SEMI-ANNUAL TECHNICAL REPORT
APRIL 1, 2000 THROUGH SEPTEMBER 30, 2000
FOR
DE-FC05-92OR22072
TO
UNITED STATES DEPARTMENT OF ENERGY

OBJECTIVES:

The Consortium for Plant Biotechnology Research, Inc. ("CPBR") continues to operate according to objectives outlined in the proposal funded through the cooperative agreement. The italicized objectives below are addressed in this report, which covers the period April 1, 2000, through September 30, 2000.

1. Update the research agenda using information obtained from member companies.

2. *Identify and implement research projects that are deemed by industrial, scientific, and sponsoring agency evaluation to address significantly the problems and future of U.S. energy resources and that are relevant to the Department of Energy's mission.*

   Specifically,
   - Announce research grants competition through a Request for Preproposals.
   - Conduct a dual-stage review process:
     Stage one: industrial and DOE review of preproposals.
     Stage two: peer review, scientific consultants' review, DOE review of full proposals and Project Recommendation Committee evaluation and recommendation for funding.
   - Board of Directors approval of recommended awards.
   - *Conduct ongoing project management.*
   - *Obtain semiannual, annual and final reports for evaluation of research goals and technology transfer.*
   - *Present reports to DOE.*

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

MAJOR ACCOMPLISHMENTS:

Governance:

No changes to report.

Administrative matters:

Ms. Debbie Thomas left the position of Fiscal Assistant.

Federal Sponsors:

U.S. Department of Energy.

Energy from Biomass Competitions:

Awards in the CPBR 2000 Energy From Biomass Competition have not yet been made pending release of the FY 2000 CPBR appropriation.

The CPBR 2001 Energy From Biomass Competition was initiated with the Request for Preproposals issued in December 1999. A total of 51 new preproposals were received, most of them transmitted electronically. An additional nine preproposals representing fundable proposals from the 2000 competition were added to the new preproposals sent for industrial review.

In order to conserve financial resources for research, the symposium scheduled for March, 2000, was postponed. Evaluation of preproposals by industry was conducted by mail with the preproposals sent on a CD in PDF format. DOE was given the opportunity to evaluate the preproposals.

Invitations issued in March resulted in the submission of 46 full proposals in the 2001 Energy from Biomass competition. Of the proposals continuing in the competition, 26 are new. The remaining 20 proposals represent top proposals considered fundable in the 2000 competition, but not yet funded.

Reviews of the submitted proposals were received

New member program:

- Several universities requested and were provided membership information.
- Membership information was provided upon request to several companies.
- Due to reduced funding, CPBR staff was unable to exhibit and/or attend national conferences to disseminate research information.

Evaluation of project results:

Fifteen semi-annual or final scientific progress reports (copies attached) were received during the reporting period. In addition, the following information was reported.

Technology transfer (e.g., inventions, patents, disclosures or licensing agreements) was reported by the following investigators during this reporting period:

Metrics Reports

Four Metrics Reports from sponsoring companies were received (copies attached).

Cumulative Technology Transfer Activities to Date:

53 inventions disclosed.
43 patents awarded or applied for.
23 inventions reported to be licensed to multiple licensees or under negotiation.

IMPLEMENTATION AND EVALUATION OF PROGRESS:

2000 PROJECTS

Energy from Biomass/Biofuels – Transportation and Power

DOE funding not released. Awards not yet made.

1999 PROJECTS

Energy from Biomass/Biofuels – Transportation and Power

Harvey D. Bradshaw, University of Washington — **IN PROGRESS**
Map-Based Cloning of Genes To Increase Poplar Biomass — *Report Enclosed*

Stanton S. Gelvin, Purdue Research Foundation — **IN PROGRESS**
Plant Genes Involved in T-DNA Integration and Radiation Sensivity — *Report Enclosed*

Jean T. Greenberg, The University of Chicago — **IN PROGRESS**
Engineering Artificial Immunity to Plant Pathogens — *Report Enclosed*

Robert Haselkorn, The University of Chicago — **IN PROGRESS**
Increasing the Energy of Plants: Molecular Genetics of Acetyl-COA Carboxylase — *Report Enclosed*

David Hildebrand, University of Kentucky — **IN PROGRESS**
Engineering Oilseeds for Epoxy Fatty Acid Accumulation

Lonnie O. Ingram, University of Florida — **IN PROGRESS**
Ethanol Production from Uronic Acid-Substituted Xylose Residues in Hemicellulose Hydrolysates — *Report Enclosed*

Jiming Jiang, University of Wisconsin — **IN PROGRESS**
Toward Cloning a Functional Rice Centromere — *Report Enclosed*

Gayle Lamppa, The University of Chicago — **IN PROGRESS**
Accumulation of Products Within the Plastid for Biomass Conversion: Test System with Cellulase — *Report Enclosed*

Yi Li, The University of Connecticut — **IN PROGRESS**
Genetic Improvement of Seed Productivity for Bioenergy Crops

John B. Ohlrogge, Michigan State University — **IN PROGRESS**
Christoph Benning — *Report Enclosed*
DNA Microarray Discovery of Gene and Networks Which Control Plant Seed Storage Products

Brenda Oppert, Kansas State University — **IN PROGRESS**
Evaluation of Insect Serpins as Biopesticides

Eric W. Triplett, University of Wisconsin-Madison — **IN PROGRESS**
Associative Nitrogen Fixation by Diazotrophic Endophytes In Switchgrass

Richard Vierstra, University of Wisconsin-Madison — **IN PROGRESS**
Development of Vectors for the Stoichiometric Accumulation of Multiple Proteins in Transgenic Crops

1998 PROJECTS

Energy from Biomass

Richard Amasino, University of Wisconsin — **IN PROGRESS**
Regulation of Plant Senescence

Zong-Ming Cheng, North Dakota State University — **IN PROGRESS**
Evaluation and Characterization of Rooting Capability of Hybrid Aspens Transformed With Rooting Genes — *Report Enclosed*

Scott A. Merkle, University of Georgia — **IN PROGRESS**
Clonal Propagation of Hybrid Southern Hardwoods — *Report Enclosed*

Basil J. Nikolau, Iowa State University — **IN PROGRESS**
How Do Plants Generate Acetyl-CoA — *Report Enclosed*

Friedrich Srienc, University of Minnesota — **IN PROGRESS**
Biodegradable Plastics from Yeast & Plants

Steven H. Strauss, Oregon State University — **IN PROGRESS**
Genes Controlling the Transition Between Vegetative and Reproductive Phases in Forest Trees — *Report Enclosed*

Eric Triplett, University of Wisconsin—Madison — **IN PROGRESS**
Engineering Sinorhizobium for Increased Alfalfa Biomass

Jack M. Widholm, University of Illinois — **IN PROGRESS**
A New Selectable Marker and Promoters Of Plant Origin — *Report Enclosed*

1997 PROJECTS

Energy from Biomass

John Davis, University of Florida — **IN PROGRESS**
Molecular Biology of Defense Responses in Populas

John Dudley, University of Illinois — **IN PROGRESS**
High Starch Adds Value to Corn for Production — *Report Enclosed*

Lee Lynd, Dartmouth College — **IN PROGRESS**
Pretreatment Process for Cellulosic Biomass

John Ohlrogge, Michigan State University — **IN PROGRESS**
Increasing Plant Oil Synthesis via Genetic Engineering of Acetyl-CoA Carboxylase

David Somers, University of Minnesota — **IN PROGRESS**
Incorporation of Value-Added Traits into Alfalfa for Biomass Energy — *Report Enclosed*

Richard Vierstra, University of Wisconsin — **IN PROGRESS**
Targeted Proteolysis: New Method for Removing Selected Intracellular Proteins

1996 PROJECTS

Energy from Biomass

Yi Li, Kansas State University (Transferred to U.Conn.) — **IN PROGRESS**
Genetic Improvement of Aspen for Wood Production

Kenneth Nickerson, University of Nebraska — **CLOSED**
Bacillus thuringiensis: Biotin Mediated Insect Toxicity Suggests Alternate Strategies for Pest Control in Energy Crops

1995 PROJECTS

Historically Black Colleges and Universities (HBCU) Initiative

Frans deBruijn, Michigan State University — **IN PROGRESS**
Edison R. Fowlks, Hampton University — *Final Report Enclosed (Fowlks)*
Frank Louws, North Carolina State University
Automated Fluorescent Genomic Fingerprinting of Bacteria

Report Period:
April 1, 2000 through
September 30, 2000

Semi-Annual Technical Report

The Department of Energy

Cooperative Agreement

DE-FC05-92OR22072

Submitted By

Dorin Schumacher, Ph.D.

President

RECEIVED JUN 15 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	H.D. Bradshaw, Jr.	
University:	University of Washington	
Agreement Number:	OR22072-89	
Project Title:	Map-based cloning of genes to increase poplar biomass	
Reporting Period and Report Type:	**From:** **1 Oct 99** **To:** **30 Jun 00**	**Check one:** [x] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. Determine the rust resistance or susceptibility phenotype of ~2000 hybrid poplars in the full sibship PMGC Family 545.
2. Use bulked segregant analysis to identify at least 100 additional AFLP markers linked to the *Mmd1* resistance gene.
3. Produce a fine structure (0.1cM) genetic map around the *Mmd1* rust resistance locus.
4. Isolate overlapping bacterial artificial chromosome (BAC) clones of *P. trichocarpa* DNA containing *Mmd1* and its flanking AFLP markers.
5. Construct a physical map of the *Mmd1* locus by making BAC contigs.

Objectives 1 and 2 were completed in Year 1 (1 Oct 98 – 30 Sep 99). Objective 3 has been completed in the first half of Year 2, along with parts of objectives 4 and 5; *i.e.*, we have a *ca.* 300kb 11-BAC contig in the vicinity of *Mmd1* containing 11 completely linked genetic markers, but as yet no markers known to flank both sides of the gene. In the remaining three months of Year 2, we will focus upon finding AFLP markers and BACs on both sides of *Mmd1*, to assure that the resistance gene is contained within a defined physical space in the genome, and therefore has been cloned. A manuscript describing the current state of the project has been written and will be submitted before the end of the funding period.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

The new science of genomics will be used to locate and isolate a gene for resistance to the poplar leaf rust disease. Poplar leaf rust is the single most important disease in limiting the biomass productivity of hybrid poplar plantations worldwide. The resistance gene, called *Mmd1*, will be located on a genetic map of the poplar chromosomes, and this positional information ultimately will be used to isolate (clone) the resistance gene. The cloned gene can be used to genetically engineer hybrid poplars for resistance to leaf rust. If this map-based (positional) cloning approach is successful for the *Mmd1* gene, the same methods may then be used to clone additional genes important for biomass yield in hybrid poplar.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

A fine-structure genetic map of the *Mmd1* locus (renamed *Mxc3* because of the specific fungal pathogen used to infect the trees has been completed (Fig. 1).

A BAC contig representing the physical structure of the poplar chromosome near the rust resistance gene is shown in Fig. 2.

Figure 1. High-resolution genetic map of the *Mxc3* rust resistance locus in hybrid poplar.

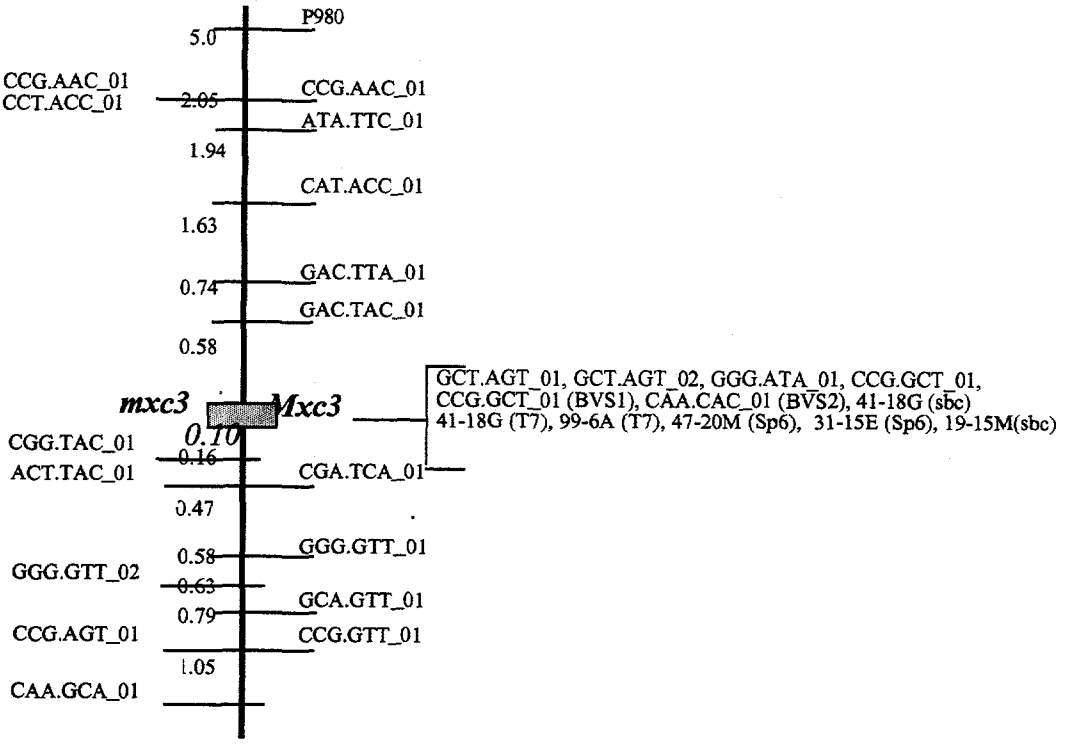


Fig. 1: The high resolution genetic linkage map of *Mxc3* consists of 20 AFLP markers and 6 markers derived from BAC-end sequences or BAC subclones. *Mxc3* and *mxc3* represent the dominant and recessive alleles, respectively. Markers are positioned to indicate their phase with respect to *Mxc3* and *mxc3*. Map distances shown for each marker represent the total distance from *Mxc3* and are based on a progeny set of 1904. Markers derived from BAC-end sequences or BAC subclones are designated by plate#-column#, row letter. T7 or Sp6 designates a marker derived from the T7 or Sp6 BAC end; sbc designates a marker from a BAC subclone. Marker P980 is an allele specific STS marker identified from previous mapping studies of the *Populus* leaf rust resistance locus *Mmd1*.

Figure 2. Putative *Mxc3* BAC contig.

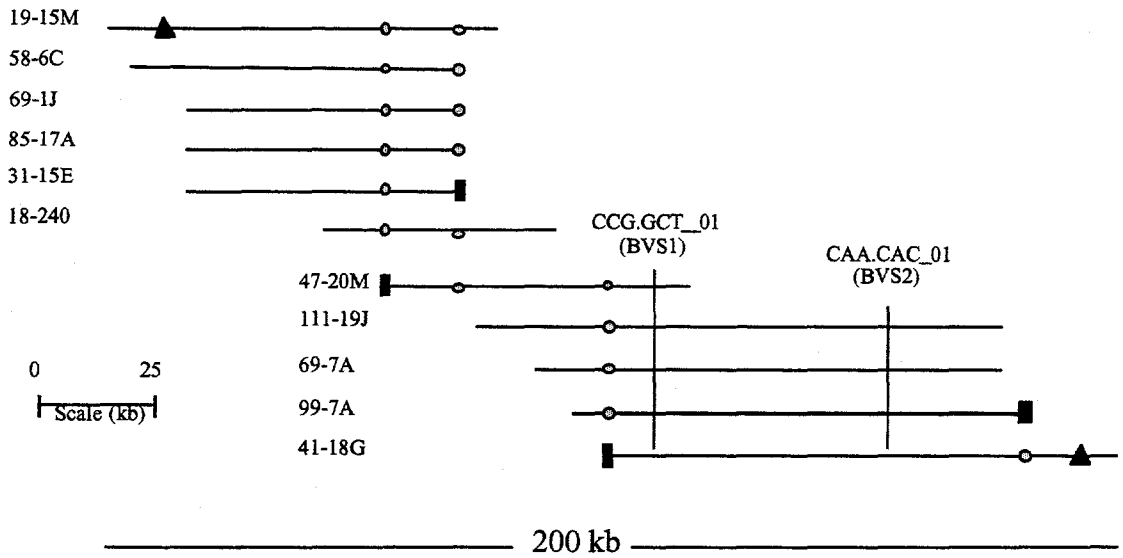


Fig. 2. Putative contig of the *Mxc3* region. BAC clones in the 200 kb contig are drawn to scale in kilobases. Horizontal arrowheads designate the he T7 side of the vector. Solid rectangles designate STS markers that were developed from BAC-end sequences and subsequently used for genetic and/or physical mapping. Shaded circles designate BAC clones that amplified with the BAC-end STS markers. Solid triangles designate markers that were developed from BAC subclones and used for genetic and/or physical mapping.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

None.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

None.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Brigid Stirling, the graduate student carrying out this project, received continued training in plant molecular genetics.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

None.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

None.

RECEIVED JUL 1 0 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Zong-Ming Cheng
University:	North Dakota State University
Agreement No:	OR22072-76
Project Title:	Evaluation and characterization of rooting capability of hybrid aspens transformed with rooting genes
Reporting Period From: 12/21/99 To: 6/30/00	**Report Type:** [X] Interim Report [ ] Final Report

Project Objectives:

- To confirm that hybrid aspen plants transformed with *iaa*M and *rol*B under three promoters by PCR amplification and Southern blotting to identify single-copy transformed plants.
- To examine the expression of these genes by GUS assay, determine the auxin contents in *iaa*M-transformed plants and sensitivity to the applied auxin in *rol*B-transformed plants.
- To evaluate the rooting capability of *in vitro* cuttings and hardwood cuttings.
- To analyze data of gene expression, auxin contents/sensitivity and results of rooting capability to determine the possible rooting mechanisms.

Layperson's Summary

One of the constraints for biomass-based fuel and electricity to be competitive is that the high cost associated with the biomass production. The goal of this project is to develop a commercially feasible hardwood cuttings propagation system for elite aspen hybrids so that these fast-growing, disease-resistant clones can be mass-produced to establish short-rotation plantations. We have genetically engineered aspen with two genes under three different regulatory promoters. The rooting tests with hardwood cuttings show that these genes, when transformed into aspen, can significantly improve the rooting capability. The rooting frequencies with hardwood cuttings increased from 15-25% of non-genetically-engineered plants to 90-100% of some genetically-engineered plants/treatments. This project is very promising for near future commercialization. The benefits of this technology include less cost for producing plant materials and establishment of plantations, maximum capture of genetic gain of the improved apsen hybrids, fast and uniform growth, and less water usage in producing plant materials and chemical inputs due to utilizing genetically improved plant materials with disease and insect resistance.

Scientific Accomplishments

- To confirm that hybrid aspen plants transformed with *iaa*M and *rol*B under three promoters by PCR amplification and Southern blotting to identify single-copy transformed plants.

  Standard PCR technique has been employed to confirm the transformation. The protocol of PCR including DNA extraction, primer design, and amplification program selection has been optimized. Forty-nine transgenic aspen hybrids with the *rol*B gene have been

confirmed by PCR. The copy number of the transferred gene has been determined in 10 plants using Southern blot analysis. One to five copies of the *rol*B gene were found in transgenic plants.

