

Legume Crop Genomics

Editors

Richard F. Wilson
USDA-ARS-NPS
Beltsville, Maryland

H. Thomas Stalker
North Carolina State University
Raleigh, North Carolina

E. Charles Brummer
Iowa State University
Ames, Iowa



Champaign, Illinois

AOCS Mission Statement

To be the global forum for professionals interested in lipids and related materials through the exchange of ideas, information science, and technology.

AOCS Books and Special Publications Committee

M. Mossoba, Chairperson, U.S. Food and Drug Administration, College Park, Maryland
R. Adlof, USDA, ARS, NCAUR, Peoria, Illinois
J. Endres, The Endres Group, Fort Wayne, Indiana
T. Foglia, USDA, ARS, ERRC, Wyndmoor, Pennsylvania
L. Johnson, Iowa State University, Ames, Iowa
H. Knapp, Deaconess Billings Clinic, Billings, Montana
A. Sinclair, RMIT University, Melbourne, Victoria, Australia
P. White, Iowa State University, Ames, Iowa
R. Wilson, USDA, REE, ARS, NPS, CPPVS, Beltsville, Maryland

Copyright © 2004 by AOCS Press. All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means without written permission of the publisher.

The paper used in this book is acid-free and falls within the guidelines established to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Legume crop genomics / editors, Richard F. Wilson, H. Thomas Stalker,
E. Charles Brummer

p. cm.

ISBN 1-893997-48-0 (alk. paper)

1. Legumes—Genome mapping. 2. Legumes—Genetics. I. Wilson, Richard F., 1947- II. Brummer, E. C. (E. Charles) III. Stalker, H. T. (Harold Thomas), 1950-

SB317.L43L435 2004

633.3'04233—dc22

2004003151

Printed in the United States of America

08 07 06 05 04 5 4 3 2 1

Preface

As we take initial steps into the 21st century, we find ourselves standing at the threshold of the greatest period of scientific discovery yet known to *Homo sapiens* L. As typically happens at the dawn of historically great eras in the advancement of knowledge, this realization comes to most of us with deliberate speed until our attention is captivated by an event of immense significance, such as the “small step for man, giant leap for mankind” on the lunar surface. In a similar vein, the Human Genome Project represents the seminal event that shines light upon the threshold at which we now stand, and signifies man’s quest to understand the mechanisms that enable “life” as we know it. The increasingly familiar term for such scientific scrutiny is *genomics*.

The advent of the genomics era did not spring upon us overnight, like some magic beanstalk. Rather, genomics is a dynamic, technology-driven process that may have originated with man’s initial efforts to accelerate nature’s selection for enhanced competence among biological organisms. Nevertheless, the success of the Human Genome Project has catapulted the evolution of genomics into a more sophisticated science that may now be synonymous with efforts to obtain the primary DNA-oligonucleotide sequence of entire genomes, within nearly every organism on Earth. Yet it is important to understand that generating DNA-sequence information from a multitude of chromosomal fragments is only the beginning of the journey to gain insight to the structure, organization, and function of genes in a given species. Related-species specific strategies must be devised to achieve useful assembly and interpretation of the derived genomic information. In many ways these necessary strategies are akin to tactics for piecing together an intricate jigsaw puzzle. However, as is evident throughout the plant kingdom, the complexity of genomic analysis of legume crops is often elevated by the presence of multiple sets of genomes, where each genomic complement represents a slightly different puzzle. Thus, assembly of legume crop genomes of varying size and complexity requires a high level of well-defined research strategies to ensure proper DNA-sequence analyses.

