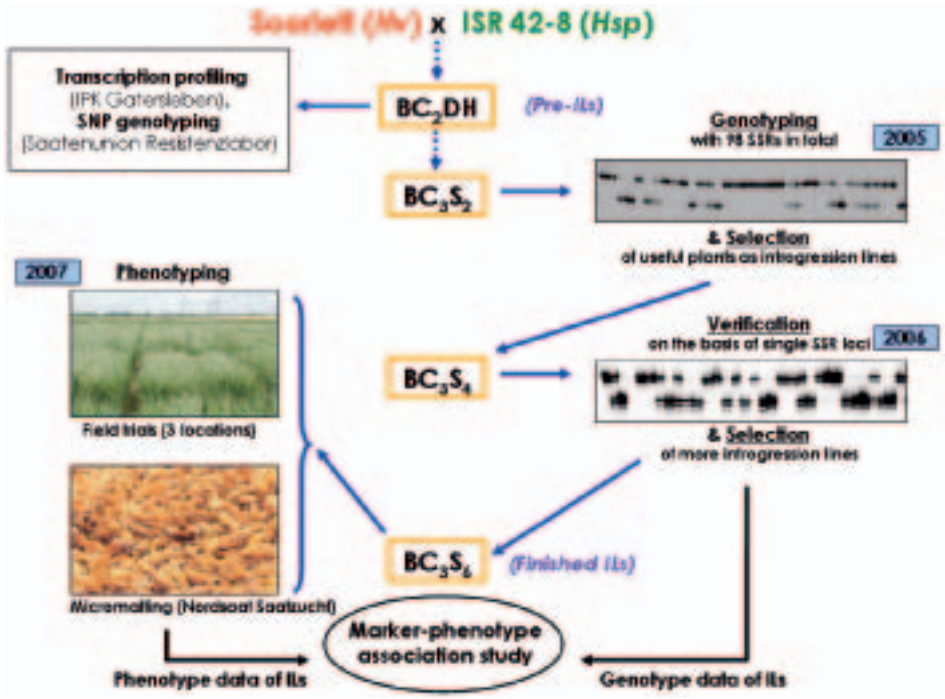


Fig. 2: *Strategy of the selection of wild barley introgression lines (ILs) in the population S42. Useful plants were selected as introgression lines on the basis of genotype data from 98 SSR markers in BC3S2. These selected ILs were selfed twice again (BC3S4). In this generation the finished ILs, which only contain the favored introgression, were verified. In addition, more finished ILs are selected from the pool of the ILs under construction, which at present still contain one or more heterozygous loci. All finished ILs were subsequently propagated until BC3S4:6 to receive seeds for field trials and micromalting. After completion of phenotyping, these data and the genotype data will be analyzed in the form of a marker-phenotype association study both to verify QTL effects and to identify new QTLs.*

locations, agronomic traits (e.g., plant height, time to flowering, and total yield) as well as disease resistance (e.g., against powdery mildew, scald, and leaf rust) will be tested. The ILs will also be subjected to micro malting (at Nordsaat) in order to investigate malting quality traits as α-amylase activity, grain protein content, and malt extract. Finally the genotype and the phenotype information will be used to verify QTL effects, which were previously detected in GABI-Diversity. Furthermore new QTLs and favorable exotic QTL alleles should be located. In addition, the ILs will be characterized at the levels of DNA and RNA. To achieve this, they will be applied both to SNP analysis of candidate genes for malting quality parameters (at Saatenunion-Resistenzlabor) and to transcription profiling at different time points of malting (at IPK Gatersleben).

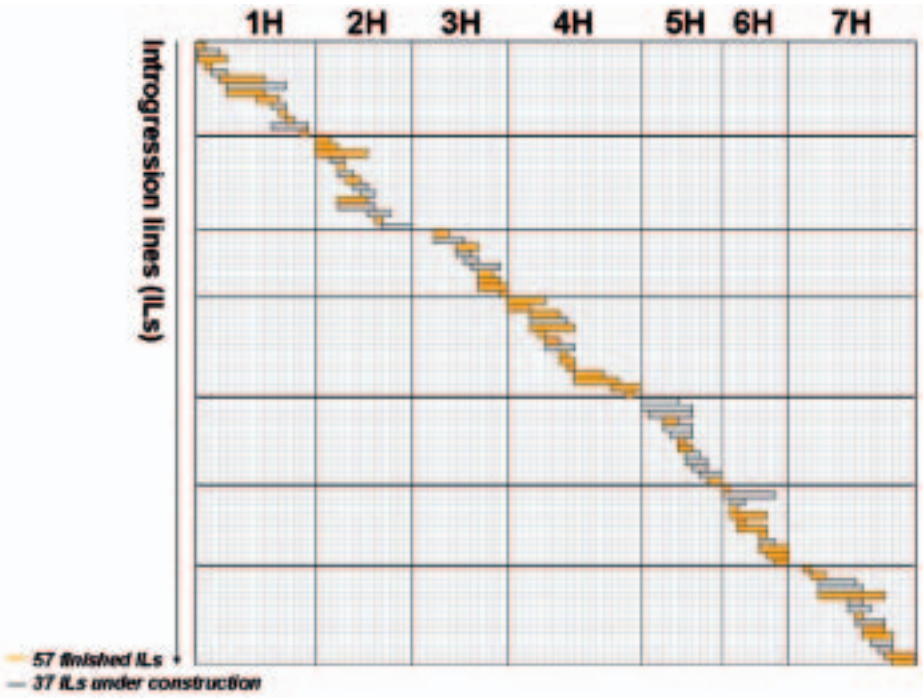


Fig. 3: *Graphical genotypes of 92 introgression lines of the population S42. The ILs are shown in rows and the seven chromosomes of barley with the SSR loci in columns. The yellow bars represent the exotic introgressions of the finished ILs and the grey bars those of the ILs still under construction.*

Identification, functional analysis and marker development of candidate genes related to malting quality by cDNA-AFLP and SSH techniques

MALT
Cereals

Sabine Mikolajewski, Kerstin Krumnacker, Günther Schweizer, Markus Herz
Bavarian State Research Center for Agriculture, Institute for Crop Science and Plant Breeding
Am Gereuth 2, D-85354 Freising-Weihenstephan, Germany

Project Goals

The malting quality of barley represents the manifestation of a well-adjusted interaction of several different genes. Due to the complex genetic basis and its quantitative inheritance the trait malting quality requires sophisticated methods to be tagged by genetic markers. To identify candidate genes related to this trait, and to generate functional genetic markers the GABI-MALT subproject at the LfL is divided into three main areas: i) development and evaluation of plant material, ii) induction and analysis of gene expression by cDNA-AFLP techniques and iii) transcriptome mapping of differential TDFs (transcript derived fragments) to develop functional markers for the selection of improved varieties

Generation of plant material

Basis of this project is a QTL map of a DH population derived from a cross between the spring barley cultivars ALEXIS and STEINA, representatives of different malting types. In four particular intervals on chromosome 1H, 4H, 5H and 6H QTL hot spots for malting quality have been identified (Hartl *et al.*, 2000). To determine the influence of these genomic regions NILs (nearly isogenic lines) have been developed during the first phase of GABI (GABI-SEED), which are now used to validate the four important QTL hot spots. Therefore corresponding pairs of NILs for each QTL hot spot as well as the ALEXIS x STEINA DH population, comprising 134 DH lines, have been grown in multilocation field trials. In addition, a reference set of 51 spring barley and 8 winter barley varieties registered between 1951 and 2004 have been multiplied and multilocation field trials were performed in 2004 and 2005.

As one of the major tasks of the LfL the harvested material of all subareas was subjected to micromalting and malt analysis. As illustrated by figure 1A-C the micro-malting process is a strictly controlled fully automated procedure performed in a closed system. The micromalting facilities at the LfL allow a throughput of 240 samples per passage. Following routinely 12 important parameters (Crude protein, Soluble N, Brabender, Friability, Kolbachindex, VZ45°C, Viscosity, pH, Colour, Malt extract, Final attenuation, Diastatic power) are analysed according MEBAK (1997) to determine the malting quality of the samples, which is summarised as malting quality index (MQI). For expression analyses induction of trait connected genes was warranted by this standardised micromalting process and performance of a temporally staggered procedure for sampling during the malting process (Fig. 1D-F). The whole micromalting is a 192 hours spanning process. Taking into account the time course of enzyme activities within the malted grains by RT-PCR analysis of key enzymes for starch mobilization and cell wall degradation, initiated during GABI-Seed, the points of time two, 24 and 72 hours have turned out as very important. Sampling was completed by inclusion the point of time 2 hours after start. That way in total 2.190 samples were generated and the respective 730 malt analyses performed. Malted material and quality data represent the basis for the functional association of genes and malting quality traits focussed in GABI-Malt.

The material of the reference variety panel is used by the project partners for expression analyses by microarray hybridization experiments as well as the development and utilization of SNP markers in association studies. The ALEXIS x STEINA DH population and QTL-NILs are the main objectives at the LfL. The combined approach will lead to the saturation and validation of the QTL map, aiming to the development of efficient functional markers of malting quality to provide future barley breeding programs.

Expression analysis by cDNA-AFLP techniques

Expression patterns derived from malted grains taken at defined points of time from the micromalting process were analysed by cDNA-AFLP technology, an efficient and sensitive method to display and compare whole transcript profiles of inducible characters or particular developmental stages and especially to identify even rarely transcribed mRNAs (Bachem *et al.*, 1996), which are not at all spotted on arrays. The potential of this technique for the investigation of malted barley was already validated during the first phase of GABI in the project GABI-Seed (Mikolajewski *et al.*, 2002).

Based on the cDNA-AFLP technique and the QTL map of the ALEXIS x STEINA DH population, a novel approach: MAGS – marker assisted genotype screening was developed likewise during GABI-Seed to assign differential TDFs to chromosomal regions carrying QTLs for malting quality (Herz *et al.*, 2003). Two contrasting pools each consisting of ten plants carrying the favourable marker alleles for a QTL interval and the respective phenotypic performance are compared to each other on the gene expression level. For the analysis of differ-

○ C.W.B. Bachem *et al.* **Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development.** Plant J. (1996) 9, 745-753. ○ L. Hartl *et al.* **Molekulargenetische Lokalisierung von QTL für die Malzqualität der Gerste.** Bericht über die Arbeitstagung 2000 der Vereinigung österreichischer Pflanzenzüchter, 21. bis 23. November 2000, 117-122. ○ M. Herz *et al.* **Identification of candidate cDNAs correlated to malting quality of barley by means of cDNA-AFLP analysis and differential genotype pooling.** Proceedings of the 29. EBC-congress, Dublin 2003 ISBN 90-70143-22-4 (10 pp.) ○ **MEBAK Brautechnologische Analysemethoden.** Methodensammlung der Mitteleuropäischen Brautechnologischen Analysenkommission (1997), Selbstverlag der MEBAK, Freising-Weihenstephan, Deutschland. ○ S. Mikolajewski *et al.* **Untersuchung differentieller Genexpression im Verlauf der Vermälzung von Gerstenkörnern – ein Vergleich von Brau- und Futtergerstensorten mittels cDNA-AFLP-Technik.** Vortr. Pflanzenzüchtung (2002), 54, 405-408.