- To examine the expression of these genes by GUS assay, determine the auxin contents in *iaa*M-transformed plants and sensitivity to the applied auxin in *rol*B-transformed plants.

  Since both genes are tailed with the GUS gene, the expression of the gene can be assayed by GUS expression. Results showed that both heat shock and wounding inducible promoters are inducible. GUS expression pattern in transgenic stem was evaluated. Among the three constructs (35s-gus, heat shock-gus, and rolC-gus). The 35S-GUS and rolC-GUS constructs have similar patterns of GUS activity in the first-branch stem. GUS activity was found in all tissues except pith. The strongest activity was in the phloem. The heat shock promoter gave rise to a very strong expression only in epidermis and phloem. In the current-growing stem, GUS activity was shown in epidermis, parenchyma, vascular cambium, and primary xylem for the 35S promoter. All tissues were stained blue in current-growing stem for rol C promoter. The heat shock promoter was mainly expressed in parenchyma in current-growing stems. Therefore, the heat shock promoter expressed more tissue-specific, especially in mature stems. Less tissue-specific expression pattern of rolC promoter was confined in this study compared to 35s and heat shock promoters. GUS expression was also determined in both young and mature leaves for three promoters, mostly in veins and mesophyll. In mature leaves, no blue staining was found in main veins. No difference in the expression pattern was found among the three promoters in leaf.

  In root tips, both 35S and heat shock promoters were expressed in columella, vascular, and root apical meristem with a very strong expression in root apical meristem. Less intensity of GUS activity was shown for the rolC promoter in root tips. No GUS expression was detected in root columella for the rol C promoter.

  The efficiency of three promoters (35S, Heat Shock, and Wound Inducible) is measured using MUG assay. The inducibility of HS and wound inducible promoters are determined. The factors such as temperature, wounding, and plant growth regulator which affect the expression of the transformed GUS gene are tested and try to optimize the best environment for the gene expression.

  Determination of the auxin contents for the iaaM-transformed plants have not been done due to the leaving of one Ph.D student. The work will be started in the fall of 2000 when a new student starts in fall semester, 2000.

  The sensitivity of the rolB-transformed plants were assayed for the cellular sensitivity to exogenous auxin. With in vitro assay, results so far indicated no differences among non-transformed plants and transformed plants, possibly due to high endogenous auxin contents.

- To evaluate the rooting capability of *in vitro* cuttings and hardwood cuttings.

Rooting with hardwood cuttings were done with both iaaM and rolB transformed plants. For iaaM-transformed plants, rooting percentages varied greatly with different clones, different promoters and different treatments. The highest rooting rates were from plants transformed with GH3 promoter, with rooting rates ranged from 50% to 91%. For rolB transformed plants, rooting were done with four HS-rolB plants, five 35S-rolB plants, two win-rolB plants. The 35S-rolB plants rooted at an average of 21.7%, showing no difference among treatments with auxin concentrations (0, 100, 1000 ppm IBA). HS-rolB plants rooted at averages between 60%-90%, cuttings in several treatments rooted at 100%. Heat shock itself did not seem to promote the rooting, but need to verified with larger experiments The non-transformed plants rooted at an average of 20%. The results with rolB genes suggest that HS-rolB seems to be the most effective gene construct.

Many plants are now growing in the outdoor for the summer to get maximum shoot growth. These plants will be placed in cooler in the fall to produce hardwood cuttings. The repeat experiments will be conducted for the rooting tests.

- To analyze data of gene expression, auxin contents/sensitivity and results of rooting capability to determine the possible rooting mechanisms.

The final analysis of rooting mechanism has not done yet because auxin quantitation has not been done and more rooting data are needed which will be availabe in thefall. This will be one of the main tasks in the remaining project period.

Publications and Presentations

1. Dai, W. H. Z.-M. Cheng, and W. A. Sargent. 2000. Efficient transformation of elite aspen clones. To be presented in the Annual meeting of the American Society for Horticultural Sciences. Olando, FL. July 23-26, 2000.
2. Dai, W. H. Z.-M. Cheng, and W. A. Sargent. 2000. Expression of Beta-glucuronidase gene in aspen under control of CaMV35S, heat shock and rolC promoters. To be presented in the Annual meeting of the American Society for Horticultural Sciences. Olando, FL. July 23-26, 2000.

Technology Transfer

Commercial Accomplishments

With the final rooting results show that the genes can truly enhance the rooting capability, we will disclose and apply for the patent(s) by the end of year 2000. After that we will start to negociate the licensing agreement with the forest companies.

Educational Accomplishments

A Ph.D student has finished research and is employed as a Postdoctoral Research Associate at the USDA Forest Sciences lab on Purdue University campus.

Additional Funding

No additional funding has been received.

Key Personnel Hiring or Turnover

One of the Ph.D students, Mike Bosela, has been taking a postdoctoral job at the USDA Forest Research Laboratory on Purdue University campus since July, 1999.

A new Ph.D student will start in late August, 2000, to continue the work.

RECEIVED JUL 0 5 2000
(email)

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	John W. Dudley, Ph.D.	
University:	University of Illinois	
Agreement Number:	OR22072-68	
Project Title:	*High starch adds value to corn for production*	
Reporting Period and Report Type:	**From:** 9/1/99 **To:** 6/30/00	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. Develop an efficient marker-assisted backcrossing method.

This is an overall objective towards which progress is being made.

2. Incorporate genomic regions into an elite stiff-stalk and an elite Lancaster line.

The Lancaster line was dropped from the conversion process because it was unexpectedly genetically heterogenous and the amount of work required to convert and evaluate NIL's in both lines was more than the funding available would support. The conversion of the stiff-stalk line will be at the BC4 at the end of the project.

3. Develop and evaluate BC^1S_1 lines with chromosome regions containing desired QTL.

$BC_1 S_2$ near isogenic lines have been developed and are in the process of being evaluated. Based on one year's data from lines per se, significant effects for starch, protein, and oil were identified in both the Syn0 and Syn4. Testcrosses are being evaluated to determine whether these effects are also expressed in hybrid combinations and whether they are also associated with yield.

4. Develop BC lines from both the Syn) and Syn 4 to measure random-mating effects.

Lines have been developed. One year's evaluation has been obtained and probes marking significant effects for starch, protein, and oil were identified in both the Syn0 and Syn4 from both per se and testcross data.

5. Develop and evaluate near-isogenic selfed lines from the Syn) and Syn 4.

Lines have been developed. One year's evaluation has been obtained and probes marking significant effects for starch, protein, and oil were identified in both the Syn0 and Syn4 from both per se and testcross data.

6. Begin development of commercially acceptable high starch lines.

The development of BC_4 lines carrying markers associated with high starch is a step toward meeting this objective.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

During this reporting period, the project moved us closer to development of high starch corn inbreds that can be used to produce higher starch corn hybrids. This was done by verifying, in selfed and backcross lines, that genetic markers previously identified as being associated with high starch in segregating generations were also associated with high starch in generations more similar to what will be used for production of commercial corn hybrids.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Since the last reporting period, we have obtained marker data on near-isogenic S_4 and BC_1 S_2 lines from the Illinois High Protein x Illinois Low Protein cross. In addition, we have obtained performance data (starch, protein, and oil) from two locations of the S_4 and BC_1 S_2 lines per se and of testcrosses of the S_4 lines. Yield data were also obtained from the testcrosses. From these results, we have tentatively confirmed the existence of QTL for high starch in regions on chromosomes 1, 2, 3, 8, and 10. The confirmation is tentative because we have only one year's data. By the time of the completion of the project, we will have a second year's data and will have additional marker data from the regions of interest. We will also have performance data from testcrosses of the backcross derived lines. With this information, we will be able to better define the importance of these regions to development of high starch corn inbreds and hybrids.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

None.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

None

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

It is too early in the life of the project to have commercial accomplishments to report.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

The educational accomplishments include the continued education of a post-doctoral associate and the training of a number of undergraduate workers in molecular genetic techniques. Hopefully some of these undergraduates will use this experience as a stepping stone to a career in molecular genetics.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

None

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

None

RECEIVED MAY 15 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Edison R. Fowlks, Ph.D.	
University:	Hampton University	
Agreement Number:	OR22072-40	
Project Title:	Automated fluorescent genomic fingerprinting (Auto-Rep-PCR) of bacteria	
Reporting Period and Report Type:	**From:** **1/15/96** **To:** **12/31/99**	**Check one:** [ ] Interim Report [X] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

I. To establish rep-PCR in the Hampton University biotechnology laboratory and further develop protocols for non-fluorescent based analysis of complex PCR-based fingerprint patterns and to participate in the agarose systems database design.

II. To use automated rep-PCR to analyze the genetic diversity of a world-wide collection of Xanthomonas that infect soybean, strawberry, zinnia, and geranium, and evaluate/develop and/or strain identification and classification.

III. To use the establishment of rep-PCR fingerprinting to enhance our laboratory educational capacities, by making it an integral part of a Biotechnology Training Program.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

The "Automated Fluorescent Genomic Fingerprinting of Bacteria" collaborative project between scientists at Michigan State University, and North Carolina State University has enabled the establishment of rep-PCR in the biotechnology laboratory at Hampton University. This new technology has had a tremendous impact on the research and educational components of our program at the university. The facilitation of the development of a repPCR workstation allows faculty and students to perform research in the identification of closely related bacterial strains related to agricultural foodstuff, renewal resources and bioremediation, as well as other areas of research. Three graduate students are using repPCR technology for the characterization of bacteria used in their research. As a result of this project a new research project is being developed in our laboratory.

Through our educational component over 700 students received training in PCR technology. The establishment of rep-PCR fingerprinting and regular PCR have become integral parts of our undergraduate and graduate training in molecular biology and biotechnology. In the classroom as well as in the research laboratories students learn the basic principles, methods and diverse applications of PCR technology. Now faculty routinely use PCR in their research and teaching as we expand our curriculum to include Biotechnology I and II as a prelude to an undergraduate Biotechnology curriculum, with a national advisory committee consisting of scientists and professors from other universities and from the biotechnology /genomics/bioinformatic industry.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Objective I: The first objective of this project was to establish rep-PCR in the Hampton University biotechnology laboratory and further develop protocols for non-fluorescent based analysis of complex PCR-based fingerprint patterns and to participate in the agarose systems database design.

Through the collaborative efforts of Dr. Frans de Bruijn at Michigan State University and Dr. Frank J. Louws at North Carolina State University, we have established rep-PCR in the biotechnology laboratory at Hampton University. Both of these scientists contributed immensely in the transfer of their technology to our laboratory, and closely provided constant support in the early beginning of the project, including visiting Hampton University and making available to us protocols for non-fluorescent-based analysis of complex PCR-based fingerprint patterns and agarose systems database design.

Objective II: The second objective of this project was to use automated rep-PCR to analyze the genetic diversity of a world-wide collection of *Xanthomonas* strains that infect soybean, strawberry, zinnia, and geranium to evaluate/develop and/or strain identification and classification.

The primary focus of our laboratory has been on *Xanthomonas* strains of zinnias and collards, which are the research focus of two graduate students who are in the final stage of their masters thesis. This geographical area is a rich source of *Xanthomonas* strains of zinnia and collards. In addition to characterizing them for the worldwide database, these strains will continue to be collected from the surrounding areas as a part of a longitudinal study. By using rep-PCR to study genetic diversity during each growing season, we may be able to understand the mechanism by which this diversity arises, and thus develop a tool for predicting its occurrence in a particular season.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

No publications and presentations have resulted from this project at this time. At least one publication should result when all of the research is completed.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

No technology transfer resulted from the Project during the reporting period

.Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

No commercial accomplishments resulted from the Project during the reporting period

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Enhancement of laboratory educational capacities

The establishment of rep-PCR fingerprinting and regular PCR have become integral parts of our undergraduate and graduate training in molecular biology and biotechnology. In the classroom as well as in the research laboratories students learn the basic principles, methods and diverse applications of PCR technology. These educational accomplishments were facilitated with the modifications in the genetics laboratory by putting in place a PCR workstation with the requisite supplies and instrumentation. Now students and faculty routinely use PCR in their teaching and research as we expand our curriculum to include Biotechnology I and II as a prelude to an undergraduate Biotechnology/Genomics/Bioinformatics curriculum, with a national advisory committee consisting of scientists and professors from other universities and from the biotechnology /genomics industry. This curriulum, of which a copy draft is included in the appendix, will be presented to the University Curriculum committee for approval during the Fall 2000 Semester.

Students' training in PCR methodology

Inclusion of PCR experiments in four courses in biology allowed a large number of students to learn about and gain experience with this technology. In Molecular Biology, Genetics, Advanced Genetics, and Developmental Genetics students wrote a paper on the relatively recent historical development of PCR and its impact on revolutionizing the approaches to biotechnological and genomic research. After they understood the history and principles involved, actual experiments were carried out individually by students or as class demonstrations. During the Project's period over 700 students have been impacted by PCR technology.

Graduate training impact

Three graduate students are using PCR in their research for the masters of science degree in biology. These students should complete and defend their research by the end of the August, 2000.

Student Internships

Two undergraduate students will work this summer in the laboratories of two members of our Biotechnology/Genomics/Bioinformatics Advisory Committee. These members are Dr. Tillman Gerngross, Darmouth; and Dr. Cheng Kao, Indiana University. Both of these students are interested in pursing the Ph.D. degree.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

No additional funding has been generated as a result of the Project. However, an application submitted to the National Institutes of Health for PCR-related research with bacteria is being revised for resubmission. Moreover, efforts are underway to find funding from the corporate sector and private foundations to expand our research and to put in place an undergraduate program in biotechnology/genomics/bioinformatics.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

There were no changes in key personnel during the reporting period.

Send completed report to :

The Consortium for Plant Biotechnology Research, Inc.
P.O. Box 20634
(Express Delivery address: 10 Sylvan Drive, Suite 21)
St. Simons Island, GA 31522
Phone: 912.638.4900 Fax: 912.638.7788

Or sent as an email attachment to: cpbr@gate.net

APPENDIX

Biotechnology, Genomics & Bioinformatics

Undergraduate Curriculum

Fall 2000

Year I

Fall	Credit Hours	Spring	Credit Hours
Math 117	3	Math 118	3
Bio 105:Molecular Biology	4	Bio 106:Molecular Biology	4
University 101	1	Eng 102: Written Communication	3
PED Activity	1	PED Activity	1
Eng 101: Written Communication	3	HEA 200: Health	1
Bio 1XX Molecular Gen Chem *	4	Bio 1XX Molecular Gen Chem*	4

Year II

Fall	Credit Hours	Spring	Credits Hours
Humanities 201	3	Humanities 202	3
Math 130	3	Bio 2XX Molecular Org. ChemII	4
Bio 2XX Molecular Org. Chem I*	4	Physics 202	4
Physics 201	4	Computer Sci I (C++)	3
Physics Lab	1	Physics Lab	1
Social Science Elective	3	Social Science Elective	3

Year III

Fall	Credit Hours	Spring	Credit Hours
Computer Sci II(C++)	3	Computer Sci III Data Structure	3
Bio 305: Genetics	3	Social Science Elective	3
Speech 103	3	Bio 3XX: Biotechnology I	4
Bio 3XX Bioinformatics I	4	Bio 3XX Bioinformatics II	4
Bio 3XX Molecular Biochem I*	4	Bio 3XX Molecular Biochem II*	4

Year IV

Fall	Credit Hours	Spring	Credit Hours
Bio 4XX Biotechnology II	4	Bio 4XX Biotechnology III	4
Bio 4XX Genomics I	4	Bio 4XX Genomics II	4
Bio 4XX Bioinformatics III	4	History 107	3
History 106 or 105	3	Bio 4XXX Genomics III	3
Free Elective	3	Free Elective	3

*Team Taught (For Discussion ONLY—not for distribution) Edison R. Fowlks

RECEIVED SEP 1 1 2000

Scientific Progress Report		
Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.		
Principal Investigator:	Stanton B. Gelvin	
University:	Purdue University	
Agreement Number:	OR22072-90	
Project Title:	Plant Genes Involved in T-DNA Integration and Radiation Sensitivity	
Reporting Period and Report Type:	**From: 12/1/98** **To: 8/31/00** **(Interim report submitted 2/29/00)**	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

OBJECTIVE 1: Screen the Feldmann and DuPont collections of T-DNA mutagenized *Arabidopsis* plants for mutants that are resistant to *Agrobacterium* transformation.

We have screened more than 3250 individual plants. Of these, 301 were putative *rat* mutants. Approximately 197 of these putative mutants have been or are now in the process of re-screening. We have confirmed at least 16 *rat* mutants from these.

OBJECTIVE 2: Screen the Feldmann and DuPont collections of T-DNA mutagenized *Arabidopsis* plants for mutants that are radiation hypersensitive.

Several reviewers pointed out that our initial *rat* mutant screening was a huge amount of work, and that we should not pursue this screen. We have therefore abandoned this line of experimentation.

OBJECTIVE 3: Perform genetic analyses on the *rat* mutants.

This is an ongoing line of experimentation. Of the confirmed new *rat* mutants, 12 are homozygous. Of four heterozygous lines, three are segregating (for the kanamycin-resistance marker on the T-DNA) 3:1, and one is segregating 7:1.

OBJECTIVE 4: Perform molecular analyses on the mutants.

We are in the process of isolating T-DNA/plant DNA junctions for several of the known new *rat* mutants. Using TAIL-PCR, we have recovered the junctions from 5 of

the mutants. DNA sequence analysis indicates that in one line, T-DNA inserted into an importin-_ gene. In another line, T-DNA inserted into a DEAD-box RNA helicase gene. Analysis is continuing on the other lines.

OBJECTIVE 5: Perform genetic complementation of the *rat* mutants with wild-type DNA.

Because we do not have the wild-type genes yet, we have not done this analysis yet. However, we have started complementation analysis with the importin-_ gene.

OBJECTIVE 6: Assay mutants for increased levels of homologous recombination.

We have not begun these analyses yet.

Layman's Summary

We are interested in the mechanism of T-DNA transfer from the bacterium *Agrobacterium tumefaciens* to plant cells. *Agrobacterium*-mediated genetic transformation of plants is the most popular method to introduce new genes into plants. However, many agronomically important crop species remain recalcitrant to *Agrobacterium* transformation. We therefore are investigating plant genes necessary for the transformation process with the hope that identification of these genes will allow us to improve the transformation process. To identify these plant genes, we have screened a mutagenized library of *Arabidopsis* for plants that cannot be genetically transformed by *Agrobacterium tumefaciens*. We call such mutant plants "*rat*" mutants (resistant to *Agrobacterium* transformation). We have screened more than 9000 mutagenized plants and have identified more than 50 *rat* mutants.

We have used genetic and molecular analyses to determine the identity of the mutant genes, and are using cell biology and biochemical analyses to investigate the functions of these genes in the transformation process. We previously identified a histone H2A gene that is important for T-DNA integration into plant chromosomes. We have now identified several other genes, including a _-importin gene and a RNA helicase gene, that are important for transformation. We shall continue to investigate the nature of the genes disrupted in the *rat* mutants, and to use this knowledge to improve *Agrobacterium*-mediated plant transformation.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

See below.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Congress on Molecular Plant-Microbe Interactions, Amsterdam, The Netherlands July 24-30, 1999. Invited lecture: "Plant genes involved in *Agrobacterium*-mediated transformation".

