

Legume Biology: Sequence to Seeds

Research on legumes is driven, to a large extent, by their importance as food crops worldwide. Some 25% of the world's major crop production is derived from legumes, and more than one-third of humanity's nutritional nitrogen requirement comes from legumes. Moreover, the ability of many legumes to establish symbiosis with nitrogen-fixing bacteria reduces the cost of growth and makes them a valuable source of soil nitrogen for other crops. Progress in legume research and advances in genomics hold the promise of improved crop yields, better nutrition, and the development of novel bio-products. Although studied in a particular context, discoveries in signal transduction, organ and seed development, microbial infection, nutritional stress, and the evolution of genomes are of broad interest since they contribute to our understanding of fundamental problems in plant biology.

Progress in legume biology has been greatly enhanced by the development of two model plant systems, *Lotus japonicus* and *Medicago truncatula*. Both organisms are diploid, have modest genome sizes, and are amenable to genetic manipulation. A *M. truncatula* handbook containing protocols and other useful information is freely accessible on the Web at <http://www.noble.org/MedicagoHandbook/>. The success of these systems is reflected in the number of articles in this issue devoted to them. Genome sequencing of these two systems is well under way, and it has facilitated positional cloning of genes of interest and given insights into legume evolution. Genome-wide synteny between *L. japonicus* and *M. truncatula* with other legumes indicates that these model plants will yield important clues into the biology of plant crops, which will be further supplemented by the soon-to-be-completed soybean (*Glycine max*) genome project.

As evidenced from recent work with *Arabidopsis thaliana*, transcription factors play a key role in plant biology, ranging from regulating organ and tissue development to mediating plant response to biotic and abiotic stress. In silico evaluation of gene indices and bioinformatics approaches have been instrumental in the recent identification of more than 1,100 putative transcription factors in legumes. Moreover, gene-specific primers have been developed for all transcription factors in *Medicago* and soybean. Although the functional role is known for less than 1% of the transcription factors in legumes, in recent years several have been identified that affect root nodule formation and function. These findings support the idea of adaptation of ubiquitous plant phenomena to the development of a phylogenetically restricted organ. For example, the NIN regulatory protein was identified several years ago from a nonnodulating *L. japonicus* mutant, but NIN is now known as the founding member of a transcription factor family found throughout the plant kingdom. Furthermore, the ubiquitous plant hor-

mone cytokinin is involved in a signaling cascade that includes NIN.

Nitrogen-fixing symbiosis of a given legume is generally limited to infection by one or a limited number of rhizobial species. Nevertheless, work in this area has shed light on plant-microbe interactions more generally. Infection by arbuscular mycorrhizae fungi is prevalent in angiosperms. These symbiotic fungi extend from the plant roots into the soil to acquire nutrients, particularly phosphorus, for both organisms. Interestingly, genetic loci originally identified as nodulation mutants are also impaired in arbuscular mycorrhizae infection, and the two symbiotic systems clearly share common signaling pathways. Understanding how legumes discriminate between a pathogen and symbiont is a developing area that is just now beginning to be understood. Once again, a theme that emerges is that despite the different outcomes of pathogenic versus symbiotic infections, legumes use similar mechanisms to recognize beneficial and disease-causing bacteria. For example, LysM receptor kinases have been shown to mediate Nod factor recognition. Orthologs of LysM receptor kinases have been identified in rice (*Oryza sativa*) and Arabidopsis. Furthermore, a rice chitin oligosaccharide elicitor-binding protein has two LysM domains.

The food value of many legumes lies, of course, in the seeds. Seeds have high nutritional value for both human consumption and animal feed. The physiology and molecular biology of legume seeds has been studied for many years, with the ultimate goal of increasing quality and food value. Genome sequencing and seed EST profiling have provided expansive resources for marker-assisted selection and QTL analysis for seed traits. Recent findings that many seed QTLs map to clusters in the genome may provide anchors for identifying specific genes controlling seed quality. Soybean seeds are especially important worldwide. Resources and technologies such as microarrays, ESTs, transformation, and RNAi are now available for soybean. Furthermore, the soybean genome sequence and the high synteny between soybean and the model legumes should facilitate positional cloning and other genetic procedures. In this issue, Le et al. describe the use of laser capture microdissection in conjunction with other molecular techniques to identify genes necessary to make a soybean seed.

Although this issue of *Plant Physiology* cannot provide an update on every facet of legume research, it should give one a sense of the pace of progress and the realization that this field spills over into many other areas of plant biology. We hope that the general readership of *Plant Physiology* will find this Focus Issue on Legume Biology interesting, informative, and relevant to their own research area.

Mark R. O'Brian
Carroll P. Vance
Focus Issue Editors