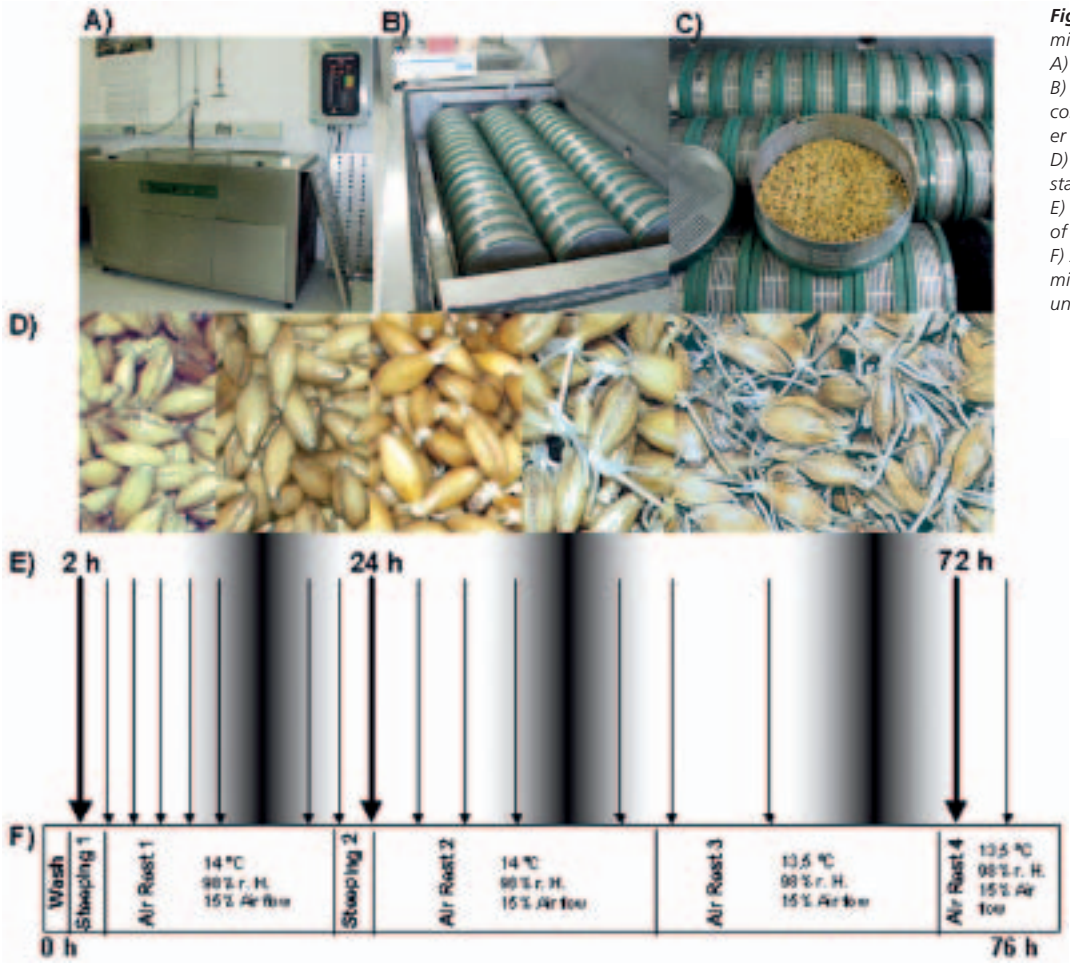


Fig. 1: *Chart of the micromalting process. A) Micromalting machine, B) Configuration of malting containers, C) Open container with malted material, D) Barley grains in selected stages of germination, E) Bold arrows mark points of time for sampling, F) According steps of the micromalting procedure until 76 h.*

ential transcripts for all QTL hot spots such contrasting MAGS pools (termed "+", carrying the genetic interval from ALEXIS, and "–" equivalent the adequate interval of STEINA) were designed from selected lines of the segregating DH progeny. As expected pools from the DH population are overlapping for the QTL regions on chromosomes 1H, 4H, 5H and 6H. However due to its bulked character this time saving method allows a preliminary association between QTL map and expression data. Up to now transcript profiles of the corresponding MAGS pools and the parents ALEXIS and STEINA, generated from malted grains taken at the points 2 h or 24 h of the micromalting process were analysed by the cDNA-AFLP method using 32 different primer combinations. The figure 2A demonstrates exemplarily the polymorphic TDF002 excised on a transcript profile generated for the point of time 24 h. TDFs which show the same segregating behaviour as the parents are defined as assignable to a QTL hot spot. For example the distribution of the fragment TDF002 as the parent STEINA would indicate association to 1H or 6H. By separate assessment of such relations in all samples of the 32 transcript profiles for the point of time two hours 36 differential TDFs could be assigned to the QTL hot spot on chromosome 1H, 27 to that on chromosome 4H, 41 to 5H, and 45 to the QTL hot spot on chromosome 6H. For samples taken at 24 hours of the malting process 41 TDFs could be assigned to the QTL hot spot on chromosome 1H, 27 to that on 4H, 35 to that on 5H and 55 to the QTL hot spot on chromosome 6H.

To investigate the influence of the particular QTL regions

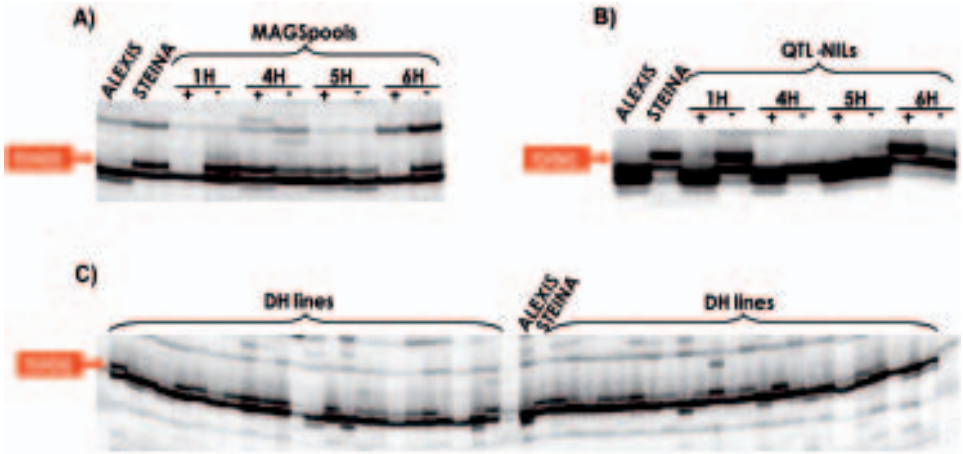


Fig. 2: *Expression analyses concerning malting quality using cDNA-AFLP techniques. Details of cDNA-AFLP profiles of malted barley grains taken at 24 h of micromalting A) of corresponding MAGS pools for the QTL hot spots on chromosome 1H, 4H, 5H, 6H and the parents ALEXIS and STEINA, B) of corresponding NILs for the QTL hot spots on chromosome 1H, 4H, 5H, 6H and parents, and C) of a subpopulation of 48 DH lines from the ALEXIS x STEINA mapping population and parents. Marked by an arrow is the polymorphic TDF002.*

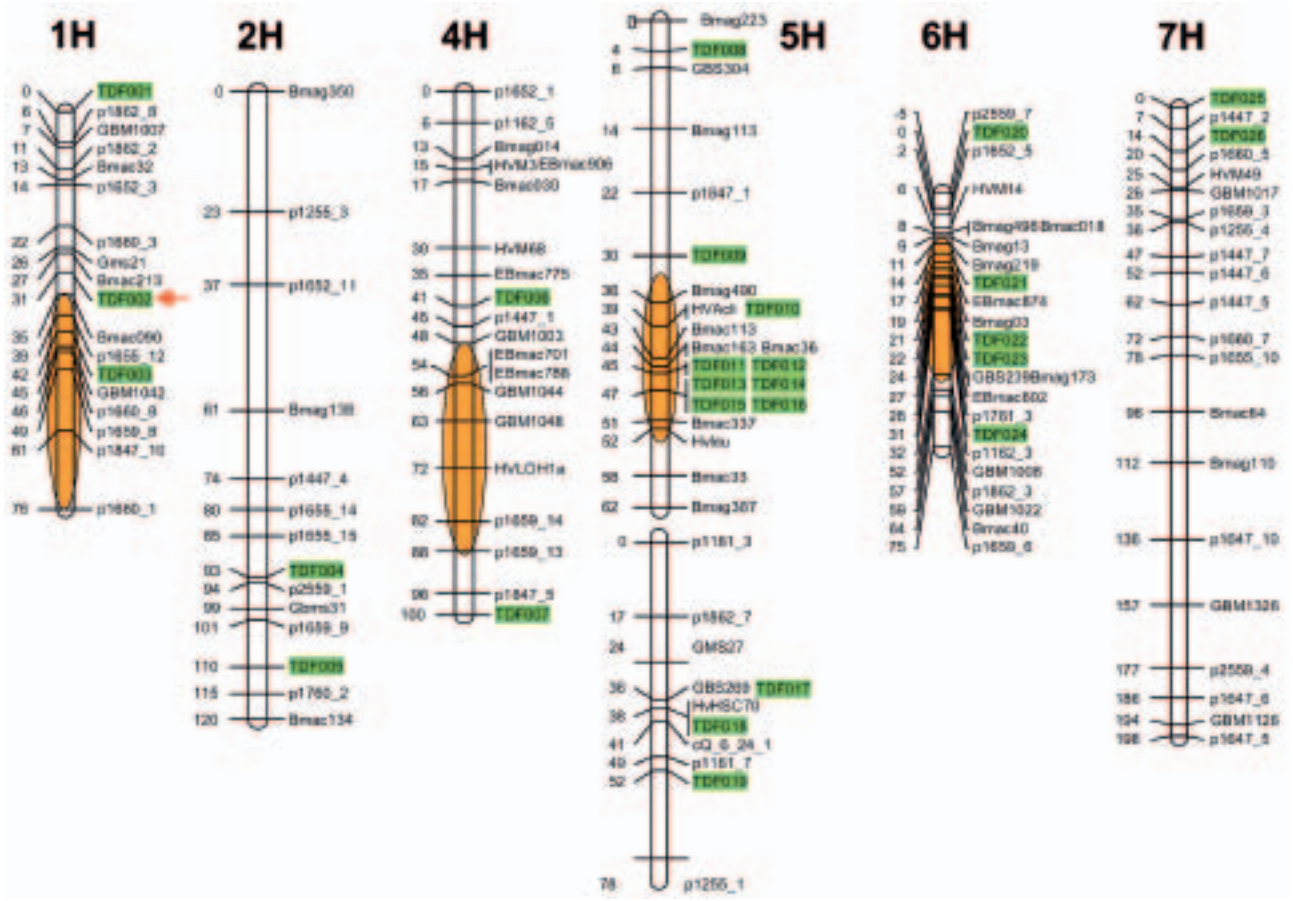


Fig. 3: *Linkage map combining a subset of genetic and transcriptome markers. The complete dataset will allow an assignment of additional markers to chromosomes and QTL intervals. Pictured are barley chromosomes 1H, 2H, 4H, 5H, 6H and 7H. Hitherto no linkage group could be matched to chromosome 3H due to less polymorphic markers within this genomic region, engendered by identity by decent. QTL hot spots are surrounded, transcriptome markers are highlighted by green color, the example TDF002 is marked by an arrow.*

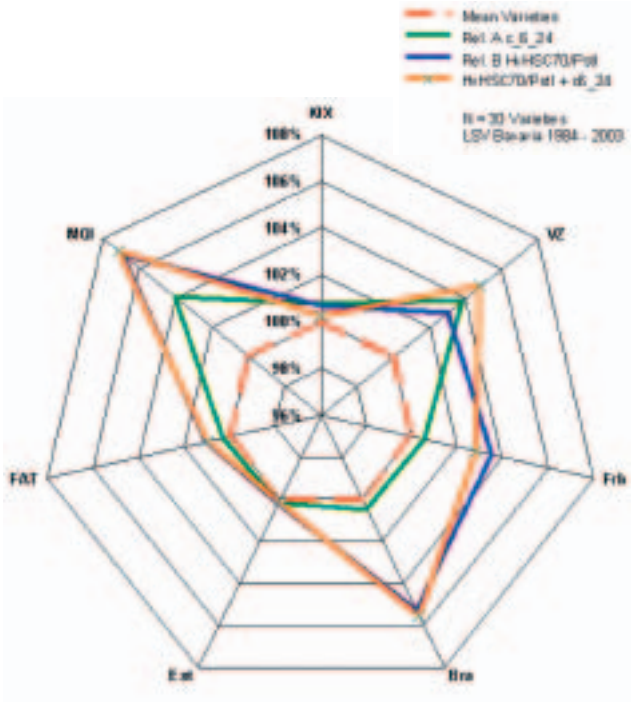


Fig. 4: *Effects of the markers HvHSC70/PstI and c6_24 on parameters for malting quality of barley varieties. Shown is the relative alteration of the average value of varieties selected by those markers compared to the average value of all varieties of this study (N=30). Average value from 8 environments. (KIX: Kolbachindex; VZ45: VZ45°C; Frb: Friability Bra: Brabender; Ext: Malt extract; FAT: Final attenuation; MQI: index of malting quality).*

on malting quality corresponding pairs of NILs for each QTL hot spot were micromalted in 2006. Samples of NILs and parents were taken at 2 h, 24 h and 72 h of the micromalting process. Expression analysis of NILs is currently in progress by cDNA-AFLP, and will be performed as well by subtractive suppression hybridisation (SSH) to develop QTL-SSH libraries. Currently ds cDNA of corresponding QTL NILs and the parents ALEXIS and STEINA is generated from malted grains taken at the points of time 24 h and 72 h of the micromalting process, and the transcript profiles by cDNA-AFLP technique generated by the same 32 primer combinations used for the MAGS approach are available. As a current result the figure 2B demonstrates the polymorphic TDF002 excised from the transcript profile generated for the point of time 24 h. Again the symbol "+" means carrying the genetic interval from ALEXIS and "–" equivalent the interval of STEINA, and TDFs which show the same segregating behaviour as the parents are defined as assignable to a QTL hot spot. Based on the distribution of the fragment TDF002 within the NILs it indicates the assignment as STEINA allele to the QTL on chromosome 1H.