20th Annual Crown Gall meeting, Houston, TX, November 5-7, 1999. Lectures and posters: "Identification of plant genes involved in Agrobacterium-mediated transformation", "An *Arabidopsis* histone H2A gene is involved in T-DNA integration".

Genetic Engineering/Physiological and Molecular Plant Biology seminar, University of Illinois, Feb. 16, 2000: "Plant genes involved in *Agrobacterium*-mediated plant transformation".

Keystone Conference on Signals and Signal Perception in Biotic Interactions in Plants, Taos NM Feb. 22-26, 2000. Presented poster "*Arabidopsis* mutants that are resistant to *Agrobacterium* infection".

University of Illinois, Chicago. April 17, 2000. Presented an invited lecture "Host genes involved in crown gall tumorigenesis by *Agrobacterium tumefaciens*.

Purdue University, Department of Horticulture. April 20, 2000. Presented an invited lecture "*Arabidopsis* mutants that are resistant to *Agrobacterium* transformation (I smell a *rat*)."

Sixth International Congress of the International Society of Plant Molecular Biology, Quebec City, Canada. June 16-23, 2000. Presented a Keynote address "Plant genes involved in *Agrobacterium*-mediated transformation."

National Science Foundation annual meeting of Plant Genome awardees. Alexandria, VA. Sept. 21-23, 2000. Presented a poster "Plant genes involved in *Agrobacterium*-mediated transformation.

University of Pennsylvania, Department of Biology. Sept. 26, 2000. Presented a invited lecture "Plant genes involved in *Agrobacterium*-mediated transformation."

Thomas Jefferson Medical School, Philadelphia. Sept. 27, 2000. Presented a invited lecture "Plant genes involved in *Agrobacterium*-mediated transformation."

North Carolina State University. October 11, 2000. Presented a invited lecture "Plant genes involved in *Agrobacterium*-mediated transformation."

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

None to date.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None to date.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Contributed to the training of three postdoctoral research fellows.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

NSF: Identification of plant genes involved in *Agrobacterium*-mediated transformation. (To Stanton B. Gelvin, Vitaly Citovsky, and Barbara Hohn). 9/1/99-8/31/04. $4,440,751

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Postdoctoral research fellows involved in the project:

1. Jaime Humara (Spanish government fellowship supplemented from this grant)
2. C.T. Ranjith Kumar (left the laboratory in May, 2000).
3. Jyothi Rajagopal
4. Ho Chul Yi (Korean KOSEF fellowship supplemented from this grant)

SCIENTIFIC ACCOMPLISHMENTS

1. SCREENING FOR *RAT* MUTANTS (Jaime Humara and Jyothi Rajagopal)

As an ongoing effort in our laboratory, we screened more than 3250 T-DNA mutagenized *Arabidopsis* plants from the Feldmann library (ecotype Ws) for resistance to *Agrobacterium* transformation (*rat* mutants). During this past year we screened an additional approximately 3000 mutagenized plants using personnel hired on other grants, and previously screened approximately 3000 plants. Our overall screening efforts have now come close to saturating the Feldmann collection (6400 total mutagenized lines; we have screened in all approximately 9000 individual plants). We had originally proposed to screen the DuPont collection of T-DNA mutagenized *Arabidopsis* plants. However, we have found that this collection has very poor germination (we have twice received this collection), and we have therefore abandoned efforts to use the DuPont collection. We have now acquired the INRA/Versailles collection of T-DNA mutagenized plants (also in the ecotype Ws) and have begun screening this collection for *rat* mutants. In addition, with the help of Dr. Ray Bressan of Purdue University, we shall soon complete construction of a new 50,000-60,000 member collection of T-DNA mutagenized *Arabidopsis* plants (ecotype Ws). We have just recently started screening this collection. (We had started to acquire the Amasino/Sussman collection of T-DNA mutagenized plants, but abandoned efforts to use this collection when we learned that the T-DNA used to mutagenize these plants contains an active *AP3* homeotic gene).

Primary screening of these lines involves germination and sterile growth of the plants in agar medium in baby food jars, sectioning of the roots into 3-5 mm segments (the shoots are placed back into medium to re-grow roots), infection of the roots for two days by an oncogenic *Agrobacterium* strain (A208), transfer of the roots to agar medium containing timentin (to kill the *Agrobacterium*) but lacking phytohormones (for selection of tumors), and incubation for 4-6 weeks. Of the approximately 3250 plants screened, we identified 301 putative *rat* mutants. Seeds of these putative mutant plants were collected and the lines are currently being re-tested for the *rat* phenotype. This secondary screening involves infection of the roots with *A. tumefaciens* A208 (to screen for tumorigenesis) and also infection with a non-oncogenic *Agrobacterium* strain (At872) containing a binary vector with a *bar* gene. The infected roots are subsequently screened for sensitivity or resistance to the herbicide phosphinothricin on callus inducing medium (CIM). We have completed secondary screening of approximately 197 putative *rat* mutants and have confirmed 16 *rat* mutants. Work during the remaining time of this grant (we shall be asking for a no-cost extension) will finish the secondary screening of these lines (as well as initiate screening of other mutagenized lines).

We have initiated genetic analysis of some of the confirmed *rat* mutants. twelve of the lines are homozygous (they do not segregate kanamycin-sensitive plants), and of three heterozygous lines tested, three segregate 3:1 (kanamycin^r:kanamycin^s), indicating a single segregating T-DNA locus, and one segregates 7:1 (an unusual segregation ratio). We shall soon initiate a co-segregation analysis of the heterozygous plants (looking for co-segregation of kanamycin-resitance with the *rat* phenotype). In addition, we shall cross the homozygous plants with a wild-type Ws plant and perform a similar co-segregation analysis of the resulting F2 progeny. Finally, we shall test the F1 progeny resulting from these crosses for dominance/recessivity of the *rat* phenotype.

We are also completing an analysis of the confirmed *rat* mutants to determine whether transformation is blocked early or late in the process. To do this, we infected cut root segments with an *Agrobacterium* strain containing a binary vector (pBISN1) with a

gusA-intron gene and assayed transient GUS activity 4-6 days after infection. Low GUS activity indicates an early transformation block-point, whereas high GUS activity indicates a late stage (probably integration) block-point. Our results to date indicate that three of the *rat* mutants are blocked early in the transformation process, whereas one is blocked late.

We have initiated a molecular analysis of some of the *rat* mutants. We have used TAIL-PCR to generate T-DNA/plant DNA junctions, and sequenced the junctions. Two of the mutants are very interesting: One contains an insertion in an importin-_ (transportin-like) gene. This type of protein is used to shuttle proteins to the nucleus of cells in an _-importin-independent fashion. We are initiating analysis of this mutant to determine whether it is defective in nuclear import of VirD2 or VirE2 proteins. A second mutant contains a T-DNA insertion in a "DEAD-box" RNA helicase. This mutant may be defective in T-DNA integration. We shall conduct further analysis on this mutant.

2. INTERACTION OF HISTONE H2A PROTEINS WITH VIRD2 PROTEIN (C.T. Ranjith Kumar)

We had previously identified a mutant, *rat5*, that contains a T-DNA insertion into the 3' untranslated region of a histone H2A gene. We recently published that this mutant is defective in T-DNA integration (Mysore, K.S., Nam, J., and Gelvin, S.B. 2000. An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration PNAS 2000 97: 948-953). Because VirD2 protein is covalently linked to the 5' end of the T-strand that enters the plant nucleus, and because we had previously published that VirD2 is involved in T-DNA integration (Mysore, K.S., Bassuner, B., Deng, X-b., Darbinian, N.S., Motchoulski, A., Ream, W., and Gelvin, S.B. 1998. Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. Mol. Plant-Microbe Interact. 11:668-683), we reasoned that VirD2 may interact with any "integration complex" of plant proteins, possibly including histone H2A.

We therefore tested VirD2, both as a bait and a prey protein, in a yeast two-hybrid system, using histone H2A (encoded by *RAT5*) as the corresponding prey or bait protein. Although we initially reported that they interacted, further analysis with additional controls indicated that this interaction was weak and probably not biologically important.

3. FUNCTIONAL EQUIVALENCE OF HISTONE H2A PROTEINS IN T-DNA INTEGRATION (Ho Chul Yi)

We had previously suggested that histone H2A genes make up a six-member multigene family in *Arabidopsis*, and that mutation of one of these genes (G1) in the *rat5* mutant results in the *rat* phenotype. (We now know that there are at least seven histone H2A gene family members). In the *rat5* mutant, the other histone H2A genes are wild-type, yet the *rat* mutant phenotype is expressed. It could thus be argued that the different histone H2A genes are not functionally redundant. However, these genes are controlled by different promoters, and the possibility exists that the proteins may be functionally equivalent if expressed in the appropriate manner.

We therefore placed 5 of the 6 histone H2A cDNAs behind a strong CaMV 35S promoter and individually introduced each histone H2A gene into the *rat5* mutant by a

flower vacuum infiltration method. Although many *Arabidopsis* ecotypes and *rat* mutants cannot be transformed by root inoculation, they are efficiently transformed by flower vacuum infiltration (Mysore, K.S., Kumar, C.T.R., and Gelvin, S.B. 2000. *Arabidopsis* ecotypes and mutants that are recalcitrant to *Agrobacterium* root transformation are susceptible to germ-line transformation. Plant J. 21:9-16). As shown in Table I below, transformation of root segments of the wild-type plant to the phenotype tumorigenesis is very efficient, whereas transformation of the *rat5* mutant to the same phenotype is negligible. When the wild-type *RAT5* histone H2A cDNA is expressed in *rat5* mutant plants, the transformation phenotype is restored. Complementation of the *rat5* plant can also be achieved by expression of cDNAs (under the control of the CaMV 35S promoter) from the other histone H2A genes in the *rat5* mutant plant. Thus, when overexpressed from a CaMV 35S promoter, all histone H2A cDNAs tested appear to be functionally equivalent with regard to transformation and T-DNA integration. Thus, the lack of histone H2A functional redundancy in the *rat5* mutant most likely results from lack of adequate expression of the other histone H2A proteins in the root tissue.

We are continuing these experiments by introducing each of the histone H2A genes (under the control of their native promoters) into the *rat5* mutant plant to determine whether when in higher copy these genes can complement the *rat* phenotype. We are also introducing each of the histone H2A genes into wild-type plants to determine whether overexpression of each of these genes results in increased transformation of infected transgenic root segments, and whether anti-sense expression of the histone H2A gene family members under CaMV 35S promoter control can reduce transformation of transgenic plants.

Table I.

Complementation of *rat5* Mutant Plants with Various Histone H2A Genes

Line	Phenotype recovery (% of Ws)
Ws	100
rat5	0
H2A1-1-1	89
H2A1-1-2	44
H2A1-5-3	69
H2A1-8-1	37
H2A2-1-6	31
H2A2-1-10	26
H2A3-2-4	94
H2A3-2-5	78
H2A3-2-8	67
H2A3-3-1	50
H2A4-1-1	45
H2A4-5-3	91
H2A4-5-5	73
H2A4-6-9	82
H2A4-6-10	82

RECEIVED APR 25 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Jean Greenberg	
University:	University of Chicago	
Agreement Number:	OR22072-91	
Project Title:	*Engineering Artificial Immunity to Plant Pathogens*	
Reporting Period and Report Type:	**From:** 10/1/98 **To:** 4/07/99	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. To identify novel secreted factors from the phytopathogen *Pseudomonas syringae.*

2. To test the feasibility of disrupting the secreted factors *in planta* through the use of artificial immunity based on selection of RNA(s) that will bind to and disrupt the action of the secreted factors.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

Plant diseases can result in significant economic losses for farmers and poor quality of food and energy crops. One of our goals is to devise a way to reduce losses due to disease by developing a disease-neutralizing technology that can be broadly applied to plants. To do this, we must first understand why diseases occur and then we can try to intervene in the disease process. Bacteria are known to cause disease in plants by producing toxins and proteins called virulence factors which the bacteria inject into the plant and disrupt their normal functions. We have investigated the properties of one of these toxic proteins that can kill plant cells with the idea that this information will help us find additional virulence proteins. Once we find all the virulence proteins, we can try our neutralizing technology. We found that our virulence protein is made of two parts: one part is used to kill plant cells and the other part is used to ensure that the protein is transported from the bacteria to the inside of plant cells where it acts. Our toxic protein is rendered harmless when the transportation part of the protein is removed. We found that we could take the killing part of our toxic protein and connect it to the transportation part of a second virulence protein and restore the lethal properties of the toxic protein. This information is very important because it means we can

use the killing part of our protein to find other virulence proteins. We plan to take the killing part of our protein and connect it to all of the bacteria's other proteins in order to discover which of the bacteria's proteins can restore the transportation of our toxic protein to the plant cells. These new proteins are likely to be important for disease, because of their transportation abilities. We are now searching for these new virulence proteins with transportation properties and we are testing a disease-neutralizing strategy.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Plant pathogens can reduce yields of important energy and food crop plants. An understanding of how pathogens modify their hosts to cause disease might result in the development of novel technologies that can prevent plant disease by interfering with pathogen virulence mechanisms. Our long term aims are to discover how pathogens modify their hosts to cause disease and to develop and/or apply new methods for preventing disease. Our short term goals for this project are: To identify novel secreted factors from the phytopathogen *P. syringae* (specific aim 1). The particular factors we are interested in are those which are secreted from the bacteria and delivered directly into host cells where they modify host functions. Our second goal is to test the feasibility of disrupting the secreted factors *in planta* through the use of artificial immunity based on the technique called SELEX (specific aim 2) which involves the iterative selection of RNA molecules with unique binding properties to find RNA(s) that will bind to and disrupt the action of the secreted factors *in planta.*

In order to achieve the first specific aim, postdoctoral fellow David Guttman decided to first characterize in detail one secreted factor called AvrRpt2 which is known to cause plant cell death when it is secreted and translocated by the type III secretion system into a plant host carrying a functional *RPS2* gene. We reasoned that if we characterized AvrRpt2, then we could exploit information about this protein to devise a genetic screen for additional secreted proteins. Such secreted proteins are likely to be important virulence factors and are thus good targets for the antimicrobial strategies we are developing in specific aim 2. David's experiments have established that this is true. David has functionally dissected the AvrRpt2 protein into two domains: an N terminal region that is necessary for protein secretion and translocation, and a C terminal region that is necessary for inducing plant cell death (the effector function). To delineate the effector function of AvrRpt2, David introduced N-terminal truncations of the protein into Arabidopsis by biolistic transformation of the truncated gene. He assayed for gene function by cotransforming in a reporter gene encoding the b-glucuronidase (GUS) protein with his AvrRpt2 truncations. Plant cells that receive a functional AvrRpt2 gene encoding the effector domain die before the GUS activity is induced, so a reduction in GUS activity indicates that AvrRpt2 (or a derivative) is triggering cell death and is therefore functional. David has also established that the N terminus from the heterologous AvrRpm1 protein can direct secretion and translocation of the C terminal effector region of AvrRpt2. This finding was critical for our subsequent experiments (see below) which involve using the C terminus of AvrRpt2 as a reporter protein to tag "secretion/translocation" sequences of novel genes that encode secreted factors. David established a rapid method for introducing AvrRpt2 or AvrRpm1 chimeras into bacteria in single copy and showed that such genes function to induce plant cell death on the appropriate host. This is important because our future experiments rely on the single copy expression of AvrRpt2 chimeric proteins to find novel targets of the Type III secretion system. Finally, David found that by fusing the N terminus of AvrRpt2 to the GAL4/VP16 transcription factor he could activate transcription of a GAL4/VP16-dependent reporter (UASGAL4-GUS) in plants in a secretion-dependent manner. In the course of his experiments, David has also established that AvrRpt2 can function as a virulence factor when it is present in single copy in *P. syringae* and the plant host lacks the RPS2 gene. Interestingly, this virulence function affects early release of the bacteria onto plant cell surfaces, suggesting that the AvrRpt2 gene functions to aid the dissemination of bacteria during

infections. We will be writing David's findings in the next one to two months to submit them for publication. These studies have recently be submitted for publication to Molecular Microbiology.

I previously described a molecular genetic screen for novel targets (effector proteins) of the type III secretion apparatus encoded by the *hrp/hrc* genes. The identification of novel effectors is a prerequisite for the experiments proposed in the application. In the last four weeks, we have used a transposon containing the C terminal cell death inducing domain of AvrRpt2 lacking a promoter and secretion/translocation signal to tag two new genes encoding secreted proteins in *PsmES4326*. We screened 10,560 bacterial isolates with independent transposition events using a pooling strategy. We knew from mock-up experiments that we could pool bacteria containing full length AvrRpt2 with bacteria lacking the protein in a ratio of 1 in 8 bacteria and still see AvrRpt2 effector function (rapid cell death- the hypersensitive response). Therefore, we pooled 8 independent transposon mutants and inoculated these pools of 8 into plants. If a pool induced rapid cell death, it identified a potential gene encoding a target of type III secretion. Positive pools were deconvoluted and inoculated individually into plants. Using this strategy, we verified that seven of the isolates contain fusions to the truncated AvrRpt2 that now allow the translocation of AvrRpt2 into plants and cause it to induce cell death. These contain insertions in the *hop1*, *hop2, hop3, hop4* genes (*hrp outer proteins*). Western blot analysis on *hop1* bacteria (other *hop* genes have not yet been tested) showed a fusion protein detected by the anti-AvrRpt2 antiserum. Sequence analysis of the 5' end of *hop1* showed that HOP1 is a novel gene. We are currently verifying the hrp-dependence of the secretion of the hop1-avrRpt2 fusion by cloning the fusion gene and integrating it at the recA locus in heterologous hrp+ and hrp- *P. syringae* strains (we previously found that this type of ectopic expression works very well). If the fusion only triggers cell death when expressed in the hrp+ strain, it will indicate that cell death is due to secretion of the chimeric protein. We are continuing the screen and we are confident that we can easily screen another 75,000 isolates within a three month period to saturate the screen in PsmES4326.

Our second goal is to test the feasibility of using SELEX to engineer artificial immunity in plants. SELEX is a procedure in which pools of random RNA sequence (called aptamers) are used in iterative binding experiments with target molecules. An aptamer that specifically binds to a target molecule could, in principal, inhibit the action of that molecule. Our goal is to express aptamers with the ability to bind bacterially secreted proteins in plants with the hope of inhibiting bacterial virulence. In practice, a SELEX experiment involves immobilizing the target molecule on a dish or a column and mixing pools of RNA with the target. After washing away the non-specific RNAs, anything that is tightly bound is eluted, converted into DNA, amplified using polymerase chain reaction and used to synthesize new RNAs so that the entire experiment can be repeated until only a few RNAs with high specificity are selected. Our strategy is to use Avr proteins as targets for SELEX and then determine if the RNAs selected have in vivo activity in plants. Postdoctoral fellow Suresh Gopalan joined the project in January 1999 and succeeded in partially solubilizing the AvrRpt2 protein in preparation for the SELEX experiment. Unfortunately, Dr. Gopalan became ill in March and had to resign his position. Dr. Sharon Hall was then recruited to continue the experiments initiated in the lab by Dr. Gopalan, but she was unable to start until August 15, 1999. In the time between Dr. Gopalan's leaving and the arrival of Dr. Hall, a part time technician was employed, Ms Priya Koduri, who successfully cloned and overexpressed an additional Avr protein, called AvrRpm1, as well as truncated AvrRpt2 derivatives so that additional targets for SELEX are available for Dr. Hall's experiments.

Dr. Hall has purified over-expressed AvrRpt2, which is visualized as a single band on SDS-PAGE gels. The entire avrRPT2 cDNA was cloned, in translational reading frame, from the ATG start codon, into pQE-60 expression vector (Quaigen). pQE60::avrRPT2, as well as vector only control, was then transformed into *E.coli* strain M15[pREP4]. Single colonies were used to inoculate overnight 10 mL cultures (LB, containing both ampicillin (100μg/mL) and kanamycin (25μg/mL), 37°C). 100 mL media was then inoculated with 5 mL of overnight culture and grown at 37°C, with rigorous shaking (300rpm) until an OD_{600} of 0.5-0.7 was reached. At this point expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactoside (IPTG). Culture were then grown for an additional 4 h at 37°C. Cells were harvested by centrifugation at 10,000g for 20

min. and resuspended in 20 mM Tris-HCl (pH 7.5), 30 mM NaCl. Cells were then lysed by freezing at -70°C and thawing in cold water. Cell lysates were sonicated using a microtip with the power level set at between 4-5, at 40-50% duty for 15-20 bursts, keeping on ice at all time. The lysate was then centrifuged at 10,000g at 4°C for 20 min. and QE60::avrRPT2 recovered was recovered as insoluble inclusion bodies.

Solubilzation and refolding of AVRRPT2 was achieved using the protocol of Eijima et al., (1999). Pelleted inclusion bodies were re-suspended in 10 mM EDTA , pH 5.5 and stored at -70°C until use. They were then thawed on ice and centrifuged at 8,000g for 15 min. The pellet of inclusion bodies was solubilized in 10 mL of 6M GdnHCL, pH 5.5 and allowed to stand for 2 h at 22°C. Reduced and oxidized glutathione were added into the solubilized AVRRPT2 at final concentrations of 0.01 mM and 0.002 mM respectively. This solution was then adjusted to pH 8.5 with 10 mM Tris-HCL and incubated for 16 h at 22°C. Refolding of the protein was completed by direct application to a size exclusion column (Sephadex G-25, 10 cm X 1 cm), equilibrated with 20 mM Tris-HCl, pH 7.5, 30 mM NaCl, using FPLC. Protein eluting from the column was detected measuring absorbance at 280 nm. Positive fractions (5 mL) were mixed with 1 mL of 50% Ni-NTA agarose, with gentle shaking at 4°C, for 1h. AVRRPT2 was then batch purified from the Ni-NTA column by elution with 250 mM imidazole in 20 mM Tris-HCl, pH 7.5, 30 mM NaCl. avrRPT2 eluted as a single band circa 29 KDa on a12% SDS-PAGE gel.

In collaboration with Andrew Ellington's lab (University of Texas, Austin), Dr. Hall synthesized a pool of random RNA aptamers. This "N30" pool consists of aptamers that have 30 random nucleotides, flanked by the T7 polymerase promoter and primer sequences. Using the solubilized protein Sharon is now using the iterative rounds of selection that make up the SELEX process, to isolate aptamers specific to AvrRpt2.

Concurrent to the work with the avrRPT2-specific aptamers, we are also investigating the possibilities of actually expressing aptamers in plants. The published kanamycin-specific aptamer sequence (Werstuck and Green, 1998), has been cloned into the bluescript vector and will be sub-cloned into binary vectors (generous gift of J. Mylne, Mylne and Botella, 1998) for transformation into Arabidopsis using agrobacterium. The binary vectors contain the bar gene for resistance to the herbicide basta, which will be used to select for transformants. Expression of the kanamycin aptamer in transformed plants will result in a kanamycin resistant phenotype.

We have also have a vector (pSP650-1, generous gift of William Folk; Bourque and Folk, 1992)) which contains the tRNA from soybean, with transcription under the control of the RNA polymerase III promoter. The tRNA methionine initiator (tRNAmeti) gene encodes the initiator tRNA utilized for virtually all cytoplasmic protein synthesis in plants and, as such, is expressed in every plant tissue (Sharp et al., 1985). The advantage of using this construct is that the fusion of a tRNA gene sequence to the aptamer sequence should provide RNA secondary structure at one end of the aptamer which may contribute to the overall stability of the RNA molecule. The kanamycin-specific aptamer in bluescript will also be sub-cloned into pSP650 and transformed into Arabidopsis using biolistic transformation. Again, the appearance of a kanamycin-resistance phenotype will indicate that aptamers can be expressed in plants and that they can impart new characteristics to plants. These will be the first examples of aptamer expression in plant tissues.

References

Bourque JE and Folk WR (1992). Suppression of gene expression in plant cells utilizing antisense sequences transcribed by RNA polymerase III. Plant Molecular Biology 19 : 641-647.

Ejima D, Watanabe M, Sato Y, Date M, Yamada N, Takahara Y (1999). High yield refolding and purification process for recombinant human interleukin-6 expressed in *Escherichia. coli.* Biotechnology and Bioengineering **62**: 301-310.

Mylne J and Botella JR (1998). Binary Vectors for sense and antisense expression of Arabidopsis ESTs. Plant Molecular Biology Reporter 16 : 257-262.

Sharp SJ, Schaack J, Cooley L, Burke DJ, Soll D (1985). Structure and transcription of eukaryotic tRNA genes. CRC Crit. Rev. Biochem 19 : 107-144.

Werstuck G and Green MR (1998). Controlling gene expression in living cells through small molecule-RNA interactions. Science 282 : 296-298

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Seminar: Molecular Analysis of a Type III Secreted Protein from a Phytopathogen. Iowa State University Graduate Student Luncheon Seminar. April, 1999.

Poster entitled: Characterization of avrRpt2- a Type III Secreted Virulence Protein from the Phytopathogen, Pseudomonas syringae. This poster was presented by Dr. David Guttman at the Microbial Population Biology Gordon Conference July, 1999 and by Dr. Jean Greenberg at the Molecular Mechanisms of Microbial Adhesion Gordon Conference August, 1999.

Guttman, D. S. and Greenberg, J. T. (2000) The *Pseudomonas syringae* AvrRpt2 type III effector is directly translocated into Arabidopsis and aids bacteria growth and transmission. submitted to Molecular Microbiology

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

No technology transfer has been attempted at this time. An invention disclosure will be made within the next months.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

No commercial accomplishments have been attempted at this time.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Postdoctoral fellow David Guttman has worked on Specific Aim 1 during this reporting period. Dr. Guttman recently left the lab for a permanent faculty position at the University of Toronto.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

Dr. David Guttman secured a postdoctoral fellowship from the NIH concurrently with this CPBR award.

A proposal entitled: "ANALYSIS OF TYPE III SECRETED EFFECTOR PROTEINS FROM PHYTOPATHOGENIC BACTERIA" has been submitted (January 10, 2000) to the National Science Foundation.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Dr. Suresh Gopalan was recruited to work on Specific Aim 2 of the project. He worked on the project from January until April 1999. Due to an illness, Dr. Gopalan resigned his position. Postdoctoral fellow Dr. Sharon Hall was hired and began working on Specific Aim 2 of the project.

Boris Vinatzer joined the project to screen for new secreted proteins from P. syringae as a visiting PhD student in November of 1999. He will continue as a postdoctoral fellow starting April 2000.

RECEIVED MAY 2 9 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Robert Haselkorn, P.D.	
University:	The University of Chicago	
Agreement Number:	OR22072-92	
Project Title:	Increasing the energy content of plants: molecular genetics of acetyl-CoA carboxylase	
Reporting Period and Report Type:	**From:** 10/1/99 **To:** 3/31/00	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. Analysis of ACCases and their genes in switchgrass.

2. Assessment of genetic variation at a molecular level to understand switchgrass phylogeny, to aid breeding programs and to further genetic studies.

3. Development of a switchgrass transformation system.

4. Development of a new selectible marker for transformation of Graminae based on herbicide resistant ACCase and engineering herbicide resistance in switchgrass to reduce weed competition.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

Switchgrass is a native grass in the US and Canada with a potential to become a source of renewable biomass for fuel production. Development of elite cultivars by classical breeding programs and development of proper agrotech practices needs to be complemented by genetic engineering of switchgrass. We have made good progress towards establishing a plant regeneration system and assessing genetic variation within and among different switchgrass populations. This knowledge will be necessary to enhance breeding programs, explore genetic diversity of germplasm and to provide the means for metabolic engineering to improve the energy content of switchgrass.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

An overview of major accomplishments in the four areas

1. Analysis of ACCases and their genes in switchgrass - **genetic analysis was geared towards the phylogenetic analysis, sensitivity to ACCase targeting herbicides (Topik, sethoxydim) has been confirmed.**

2. Assessment of genetic variation at a molecular level to understand switchgrass phylogeny, to aid breeding programs and to further genetic studies - **phylogenetic analysis of six important switchgrass cultivars has been completed and incorporated into a more general analysis of the grass family.**

3. Development of a switchgrass transformation system - **an efficient system to regenerate switchgrass plants from callus obtained from different plant organs has been established, transient expression of the GUS reporter gene has been achieved in different calus types, attempts to stably transform switchgrass plants continues.**

4. Development of a new selectible marker for transformation of Graminae based on herbicide resistant ACCase and engineering herbicide resistance in switchgrass to reduce weed competition - **first set of constructs have been made to test the idea.**

Key results

Phylogenetic analysis of selected switchgrass cultivars.

Results of phylogenetic analysis of selected tetraploid and hexaploid switchgrass cultivars are presented in Fig. 1. The number of nucleotide substitutions suggests that switchgrass genomes as well as different cultivars included in our study and representing both ploidy levels and both major ecotypes (lowland and highland) diverged very recently. If the molecular clock calibrated for the *Triticae* tribe applies here, the earliest divergence time of the taxons shown on the phylogenetic tree is approximately 1 MYA (Million Years Ago). The phylogenetic analysis of Fig. 1 provides a first view of the relationships at the gene sequence level. The plastid ACCase gene is well suited for this analysis. First, our results suggest that it is present in only one copy per diploid chromosome in grasses including switchgrass, making establishment of orthology much more reliable. Second, we have already obtained significant amount of information on phylogenetic relationships within the *Poacae* family based on the plastid ACCase gene sequence and our conclusions have been reinforced by similar sequence analysis of other genes (unpublished results). We have a good tool to study the flow of genetic information between switchgrass species. The results will be described in detail in an upcoming publication (manuscripts in preparation):

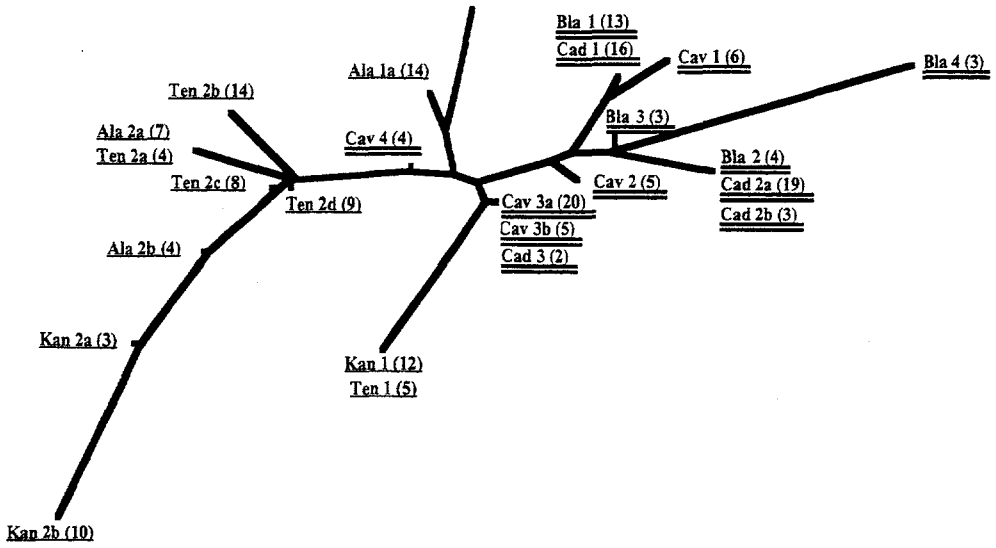


Fig. 1. Phylogenetic analysis of switchgrass (*P. virgatum*) cultivars based on a partial sequence of the plastid ACCase gene (*Acc-1*).

The taxon names include cultivar names abbreviated as shown below (tetraploids, underlined; octaploids, doubleunderlined), sequence ID number and number of clones analyzed (in brackets).). Information on geographical origin and ecotype was compiled from published data. Morphology of cultivar Tennessee 104 is that of an upland ecotype although the site of collection was marginally upland and flow cytometry results indicated that it is a tetraploid (C.M. Taliaferro, personal communication). The total length of the alignment is 1828 nucleotides in 8 exons and 7 introns. All polymorphic sites including two gaps (one two-nucleotide and one one-nucleotide) and two insertions (one four-nucleotide and one eight-nucleotide) are shown below. Dots in the alignment indicate identical nucleotides and dashes indicate gaps. Multiple alignment and phylogenetic tree (neighbor joining method) were created using Clustal X. Seeds were provided by C.M. Taliaferro, Oklahoma State University.

			Polymorphic sites	
Tetraploids (2n=4x=36)			exons	introns
Kanlow	Kan 1	(12)	TCCC	GCCTAACCAAC----TCTAG--------C
central OK	Kan 2a	(3)	CT..	....--...TT-----..C.A--------.
lowland	Kan 2b	(10)	CT..	A...--.A.TT-----..C.A--------.
Tennessee 104	Ten 1	(5)	....	..........----.....--------.
southwestern TN	Ten 2a	(4)	CT..	..T......T.----..C..--------.
upland	Ten 2b	(14)	CT..	...C.....T.----..C..--------.
	Ten 2c	(8)	CT..	....--...T.----..C..--------.
	Ten 2d	(9)	CT..	.........T.----..C..--------.
Alamo	Ala 1a	(14)	.T..	.........T.----.TC..--------.
southern TX	Ala 1b	(3)	.T..	.T.........----.TC..--------.
lowland	Ala 2a	(7)	CT..	..T......T.----..C..--------.
	Ala 2b	(4)	CT..	.........T.----..C.A--------.
Octaploids (2n=8x=72)				
Blackwell	Bla 1	(13)	.T..	.........T.----..CG.--------T
northern OK	Bla 2	(4)	.TT.	.........T.----..C..--------T
upland	Bla 3	(3)	.T..	.........T.----..C..--------T
	Bla 4	(3)	.T.T	......T..T.----G.C..--------T
Cave-in-Rock	Cav 1	(6)	.T..	.........-..----..CG.--------T
southern IL	Cav 2	(5)	.T..	.........-..----..C..--------T
upland	Cav 3a	(20)	.T..	.........-..----..C..ACTGCAGA.
	Cav 3b	(5)	.T..	.........-..----..C..--------.
	Cav 4	(4)	.T..	.........-T.----..C..--------.
Caddo	Cad 1	(16)	.T..	.........T.----..CG.--------T
central OK	Cad 2a	(19)	.TT.	.........T.----..C..--------T
upland	Cad 2b	(3)	.TT.	.........T.CCAC..C..--------T
	Cad 3	(2)	.T..	..........----..C..--------.

Our preliminary results of switchgrass tissue culture, regeneration and transformation.

Thus far, most transgenic cereal plants have been generated by bombarding isolated pre-mature embryos with DNA-coated gold particles. We use this approach in our laboratory for wheat (cv. "Bob White") transformation. This experience in wheat transformation is very helpful in the proposed work with switchgrass.

i.) Plant regeneration from calli derived from mature seeds. In one approach, we used microprojectile bombardment to deliver DNA to callus derived from mature embryos of cv. "Alamo". Mature seeds were surface sterilized and placed on callus induction medium (MS salt and vitamin-supplemented with 2.5 mg/L 2,4-D and 10 mg/L BA) for 3-6 weeks. Calli derived from explants were transferred to hormone-free MS medium for plant regeneration. Plantlets of 6-10 cm height were then transferred to soil in greenhouse. Regenerated switchgrass plants growing in the greenhouse are shown in Fig. 2 . Hundreds of regenerated plants can be obtained.

Highly efficient transient transformation of calli described above was achieved by particle bombardment under conditions developed for wheat transformation -Fig. 3 a & b. The pAHC25 and pUB-AG plasmids used in these experiments contain both the GUS reporter gene (for detection with X-Gluc) and the*bar* gene as a selectable marker (for selection with phosphinothricin, Basta). The two constructs use maize ubiquitin promoter and rice actin promoter, respectively, to drive expression of GUS and *bar* genes.

ii.) Plant regeneration from calli derived from leaf base explants. In another approach, embryogenic callus was induced from switchgrass leaf base segments -Fig. 4. Our preliminary results showed that plants can be regenerated from such callus as well. Conditions for callus transformation by particle bombardment and for selection of transformants are being worked out. Again, our approach is to use existing protocols developed for other grasses as the starting point.

Sensitivity of switchgrass to herbicides targeting ACCase.

We have selected one representative compound from each of the two grass specific herbicide classes: Topik an aryloxyphenoxypropionate and sethoxydim a cyclohexanedione. The result of a two-time herbicide spraying experiment are shown in Fig, 5. Switchgrass is highly sensitive to these herbicides.

Fig. 2 Fig. 3

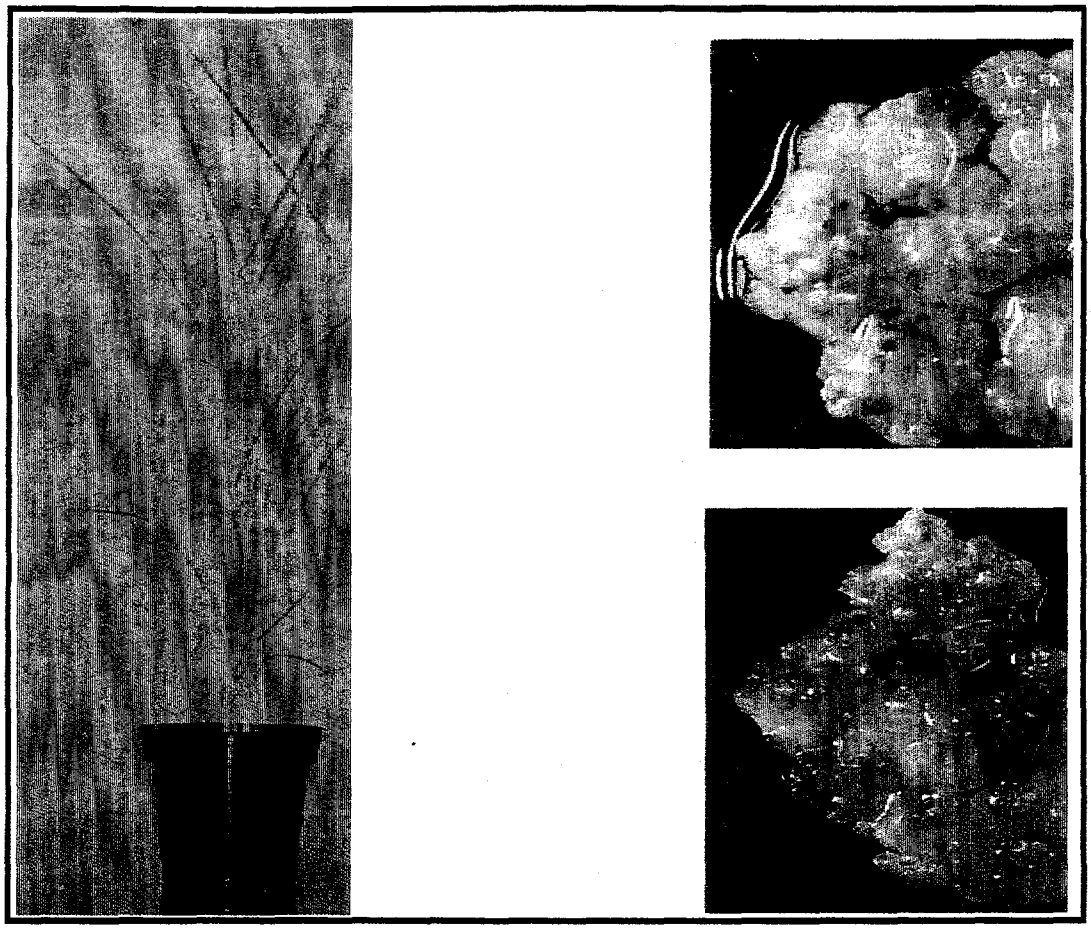


Fig.2. Switchgrass plants regenerated from callus derived from mature embryos.
Mature seeds of switchgrass cultivar 'Alamo' were surface-sterilized and placed on callus induction medium (MS salt and vitamin supplemented with 2.5 mg/L 2,4-D and 10 mg/L BAP). After 3-6 weeks, calli derived from explants were transferred to hormone free MS medium for plant regeneration. Plantlets 6-10 cm high were transferred to soil and are currently being grown in a greenhouse, shown in this picture, to evaluate their development and fertility. Hundreds of regenerated plants were obtained.

Fig.3. Transient expression of GUS reporter gene in switchgrass callus after particle bombardment.
Switchgrass calli derived from mature embryos were bombarded with gold particles carrying pUB-AG (A) and pAHC25 (B) plasmid, respectively. The GUS reporter gene is fused to the maize ubiquitin promoter in pAHC25 and to the rice actin promoter in pUB-AG. Bombardment was performed using the Biolistic Particle System (PDS-1000/He) according to the protocol we developed for wheat.

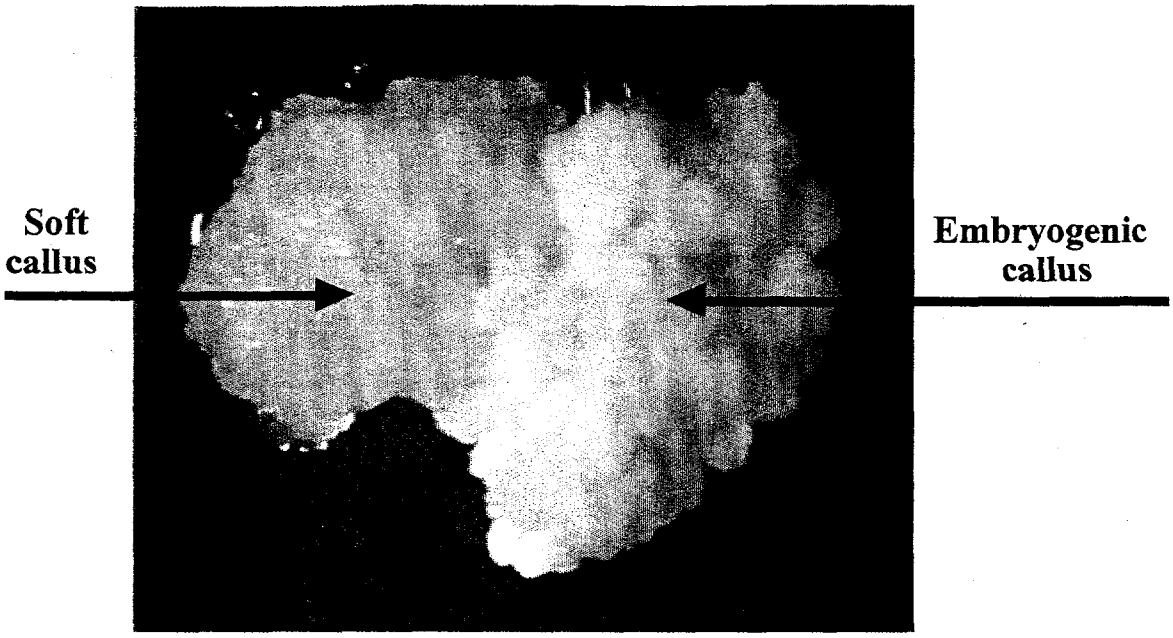


Fig.4.Multiple somatic embryogenic callus derived from a Switchgrass leaf base explant.
Leaf explants were obtained from sterile 2-week old switchgrass plants (cultivar 'Alamo') and placed on callus induction medium (MS salt and vitamin supplemented with 2 mg/L 2,4-D and 0.5 mg/L BAP). After 3-6 weeks, derived hard calli were transferred to hormone free MS medium for plant regeneration. Some regenerated plants were obtained.

Fig. 5A

Fig. 5B

Fig. 5. Effecti of Topik (A) and sethoxydim (B) on switchgrass.
Two-week old switchgrass plants (cv. Alamo) were sprayed with 25μM solution of herbicide in water containing 0.1% Tween 20. One week later the herbicide was applied again. Picture of treated (left) and control (right) plants was taken after one week after the second treatment.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Shaping the grass genome: acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci. Shaoxing Huang, Anchalee Sirikhachornkit, Justin Faris, Bikram Gill, Robert Haselkorn and Piotr Gornicki - publication in preparation and poster to be presented at the Annual Meeting of the American Society of Plant Physiologists in San Diego, July 2000.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

None yet

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None yet

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

One postdoctoral fellow, one graduate and one undergraduate student, directly participating in the project, receive an extensive training in plant physiology, genetics and evolution.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

None

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Same personnel

RECEIVED JUL 1 0 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	L.O. Ingram, J.F. Preston, III., and K.T. Shanmugam	
University:	University of Florida	
Agreement Number:	OR22072-94	
Project Title:	*Ethanol production from uronic acid-substituted xylose residues in hemicellulose hydrolysates*	
Reporting Period and Report Type:	**From:** **9/16/99** **To:** **6/30/00**	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. To determine the amounts of acid resistant glucuronic acid-substituted sugars in acid hydrolysates from candidate biomass species.

2. To investigate the limitations for concurrent sugar utilization in strains LY01 of *E. coli* and P2 of *Klebsiella oxytoca* and develop new and improved strains which more rapidly complete the conversion of hemicellulose-derived sugars to ethanol.

Layperson's Summary

The lignocellulose of woody biomass represents a significant underutilized resource for fermentative production of alternative fuels. The hemicellulose components, which constitute as much as 40% of the lignocellulosic biomass, must be converted to fermentable substrates for efficient and cost-effective processes for generating ethanol from the biomass. Pretreatment with dilute acid (0.5% sulfuric acid) at elevated temperatures (greater than 120C) is used to hydrolyze the hemicellulose fraction in commercial processes involved in the production of ethanol from biomass. The free sugars, principally xylose, generated by this process may be fermented to ethanol by recombinant strains of *Escherichia coli* and *Klebsiella oxytoca* that have been developed by our research group. A significant amount of the carbohydrate components of the hemicellulose

are resistant to hydrolysis in dilute acid. These contain glucuronic acid residues linked to xylose, and require enzymatic digestion for their conversion to fermentable substrate.

As noted in the previous progress report, we have established that the levels glucuronic acid in the hemicellulose fraction of sweetgum and cottonwood, both of which are a significant biomass resource for production of ethanol, are at levels that render as much as 25% of the sugars resistant to release and fermentation. Sugarcane bagasse and rice hulls, other important resources for bioconversion to ethanol, also have acid-resistant components in their hemicellulose fractions. We have now established that the recombinant ethanologenic strains of *E. coli* and *Klebsiella oxytoca* that have been developed for conversion of glucose and xylose to ethanol can efficiently convert these free sugars to ethanol in acid hydrolysates of hemicellulose. However, the oligosaccharides containing glucuronic acid linked to xylose or xylobiose are not metabolized by these recombinant strains. This work has established that the efficient conversion of acid hydrolysates of hemicellulose to ethanol will require biocatalysts that are able to cleave the bond linking glucuronic acid to xylose in the xylose-containing oligosaccharides that are resistant to acid hydrolysis. The ability to utilize the all of the xylose in the acid hyrolysates would increase the conversion of hemicellulose to ethanol by 5 to 10%, depending on the amount of glucuronic acid substitution. The genetic engineering of ethanologenic bacteria that can utilize the glucuronic acid as well as the xylose could increase the conversion of hemicellulose by 10 to 20%. Our current efforts are expected to lead to cost-effective processes for the commercial production of ethanol and bio-based products from underutilized lignocellulosic biomass.

Scientific Accomplishments

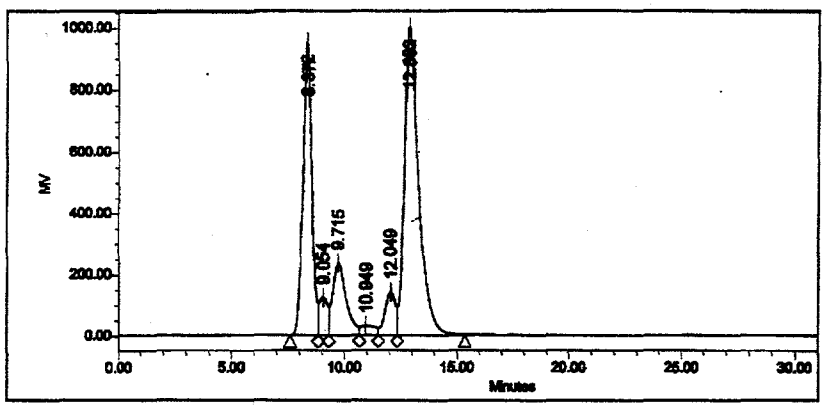
As noted in the previous progress report, 4-O-methyl-D-glucuronopyranosyl residues (MeGA) linked α-(1-2) to xylose (X) were detected and quantified in hemicellulose fractions of different hardwood biomass candidates, i.e. sweetgum (*Liquidambar styracifula*), cottonwood (*Populus deltoides*), and yellow poplar (*Lirodendron tulipifera*). The MeGA and X residues are part of a polysaccharide designated glucuronoxylan (GAXn) in which n is the number of xylose residues in the β-1.4-xylan polymer. The ratio of MeGA to X in the hardwood hemicellulose fractions was between 1:8 and 1:7, indicating that more than 20% of the monosaccaride equivalents, including MeGA and X residues, would not be released as fermentable monosccharides during the dilute acid pretreatment applied to the hemicellulose fractions for commercial ethanol production. This observation was confirmed with both ^{13}C-NMR analysis and HPLC quantitation of xylose and other sugars released after hydrolysis of GAXn fractions in 0.5% sulfuric scid. Similar analyses of the GAXn fractions from sugarcane bagasse and rice hull indicated the presence of MeGA residues comprising lesser amounts of the total monosaccharide equivalents, although the GAX produced during acid hydrolysis still represented approximately 10% or more of the monosaccharide equivalents.

We have now evaluated the fermentability of the GAX and xylose produced during the acid hydrolysis of GAXn derived from sweetgum sawdust, sugarcane bagasse, and rice hulls.

A chromatographic system has been developed to resolve and quantify the products of dilute acid hydrolysis of GAXn from hemicellulose fractions. Figure 1 depicts the HPLC analysis of LB medium containing sweetgum GAXn acid hydrolysate prior to inoculation. LB medium was prepared containing 1 % xylose equivalents from a calcium hydroxide-neutralized acid digest (0.5% H_2SO_4, 122 C, 1 h) of glucuronoxylan prepared from sweetgum sawdust. Samples were centrifuged and filtered, and 0.05 ml was delivered onto an Aminex HPX87H column eluted with 0.01 N H_2SO_4 at 65 C, 0.5 ml/min. Detection of components was made with a Waters R401 refractive index monitor. Peaks detected with retention times of 8.37, 9.05, 9.71, 12.05, and 12.88

min contain components of LB medium (yeast extract and tryptone), GAX2, GAX1, glucose, and xylose respectively. Using the same chromatographic conditions, standards of lactate, acetate, and ethanol eluted at 16.2, 19.02, and 26.7 min, respectively.

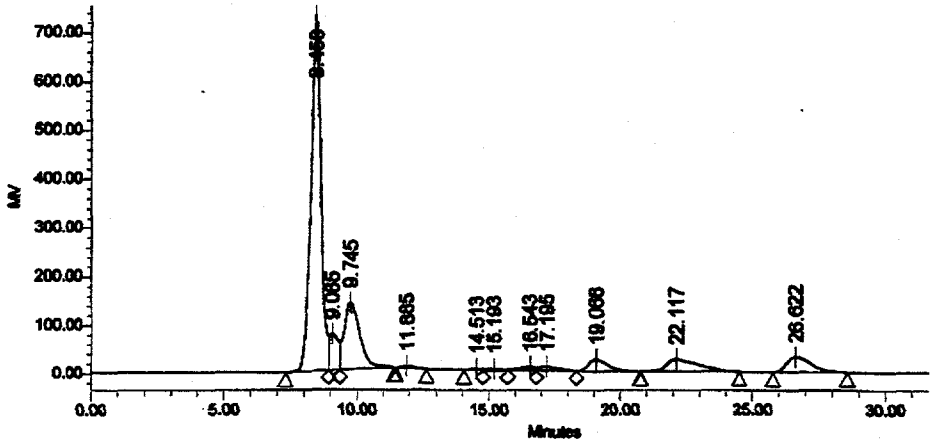
FIG 1



This HPLC protocol gave resolution of GAX2, GAX1, monosaccharide substrates, and fermentation products

Fig. 2 shows the HPLC profile of sweetgum glucuronoxylan hydrolysate in LB medium inoculated with *E. coli* KO11 and incubated for 24 h. *E. coli* KO11 was inoculated into 25 ml LB medium containing 1 % xylose equivalents and incubated in 125 ml Erlenmyer flasks at 35 C with rotary shaking at 120 rpm. Samples were removed at 24 h after inoculation, centrifuged, filtered, and delivered onto an Aminex HPX87H column eluted with 0.01 N H_2SO_4 at 65 C, 0.5 ml/min. Components with RT values of 9.065, 9.75, 16.45, 19.07, and 26.62 are GAX2, GAX1, lactate, acetate, and ethanol, respectively. The component with an RT of 22.12 min was seen after incubation of uninoculated medium, and presumably represents one or more components derived from the yeast extract or tryptone. The xylose has been completely utilized in 24 h.

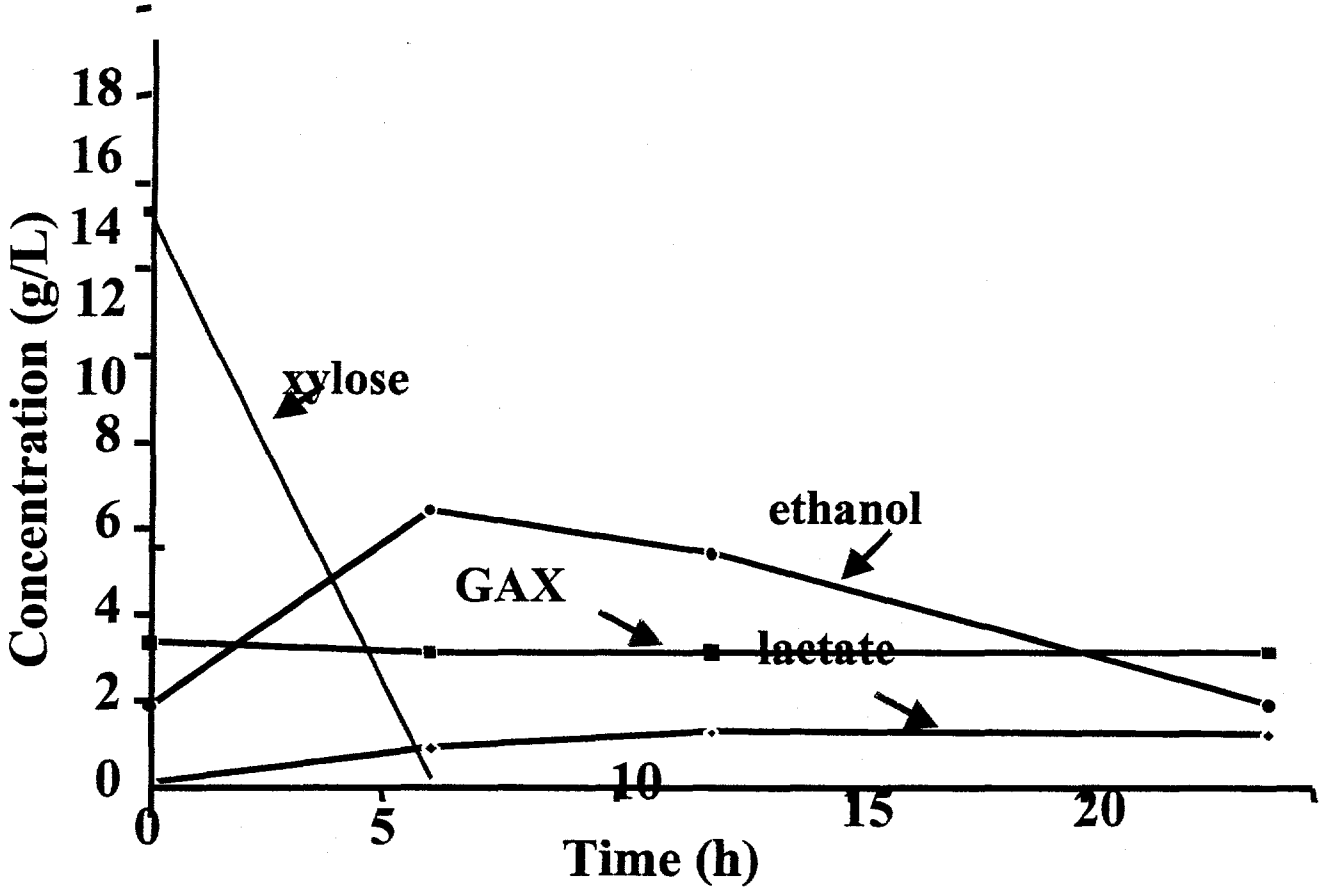
FIG 2



Cultures of of sweetgum g same time the GAX1 a have the α-glucuronidase required for the cleavage of 4-O-methyl-D-glucuronosyl-α-(1-2)-D-xylopyranose.

The progress of the formation of ethanol in cultures of *E. coli* KO11 growing on the acid hydrolysates of glucuronoxylan from sweetgum is depicted in Fig. 3

FIG. 3



This analysis demonstrates that the ethanologenic strain of *E. coli* KO11 is able to convert all of the free xylose to fermentation products, with ethanol as the predominant product. Lactate was also detected, although at 6 h, when all of the xylose had been metabolized, the ratio of ethanol to lactate was 13.3 on a molar basis. Small amounts of acetate and succinate were also detected, although these were less than 10 % of the lactate on a molar basis. The most notable result from this study was the persistence of the GAXn, with n equivalent to 1 or 2. Thus GAX1 and GAX2, which together comprise 25% of the hemicellulose monosaccharides of sweetgum, are refractory to metabolism by *E. coli* KO11 and unavailable as substrates for production of ethanol.

Similar studies were carried out with *E. coli* LY01, a strain derived from KO11 that is more tolerant of ethanol concentrations in the medium; and with *Klebsiella oxytoca* P2, a strain that had been engineered to express the *pdc* and *adh* genes from *Zymomonas mobilis* to produce ethanol instead of butanediol. As in the case of *E. coli* KO11, these recombinant ethanologenic strains were also unable to metabolize GAX1 or GAX2, and consequently were able to convert less than 80% of the carbohydrate of the sweetgum hemicellulose to ethanol.

Both *E. coli* KO11 and LY01 were evaluated for their ability to convert the acid hydrolysates of the glucuroxylan fractions obtained from sugarcane bagasse and rice hulls. As in the case of the hydrolysate from sweetgum, the xylose was readily converted to ethanol, while the GAX1 and GAX2 were not metabolized. The results of all of these studies with respect to the metabolism of GAX and xylose are summarized in Table 1. In all cases the GAX (including GAX1 and GAX2) are persistant components of the fermentation broth, or stillage.

Table 1 . Comparative utilization of glucuronoxylan-derived xylose and GAX by ethanologenic strains of *E. coli* and *K. oxytoca.*

GAXn	Bacterial	Xylose[c]	GAX[d]

Hydrolysate[a]	Strain[b]	t_0	t_{24}	remainder, %	t_0	t_{24}	remainder, %
Sweetgum	EcKO11	13.9	0	0	3.3	3.2	97
Sweetgum	EcLY01	13.9	0	0	3.1	3.3	106
Sweetgum	KoP2	15.4	0	0	3.5	3.3	94
Sugarcane	EcKO11	16.9	0	0	0.66	0.65	97
Sugarcane	EcLY01	17.3	0	0	0.66	0.66	100
Rice hulls	EcKO11	9.78	0	0	0.56	0.44	74
Rice hulls	EcLY01	10.5	0	0	0.56	0.53	95

a- GAXn hydrolysates were prepared by autoclaving 1 % glucuronoxylan from sweetgum, sugarcane bagasse, or rice hulls in 0.5% H_2SO_4 for 30 min. The hydrolysates were neutralized with $Ca(OH)_2$, centrifuged to remove $CaSO_4$, and 0.05 ml samples subjected to HPLC on BioRad Aminex HPX87H column eluted with 0.01 N H_2SO_4 at 65 C. Quantitative detection was made by refractive index.
b- The bacterial strains used in this study were *Escerichia coli* KO11, *Escherichia coli* LY01, and *Klebsiella oxytoca* P2, each expressing the *pdc* and *adh* genes derived from *Zymomonas mobilis*.
c- Xylose was quantified as a single peak detected by RI,
d- GAX represents the sum of GAX1 and GAX2, each quantified by RI.

These studies establish the need to develop ethanologenic strains that are able to cleave the α-(1-2) linkage between the 4-O-methyl-D-glucuronic acid moiety and D-xylopyranose residues in the β-(1-4) xylan backbone of the hemicellulose. The additional xylose that would then be available for fermentation and would increase the ethanol yield by an additional 10 % or more with hardwood lignocellulosics serving as the biomass resource, and by as much as 5 % or more with lignocellulosic biomass from sugarcane bagasse and rice hulls. Engineering ethanologenic strains to convert the glucuronic acid residues to ethanol could double these recovery values. This enhanced recovery of fermentable substrate should make it possible to develop cost-effective and competitive processes with which to convert underutilized lignocellulosics to ethanol and other bio-based products.

Publications and Presentations

The following publication is nearly completed for submission to Biotechnology and Bioengineering.

Properties of the Hemicellulose Fractions of Lignocellulosic Biomass Affecting Bacterial Ethanol Production.

M. Rodriguez, A. Martinez, S. York, L.O. Ingram and J.F. Preston

Abstract

The compositions of the 4-O-methyl-glucuronoxylan fractions from sweetgum sawdust, sugarcane bagasse, and rice hulls have been determined before and after dilute acid hydrolysis.

Based upon ^{13}C-NMR spectrometry, the sweetgum glucuronoxylan contained 4-O-methyl-D-glucuronic acid residues and xylose residues in a ratio of 1.2 to 10. Xylan preparations from sugar cane bagasse and rice hulls also showed the presence of 4-O-methyl-D-glucuronic acid residues, although in amounts less than 5% of the total carbohydrate. Fermenation of dilute acid hydrolysates prepared from each of these lignocellulosic sources was evaluated in three recombinant ethanologenic strains of bacteria, including *Escherichia coli* strains KO11 and LY01, and *Klebsiella oxytoca* strain P2. All strains completely converted free xylose to ethanol and small amounts of other fermentation products, and accumulated glucuronoxylose as a limit product in the fermentation broth. The resistance of glucuoronoxylose to further catabolism establishes the need for genetic modification(s) to allow the cleavage of the glucosiduronic bond that is linked α(1-2) to xylose, as well as the metabolism of the 4-O-methylglucuronic acid, for the efficient conversion of the hemicellulose fractions of lignocellulosic biomass to ethanol.

Technology Transfer

None at this time

.

Commercial Accomplishments

Since the last reporting period, the most significant accomplishment is the demonstration that GAX1 and GAX2, which comprise 10 to 25% of the carbohydrate biomass of the hemicellulose fractions of underutilized lignocellulosic biomass resources, are produced during commercial acid pretreatment processes, and are not utilized by ethanologenic strains of bacteria. From this information, a set of objectives has been defined to genetically engineer bacteria that will be able to use the xylose and glucuronic acid residues in the GAX fractions for the efficient conversion of hemicellulose fractions to ethamol and other bio-based products.

Educational Accomplishments

Educational accomplishments include the training of post-doctoral research associate, Jason Hurlbert, and Maria Rodriguez as a research associate with plans to pursue a Ph.D.

Additional Funding

A grant for $96,008 has been awarded from DOE from 09/30/99 to 09/30/01 for Enzymatic Complementation of Acid Hydrolysis for Hemicellulose Bioconversion, J.F. Preston, PI; L.O. Ingram, Co-PI; K.T. Shanmugam, Co-PI. This will support efforts to develop endoxylanases and acetyl esterases to mitigate the production of inhibitors, e.g. furans, which are formed during the acid hydrolysis of hemicellulose. The work supported by CPBR is indirectly related to this

effort in that enzymatic pretreatment protocols may assist in the production of higher yields of xylose for fermentative conversion to ethanol.

Key Personnel Hiring or Turnover

Maria Rodriquez has returned to Mexico. Yolanda Salinas Moreno, a Ph.D. candidate in the Department of Plant Pathology at Colegio de Postgraduados en Ciencias Agricolas, Mexico, has taken her position as a research associate and is evaluating the co-metabolism of pentoses and hexoses by recombinant ethanologenic strains of *E. coli* and *K. oxytoca*.

RECEIVED APR - 3 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Jiming Jiang, Ph.D.	
University:	University of Wisconsin-Madison	
Agreement Number:	OR22072-95	
Project Title:	Toward cloning a functional rice centromere	
Reporting Period and Report Type:	**From:** 12/1/98 **To:** 5/31/00	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

Fully characterize the DNA composition of rice centromeres (Year One).

1. Sequenced a 68-kb rice centromeric bacterial artificial chromosome (BAC) clone.
2. Analyzed the cytological locations of more than 30 rice BAC clones containing centromeric repeats.
3. Developed a special fluorescence in situ hybridization (FISH) technique to visualize the chromosomal origin of centromeric BAC clones.
4. Several BAC clones derived from centromere 11 were identified. DNA sequences specific to this centromere were isolated.
5. Discovered a rice neocentromere.

Develop physical maps which span the centromeric regions and characterize the DNA composition of individual centromeres (Year Two).

We analyzed the fingerprinting data of ~400 rice BAC clones containing centromeric repeats.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

The main goal of the current CPBR project is to determine the DNA sequences located in the rice centromeres. The centromeres of eukaryotic chromosomes contain mainly repetitive DNA sequences. It is very difficult to localize a large DNA fragment to a specific centromere because of the presence of the repetitive DNA sequences. Using specially designed techniques and research strategies we were able to assign several large DNA fragments to specific rice centromeres. This is the first such achievement in plant species. This progress will provide the foundation for us to discover the DNA sequences spanning an entire rice centromere. We also discovered a rice "neocentromere", which is a new functional centromere originated from a non-centromeric chromosomal fragment. This is the first *de novo* neocentromere reported in plant species. Neocentromeres are extremely important tools in humans and other model organisms for centromere studies. The rice neocentromere will provide us an unprecedented tool to study the structure and function of plant centromeres in the future.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

We sequenced a BAC clone containing a 68-kb DNA fragment derived from a rice centromere. The most dominant component of this BAC clone, 17p22, is a 155-bp satellite repeat. The high sequence similarity between the monomers of this repeat made sequence assembly difficult. We developed a cytology-based digital mapping technique. Using this technique we were able to determine the distribution of the satellite repeat and other low copy DNA elements within the BAC insert (Jackson et al. 1999). The locations of the low copy DNA elements provided anchor markers for sequence assembly. A single contig was successfully constructed from the sequencing data. DNA sequences homologous to retrotransposons were discovered in 17p22, indicating that these sequences may be derived from a degenerated retrotransposon. Part of the 155-bp rice centromeric satellite repeat is partially homologous to the maize centromeric satellite repeat Cent-C. This dramatic discovery suggests that the major DNA components of the rice and maize centromeres are at least partially conserved after over than 100 million years of divergence of these two species.

We screened a rice BAC library using several repetitive DNA elements isolated from BAC 17p22. More than 30 BAC clones were analyzed by FISH. These BAC clones can be divided into two groups. Clones in the first group hybridized specifically to the centromeric regions (Fig. 1a). Clones in the second group hybridized all over the chromosomes (Fig. 1b), indicating that these clones contain repetitive DNA elements which are not specific to the centromeres. We have selected several clones in the first group to isolate DNA elements which are specific to rice centromeres but are different from those in 17p22. The research strategy is to digest the DNA from these clones with several restriction enzymes and to blot the DNA to nylon membranes. DNA fragments, which do not hybridize to 17p22, will be subcloned into plasmid clones and characterized using molecular and cytogenetic approaches.

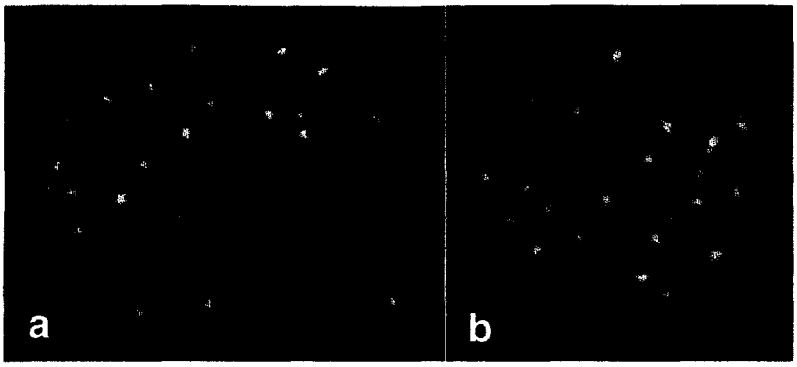


Fig. 1. Two different types of rice centromeric BAC clones. (**a**) A BAC clone hybridized only to the centromeric regions. (**b**) A BAC clone hybridized to all over the chromosomes.

Although the clones in the first group hybridize only to the centromeric regions, the chromosomal origins of the BACs cannot be determined because these clones hybridized to all the 12 rice centromeres. We developed a FISH technique which includes a special preanneal procedure with rice C_0t-1 DNA. Using this technique we found that the signals derived from several BAC clones in the first group are much stronger in one pair of centromeres than those in other centromeres (Fig. 2). Such signals from three centromeric BAC clones were localized to specific chromosomes by co-hybridization with rice chromosome-specific DNA markers developed recently in our laboratory. This is the first time that large insert DNA clones containing centromeric repeats are mapped to specific chromosomes in plants. The chromosome-specific centromeric DNA sequences from such BACs will provide anchor markers to construct contigs spanning complete rice centromeres. We have developed sublibraries from two BAC clones which were mapped to the centromeres of chromosomes 3 and 11, respectively. Plasmid clones containing sequences different from these of 17p22 will be isolated and sequenced.

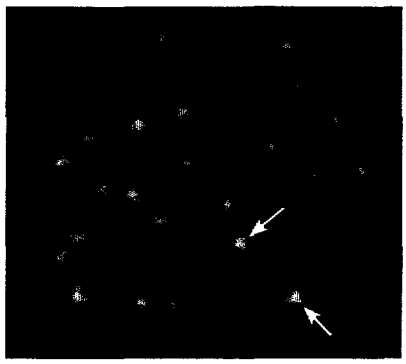


Fig. 2. A rice centromeric BAC showed stronger signals in the centomeres of chromosome 11 (arrows) than those in other centromeres

The 5S ribosomal RNA genes were previously mapped on the short arm very close to the centromere of rice chromosome 11. We did a high resolution FISH mapping on rice pachytene chromosomes using a 5S rDNA probe and a satellite DNA probe, pRCS2, which is specific to the centromeres of rice. Both the 5S rRNA genes and the RCS2 satellite DNA are organized into two separate domains on early pachytene chromosomes (Fig. 3b, c). One of the two 5S rRNA gene domain is clearly located between the two satellite DNA domains (Fig. 3d). This result revealed that the 5S rRNA genes are part of the rice centromere 11 because the RCS2 satellite DNA is only located in the primary constriction of rice chromosomes. We isolated four 11L (the long arm of chromosome 11) telocentric chromosomes which are derived from independent centromere misdivision events of chromosome 11. All four 11L telocentric chromosomes carry 5S rRNA genes in the centromere based on FISH analysis (Fig. 3e). The 5S rDNA signals are clearly located outside of the RCS2 signals on prometaphase 11L (Fig. 3e), confirming that the 5S rDNA is part of the centromeres of 11L telocentric chromosomes.

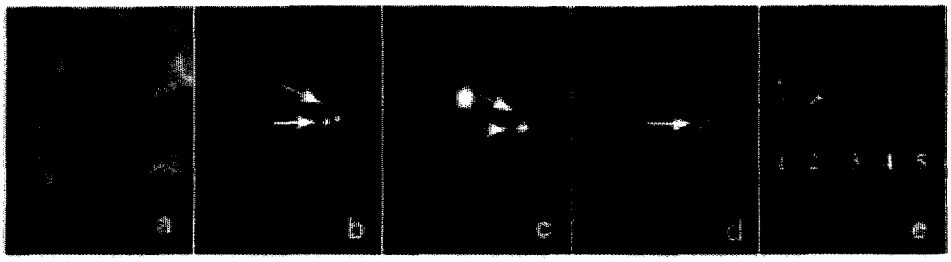


Fig. 3. High resolution FISH mapping of the 5S rRNA genes on early pachytene chromosomes of rice. (**a**) Part of a pollen mother cell at early pachytene. (**b**) The FISH signals derived from a 5S rDNA probe are separated into two domains. (**c**) The FISH signals derived a rice centromere-specific satellite DNA probe pRCS2 are also separated into two domains on the same chromosome. (**d**) One of the two 5S rDNA domains (green signal, arrow) is located between the two RCS2 domains, indicating that this domain is located within the primary constriction. (**e**) Locations of the 5S rRNA genes on telocentric chromosomes derived from rice chromosome 11. **1**, An isochromosome 11S.11S shows 5S rDNA (green) and pRCS2 (red) signals on both arms; **2-5**, All four telocentric 11L chromosomes, which were derived from different centromere misdivision events, contain 5S rRNA genes. The 5S rDNA signals (green) always locate outside of the pRCS2 signals (red) on prometaphase chromosomes.

We screened a rice BAC library using a 5S rDNA probe and probe pRCS1 that is dispersed in the rice centromeric regions. Several BAC clones hybridizing to both pRCS1 and the 5S rDNA probe are identified. Sequencing of the ends of one of these BAC clones, 22B9, indicated that only one end of the BAC insert is derived from 5S rRNA genes. About 30% of the insert of 22B9 contains non-rDNA sequences. Characterization of the non-rDNA sequences is underway. The discovery of the association of 5S ribosomal RNA genes with the centromere of chromosome 11 provides us another approach to isolate DNA sequences specific to centromere 11. Thus, rice chromosome 11 will be the best target to construct a DNA contig spanning its complete centromere.

We have isolated numerous cytogenetic stocks containing rearranged rice chromosomes from a triploid rice. The rearrangements in some of the rice chromosomes presumably involve in the centromeres based on morphology of pachytene chromosomes. We recently made an extensive survey on these rearranged chromosome by FISH analysis using rice centromeric DNA probes. An abnormal metacentric chromosome in a trisomic stock, CZ37, is completely devoid of any rice centromeric DNA elements isolated in our lab. All the root tip cells analyzed from a CZ37 plant contained one copy of this chromosome, indicating that this chromosome has a normal transmission in somatic cells. However, this chromosome lags in majority of the anaphase I cells in meiosis, suggesting that the centromere of this chromosome is probably only partially functional. It is highly likely that this chromosome contains a neocentromere, possibly the first *de novo* neocentromere discovered in plant species. This neocentromere will provide us an unprecedented tool to study the structure and function of plant centromeres.

By collaborating with Dr. Rod Wing's lab at the Clemson University we have started to study the fingerprinting data of BAC clones containing centromeric DNA repeats. Several BAC contigs including multiple clones were identified by grouping approximately 400 BAC clones with at least one end homologous to previously identified rice centromeric repeats.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Publications:

Jackson, S. A., F. Dong, and J. Jiang (1999) Digital mapping of bacterial artificial chromosomes by fluorescence in situ hybridization. Plant J. 17: 581-587.

Invited Seminars:

Department of Botany and Plant Pathology, Purdue University. Molecular structure of grass centromeres. February 10, 1999.

1999 Symposium of the Consortium for Plant Biotechnology Research, Washington, D.C. Toward cloning a functional rice centromere. March 3, 1999.

Department of Plant Breeding, Cornell University. Molecular structure of grass centromeres. July 22, 1999.

Department of Agronomy, Yangzhou University, P. R. China. Fluorescence-based DNA *in situ* hybridization techniques and their applications in plant genome mapping. August 11, 1999.

Institute of Nuclear Agricultural Science, Zhejiang University, P. R. China. Molecular structure of plant centromeres. August 18, 1999.

Department of Plant Pathology, Kansas State University. Centromeres of plant chromosomes. October 21, 1999.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

Patent Submitted: Jiang, J. and F. Dong (1999) DNA sequences specific to rice centromeres.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

One graduate student, Fenggao Dong, completed his Ph.D. degree in March, 2000. Fenggao was partially supported by the current CPBR grant during the past year and played the major role in characterizing a large DNA fragment derived from a rice centromere. Fenggao will join Cereon Genomics (Monsanto) in April of 2000.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

None

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Dr. Zhukuan Cheng, hired in July of 1999 as a new postdoctoral associate by the current CPBR grant.

Dr. Alexander S. Parokonny, hired in March, 2000 as a new postdoctoral associate by the current CPBR grant.

RECEIVED JUL 19 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Gayle Lamppa, Ph.D.	
University:	The University of Chicago	
Agreement Number:	OR22072-96	
Project Title:	*Accumulation of products within the plastid of biomass conversion: test system with cellulase*	
Reporting Period and Report Type:	**From:** 1/1/2000 **To:** 6/30/2000	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. Design gene constructs that code for a precursor fusion protein with endoglucanase E1 fused to a transit peptide with a site cleavable by the stromal processing peptidase (SPP) at the junction.

2. Expression of gene constructs and production of endoglucanase E1 *in vivo*.

3. Determine if an active endoglucanase can be recovered by *in vitro* translation, and if the transit peptide affects this activity.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

Please complete.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Please complete.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Please complete.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

Please complete.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

Please complete.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Please complete.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

Please complete.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Please complete.

GENERAL SUMMARY OF PROJECT OBJECTIVES AND PROGRESS

Our immediate goal is to determine if a cellulase, endoglucanase E1 from Acidothermus cellulocystis which is a thermophilic bacterium, can be accumulated in plastids in an active form. It seems reasonable that the cytosol would not tolerate high levels of a cellulase during plant growth and development, and hence, an effective means of sequestering it may be crucial for its overexpression in vivo. We propose that the numerous chloroplasts per cell will serve as an excellent site for this purpose. Our specific aims should contribute to the production of usable and relatively inexpensive amounts of a cellulase in transgenic plants, which subsequently can be used for cellulose degradation and glucose fermentation to ethanol. **Our specific research objectives are described below, and a brief summary of the progress we have made is provided.** Our results relating to these goals are presented in greater detail in the section "Scientific Accomplishments".

Objective 1. Design gene constructs that code for a precursor fusion protein with endoglucanase E1 fused to a transit peptide with a site cleavable by the stromal processing peptidase (SPP) at the junction.

The goal of these experiments was to test the general hypothesis that transit peptides and any protein linked to them must be compatible for efficient processing and import into chloroplasts, and specifically, this would be the case for endoglucanase E1. Hence, we have separately linked E1 and its catalytic domain to two different transit peptides, and have attempted to optimize the transit peptide-mature protein junction for cleavage by SPP and import.

An array of fusion proteins (see Figure 1) was made using the transit peptide of ferredoxin (FD). We found that the efficiency of proteolytic processing and import is significantly affected by the number of amino acids immediately following the cleavage site, before the start of E1. Hence, it appears that the spacing at the transit peptide-E1 junction is critical. This is supported by the observation that an insertion after the cleavage site of 10 amino acids from the signal peptide of E1, which is normally synthesized as a precursor that is exported from Acidothermus, also promotes processing and import.

Another important observation is that a fusion protein containing only the catalytic domain (CD) of E1 is more readily cleaved and imported than a comparable precursor with full length E1.

We have extended our results with the transit peptide of FD by using the transit peptide of Rubisco activase (RBCA) fused to E1 and CD. In comparable constructs, each transit peptide confers different properties, which are discussed at greater length in "Scientific Accomplishments". Nevertheless, precursors with the transit peptide of RBCA are also cleaved by SPP and imported.

Objective 2. Expression of gene constructs and production of endoglucanase E1 in vivo.

The results presented above identified several important gene constructs encoding precursor fusion proteins that are efficiently cleaved and imported in vitro. Other constructs that are not cleaved or imported serve as useful controls for the in vivo analyses. We selected five of these for transformation of tobacco. The genes were placed downstream of the nearly constitutive CaMV 35S promoter with 5' and 3' nontranslated regions as part of a cassette. This promoter was selected to assure fairly high transcription levels. The transformants are described below in the section on "Scientific Accomplishments". Leaf disks were inoculated with Agrobacterium and regenerated plants---the primary transformants---are being analyzed. The level of E1 and

endoglucanase activity is being examined in different tissues of the transgenic plants, starting with leaves. We have started with assays of total cellular protein, and now are isolating chloroplasts to determine the success of targeting E1 to this organelle. We already know that very little E1 activity can be detected in the transgenic plants if E1 is synthesized without a transit peptide for chloroplast localization.

Objective 3. Determine if an active endoglucanase can be recovered by in vitro translation, and if the transit peptide affects this activity.

Based on the recommendations of the reviewers of our original proposal, our revised research objectives de-emphasized this goal. However, as we began to make more complicated constructs, we realized that the addition or deletion of amino acids at the N-terminus of E1 might influence endoglucanase activity. Therefore, there was good reason to attempt in vitro expression. Furthermore, we foresaw the possibility of manipulating cellulase activity genetically if we had a relatively rapid assay for analyzing mutant constructs. These experiments have been successful. We have synthesized active E1 as well as its CD by in vitro transcription of their respective genes followed by translation of their transcripts. Unexpectedly, the precursor fusion proteins are active. Activity is inhibited by a specific mutation at the N-terminus of E1 in the precursor, and optimal activity depends on high temperature. Our studies on the import efficiency of E1 and its CD into isolated chloroplasts, using the array of constructs already created (see Objective 1 and "Scientific Accomplishments"), can now be extended to include assays for a functional enzyme after transport into the organelle in vitro.

IN SUMMARY, based on the Time Table we proposed for the last 18 months of our experiments, our studies are right on schedule. An illuminating series of fusion proteins have been synthesized and analyzed. Importantly, they demonstrate the specificity of SPP's recognition of different substrates. Further, the translocation apparatus of the chloroplast imports these substrates with different efficiency. We have analyzed E1 expression in the primary tobacco transformants. Seeds from these transgenic plants are now available for producing the next generation of plants (and sufficient amounts of tissue) for detailed biochemical studies. In addition, we have developed a new in vitro system for E1---and CD---expression that should help us explore in the future some of the structural features of E1 needed for activity. The in vitro expression system should facilitate testing of some of the predictions that have been made based on structural information and evolutionary comparisons of different cellulases.

LAYPERSON'S SUMMARY

Our goal is to try to accumulate large amounts of active cellulase in the chloroplasts of transgenic plants, where it will be sequestered and thus unlikely to have deleterious effects on plant growth and development. Furthermore, it might be possible to rapidly enrich for the enzyme by chloroplast isolation. We chose for our study endoglucanase E1 from Acidothermus cellulocystis, a thermophilic bacterium (courtesy of Steven Thomas, NREL). We proposed to synthesize precursor fusion proteins with signals that would specifically target E1 to the chloroplast, where the signal---called the transit peptide---would then be cleaved, releasing mature E1 within the organelle. We proposed to intially test our fusion proteins in vitro for proteolytic processing and import into isolated chloroplasts in order to select the best candidates for more costly in vivo studies using transgenic plants. Thus far we have made important steps forward in designing constructs that promote efficient cleavage of a subset of precursor fusion proteins and we also demonstrate that these precursors are imported. The junction between the transit peptide and E1 is a critical region within the structure of the precursor. In addition, different transit peptides affect processing and import to different degrees. These studies have

guided our choice of constructs for transformation of tobacco. Primary transformants have been regenerated, and are being analyzed. We are starting a comprehensive analysis of E1 levels and activity within chloroplasts of transgenic plants.

In parallel studies, we attempted to express active E1 in an in vitro system starting with the E1 gene. This includes in vitro transcription followed by in vitro protein synthesis. We are excited that we have obtained an active enzyme. In the future, this allows us to extend our in vitro studies to explore the import and release of active E1 in the chloroplast. We can test the many constructs that we have already synthesized in the import assay and these findings can be compared with our in vivo results. Further, we foresee that this should allow us to genetically manipulate E1 structure and rapidly analyze what changes influence its activity. Many predictions have been made based on structural analyses and evolutionary comparisons of cellulases that can be tested using the in vitro expression system, which should facilitate a comprehensive study on structure-function relationships.

SCIENTIFIC ACCOMPLISHMENTS

In this section, I describe the nature of the constructs we have designed, and discuss in greater detail our experimental results. Representative original data are presented.

Objective 1. Design gene constructs that code for a precursor fusion protein with endoglucanase E1 fused to a transit peptide with a site cleavable by the stromal processing peptidase (SPP) at the junction.

A. Precursor fusion proteins using the ferredoxin transit peptide.

1. Constructs using full-length endoglucanase E1 with insertions or a specific deletion.

We initially designed a series of constructs that encoded precursor fusion proteins with E1 linked to the transit peptide of ferredoxin (FD). The FD transit peptide is considered to possess considerable structural flexibility, and has been used previously to import some foreign proteins into chloroplasts (de Boer et al., 1991, EMBO J. 10: 2765-2772 and Pilon et al., 1992, J. Biol. Chem. 267: 19907-19913), and hence it was chosen for our first studies. Figure 1 lists each construct (#1-#11) used in our experiments to-date, and shows the structure at the transit peptide-mature protein junction of each precursor fusion protein. The results from our assays are also tabulated in Figure 1.

In the first three precursors, the transit peptide and cleavage site were left intact and an increasing number of amino acids---one, five and fifteen residues---from mature ferredoxin were included as a spacer before the start of mature E1. Construct #2 (FD+1::E1, transit peptide plus one amino acid fused to E1) was not proteolytically cleaved by recombinant SPP from E. coli or by a soluble chloroplast extract. Construct #3 (FD+5::E1) was cleaved very weakly by SPP, and not at all by the chloroplast extract (Figure 2). For construct #2 import into pea chloroplasts was "poor", whereas we characaterized the import of construct #3 as "good". On the other hand, an increase in the spacer region to 15 residues in construct #4 (FD+15::E1) yielded a precursor that was efficiently cleaved by SPP and the chloroplast extract (Figure 2). FD+15::E1 import was characterized as "very good"; it imported into the chloroplast, processed, and most of mature E1 was found in the stromal fraction (Figure 3). Lack of degradation by thermolysin treatment of the chloroplasts demonstrated that E1 was indeed sequestered within the organelle.

When we observed that construct #2 was not processed, and #3 was processed inefficiently, and further that import was relatively low, we became concerned that the unusual residues at the start of E1---Ala-Gly-Gly-Gly-Tyr (AGGGY)---might prevent or compromise recognition of the transit peptide by SPP and the chloroplast import machinery. Glycines are

well-known to alter protein conformation. Therefore, in construct #5 (FD+5::Δ5E1) these residues were deleted. Although FD+5::E1 itself was not efficiently processed or imported, the AGGGY deletion resulted in a precursor that was cleaved by SPP. However, the chloroplast extract did not remove the transit peptide. (One explanation for the discrepancy is that recombinant SPP is significantly more "robust" than SPP in the chloroplast extract, and there is thus a difference between the amount of active SPP relative to the substrate. Another possibility to consider is that features of some substrates are not recognized by SPP in the context of other factors in the chloroplast extract. FD+5::Δ5E1 was also imported into chloroplasts, and cleaved. We conclude that the very N-terminal AGGGY of E1 can have a negative effect on different steps in the import pathway if located in close proximity to the transit peptide.

To test the hypothesis that a spacer is needed between the transit peptide and E1 for efficient cleavage, as suggested from our results with construct #4 (FD+15::E1), we generated construct #6 (FD+5::sp10E1) which contains, besides the 5 residues from mature FD, an additional 10 amino acids from E1's own signal peptide. The introduction of this spacer sequence yielded a precursor that was efficiently cleaved by SPP and the chloroplast extract. It was also imported into the chloroplast, processed and 40% was found in the stroma and 60% in the the membrane fraction after thermolysin treatment. That $_{sp10}$E1 is found in the membrane fraction is probably due to the presence of the region from the signal peptide, which contains a number of hydrophobic residues (Ala, Val, and Pro). It will be interesting to determine if suborganellar location influences E1 activity.

2. Constructs using the catalytic domain of E1.

E1 is comprised of three domains: the N-terminal catalytic domain (CD), the serine-proline rich linker, and the C-terminal cellulose binding domain (CB), as illustrated in Figure 1A. We entertained the idea that other structural features of E1, in addition to the AGGGY sequence, were in some way affecting processing by SPP and transport into the chloroplast. To investigate this question, we decided to determine if CD alone---separated from the linker and CB---fused to a transit peptide would be a better substrate in these reactions. Hence, three new precursor fusion proteins were synthesized:

-Construct #7 (FD+1::E1CD, i.e. FD transit peptide plus 1 amino acid fused to CD),
-Construct #8 (FD+5::E1CD) and
-Construct #9 (FD+5::Δ5E1CD).

Once again, when only one amino acid followed the transit peptide in construct #7 (FD+1::E1CD), the precursor was not processed by SPP, and import was "poor". In contrast, both construct #8 (FD+5::E1CD, Figure 2) and construct #9 (FD+5::Δ5E1CD) were processed, albeit not as well as several other constructs listed in Figure 1B, or presented in Figure 2. Import was quite efficient for both precursors. The results for FD+5::E1CD are shown in Figure 3.

B. Precursor fusion proteins using the transit peptide of Rubisco activase.

Transit peptides do not share a common primary sequence. Therefore, it seemed important to extend our study and determine if an alternative to the FD transit peptide would yield a precursor containing E1 that was more efficiently recognized in the processing and import assays. We selected ribulose-bisphosphate carboxylase/oxygenase activase (RBCA) because we had found previously that at least in vitro, the RBCA precursor itself is very efficiently cleaved by SPP. Two constructs were made with the RBCA transit peptide plus five amino acids from mature RBCA:

-Construct #10 (RBCA+5::E1) and
-Construct #11 (RBCA+5::E1CD).

Construct #10 (RBCA+5::E1) was efficiently cleaved by recombinant SPP from E. coli and by the chloroplast extract. However, import for RBCA+5::E1 was poor. Construct #11 (RBCA+5::E1CD) contained the CD fused to the transit peptide, and it was processed very efficiently (Figure 2) and imported (Figure 3). From a comparison of the results using the RBCA+5 constructs (#10 and #11) versus the FD+5 constructs (#3 and #8), we conclude that RBCA+5 promotes more efficient processing of the E1 precursors by SPP.

Objective 2. Expression of gene constructs and production of endoglucanase E1 in vivo.

In the General Summary, I described the basic expression cassette used for transformation. Figure 1 (far right column) lists the 5 different constructs that we have introduced into tobacco leaf disks using this cassette. We selected the FD transit peptide::E1 constructs (#4, #6 and #8) that were positive in our in vitro processing and import assays, as well as construct #3 that was not well processed. In addition, we included construct #1; that is, the gene for E1 without a linked transit peptide. The aim was to assess the importance of targeting to the chloroplast for detection of E1 activity in vivo.

Based on our in vitro results with the RBCA transit peptide constructs, we have also introduced construct #11 containing the catalytic domain of E1 into tobacco as well. We have restricted ourselves to a careful analysis of lines of these transgenic plants because of the large number of individuals that must be examined for a thorough analysis and accurate interpretation of what happens to E1 expression in vivo. Thus far, 137 plants have been regenerated which represent , on average, 27 independent transformants for each of the 5 precursor fusion protein constructs.

An initial analysis of total cellular extracts from the primary transformants has been performed. Expression of E1 was monitored using a MUC assay, as decribed in the next section. The range of E1 expression varies, as expected given that "position effect" upon gene insertion into the nuclear genome can significantly influence the level of gene expression. The lowest levels of expression are found for E1 lacking a transit peptide, indicating the importance of mobilizing E1 out of the cytosol. The remainder of the transgenic plants, with either the full-length E1 construct or CD alone targeted to the chloroplast, are under investigation, and the results are considered confidential until a comprehensive study is completed. Seeds have been collected from the primary transformants and the next generations will be analyzed.

Objective 3. Determine if an active endoglucanase can be recovered by in vitro translation, and if the transit peptide affects this activity.

We have investigated whether E1 can be expressed in an active form following in vitro transcription and translation. Nine of the constructs shown in Figure 1 have been tested, and our results are presented in Table 1. Each one of these genes was inserted dowstream of the T7 bacteriophage promoter, allowing for in vitro transcription and coupled translation in a TNT system (Promega). This system includes a reticulocyte lysate for synthesis of radiolabeled protein. The activity of the translation products was monitored using a MUC assay that is very similar to a ß-glucuronidase (GUS) assay. Experiments were first carried out at 65° C and fluorescence measured. A complete description of the assay is given in the legend of Table 1. Table 2 demonstrates that the values obtained are within the linear range of detection.

Significantly, we were able to synthesize active E1 in the in vitro expression system. To our surprise, however, six of the precursor fusion proteins exhibited nearly as much activity as endoglucanase E1without a transit peptide. Construct #3 (FD+5::E1) showed 85% of the activity

found for E1 alone, and construct #10 (RBCA+5::E1) showed 93% of E1 activity. Precursors with only the catalytic domain of E1 were also tested. Construct #8 (FD+5::E1CD) contained 67% of E1 activity. On the other hand, construct #11 (RBCA+5::E1CD) was significantly more active, with 126% of E1 activity.

The two precursors---constructs #5 and #9---with the AGGGY deletion near the N-terminus of E1 showed very low levels of activity (9% and 11% of E1 activity, respectively), yet both were higher than the control reactions, where a background of 2% activity was found using a vector containing the gene for the native precursor of RBCA (without the E1 coding region), or no vector at all in the MUC assay.

We performed the MUC reactions at 65° C again, only we first treated the precursor fusion proteins with recombinant SPP (see Table 1, legend) There was an average of ~21% increase in endoglucanse activity after processing by SPP (Table 1). We tentatively conclude from these experiments that the transit peptide can function as a separate domain from the mature protein, and does not strongly interfere with the enzymatic activity of E1 under these in vitro conditions. Nevertheless, these are still preliminary experiments and conditions may be identified where the transit peptide has an inhibitory effect on the ability of E1 to carry out its role in cellulose degradation. Further, it might be crucial that transit peptide removal is accomplished within the organelle after import, since proteins with transit peptides may be a target for degradation, or interfere with organelle biogenesis if accumulated.

Next, MUC assays were carried out at 37° C for a subset of the constructs. Since E1 is from a thermophilic bacterium, the native enzyme is most active at elevated temperatures. Indeed, we found about a 50% drop in endoglucanase activity in the MUC assay at the lower temperature compared to 65° C.

Taken together, these results demonstrate the specificity of the reaction carried out by E1 expressed in the in vitro eukaryotic (reticulocyte lysate) system. Based on these findings, we are in an excellent position to further optimize the reaction and explore the importance of different domains and features of E1 that might be necessary for its function. Most importantly, we are currently investigating whether E1 activity can be detected in chloroplasts after in vitro import, exploiting the different constructs we have already designed. Preliminary results indicate that indeed active E1 can be detected. These results may help to direct our in vivo studies geared to accumulating high levels of E1 in the chloroplast.

IN SUMMARY, in the last 18 months we have developed the tools to investigate if endoglucanase E1 can be sequestered in an active form within the chloroplast. We have utilized an in vitro approach that has served as a guide for transgenic plant studies, which are in progress. Populations of transgenic plants are nearly ready for a comprehensive biochemical analysis. Furthermore, we have developed a new in vitro expression system that can be used to study E1 activity after in vitro import. Hence, in the future, we can explore parameters that are important for enzyme activity, including determinants within the structure of E1 itself.

PUBLICATIONS AND PRESENTATIONS

Based on the results presented in our Progress Report, a manuscript has been written on the in vitro import of E1 into chloroplasts. This was submitted to Novartis for review per our Research Agreement on April 10, 2000. The manuscript has been submitted to a refereed journal for publication. I presented the experimental plan and goals of this project at the CPBR Symposium, March 1999. Rong Guan Jin presented our current results at a Division of Biological Sciences Annual Retreat at the University of Chicago, fall, 1999.

TECHNOLOGY TRANSFER

Research Agreement was finalized with Novartis Inc. and the University of Chicago (ARCH Development Inc).

COMMERCIAL ACCOMPLISHMENTS

None to-date.

EDUCATIONAL ACCOMPLISHMENTS

Training of three postdoctoral research associates in plant molecular biology and biochemistry.

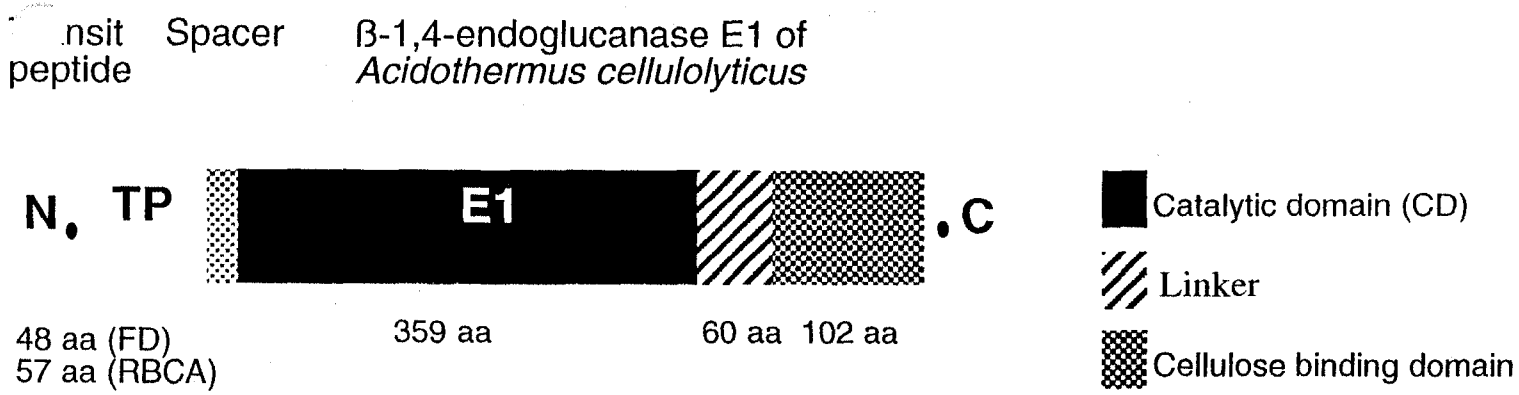
ADDITIONAL FUNDING

None for this project.

KEY PERSONNEL HIRING or TURNOVER

1) Dr. Rong Guan Jin joined my laboratory in January, 1999, to work on this project, 100% effort.
2) Dr. Stefan Richter, in my laboratory for 4 years, is training and working with Dr. Jin, 40% effort.
3) Dr. Rong Zong joined my laboratory in July, 1999, to work on this project, 60% effort.
All of these individuals are highly skilled scientists with backgrounds and expertise that complement one another. Dr. Richter is a microbiologist, molecular biologist and has received training in my laboratory in protein biochemistry. Dr. Jin is an expert in working with transgenic plants and tissue culture. Similarly, this is Dr. Zong's major area of expertise, although both of these individuals have research broad training.

A



B

	Construct	Transit Peptide-E1 Junction		Processing SPP[1)]	Processing Extr.[2)]	Import	Activity	Transgene started
1	E1		M AGGGYW-	N.T.	N.T.	N.T.	Yes	Yes
2	FD+1::E1	-RVTAM A.....	.. AGGGYW-	-	-	Poor	Yes	No
3	FD+5::E1	-RVTAM ATYKV.....	.. AGGGYW-	+	-	Good	Yes	Yes
4	FD+15::E1	-RVTAM ATYKVTLITKESGTV	.. AGGGYW-	++	+++	Very Good	Yes	Yes
5	FD+5::Δ5E1	-RVTAM ATYKV.....	.. ..W-	++	-	Good	No	No
6	FD+5::$_{SP10}$E1[3)]	-RVTAM ATYKV.....	$_{SP10}$AGGGYW-	++	++	V. Good	Yes	Yes
7	FD+1::E1CD	-RVTAM A.....	.. AGGGYW-	-	-	Poor	Yes	No
8	FD+5:: E1CD	-RVTAM ATYKV.....	.. AGGGYW-	++	-	V. Good	Yes	Yes
9	FD+5::Δ5 E1CD	-RVTAM ATYKV.....	.. ..W-	++	-	V. Good	No	No
10	RBCA+5::E1	-SMTVK AAENE.....	.. AGGGYW-	++	+++	Poor	Yes	No
11	RBCA+5:: E1CD	-SMTVK AAENE.....	.. AGGGYW-	+++	+++	Good	Yes	Yes
		Transit Peptide	Spacer[4)] E1					

[1)] Recombinant SPP; [2)] Chloroplast extract; [3)] SP10: 10 aa of C-terminus of E1 signal peptide; [4)] N-terminus of mature FD or RBCA

Figure 1. Expression constructs for accumulation of E1 in chloroplasts. (A) Schematic representation of the expression construct. (B) Characterization of E1 expression constructs. Relative amount of precursor processed: >75%, +++; 75--25%, ++; 25--5%, +; <5%, -. N.T., not tested; FD, ferredoxin; RBCA, Rubisco activase.

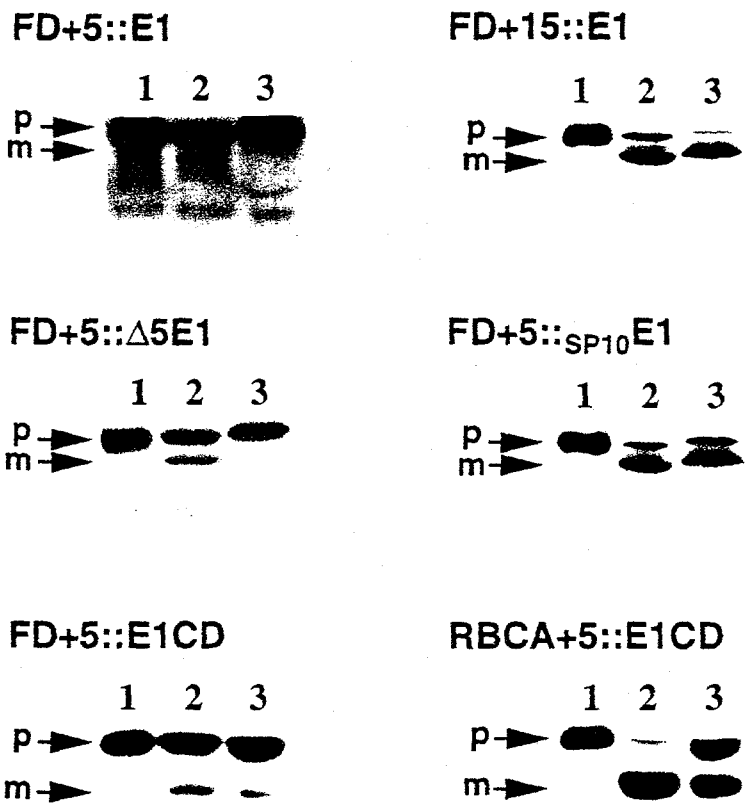


Figure 2. Processing of each E1 construct by recombinant SPP (Richter and Lamppa, 1998, *PNAS* 95, pp. 7463-7468) and a chloroplast extract from pea (Abad et al., 1989, *J. Cell Biol.* 90, pp. 117-124). The name of each construct is given at the top of each panel. The results are tabulated in Figure 1B. Lanes 1, [^{35}S]methionine-labeled precursor fusion proteins. Lanes 2, processing with recombinant SPP. Lanes 3, processing with chloroplast extract. p, precursor; m, mature E1 or E1CD.

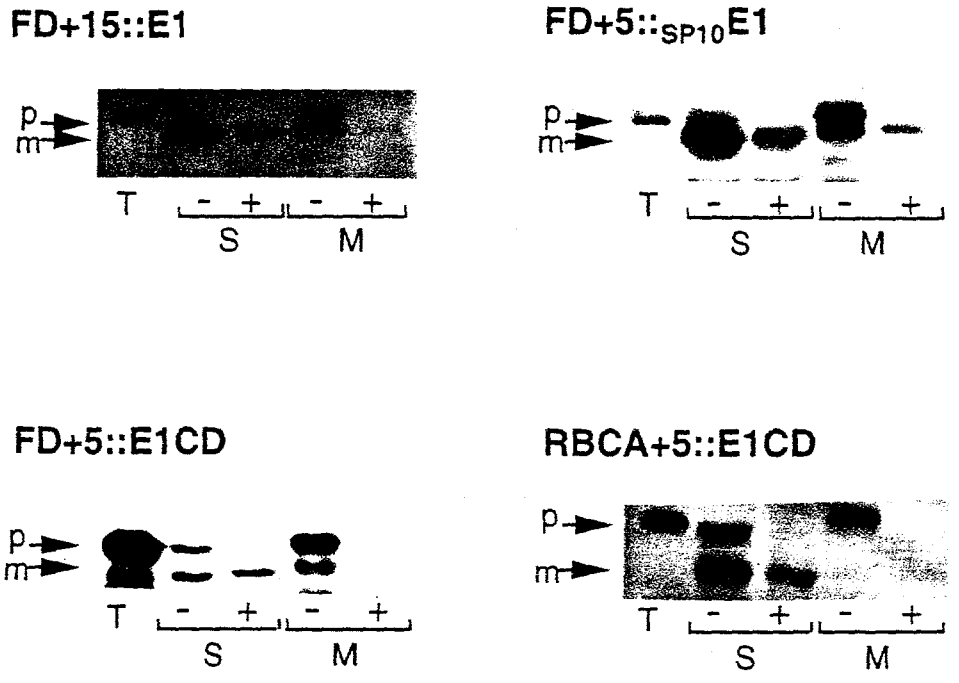


Figure 3. In vitro chloroplast import of E1 precursor fusion proteins. [^{35}S] methionine-labeled precursors were incubated with pea chloroplasts for 20 min, and then stromal and membrane fractions of chloroplasts with or without thermolysin treatment (indicated as -/+) were seperated and analyzed by SDS-PAGE (Lamppa, 1995, *Meth. Plant Mol. Biol.*, pp. 141-171). The name of each construct is given at the top of each panel. The results are included in Figure 1B. T, 1 µl translation product; S and M, stromal and membrane fraction, respectively; p, E1 fusion precursor; m, mature E1 or E1CD.

Table 1. Relative activities of E1 expression constructs in vitro[1].

Constructs	Relative amount of released MU					
	at 65 °C		at 65 °C upon SPP processing[2]		at 37 °C	
	Units[3]	%[4]	Units[3]	%[4]	Units[3]	%[5]
1 E1	54.98±0.34	100			25.14±0.59	46
3 FD+5::E1	46.76±0.73	85	57.53±0.61	105	22.36±0.41	48
4 FD+15::E1	41.60±0.62	76	54.66±1.36	99		
5 FD+5::Δ5E1	6.30±0.82	11	7.34±0.08	13		
6 FD+5::$_{SP10}$E1	50.11±0.72	91	58.87±0.77	107		
8 FD+5:: E1CD	37.00±0.16	67	41.92±0.69	76	15.37±0.20	41
9 FD+5::Δ5 E1CD	5.22±0.14	9	11.38±0.30	21		
10 RBCA+5::E1	51.13±0.41	93	65.48±0.34	119	23.50±0.39	36
11 RBCA+5:: E1CD	69.06±0.45	126	83.64±0.28	152	24.75±0.28	30
PreRBCA	1.00±0.09	2				
Control[6]	1.07±0.10	2				

Table 2. MU standard.

[MU] in nM	Units[3]
40.0	156.74±1.70
16.0	62.31±0.53
8.0	30.37±0.30
4.0	15.15±0.16
2.0	7.73±0.18
1.0	3.91±0.12
0.5	2.08±0.13
0.25	1.02±0.07
0.125	0.56±0.05

[1] Relative activity was examined by hydrolysis of the substrate 4-methylumbelliferyl ß-D-cellobioside (MUC) which liberates the fluorescent product 4-methylumbelliferone (MU)(MUC assay, protocol kindly provided by S.R. Thomas, National Research Energy Laboratories). All constructs are under T7 promoter control and were expressed in coupled transcription/translation reactions (50 µl standard reactions, TNT System, Promega). Translation product (5 µl) was added to 200 µl 0.5 mM MUC in reaction buffer (100 mM sodium chloride, 50 mM sodium acetate, pH 5) and incubated at 65 °C or 37 °C for 30 min. To stop a reaction, 200 µl of 150 mM glycine•NaOH, pH 10, were added. The relative amount of released MU was measured as fluorescence using 365 nm excitation and 456 nm emission filters. All translation products were also radiolabeled using [^{35}S]methionine in TNT standard reactions and quantified upon SDS-PAGE. A specific factor was calculated for each E1 construct and used to normalize fluorescence values.

[2] Translation products were processed by recombinant SPP before MUC assay (Richter and Lamppa, 1998, *PNAS* 95, pp. 7463-7468).

[3] Fluorescence units, means of five measurements.

[4] Relative amount of MU released by E1 was taken as 100%.

[5] Relative amount of MU released at 65 °C was taken as 100 %.

[6] Master mix of TNT System.

RECEIVED JUN 2 2 2000
email

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Dr. Scott Merkle	
University:	University of Georgia	
Agreement Number:	OR22072-72	
Project Title:	*Clonal Propagation of Hybrid Southern Hardwoods*	
Reporting Period and Report Type:	**From:** **01/01/00** **To:** **06/30/00**	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

Objective: To demonstrate the feasibility of using somatic embryogenesis for production of southern hardwood clones by generating embryogenic cultures from hybrid seeds of *L. tulipifera* x *L. chinense* and *L. styraciflua* x *L. formosana* and testing somatic seedlings derived from the hybrid cultures for their performance on intensively-managed southeastern U.S. Coastal Plain and/or Piedmont sites.

Specific Aim 1: Conduct controlled pollinations of *L. tulipifera* x *L. chinense* , *L. styraciflua* x *L. formosana,* and crosses within the North American parent species, collect hybrid and nonhybrid seeds and initiate embryogenic cultures from them.

From the 1100 hybrid *Liquidambar* and 1000 hybrid *Liriodendron* seeds cultured during summer, 1999, 21 embryogenic *Liquidambar* and 6 embryogenic *Liriodendron* cultures were derived.

Specific Aim 2: Use RAPD markers to verify hybrid genotypes and fingerprint hybrid clones.

Using RAPDs markers with DNA from leaves of the *Liquidambar styraciflua* and *L. formosana* parents and from the putative hybrid embryogenic cultures, we have evidence that 4 of the 5 cultures tested so far have hybrid genotypes. Other putative hybrids are in the process of being tested.

Specific Aim 3: Optimize cryopreservation protocols for long-term storage of embryogenic cultures.

A cryopreservation protocol optimized for *Liriodendron tulipifera* and *Liquidambar styraciflua* embryogenic cultures is currently being tested with the new hybrid *Liriodendron* and *Liquidambar* cultures. We have no reason to expect that our protocol will not work with the hybrid cultures. We have submitted for publication a manuscript describing the results of our cryopreservation experiments.

Specific Aim 4: Select prolific embryogenic cultures from each hybrid and parental species, and generate somatic embryo-derived somatic seedlings from the selected hybrid and parental lines.

We have selected 6 hybrid *Liquidambar* cultures and already regenerated over 200 somatic seedlings from them that are ready to be transferred to the greenhouse. Other hybrid *Liquidambar* cultures are being bulked up form somatic seedling production now. Somatic embryos have been produced from 3 of the hybrid *Liriodendron* cultures, and about 100 of these somatic embryos have germinated, although none of these are in soil yet.

Specific Aim 5: Install replicated field tests of hybrid and parental species somatic seedlings and score them for survival and growth in comparison to each other and to *L. styraciflua* and *L. tulipifera* seedlings.

Field test planning has begun in collaboration with International Paper Co. cooperators.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

Since our December, 1999 Progress Report for our application for Year 2 funding, we have accomplished the following objectives: (1) Generated at over 20 putative hybrid *Liquidambar* embryogenic cultures with demonstrated somatic embryo-producing ability. (2) Regenerated over 200 somatic seedlings representing 6 of the putative hybrid *Liquidambar* lines, (3) Generated 6 putative hybrid *Liriodendron* embryogenic cultures, 3 of which have produced embryos and somatic seedlings, (4) Used DNA markers to verify that four of the *Liquidambar* cultures tested to date have hybrid genotypes, and (5) Submitted a manuscript for publication detailing our proven protocols for cryostoring and recovering *Liriodendron* and *Liquidambar* embryogenic cultures.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Since our December, 1999 Progress Report, we have concentrated our efforts in four areas: (1) Maintaining and bulking-up putative hybrid *Liquidambar* and hybrid *Liriodendron* cultures initiated during Summer, 1999, (2) Testing putative hybrid cultures for somatic embryo and somatic seedling production, (3) Genotyping putative hybrid cultures using RAPD markers, and (4) Cryostoring hybrid cultures.

Maintenance and bulking-up of putative hybrid *Liriodendron* and hybrid *Liquidambar* cultures. Details of breeding of *L. styraciflua* x *L. formosana* and *L. tulipifera* x *L. chinense* , followed by culture initiation from the resulting seeds, are described in our December, 1999 Progress Report. Briefly, for hybrid *Liquidambar* breeding, pollen from three *L. formosana* trees growing in a USDA Forest Service test planting in Saucier, MS, was used to pollinate pistillate

flowers on four *L. styraciflua* trees growing at International Paper Company's Southlands Experiment Forest in Bainbridge, GA, during March, 1999. A total of 72 crosses were conducted in a half-diallel design whereby all three *L. formosana* trees served as male parents and all four *L. styraciflua* parents served as female parents. Fruits were collected from the female parents and culture initiation from the hybrid *Liquidambar* seeds began on June 17, 1999. Two induction media were tested: (1) Yellow-poplar induction/ maintenance (IM) medium (Merkle and Sommer 1986) and Lloyd and McCown's (1980) woody plant medium (WPM). Both media were supplemented with 2 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D) and 0.25 mg/l benzyladenine (BA). In all, over 1100 explants were cultured, representing 23 full-sib (hybrid) families. To date, 21 embryogenic cultures (i.e. observed to actually produce somatic embryos) representing 6 crosses have been generated. No statistically significant differences in culture initiation frequency were found between the two basal media. Cultures are being maintained by monthly transfer to fresh medium of the same formulation as that on which they were initiated. In March, 2000, we began to subdivide the cultures onto more plates in order to bulk-up the material for each clone in preparation for somatic seedling production for field test establishment (see below).

For hybrid *Liriodendron* breeding, pollen from five *L. chinense* parents was used to pollinate ten *L. tulipifera* trees in the Chapel Hill, NC area during April and May, 1999, using a partial diallel design. Over 200 crosses were performed. Culture initiation from hybrid *Liriodendron* seeds began on June 30, 1999. Embryos and endosperm were cultured together on Petri plates of the same media used for the *Liquidambar* cultures. