Summaries of Legume Genomics Projects from around the Globe. Community Resources for Crops and Models

Edited by Kathryn A. VandenBosch and Gary Stacey, Monitoring Editors, *Plant Physiology*

Genomic research has and will continue to revolutionize plant biology. It is clear that the adoption of Arabidopsis as a model species has done much to speed the development of plant genomics and to hasten our increased understanding of basic plant biology. However, Arabidopsis is not an "omniscient" model because this plant does not encompass all of the diverse physiological, developmental, and environmental processes seen throughout the plant kingdom. Thus, to study these other processes and to bring the genomic revolution to crop species, additional genomic resources must be developed in other plants.

Over the past several years, this realization has led to the adoption of the model species concept to the study of legumes. Unlike Arabidopsis, legumes develop important and interesting symbioses with nitrogen (N)-fixing rhizobia and with mycorrhizal fungi. They also exhibit interesting differences in secondary metabolism, pod development, and other processes that cannot be adequately modeled with Arabidopsis. The impetus for the development of legume models has come primarily from researchers interested in the rhizobium-legume symbiosis. Because of this, two models, not one, have been developed: *Lotus japonicus* and *Medicago truncatula*. In reality, these two models have evolved due to the energy of their proponents but, scientifically, they can also be justified because they exhibit two developmental systems for nodulation as well as other differences. *L. japonicus* forms determinate nodules, in which the root subepidermal cortical cells initiate nodule formation and a persistent, terminal nodule meristem does not develop. In contrast, *M. truncatula* nodules initiate from the division of inner cortical cells and continue to grow from a terminal, persistent meristem. As can be seen by the summaries below, both legume model species are now well established with a large number of laboratories involved. Therefore, in the long run, legume biology can only benefit by a comparison of the results between these models and legume crop plants.

In contrast to the effort focused on the legume models, with the possible exception of soybean (*Glycine max*), significantly smaller efforts exist to study the genomics of legume crop species. This is unfortunate because it remains to be seen just how much of the information developed from legume models can be directly applied to the improvement of legume crops.

It is clear that legume biology is rapidly undergoing a revolutionary transformation due to the application of genomic methods. The future is exceedingly bright, and one would expect rapid progress in our understanding of basic plant processes and the unique aspects of legume physiology and development.

The edited summaries below represent the currently funded genomic activities focused on legume models and crops. For convenience, these are listed by relevant species, although similar trends and interests are apparent throughout.

L. JAPONICUS

A Domestic Weed Goes Worldwide. Recent Progress on *Lotus* Research in Japan

Makoto Hayashi, Osaka University, Japan, hayashi@bio.eng.osaka-u.ac.jp; Toshio Aoki, Nihon University, Japan, taoki@brs.nihon-u.ac.jp; Sachiko Isobe, National Agricultural Research Center for Hokkaido Region, Japan, sisobe@affrc.go.jp; Kyuya Harada, Chiba University, Japan, haradaq@faculty.chiba-u.jp; Hiroshi Kouchi, National Institute of Agrobiological Sciences, Japan, kouchih@nias.affrc.go.jp; Kiwamu Minamisawa, Tohoku University, Japan, kiwamu@ige.tohoku.ac.jp; Kazuhiko Saeki, Osaka University, Japan, ksaeki@bio.sci.osaka-u.ac.jp; Shusei Sato, Kazusa DNA Research Institute, Japan, ssato@kazusa.or.jp; Satoshi Tabata, Kazusa DNA Research Institute, Japan, tabata@kazusa.or.jp; and Masayoshi Kawaguchi, Niigata University, Japan, masayosi@env.sc.niigata-u.ac.jp

This work was supported by grants from the Japan Science and Technology Corporation; the Japanese Ministry of Education, Culture, Sports, Science and Technology; the Kazusa DNA Research Institute Foundation; and the Bio-oriented Technology Research Advancement Institution.

L. japonicus was first recognized at the ancient capital of Japan, Kyoto, centuries ago. Its Japanese name, "Miyakogusa," means "capital weed," but the reason for this is not clear. It might be due to the fact that this weed was common in Kyoto, or the bright and

www.plantphysiol.org/cgi/doi/10.1104/pp.103.020388.

 © 2003 American Society of Plant Biologists

showy color of its flowers might evoke the luxury ofthe capital city. In Japan, people used the weed as a remedy. In the 1950s, Professor Isao Hirayoshi (Kyoto University) collected *L. japonicus* plants growing on a riverbank in Gifu. Professor William F. Grant (McGill University, Montreal) collected its progeny as the accession B-129. In 1992, Kurt Handberg and Jens Stougaard of University of Aarhus (Denmark) obtained B-129 and established the weed as a valuable tool for modern legume research. Now, *Lotus* is regarded as one of the most useful plants for legume study. Four genes that condition nodulation phenotypes have already been cloned. Researchers who have interests in nodulation and other aspects of legume biology use it worldwide.

The research activity of *Lotus* in Japan is mostly carried out by a nonprofit organization, the Miyakogusa Consortium. Started at the end of 1999, it develops and maintains public resources essential for *Lotus* research: linkage maps, expression arrays (of both plant and endosymbiont genes), transformation techniques, as well as making accessions available and fostering communication among researchers. Approximately 30 laboratories all over Japan are involved in *Lotus*.

An advantage to researchers in Japan is that many ecotypes of *Lotus* can be found growing in the wild. Genotypic variety is necessary to investigate actual and potential traits of agronomic importance, such as seed yield, plant height, cold tolerance, and disease resistance, which can be identified by quantitative trait locus (QTL) mapping. Dozens of accessions have been collected, from the northernmost Hokkaido to the southernmost island Miyakojima. The *L. japonicus* Seed Center was recently established at the National Agricultural Research Center for Hokkaido Region (http://cryo.naro.affrc.go.jp/sakumotu/mameka/lotus-e.htm). As of October 2002, more than 60 accessions are in distribution, and almost 80 accessions are under production. Also available in the near future will be recombinant inbred lines (RILs) between "Gifu" B-129 and "Miyakojima" MG-20 (Kawaguchi, 2000), the most divergent accession of *L. japonicus* from "Gifu" known to date (Kawaguchi et al., 2001). The lines, established by Kazusa DNA Research Institute, consist of 221 lines at the F_8 stage. Almost one-half of the RILs will be distributed by March 2003. Using RILs, stable assessment of mapping loci and sharing of the mapping population will be possible.

Genetic transformation is a prerequisite to modern molecular biology and molecular genetics. Leguminous plants are relatively recalcitrant to transformation, although this is highly dependent on species. Since the first report of *Lotus* transformation, several articles have dealt with the improvement of transformation, and the technique is now readily at hand. We established a new *Agrobacterium tumefaciens*-mediated transformation technique for "Gifu" (Aoki et al., 2002) and "Miyakojima" with minimal contamination by non-transformants, thus largely eliminating the need for confirmation of the transgene by PCR or other time-consuming tests. This reliable transformation method is efficient enough for large-scale experiments using insertional mutagenesis or gene tagging. More than 1,300 T_1 plants of "Gifu" have been generated for gene tagging. The tagged lines were constructed using *A. tumefaciens* EHA101 and a tagging vector pEKB35SEXTra, which can be used for both activation tagging and exon trapping (I. Nakamura, unpublished data). Genomic DNA gel-blot analysis of arbitrarily selected T_1 plants showed that more than 60% had single copy of T-DNA, implying a relatively low frequency of multicopy insertion and genomic rearrangement. Some putative mutants show modified morphology and/or nodulation (T. Aoki, unpublished data).

Although molecular genetic studies mainly deal with the traits of whole plants, cell suspension cultures can serve as an alternative for investigating cell biology and physiology of metabolism. Cultured cell lines were established from "Gifu" and "Miyakojima" (K. Syono, unpublished data) and will be used for comprehensive profiling of metabolites in future studies. Cultured cells under various conditions will also provide sources for new cDNA libraries, which could be mined for unusual and invaluable gene transcripts.

With the aim of understanding the whole genome of *Lotus*, both cDNA and genome sequencing are in progress at the Kazusa DNA Research Institute. As of October 2002, 93,000 5′ and 3′ expressed sequence tags (ESTs) have been obtained from normalized and size-selected cDNA libraries constructed from seven different organs, such as nodules, pods, and flower buds (Asamizu et al., 2000). A total of 70,137 3′ ESTs have been clustered into 20,127 nonredundant groups. The sequence data from these ESTs are available at the Web site: http://www.kazusa.or.jp/en/plant/lotus/EST/. For initiating genome sequencing, genomic clones corresponding to multiple seed points were selected using sequence information from ESTs and cDNA markers of *Lotus* and other legumes. As of October 2002, a total of 975 seed clones have been selected. Ninety-two of them are in the library phase, seven are being sequenced, 411 are in the finishing phase, 88 are being annotated, and 199 have been annotated. The annotated sequences are being made available on the public databases and on our Web database at http://www.kazusa.or.jp/lotus.

A large-scale cDNA macroarray was constructed using *Lotus* ESTs. This contains 18,144 nonredundant ESTs on a set of nylon membranes, which were selected from the EST resources (about 69,000 clones) established in the Kazusa DNA Institute. By way of example, we have analyzed comprehensive gene expression during early stages of *Lotus* nodule forma-

tion by means of the cDNA array. These studies detected more than 1,000 genes that are significantly up-regulated during nodulation.

The isolation of plant mutants will be required to fully elucidate the molecular mechanisms underlying plant-microbe symbioses. In our case, 33 stable mutant lines affecting nodule number and organogenesis were isolated by ethyl methanesulfonate (EMS) or ion beam mutagenesis. They include Nod^-, $Hist^-$, Fix^-, Myc^-, Nod^{2+}, and Myc^{2+} phenotypes. We found three loci conferring an increased nodule number, i.e. *astray* (*Ljsym77*), *har1* (*Ljsym78*), and *beading nodule* (*bel*). Among them, two genes have been cloned. The *Astray* gene was demonstrated to encode a bZIP protein similar to Arabidopsis *HY5* that is known as a key regulator of photomorphogenesis (Nishimura et al., 2002b). On the other hand, the *Har1* gene encodes a Leu-rich repeat (LRR) receptor-like kinase having the highest identity with Arabidopsis *CLAVATA1* (Nishimura et al., 2002a). Sequence analysis of the *Har1* ortholog in soybean revealed that the hypernodulating mutant En6500 that is allelic to *nts1* has a stop codon near the transmembrane domain. These cloned genes would serve as starting points to understand light and systemic regulation of nodule development at a molecular level.

Some symbiotic mutants were isolated by means of possible somaclonal variation during tissue culture (Y. Umehara and H. Kouchi, unpublished data). Because active retrotransposons were found in the process of positional cloning of *LjSYMRK*, their use would facilitate the cloning of symbiotic genes. Besides nodulation, mutants affecting nyctinastic movement (*sleepless*), root hair (*slippery root*; Kawaguchi et al., 2002), and anthocyanin accumulation (*viridicaulis*; T. Aoki, unpublished data) were also isolated. For the positional cloning of EMS and other mutants lacking tagged genes, linkage maps were constructed between "Gifu" and "Miyakojima" (Hayashi et al., 2001). Using amplification length polymorphisms (AFLP; recombinations of *Eco*RI-*Mse*I and *Hin*dIII-*Taq*I), simple sequence repeat (SSR), and derived cleaved amplified polymorphic sequences, almost 480 markers covering roughly 500-cM distance are mapped in both linkages. The Kazusa DNA Research Institute is now generating more SSR markers (Sato et al., 2001; Nakamura et al., 2002). This will make it possible to map and isolate mutated genes by simple PCR.

In summary, a wide spectrum of work has been initiated to develop resources for *Lotus* research. As a result, research on the molecular genetic, functional genomics, molecular breeding, and metabolomics of *Lotus* will be facilitated by the availability of linkage maps, genome sequences, ESTs, expression arrays, mutant lines, and accessions. It is now up to the choice of individual investigators to find a new edge in legume research using *Lotus* as a model plant.

Research Training Using *Lotus japonicus*. A Model Legume for Functional Genomics

Principal Investigator (PI): Michael Udvardi, Max Planck Institute of Molecular Plant Physiology, Germany, udvardi@mpimp-golm.mpg.de; Co-PI: Maurizio Chiurazzi, Institute of Genetics and Biophysics, Naples, Italy, chiurazz@iigbna.iigb.na.cnr.it; Co-PI: Panagiotis Katinakis, Agricultural University of Athens, Greece, bmbi2kap@auadec.aua.gr; Co-PI: Antonio Marquez, University of Seville, Spain, cabeza@cica.es; Co-PI: Martin Parniske, John Innes Centre, UK, martin.parniske@bbsrc.ac.uk; Co-PI: Gerhard Saalbach, Risoe National Laboratory, Denmark, g.saalbach@risoe.dk; Co-PI: Herman Spaink, Leiden University, The Netherlands, spaink@rulbim.leidenuniv.nl; Co-PI: Jens Stougaard, University of Aarhus, Denmark, stougaard@mbio.aau.dk; and Co-PI: Judith Webb, Institute of Grassland and Environmental Research (IGER), UK, judith.webb@bbsrc.ac.uk)

This work was supported by the European Union (FP5 Project No. HPRN–CT–2000–00086; http://improving.cordis.lu/rtn/).

Beneficial plant-microbe interactions are extremely important to agriculture and the ecology of our planet. Root symbioses between plants (specifically legumes) and bacteria of the family Rhizobiaceae (called simply rhizobia), and between plants and arbuscular mycorrhizal (AM) fungi are perhaps the most important of all such interactions. N-fixing symbioses between legumes and rhizobia enable the plants to grow in the absence of fertilizer N. AM symbioses, on the other hand, often play a crucial role in the phosphorous (P) nutrition of plants.

N-fixing and AM symbioses have played an important role in sustainable agricultural systems for hundreds, if not thousands, of years and have the potential to play an even greater role in the future. Realization of this potential requires further fundamental research on these symbioses. The legume *L. japonicus* is a valuable model species for symbiosis research because it has a relatively small diploid genome, it is self-fertile, and it can be transformed easily. Thus, it is amenable to forward and reverse genetics, genomics, and functional genomics.

The *Lotus* project is an international, multidisciplinary effort funded by the European Union to promote research training at the cutting edge of plant science. The two principal research objectives of the project are to develop resources for functional genomics of *L. japonicus*, and to use these resources to understand better how N-fixing and AM symbioses develop and how they function to provide plants with N, P, and other nutrients. With respect to the first objective, a number of essential public resources are being developed. These include: large populations of genetically tagged and untagged mutants; a

high-resolution genetic map of the *L. japonicus* genome; large libraries of root and nodule ESTs; facilities for high-throughput analysis of mRNA, protein, and metabolites; standardized protocols for growth and physiological analysis of plants; advanced microscopy protocols for cell biology; and, finally, capabilities to collect, store, analyze, and distribute large amounts of data. With respect to the second objective, work will focus on the identification of genes and signals involved in development of N-fixing nodules and AMs, as well as on metabolism and transport in nodules.

There are several highlights of progress after 2 years. To facilitate transcriptome and classical molecular/cell biological studies in *Lotus*, approximately 10,000 ESTs have been obtained from rhizobium-infected roots and mature nodules of *Lotus*. Using bioinformatics, we have ascribed putative functions to the proteins encoded by many of these genes, and classified them into various functional categories including putative signaling proteins, putative transcription factors, enzymes involved in primary and secondary metabolism, and many different types of transporters.

Full-length cDNAs encoding enzymes of carbon metabolism (phospho*enol*pyruvate carboxylase, Suc synthase, Suc transporters, trehalose phosphatase, and carbonic anhydrase-α, -β, and -γ type) and N metabolism (Asn synthetase, Gln synthetase, Asp aminotransferase, Orn decarboxylase, Arg decarboxylase, and Glu decarboxylase) were identified in our EST database, and RNA in situ hybridization analysis was used to localize expression of these genes in nodules. Immunolocalization of phospho*enol*pyruvate carboxylase and carbonic anhydrase (β-type) protein confirmed the gene expression data. This information will contribute to a better understanding of how nodule metabolism is organized and regulated.

Genes that are up-regulated during nodule or AM development in *Lotus*, and which may be essential for these processes, were identified by two complementary approaches. The first approach utilized DNA arrays, produced by spotting 2,000 EST clones onto nylon membranes, to identify genes that were differentially expressed in nodules compared with roots. In this way, 83 genes were identified that may play important roles in nodule development or function of mature nodules (Colebatch et al., 2002). Among these were genes for primary C and N metabolism, metabolite transport, hormone metabolism, cell wall biosynthesis, signal transduction, and regulation of transcription. The size of arrays was recently increased to 5,000 clones, and this will be increased to over 10,000 in the near future, using additional clones obtained from Japanese colleagues (see Hayashi et al., above). The second approach employed cDNA-AFLP analysis to identify genes that were induced or repressed in roots shortly after inoculation with *Mesorhizobium loti* or *Glomus intraradices*. In total, 1,200 differentially regulated cDNA fragments were isolated. Research has been focused on those genes that are induced by both microsymbionts ("symbiosins"). RNA interference (RNAi) construct design for silencing of selected symbiosins has been initiated. Some of these genes may play regulatory roles during early stages of symbiosis. Differentially expressed genes will be used to produce a temporal map of the molecular events that occur during symbiosis development. This map will be useful in establishing a hierarchy of mutants affected in nodulation and/or AM development.