Rising scientific enthusiasm for genomic investigations among living organisms also creates an extraordinarily competitive environment for appropriations that would finance these rather expensive ventures. Hence, the most successful genomic strategies not only develop and deploy innovative fundamental technologies that help move genomic science forward, but also build research coalitions among scientists to improve the efficiency of genomic investigations among related species. For example, the U.S. Legume Crops Genomics Initiative (USLCGI) was organized under the auspices of the American Soybean Association, United Soybean Board, National Peanut Foundation, USA Dry Pea and Lentil Council, the National Dry Bean Council, and the Alfalfa Council to facilitate communication and cooperation among scientists with an interest in genomic research on soybeans, peanuts, pea and lentils, common bean, alfalfa, and model-legume crops. USLCGI was founded on the premise that the development of an integrated legume genomics research system

would enhance the ability to leverage information across legume crops and model species. In keeping with that mission, demonstrated cooperative interaction of this nature not only ensures timely research progress in all legume crops associated with USLCGI, but also enhances the competitive position of USLCGI within the framework of the National Plant Genome Initiative, which is coordinated by the Interagency Working Group on Plant Genomes, Committee of Science, National Science & Technology Council.

With that background, one purpose of this volume is to document the initial research strategies, the development of genomic tools and resources, and the legume-community consensus on the research objectives that will guide the genomic characterization of major legume crops. However, another goal is to present this technical theme in a manner that should help many readers answer the question, “What is genomics?” In that regard, [Chapter 2](#), by Weissinger and Allina, provides a strong foundation for developing an understanding of the scope and of the jargon of genomics. [Chapter 3](#), by Doyle, shows how the many different legume species are related to one another; and [Chapters 9](#) and [10](#), by Young and by Schlueter *et al.*, delve into the question of how the genomes of major legume crops have evolved. These chapters establish a broad context for the following discussions of how DNA sequences are used to identify genes, to determine gene function, and to reconstruct major portions of gene-rich regions in legume chromosomes. Then, perhaps most important of all, the [Chapters 17–19](#) give insight into some of the practical applications of legume crop genomics. These examples include ways to (i) expedite genetic enhancement of seed-constituent quality in legume crops, (ii) control major pests that limit legume crop productivity, such as soybean cyst nematode, and (iii) abate allergenic response among individuals who may be sensitized to specific proteins in legume crop species. These final chapters should help readers formulate an opinion on the question, “Why is genomic research needed?” In considering that query, we hope this volume has shown that genomics represents a quantum leap in technology; and that the application of this technology in legume-crop enhancement will ensure that U.S. agriculture remains competitive in domestic and global markets for legumes and legume crop products.

*Richard F. Wilson
H. Thomas Stalker
E. Charles Brummer*

Contents

Preface

- Chapter 1 **Development and Status of the U.S. Legume Crops Genomics Initiative**
H. Roger Boerma and Marc Curtis
- Chapter 2 **An Introduction to Plant Genomics**
Arthur K. Weissinger and Sandra M. Allina
- Chapter 3 **Legume Phylogeny: Context for the Family, Major Groups, and the Relationships of Crop and Model Species**
Jeff J. Doyle
- Chapter 4 **Genomics and Genetic Diversity in Common Bean**
Phillip McClean, Paul Gepts, and James Kamir
- Chapter 5 **Genomics and Genetic Improvement in the Cool Season Pulse Crops Pea, Lentil, and Chickpea**
N.F. Weeden and F.J. Muehlbauer
- Chapter 6 **Genomics and Genetic Enhancement of Peanut**
Andrew H. Paterson, H. Thomas Stalker, Maria Gallo-Meagher, Mark D. Burow, Sangam L. Dwivedi, Jonathan H. Crouch, and Emma S. Mace
- Chapter 7 **Genomics Research in Alfalfa, *Medicago sativa* L.**
E. Charles Brummer
- Chapter 8 ***Medicago truncatula* as a Model Legume**
Sergey Ivashuta, Steven Gantt, and Carroll Vance
- Chapter 9 **Comparative Genomics of *Glycine max*, *Medicago truncatula*, Other Legumes, and *Arabidopsis thaliana***
Nevin Dale Young
- Chapter 10 **EST Projects for the Study of Genome Evolution in Legumes.**
J. Schlueter, P. Dixon, and R. Shoemaker
- Chapter 11 **The Gene Space of the Soybean Genome**
Scott Jackson, Barbara Hass Jacobus, and Janice Pagel
- Chapter 12 **Targeting Induced Local Lesions in Genomes—TILLING**
Rae Ritchie, Hunt Wiley, Dan Szymanski, and Niels Nielsen
- Chapter 13 **Genetic Engineering of Soybean: Strategies and Utility**
Tom E. Clemente and Ted M. Klein