Mapping of the barley transcriptome concerning malting quality

To validate and localise the indicated candidate TDFs, in parallel transcriptome mapping regarding malting quality is initiated. Up to now for the transcriptome mapping a subpopulation of 48 DH lines from the QTL mapping population was used to integrate differential TDFs generated by the cDNA-AFLP method into the existing linkage map. Transcript profiles of the single lines were generated using the same 32 cDNA-AFLP primer combinations as above. Fig. 2C shows exemplarily the distribution of TDF002 within the subpopulation. Altogether from the 32 profiles 197 differential TDFs have already been analysed in the subpopulation using samples from the point of time 24h during malting. Out of these 26 TDFs could already be integrated into the existing linkage map of the ALEXIS x STEINA cross (Fig. 3). A number of cDNA-AFLP markers are located closely to or within the most interesting QTL hot spots on the map. For example the locus TDF002 is confirmed nearby the QTL hot spot on chromosome 1H. The planned extension of the linkage map of TDFs using the complete data set of more than 600 markers and the entire ALEXIS x STEINA DH population of 134 DH lines will clarify the collocation of the transcriptome markers to the four QTL intervals of particular interest. Statistical analyses like the estimation of the R^2 value of the already mapped transcriptome markers has demonstrated that single TDFs explain up to 40% of the variance of malting quality parameters across several environments. These markers are the most valuable candidates for sequence analysis and genetic marker development. Until now 16 fragments were cloned, sequenced and PCR primers were developed to transform the results in functional genetic markers for the selection of barley accessions concerning their malting quality. Two examples of currently developed functional markers derived from differential TDFs demonstrate the success of this strategy. The CAPS marker HvHSC70/PstI and the InDel marker c6_24. are already retroceded from expression to genomic level and are useful for the selection of barley accessions like the diagram shown in figure 4 illustrates for 30 barley accessions from the Bavarian regional variety trials 1983-2004. Varieties selected by these markers give evidence that there is a clear correlation to the malting quality parameters VZ45°C, brabender, friability and the resulting MQI. An improvement up to 7% compared to the average of all varieties is reached.

Brassica napus: Allelic Diversity in Candidate Genes Molecular and Functional Characterisation using Genomics Resources

BRIDGE
Rapeseed

Wolfgang Ecke 1; Martin Lange 1; Renate Schmidt 2; Katrin Bach 2; Eva Dietrich 2; Rod Snowdon 3; Katrin Link 3; Milena Ouzunova 4; Frank Breuer 4; Dieter Stelling 5; Detlef Hauska 5; Gunhild Leckband 6; Felix Dreyer 6; Jutta Detering 7; Peter Duchscherer 7; Werner Horn 8 and Stine Tuvesson 8

1 Department of Crop Sciences, Georg-August-University Göttingen, Göttingen
2 Max Planck Institute of Molecular Plant Physiology, Golm
3 Department of Plant Breeding, Justus-Liebig-University Giessen, Giessen
4 KWS SAAT AG, Einbeck
5 Deutsche Saatveredelung AG, Lippstadt
6 Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Hohenlieth
7 RAPS GbR, Grundhof
8 SW Seed Hadmersleben GmbH, Hadmersleben

Introduction

Rapeseed is the major oil crop in Germany and one of the most important oilseed crops worldwide. The oil is used in human nutrition and as raw material for industrial purposes and the meal is used in animal feeding. A further improvement of rapeseed as an economically optimised and ecologically well-adapted production system will depend on a better utilisation of the genetic diversity available within this crop plant.

The last two decades have seen a tremendous increase in knowledge of plant molecular biology. Many genes for enzymes catalysing the reactions of the biochemical pathways plants use to synthesise the special primary and secondary metabolites that are the products of crop plants have been isolated and characterised. In addition, genes controlling the interactions of plants with biotic stress factors like fungal and bacterial diseases have been cloned and analysed.

With the information on basic gene structure and function in place the objective of this project was to analyse the specific allelic diversity that is the basis for the improvement of rapeseed as a crop plant in two types of genes: (1) in genes encoding key enzymes and regulatory proteins involved in lipid biosynthesis as candidate genes for oil content, and (2) in resistance gene analogues as candidates for genes involved in disease resistance. Allelic diversity in these genes was to be characterized at the DNA sequence level and for the candidate genes for oil content also at the transcriptional level. Furthermore, phenotypic effects of the allelic diversity observed at the molecular level were analysed in order to identify loci that are involved in the control of oil content and resistance to *Leptosphaeria maculans* (anamorph *Phoma lingam*: blackleg disease).

Partners and Subprojects

Eight partners collaborated to carry out the following three interrelated subprojects that constituted the entire project:

I. Analysis of genetic diversity of candidate genes for oil content and oil quality in rapeseed (*Brassica napus* L.).
Department of Crop Sciences,
Georg-August-University Göttingen;
Max Planck Institute of Molecular Plant Physiology, Golm

II. Development of SNP markers for high-throughput genotyping of resistance gene candidates in rapeseed (*Brassica napus* L.)
Department of Plant Breeding,
Justus-Liebig-University Giessen

III. Phenotypic and genotypic characterisation of the genetic diversity of oil content, oil quality and disease resistances in winter rapeseed material
Deutsche Saatveredelung AG, Lippstadt (DSV)
KWS SAAT AG, Einbeck (KWS)
Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Hohenlieth (NPZ)
SW Seed Hadmersleben GmbH, Hadmersleben (SH)
RAPS GbR, Grundhof

Results

Analysis of genetic diversity in candidate genes for oil content

Based on information from the literature and public databases nine candidate genes were selected (Table 1). Included were genes directly involved in fatty acid biosynthesis along with genes involved in pathways providing substrates for fatty acid biosynthesis or regulating such pathways. Using a BAC library of the rapeseed variety 'Express' and *Arabidopsis* genes as probes a total of 44 loci was isolated from the polyploid rapeseed genome for the nine candidate genes, corresponding to three to nine loci per *Arabidopsis* single copy gene. Se-

quence comparisons among the homeologous rapeseed loci and with the corresponding Arabidosis sequences showed conserved exon/intron structures at 32 of the loci; at the remaining 12 pseudogenes were found.

The allelic diversity of 29 loci of the nine candidate genes was analysed in a core set of 57 genetically diverse rapeseed genotypes, including old and new winter rapeseed varieties as well as resynthesized and other exotic genotypes. Between one and five locus-specific primer pairs were developed for each locus based on the sequence alignments of the homeologous rapeseed loci. All corresponding amplicons were first screened for polymorphisms in eight genotypes representing the genetic diversity of the core set. Four loci failed to show any polymorphisms in the prescreening; for the remaining 25 loci the amplicon showing the highest number of polymorphisms was sequenced in the full core set. This resulted in the identification of 332 SNP and 75 InDel markers with a range of one to 47 markers per locus (Table 1). Generally, the number of haplotypes defined by the markers at a locus was smaller than the number of markers. This was due to the fact that many markers at a locus were often completely linked. This indicates that, while not being absolute, linkage disequilibrium is quite high in the rapeseed genome at a scale of several hundred base pairs. Using five different mapping populations it was possible to map 18 of the candidate gene loci on 10 different linkage groups of the genetic map of rapeseed (Fig. 1)

For each of 28 loci with an intact exon-intron structure two locus-specific primer pairs complementary to exon sequences were established in oder to analyse the allelic diversity of the candidate genes at the transcriptional level. Semiquantitative expression analysis by RT-PCR using RNA from leaf, shoot, root, flower, and five seed developmental stages showed that all loci were expressed in different tissues including seeds. Thirteen loci were further analysed in 11 genotypes from the core set, selected to represent the genetic diversity present in this material. Expression analysis with RNA from seeds harvested 30 days after flowering showed significantly different expression levels in one or several genotypes for eight of the loci when compared to the remainder of the genotypes analysed. This clearly indicates that in addition to allelic diversity at the DNA level, many of the candidate gene loci also show diversity at the transcriptional level.

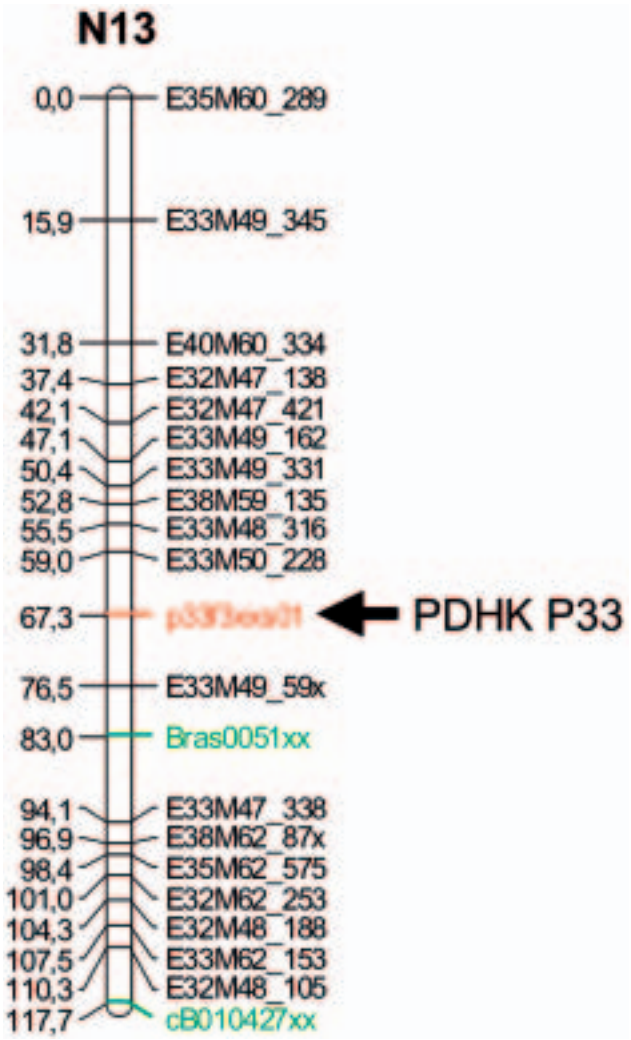


Fig. 1: *Genetic mapping of the pyruvate dehydrogenase kinase locus P33 on linkage group N13 of Brassica napus. The PDHK locus P33 was mapped in a doubled haploid population from the cross 'Express' x '1012-98' using p33f3xxs01, a marker developed based on the comparative sequencing of this locus in the core set.*

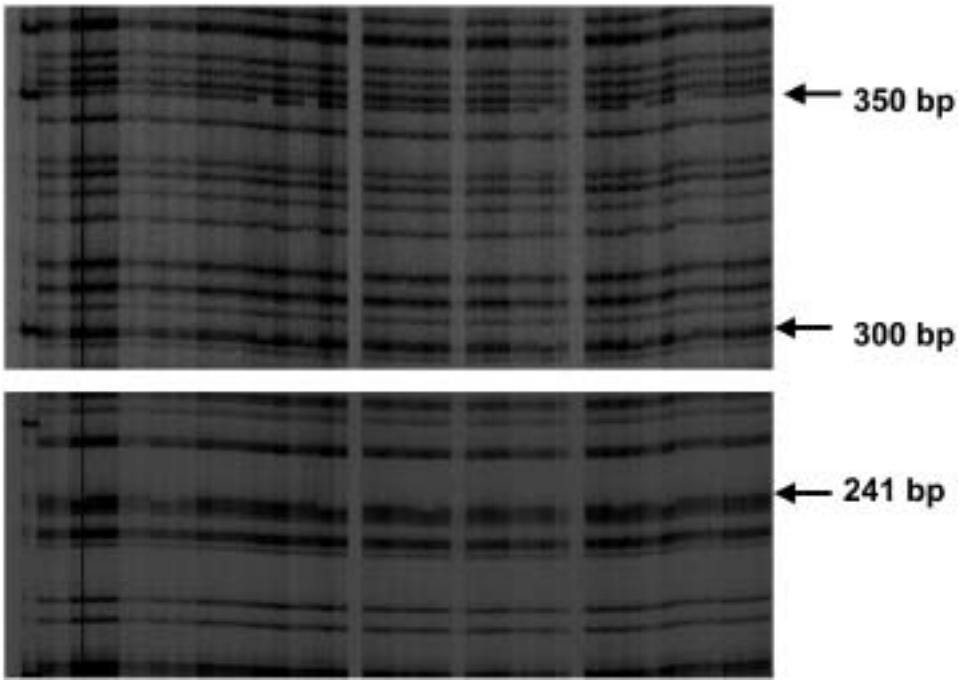


Fig. 2: *Identification of SNPs in the resistance gene candidate RGC MB73-12G by BESS-T. Three SNPs were identified between core set genotypes in the gene for the ethylene-salicylic acid-inducing protein RGC MB73-12G at 241, 300 and 350 bp.*

Isolation and analysis of allelic diversity of resistance gene candidates

For the development of primer pairs specific for resistance gene candidates in rapeseed two different approaches were used. In the first approach, degenerated primers complementary to conserved motifs of resistance genes (NBS, TIR, LRR) were used to amplify and clone resistance gene analogous sequences (RGAs) from one blackleg resistant and one susceptible rapeseed variety. After sequencing of the resulting clones, which ranged in length from 100 to 1000 bp, specific primer pairs could be designed for 36 RGAs. The remaining sequences belonged to large gene families with high sequence identities, effectively preventing the design of specific primer pairs.

In a second approach 277 publicly available ESTs from genes that had shown overexpression in rapeseed after *Leptosphaeria maculans* infection were analysed. About 30% of these genes displayed sequence similarities to known defence, stress and resistance genes. After sequence comparisons an additional 47 primer pairs that were not complementary to known consensus sequences of resistance genes could be designed, giving a total of 83 specific primer pairs.

Allelic diversity was analysed in amplicons derived from 79 of the primer pairs in the core set of rapeseed genotypes. Because of the large number of primer pairs that were developed, the sequencing-independent approach of base-excision sequence-scanning (BESS-T) was used to detect polymorphisms (Fig. 2). In total, 58 SNPs were detected in 32 of the amplicons, with a range of one to five SNPs per amplicon. Two of the loci corresponding to amplicons from two genes, RGA10 and RGC3, could be mapped on linkage groups N3 and N16 of the rapeseed genetic map.

Analysis of phenotypic effects of the candidate gene loci

The core set was evaluated for oil content and blackleg resistance in field trials at eight locations over three years. In addition, blackleg resistance was also analysed by artificial inoculation in greenhouse tests. With these data the phenotypic effects of the allelic diversity observed at the DNA level were evaluated by association analysis using the programs TASSEL and PowerMarker.

A total of 14 loci showed significant ($p < 0.05$) association with resistance in at least one of the blackleg disease datasets with one or two markers in a single-locus F-test. For 11 loci this association could be confirmed by haplotype trend regression. For two of these loci, RGA10 and RGC3, it was possible to develop locus specific primer pairs as the basis for the development of allele specific SNP assays.

Among the candidate genes for oil content, seven loci of five genes were found to show significant association to oil content in the core set using TASSELs GLM procedure. The phenotypic effects at these loci ranged from 0.58 – 1.55% shift in oil con-

tent with a mean of 1%. The positive association at three of these loci could be confirmed in a verification set of around 200 rapeseed genotypes that were genotyped using SNP and InDel markers derived from the comparative sequencing in the core set.

Conclusions

For the first time the allelic diversity of a large number of candidate gene loci for oil content and disease resistance has been analysed in a broad set of rapeseed genotypes. By association analysis a number of loci could be identified that are actually involved in the genetic control of oil content and blackleg resistance, respectively. Locus-specific primer pairs and marker assays were developed that will allow breeding companies to characterize their breeding materials and can be used in marker assisted selection. During the project BESS-T was found to be an efficient method to screen many amplicons in a large number of genotypes for polymorphisms, although the development of locus-specific primer pairs and marker assays based on this approach requires fine-tuning for highly-duplicated sequences in the polyploid rapeseed genome. In contrast, the strategy to isolate all rapeseed loci corresponding to *Arabidopsis* candidate genes from a BAC library has allowed a very efficient development of locus specific primer pairs for an analysis of allelic diversity by comparative sequencing and RT-PCR.

Tab. 1: *Number of polymorphisms and haplotypes detected in the core set in loci of candidate genes for oil content*

Gene coding for	Locus	No. of		
		SNPs	InDels	Haplotypes
Pyruvate kinase (PK)	K14	2	0	3
	K48	4	0	3
	K141	8	0	5
Pyruvate dehydrogenase (PDH)	H40	5	5	4
	H81	14	3	3
	H29	16	2	6
	H71	9	0	3
Pyruvate dehydrogenase kinase E1α	P12	7	1	4
(PDHK)	P133	19	8	2
	P33	21	5	3
	P57	17	0	3
Biotin carboxylase (CAC2)	A24	22	3	2
	A22	11	2	2
	A78	15	3	5
α-Carboxyltransferase (CAC3)	C10	17	0	4
	C28	14	2	3
β-Ketoacyl-ACP-synthase III (KAS III)	S13	20	4	4
Biotin carboxyl carrier protein (BCCP2)	B51	2	1	3
	B62	17	11	6
	B114	24	7	3
	B173	24	7	4
Diacylglycerol acyltransferase (DAGAT)	D30	1	0	2
	D120	37	10	4
Wrinkled (WRI)	W80	1	1	3
	W102	5	0	4
Sum		**332**	**75**	**88**

Reducing Fusarium Toxins in Wheat Through Genomics – Guided Strategies

WHEAT
Cereals

Erhard Ebmeyer 4, Lorenz Hartl 1, Viktor Korzun 4, Thomas Miedaner 7, Frank Ordon 2, Michael Schmolke 6, Jörg Schondelmaier 5, Patrick Schweizer 3, Jens Weyen 5, Gerhard Wenzel 6

1 Bavarian State Research Center for Agriculture, Freising
2 Federal Centre for Breeding Research on Cultivated Plants, Quedlinburg
3 Institute of Plant Genetics and Crop Plant Research, Gatersleben
4 Lochow-Petkus GmbH, Bergen
5 Saaten-Union Resistenzlabor GmbH, Leopoldshoehe
6 Technical University Munich, Freising
7 University of Hohenheim, State Plant Breeding Institute, Stuttgart

Summary

Fusarium head blight (FHB) of wheat is a major threat to sustainable wheat production worldwide. Infection with Fusarium leads to severe yield losses and mycotoxin contamination in infected grain that is deleterious to animals and humans. Growing of resistant cultivars is the best means to reduce the threat of mycotoxin contamination of cereal food and feed. Genetic variation for FHB resistance in wheat was described in several gene pools. In spring wheat highly effective quantitative trait loci (QTL) were found in exotic stocks from China and Brazil. In European winter wheat, the genetic basis seems to be much more complex. Resistance to FHB in wheat is of quantitative, oligo- or polygenic nature and therefore it is a tedious and time-consuming task for breeders to develop cultivars adapted to local environmental conditions with a high level of FHB resistance. In this project molecular marker techniques and functional genomics as well as mutagenesis and conventional breeding strategies will be used to develop FHB resistant wheat germplasm.

Objectives

The overall goal of this project is to develop FHB resistant wheat germplasm to minimize mycotoxin contamination. This project concentrates on marker assisted selection and subsequent use of selected lines in pre-breeding for FHB resistance in wheat. A comprehensive study of known FHB resistance sources from Europe, China and Brazil will be conducted to identify genes and metabolites involved in FHB resistance and understand where in the wheat genome FHB resistance genes reside. Additionally, lasting resources will be created including genomic tools to understand which genes and protein products are altered during infection and how these changes in metabolism can make wheat plants susceptible or resistant to FHB. From the Canadian side (Daryl J. Somers, Agriculture and Agri-Food Canada) the central contributions are in fields of transcriptomics, proteomics and metabolomics in addition to the German activities concentrating on molecular breeding and mapping. The German project is structured into three modules:

Module 1

Mapping new adapted winter wheat populations and validation of previously identified QTL: Only a few winter wheat populations of European origin have been mapped. New effective resistance sources from German breeding companies are available that have not yet been genetically analysed.

For the validation of individual QTL, sets of truncated QTL lines based on mapping information will be developed via backcross and selfing generations and phenotyped in field trials. This will allow determining resistance of the donor in general and the effect of specific QTL. The most prominent QTL regions will be enriched by diagnostic molecular markers (fine mapping).

Module 1	Module 2	Module 3
Molecular characterization of adapted resistance sources	**Introgression breeding of exotic resistance sources**	**Functional Genomics**
Topics		
• Production: TILLING population • QTL mapping & fine mapping • Verification of QTL effects • Association mapping	• Linkage drag • Dissection of exotic QTL • Effect on fungal populations • New resistance sources	• Expression profiling • Establishment of VIGS • Use of TILLING • Mapping of candidate genes
Partners		
LfL, TUM, UHOH, SU-RL	UHOH, LP, BAZ	TUM, IPK

Bavarian State Research Center for Agriculture (LfL); Federal Centre for Breeding Research on Cultivated Plants (BAZ); Institute of Plant Genetics and Crop Plant Research (IPK); Lochow-Petkus GmbH (LP); Saaten-Union Resistenzlabor GmbH (SU-RL); Technical University Munich (TUM); University of Hohenheim (UHOH)

○ H. Buerstmayr *et al.* **Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat.** II.Resistance to fungal penetration and spread. Theor Appl Genet (2003) 107, 503–508 ○ T. Miedaner: **Plant breeding as a tool for reducing mycotoxins in cereals.** In: Barug D, H van Egmond, R López-García, R van Osenbruggen, A Visconti (Eds) Meeting the Mycotoxin Menace (2004). Wageningen Acad. Publ., NL. ○ T. Miedaner *et al.* **Stacking quantitative trait loci (QTL) for Fusarium head blight resistance from non-adapted sources in an European elite spring wheat background and assessing their effects on deoxynivalenol (DON) content and disease severity.** Theor Appl Genet (2006) 112, 562-569. ○ M. Schmolke *et al.* **Molecular mapping of Fusarium head blight resistance in the winter wheat population Dream/Lynx.** Theor Appl Genet (2005) 111, 747-756. ○ F. Wilde *et al.* **Comparison of phenotypic and marker-based selection for Fusarium head blight resistance and DON content in spring wheat.** Mol Breed (2007) DOI 10.1007/s11032-006-9067-5.

Association mapping of adapted FHB resistance donors in winter wheat: Winter wheat cultivars and lines from an international set with documented FHB resistance along with highly susceptible lines will be characterised by molecular marker analysis spread across the whole genome. Reaction to FHB of the lines will be assessed in field trials by inoculation. Identified linkage blocks that can be assigned to FHB resistance will be fine mapped and verified by SSR markers.

Development of a TILLING resource for wheat and identification of variant alleles associated with FHB resistance: The use of targeting induced local lesions in genomes (TILLING) is a high-throughput reverse genetics method that combines random chemical mutagenesis with PCR based screening of candidate genes. A mutagenised hexaploid wheat population (donor 'Dream') will be generated to provide a resource of trait diversity to researchers. This population will allow identification of allelic series of known candidate genes by TILLING.

Module 2

Introgressing non-adapted sources into elite wheat material might result in negative side effects, like linkage drag. UHOH and LP have already developed large elite winter wheat populations with exotic resistance QTL on chromosomes 3BS and 5A from Chinese sources. The introgression of these QTL were monitored by linked SSR markers across several backcross (BC) and selfing generations. In this module we aim for:

(1) Dissection of QTL from non-adapted sources: Haplotypes with sub-QTL will be obtained by selective high-resolution mapping and analysed for their FHB resistance and DON content in multi-locational field trials.

(2) Stacking of QTL and estimation of side-effects (linkage drag) of exotic QTL on agronomic traits: BC lines with none, one or two target QTL will be tested in large-plot trials to estimate their effects on FHB resistance, other diseases, quality traits, and yield components in a similar adapted winter wheat background.

Risk analysis for introgression of highly effective exotic QTL: Using highly effective resistance sources might change the fitness of Fusarium populations. *F. graminearum* is a genetically highly flexible pathogen and might be able to unspecifically adapt to higher resistance levels of wheat, when the respective varieties are grown on large acreages. We, therefore, analyse the effect of the two aforementioned QTL individually and in combination on parasitic fitness of *F. graminearum* populations. Artificially established fungal populations of different complexity, aggressiveness, mycotoxin type (DON vs. NIV) and mycotoxin production will be field inoculated on BC lines with none, one or two target QTL, re-isolated, and genotyped by molecular markers to monitor changes in the composition of the original populations caused by host genotype.

Screening of genetic resources (Triticum spp.) and marker-based introgression: For broadening the genetic basis of FHB resistance germplasm of diploid and tetraploid wheat is tested. Parental lines will be screened for polymorphic SSR markers and genes or QTL involved in FHB resistance will be analysed for allelic/haplotype diversity in the above mentioned relatives and included into mapping. Using single seed descent, lines will be advanced to F5 generation in the greenhouse and repetitively tested for resistance to FHB. Based on skeleton maps and phenotypic data QTL analyses will be performed. Marker saturation of QTL regions will be conducted by AFLP. In parallel, crosses of resistant genotypes to durum wheat will be carried out.

Module 3

The mechanisms of quantitative resistance to FHB are largely not known at the molecular level. A promising approach to discriminate between different resistance mechanisms would be the comparison of host transcriptome with respect to absence/presence in near-isogenic lines of important genes/alleles at the most important loci conferring quantitative resistance. We will carry out a medium-throughput approach for the identification of gene function in Fusarium-attacked resistant wheat spikes. The approach is split up in several parts:

- Expression profiling in BC lines differing in absence/presence of important QTL for resistance on chromosomes 3BS and 5A
- Functional analysis by inoculation in the greenhouse
- Virus induced gene silencing (VIGS) of candidate genes whose expression is correlated to QTL-mediated resistance
- Sequence haplotypes of functionally verified candidate genes will be determined among wheat lines carrying/not carrying resistance QTL

Go to Content Page

Appendix

Improved interaction between plant scientists is a core aspect of the GABI programme. The programme provides the framework for contacts and communication. Scientific publications are the key output of each research project. Not only GABI project proceedings were released in the past three years, also many third party publications took advantage of GABI resources and technologies.

Go to Content Page

Authors Index

Authors Index

Authors Index

R

S

Authors Index

GABI Publications

GABI Project Publications

A. Ballvora, A. Jocker, P. Viehover, H. Ishihara, J. Paal, K. Meksem, R. Bruggmann, H. Schoof, B. Weisshaar and C. Gebhardt
Comparative sequence analysis of Solanum and Arabidopsis in a hot spot for pathogen resistance on potato chromosome V reveals a patchwork of conserved and rapidly evolving genome segments
2007, BMC Genomics 8, 112

A. Baudry, M. Caboche and L. Lepiniec
TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*
2006, The Plant Journal 46 (5), 768-779

D. Bellin, B. Schulz, T.R. Soerensen, F. Salamini and K. Schneider
Transcript profiles at different growth stages and tap-root zones identify correlated developmental and metabolic pathways of sugar beet
2007, J. Experimental Botany 58 (3), 699-715

O. Bläsing, Y. Gibon, M. Günther, M. Höhne, D. Osuna, O. Thimm, W.-R.Scheible, R. Morcuende and M. Stitt
Sugars and Circadian Regulation make Major Contributions to the Global Regulation of Diurnal Gene Expression in Arabidopsis
2005, Plant Cell 17, 3257-3281

C.A. Bormann, A.M. Rickert, R.A. Ruiz, J. Paal, J. Lübeck, J. Strahwald, K. Buhr and C. Gebhardt
Tagging quantitative trait loci for maturity-corrected late blight resistance in tetraploid potato with PCR-based candidate gene markers
2004, Mol Plant Microbe Interactions 17, 1126-1138

R. Bruggmann, A.K. Bharti, H.Gundlach, J. Lai, A.C. Pontaroli, G. Haberer, G. Fuks, C. Du, F. Wei, S. Young, C. Raymond, M. Estep, R. Liu, J.L. Bennetzen, A.P. Chan, P.D. Rabinowicz, J.Quackenbush, W.B. Barbazuk, R.A. Wing, S. Rounsley, B. Birren, C. Nusbaum, K.F.X. Mayer and J. Messing
Uneven Chromosome Contraction and Expansion in the Maize Genome
2006, Genome Research 16 (10), 1241-1251

L. Camus-Kulandaivelu, J.-B. Veyrieras, D. Madur, V. Combes, M. Fourmann, S. Barraud, P. Dubreuil, B. Gouesnard, D. Manicacci and A. Charcosset
Maize adaptation to temperate climate: relationship between population structure and polymorphism of Dwarf8 gene
2006, Genetics 172, 2449-2463

F. Carrari, D. Coll-Garcia, N. Schauer, A. Lytovchenko, N. Palacios-Rojas, I. Balbo, M. Rosso M and A.R. Fernie
Deficiency of a Plastidial Adenylate Kinase in Arabidopsis Results in Elevated Photosynthetic Amino Acid Biosynthesis and Enhanced Growth
2005, Plant Physiology 137 (1), 70-82

F. Carrari, C. Baxter, B. Usadel, E. Urbanczyk-Wochniak, M.I. Zanor, A. Nunes-Nesi, V. Nikiforova, D. Centero, A. Ratzka, M. Pauly, L.J. Sweetlove and A.R. Fernie
Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior
2006, Plant Physiol 142 (4), 1380-1396

J.M. Cross, M. von Korff, T.Altmann, L. Bartetzko, R. Sulpice, Y. Gibon, N. Palacios and M. Stitt
Variation of enzyme activities and metabolite levels in 24 Arabidopsis thaliana accessions growing in carbon-limited conditions
2006, Plant Physiol 142 (4), 1574-1588

GABI Publications

K.A. Dekker, T.R. Sörensen, B. Weisshaar and H. Saedler
Functional Genomic Tools for Arabidopsis available at the MPIZ
2003, Biotechnologia 4 (63), 19-31

A. Eudes, B. Pollet, R. Sibout, C.T. Do, A. Seguin, C. Lapierre and L. Jouanin
Evidence for a role of AtCAD 1 in lignification of elongating stems of *Arabidopsis thaliana*
2006, Planta 225 (1), 23-39

T. Feilner, C. Hultschig, J. Lee, S. Meyer, R.G. Immink, A. Koenig, A. Possling, H. Seitz, A. Beveridge, D. Scheel, D.J. Cahill, H. Lehrach, J. Kreutzberger and B. Kersten
High-throughput identification of potential Arabidopsis mitogen-activated protein kinases substrates
2005, Molecular & Cellular Proteomics 4 (10), 1558-1568

T. Feilner and B. Kersten
Phosphorylation studies using plantprotein microarrays
In: H. Thiellement, V. Méchin, C. Damerval and M. Zivy (eds.): Plant Proteomics
2007, Methods in Molecular Biology 355, 379-390

R.M. Gaafar, U. Hohmann, and C. Jung
BAC-derived molecular markers for early bolting in sugar beet
2005, Theor. Appl. Genet. Online early publication 110, 1027-1037

S. Glombitza, P.-H. Dubuis, O. Thulke, G. Welzl, L. Bovet, M. Götz, M. Affenzeller, B. Geist, A. Hehn, C. Asnaghi, D. Ernst, H. Seidlitz, H. Gundlach, K. Mayer, E. Martinoia, D. Werck-Reichhart, F. Mauch and A. Schäffner
Crosstalk and differential response to abiotic and biotic stressors reflected at the transcriptional level of effector genes from secondary metabolism
2004, Plant Molecular Biology 54, 817-835

T. Goujon, R. Sibout, B. Pollet, B. Maba, L. Nussaume, N. Bechtold, F. Lu, J. Ralph, I. Mila, Y. Barriere, C. Lapierre and L. Jouanin
A new Arabidopsis thaliana mutant deficient in the expression of O-methyltransferase impacts lignins and sinapoyl esters
2003, Plant Molecular Biology 51, 973-989

G. Haberer, M. Mader, P. Kosarev, M. Spannagl, L. Yang and K.F.X. Mayer
Large-scale cis-element detection throughby analysis of correlated expression and sequence conservation between Arabidopsis and Brassica oleracea
2006, Plant Physiology 142 (4), 1589-1602

G. Haberer, S. Young, A. Bharti, H. Gundlach, C. Raymond, G. Fuks, E. Butler, R. Wing, S. Rounsley, C. Nussbaum, B. Birren, K.F.X. Mayer and J. Messing
Structure and Architecture of the Maize Genome
2005, Plant Physiology, 139, 1612-1624

G. Haberer, T. Hindemitt, B. C. Meyers and K. Mayer
Evolutionary and transcriptional dynamics of duplicated genes in *Arabidopsis thaliana*: Conservation, Sub- and Neofunctionalisation
2004, Plant Physiology, 136, 3009-3022

M. Heckenberger, A.E. Melchinger, H.P. Maurer and M. Frisch
A comprehensive database management system for integrating phenotypic and genomic data in scientific and applied plant breeding programs
2006 (submitted), Euphytica

T. Hindemitt and K.F.X. Mayer
CREDO: a web-based tool for computational detection of conserved sequence motifs in noncoding sequences
2005, Bioinformatics, 21 (23), 4304-4306

GABI Publications

U. Hohmann, G. Jacobs, and C. Jung.
An EMS mutagenesis protocol for sugar beet and isolation of non-bolting mutants
2005, Plant Breeding 124: 317-321

A.K. Jacobs, V. Lipka, R.A. Burton, R. Panstruga, N. Strizhov, P. Schulze-Lefert and G.B. Fincher
An Arabidopsis Callose Synthase, GSL5, Is Required for Wound and Papillary Callose Formation
2003, Plant Cell 15 (11), 2503-2513

M. Jakoby, H.Y. Wang, W. Reidt, B. Weisshaar and P. Bauer
FRU (BHLH029) is required for induction of iron mobilization genes in *Arabidopsis thaliana*
2004, FEBS Letters 577(3), 528-534

C. Jansen, D. von Wettstein, W. Schäfer, K.H. Kogel, A. Felk and F.J. Maier
Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*
2005, Proc. Nat. Acad. Sci. USA 102, 16892-16897

B.H. Junker, C. Klukas and F. Schreiber
VANTED: a system for advanced data analysis and visualization in the context of biological networks
2006, BMC Bioinformatics 7: e109

U. Kanter, B. Usadel, F. Guerineau, Y. Li, M. Pauly and R. Tenhaken
The inositol oxygenase gene family of Arabidopsis is involved in the biosynthesis of nucleotide sugar precursors for cell-wall matrix polysaccharides
2005, Planta 221 (2), 243-254

B. Kersten, G.K. Agrawal, H. Iwahashi and R. Rakwal
Plant phosphoproteomics: A long road ahead
2006, Proteomics 6 (20), 5517-5528

B. Kersten and T. Feilner
Generation of plant protein micorarrays and investigation of antigen-antibody interactions
In: H. Thiellement, V. Méchin, C. Damerval and M. Zivy (eds.): Plant Proteomics
2007, Methods in Molecular Biology 355, 365-378

M. von Korff, H. Wang, J. Léon and K. Pillen
AB-QTL Analyse in Sommergerste
2004, Vorträge Pflanzenzüchtung 64, 58-60

M. von Korff, H. Wang, J. Léon and K. Pillen
Development of candidate introgression lines using an exotic barley accession (*Hordeum vulgare* ssp. *spontaneum*) as donor
2004, Theor Appl Genet 109 (8), 1736-1745

M. von Korff, H. Wang, J. Léon, K. Pillen
Detection and localisation of resistance genes against powdery mildew and leaf rust introgressed from wild barley (*H. vulgare* ssp. *spontaneum*)
2004, Proceedings of the IX. International Barley Genetics Symposium, 20.-26.6.2004, in Brno, Czech Republic, Part II, 686-692

M. von Korff, H. Wang, J. Léon and K. Pillen
AB-QTL analysis in spring barley: I. Detection of resistance genes against powdery mildew, leaf rust and scald introgressed from wild barley
2005, Theor Appl Genet 111, 583-590

M. von Korff, H. Wang, J. Léon and K. Pillen
Die Effekte der Wildgerste auf die agronomische Leistung von zwei Sommergerstenpopulationen
2005, Vorträge Pflanzenzüchtung 67, 35-39

GABI Publications

M. von Korff, H. Wang, J. Léon and K. Pillen
AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (*H. vulgare* ssp. *spontaneum*)
2006, Theor Appl Genet 112, 1221-1231

M. von Korff, H. Wang, J. Léon and K. Pillen
AB-QTL analysis in spring barley: III. Identification of exotic alleles for the improvement of malting quality in spring barley (*H. vulgare* ssp. *spontaneum*)
submitted

A. Kunert , A. Naz, J. Léon and K. Pillen
AB-QTL-Analyse in Winterweizen: Detektion von vorteilhaften Genen für Qualität und Ertrag aus Wildweizenarten
2006, Vorträge Pflanzenzüchtung (in press)

J. Lai, N. Dey, C.-S. Kim, S. Rudd, K.F.X. Mayer, B. Larkins, P. Becraft and J. Messing
Characterization of the maize endosperm transcriptome and its comparison to the rice genome
2004, Genome Research, 14, 1932-1937

J.C. Lein, K. Asbach, Y.Y. Tian, D. Schulte, C. Li, G. Koch, C. Jung, and D. Cai
Resistance gene analogues are clustered on chromosome 3 of sugar beet and co-segregate with QTL for rhizomania resistance
(2007)2006, Genome 50, 61-71

L. Lepiniec , I. Debeaujon, J.M. Routaboul, A. Baudry, L. Pourcel, N. Nesi and M. Caboche
Genetics and biochemistry of seed flavonoids
2006, Annu. Rev. Plant Biol. 57, 405-430

L. Li, J. Strahwald, H.R. Hofferbert, J. Lubeck, E. Tacke, H. Junghans, J. Wunder and C. Gebhardt
DNA variation at the invertase locus invGE/GF is associated with tuber quality traits in populations of potato breeding clones
2005, Genetics 170, 813-821

W. Li, C. Chen, U. Markmann-Mulisch, L. Timofejeva, E. Schmelzer, H. Ma and B. Reiss
The Arabidopsis AtRAD51 gene is dispensable for vegetative development but required for meiosis.
2004, Proc. Nat. Acad. Sci. USA 101 (29), 10596-10601
Y. Li, M.G. Rosso, B. Ulker and B. Weisshaar
Analysis of T-DNA insertion site distribution patterns in *Arabidopsis thaliana* reveals special features of genes without insertions
2006, Genomics 87 (5), 645-652

Y. Li, M.G. Rosso, P. Viehoever and B. Weisshaar
GABI-Kat SimpleSearch: an *Arabidopsis thaliana* T-DNA mutant database with detailed information for confirmed insertions
2007, Nucleic Acids Res. 35 (Database Issue), D874-D878

S. Lorrain, B. Lin, M.C. Auriac, T. Kroj, P. Saindrenan, M.Nicole, C. Balague and D. Roby
VASCULAR ASSOCIATED DEATH 1, a novel GRAM domain-containing protein, is a regulator of cell death and defense responses in vascular tissues
2004, Plant Cell 16, 2217-2232

H.P. Maurer, A.E. Melchinger and M. Frisch
Population genetical simulation and data analysis with Plabsoft
2006 (submitted), Euphytica

GABI Publications

S.A. Mebrate, K. Pillen, H.W. Dehne and E.C. Oerke
Molecular diversity and long distance dispersal of Puccinia triticina isolates collected from Ethiopia and Germany
2006, Journal of Phytopathology (accepted)

J. Messing, A.K. Bharti, W.M. Karlowski, H. Gundlach, H.R. Kim, Y. Yu, F. Wei, G. Fuks, C. Soderlund, K.F.X. Mayer and R.A. Wing
Map-based Analysis of Sequence Composition and Genome Organization of Maize
2004, Proc. Natl. Acad. Sci. USA, 101, 14349-14359

H.W. Mewes, D. Frishman, K.F.X. Mayer, M. Münsterkötter, O. Nobibou, P. Pagel, T. Rattei, M. Oesterheld, A. Ruepp and V. Stümpflen
MIPS: analysis and annotation of proteins from whole genomes – progress report
2006, Nucleic Acids Research, 34 (Database Issue), D169-172

D. Meyer, E. Lauber, D. Roby, M. Arlat and T. Kroj
Optimization of pathogenicity assays to study the Arabidopsis thaliana *Xanthomonas campestris* pv. *campestris* pathosystem
2005, Mol Plant Pathol 6, 327-334

K.A.H. Mohammed, M. von Korff, Wang H, G.A.R. El-Sherbeny, Pillen K and J. Léon
AB-QTL Analysis for Osmotic Adjustment and Relative Water Content of Drought Tolerance in Spring Barley
2004, Proceedings of the IX. International Barley Genetics Symposium, 20.-26.6.2004, in Brno, Czech Republic, Part II, 978-984

K.A.H. Mohammed, M. von Korff, H. Wang, K. Pillen and J. Léon
Improving crop varieties of spring barley for drought tolerance with AB-QTL-analysis
2005, Vorträge Pflanzenzüchtung 67, 46-52

G. Neuhaus, R. Eibach, E. Maul, R. Töpfer, E. Zyprian, A.F. Adam-Blondon and J.M. Martínez Zapater
Exploitation of the Genetic Diversity: an approach to elucidate resistance
Acta Horticulturae, submitted

G. Neuhaus, R. Eibach, E. Maul, R. Töpfer and E. Zyprian
Exploitation of the Genetic Diversity: an approach to elucidate resistance
2006, 9th International Conference on Grape Genetics and Breeding, Udine, Italy 2-6 July, 2006

G. Neuhaus, R. Eibach, E. Maul, R. Töpfer and E. Zyprian
Nutzung der natürlichen genetischen Diversität: ein Instrument zur funktionellen Genomik der Weinrebe
2006, 8. GPZ Tagung Freising-Weihenstephan, 14.03.-16.03.2006

G. Neuhaus
Neue Ansätze bei der Nutzung molekularer Markersysteme in der Züchtung
2006, Mitteilungen des Institut für Rebenzüchtung Geilweilerhof (1)

G. Neuhaus, R. Eibach, E. Maul, R. Töpfer and E. Zyprian
Nutzung der natürlichen Diversität der Weinrebe als Grundlage für eine verbesserte Resistenz in *Vitis vinifera*
2006, Vortr. Pflanzenzüchtg. 70, 2006, 187-189

G. Neuhaus, R. Eibach, E. Maul, R. Töpfer and E. Zyprian
Nutzung der natürlichen Diversität der Weinrebe durch funktionelle Genomik
2006, Forschungsring Deutscher Weinbau, AK Rebenzüchtung, 03.-04.05.2006

GABI Publications

G. Neuhaus, R. Eibach, E. Maul, R. Töpfer and E. Zyprian
Nutzung der natürlichen Diversität der Weinrebe als Grundlage für eine verbesserte Resistenz in *Vitis vinifera*
2005, GPZ-Tagung AG Obst, Gehölze, Reben: "Phänotypische und genotypische Analyse von wichtigen Merkmalen in Zuchtmaterial von Obst, Gehölzen und Reben", Großhansdorf, 18.-19.10.2005

P. Niewiadomski, S. Knappe, S. Geimer, K. Fischer, B. Schulz, U.S. Unte, M.G. Rosso, P. Ache P, U.I. Flügge and A. Schneider
The Arabidopsis Plastidic Glucose 6-Phosphate/ Phosphate Translocator GPT1 Is Essential for Pollen Maturation and Embryo Sac Development
2005, Plant Cell 17(3), 760-75

E. Olmos, B. Reiss and K. Dekker
The ekeko mutant demonstrates a role for tetraspanin-like protein in plant development.
2003, Biochem Biophys Res Commun. 310 (4), 1054-1061

D. Osuna, B. Usadel, R. Morcuende, W.-R. Scheible, Y. Gibon, O-E. Bläsing, O. Thimm, M. Höhne, M. Günter, M.K. Udvardi, B. Kamlage, R. Trethewey and M. Stitt
Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings
2007, The Plant Journal 49 (3), 463-491

X. Pan, Y. Li and L. Stein
Site Preferences of Insertional Mutagenesis Agents in Arabidopsis.
2005, Plant Physiology 137(1), 168-75

H.P. Piepho, A. Büchse and B. Truberg
On the use of multiple lattice designs and a-designs in plant breeding trials
2006, Plant Breeding 125, 523-528

K. Pillen, A. Zacharias and J. Léon
Lokalisation und Transfer von vorteilhaften Genen der Gerstenwildform H. spontaneum mit Hilfe der Advanced Backcross QTL-Analyse
2001, Vorträge Pflanzenzüchtung 52, 91-101

K. Pillen
The exploitation of genetic variation in barley
2004, Proceedings of the Brazilian Oat Breeding Conference in Pelotas, Brazil, 34-45

K. Pillen , A. Zacharias , A. Taheri and J. Léon
AB-QTL Analysis in spring barley and three strategies of QTL verification
2004, Proceedings of the IX. International Barley Genetics Symposium, 20.-26.6.2004, in Brno, Czech Republic, Part II, 314-320

K. Pillen, M. von Korff, H. Wang and J. Léon
AB-QTL analysis in barley: Detection of favorable genes for quality and yield traits from barley wild species and their introgression into the elite germplasm of cultivated barley by means of DNA markers
2005, GABI – The German Plant Genome Program - Progress Report 1999 – 2003/04, ISBN: 3-00-011587-0, 130-131

K. Pillen, M. von Korff, A. Zacharias, A. Taheri and J. Léon
Die AB-QTL-Strategie zur Nutzung von genetischen Ressourcen in Gerste
2005, Vorträge Pflanzenzüchtung 67, 26-34

K. Pillen, A. Kunert, A. Schubert and J. Léon
AB-QTL analysis in wheat: Detection of favorable genes for quality and yield traits from wheat wild species and their introgression into the elite germplasm of cultivated wheat by means of DNA markers
2005, GABI – The German Plant Genome Program - Progress Report 1999 – 2003/04, ISBN: 3-00-011587-0, 160-161

GABI Publications

V. Radchuk, N. Sreenivasulu, R. Radchuk,
U. Wobus and W. Weschke
The methylation cycle and its possible function in barley endosperm development
2005, Plant Mol. Biol. 59, 289-307

J.M. Routaboul, I. Debeaujon, L. Pourcel,
M. Caboche and L. Lepiniec
Flavonoid biosynthesis and accumulation in seed of *Arabidopsis thaliana*
2006, Planta 224 (1), 96-107

S. Rudd, H. Schoof and K.F.X. Mayer
PlantMarkers – a database of predicted molecular markers from plants
2005, Nucleic Acids Research; 33 Database Issue: D628-32

A. Sattarzadeh et al.
Single nucleotide polymorphism (SNP) genotyping as basis for developing a PCR-based marker highly diagnostic for potato varieties with high resistance to Globodera pallida pathotype Pa2/3.
2006, Mol Breed 18, 301-312.

H. Schoof, M. Spannagl, L. Yang, R. Ernst,
H. Gundlach, D. Haase, G. Haberer and K.F.X. Mayer
MIPS plant genome resurces: A framework for integrative and comparative analyses
2005, Plant Physiology, 138, 1301-1309

R. Schwacke, U.I. Flügge and R. Kunze
Plant membrane proteome databases
2004, Plant Physiol. Biochem. 42 (12), 1023-1034.

K. J. Schmid, S. Ramos-Onsins, H. Ringys-Beckstein,
B. Weisshaar and T. Mitchell-Olds
A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome- wide departure from a neutral model of DNA sequence polymorphism
2005, Genetics 169 (3), 1601-1615

R. Sibout, A. Eudes, B. Pollet, T. Goujon, I. Mila,
F. Granier, A. Seguin, C. Lapierre and L. Jouanin
Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenase in Arabidopsis. Isolation and characterization of the corresponding mutants
2003, Plant Physiology 132, 848-860

R. Sibout, A. Eudes, G. Mouille, B. Pollet,
C. Lapierre, L. Jouanin and A. Seguin
Cinnamyl alcohol dehydrogenase C and D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis
2005, Plant Cell 2059-2076

V. Siewers, M. Viaud, D. Jimenez-Teja, I.G. Collado, C.S.
Gronover, J.M. Pradier, B. Tudzynski and P. Tudzynski
Functional analysis of the cytochrome P450 monooxygenase gene bcbot1 of Botrytis cinerea: role in botrydial biosynthesis and impact on virulence in different wild strains
2005, Mol. Plant-Microbe Inter. 18, 602-12

M. Spannagl, O. Noubibou, D. Haase, L.Yang, H. Gundlach,
K. Klee, G. Haberer, H. Schoof and K.F.X. Mayer
MIPSPlantsDB – Plant database resource for integrative and comparative plant genome research
2007, Nucleic Acids Research 35 (Database Issue), D834-D840

N. Sreenivasulu, V. Radchuk, M. Strickert,
O. Miersch, W. Weschke and U. Wobus
Gene expression patterns reveal tissue-specific signaling networks controlling programmed cell death and ABA-regulated maturation in developing barley seeds
2006, Plant J. 47, 310-327

GABI Publications

N. Stein, M. Prasad, U. Scholz, T. Thiel, H. Zhang, M. Wolf, R. Kota, R. Varshney, D. Perovic, I. Grosse and A. Graner
A 1000 loci transcript map of the barley genome – new anchoring points for integrative grass genomics
2007, Theor. Appl. Genet. 114 (5), 823-839

B. Stich, A.E. Melchinger, H.-P. Piepho, M. Heckenberger, H.P. Maurer and J.C. Reif
A new test for family-based association mapping with inbred lines from plant breeding programs
2006, Theor. Appl. Genet. 113, 1121-1130

B. Stich, J. Yu, A.E. Melchinger, H.-P. Piepho, H.F. Utz, H.P. Maurer and E.S. Buckler
Power to detect higher-order epistatic interactions in a metabolic pathway using a new mapping strategy
2006, Genetics 176 (1), 563-570

B. Stich, H.P. Maurer, A.E. Melchinger, M. Frisch, M. Heckenberger, J. Rouppe van der Voort, J. Peleman, A. P. Sørensen and J.C. Reif
Comparison of linkage disequilibrium in elite European maize inbred lines using AFLP and SSR markers
2006, Molecular Breeding 17, 217-226

B. Stich, A.E. Melchinger, M. Frisch, H.P. Maurer, M. Heckenberger and J.C. Reif
Linkage disequilibrium in European elite maize germplasm investigated with SSRs
2005, Theor. and Appl. Genet. 111, 723-730

M. Strickert, U. Seiffert, N. Sreenivasulu, W. Weschke, T. Villmann and B. Hammer
Generalized Relevance LVQ (GRLVQ) with Correlation Measures for Gene Expression Data
2006, Neurocomputing 69, 651-659

I.V. Tetko, G. Haberer, S. Rudd, B.C. Meyers, H.-W. Mewes and K.F.X. Mayer
Spatio-/Temporal Expression Control mediated through intragenic Scaffold Matrix Attachment Regions (S/MARs) in *Arabidopsis thaliana*
2006, PloS Computational Biology, 2 (3) e21

E. Urbanczyk-Wochniak, B. Usadel, O. Thimm, A. Nunes-Nesi, F. Carrari, M. Davy, O. Bläsing, M. Kowalczyk, D. Weicht, A. Polinceusz, S. Meyer, M. Stitt and A.R. Fernie
Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: effects of genetic and environmental alterations in energy metabolism in the leaf
2006, Plant Molecular Biology 60, 773-792

B. Usadel, A. Nagel, O. Thimm, H. Redestig, O.E. Blaesing, N. Palacios-Rojas, J. Selbig, J. Hannemann, M.C. Piques, D. Steinhauser, W.R. Scheible, Y. Gibon, R. Morcuende, D. Weicht, S. Meyer and M. Stitt
Extension of the Visualisation Tool MapMan to Allow Statistical Analysis of Arrays, Display of Co-responding Genes and Comparison with Known Responses
2005, Plant Physiology 138, 1195-1204

B. Usadel, A.M. Kuschinsky, M.G. Rosso, N. Eckermann and M. Pauly
RHM2 Is Involved in Mucilage Pectin Synthesis and Is Required for the Development of the Seed Coat in Arabidopsis
2004, Plant Physiology 134 (1), 286-295

M. Viaud, A. Brunet-Simon, Y. Brygoo, J.M. Pradier and C. Levis C
Cyclophilin A and calcineurin functions investigated by gene inactivation, cyclosporin A inhibition and cDNA arrays approaches in the phytopathogenic fungus *Botrytis cinerea*
2003, Mol. Microbiol. 50, 1451-1465.

GABI Publications

Wang H, von Korff M, Léon J and Pillen K
Preliminary results on AB-QTL analysis of cold damage tolerance and pathogen resistances in winter barley
2005, Vorträge Pflanzenzüchtung 67, 40-45

Y. Wang, T. Hindemitt and K.F.X. Mayer
Significant sequence similarities in promoters and precursors of Arabidopsis thaliana non-conserved microRNAs sequence
2006, Bioinformatics 22 (21), 2585-2589

A.P. Weber, R. Schwacke and U.I. Flügge
Solute transporters of the plastid envelope membrane
2005, Annu. Rev. Plant Biol. 56, 133-164.

S. Weise, I. Grosse, C. Klukas, D. Koschützki,
U. Scholz, F. Schreiber and B.H. Junker
Meta-All: a system for managing metabolic pathway information
2006, BMC Bioinformatics 7, e465

B. Weisshaar and H. Saedler
Freier Zugang zu genetischen Pflanzenressourcen
2005, GenomXPress (3), 7-9

S. Xing, M.G. Rosso and S. Zachgo
ROXY1, a member of the plant glutaredoxin family, is required for petal development in *Arabidopsis thaliana*
2005, Development 132 (7), 1555-1565

Publications made with the aid of GABI Ressources (examples)

K. Baba, J. Schmidt, A. Espinosa-Ruiz, A. Villarejo,
T. Shiina, P. Gardestrom, A.P. Sane and R.P. Bhalerao
Organellar gene transcription and early seedling development are affected in the rpoT; 2 mutant of Arabidopsis
2004, Plant Journal 38 (1), 38-48

A. Chalfun-Junior, J. Franken, J.J. Mes,
N. Marsch-Martinez, A. Pereira and G.C. Angenent
ASYMMETRIC LEAVES2-LIKE1gene, a member of the AS2/LOB family, controls proximal-distal patterning in Arabidopsis petals
2005, Plant Molecular Biology 57 (4), 559-575

L. Corbesier, C. Vincent, S. Jang, F. Fornara, Q. Fan,
I. Searle, A. Giakountis, S. Farrona, L. Gissot,
C. Turnbull and G. Coupland
FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis
2007, Science 316 (5827), 1030-1033

C. Ebel, L. Mariconti and W. Gruissem
Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte
2004, Nature 429 (6993), 776-780

A. Feechan, E. Kwon, B.W. Yun, Y. Wang,
J.A. Pallas and G.J. Loake
A central role for S-nitrosothiols in plant disease resistance
2005, Proc. Nat. Acad. Sci. USA 102 (22), 8054-8059

S. Feng, Y. Shen, J.A. Sullivan, V. Rubio, Y. Xiong, T.P. Sun and X.W. Deng
Arabidopsis CAND1, an Unmodified CUL1-Interacting Protein, Is Involved in Multiple Developmental Pathways Controlled by Ubiquitin/Proteasome-Mediated Protein Degradation
2004, Plant Cell 16 (7), 1870-1882

M. Fiers, G. Hause, K. Boutilier, E. Casamitjana-Martinez, D. Weijers, R. Offringa, L. van der Geest, M. van Lookeren Campagne and C.M. Liu
Mis-expression of the CLV3/ESR-like gene CLE19 in Arabidopsis leads to a consumption of root meristem
2004, Gene 18 (327), 37-49

I. Finkemeier, M. Goodman, P. Lamkemeyer, A. Kandlbinder, L.J. Sweetlove and K.J. Dietz
The Mitochondrial Type II Peroxiredoxin F Is Essential for Redox Homeostasis and Root Growth of Arabidopsis thaliana under Stress
2005, Journal of Biological Chemistry 280 (13), 12168-12180

C.M. Ha, J.H. Jun, H.G. Nam and J.C. Fletcher
BLADE-ON-PETIOLE1 Encodes a BTB/POZ Domain Protein Required for Leaf Morphogenesis in *Arabidopsis thaliana*
2004, Plant Cell Physiology 45 (10), 1361-1370

Y. Hanzawa, T. Money and D. Bradley
A single amino acid converts a repressor to an activator of flowering
2005, Proc. Nat. Acad. Sci. USA 102 (21), 7748-7753

S. Kubis, R. Patel, J. Combe, J. Bedard, S. Kovacheva, K. Lilley, A. Biehl, D. Leister, G. Rios, C. Koncz and P. Jarvis
Functional Specialization amongst the Arabidopsis Toc159 Family of Chloroplast Protein Import Receptors
2004, Plant Cell 16 (8), 2059-2077

S. Lahmy, J. Guilleminot, C.M. Cheng, N. Bechtold, S. Albert, G. Pelletier, M. Delseny and M.S. Devic
DOMINO1, a member of a small plant-specific gene family, encodes a protein essential for nuclear and nucleolar functions
2004, Plant Journal 39 (6), 809-820

S. Laubinger, K. Fittinghoff and U. Hoecker
The SPA Quartet: A Family of WD-Repeat Proteins with a Central Role in Suppression of Photomorphogenesis in Arabidopsis
2004, Plant Cell 16 (9), 2293-2306

Y. Li, K. Sorefan, G. Hemmann and M.W. Bevan
Arabidopsis NAP and PIR Regulate Actin-Based Cell Morphogenesis and Multiple Developmental Processes
2004, Plant Physiology 136 (3), 3616-3627

C. Lurin, C. Andres, S. Aubourg, M. Bellaoui, F. Bitton, C. Bruyere, M. Caboche, C. Debast, J. Gualberto, B. Hoffmann, A. Lecharny, M. Le Ret, M.L. Martin-Magniette, H. Mireau, N. Peeters, J.P. Renou, B. Szurek, L. Taconnat and I. Small
Genome-Wide Analysis of Arabidopsis Pentatricopeptide Repeat Proteins Reveals Their Essential Role in Organelle Biogenesis
2004, Plant Cell 16 (8), 2089-2103

GABI Publications

R.F. Mills, A. Francini, P.S. Ferreira da Rocha, P.J. Baccarini, M. Aylett, G.C. Krijger and L.E. Williams
The plant P(1B)-type ATPase AtHMA4 transports Zn and Cd and plays a role in detoxification of transition metals supplied at elevated levels
2005, FEBS Letters 579 (3), 783-791

R. Morcuende, R. Bari, Y. Gibon, W. Zheng, B.D. Pant, O. Bläsing, B. Usadel, T. Czechowski, M.K. Udvardi, M. Stitt M and W.R. Scheible
Genome-wide reprogramming of metabolism and regulatory networks of Arabidopsis in response to phosphorus
2007, Plant Cell Environ 30 (1): 85-112

S. Ren, J.S. Johnston, D.E. Shippen and T.D. McKnight
TELOMERASE ACTIVATOR1 Induces Telomerase Activity and Potentiates Responses to Auxin in Arabidopsis
2004, Plant Cell 16 (11), 2910-2922

P.S. Rocha, M. Sheikh, R. Melchiorre, M. Fagard, S. Boutet, R. Loach, B. Moffatt, C. Wagner C and H. Vaucheret
The Arabidopsis HOMOLOGY-DEPENDENT GENE SILENCING1 Gene Codes for an S-Adenosyl-L-Homocysteine Hydrolase Required for DNA Methylation-Dependent Gene Silencing
2005, Plant Cell 17 (2), 404-417

R. Saedler, I. Zimmermann, M. Mutondo and M. Hulskamp
The Arabidopsis KLUNKER gene controls cell shape changes and encodes the AtSRA1 homolog
2004, Plant Molecular Biology 56 (5), 775-782

T. Shimada, K. Fuji, K. Tamura, M. Kondo, M. Nishimura and I. Hara-Nishimura
Vacuolar sorting receptor for seed storage proteins in *Arabidopsis thaliana*
2003, Proc. Nat. Acad. Sci. USA 100 (26), 16095-16100

J. Stockel and R. Oelmuller
A novel protein for photosystem I biogenesis
2004, Journal of Biological Chemistry 279 (11), 10243-10251

K. Vlieghe, V. Boudolf, G.T. Beemster, S. Maes, Z. Magyar, A. Atanassova, J. de Almeida Engler, R. De Groodt, D. Inze and L. De Veylder
The DP-E2F-like Gene DEL1 Controls the Endocycle in *Arabidopsis thaliana*
2005, Current Biology 15 (1), 59-63

M.K. Wall, L.A. Mitchenall and A. Maxwell
***Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria**
2004, Proc. Nat. Acad. Sci. USA 101 (20), 7821-7826

E. Zuther, K. Buchel, M. Hundertmark, M. Stitt, D.K. Hincha and A.G. Heyer
The role of raffinose in the cold acclimation response of *Arabidopsis thaliana*
2004, FEBS Letters 576 (1-2), 169-173

GABI Contacts

Achenbach, Ute
Max-Planck-Institute for Breeding Research, Cologne
achen@mpiz-koeln.mpg.de

Altmann, Thomas
University of Potsdam
taltmann@uni-potsdam.de

Aranda García, Ismael
INIA-CIFOR, Madrid, Spain
aranda@inia.es

Baliashvilli, Nino
Heinrich-Heine-Universität, Düsseldorf
nino.baliashvili@uni-duesseldorf.de

Basekow, Rico
RZPD German Resource Center
for Genome Research, Berlin
basekow@rzpd.de

Biemelt, Sophia
Friedrich-Alexander-University, Erlangen-Nürnberg
sbiemelt@biologie.uni-erlangen.de

Böttcher, Christoph
Free University of Berlin
bottcher@chemie.fu-berlin.de

Boyazen, Mondheur
INRA Toulouse, France
mondheur.boyazen@toulouse.inra.fr

Breun, Josef
Saatzucht Josef Breun GdbR, Herzogenaurach
saatzucht@breun.de

Brunel, Dominique
INRA Evry, France
brunel@versailles.inra.fr

Buffard, Dominique
INRA Versailles, France
buffard@versailles.inra.fr

Carbonero, Pilar
Ciudad Universitaria, Madrid, Spain
pcarbonero@bit.etsia.upm.es

Catusse, Julie
CNRS/Bayer CropScience joint laboratory, Lyon, France
julie.catusse@bayercropscience.com

Cervera, Maria Teresa
INIA, Madrid, Spain
cervera@inia.es

Charcousset, Alain
INRA Ferme du Moulon, Gif sur Yvette, France
alain.charcosset@moulon.inra.fr

Clemens, Stephan
Institute of Plant Biochemistry, Halle
sclemens@ipb-halle.de

Collada, Carmen
Ciudad Universitaria, Madrid, Spain
collada@montes.upm.es

Conrad, Udo
Institute of Plant Genetics and Crop Research, Gatersleben
conradu@IPK-gatersleben.de

Coupland, George
Max-Planck-Institute for Breeding Research, Cologne
coupland@mpiz-koeln.mpg.de

de Maria, Nuria
CIFOR-INIA, Madrid, Spain

Detering, Jutta
RAPS GbR, Grundhof
service@rapsgbr.com

GABI Contacts

Díaz, Luis
Madrid, Spain

Do, Cao-Trung
INRA Versailles, France

Dreyer, Felix
Norddeutsche Pflanzenzucht
Hans-Georg Lembke KG, Hohenlieth
f.dreyer@npz.de

Duchscherer, Peter
RAPS GbR, Grundhof
php.saat@t-online.de

Durand-Tardif, Mylène
INRA Versailles, France
durand-t@versailles.inra.fr

Ecke, Wolfgang
Georg-August-University, Göttingen
wecke@gwdg.de

Einfeldt, Claus
Dr. J. Ackermann & Co., Irlbach
einfeldt@sz-ackermann.de

Etienne, Catherine
UMR Interactions Plantes-Microorganismes et Santé végétale, Sophia Antipolis
c.etienne@bu.u-bordeaux1.fr

Fernandez, Rafael
EELM Almeria, Spain
rfern@eelm.csic.es

Finkeldey, Reiner
University of Göttingen
rfinkel@gwdg.de

Flügge, Ulf-Ingo
University of Cologne
ui.fluegge@uni-koeln.de

Frisch, Matthias
University of Hohenheim
frisch@uni-hohenheim.de

Garnier-Géré, Pauline
INRA Cestas, France
pauline@pierroton.inra.fr

Geiger, Hartwig H.
University of Hohenheim
geigerhh@uni-hohenheim.de

Gierl, Alfons
Technical University of Munich
gierl@wzw.tum.de

Goldringer, Isabelle
INRA Ferme du Moulon, Gif sur Yvette, France
isa@moulon.inra.fr

Gottwald, Sven
Institute of Plant Genetics and Crop Research, Gatersleben
gottwald@ipk-gatersleben.de

Graner, Andreas
Institute of Plant Genetics and Crop Research, Gatersleben
graner@ipk-gatersleben.de

Guevara, Maria-Angeles
Universidad Politecnica de Madrid, Spain

Haberer, Georg
GSF - National Research Center for Environment and Health, Neuherberg
g.haberer@gsf.de

Hannemann, Jan
Max-Planck-Institute for Molecular Plant Physiology, Potsdam
hannemann@mpimp-golm.mpg.de

Haseneyer, Grit
Institute of Plant Genetics and Crop Research, Gatersleben
haseneyr@ipk-gatersleben.de

GABI Contacts

Heckenberger, Martin
University of Hohenheim
heckenb@pz.uni-hohenheim.de

Himmelbauer, Heinz
Max-Planck-Institute for Molecular Genetics, Berlin
himmelb@molgen.mpg.de

Hohmann, Uwe
University of Kiel
uhohmann@plantbreeding.uni-kiel.de

Horn, Werner
SE Seed Hadmersleben GmbH, Hadmersleben

Job, Dominique
CNRS/Bayer CropScience joint laboratory, Lyon, France
dominique.job@bayercropscience.com

Jung, Christian
University of Kiel
c.jung@plantbreeting.uni-kiel.de

Keller, Harald
UMR Interactions Plantes-Microorganismes et Santé végétale, Sophia Antipolis
keller@antibes.inra.fr

Knopf, Erich
Pflanzenzucht Dr. h.c. Carsten - Inh. Erhardt Eger KG
e.knopf@pflanzenzucht-carsten.de

Kopka, Joachim
Max-Planck-Institute for Molecular Plant Physiology, Potsdam
kopka@mpimp-golm.mpg.de

Kremer, Antoine
INRA Cestas, France
antoine.kremer@pierroton.inra.fr

Kroymann, Jürgen
Max-Planck-Institute of Chemical Ecology, Jena
kroymann@ice.mpg.de

Kunert, Antje
University of Bonn
a.kunert@uni-bonn.de

Lange, Martin
Georg-August-University, Göttingen

Leckband, Gunhild
Norddeutsche Pflanzenzucht
Hans-Georg Lembke KG, Holtsee
g.leckband@npz.de

Léon, Jens
University of Bonn
j.leon@uni.bonn.de

Lepiniec, Loïc
INRA Versailles, France
lepiniec@versailles.inra.fr

Li, Yong
Max-Planck-Institute for Breeding Research, Cologne
yli@cebitec.uni-bielefeld.de

Link, Katrin
Justus-Liebig-University, Giessen
katrin.link@agrar.uni-giessen.de

Lunn, John
Max-Planck-Institute for Molecular Plant Physiology, Potsdam
lunn@mpimp-golm.mpg.de

Matthies, Inge
Institute of Plant Genetics and Crop Research, Gatersleben
matthies@ipk-gatersleben.de

Maurer, Hans-Peter
University of Hohenheim
hpmaurer@uni-hohenheim.de

Masur, Clarissa
RWTH Aachen
masur@bio3.rwth-aachen.de

GABI Contacts

Melchinger, Albrecht E.
University of Hohenheim
melchinger@uni-hohenheim.de

Mewes, Hans Werner
GSF - National Research Center for Environment and Health, Neuherberg
w.mewes@gsf.de

Miedaner, Thomas
University of Hohenheim
miedaner@uni-hohenheim.de

Mock, Hans-Peter
Institute of Plant Genetics and Crop Research, Gatersleben
mock@ipk-gatersleben.de

Möhring, Jens
University of Hohenheim
moehring@uni-hohenheim.de

Mönke, Gudrun
Institute of Plant Genetics and Crop Research, Gatersleben
moenke@ipk-gatersleben.de

Muminović, Jasmina
University of Hohenheim
muminov@uni-hohenheim.de

My Linh, Tran
Institute of Plant Genetics and Crop Research, Gatersleben
linh@ipk-gatersleben.de

Naz, Ali Ahmad
Max-Planck-Institute for Breeding Research, Cologne
naz@mpiz-koeln.mpg.de

Neuhaus, Gisela
Institute for Grapewine Breeding, Geilweilerhof
g.neuhaus@bafz.de

Ordon, Frank
Federal Centre for Breeding Research on Cultivated Plants, Quedlinburg
f.ordon@bafz.de

Parker, Jane
Max-Planck-Institute for Breeding Research, Cologne
parker@mpiz-koeln.mpg.de

Pégard, Anthony
UMR Interactions Plantes-Microorganismes et Santé végétale, Sophia Antipolis, France
pegard@antibes.inra.fr

Pestova, Elena
Heinrich-Heine-Universität, Düsseldorf
elena.pestsova@uni-duesseldorf.de

Pietsch, Christoph
Institute of Plant Genetics and Crop Research, Gatersleben
pietsch@ipk-gatersleben.de

Plomion, Christophe
INRA Cestas, France
plomion@pierroton.inra.fr

Presterl, Thomas
KWS Saat AG, Einbeck
t.presterl@kws.com

Ravel, Catherine
INRA Clermont-Ferrand, France
ravel.clermont.inra.fr

Roby, Dominique
CNRS, Toulouse, France
dominique.roby@toulouse.inra.fr

Rogowsky, Peter
INRA Lyon, France
peter.rogowsky@ens-lyon.fr

GABI Contacts

Rosso, Mario
Max-Planck-Institute for Breeding Research, Cologne
mrosso@cebitec.uni-bielefeld.de

Rousset, Michel
INRA Ferme du Moulon, Gif sur Yvette, France
michel.rousset@moulon.inra.fr

Sagasser, Martin
University of Bielefeld
sagasser@CeBiTec.Uni-Bielefeld.DE

Sauer, Sascha
University of Hohenheim
sauer@molgen.mpg.de

Scheel, Dierk
Institute for Plant Biochemistry, Halle
dscheel@ipb-halle.de

Schmalenbach, Inga
Max-Planck-Institute for Breeding Research, Cologne
i.schmalenbach@uni-bonn.de

Schmidt, Thomas
Technical University of Dresden
thomas.schmidt@tu-dresden.de

Schmitz, Gregor
Max-Planck-Institute for Breeding Research, Cologne
schmitz@mpiz-koeln.mpg.de

Schneider, Katharina
GSF - National Research Center
for Environment and Health, Neuherberg
katharina.schneider@gsf.de

Schondelmaier, Jörg
SAATEN-UNION Resistenzlabor, Leopoldshöhe
schondel@saaten-union-labor.de

Schubert, Anne
University of Bonn
a.schubert@uni-bonn.de

Schwacke, Rainer
University of Cologne
rainer.schwacke@uni-koeln.de

Schweizer, Patrick
Institute of Plant Genetics and Crop Research, Gatersleben
schweiz@ipk-gatersleben.de

Selbig, Joachim
University of Potsdam
selbig@mpimp-golm.mpg.de

Snowdon, Rod
Justus-Liebig-University, Giessen
rod.snowdon@agrar.uni-giessen.de

Soto, Alvaro
CIFOR-INIA, Madrid, Spain
asdv@inia.es

Sreenivasulu, Nese
Institute of Plant Genetics and Crop Research, Gatersleben
srinivas@ipk-gatersleben.de

Stelling, Dieter
Deutsche Saatveredelung Lippstadt-Bremen GmbH, Lippstadt
stelling@dsv-saaten.de

Stich, Benjamin
University of Hohenheim
stich@uni-hohenheim.de

Stracke, Silke
Institute of Plant Genetics and Crop Research, Gatersleben
stracke@ipk-gatersleben.de

GABI Contacts

Streng, Stefan
Saatzucht Streng GmbH & Co. KG, Uffenheim
s.streng@aspachhof.de

Strittmatter, Günter
KWS Saat AG, Einbeck
g.strittmatter@kws.com

Sulpice, Ronan
Max-Planck-Institute for Molecular Plant Physiology, Potsdam
sulpice@mpimp-golm.mpg.de

Theres, Klaus
Max-Planck-Institute for Breeding Research, Cologne
theres@mpiz-koeln.mpg.de

Tronchet, Maurine
Interactions Plantes-Microorganismes, Castanet-Tolosan, France
maurine.tronchet@toulouse.inra.fr

Usadel, Björn
Max-Planck-Institute for Molecular Plant Physiology, Potsdam
usadel@mpimp-golm.mpg.de

v. Roepenack-Lahaye, Edda
Institute for Plant Biochemistry, Halle
vonroep@ipb-halle.de

Viehoever, Prisca
University of Bielefeld
viehoeve@CeBiTec.Uni-Bielefeld.DE

Vorwieger, Astrid
Institute of Plant Genetics and Crop Research, Gatersleben
vorwiege@ipk-gatersleben.de

Wang, Yu
GSF - National Research Center
for Environment and Health, Neuherberg
yu.wang@gsf.de

Weigel, Detlef
Max-Planck-Institute for Developmental Biology, Tübingen
weigel@weigelworld.org

Wenzel, Gerhard
Technical University of Munich, Freising
gwenzel@wzw.tum.de

Weschke, Winfriede
Institute of Plant Genetics and Crop Research, Gatersleben
weschke@ipk-gatersleben.de

Weyen, Jens
SAATEN-UNION Resistenzlabor, Leopoldshöhe
weyen@saaten-union-labor.de

Witucka-Wall, Hanna
Univerity of Potsdam
wall@mpimp-golm.mpg.de

Wobus, Ulrich
Institute of Plant Genetics and Crop Research, Gatersleben
wobusu@ipk-gatersleben.de

Zacharias, Arndt
KWS Saat AG, Einbeck
a.zacharias@kws.de

Zyprian, Eva
Institute for Grapewine Breeding, Geilweilerhof
e.zyprian@bafz.de

Go to Content Page

www.gabi.de