To accelerate reverse genetics in *Lotus*, we have improved methods for *Lotus* transformation. An optimized, in vitro transformation regeneration protocol using root explants has been developed that increases transformation and regeneration efficiencies and decreases the plant regeneration time. Other transformation techniques are being tested to find the most time- and labor-efficient method for generation of transgenic plants/roots.

To facilitate forward genetics in *Lotus*, populations of transposon and T-DNA insertion mutants are being developed. To create *Ds* insertion mutants in *L. japonicus*, two gene trap constructs and two activation-tagging constructs were transferred into *Lotus* by *Agrobacterium* transformation. At present, 200 lines are being raised for seed production. From this material, selection of double resistant lines was initiated for isolation of *Ds* lines for nodulation mutant screening. An additional 800 transformed calli are going through the regeneration procedure and will be added to the collection of *Ds* launching lines. A large collection of independent transformants obtained with two different promoterless reporter gene T-DNA constructs is also in preparation. Most recently, an active retrotransposon, LORE1, has been found in *Lotus*, and simple conditions for activation of this element are being investigated. Activation of this endogenous element by a controllable environmental condition would facilitate efficient tagging procedures.

Using polymorphic markers (AFLP, RFLP, and sequence/gene-specific PCR), the genetic map of *Lotus* has been developed into a very effective tool for map-based cloning of symbiotic genes and other genes of interest. We have developed a series of codominant sequence-known and gene-based markers that resulted in a consolidated genetic map of *L. japonicus* (Sandal et al., 2002). This collaboration has already facilitated the map-based cloning of a gene indispensable for root symbioses (Stracke et al., 2002) and another that controls nodule number (Krusell et al., 2002). Cloning of three other symbiotic loci involved in Nod-factor signal perception or immediate downstream signal transduction is in a very advanced stage.

A proteomics approach has been taken to identify proteins at the symbiotic interface in N-fixing nodules. A method to isolate the peribacteroid mem-

brane from *Lotus* nodules was developed, which utilizes aqueous two-phase partitioning of membrane fractions. Using this method together with mass spectrometry (MS), several novel putative PBM proteins have been identified, including a sulfate transporter that matches an EST, which showed nodule-induced expression pattern on DNA arrays (Weinkoop and Saalbach, 2003).

Metabolite profiling, using gas chromatography (GC)-MS, has also commenced in two of our groups and has revealed quantitative changes in flavonoids after mycorrhizal infection. Comparisons were made of wild-type leaf, stem, and root tissues of *L. corniculatus* and *L. japonicus*, using HPLC-photodiode array, HPLC-photodiode array/MS and GC/MS. This comparison showed that although shoot profiles are similar, the roots of these two species are different. GC-MS analysis has also been used to profile changes in metabolites in nodules of mutant, non-N-fixing plants and wild-type nodules containing mutant rhizobia.

In summary, the *Lotus* consortium has made significant progress in legume functional genomics. The tools for rapid map-based cloning of genes have been developed to the point that *Lotus* is now a premier model legume for forward genetics. As a result, members of the *Lotus* consortium have been among the first to identify several genes that are essential for beneficial symbiosis in plants. The state-of-the-art of legume reverse genetics has also been advanced by the *Lotus* project, which has developed insertion mutant populations, protocols for RNAi suppression, or overexpression of genes in *Lotus*, and a large population of EMS mutants and facilities for TILLING. Advances have also been made in transcriptome, proteome, and metabolome analysis for legume research.

A TILLING Reverse Genetics Tool for *L. japonicus*

PI: Martin Parniske, The Sainsbury Laboratory, UK, martin.parniske@bbsrc.ac.uk; Co-PI: Trevor Wang, John Innes Centre, UK, trevor.wang@bbsrc.ac.uk; Co-Investigators: Jillian Perry, Sarah Gardener, and Jodie Pike, The Sainsbury Laboratory, UK; and Tracey Welham, John Innes Centre, UK

This work was supported by the Biotechnology and Biological Science Research Council (grant no. 83/D15167), by the John Innes Centre (two rapid response interdepartmental research grants), and by the Gatsby Charitable foundation (http://www.lotusjaponicus.org; to The Sainsbury Laboratory).

A strategy for reverse genetics that is based on EMS-mutagenesis was first described by McCallum et al.(2000) using the acronym TILLING (targeted induced local lesions in genomes). A specific advantage of EMS mutagenesis is that the series of allelic mutations can serve as the basis of detailed structure-function studies. In addition, this has the potential to recover weak alleles with subtle changes in functionality of genes that would be lethal when more strongly affected. TILLING identifies individuals carrying point mutations in any gene of interest within a large population of EMS-mutagenized plants.

We have established a TILLING reverse genetics tool for the legume *L. japonicus* with the objective of establishing a resource for the scientific community. The methods and early results of this endeavor are described in detail in this issue (Perry et al., 2003). Taking into account that root symbiosis is a large and legume-specific area of interest, we generated two populations directly accessible for TILLING: a general TILLING population, biased against the occurrence of severe developmental phenotypes, and a series of smaller populations of siblings exhibiting defects in the root nodule symbiosis. For the general TILLING population, our aim was to include fertile individuals only, so that progeny of a plant carrying a mutant allele can be directly recovered from progeny of the TILLed plant.

Within a population of preselected symbiotic mutants, a series of functionally impaired alleles of the *SYMRK* gene could be identified. This gene is required for the formation of root symbioses (Stracke et al., 2002). Only a fraction of these alleles were represented in the corresponding siblings of the general TILLING population. We concluded that in the case where a gene of interest has already been implicated in a particular biological process, the inclusion of a forward screen for that particular trait can increase the frequency at which functionally affected alleles can be recovered.

To cover research interest in other aspects of legume biology, mutant siblings were isolated that exhibited abnormal root branching patterns, abnormal growth habit (dwarf and stature mutants), abnormal leaf or flower development, or were affected either in starch synthesis or breakdown. The mutant phenotypes including photographs were entered into a Web-accessible database (www.lotusjaponicus.org/finder.htm). We have collected seed from these developmental mutants, and trait-specific TILLING populations could be set up on demand.

Plant-Insect Interactions as a Response to Metabolic Engineering of Natural Product Synthesis Studied by Functional Genomics in *L. japonicus*

PI: Søren Bak, Royal Veterinary and Agricultural University, Denmark, bak@kvl.dk; Co-PI: Karin Forslund, Royal Veterinary and Agricultural University, Denmark, kaf@kvl.dk; Co-PI: Birger Lindberg Møller, Royal Veterinary and Agricultural Univer-

sity, Denmark, blm@kvl.dk; Co-PI: Bodil Jørgensen, Danish Institute of Agricultural Sciences, Denmark, b.jorgensen@dias.kvl.dk; International Collaborators: David Galbraith, University of Arizona, Tucson, galbraith@arizona.edu, Kazusa DNA Research Institute, Japan, lotus@kazusa.or.jp; Michel Udvardi, Max-Planck-Institut, Golm, Germany, udvardi@mpimp-golm.mpg.de; and Clas M. Naumann, Leibniz Institute for Research in Terrestrial Biology, Germany, c.naumann.zfmk@uni-bonn.de

This work was supported by the Danish National Research Foundation (grant to the Center for Molecular Plant Physiology, [PlaCe]) and by the Danish Agricultural and Veterinary Research Council (http://www.place.kvl.dk; grant no. 23-02-0095).

We are introducing *L. japonicus* as a genetic model system to study cyanogenic glucosides. Cyanogenic glucosides are β-glucosides of α-hydroxynitriles and are classified as phytoanticipins. Upon disruption of plant tissue containing cyanogenic glucosides, these are degraded by β-glucosidases and α-hydroxynitrilases, resulting in the release of toxic hydrogen cyanide as well as of Glc and an aldehyde or ketone. This binary system—two sets of components that separately are chemically inert—provides plants with an immediate chemical defense response to herbivores and pathogens that cause tissue damage (Møller and Seigler, 1999). To study the specific effect of cyanogenic glucosides on plant-insect interactions, we have previously transferred the entire pathway for synthesis of the Tyr-derived cyanogenic glucoside dhurrin into the model plant Arabidopsis using the sorghum (*Sorghum bicolor*) genes *CYP79A1*, *CYP71E1*, and *UGT85B*. In free-choice tests, the crucifer specialist flea beetle (*Phyllotreta nemorum*) was found to avoid the dhurrin-containing Arabidopsis plants (Tattersall et al., 2001). Arabidopsis belongs to the Brassicaceae, which do not contain cyanogenic glucosides but produce a related group of amino acid-derived natural products classified as glucosinolates. The two pathways have aldoximes as common intermediates and, in accordance, the introduction of the Tyr-derived cyanogenic glucoside dhurrin also results in simultaneous accumulation of a new Tyr-derived glucosinolate, *p*-hydroxybenzyl glucosinolate (Bak et al., 1999, 2000). Furthermore, the level of endogenous β-glucoside activity available to hydrolyze dhurrin is low in Arabidopsis, rendering cyanide release slow (Tattersall et al., 2001). To overcome these experimental problems, we are introducing *L. japonicus* as a new experimental model system, and taking a functional genomics approach.

L. japonicus contains the two cyanogenic glucosides, linamarin and lotaustralin, derived from Val and Ile, respectively. Lotaustralin constitutes the major glucoside. The β-glucosidase activity is high, causing rapid cyanide release upon tissue damage. In collaboration with the Kazusa DNA Research Institute (http:www.kazusa.or.jp/lotus/), the *L. japonicus* genome has been found to contain two *CYP79* orthologs assigned as *CYP79D3* and *CYP79D4*, which show different expression patterns at the tissue level. Bioinformatic approaches are currently being used to identify the *L. japonicus* genes orthologous to *S. bicolor CYP71E1* and *UGT85B1* (Paquette et al., 2000, 2003). Promoter fusion constructs of key regulatory enzymes in biosynthesis and degradation of cyanogenic glucosides will be generated in collaboration with Dr. David Galbraith to delineate the tissue-specific expression pattern in planta and facilitate in vivo expression studies after abiotic and biotic stress.

Metabolic engineering of cyanogenic glucoside synthesis in *L. japonicus* proceeds following two parallel approaches. Transgenic plants overexpressing pathway enzymes from cassava (*Manihot esculenta* Crantz.; Andersen et al., 2000) or sorghum are being analyzed for altered patterns of glucoside accumulation. Second, *L. japonicus* genes involved in cyanogenic glucoside synthesis or degradation will be silenced by either RNAi (posttranslational gene silencing), or mutants will be isolated by screening the collection of TILLING mutants available at The Sainsbury Laboratory (see Parniske et al., above; Perry et al., 2003). Metabolite profiling based on HPLC/MS/MS has been established. Transcriptome analyses are being carried out using DNA macroarrays provided by Dr. Michael Udvardi, according to previously established protocols (Xu et al., 2001).

L. japonicus has co-evolved with Zygaenae moths. Members of the Zygaenae sequester cyanogenic glucosides in special glands and utilize them in defense against its predators. Together with Clas Naumann, who is able to rear *Zygaena trifolii*, we will investigate interplay between *Z. trifolii* and the transgenic *L. japonicus* plants engineered to have no or different cyanogenic glucoside profiles or to be unable to degrade such glucosides. The *L. japonicus* plants expressing promoter fusions of the key regulatory enzymes in biosynthesis, and degradation of cyanogenic glucosides will be included in these studies. The plant/insect studies will be carried out with additional insects with a focus on insects for which DNA macroarray chips are available. In this way, the chemical warfare between plants and insects can be followed at the transcriptional level by transcriptome analyses and by promoter reporter gene fusions, as well as through metabolite profiling, thereby providing a detailed understanding of the relative importance of complete metabolism, detoxification, and sequestering.

Cytochromes P450, UDPG-glycosyltransferases, and β-glucosidases are key biocatalysts in the formation and degradation of natural products in plants (Vogt and Jones, 2000) and enable recruitment of new functions to maintain competitiveness in the chemical warfare toward herbivores and pathogens (Paquette et al., 2003). We will annotate *L. japonicus* genes from these

three multigene families and include them on our bioinformatics Web site (The Arabidopsis P450, cytochrome b5, P450 reductase, and Glycosyltransferase Site at Center for Molecular Plant Physiology (PlaCe) at http:/www.biobase.dk/P450/) to ensure correct annotation and naming according to the nomenclature rules. We will carry out comparative genomic phylogenetic analyses of cytochrome P450s and UDPG-glycosyltransferase from green algae, *Lotus*, and Arabidopsis to obtain a better understanding of the expansion of these two multigene families during plant evolution.

Identifying Symbiotic Genes in Model and Crop Legumes

Co-PI: K. Judith Webb, IGER, Aberystwyth, UK, judith.webb@bbsrc.ac.uk. Co-PI: Leif Skøt, IGER, leif.skot@bbsrc.ac.uk; and Co-PI: William Eason, IGER, william.eason@bbsrc.ac.uk

This work was supported by the Biotechnology and Biological Sciences Research Council, UK (Competitive Strategic Grant).

We aim to increase understanding of the genetic basis of interactions between legumes and microorganisms that affect plant health and performance. We are exploiting a model legume, *L. japonicus*, and a crop legume, white clover (*Trifolium repens*), to study both rhizobium bacteria (free-living symbionts) and AM fungi (AMF; obligate symbionts). This project has two main aims: (a) to identify and analyze plant genes involved in early recognition events in interactions of legumes with their rhizobium and AMF symbionts, and (b) to identify and analyze plant genes involved in efficient functioning in rhizobium and AMF symbioses.

Gene expression analysis has provided evidence of similarities between nodulation and mycorrhizal colonization. We have created and are exploiting populations of mutants (including Nod^- and Myc^-; Bonfante et al., 2000; Novero et al., 2002) and genetically tagged transformants (promoterless-*GUS*) of the model *L. japonicus* to dissect common early stages in development of both symbioses. We have identified putative promoter regions and genes from these promoter-trapped plants (Webb et al., 2000). In one of these lines, β-glucuronidase (GUS) is expressed in mature and senescing nodules, whereas in another, GUS expression is apparent soon after challenge with rhizobium. This line is providing a tool for further analysis of symbiotic signaling. We have also identified plants expressing GUS in other tissues, such as embryos, from the population of 284 independently tagged lines.

We are also exploiting unique material generated at IGER: near-isogenic lines (NILs) of white clover that show phenotypic differences in plant response to AM infection (Eason et al., 2001). We have isolated functional phenotypes of NILs of white clover, which vary in AMF colonization and effectiveness. We have identified potential differences in gene expression in both leaves and roots, using differential display. Some of these sequences have homology to EST sequences reported by other researchers working with AMF. We plan to generate ESTs from these white clover lines, targeting NILs with contrasting phenotypes and specific cells involved in the symbiotic interaction, to yield more precise genetic information on establishment and maintenance of successful symbioses.

M. TRUNCATULA

Toward the Complete Gene Inventory and Function of the *M. truncatula* Genome

PI: Douglas Cook, University of California, Davis, drcook@ucdavis.edu; Co-PI: Steve Gantt, University of Minnesota, steve@cbs.umn.edu; Co-PI: Michael G. Hahn, University of Georgia, hahn@ccrc.uga.edu; Co-PI: Maria Harrison, The Samuel Roberts Noble Foundation (SRNF), harrison@noble.org; Co-PI: Dongjin Kim, University of California, Davis, djim@ucdavis.edu; Co-PI: Ernest Retzel, University of Minnesota, ernest@mail.ahc.umn.edu; Co-PI: Deborah Samac, U.S. Department of Agriculture (USDA) and University of Minnesota, dasamac@tc.umn.edu; Co-PI: Christopher Town, The Institute for Genomics Research (TIGR), cdtown@tigr.org; Co-PI: Kathryn A. VandenBosch, University of Minnesota, kvandenb@cbs.umn.edu; Co-PI: Carroll Vance, USDA and University of Minnesota, vance004@maroon.tc.umn.edu; Co-PI: Nevin Young, University of Minnesota, neviny@tc.umn.edu; Collaborators: David Bird, North Carolina State University, david_bird@ncsu.edu; Julia Frugoli, Clemson University, jfrogol@clemson.edu; Michael Grusak, USDA-Agricultural Research Service (ARS) Children's Nutrition Research Center, mgrusak@bcm.tmc.edu; and Gary Stacey, University of Missouri, staceyg@missouri.edu

This work was supported by the National Science Foundation (NSF; Plant Genome Project no. DBI–0110206, a continuation of DBI–0196179; http://www.medicago.org/).

This project involves the large-scale genomic analysis of *M. truncatula*, a model legume with a small genome for efficient molecular, genetic, and reverse genetic analyses (Barker et al., 1990; Cook, 1999). This multiinstitutional U.S. effort is part of an international collaboration to develop a complete inventory and functional analysis of the *Medicago* genome. The emphases of this project include: (a) construction of genetic and physical maps (involving participants Douglas Cook, Dongji Kim, and Nevin Young), (b) analysis of gene function in legume biology (Steve Gantt, Michael G. Hahn, Maria Harrison, Deborah

Samac, Christopher Town, Carroll Vance, Kathryn A. VandenBosch, Nevin Young, and collaborators), and (c) analysis and public distribution of data (Ernest Retzel and Christopher Town). The results are expected to accelerate the discovery of agronomically important genes, in both *Medicago* and crop legumes, and to enhance understanding of gene and genome evolution within the Leguminosae.

A Sequence-Based M. truncatula Genetic Map Facilitates Comparisons between Species

To establish a comparative genetic map, we mapped conserved markers in the cool season legumes *M. truncatula*, alfalfa (*Medicago sativa*), and pea (*Pisum sativum*) in the galegoid clade, and in mung bean (*Vigna radiata*) and soybean, which are tropical legumes in the tribe Phaseoleae (for classification, see Doyle and Luckow, 2003). The use of markers developed from genes or gene-containing bacterial artificial chromosome (BAC) clones facilitated comparison of homologous sequences across species. Among the sequence-based markers used, EST-based microsatellite (SSR) markers were developed in collaboration with T. Huguet and are instrumental for coordinating the U.S. and French *M. truncatula* maps, which use different mapping populations. A collaboration with T. Bisseling integrated the genetic and cytogenetic maps by hybridizing BAC clones to pachytene chromosomes (Kulikova et al., 2001).

Extensive nucleotide conservation facilitated comparative mapping of *M. truncatula* and alfalfa (a collaboration with G. Kiss), which demonstrated a high level of synteny between the *Medicago* spp. A comparative map between *M. truncatula* and pea was generated in collaboration with N. Ellis. The colinearity of genes was well conserved in most linkage groups, with several inferred rearrangements. Comprehensive analysis between MtLG6 and the pea genome was hampered by a large proportion of heterochromatin (Kulikova et al., 2001) and the low frequency of non-RGA genetic markers on MtLG6. Together with T. Bisseling, we compared the *SYM2* region of pea with the orthologous region in *M. truncatula* and constructed a BAC contig bridging the region in *M. truncatula* (Gualtieri et al., 2002). The utility of the *M. truncatula* physical and genetic tools has implications for gene cloning experiments in legumes with high levels of synteny to *M. truncatula* in the region of interest.

Comparative analysis between *Medicago* and the more distantly related soybean was undertaken. Experiments exploring macrosynteny identified at least eight syntenic blocks. BAC hybridization was used to identify microsyntenic regions shared throughout the genomes (Yan et al., 2002). To further compare genome organization between *M. truncatula* and members of the Phaseolidae tribe, we focused on the diploid mung bean. To date, nine syntenic blocks are apparent. Ongoing work comparing these three species emphasizes markers derived from BAC contigs anchored to the *M. truncatula* map.

Syntenic relationships between *M. truncatula* and Arabidopsis also have been explored. A detailed analysis of findings is presented elsewhere in this issue (Zhu et al., 2003).

M. truncatula Resistance Gene Evolution and Genomic Organization

Most plant disease resistance genes belong to the nucleotide-binding site (NBS) LRR family. These genes can be classified by the presence or absence of a *Toll*/interleukin receptor domain (TIR) region. Retrieval of *M. truncatula* sequences homologous to the NBS domain of resistance genes identified at least 150 resistance gene analogs, of which more than 100 have been mapped (Cannon et al., 2002; Zhu et al., 2002). Phylogenetic analysis classified these sequences into several clades within the TIR and non-TIR subfamilies. Comparison of legume and nonlegume resistance gene homologs indicates that legume genes possess a unique evolutionary history, with many clades either unique to legumes or expanded within legumes (Cannon et al., 2002). We found that the origins of the major resistance gene clades appear to predate radiation of these Papilionoid species, and their diversification mirrors predicted speciation events.

The sequence diversity of resistance gene homologs in *M. truncatula* is similar to diversity found in other legume crops. There is strong evidence for a conserved location of several loci between *M. truncatula* and soybean, and between *M. truncatula* and pea. In each case not only is synteny evident, but also the resistance gene homologs themselves have a conserved phylogenetic position. These results indicate that *M. truncatula* will be a useful reference for NBS-LRR genes within the legume subfamily Papilionoideae.

Development of a Medicago Physical Map to Aid Positional Cloning and Full Genome Sequencing

A DNA fingerprinting approach is being used to assemble the physical map of *M. truncatula* based on BAC libraries (Nam et al., 1999; D.J. Kim and D.R. Cook, unpublished data; F. Debelle and J. Dénarié, unpublished data). The existing contig map covers approximately 480 Mbp, or 95% of the genome. Several hundred BAC clones have been correlated with ESTs, and the physical map is linked to the genetic map by means of ESTs and SSR markers on BAC contigs.

A consortium in Europe and the U.S. has established the following objectives for constructing the sequence-ready physical map by August 2003: (a) completion of a 20× genome coverage BAC finger-

print, at University of California (Davis); (b) extensive integration of genetic and physical maps, involving investigators in the U.S., Hungary, and France; (c) acquisition of BAC end sequence data for all clones within the physical map; and (d) cytogenetic analysis of pachytene chromosomes to define transition points between euchromatic and heterochromatic regions (to be completed in Wageningen). Mapped clones are currently being sequenced at the University of Oklahoma (B. Roe; see below). The sequenced BAC clones will establish seed points for an international effort to complete the sequence of the *M. truncatula* genome.

Analysis of Medicago Gene Function during Interactions with Microbes

Analysis of gene expression patterns is being used to gain insight into *Medicago* genome function. We have sequenced random cDNAs from a wide variety of organs and conditions. Of the approximately 180,000 *M. truncatula* ESTs now publicly available, about 80,000 were derived from the 17 project libraries. This work emphasizes interactions of *M. truncatula* with microbes, and two-thirds of the project's ESTs are from tissues responding to symbionts, pathogens, or elicitors.

At TIGR, the *Medicago* gene index (http://www.tigr.org/tdb/tgi/mtgi/) groups ESTs from *M. truncatula* into contigs to produce tentative consensus sequences (TCs). In silico analysis of gene expression was used to assess transcription patterns of genes that are highly expressed. We sorted TCs to identify predicted genes with particular expression patterns, based on the libraries of origin of the ESTs in the TCs. For example, Fedorova et al. (2002) identified nodule-specific expressed genes, including some unique classes of genes. We have recently released MtDB, a new database that clusters all public *M. truncatula* ESTs into about 17,000 contigs plus a similar number of singletons (Lamblin et al., 2003; http://www.medicago.org/MtDB). MtDB offers unrestricted access to novel online query tools that require no programming or database knowledge.

A more detailed understanding of gene expression patterns is being obtained using microarrays. We constructed an array of approximately 1,000 nonredundant cDNA clones that has been used in pilot microarrays to standardize hybridization conditions and data analysis. A broader set of 6,000 clones, obtained from all project libraries, is now in use as a step toward construction of a comprehensive unigene set. The 6-K set has been resequenced, and after annotation, it will be available for public distribution. Current gene expression profiling experiments monitor responses to elicitors, the symbiotic microbes *Sinorhizobium meliloti* and *Glomus versiforme*, and the pathogens *Phytophthora medicaginis*, *Colletotrichum trifolii*, *Erysiphe pisi*, and *Xyllela fastidiosa*. In addition, vegetative and reproductive development and responses to nutrient stress will be examined in wild-type and developmental mutants (e.g. Penmetsa and Cook, 2000).

Sifting for Novel Expressed Sequences among Legume ESTs

Early inspection of *Medicago* ESTs identified many sequences that appeared to be unique to legumes. Recently, we have made a comprehensive search among *M. truncatula*, *L. japonicus*, and *Glycine* spp. ESTs for sequences with no known homologs outside the Leguminosae. This was done by comparing, using BLAST algorithms, legume EST contigs with ESTs of other angiosperms, the National Center for Biotechnology Information (NCBI) nonredundant database, and the genome sequences of Arabidopsis and rice (*Oryza sativa*). In *Medicago*, over 500 apparently legume-specific TCs, or "leguminosins," were identified. Some of these sequences appear to be members of gene families, based on clustering analyses. One spectacular example is a group of more than 300 putatively secreted proteins that are Cys rich and were previously identified as nodule specific (Fedorova et al., 2002). Despite the lack of apparent sequence similarity of leguminosins to known sequences from other organisms, motif searching has provided clues to the functions for some members of this group. Full-length sequencing of about 400 cDNA clones and identification of genomic clones are currently under way. Future work will highlight expression profiling, promoter analysis, and phylogenetic relationships of these genes.

The Institut National de la Recherche Agronomique (INRA) Project. Genetics and Genomics of the Model Legume *M. truncatula*

Coordinator: Jean Dénarié, Centre National de la Recherche Scientifique (CNRS)-INRA, Castanet-Tolosan, France, denarie@toulouse.inra.fr; PI: J.M. Prosperi, INRA, Montpellier, France, prosperi@ensam.inra.fr; PI: T. Huguet, CNRS-INRA, Castanet-Tolosan, France, thuguet@toulouse.inra.fr; PI: C. Rameau, INRA, Versailles, France, rameau@versailles.inra.fr; PI: P. Gamas, CNRS-INRA, Castanet-Tolosan, France, gamas@toulouse.inra.fr; PI: J. Gouzy, CNRS-INRA, Castanet-Tolosan, France, gouzy@toulouse.inra.fr; PI: B. Tivoli, INRA, Rennes, France, tivoli@rennes.inra.fr; and PI: R. Thompson, INRA, Dijon, France, thompson@epoisses.inra.fr

This work was supported by INRA (AIP 242; http://medicago.toulouse.inra.fr/ATS/).

This project has two main objectives: (a) to contribute to the development of *M. truncatula* genetic and

genomic resources, and (b) to initiate studies to prepare for transfer of genetic and genomic information collected on this model species to cultivated legumes important in France such as pea and alfalfa. The project involves 17 laboratories, will last 3 years (2001–2003), and is organized in the following areas.

Genetic Resources and Development of Genetic Maps

It is essential for a model species to develop important core collections representative of a broad genetic diversity and containing reference material that is well defined for a number of biological and molecular characters (Bonnin et al., 2001). To this end, identification of reference populations and lines representative of populations, characterization of the lines with molecular markers, and construction of RILs is being coordinated by J.M. Prosperi.

Genetic maps are important tools for the broader scientific community for positional cloning of genes and for interspecific comparisons of genome structure (Thoquet et al., 2002). T. Huguet is leading the development of standard tools, including a high-resolution genetic map and a collection of microsatellite markers (SSR) derived from ESTs. The SSR markers will be used for mapping ESTs.

Our objective in comparative mapping is to align genetic maps of *M. truncatula*, alfalfa, and pea by using *M. truncatula* ESTs corresponding to pea genes that have already been mapped, using *M. truncatula* ESTs containing SSR markers. The goal of this work, overseen by C. Rameau, is to map loci affecting traits of agronomic interest, such as plant architecture, resistance to selected pathogens, flowering time, and nutritional value of seeds. J. Gouzy leads the development of bioinformatic tools for genetic mapping of *M. truncatula* and for comparative mapping with pea and alfalfa, as well as the development of a Web site for integration of heterogeneous data banks.

Functional Genomics

Our transcriptomics program, coordinated by P. Gamas, makes use of a collection of 25,000 3′ and 5′ ESTs obtained by a collaboration between INRA and Génoscope, as well the international *M. truncatula* ESTs that are available. Tools for EST analysis and annotation and for expression profiling have been adapted or developed by Jerome Gouzy. The EST database (Journet et al., 2002) may be accessed at http://medicago.toulouse.inra.fr/Mt/EST. Macro- and microarrays are being used to analyze gene expression in various conditions, including symbiotic interactions with *Rhizobium* and endomycorrhizal fungi, interactions with fungal pathogens and insects, seed development, and N assimilation.

Analysis of interactions of *M. truncatula* with pathogens and pests isolated from legume crops will be used to identify new possible sources of resistance, in an effort coordinated by B. Tivoli. The pests under study include aerial fungal pathogens (*Mycosphaerella pinodes*, *Ascochytapisi* sp., *C. trifolii*, and *Ascochyta fabae*), root pathogens (*Aphanomyces euteiches*, *P. medicaginis*, *Phoma medicaginis*, and *Fusarium oxysporum*), the stem nematode *Ditylenchus dipsaci*, aphids (*Acyrthosiphon pisum*), and the seed-eating insect *Sitophilus* sp. The interactions will be studied on genotypes selected from the Montpellier collection.

As a model for seed development and storage protein accumulation, *M. truncatula* will facilitate the study of seed biology of grain legumes such as pea. Likewise, *M. truncatula* is also an effective model for stem and leaf development in forage legumes such as alfalfa. These efforts, coordinated by R. Thompson, will also analyze the genetic variability of seed composition and use proteomics to characterize seed formation in both *M. truncatula* and pea.

The Integrated Structural, Functional, and Comparative Genomics of the Model Legume *M. truncatula*

Coordinator: Jean Dénarié, INRA-CNRS, Toulouse, France, denarie@toulouse.inra.fr; Vivienne Gianinazzi-Pearson, INRA-CMSE, Dijon, France, gianina@epoisses.inra.fr; Alfred Puehler, Universität Bielefeld, Germany, puehler@genetic.uni-bielefeld.de; Helge Kuester, Universität Bielefeld, Germany, helge.kuester@genetik.uni-bielefeld.de; Philipp Franken, Max Planck Institute for Terrestrial Microbiology, Germany, frankenp@mailer.uni-marburg.de; Jörn Kalinowski, Universität Bielefeld, Germany, joern.kalinowski@genetic.uni-bielefeld.de; Adam Kondorosi, CNRS, Gif-sur-Yvette, France, kondorosi@isv.cnrs-gif.fr; Michael Schultze, University of York, UK, ms47@york.ac.uk; Noel Ellis, John Innes Centre, UK, noel.ellis@bbsrc.ac.uk; György Kiss, Biological Research Center of the Hungarian Academy of Sciences, Hungary, kgb@nucleus.szbk.u-szeged.hu; Ton Bisseling, Wageningen University, The Netherlands, ton.bisseling@wur.nl; and Anne Schneider, European Association for Grain Legume Research, France, a.schneider-aep@prolea.com

This is project no. QLG2–CT–2000–00676 of the European Commission Fifth Framework Program (http://medicago.toulouse.inra.fr/EU/).

This large consortium includes 11 participating sites in five European countries (France, Germany, Hungary, The Netherlands, and UK). The principal investigators at the project's 11 sites are listed above. The diverse expertise of the participants allows a comprehensive approach to genomics of *M. truncatula*.

Comparative Analysis of Legume Genome Structure

Investigators in Toulouse, France and Szeged, Hungary are constructing a consensus genetic map of *M. truncatula*, using several F_2 mapping populations and RILs (Thoquet et al., 2002). The mapping of EST-based microsatellites is facilitated by their strong polymorphism among *M. truncatula* genotypes. The conservation of EST sequences among closely related crop legumes, such as pea, fava bean (*Vicia faba*), alfalfa, and clovers, will further the comparative mapping of these species with *M. truncatula*. Recent comparisons of the maps of *M. truncatula* and diploid alfalfa, which both have a haploid chromosome number of 8, showed a high degree of synteny between the two species. The chromosomal location of a few loci was found to differ, including the nucleolar-organizing regions.

Pea, a major European grain legume, has a genome size about 10 times larger than *Medicago*, with one fewer chromosome per haploid genome than *M. truncatula* (Ellis and Poyser, 2002). Participants in the UK, Hungary, and The Netherlands are cooperating to take advantage of the close phylogenetic relationship of these two legumes and the simpler genome of *M. truncatula* to advance the genomic analysis of pea. Five of seven linkage groups in pea were found to be largely syntenic to five of the eight chromosomes in *Medicago*. The remainder of the pea genome shows evidence that gene rearrangements and one large duplication event have taken place since the divergence from *Medicago*. Microsynteny between *M. truncatula* and pea is studied using DNA clone cross hybridization and fluorescent in situ hybridization technology (Gualtieri et al., 2002). Other temperate grain legumes, including fava bean, lentil (*Lens culinaris*), and chickpea (*Cicer arietinum*), show similar gene arrangements, and, therefore, can also benefit from advances in *Medicago* genomics.