- Chapter 14 **Mapping and Sequencing the Soybean Genome**
Gary Stacey and Henry Nguyen
- Chapter 15 **Bioinformatics: The Interpretation of Genomic Information**
Volker Brendel, Xiaokang Pan, and Michael E. Sparks
- Chapter 16 **Soybean Microarrays: A Genomic Tool for Crop Improvement**
Steven J. Clough and Lila O. Vodkin
- Chapter 17 **DNA Marker-Assisted Selection for Improvement of Soybean Oil Concentration and Quality**
Vincent R. Pantalone, David R. Walker, Ralph E. Dewey, and Istvan Rajcan
- Chapter 18 **Genomic Approaches for Developing Soybeans with Resistance to Pests (Soybean Cyst Nematode)**
Benjamin F. Matthews
- Chapter 19 **Allergenic Responses to Legume Proteins**
Eliot M. Herman

Chapter 1

Development and Status of the U.S. Legume Crops Genomics Initiative

H. Roger Boerma^a and Marc Curtis^b

^a Center for Soybean Improvement, University of Georgia, Athens, GA 30602-6810;

^b American Soybean Association, 12125 Woodcrest Executive Dr., Suite 100, St. Louis, MO 63141-5009.

Introduction

Legumes, together with cereals, have been fundamental to the development of modern agriculture. Since the dawn of civilization, many legume species have been instrumental in supplying human food (e.g., soybean, *Glycine max* (L.) Merr.; common bean, *Phaseolus vulgaris* L.; pea, *Pisum sativum* L.; peanut, *Arachis hypogaea* L.; lentil, *Lens culinaris* Medik; and chickpea, *Cicer arietinum* L.), edible oils (peanut, soybean), and animal fodder and forage (alfalfa, *Medicago sativa* L.; and clovers, *Trifolium* spp.) (1). Legumes are second only to grasses in importance for human and animal dietary needs. Worldwide, grain and forage legumes are grown on 12–15% of the arable land (about 180 million ha) (2). They provide 33% of the nutritional nitrogen requirements of humans (3). In 2002, U.S. soybean production was 79 million metric tons and alfalfa production was 74 million metric tons (4, 5). The direct value of U.S. production of soybean and alfalfa in 2002 was \$20 billion.

Agriculturally important legumes provide added value through their symbiosis with bacteria that annually fix 40 to 60 million metric tons of atmospheric nitrogen (6). This unique ability of legumes reduces the dependence of farmers on expensive chemical fertilizer, reduces our dependence on petroleum products, and improves soil and water quality. One of the driving forces behind sustainable agriculture and protection of the environment is effective management of nitrogen in farming systems. Intensive farming practiced in developed countries is predicated on using large amounts of nitrogen fertilizer. This practice has led to significant deterioration of water, soil, and air quality. As the world's population approaches 10 billion within the next half century, nitrogen needs for increased crop production will exacerbate current environmental problems. Increased cultivation and productivity of legumes would ameliorate environmental degradation, reduce the depletion of nonrenewable resources, and provide adequate nitrogen for the population (1). Legumes clearly play a significant role in protecting human health, increasing farm profitability, and mitigating environmental problems.

The legume species, pea, was the experimental organism for Mendel's pioneering genetics research. Since the rediscovery of Mendel's findings, crop legumes have been the focus of intensive genetic studies to improve yield, quality, resistance to biotic and abiotic stresses, and to extend the geographic range of adaptation. As a result, certain legume crops have well-studied genetic systems characterized by classical biochemical and physical markers, cytogenetic analysis, chemically induced mutations, and DNA marker-based genome linkage maps. Yet in many of these well-defined systems, comprehensive genetic analysis is limited due to the large size of the genomes of legume crops. Furthermore, few of the basic tools required for modern genome analysis, including polymerase chain reaction (PCR) based DNA markers, expressed sequence tag (EST) databases, or bacterial artificial chromosome (BAC) libraries, have been developed for most legume species. No crop legume has an integrated genetic, physical, and transcript map (7). Furthermore, although efficient transformation has recently been achieved in several of the crop legumes, some of the important crop species remain recalcitrant to transformation (8).