Cytogenetic mapping by investigators in The Netherlands has demonstrated that *M. truncatula* has a relatively simple organization where the condensed chromatin, which comprises more than 50% of the DNA, is clustered around the centromeres. Therefore, the remaining gene-rich euchromatin occurs in long stretches on the chromosome arms that are largely uninterrupted by repetitive DNA-containing heterochromatin (Kulikova et al., 2001). These same investigators have developed fluorescent in situ hybridization methods for *M. truncatula*, using pachytene chromosomes. Using BAC clones (in cooperation with laboratory of Douglas Cook; see above) as probes, a high-resolution cytogenetic map has been obtained and integrated with the genetic map. The integrated map is a valuable tool for positional cloning. Moreover, the simple genome organization of the *M. truncatula* genome suggests that an efficient strategy for sequencing the majority of the protein coding regions would be to make the sequencing of the gene-rich euchromatic regions the first priority.

Positional cloning of *M. truncatula* genes has become reality with the cloning by the Hungarian participants of a receptor kinase required for nodulation and mycorrhizal development (NORK; Endre et al., 2002). Several nodulation-defective mutants identified in alfalfa, *M. truncatula*, and pea turned out to be conditioned by orthologous genes. Comparative mapping of these loci aided the physical mapping and verification of the identity of the NORK gene. Map-based cloning of other symbiotic genes has been initiated (Ané et al., 2002).

Functional Genomics

Participants in the *Medicago* project at six sites, including Toulouse, Dijon, and Gif-sur-Yvette in France, and Marburg and two groups in Bielefeld, Germany, are cooperating in transcriptional profiling. For analyzing gene expression during root development and root symbioses, a 6,000-element array has been constructed using cDNA clones. An additional 1,700 consensus sequences have been obtained from ESTs derived from flowers and developing seeds. Together, the nearly 8,000 cDNAs are being used to construct microarrays that will be useful for analyzing gene expression under a broad variety of conditions (see Weidner et al., 2003). In parallel, specific macroarrays have been developed to refine the analysis of gene profiling in certain conditions (Wulf et al., 2003). The functional genomics team has worked with their counterparts in the U.S. (see summaries on the projects of Cook et al. and May et al.) to standardize a set of control clones.

Researchers in Gif-sur-Yvette and Toulouse have developed two efficient systems for transformation. The first uses infiltration of *A. tumefaciens* into leaf or flower explants, followed by somatic embryogenesis and regeneration. This method is being used to construct reporter lines, to produce plants with overexpression or inactivation of candidate genes, and for gene tagging strategies. The second method uses *Agrobacterium rhizogenes* to produce transformed hairy roots (Boisson-Dernier et al., 2001). The *Medicago* groups at Gif and York, UK, are cooperating in the development of gene tagging systems, using T-DNA (DNA; Scholte et al., 2002) and the tobacco (*Nicotiana tabacum*) retrotransposon Tnt1. This transposable element has been found to be active in *M. truncatula* (d'Erfurth et al., 2002), and about 1,000 genes have already been tagged. Together, these two approaches will complement transcriptional profiling for more efficient gene discovery in this species.

Center for *Medicago* Genomics Research

Coordinator: Gregory D. May, Samuel Roberts Noble Foundation (SRNF), gdmay@noble.org; Principal

Investigators: Lloyd W. Sumner, SRNF, lwsumner@noble.org; Richard A. Dixon, SRNF, radixon@noble.org; Richard S. Nelson, SRNF, rsnelson@noble.org; Maria J. Harrison, SRNF, mjharrison@noble.org; Robert A Gonzales, SRNF, ragonzales@noble.org; Liangjiang Wang, SRNF, lwang@noble.org; Xiaoqiang Wang, SRNF, xwang@noble.org; Nancy L. Paiva, SRNF, nlpaiva@noble.org; Kiran Mysore, SRNF, ksmysore@noble.org; Rujin Chen, SNRF, rchen@noble.org; and Elison Blancaflor, SRNF, eblancaflor@noble.org

This work was supported by the SRNF (http://www.noble.org/medicago/index.htm).

A Center for *Medicago* Genomics Research was established at the SRNF in the fall of 1999. We have taken a global approach in the study of the genetic and biochemical events associated with the growth, development, and environmental interactions of *M. truncatula*. Our approach includes large-scale EST sequencing, gene expression profiling, and high-throughput metabolite and protein profiling. We are interfacing these multidisciplinary data types to provide an integrated set of tools to address fundamental questions pertaining to legume biology. These questions include the analysis and understanding of: (a) the biosynthesis of natural products that affect forage quality and human health, (b) the cellular and molecular basis for the directional growth response of roots to gravitropism and the role of the cytoskeleton in this process, (c) legume root development and elucidating molecular mechanisms of polar auxin transport, (d) non-host pathogen resistance, (e) the RNA silencing pathway, and (f) the use of *M. truncatula* in combination with an AM fungus *G. versiforme* for analyses of the AM symbiosis.

The *Medicago* Genome Initiative, established at the National Center for Genome Resources, is a database of EST sequences of the model legume *M. truncatula* (Bell et al., 2001). From January 2000 to December 2002, almost 100,000 *M. truncatula* ESTs have been characterized at SRNF, and a total of approximately 175,000 have been characterized worldwide. Unidirectional cDNA libraries representing different stages of *M. truncatula* development and exposure to biotic and abiotic stresses have been generated. The international *Medicago* research community has characterized ESTs from more than 24 different cDNA libraries. The goal of the SRNF's EST project is to identify and characterize 20,000 to 40,000 unique *Medicago* cDNA isolates.

Changes in gene expression underlie many biological phenomena. The use of DNA microarrays and serial analysis of gene expression (SAGE) will provide insights into tissue- and developmental-specific expression of genes and the response of gene expression to environmental stimuli. Qiagen Operon, in collaboration with SRNF, Chris Town (The Institute for Genomic Research), and Kathryn VandenBosch (University of Minnesota), are developing a commercially available Array Ready Genome Oligonucleotide Set for *M. truncatula*. This set of 16,000 bioinformatically optimized oligonucleotides will be used as probes in our microarray analysis, and will provide a uniform platform for gene expression analysis around the globe.

The protein complement of the genome, the proteome, serves as a biological counterpart to the *Medicago* EST and gene expression analyses. Given that many biological phenomena lack the requirement for de novo gene transcription, proteomics studies provide a mechanism to study proteins and their modifications under developmental changes and in response to environmental stimuli. An automated system has been established for the electrophoretic separation of complex protein mixtures and differential analysis to discover changes in proteome content (Asirvatham et al., 2002; Watson et al., 2003).

A state-of-the-art biological MS laboratory has been established as part of the *Medicago* genomics activities (Sumner et al., 2002). Instrumentation within the laboratory included liquid chromatography/MS, GC/MS, and matrix-assisted laser-desorption ionization time of flight MS. *M. truncatula* ecotypes and elicited cell cultures will be screened for changes in the levels of a wide range of primary and secondary metabolites.

Our aim is to develop a program that will integrate gene expression and protein and metabolite profiling in conjunction with *M. truncatula* genetics to provide a global view of *Medicago* biology. *Medicago* resources are publicly available through the SRNF.

BAC-Based Sequencing of the Gene-Rich Euchromatic Regions of *M. truncatula*

PI: Bruce A. Roe, The University of Oklahoma, broe@ou.edu

This work was supported by the SRNF (http://www.genome.ou.edu).

To better understand biological processes associated with legumes, as a pilot project we initially cloned and end sequenced 25,000 random 2-to 4-kb whole-genome shotgun fragments. Although this data represented only 2% of the approximately 500 million-bp *M. truncatula* genome, approximately 1,000 individual sequence reads match the *M. truncatula* EST database, including an approximately 20-kb contig matching 26S rRNA genes and several classes of extensive repeated sequences representing between 60% and 80% of the genome. In addition, we now have completed the 124,034-bp sequence of the *M, truncatula* chloroplast genome, which has been deposited into GenBank (accession no. AC093544).

Funding from the SRNF is allowing us to obtain the working draft (4–6-fold shotgun coverage) of mapped BACs that are provided to us by Douglas

Cook and Dongjin Kim (University of California). To date, we have approximately 480 BACs in our sequencing pipeline and anticipate that by September 2003, we will have obtained the working draft sequence of an additional 480 BACs representing approximately 10% of the *M. truncatula* genome and over one-half of the gene coding regions in the genome. By September 2003, we also propose to have sequenced approximately one-half of the approximately 960 BACs to phase 2, i.e. ordered and oriented contigs.

Our present sequencing strategy entails random shotgun cloning of isolated and mapped BAC clones into pUC vectors, followed by colony picking on a Q-pick system and cell growth in a 384-well format in HiGrow shaker-incubators. The template isolations now are fully automated and performed on a Zymark 384 well pipetting SiClone robot with Twister II arm designed with four 384-well plate shakers, a newly designed, key component in our work flow. The cycle sequencing reactions with non-fluorescent universal forward and reverse primers, and fluorescent-labeled terminators are pipetted on a Velocity 11 V-Prep and incubated in 384-well format ABI Viper thermocyclers. After electrophoresis and data collection on the ABI 3700 and data transfer, the individual contigs assembled by Phrap are analyzed by several automated scripts to predict efficient closure strategies. An Oracle database has been developed to facilitate automated data analysis and preliminary sequence annotation. Implementation of these robotic procedures and automated data analysis protocols has reduced the cost of finished DNA sequence to less than 10 cents/base while improving the overall efficiency from cells to sequence to over 95%.

In accord with the Bermuda Rules for data access, all our sequence data is being deposited into GenBank within 24 h after it is assembled into contigs greater than 2 kb. A genome browser showing the results of our automated annotation pipeline is available from links on our Web site (http://www.genome.ou.edu).

An Integrated Approach to Functional Genomics and Bioinformatics in a Model Legume

PI: Pedro Mendes, Virginia Bioinformatics Institute, mendes@vt.edu; Co-PI: Richard A. Dixon, SRNF, radixon@noble.org; Co-PI: Lloyd W. Sumner, SRNF, lwsumner@noble.org; Co-PI: Gregory D. May, SRNF, gdmay@noble.org; and Co-PI: J. Tim Smith, Southeastern Oklahoma State University, tsmith@sosu.edu

This work was supported by the NSF (Plant Genome Project No. 0109732; http://medicago.vbi.vt.edu).

This project studies the responses of *M. truncatula* root cell cultures to three elicitations: exposure to high-UV light, methyl-jasmonic acid, and yeast elicitor. During the course of this project, methodologies are being developed to produce, store, and analyze integrated functional genomics data sets including the transcriptome, the proteome, and the metabolome. Analysis of these data sets is expected to facilitate the functional identification of many new genes and proteins in *M. truncatula* and other legumes, many of which will be associated with natural product biosynthesis. We aim to explain in a quantitative way, through computer models and simulations, the working of the cellular machinery as plant cells reprogram themselves in response to the three stresses. This information will facilitate future manipulations of legumes to improve beneficial traits such as disease resistance, radiation protection, and nutritional content by decreasing the levels of toxic natural products and increasing beneficial ones.

There is currently little knowledge available concerning the biosynthesis of the large majority of plant natural products, many of which are lead compounds for pharmacological drug development. This project is aimed at identification of new natural products and, importantly, also the genes involved in their biosynthesis. The data generated consist of hundreds of microarrays, two-dimensional gels, chromatography and electrophoresis runs, and various types of mass spectra. These data, together with processed data, consisting of ratios of protein, mRNA, and metabolite levels, will be available through a public database system that will provide and integrated view of the transcriptome, proteome, and metabolome.

MolMyk. Molecular Basics of Mycorrhizal Symbioses

PI: Alfred Pühler, Universität Bielefeld, Germany, puehler@genetik.uni-bielefeld.de; Co-PI: Helge Küster, Universität Bielefeld, helge.kuester@genetik.uni-bielefeld.de; Co-PI: Karsten Niehaus, Universität Bielefeld, Karsten.Niehaus@Genetik.Uni-Bielefeld.DE; Co-PI: Andreas Perlick, Universität Bielefeld, Andreas.Perlick@Genetik.Uni-Bielefeld.DE; Co-PI: Philipp Franken, Institute for Vegetable and Ornamental Plants, Grossbeeren, franken@igzev.de; Co-PI: Thomas Fester, Leibniz-Institut für Pflanzenbiochemie, Halle, tfester@ipb.uni-halle.de; Co-PI: Wilibald Schliemann, Leibniz-Institut für Pflanzenbiochemie, Halle, wschliemann@ipb-halle.de; Co-PI: Dieter Strack and Bettina Hause, Leibniz-Institut für Pflanzenbiochemie, Halle, dstrack@Ipb.uni-Halle.DE, bhause@ipb.uni-halle.de; Co-PI: Michael H. Walter, Leibniz-Institut für Pflanzenbiochemie, Halle, mhwalter@ipb-halle.de; Co-PI: Franziska Krajinski, Universität Hannover, krajinski@lgm.uni-hannover.de; Co-PI: Hermann Bothe, Universität zu Köln, hbothe@novell.biolan.uni-koeln.de; Co-PI: Natalia Requena-Sanchez, Universität Tübingen, Natalia.Requena@Uni-Tuebingen.de; Co-PI: Rainer Hedrich

and Peter Ache, Universität Würzburg, hedrich@botanik.uni-wuerzburg.de, ache@botanik.uni-wuerzburg.de; and Co-PI: Rainer Kaldenhoff, Universität Würzburg, kaldenhoff@botanik.uni-wuerzburg.de

This is Deutsche Forschungsgemeinschaft Focus Program no. SPP1084 (MolMyk; http://MolMyk.Genetik.Uni-Bielefeld.DE).

More than 80% of flowering plants are able to form either endomycorrhiza or ectomycorrhiza symbioses with certain groups of fungi. In return for carbon compounds, the symbiotic fungi supply phosphate, N, potassium, and trace elements to the plant, resulting in an increased plant growth and an improved resistance against biotic and abiotic stress.

The targeted molecular analysis of plant endo- and ectomycorrhiza is difficult due to the complex development of the interaction. The Deutsche Forschungsgemeinschaft Focus Program MolMyk was established in the year 2000 to apply untargeted techniques of plant genomics to identify plant genes essential for the establishment and functioning of mycorrhizal symbioses. We have chosen the interaction of *M. truncatula* with *Glomus* spp. as a model for endomycorrhiza and the interaction of poplar (*Populus tremula*) with *Amanita muscaria* as a model for ectomycorrhiza.

Within MolMyk, the collaboration of groups working on *M. truncatula* endomycorrhiza is supported primarily by establishing and applying modern methods and technologies of plant genomics and bioinformatics. As a basis for this collaboration, we have carried out and continue to perform an EST sequencing project using the interaction between *M. truncatula* and *G. intraradices* as a model. During this project, two cDNA libraries of *M. truncatula* endomycorrhiza were constructed and sequenced: a general cDNA library of mycorrhizal roots harvested at different stages of development and a suppression-subtractive hybridization cDNA library that was enriched for mycorrhiza-specific or mycorrhiza-amplified transcript sequences. To date, about 6,500 ESTs from these two libraries were generated and deposited in the EMBL and GenBank databases. Using the software tool BioMake that was developed within the MolMyk project, all EST sequences obtained were clustered and automatically annotated. Using electronic northern approaches, in silico patterns for the genes represented by the MolMyk EST collection were obtained based on comparisons with the TIGR *M. truncatula* Gene Index (http://www.tigr.org/tdb/mtgi/).

In collaboration with the French EST sequencing project "Functional Genomics in *M. truncatula* EST Analysis as a Tool to Explore *M. truncatula* Root Symbiotic Program" (http://medicago.toulouse.inra.fr/Mt/EST/DOC/MtB.html), we constructed macro- and microarrays covering the root interaction transcriptome of *M. truncatula*. These arrays, designated 6k-RIT, comprise about 6,000 EST clusters that are derived from cDNA libraries from root nodules, root endomycorrhiza, and uninfected *M. truncatula* roots. To widen the scope of our DNA arrays, we are switching to the construction and hybridization of 70-mer oligonucleotide microarrays covering about 16,000 *M. truncatula* EST clusters that are based on the TIGR *M. truncatula* Gene Index (see May et al., above). Both the 6k-RIT cDNA and the 16k oligo arrays are and will be used within MolMyk to obtain expression profiles of genes relevant for root endomycorrhiza under different symbiotic or physiological conditions. To evaluate expression profiles generated from DNA arrays, the EMMA ("ESTs meet microarrays") software was developed as an integral part of the bioinformatics section of MolMyk. The EMMA software is part of a relational database that allows the linkage of EST annotations with expression profiles obtained on the DNA arrays. The information gained by this integrated genomics approach will be complemented by functional analyses using gene silencing techniques in transgenic plants and transgenic roots of *M. truncatula*.

EST sequence and expression data obtained within the MolMyk genome project are mined by the project participants to address biological questions related to different aspects of legume endomycorrhiza. These questions can be grouped into three main sections. The section "development of mycorrhizae" covers the identification of signals involved in the recognition of symbiotic partners, the analysis of signal transduction pathways playing a role in the establishment of the symbiosis, and the identification of legume symbiosis-specific genes (symbiosins) that are common to the root nodule and the root endomycorrhiza symbiosis. The second section, "fluxes," addresses the transport of monosaccharides, amino acids, N and sulfur compounds, and the adaptation of the plant's and the fungal metabolism to the symbiotic status. Finally, applied aspects, such as mycorrhiza-induced plant resistance, are considered.

We envisage that the MolMyk project will deliver an integrated view of sequence properties, expression patterns, and functions for *M. truncatula* genes that are relevant for different symbiotic and physiological conditions of an arbuscular mycorrhiza.

Progress Report on TILLING. High-Throughput Genotyping in *M. truncatula*

PI: Douglas Cook, University of California, Davis, drcook@ucdavis.edu; and Co-PI: R. Varma Penmetsa, University of California, Davis, rvpenmetsa@ucdavis.edu

This work was supported by a U.S. DA-ARS cooperative agreement with the University of California (Davis).

For decades, biologists have used genetic screens to link the function of individual genes to specific

biological processes. Although powerful, such "phenotype-first" screens are limited in the types of individual genes and range of alleles that can be identified. In contrast, "genotype-first" screens identify allelic variation in genes independent of phenotype. Genotype-first screens allow a systematic search for variation in any gene whose sequence is known and identify a greater diversity of alleles than do phenotype-first screens. Moreover, genotype-first screens are a logical follow-on from the massive EST sequencing projects that have been conducted for numerous plant species.

For reverse genetic analysis in *M. truncatula*, we have adapted the TILLING technology originally developed in Arabidopsis (McCallum et al., 2000). We have established an efficient pipeline for germplasm curation and DNA extraction from EMS mutagenized populations. To date, approximately 2,000 individual germlines have been prepared for genotyping, and an additional set of several thousand mutagenized, single-seed decent germlines are being prepared for incorporation into the screening pipeline. We have developed a high-throughput genotyping procedure based on the Cel I mismatch enzyme detection system to enable screening for single-nucleotide polymorphisms (SNPs) in genes of interest. DNA mismatches induced by EMS mutagenesis are cleaved by the Cel I endonuclease to identify plant lines containing SNPs (Oleykowski et al., 1998; Colbert et al., 2001). We have instituted a DNA multiplex strategy that involves pooling DNA from four unique germlines. The pooling increases throughput of screening and reduces cost, without compromising sensitivity of SNP detection.

To date, we have interrogated approximately 2,000 germlines for SNPs in 1-kb regions in four genes of interest. This screening has identified an average of approximately one SNP lesion per 1 kb per 500 germlines. After the incorporation of the remaining germlines to the existing group, it will be possible to screen an anticipated 4,000 unique germlines. We estimate that we can identify approximately eight SNPs in genes of interest, and this number would be augmented by screening of additional germplasm as this material is developed.

Proteome Analysis of Meristematic Development of *M. truncatula* and the Investigation of the Molecular Interactions between *Rhizobium*, Phytohormones, and Plant Nodulation Mutants

PI: Barry Rolfe, The Australian National University (ANU), Canberra, Australia, rolfe@rsbs.anu.edu.au; Collaborators and Co-Pis: Ulrike Mathesius, ANU, ulrike.mathesius@anu.edu.au; Michael Djordjevic, ANU, michael@rsbs.anu.edu.au; Jeremy Weinman, ANU, weinman@rsbs.anu.edu.au; Charles Hocart, ANU, hocart@rsbs.anu.edu.au; Georg Weiller, ANU, weiller@rsbs.anu.edu.au; Peter Gresshoff, University of Queensland, Australia; Ray Rose, University of Newcastle, Australia, birjr@alinga.newcastle.edu.au; Mohan Singh, University of Melbourne, Australia, mohan@unimelb.edu.au; and W. Dietz Bauer, Ohio State University, bauer.7@osu.edu

This work was supported by the ANU (block grant), by the U.S. Department of Agriculture (project no. 741734), by the University of Newcastle (project grant), and by the ARC Centre of Excellence (5-year support for CILR; http://biology.anu.edu.au/research-groups/gig/gig.html; http://semele.anu.edu.au/2d/2d.html).

Meristems Are the Key to Plant Growth and Adaptation

Common features of meristem development are: (a) a trigger for the activation of cell cycle genes, thought to be initiated by the phytohormones auxin and cytokinin; (b) the regulation of meristem maintenance; and (c) phenotypic plasticity of meristem development in response to external signals including nutrient availability. Plant architecture is regulated by the activity of apical and lateral meristems in the shoot and root, including the apical meristem and vegetative lateral meristems that control shoot branching and flower meristems. Plants must coordinate the growth of root and shoot meristems to maintain an appropriate balance of root and shoot organs and to respond and adapt to various environmental conditions. This balance is achieved by an inter-meristem coordination of growth and development of the plant and involves the interplay of several long-range signals (Jiang and Gresshoff, 2002; Maguire et al., 2002; Searle et al., 2003), and we are examining this coordination in both *M. truncatula* and *L. japonicus*.

Integrated Legume Research Program

We have instituted a joint program to investigate systems biology of legumes. Our hypothesis is that, with the likely exception of initial elicitors, meristem development has been evolutionarily conserved in the whole plant, and, thus, knowledge gained on the ontogeny of one meristem type will be applicable to the regulation of other meristems such as root apical, shoot apical, and floral meristems. Our approach is to use plant mutants along with proteome, transcriptome, and metabolome technologies, linked through bioinformatics, to analyze the dynamic interactions of genes and their products that are occurring during organogenesis in the model legumes.

Research Tools

Using a combination of two-dimensional gel electrophoresis (2-DE) and matrix-assisted laser-desorption ionization time of flight MS, we have established a proteome reference map of roots and leaves for the model legume, *M. truncatula*. Proteins were separated on 2-DE gels, yielding reproducibly 2,500 proteins per gels from root extracts. More than 200 root (Mathesius et al., 2001) and 100 leaf proteins have been identified by peptide mass fingerprinting using the *M. truncatula* EST or TC databases for reliable identification of the proteins. A database has been compiled (http://semele.anu.edu.au/2d/2d.html). The use of this proteome reference map was also tested against the proteomes of related legumes (Mathesius et al., 2002). Defined tissues and developmental stages are analyzed and compared for their protein patterns. We use statistical analysis of quantified protein spot volumes from 2-DE gels to quantify changes in protein expression between different tissues or treatments.

Meristem Development of M. truncatula Cells Grown in Culture

A valuable somatic embryogenesis system for *M. truncatula* has been developed using leaf cells into culture (Rose and Nolan, 1995). The first appearance of embryos from mesophyll protoplasts occurs between 5 to 7 weeks and has a reasonable degree of synchrony, thus enabling a developmental study of the molecular changes taking place in the dividing cells. Seven recognizable stages have been documented during the regeneration. This meristematic system has ideal attributes: the regenerative capacity of the mutant line 2HA, access to a genome sequence, conditional induction of meristematic growth exerted by abscissic acid or light, and the possibility of embryo formation, which is fundamental for legume transformation and regeneration. The collaborative group is using this system to investigate meristematic growth and differentiation in culture and have identified several differentially expressed proteins during the first developmental stages of somatic embryogenesis.

Legume Roots, Nodules, and Induction of Meristems

Legume nodules are a model not only for plant-microbe communication but also for meristem induction. We are comparing the meristem development of nodules and lateral roots because we hypothesized that similar developmental events are initiated in both cases (Mathesius et al., 2000). Unlike other meristems, the development of the nodule meristem is unique because its site, timing of initiation, the target cell type, and the ontogeny can be defined. We have established a baseline, protein expression during nodule and lateral root formation to identify infection-specific and root developmental proteins. To further characterize the nodule meristem, we are comparing nodulation-specific protein expression in the wild type with that of super-nodulating and non-nodulating *M. truncatula* mutants, including *sunn* (Penmetsa et al., 2003) and *skl* (Penmetsa and Cook, 1997). These two mutants show changes in auxin transport and ethylene sensitivity, respectively, and changes in protein expression have been correlated to their phenotypes.

In addition, we are interested in the signals of microbes that influence meristem development in *M. truncatula*. We are currently testing the effect of Nod factors, chitin oligosaccharides, exopolysaccharides, and *N*-acyl homo-Ser lactones, compounds used by bacteria for quorum sensing. We found multiple (>150) changes in protein accumulation after different treatments, compared with control root segments of the same developmental stage. Significant changes were found in proteins of diverse function, including defense, flavonoid metabolism, hormonal regulation, and protein processing. Interestingly, many protein isoforms changed in relative abundance, a result not predictable from genomic studies. We plan to characterize protein function with expression and localization studies to determine the regulatory networks underlying root responses to microbes.

Communication between Meristems

Genetic and physiological evidence indicates the control of meristem development by genes and signals acting locally or at a distance. In plants, signaling molecules fall into several classes. The five classical plant hormones (auxin, cytokinin, ethylene, GA, and abscissic acid) have been studied intensively. More recently, small peptide molecules that are candidates for meristem regulation have been investigated. The collaborative group has a program of examining the regulatory systems involved in lateral shoot, root growth and nodule development, and the isolation of peptide signal molecules. These programs will define mechanisms of systemic autoregulation control and functional linkages between lateral root formation and nodule initiation and identify novel regulatory molecules.

Resistance of *M. truncatula* to Pathogens and Insect Pests. *R* Gene Isolation, Defense Gene Expression Profiling, Proteomics, and Comparative Mapping within the Papillonidae

PI: Karam B. Singh, Commonwealth Scientific and Industrial Research Organization (CSIRO) Centre for Environment and Life Sciences, Australia,

karam.singh@csiro.au; Co-PI: Owain Edwards, CSIRO Centre for Environment and Life Sciences, Australia, owain.edwards@csiro.au; Co-PI: Richard P. Oliver, Murdoch University, Australia, roliver@central.murdoch.edu.au; Co-PI: John Klingler, Centre for Environment and Life Sciences, Australia, john.klingler@csiro.au; and Co-PI: Simon Ellwood, Murdoch University, Australia, sellwood@central.murdoch.edu.au

Disease and pests are major problems for legume crops worldwide. The development of *M. truncatula* as a model system for plant research opens up new opportunities to study biotic stress responses in these important crops. Our research program is using genomic approaches in *M. truncatula* to study some of the key biotic stresses impacting on legume crops in Australia. Our program is linked with legume breeders through the Centre for Legumes In Mediterranean Agriculture and the South Australian Research and Development Institute to facilitate the transfer of results obtained with *M. truncatula* into other agronomically important species. The accumulating evidence of extensive synteny among legumes will aid rapid transfer between species. Of particular relevance to our research program is the fact that *M. truncatula* is grown in Australia as an important pasture crop and the Australian *Medicago* Genetic Resource Centre in Adelaide, South Australia has one of the world's largest collections of annual and perennial *Medicago* germplasm (http://www.sardi.sa.gov.au/pages/pastures/pas_genetres.htm).

Molecular Genetic and Genomic Approaches to Study Aphid-Plant Interactions

Aphids and other sap-feeding insects are important pests in many legume crops due to direct feeding damage and their role as vectors of plant viruses. Little is known about molecular mechanisms of defense against these insects. Arabidopsis has had limited use as a model for studying plant resistance because simply inherited variation for aphid performance has not been reported in this species. Dominant, monogenic resistance to a major aphid pest of legume crops has been identified in *M. truncatula* germplasm at the Australian *Medicago* Genetic Resource Centre in Adelaide, South Australia. Plant breeders at South Australian Research and Development Institute have backcrossed resistance to the blue-green aphid (BGA; *Acyrthosiphon kondoi*) to three widely grown *M. truncatula* cultivars, creating three pairs of NILs in which members of each pair differ with respect to BGA performance. One of the susceptible recurrent parents in these pairs is Jemalong, the progenitor of line A17. This is a tremendous advantage for the genetic analysis of this interaction, because the genome of A17 is currently being sequenced.

Using a wide range of approaches, this CSIRO-funded project is exploiting these three pairs of lines to assist in the elucidation of the resistance mechanism to BGA. We have employed an electrophysiological technique to compare the stylet probing behavior of aphids between resistant and susceptible NILs. Our results suggest that the resistance trait exerts its effect at the level of the phloem tissue. Prior aphid infestation significantly reduces phloem sap ingestion, indicating that the resistance trait is induced by BGA. We are currently phenotyping F_2 populations segregating for BGA resistance, with the goal of map-based cloning of the BGA resistance gene. To perform fine mapping, we have developed an efficient and nondestructive bioassay to determine the resistance phenotype of F_2 plants. In a parallel approach, we are using differential screens and genomic/proteomic approaches to conduct a global analysis of aphid-responsive gene expression, comparing the responses of resistant and susceptible NILs. The results are expected to shed light on molecular signals critical to the elicitation of an effective defense response to BGA, and may contribute to enhanced aphid resistance in other crop species.

Genetic Dissection of Necrotrophic Fungal Disease Resistance in Legumes Using M. truncatula

Fungal necrotrophic diseases are a key production constraint for Australian legume production. This Grains Research and Development Corporation-funded project aims to improve our understanding of legume disease resistance mechanisms by initiating multidisciplinary studies that combine genetics, genomics, biochemistry, and cell biology using *M. truncatula*. To date, research into the underlying genetic basis of plant resistance to necrotrophic fungal pathogens has been limited by a lack of genetic tools and plant models, particularly because Arabidopsis is host to only a few necrotrophic species. Using a collection of fungal pathogens isolated from diseased legumes, we have identified several isolates that cause disease on *M. truncatula*. These include *A. fabae*, *Botrytis fabae*, *Colletotrichum gloeosporioides*, *Diaporthe toxica*, *F. oxysporum*, *Phoma pinodella*, *Pleiochaeta setosa*, *Pythium* sp., *Rhizoctonia solani*, *Stemphyllium botryosum*, and *S. vesicarium*. Where clear differences in ecotype phenotypes are evident, we intend to map-base clone resistance genes in segregating F_2 populations, RILs, or mutagenized populations.

A more global view of necrotrophic resistance in *M. truncatula* is being developed by analyzing gene expression differences using quantitative PCR and microarrays. We will compare responses to different pathogens between susceptible and resistant inbreds. Due to the paucity of genome information in several crop legumes, notably lupins and chickpea, we also plan to harness the wealth of sequence information in *M. truncatula* to enable comparative mapping of re-

sistance genes within the Papilionoideae by virtue of common microsynteny. We have also initiated forward genetic screens involving pathogen-inducible promoters linked to the luciferase gene to identify mutants in defense gene expression. Mutants with altered luciferase expression (constitutive and absent) will be sought and characterized for their disease resistance. Of particular interest is the identification of mutants in defense signaling in legume roots and how these mutants affect symbiotic interactions with rhizobium and/or mycorrhiza.

G. MAX

The Public Soybean EST Project

PI: Randy Shoemaker, USDA-ARS, Iowa State University, rcsshoe@iastate.edu; Co-PI: Ernest Retzel, University of Minnesota, ernest@ahc.umn.edu; Co-PI: Lila Vodkin, University of Illinois, l-vodkin@uiuc.edu; Co-PI: Paul Keim, northern Arizona University, paul.keim@nau.edu; and Co-PI: Sandra Clifton, Washington University School of Medicine, sclifton@watson.wustl.edu

This work was jointly supported by the United Soybean Board and the North Central Soybean Research Program (http://soybean.ccgb.umn.edu/).

The size and complexity of the soybean genome currently make whole-genome sequencing a nontrivial problem to be solved. Because it is primarily the coding regions of genes that provide the information many geneticists desire, transcript sequencing provides an efficient and cost-effective method of sampling the coding portion of the genome.

The Public Soybean EST Project began in 1998 and was funded by check-off dollars through the North Central Soybean Research Program and the United Soybean Board, and by the USDA-ARS. The project was developed as a truly "public" project with input from many members of the soybean community, and data placed immediately into the public sector.

More than 80 cDNA libraries were generated through the project. These libraries represent transcripts from a very broad range of genotypes, developmental and reproductive stages, organs, tissues, and abiotic and biotic stresses. Libraries are sampled deeply when the frequency of discovery of "new" ESTs remains high, and less deeply when redundant sampling significantly decreases the efficiency of discovery. Most cDNA libraries were generated by members of the PI's or CoPIs' laboratories. Other libraries were created and donated by scientists with unique specialization with particular soybean diseases or stresses. Still other scientists contributed tissues representing unique mutants or unique soybean genotypes.

The success of the project has been gratifying. More than 285,000 EST sequences have been deposited in dbEST, and more are being processed. The soybean ESTs represent the largest collection of ESTs for any plant. The collection currently coalesces into approximately 50,000 unigenes and singletons. The average length of each sequence is greater than 400 bases. Because of the size of the EST collection, contigs are often comprised of more than six ESTs. On average, each contig spans more than 750 bases.

The project is providing more to the legume community than simply "gene discovery." When the EST sample is large and random, the frequency of recurrence of any EST provides information on the level of expression of the corresponding gene in the tissue or organ sampled. As a consequence, we are able to draw inferences about gene relationships and gene ontology based upon similarities in expression patterns. In addition, we are learning much about the evolution of duplicated genes in this ancient polyploid and can estimate a coalescence time for the duplicates.

This project has already resulted in publications based directly upon analyses of the EST collections (Granger et al., 2002; Shoemaker et al., 2002), and led to a project funded by the NSF Plant Genome Program (A Functional Genomics Program for Soybean; L. Vodkin, PI; see below).

A Functional Genomics Program for Soybean

PI: Lila Vodkin, University of Illinois, l-vodkin@uiuc.edu; Co-PI: Paul Keim, Northern Arizona University, paul.keim@nau.edu; Co-PI: Joseph Polacco, University of Missouri, bcjoecpo@muccmail.missouri.edu; Co-PI: Ernest Retzel, University of Minnesota, ernest@mail.ahc.umn.edu; Co-PI: Randy Shoemaker, USDA/ARS and Iowa State University, rcsshoe@iastate.edu; and Co-PI: Nevin Young, University of Minnesota, neviny@umn.edu

This work was supported by the NSF (Plant Genome Project No. 9872565; http://soybeangenomics.cropsci.uiuc.edu/and http://soybean.ccgb.umn.edu/).

This project is a collaborative effort to stimulate basic and applied research in functional genomics of soybean through development of tools for global gene expression analysis and physical mapping. The main objectives are to build a soybean "unigene" set defined by 5′ and 3′ sequence data, to construct and use microarrays for global expression, to generate and sequence SAGE-tagged libraries, and to produce BAC end sequences that are anchored on the soybean genetic map. The objectives of this NSF-sponsored project and progress toward those goals are described briefly below:

Development of a Soybean "Unigene Set" and Validation by 3′ Sequencing

Our objective is to assemble a "unigene set" of approximately 36,000 sequences that have been verified by 3′ sequencing and used for microarrays. This set is being built by collapsing the 5′ sequence data from the companion "The Public Soybean EST Project" (see above; Shoemaker et al., 2002). The ESTs were clustered within each library and then across libraries to yield unique sequences. The singletons and a representative of each contig were selected for the unigene sets.

The current "unigene" collection (or tentatively unique sequences) represents low-redundancy sets of cDNA clones. These include reracked libraries Gm-r1070 (a set of 9,216 cDNA clones from various stages of immature cotyledons, flowers, pods, and seed coats), Gm-r1021 plus Gm-r1083 (a set of approximately 9,216 cDNA clones from 8-d-old seedling roots, seedling roots inoculated with *Bradyrhizobium japonicum*, whole seedlings, and 2-month-old roots), and Gm-c1088 (a collection of 9,216 cDNA clones from a number of libraries made from cotyledons and hypocotyls of germinating seedlings and leaves and other plant parts subjected to various pathogens or environmental stress conditions).

After cluster analysis, the 3′ ends of the reracked set of cDNA were sequenced. The 5′ and 3′ sequences were then compared and functional assignments made using BLASTX. Contig analyses can be found at http://soybean.ccgb.umn.edu. Bioinformatics tools as MetaFam for a protein database were developed as part of this and other projects (Shoop et al., 2001; Silverstein et al., 2001a, 2001b). The cDNA clones are available from the American Type Culture Collection (Manassas, VA).

Development and Use of Global Expression Methods for Soybean

The "unigene" sets were processed for use in microarray experiments. Over 27,000 PCR reactions were processed. We have conducted initial hybridizations with arrays containing cDNAs mostly from the young roots. Experiments include transcript profiling during the process of inoculation by *B. japonicum* (in collaboration with Gary Stacey), phosphate application, or tissue profiles. We also utilized microarrays containing 9,216 clones of the Gm-r1070 set (representing many cDNAs from developing seeds, seed coats, flowers, and pods) in a number of studies funded by additional projects. One was a detailed analysis of induction of somatic embryos during culture of cotyledons on auxin-containing media. These transcript profiles were subjected to a cluster analysis and revealed the process of reprogramming of the cotyledons cells during the induction process. Examples of array data are on our Web site (http://soybeangenomics.cropsci.uiuc.edu). The soybean microarray methods were communicated to the community through a workshop organized at the University of Illinois (May 16–18, 2000) and attended by 29 participants from 15 universities. The detailed protocols are also presented on the Web site at the University of Illinois (http://soybeangenomics.cropsci.uiuc.edu). Contact Lila Vodkin for further information on soybean microarrays.

Another goal was to develop and apply SAGE technology for soybean. SAGE protocols have been optimized for soybean by the laboratory of Paul Keim. To date, we have constructed 20 libraries and generated a total of 132,992 SAGE tags, 40,121 of which are unique. Cluster analysis of transcriptomes from various tissues shows relatedness among tissues, which fits expectations. The SAGE tags are blasted against the soybean EST collection and SAGE data can be viewed at http://soybean.ccgb.umn.edu. Contact Paul Keim or James Schupp (James.Schupp@nau.edu) for further information on soybean SAGE data.

Building an Infrastructure of BAC Contigs Anchored to the Genetic Map

This objective was addressed by the Shoemaker and Young laboratories. More than 750 SSR and RFLP markers were used to identify BAC clones in soybean. The markers anchor these BACs to the consensus molecular genetic map for soybean (Marek et al., 2001; Foster-Hartnett, et al., 2002). Almost 60% of the end sequences from BACs identified by both types of markers produce significant similarity in BLAST database searches. Of the significant hits, the largest single category is repetitive sequences. Repetitive sequences have similarity to previously characterized long-terminal repeat and non-long-terminal repeat retrotransposons. These BACs delimit contigs that represent roughly 200 Mb or close to 20% of the soybean genome.

Integrative Physical Mapping of the Soybean Genome

PI: David A. Lightfoot, Southern Illinois University, ga4082@siu.edu; Co-PI: Khalid Melssem, Southern Illinois University, kmeksem@siu.edu; and Co-PI: Hongbin Zhang, Texas A&M University, hbz7049@tamu.edu

This work was supported by the NSF (Plant Genome Award no. 9872635; http://www.siu.edu, http://hbz.tamu.edu/info).

The development of a physical map of the soybean genome has the potential to accelerate the rate of discovery and cloning of economically important genes by severalfold. A detailed physical map would also provide for the future sequencing of the whole soybean genome. The objectives of this research are to develop a physical map of the soybean genome

and to integrate it with the genetic map. The ultimate goal is to facilitate discovery, cloning, manipulation, and utilization of soybean genes for genetic improvement and agricultural production. The research also provides soybean researchers with electronic access to BAC clones encompassing regions likely to contain genes of economic importance. This report overviews the public resources created by the project and gives a progress report on results to date.

Community Access to Project Resources

A marker-anchored (Zobrist et al., 2000; Marek et al., 2001; Shultz et al., 2001) physical map of the soybean genome derived from large insert clones (Marek and Shoemaker, 1997; Danesh et al., 1998; Meksem et al., 2000) has been constructed and refined (C. Wu, personal communication). There are currently four routes available to access the data sets. One route of access is provided by genetic markers. There is a search tool for individual markers at www.siu.edu/~pbgc/htmls/search.html. The genetic map location of a marker on the USDA linkage group can be viewed at www.siu.edu/~pbgc/htmls/link.htm. Selecting http://www.siu.edu/~pbgc/htmls/mapa1.htm allows the genetic map of a particular linkage group to be displayed. Selecting a marker within that group allows the contig(s) identified by that marker to be displayed with clone plate addresses. The sequences associated with a contig represent a second route of entry to the physical map. Selecting the NCBI button next to a marker allows the DNA sequences associated with that marker to be displayed (Iqbal et al., 2001; Marek et al., 2001). These include primer sequences, allele sequences, BAC end sequences, and sequence-tagged sites linked to clone plate addresses within that contig. The sequences associated with a contig represent a second route of entry to the physical map.

A third route of access is provided by the plate addresses for clones that can be used to view individual BAC fingerprints (http://hbz.tamu.edu/). At this site, fingerprints were used to create a fingerprint contig (FPC) database (Soderlund et al., 1997, 2000) for investigator-driven contig assembly (http://hbz.tamu.edu/bacindex1.html). Contigs representing the soybean physical map have been assembled and posted at this site. Three search tools are available: BAC clone plate address to FPC fingerprint, BAC clone plate address to contig, and contig number to contig map. All contigs show plate addresses that identify libraries correctly. The fourth route of entry is provided by a downloadable FPC database containing 78,001 high-quality fingerprints (http://www.siu.edu/~pbgc/contig/soy_fpc_data/). Members of the community can generate their own builds and merges using fingerprint data. Distributed access will allow for community builds to be incorporated into the central database by author resubmission. The whole database structure will be mirrored at two sites to avoid loss of access. The databases are being incorporated into a GMOD (genome browser for model organisms) browser format with an FPC plug that allows distributed annotation system applications. These resources are made available so that community members can use the database to build their own overlapping clone tiles, to provide a framework for EST mapping, to allow comparative genome analysis between cultivars, to support large-scale genome sequencing, and to direct targeted DNA marker development.

Summary of the Results of the Completed Work

Precisely 469 microsatellite markers and 105 RFLP markers have been anchored to contigs in cooperation with another NSF project (no. 9872565, “A Functional Genomics Program for Soybean”; see Vodkin et al., above). The genetic map location for all markers and plate addresses for all clones can be viewed at http://www.siu.edu/~pbgc/. The BAC end sequences are deposited at NCBI along with BAC subclone sample sequences. There are useful fingerprints for about nine soybean genomes. Wu et al. (C. Wu, S. Sun, N. Padmavathi, F.A. Santos, R. Springman, K. Meksem, D. Lightfoot, and H.-B. Zhang, unpublished data) have assembled and edited the map contigs from the fingerprint database, resulting in two data releases. An automated build, released in September 2001, contained more than 78,000 BACs, forming nearly 5,500 contigs, for a total length of 1,664 Mb. The subsequent build that was manually edited was released in October of 2002. The approximately 58,000 BACs in the 2,907 contigs of the second build totaled 1,451 Mb in length. In the future, distributed community-based contig editing will refine the map further, using a downloadable FPC database available at http://hbz.tamu.edu.

To support functional genomics, we have developed a best tiling path of 9,600 clones (25 plates) that encompasses 99% of the cloneable genome. This resource is available from Southern Illinois University (jlshultz@siu.edu). There is a renewable collection of 100 RILs derived from the cross of Essex and Forrest (Lightfoot et al., 2002; Njiti et al., 2002) that is available from Southern Illinois University on request. An immortal collection of 100 NIL pairs is available, where each pair derived from a single RIL (Njiti et al., 1998). An EMS-mutagenized library suitable for tilling has been developed, and a viral-induced gene silencing vector is under development.

Conclusions

The Forrest physical map and best tiling path provide a useful first step toward genome access for soybean. The map builds will be used to compare genome structures among legumes (Men et al., 2001; Rajesh et al., 2002), to anchor the proposed high information content physical map to the genetic map,

to anchor selected sequences to contigs (Wu et al., 2003), to generate new genetic markers (Meksem et al., 2001), to identify genes underlying QTL (Lightfoot, 2001; Cho et al., 2002; Lightfoot and Meksem, 2002; Yuan et al., 2002), and to target gene-rich regions of linkage group G for pooled genomic clone genome sequencing (D. Lightfoot, unpublished data). The end product of the related projects will be a high-density gene map of soybean from which the position, order, and function of genes can be determined.

Dissecting *Phytophthora* Resistance in Soybean Using Expression Profiling and Analysis of QTLs

PI: Brett M. Tyler, Virginia Polytechnic and State University, bmtyler@vt.edu; Co-PI: Glenn Buss, Virginia Polytechnic and State University, gbuss@vt.edu; Co-PI: Anne E. Dorrance, Ohio State University, dorrance.1@osu.edu; Co-PI: Ina Hoeschele, Virginia Polytechnic and State University, inah@vt.edu; Co-PI: M. A. Saghai Maroof, Virginia Polytechnic and State University, smaroof@vt.edu; Co-PI: Steven St. Martin, Ohio State University, stmartin+@osu.edu; and Co-PI: Keying Ye, Virginia Polytechnic and State University, keying@vt.edu

This work was supported by the NSF (Plant Genome Project No. 0211863; initiated October 1, 2002; http://www.vbi.vt.edu/staff/perspage_fcty/fcty_tyler_bret.htm).

Much has been learned over the last decade regarding the mechanism by which plants are protected from disease by major resistance genes. Much less is known about mechanisms of quantitative resistance, which protect a plant species against all genetic forms of a pathogen, albeit not completely. However, this form of durable resistance often retains its effectiveness over time against pathogen populations that can change rapidly to overcome major resistance genes in the host.

This project will use functional genomics to dissect mechanisms of quantitative resistance in soybean against one of its most destructive pathogens, the oomycete *Phytophthora sojae*. cDNA and oligonucleotide microarrays will be used to examine gene expression profiles of both the host and the pathogen during *P. sojae* infection of soybean lines containing varied levels of quantitative resistance and in progeny segregating for quantitative resistance. QTL analysis of the expression profiles in a segregating population will enable us to identify soybean loci that contribute to quantitative resistance.

Thirteen major resistance alleles (*Rps* genes) have been identified in soybeans at seven loci that provide protection against *P. sojae* (Diers et al., 1992; Lohnes and Schmitthenner, 1997), and matching avirulence genes have been identified in *P. sojae* for 11 of these (May et al., 2002). Many populations of *P. sojae* have adapted to overcome the *Rps* genes that have been deployed in commercially grown soybean cultivars. Quantitative resistance to *P. sojae* is also present in soybeans and appears to be controlled by multiple genes (Walker and Schmitthenner, 1984; Glover and Scott, 1998). It is an important means for controlling the disease, especially in locations where *Rps* genes are no longer effective. Preliminary data from an analysis of three soybean populations indicate that three or more loci may be involved with this trait (A.E. Dorrance and S. St. Martin, unpublished data). Heritability estimates range from 0.57 to 0.87 for root lesion length in these F_4:F_6 populations.

The soybean-*P. sojae* system is one of very few pathosystems in which extensive genomic resources are available to examine both plant and pathogen simultaneously (Tyler, 2001; see also summaries by Lightfoot et al., Shoemaker et al., and Vodkin et al.). The overall design of this project is to screen a large population of soybean genetic progeny segregating for quantitative resistance against *P. sojae*, using soybean and *P. sojae* microarrays. QTL analysis of the plant and pathogen profiles will identify soybean loci that contribute to quantitative resistance, and the mechanisms of resistance conferred by genes at those loci. By assaying pathogen expression profiles in addition to plant expression profiles, we should more readily identify minor resistance genes, especially if there are several such genes segregating. For example, a small decrease in lesion size caused by one minor gene may be very difficult to assay and map using conventional QTL analysis, but if that gene triggers large-scale changes in the levels of certain pathogen mRNAs, then the resistance gene would be easily detected. Further, just as expression profiles can reveal how antimicrobial drugs debilitate human pathogens, they should reveal mechanisms by which plant defenses debilitate plant pathogens. Therefore, analyzing plant and pathogen expression profiles simultaneously will provide unique insights into the interplay between plant and pathogen genes that regulate the outcome of an infection.

PHASEOLUS VULGARIS

The *Phaseomics* International Consortium

Coordinators: William J. Broughton, University of Geneva, Switzerland, william.broughton@bioveg.unige.ch; Gina Hernández, C. de Investigación sobre Fijación de Nitrógeno-Universidad Nacional Autónoma de México (UNAM), Mexico, gina@cifn.unam.mx; Matthew Blair, International Center for Tropical Agriculture (CIAT), Colombia, m.blair@cgiar.org; Paul Gepts, University of California, Davis, plgepts@ucdavis.edu; Jos Vanderleyden, Katholieke University

Leuven, Belgium, jozef.vanderleyden@agr.kuleuven.ac.be; and Nancy Terryn, Ghent University, Belgium, nater@gengenp.rug.ac.be

One-half the dry edible legumes consumed worldwide as human food are common beans. Beans are one of the world's most ancient crops, originating in Central America and the Andean region of South America. Beans and their relatives (*Phaseolus* spp.) are extremely diverse crops in terms of cultivation methods, uses, the range of environments to which they have been adapted, and morphological variability. Their genetic resources exist as a complex array of major and minor gene pools, races, and intermediate types, with occasional introgression between wild ancestors and domesticated types. Thus, beans are a crop that is adapted to many niches, both in agronomic and consumer preference terms. Although many diverse associations contribute to symbiotic N fixation, in most agricultural settings the primary source (80%) of biological fixed N is through the *Rhizobium*-legume symbiosis. Important agricultural goals include enhancing the use of and improving the management of biologically fixed N, as well as increasing the adaptation to abiotic stresses. The original microsymbiont of common bean is the gram-negative bacterium *Rhizobium etli* that also originated in the Americas and has co-evolved with beans for millennia. However, other rhizobia, such as *Rhizobium tropici*, which is resistant to acid soils, *M. loti*, or *Rhizobium leguminosarum* bv *phaseoli*, can also nodulate the species.

Genomics, transcriptomics, and proteomics permit the study of many (and sometimes all) genes of a particular organism. Significant discoveries concerning interrelationships between some of the basic metabolic functions of the organisms have been made this way for model plant species. As a consequence, an integrated, almost holistic view of the organism is evolving. An international consortium called *Phaseomics* (*Phaseolus* gen*omics*) has been formed to establish the necessary framework of knowledge and materials for the advancement of genomic studies of bean. *Phaseomics* includes around 80 scientists from 20 different countries. The principal goal of the consortium is to increase the genetic resources and tools available for the crop, especially large insert and cDNA libraries, genomic sequences, ESTs, and genetic markers. The ultimate objective of the research carried out within *Phaseomics* is to assist in the generation of new common bean varieties that are not only suitable for but also desired by local farmer and consumer communities. An additional long-term goal is to more rationally use the large germplasm collections held for the crop at the CIAT (36,000 accessions), and at national germplasm repositories in the U.S. (USDA), Brazil (Brazilian Agricultural Research Corporation), and Mexico (National Institute for Forest and Agricultural Research). A detailed description of the global project proposed by the *Phaseomics* partners will be published this year (Broughton et al., 2003). *Phaseomics* partners have met in two occasions: in 2001 in Cuernavaca, Mexico, and in 2002 in Geneva. During the second *Phaseomics* meeting, different working groups that concentrate on certain of the *Phaseomics* goals were formed, with the aim of providing efficient exchange of information among the scientists performing related individual projects and to avoid duplication.

One of the working groups is the EST sequencing group that is coordinated by G. Hernández. EST projects provide an inexpensive and efficient method to get information on the genes that are expressed in a certain tissue or organ of the plant and when converted to genetic markers, breeders can use them to position genes on the genetic map of the species. Judicious selection of the type of tissue from which to isolate the mRNA (and hence prepare a cDNA library) provides valuable information not only on the type of genes found in a particular plant, but also on the conditions in which they are expressed. Thus, EST projects permit "skimming" gene expression information and, by extension, the genome itself. The following *Phaseomic* partners are performing projects on common bean ESTs that are funded by different sources depending on the groups and/or country. M. Blair, S. Beebe, and J. Tohme from (CIAT, Colombia) have sequenced approximately 4,000 ESTs from leaves of an Andean genotype G19833 and from young lateral and basal roots treated with limited or unlimited phosphorus of the Mesoamerican genotype DOR364. M. Melotto et al. (University of São Paulo, Brazil) are preparing two cDNA libraries from seedling shoots of the anthracnose-resistant genotype SEL1308 with and without inoculation with *Colletotrichum lindemuthianum*, and they will sequence a total of 5,000 ESTs from each library. J. Vanderleyden et al. are sequencing ESTs from a cDNA library from nodules (10 d postinoculation) of the BAT477 genotype inoculated with *R. etli* CNAPF512. G. Hernández, M. Lara, and M. Ramírez (Nitrogen Fixation Research Center-UNAM, Mexico), in collaboration with C.P. Vance (University of Minnesota, USDA) have sequenced 3,000 ESTs from a cDNA library of nodules (18 d postinoculation) from the variety Negro Jamapa 81 inoculated with *R. tropici* 899. Libraries were also prepared from mRNA isolated from developing pods and phosphorus-stressed roots for which 3,000 cDNA clones will be sequenced each. W.J. Broughton et al. (University of Geneva) are currently isolating mRNA from root hairs (BAT93) treated with *Rhizobium* sp. NGR234 and its Nod factors. Each group has or will have their internal database and will share the EST data with the public through GenBank. The research from the several groups is moving toward functional genomics analyses, through macro- and microarray technologies.

An efficient and reliable genetic transformation system is crucial to any genomic project. The work-

ing group coordinated by N. Terryn focuses on *Phaseolus* transformation mediated by *Agrobacterium*. It is now clear that two *Phaseomics* groups can routinely transform *Phaseolus acutifolius*: the group in Ghent (N. Terryn) and the group at CIAT (A. Mejía-Jiménez and J. Tohme). Most work is done with NI576, a wild *P. acutifolius* line, whereas the Ghent group is now also exploring the use of a cultivated variety. The problem is that the protocol is rather laborious and time consuming. As for *P. acutifolius* crosses to common bean, embryo rescue is needed, and crosses are not always successful. Therefore, CIAT is working on hybrid lines between *P. acutifolius* and common bean that will have greater transformability. This could make it easier to cross transgenics obtained in these lines with common bean varieties of interest. Although it has been shown that *A. tumefaciens* can transfer DNA to common bean, an efficient transformation system for common bean has not yet been established. The groups of M. Lara and G. Hernandez (Mexico) are working in trying to develop a transformation system for common bean based in *A. tumefaciens* and in vitro regeneration protocols that they have developed for several cultivars.

Other *Phaseomics* working groups and their coordinators are: resources and libraries, P. Gepts (U.S.); biological N fixation, G. Hardarson (Austria); germplasm, M. Blair (Colombia); genome sequencing, E. Triplett (U.S.); bioinformatics, J.P. Nap (The Netherlands); and coordination, W.J. Broughton (Switzerland).

Development of Genetic and Genomic Tools to Study Nutritional Quality and Aluminum (Al) Tolerance in Common Bean

PI: M.W. Blair, CIAT, Colombia, m.blair@cgiar.org; Co-PI: S. Beebe, CIAT, Colombia, s.beebe@cgiar.org; Co-PI: P. Beyer, University of Freiburg, Germany, beyer@uni-freiburg.de; Co-PI: D. Dellapena, Michigan State University, dellapen@msu.edu; Co-PI: M. Grusak, USDA-ARS Children's Nutrition Research Center,BaylorCollegeofMedicine,mgrusak@bcm.tmc.edu; Co-PI: W. Horst University of Hannover, Germany, horst@mbox.pflern.uni-hannover.de; Co-PI: J. Tohme, CIAT, Colombia, j.tohme@cgiar.org; and Co-PI: I. Rao, CIAT, Colombia, i.rao@cgiar.org

This work was supported by the U.S. Agency for International Development (grant no. DAN–G–IN–89–00048–00) and by the German Federal Ministry for Economic Cooperation and Development (Project no. 2000.7860.0–001.11,contractno.81043147;http://www.ciat.cgiar.org/biotechnology/index.htm).

Micronutrient deficiencies in human populations, especially for iron and zinc, are a serious health concern in the developing world, where more than 1 billion people live in a state of abject poverty and have limited access to sufficient quantities of these minerals in their daily diet. Common beans are by far the world's most important food grain legume and, therefore, have a role in addressing this problem (Graham et al., 1999). Genetic variability exists for mineral content in common bean and, surprisingly, seed iron and zinc concentrations tend to be correlated, leading to the promise of incorporating these traits into a bean improvement program (Beebe et al., 2000). However, the physiology and genetic mechanisms controlling mineral transport pathways are still being worked out and remain a priority if nutritional breeding is to be effective.

Breeding for tolerance to Al phytotoxicity is another important objective of bean breeding programs in the tropics, because this is a serious edaphic constraint affecting over 40% of the world's arable land, especially acid soils of tropical and subtropical regions (Von Uexkull and Mutert, 1995). New bean varieties with Al tolerance would reduce the dependence of farmers on lime and fertilizer (inorganic/organic) inputs that are used to solve these soil problems. Genetic variation exists for Al resistance among common bean genotypes (Thung and Rao, 1999), and screening procedures have been or are being established. However, plant breeders are only in the early stages of developing Al-resistant cultivars for farmers in areas that need them, and much work is needed to understand the genetic control of Al toxicity tolerance and how best to select for this trait.

At the CIAT, two projects are investigating the genes that control micronutrient content and Al stress tolerance in common bean. The issue of seed micronutrient accumulation and abiotic stress tolerance are related given the link between soil and plant nutrition and their affect on grain quality and human nutrition. Therefore, in the research for these two projects, we are asking similar questions, such as how plant roots adapt to acid soils or how root uptake of minerals translates into increased accumulation in the grain, and for both projects, we are developing similar genetic tools that are described below.

Marker Development

A large priority has been placed on the development of PCR-based markers. One of the main marker types emphasized has been microsatellites or SSRs. A set of microsatellites has been put together to efficiently tag the QTLs controlling the characteristics of interest for the two projects.

Genetic Mapping

Identification of contrasting parents and the development of RILs are permitting us to identify molecular markers linked to QTLs associated with micronutrient accumulation or Al resistance and/or

individual physiological mechanisms of each trait. Four RIL populations have been developed for the micronutrient project and five for the Al tolerance project. The genetic maps for these populations are integrated with CIAT's principal mapping population (DOR364 × G19833), which now contains over 500 markers including RFLPs, random-amplified polymorphic DNAs, microsatellites, and AFLPs.

The first is a leaf cDNA library constructed from total mRNA extracted from leaves of adult plants of the Andean variety G19833. A total of 64,000 clones have been plated from this library and picked into 384-well plates that have been arrayed onto high-density filters for clone hybridization and gene discovery. Root cDNA libraries have also been made for adventitious and basal roots grown with and without P deficiency stress for the genotype DOR364. An additional 32,000 clones have been picked from each of these libraries. A total of 4,200 clones from all three of the libraries have been sequenced and compared for homologs in the soybean database.

Future Plans

CIAT is part of a consortium of centers within the Consultative Group on International Agricultural Research to develop bioinformatics tools linking mapping, QTL analysis, and germplasm evaluation. Emphasis has been placed on creating databases for managing genotype and genetic mapping information and establishing sequence storage and processing capacities. In addition, CIAT has established a microarray facility that will be used to study gene expression in common beans using clones from the cDNA libraries described above. Gene expression studies will aim to identify the genes that control and characterize mechanisms of mineral accumulation and Al resistance.

Results from these projects will be applied to the genetic improvement of common bean for small farmers and consumers in the developing world. CIAT has a strong record in developing common bean varieties for tropical production zones in Africa and Latin America. CIAT also holds a global mandate for conserving common beans and their respective wild relatives. This bean collection represents critical agrobiodiversity for hundreds of millions of rural and urban people throughout the tropics because the GenBank is a source of novel traits for breeding improved genotypes. It is hoped that the collections' active use in genomic analysis and breeding will help to address the constraints represented by the projects described above and will lead to the development of improved varieties that contribute to food security, poverty alleviation, and economic development.

Received January 12, 2003; returned for revision January 14, 2003; accepted January 14, 2003.

LITERATURE CITED

Andersen MD, Busk PK, Svendsen I, Møller BL (2000) Cytochromes P-450 from cassava (*Manihot esculenta* Crantz) catalyzing the first steps in the biosynthesis of the cyanogenic glucosides linamarin and lotaustralin. cloning, functional expression in *Pichia pastoris*, and substrate specificity of the isolated recombinant enzymes. J Biol Chem **275:** 1966–1975

Ané JM, Lévy J, Thoquet P, Kulikova O, de Billy F, Kim DJ, Penmetsa V, Debellé F, Rosenberg C, Cook D et al. (2002) Genetic and cytogenetic mapping of *DMI1, DMI2 and DMI3* genes of *Medicago truncatula* involved in Nod factor transduction, nodulation and mycorrhization. Mol Plant-Microbe Interact **15:** 1108–1118

Asamizu E, Nakamura Y, Sato S, Tabata S (2000) Generation of 7137 non-redundant expressed sequence tags from a legume, *Lotus japonicus*. DNA Res **7:** 127–130

Asirvatham VS, Watson BS, Sumner LW (2002) Analytical and biological variances associated with proteomic studies of *Medicago truncatula* by 2-DE. *Proteomics* **2:** 960–968

Aoki T, Kamizawa A, Ayabe S (2002) Efficient *Agrobacterium*-mediated transformation of *Lotus japonicus* with reliable antibiotic selection. Plant Cell Rep **21:** 238–243

Bak S, Olsen CE, Halkier BA, and Møller BL (2000) Transgenic tobacco and Arabidopsis plants expressing the two multifunctional sorghum cytochromes P450, CYP79A1 and CYP71E1, are cyanogenic and accumulate metabolites derived from intermediates in dhurrin biosynthesis. Plant Physiol **123:** 1437–1448

Bak S, Olsen CE, Petersen BL, Møller BL, Halkier BA (1999) Metabolic engineering of *p*-hydroxybenzylglucosinolate in *Arabidopsis* by expression of the cyanogenic CYP79A1 from *Sorghum bicolor*. Plant J **20:** 663–671

Barker DG, Bianchi S, London F, Dattee Y, Duc G, Essad S, Flament P, Gallusci P Genier G, Guy P et al. (1990) *Medicago truncatula*, a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. Plant Mol Biol **8:** 40–49

Beebe S, Gonzalez AV, Rengifo J (2000) Research on trace minerals in the common bean. Food Nutr Bull **21:** 387–391

Bell CJ, Dixon RA, Farmer AD, Flores R, Inman J, Gonzales RA, Harrison MJ, Paiva NL, Scott AD, Weller JW et al. (2001) The *Medicago* genome initiative: a model legume database. Nucleic Acids Res **29:** 114–117

Boisson-Dernier A, Chabaud M, Garcia F, Becard G, Rosenberg C, Barker DG (2001) *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Mol Plant-Microbe Interact **14:** 695–700

Bonfante P, Genre A, Faccio A, Martini I, Schauser L, Stougaard J, Webb J, Parniske M (2000) The *Lotus japonicus LjSym4* gene is required for the successful symbiotic infection of root epidermal cells. Mol Plant-Microbe Interact **13:** 1109–1120

Bonnin I, Ronfort J, Wozniak F, Olivieri I (2001) Spatial effects and rare outcrossing events in *Medicago truncatula* (Fabaceae). Mol Ecol **10:** 1371–1383

Broughton WJ, Hernandez G, Blair M, Beebe S, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.): model food legumes. Plant Soil **249:** 291–363

Cannon SB, Zhu HY, Baumgarten A, Spangler R, May G, Cook DR, Young ND (2002) Diversity, distribution and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. J Mol Evol **54:** 548–562

Cho Y, Njiti VN, Chen X, Triwatayakorn K, Kassem MA, Meksem K, Lightfoot DA, Wood AJ (2002) Quantitative trait loci (QTLs) associated with foliar trigonelline accumulation in *Glycine max* L. J Biomed Biotechnol **2:** 151–157

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High-throughput screening for induced point mutations. Plant Physiol **126:** 480–484

Colebatch G, Kloska S, Trevaskis B, Freund S, Altmann T, Udvardi M (2002) Novel aspects of symbiotic nitrogen fixation uncovered by transcript profiling with cDNA arrays. Mol Plant-Microbe Interact **15:** 411–420

Cook D (1999) *Medicago truncatula*: a model in the making! Curr Opin Plant Biol **2:** 301–304

Danesh D, Penuela S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND (1998) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. Theor Appl Genet **96:** 196–202

d'Erfurth I, Cosson V, Eschstruth A, Lucas H, Kondorosi A, Ratet P (2003) Efficient transposition of the *Tnt1* tobacco retrotransposon in the model legume *Medicago truncatula*. Plant J **33:** 1–12

Diers BW, Mansur L, Imsande J, Shoemaker RC (1992) Mapping *Phytophthora* resistance loci in soybean with restriction fragment length polymorphism markers. Crop Sci **32:** 377–383

Doyle JJ, Luckow MA (2003) The rest of the iceberg: legume diversity and evolution in a phylogenetic context. Plant Physiol **131:** 900–910

Eason WR, Webb KJ, Michaelson-Yeates TPT, Abberton MT, Griffith GW, Culshaw CM, Hooker JE, Dhanoa MS (2001) Effect of genotype of *Trifolium repens* on mycorrhizal symbiosis with *Glomus mosseae*. J Agric Sci **137:** 27–36

Ellis THN, Poyser SJ (2002) An integrated and comparative view of pea genetic and cytogenetic maps. New Phytol **153:** 17–25

Endre G, Kereszt A, Kevei Z, Mihacea S, Kaló P, Kiss GB (2002) A receptor kinase gene regulating symbiotic nodule development. Nature **417:** 962–966

Fedorova M, van de Mortel J, Matsumoto PA, Cho J, Town CD, VandenBosch KA, Gantt JS, Vance CP (2002) Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula*. Plant Physiol **30:** 519–537

Foster-Hartnett D, Mudge J, Danesh D, Yan H, Larse D, Denny R, Young ND (2002) Comparative genomic analysis of sequences sampled from a small region on soybean molecular linkage group "G." Genome **45:** 634–645

Glover KD, Scott RA (1998) Heritability and phenotypic variation of tolerance to phytophthora root rot of soybean. Crop Sci **38:** 1495–1500

Graham R, Senadhira D, Beebe S, Iglesias C, Monasterio I (1999) Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. Field Crops Res **60:** 57–80

Granger C, Coryell V, Khanna A, Keim P, Vodkin L, Shoemaker RC (2002) Identification, structure and differential expression of members of a BURP domain containing protein family in soybean. Genome **45:** 693–701

Gualtieri G, Kulikova O, Limpens E, Kim DJ, Cook DR, Bisseling T, Geurts R (2002) Microsynteny between pea and *Medicago truncatula* in the SYM2 region. Plant Mol Biol **50:** 225–235

Hayashi M, Miyahara A, Sato S, Kato T, Yoshikawa M, Taketa M, Hayashi M, Pedorosa A, Onda R, Imaizumi-Anraku H et al. (2001) Construction of a genetic linkage map of the model legume *Lotus japonicus* using an intraspecific F2 population. DNA Res **8:** 301–310

Iqbal J, Meksem K, Njiti V, Kassem M, Lightfoot DA (2001) Microsatellite markers identify three additional quantitative trait loci for resistance to soybean sudden death syndrome (SDS) in Essex × Forrest RILs. Theor Appl Genet **102:** 187–192

Jiang Q, Gresshoff PM (2002) Shoot-control and genetic mapping of the har1–1 (hypernodulation and aberrant root formation) mutant of *Lotus japonicus*. Funct Plant Biol **29:** 1371–1376:

Journet EP, van Tuinen D, Gouzy J, Crespeau H, Carreau V, Farmer MJ, Niebel A, Schiex T, Jaillon O, Chatagnier O et al. (2002) Exploring root symbiotic programs in the model legume *Medicago truncatula* using EST analysis. Nucleic Acids Res **30:** 5579–5592

Kawaguchi M (2000) *Lotus japonicus* "Miyakojima" MG-20: an early-flowering accession suitable for indoor handling. J Plant Res **113:** 507–509

Kawaguchi M, Imaizumi-Anraku H, Koiwa H, Niwa S, Ikuta A, Syono K, Akao S (2002) Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant-Microbe Interact **15:** 17–26

Kawaguchi M, Motomura T, Imaizumi-Anraku H, Akao S, Kawasaki S (2001) Providing the basis for genomics in *Lotus japonicus*: The accessions Miyakojima and Gifu are appropriate crossing partners for genetic analyses. Mol Gen Genome **266:** 157–166

Krusell L, Madsen LH, Sato S, Aubert G, Genua A, Szczyglowski K, Duc G, Kaneko T, Tabata S, de Bruijn F et al. (2002) Shoot control of root development and nodulation is mediated by a receptor-like kinase. Nature **420:** 422–426

Kulikova O, Gualtieri G, Geurts R, Kim DJ, Cook D, Huguet T, de Jong JH, Fransz PF, Bisseling T (2001) Integration of the FISH pachytene and genetic maps of *Medicago truncatula*. Plant J **27:** 49–58

Lamblin AFJ, Crow JA, Johnson JE, Silverstein KAT, Kunau TM, Kilian A, Benz D, Stromvik M, Endre G, VandenBosch KA et al. (2003) MtDB: a database for personalized data mining of the model legume *Medicago truncatula* transcriptome. Nucleic Acids Res **31:** 196–201

Lightfoot DA (2001) Genomic techniques for isolating genes underlying QTL: disease resistance. Agric Biotechnol Net **3:** 1–4

Lightfoot DA, Meksem K, investors. January, 2001, 2002. Isolated polynucleotides and polypeptides relating to loci underlying resistance to soybean cyst nematode and sudden death syndrome and methods employing same. Jenkins and Wilson PA, Durham NC, Patent Application No. 60/178,811

Lightfoot DA, Njiti VN, Schmidt ME, Myers O, Meksem K, Gibson PT (2003) Registration of Essex × Forrest recombinant inbred line (RIL) population. Crop Sci (in press)

Lohnes DG, Schmitthenner AF (1997) Position of the phytophthora resistance gene Rps7 on the soybean molecular map. Crop Sci **37:** 555–556

Maguire TL, Grimmond S, Forrest A, Iturbe-Ormaetxe I, Meksem K, Gresshoff PM (2002) cDNA microarray analysis of soybean (*Glycine max*): tissue-specific gene expression. J Plant Physiol **159:** 1361–1374:

Marek L, Shoemaker R (1997) BAC contig development by fingerprint analysis in soybean. Genome **40:** 420–427

Marek LF, Mudge J, Darnielle L, Grant D, Hanson N, Paz M, Huihuang Y, Denny R, Larso K, Foster-Hartnett D et al. (2001) Soybean genomic survey: BAC-end sequences near RFLP and SSR markers. Genome **44:** 472–581

Mathesius U, Keijzers G, Natera SHA, Weinman JJ, Djordjevic MA, Rolfe BG (2001) Establishment of a root proteome reference map for the model legume *Medicago truncatula* using the EST database for peptide mass fingerprinting. Proteomics **1:** 1424–1440

Mathesius U, Imin N, Chen H, Djordjevic MA, Weinman JJ, Natera SHA, Morris AC, Kerim T, Paul S, Menzel C et al. (2002) Evaluation of proteome reference maps for cross-species identification of proteins by peptide mass fingerprinting. Proteomics **2:** 1288–1303

Mathesius U, Weinman JJ, Rolfe BG, Djordjevic MA (2000) Rhizobia can induce nodules in white clover by "hijacking" mature cortical cells activated during lateral root development. Mol Plant-Microbe Interact **13:** 170–182

May KJ, Whisson SC, Zwart RS, Searle IR, Irwin JAG, Maclean DJ, Carroll BJ, Drenth A (2002) Inheritance and mapping of eleven avirulence genes in *Phytophthora sojae*. Fungal Genet Biol **37:** 1–12

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiol **123:** 439–442

Meksem K, Ruben E, Hyten D, Triwitayakorn K, Lightfoot DA (2001) Conversion of AFLP bands to high-throughput DNA markers. Mol Gen Genet **265:** 207–214

Meksem K, Ruben E, Zobrist K, Zhang H-B, Lightfoot DA (2000) Two large insert libraries for soybean: applications in cyst nematode resistance and genome wide physical mapping. Theor Appl Genet **101:** 747–755

Men A, Meksem K, Kassem MA, Lothar D, Stiller J, Lightfoot DA, Gresshoff PM (2001) A bacterial artificial chromosome library of *Lotus japonicus* constructed in an *Agrobacterium tumefaciens*-transformable vector. Mol Plant-Microbe Interact **14:** 422–425

Møller BL, Seigler DS (1999) Biosynthesis of Cyanogenic Glycosides, Cyanolipids, and Related Compounds. *In* BK Singh, ed, Plant Amino Acids. Marcel Dekker, New York, pp 563–609

Nakamura Y, Kaneko T, Asamizu E, Kato T, Sato S, Tabata S (2002) Structural analysis of a *Lotus japonicus* genome: II. Sequence features and mapping of sixty-five TAC clones which cover the 6.5-Mb regions of the genome. DNA Res **9:** 63–70

Nam Y-W, Penmetsa RV, Endre E, Kim D, Cook DR (1999) Construction of a bacterial artificial chromosome library of *Medicago trucatula* and identification of clones containing ethylene response genes. Theor Appl Genet **98:** 638–646

Njiti V, Doubler TW, Suttner RJ, Gray L, Gibson P, Lightfoot DA (1998) Loci underlying resistance to *Fusarium solani* and soybean sudden death syndrome (SDS) mapped in near isogenic lines. Crop Sci **38:** 472–477

Njiti VN, Meksem K, Iqbal MJ, Johnson JE, Zobrist KF, Kilo VY, Lightfoot DA (2002) Common loci underlie field resistance to soybean sudden death syndrome in Forrest, Pyramid, Essex, and Douglas. Theor Appl Genet **104:** 294–300

Nishimura R, Hayashi M, Wu G-J, Kouchi H, Murakami Y, Imaizumi-Anraku H, Kawasaki S, Akao S, Ohmori M, Harada K et al. (2002a) HAR1 mediates systemic regulation of symbiotic organ development. Nature **420:** 426–429

Nishimura R, Ohmori M, Fujita H, Kawaguchi M (2002b) A Lotus basic leucine zipper protein with a RING-finger motif negatively regulates the developmental program of nodulation. Proc Natl Acad Sci USA **99:** 15206–15210

Novero M, Faccio A, Genre A, Stougaard J, Webb KJ, Mulder L, Parniske M, Bonfante P (2002) Dual requirement of the *LjSym*4 gene for mycorrhizal development in epidermal and cortical cells of *Lotus japonicus* roots. New Phytol **154:** 741–749

Oleykowski CA, Bronson Mullins CR, Godwin AK, Yeung AT (1998) Mutation detection using a novel plant endonuclease. Nucleic Acids Res **26:** 4597–4602

Paquette SM, Bak S, Feyereisen R (2000) Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of *Arabidopsis thaliana*. DNA Cell Biol **19:** 307–317

Paquette SM, Møller BL, Bak S (2003) On the origin of family 1 plant glycosyltransferases. Phytochemistry **63:** 399–413

Penmetsa RV, Cook DR (1997) A legume ethylene-insensitive mutant hyperinfected by is rhizobial symbiont. Science **275:** 527–530

Penmetsa RV, Cook DR (2000) Production and characterization of diverse developmental mutants of *Medicago truncatula*. Plant Physiol **123:** 1387–1397

Penmetsa RV, Frugoli JA, Smith L, Long SR, Cook DR (2003) Dual genetic pathways controlling nodule number in *Medicago truncatula*. Plant Physiol **131:** 998–1008

Perry J, Wang TL, Welham T, Gardner S, Pike J, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. Plant Physiol **131:** 866–871

Rajesh PN, Meksem K, Coyne CJ, Lightfoot DA, Muehlbauer FJ (2002) Construction of the first bacterial artificial chromosome library in chickpea. Int Chickpea Pigeonpea Newslett **9:** 29–30

Rose RJ, Nolan KE (1995) Regeneration of *Medicago truncatula* from protoplasts isolated from kanamycin-sensitive and kanamycin-resistant plants. Plant Cell Rep **14:** 349–353

Sandal N, Krusell L, Radutoiu S, Olbryt M, Pedrosa A, Stracke S, Sato S, Kato T, Tabata S, Parniske M et al. (2002) A genetic linkage map of the model legume *Lotus japonicus* and strategies for fast mapping of new loci. Genetics **161:** 1673–1683

Sato S, Kaneko T, Nakamura Y, Asamizu E, Kato T, Tabata S (2001) Structural analysis of a *Lotus japonicus* genome: I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4-Mb regions of the genome. DNA Res **8:** 311–318

Scholte M, D'Erfurth I, Ripa S, Mondy S, Jean V, Durand P, Breda C, Trinh TH, Rodriguez-Llorente I, Kondorosi E et al. (2002) T-DNA tagging in the model legume *Medicago truncatula* allows efficient gene discovery. Mol Breed **10:** 203–215

Searle IR, Men AE, Laniya TS, Buzas DM, Iturbe-Ormaetxe I, Carroll BJ, Gresshoff PM (2003) Long distance signalling for nodulation control in legumes requires a clavata-1 like receptor kinase. Science **299:** 109–112

Shoemaker R, Keim P, Vodkin L, Retzel E, Clifton SW, Waterston R, Smoller D, Coryell V, Khanna A, Erpelding J et al. (2002) A compilation of soybean ESTs: generation and analysis. Genome **45:** 329–338

Shoop E, Silverstein KAT, Johnson JE, Retzel EF (2001) MetaFam: a unified classification of protein families: II. Schema and query capabilities. Bioinformatics **17:** 262–271

Shultz J, Wu C, Santos FA, Nimmakayala P, Springman R, LaMontague C, Zobrist K, Meksem K, Zhang H-B, Lightfoot DA (2001) A physical gene map for the soybean genome. Soybean Genet Newslett **28:** 5–10

Silverstein KAT, Shoop E, Johnson JE, Kilian A, Freeman JL, Kunau TM, Awad IA, Mayer M, Retzel EF (2001a) The MetaFam server: a comprehensive protein family resource. Nucleic Acids Res **29:** 49–51

Silverstein KAT, Shoop E, Johnson JE, Retzel EF (2001b) MetaFam: a unified classification of protein families: I. Overview and statistics. Bioinformatics **17:** 249–261

Soderlund C, Humphray S, Dunham A, French L (2000) Contigs built with fingerprints, markers, and FPC V4.7. Genome Res **10:** 1772–1787

Soderlund C, Longden I, Mott R (1997) FPC: a system for building contigs from restriction fingerprinted clones. CABIOS **13:** 523–535

Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Szczyglowski K et al. (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature **417:** 959–962

Sumner LW, Duran AL, Huhman DL, Smith JT (2002) Metabolomics: a developing and integral component in functional genomic studies of *Medicago truncatula*. *In* JT Romeo, RA Dixon, eds, Recent Advances in Phytochemistry, Vol 36: Phytochemistry in the Genomics and Post-Genomics Era. Pergamon, Oxford, pp 31–61

Tattersall DB, Bak S, Jones PR, Olsen CE, Nielsen JK, Hansen ML, Høj PB, Møller BL (2001) Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science **293:** 1826–1828

Thoquet P, Ghérardi M, Journet EJ, Kereszt A, Ané JM, Prospéri J-M, Huguet T (2002) The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. BioMedCentral Plant Biol **2:** 1–13

Thung M, Rao IM (1999) Integrated management of abiotic stresses. *In* SP Singh, ed, Common Bean Improvement in the Twenty-first Century. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 331–370

Tyler BM (2001) Genetics and genomics of the oomycete-host interface. Trends Genet **17:** 611–614

Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci **5:** 380–386

Von Uexkull HR, Mutert E (1995) Global extent, development and economic impact of acid soils. *In* RA Date, NJ Grundon, GE Rayment, ME Probert, eds, Plant-Soil Interactions at Low pH: Principles and Management. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 5–19

Walker AK, Schmitthenner AF (1984) Heritability of tolerance to *Phytophthora* rot in soybean. Crop Sci **24:** 490–491

Watson BS, Asirvatham VS, Wang L, Sumner LW (2003) Mapping the proteome of *Medicago truncatula*. Plant Physiol **131:** 1104–1123

Webb KJ, Skøt L, Nicholson MN, Jørgensen B, Mizen S (2000) *Mesorhizobium loti* increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in *Lotus japonicus*. Mol Plant-Microbe Interact **13:** 606–616

Weidner S, Pühler A, Küster H (2003) Genomics insights into symbiotic nitrogen fixation. Curr Opin Biotechnol **14:** 1–6

Wienkoop S, Saalbach G (2003) Proteome analysis: novel proteins identified at the peribacteroid membrane from *Lotus japonicus* root nodules. Plant Physiol **131:** 1080–1090

Wu C, Tao FA, Santos Q, Nimmakayala P, Springman R, Zhang HB (2003) A bacterial artificial chromosome library for "Forrest" soybean. Theor Appl Genet (in press)

Wulf A, Manthey K, Doll J, Perlick AM, Linke B, Bekel T, Meyer F, Franken P, Küster H, Krajinski F (2003) Detection of highly specific transcriptional changes in response to arbuscular mycorrhiza development in the model plant *Medicago truncatula*. Mol Plant-Microbe Interact (in press)

Xu W, Bak S, Decker A, Paquette SM, Feyereisen R, Galbraith DW (2001) Microarray-based analysis of gene expression in very large gene families: the Cytochrome P450 gene superfamily of *Arabidopsis thaliana*. Gene **272:** 61–74

Yan H, Mudge J, Kim DJ, Shoemaker RC, Cook DR, Young ND (2003) Estimates of conserved microsynteny among the genomes of *Glycine max*, *Medicago truncatula*, and *Arabidopsis thaliana*. Theor Appl Genet (in press)

Yuan Z, Njiti V, Meksem K, Iqbal MJ, Triwitayakorn K, Kassem MA, Davis GT, Schmidt ME, Lightfoot DA (2002) Identification of yield loci in soybean populations that segregate for disease resistance. Crop Sci **42:** 271–277

Zhu H, Cannon SB, Young ND, Cook DR (2002) Phylogeny and genomic organization of the TIR and non-TIR NBS-LRR resistance gene family in *Medicago truncatula*. Mol Plant-Microbe Interact **15:** 529–539

Zhu H, Kim DJ, Baek FM, Choi HK, Ellis L, Kuester H, McCombie R, Peng HM, Cook DR (2003) Syntenic relationships between *Medicago truncatula* and *Arabidopsis thaliana* reveal extensive divergence of genome organization. Plant Physiol **131:** 1018–1026

Zobrist K, Meksem K, Wu C, Tao Q, Zhang H, Lightfoot DA (2000) Integrated physical mapping of the soybean genome: a tool for rapid identification of economically important genes. Soybean Genet Newslett **27:** 10–15