The genomes of most crop legumes are large and relatively complex. For example, soybean is an ancient polyploid and alfalfa an autotetraploid with genome sizes of approximately 1,200 megabases (Mb) and 1,600 Mb, respectively (9). The genome size of cultivated peanut is 2,800 Mb and pea is 4,000 Mb. These larger sizes significantly complicate the development of ordered physical maps of the genome, as well as the identification and location of important genes. Legume genomes contain considerable regions of repetitive DNA, making their complete sequencing financially tenuous at the current cost per finished base. Syntenic relationships within botanical families, however, make it possible to use plant species with much smaller genomes to facilitate understanding of those with large genomes. For example, the recent complete sequencing of the smaller genomes of *Arabidopsis thaliana* (L.) Heyn. (128 Mb) and rice, *Oryza sativa* L. (425 Mb), provide the platform for genome analysis of more complex species such as canola (*Brassica napus* L.), broccoli (*Brassica oleracea* L. var. *italica*), corn (*Zea mays* L.), and wheat (*Triticum aestivum* L. Thell.). Information from the *Arabidopsis* and rice genomes is rapidly being translated across the complex species to enhance disease and pest resistance, yield, and compositional quality of the seed.

To expedite and simplify genome analysis of crop legumes, it has been proposed that parallel analysis of a legume with a smaller genome be considered (10). Recently, studies sponsored by the NSF Plant Genome Program have shown that the barrel medic, *Medicago truncatula* Gaertner, is an ideal candidate for parallel analysis with crop legumes. Barrel medic is a diploid; has a relatively small genome (about 450 Mb); has a rapid generation time; is self-compatible; and appears to have synteny with alfalfa and also to a somewhat lesser degree with pea and soybean (11–13). Comparative analysis with other legume crops would provide additional advantages by complementing genetic knowledge available in the different legume species. For example, soybean is a major crop with significant prior study of genetics and crop and seed physiology. Common bean benefits from relatively well-developed

genetic studies and ample polymorphism within the cultigen. Peanut possesses a unique reproductive physiology, which can contribute to a greater understanding of plant reproductive biology.

Development of the Initiative

The U.S. Legume Crops Genomics Initiative (USLCGI) was organized to facilitate communication and cooperation among growers and scientists with an interest in genomics research on U.S. legume crops. This initiative was founded on the premise that the development of an integrated legume genomics research plan would enhance scientists' ability to leverage information across legume crops and model species. In addition, the initiative had as its goal the enhancement of federal funding for genomics research for all U.S. legume crops. During the late 1990s independent groups of legume researchers developed crop-specific genomic plans (14). From a review of these plans, it became clear that achieving the proposed research was an enormous and expensive challenge given the large number of important U.S. legume crops. In addition, few of these legume crops possessed the necessary genetic and genomics resources to effectively utilize existing genomics data from model organisms and the rapidly evolving approaches to whole genome evaluations.

Given the importance of legume crops, their limited genomic and genetic resources, and the inadequate level of research funding, it was decided to convene a workshop of U.S. legume genomic scientists to review the current status of legume genomics research and determine the highest priority research for the major U.S. legume crops (15). In July 2001, 26 legume scientists with knowledge of structural and functional genomics, DNA markers, transformation, bioinformatics, and legume crop improvement were invited to participate in a workshop held at Hunt Valley, Maryland. These scientists came from public and private organizations in 20 states and represented many of the leading geneticists in the United States working with legumes. During this 2-day workshop, scientists reached consensus on six critical areas of genomic research in which progress was needed across all legume species: (i) genome sequencing of strategic legume species, (ii) physical map development and refinement, (iii) functional analysis, (iv) development of DNA markers for comparative mapping and breeding, (v) characterization and utilization of legume biodiversity, and (vi) development of legume data resources (see [Table 1.1](#)). At the same time that the critical areas were defined, it was recognized that the level of resource development for the various crop species was significantly different, with no single species being advanced in all areas. Because these six areas encompass a wide research spectrum, it was determined that improved coordination among scientists was critical to achieve significant progress across crop legume species.

Following the Hunt Valley priority-setting workshop, leaders of six U.S. commodity organizations (American Soybean Association, United Soybean Board, National Peanut Foundation, USA Dry Pea and Lentil Council, the National Dry

TABLE 1.1Priority Genomics Research Goals Established at the 2001 Hunt Valley Workshop

- Genome Sequencing of Strategic Legume Species
 - Sequence gene-rich regions of soybean, common bean, and peanut
 - Produce whole-genome sequencing of barrel medic (*Medicago truncatula*)
 - Physical Map Development and Refinement
 - Develop physical maps in peanut and common bean
 - Refine and/or complete physical maps in barrel medic and soybean
 - Develop transcript maps of known genes (ESTs)
 - Integrate maps among taxonomically key species
 - Functional Analysis: Transcriptional and Genetic
 - High-throughput stable and transient gene transfer systems
 - Assignment of gene function
 - Expression analysis
 - Proteomics, metabolomics, and metabolic reconstruction
 - Gene knockout systems
 - Development of DNA Markers for Comparative Mapping and Breeding
 - Universal set of PCR-based legume Sequence Tag Sites
 - Establishment of local and global similarities of different gene arrangements
 - Quantitative trait loci (QTL) discovery
 - Determination of levels of genome conservation among the legume crops
 - Characterization and Utilization of Legume Biodiversity
 - Broadening and refining legume phylogeny
 - Establishing levels of biodiversity within the different legume crops
 - Comparing closely related species for gene discovery
 - Crop-microbe co-evolution
 - Multigene family evolution
 - Domestication of legume crops
 - Preserving and utilizing germplasm
 - Development of a Legume Data Resource
 - Development of a legume-wide database (molecular, genetic, expression, diversity, and breeding) publicly available through the Web
 - Leveraging data from each legume species for greater efficiency
-

Bean Council, and the Alfalfa Council) met to review the newly established research priorities. The grower leaders agreed to accept the scientists' consensus on research priorities and to request formal approval from their grower organizations to work together as a unified group in support of the priority genomics research on legume crops as part of a U.S. Legume Crops Genomics Initiative (15). During the fall of 2002, the outcome of the Hunt Valley Workshop was presented to the various legume crops grower organizations. Each organization voted to accept the genomic research priorities and to participate as a formal member of the planned USLCGI.

Activities of the USLCGI

At the March 2002 USLCGI meeting, the grower leaders and species coordinators (researchers selected at the Hunt Valley workshop to represent each major legume crop) agreed on the formal structure of the USLCGI Steering Committee. The committee consisted of two grower members and one research member for each of the following legume commodities: (1) alfalfa and clovers, (2) dry bean, (3) dry pea and lentil, (4) peanut, and (5) soybean. In addition, one research scientist was selected to represent model legumes. Each commodity polled its researchers to nominate their research representative for the USLCGI Steering Committee. The grower members for each crop were selected by their respective grower organization(s), and the nominated research representatives were approved by their respective grower organization. The current membership of the USLCGI Steering Committee is shown in Table 1.2.

TABLE 1.2
U.S. Legume Crops Genomics Steering Committee Membership
and Area of Representation

- Alfalfa and clovers
 - Mr. Mark McCaslin, Forage Genetics, Prior Lake, MN
 - Dr. Charlie Brummer,^{a,b} Iowa State University, Ames, IA
 - Open position
 - Dry pea and lentil
 - Mr. George Anderson, USA Dry Pea & Lentil Council, Moscow, ID
 - Mr. Todd Scholz, USA Dry Pea & Lentil Council, Moscow, ID
 - Dr. Norman Weeden,^b Montana State University, Bozeman, MT
 - Dry bean
 - Mr. Tom Grebb, Central Bean Co., Inc., Quincy, WA
 - Mr. Steve Antonius, ADM Edible Bean Research, Caldwell, ID
 - Dr. Paul Gepts,^b University of California, Davis, CA
 - Model legumes
 - Dr. Nevin D. Young,^b University of Minnesota, St. Paul, MN
 - Peanut
 - Mr. Joe Boddiford, Georgia Peanut Commission, Tifton, GA
 - Mr. Howard Valentine, American Peant Council, Big Canoe, GA
 - Dr. Tom Stalker,^b North Carolina State University, Raleigh, NC
 - Soybean
 - Mr. Jim Stallstrom, United Soybean Board, Winthrop, MN
 - Mr. Marc Curtis,^c American Soybean Association, Leland, MS
 - Dr. Randy Shoemaker,^b USDA-ARS, Ames, IA
-

^a Vice-chair of the USLCGI Steering Committee.

^b Research member of the USLCGI Steering Committee.

^c Chair of the USLCGI Steering Committee.

While agricultural research has been exemplary in linking basic research with the needs of agriculturalists, the USLCGI is innovative in bringing grower leadership and genetic and genomic researchers of U.S. legume crops together to develop the research agenda cooperatively. Historically, the grower leadership of each crop has acted independently of other commodities to move its own genomics research agenda forward. Given the apparent similarity of the genomes of legumes, however, there are opportunities to focus a research agenda that would not only benefit individual legume crops, but would also effectively move forward the science of this important plant family and result in the enhanced sustainability and profitability of U.S. agriculture.

The USLCGI Steering Committee has continued to meet regularly to formulate bylaws, elect its leadership, develop additional plans, and communicate the importance of the USLCGI to Federal funding organizations (Table 1.3). For example, meetings with Under Secretary Joe Jen and Deputy Under Secretary Rodney Brown of the U.S. Department of Agriculture and representatives from the National Science Foundation have served to communicate the purpose of the USLCGI and the high-

TABLE 1.3

Major Activities Related to the Development and Organization of the U.S. Legume Crops Genomics Initiative (USLCGI)

Date	Activity
July 2001	Legume Genomics Workshop and associated grower leader meeting, Hunt Valley, MD
December 2001	Approval of Hunt Valley research priorities by individual grower organizations
January 2002	Presentation of Hunt Valley research priorities in the Legumes Workshop at the Plant & Animal Genome X, San Diego, CA
March 2002	USLCGI organizational meeting of grower leaders and species coordinators, Baltimore, MD
April 2002	Election of USLCGI Steering Committee by grower organizations and researchers
June 2002	USLCGI Steering Committee meeting to initiate development of bylaws and meeting with USDA administrators, Washington, DC
October 2002	USLCGI Steering Committee meeting for development of bylaws, election of chair, vice-chair, and executive committee, and for meeting with NSF administrators, Washington, DC
November 2002	Initial issue of <i>USLCGI Update</i> (Diane Bellis, editor), a newsletter to keep USLCGI Steering Committee members, researchers, and grower leaders informed on important legume genomics issues and opportunities
January 2003	USLCGI Steering Committee meeting to plan the NSF Research Coordination Network Grant and development of a Legume Genomics Monograph, San Diego, CA
June 2003	Submitted NSF Research Coordination Network Grant (PIs: T. Stalker, C. Brummer, P. Gepts, R. Shoemaker, N. Weeden, and N. Young)
November 2003	USLCGI Steering Committee meeting to develop legislative strategy and meet with NIH, NSF, and USDA staff, Washington, DC

priority genomics research needs for legume crops. In addition, the Steering Committee has served as the conduit through which research priorities are related to commodity-funding organizations at both the state and national levels. In fact, one can see the effect of the USLCGI in revised genomics plans of each individual legume crop (16).

Because the research members of the USLCGI were nominated by their peers and approved by their respective grower organizations, they have credibility with their fellow researchers and a good working relationship with the grower members on the committee. By partnering researchers with grower leaders, the USLCGI has created broad-based support for legume crops genomics research in many states. Continued communication among commodity leaders of the major U.S. legume crops and the scientific community will be critical for applying basic scientific information to solve real-world problems in legume crop improvement.

The USLCGI has already shown that a diverse group of commodity organizations can work together to achieve a common goal. The steering committee has established its research goals and organizational structure, elected its leadership, developed previously nonexistent lines of communication, and worked together to submit a networking grant to the National Science Foundation (Table 1.3). The final impact of the USLCGI on the coordination and funding of legume crops research has not yet been determined, but the initiative is positioned to benefit society by more efficiently meeting the nutritional needs of humans while protecting the environment and increasing the sustainability and profitability of U.S. agriculture.

References

1. Graham, P.H., and C.P. Vance, Legumes: Importance and Constraints to Greater Use, *Plant Physiol.* 131:872–877 (2003).
2. Food and Agricultural Organization of the United Nations, FAOSTAT Agricultural Data, <http://apps.fao.org/page/collections?subset=agriculture> (2003).
3. Vance, C.P., P.H. Graham, and D.L. Allan, Biological Nitrogen Fixation: Phosphorus a Critical Future Need, in *Nitrogen Fixation: From Molecules to Crop Productivity*, edited by F.O. Pedrosa, M. Hungria, M.G. Yates, and W.E. Newton, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000, pp. 506–514.
4. Golbitz, P., 2003 *Soya & Oilseed Bluebook*, Soyatech, Inc., Bar Harbor, Maine, 2003, pp. 321–364.
5. United States Department of Agriculture, National Agricultural Statistics Service, Crop production 2002 summary, <http://usda.mannlib.cornell.edu/reports/nassr/field/pcp-bban/cropan03.txt> (2003).
6. Smil, V., Nitrogen in Crop Production, *Global Biogeochem Cycles* 13:647–662 (1999).
7. VandenBosch, K.A., and G. Stacey, Summaries of Legume Genomics Projects from around the Globe: Community Resources for Crops and Models, *Plant Physiol.* 131:1–26 (2003).
8. Somers, D.A., D.A. Samac, and P.M. Olhoft, Recent Advances in Legume Transformation, *Plant Physiol.* 131:892–899 (2003).
9. Arumuganathan, K., and E.D. Earle, Nuclear DNA Content of Some Important Plant Species, *Plant Mol. Biol. Rep.* 9:211–215 (1991).

10. National Research Council, The National Plant Genomics Initiative: Objectives for 2003–2008, www.nap.edu/books/0309085217/html/ (2002).
11. Endre, G., A. Kereszt, Z. Kevei, S. Mihacea, P. Kaló, and G.B. Kiss, A Receptor Kinase Gene Regulating Symbiotic Nodule Development, *Nature* 417:962–966 (2002).
12. Gualtieri, G., O. Kulikova, E. Limpens, D.J. Kim, D.R. Cook, T. Bisseling, and R. Geurts, Microsynteny between Pea and *Medicago truncatula* in the SYM2 Region, *Plant Mol. Biol.* 50:225–235 (2002).
13. Yan, H.H., J. Mudge, D.J. Kim, D. Larsen, R.C. Shoemaker, D.R. Cook, and N.D. Young, Estimates of Conserved Microsynteny among the Genomes of *Glycine max*, *Medicago truncatula* and *Arabidopsis thaliana*, *Theor. Appl. Genet.* 106:1256–1265 (2003).
14. Boerma, H.R., D. Buxton, M. Kelly, and K. Van Amburg, Soybean Genomics White Paper January 2000, http://129.186.26.94/Genomics/Soybean_Genomics.html (2000).
15. Boerma, H.R., J. St. John, and J. Yezak Molen, U.S. Legume Crops Genomics Workshop White Paper, www.legumes.org/ (2001).
16. Bellis, D., H.R. Boerma, E. Ready, and R. Wilson, Strategic Plan for Soybean Genomics 2003 to 2007, http://129.186.26.94/external_links.html (2003).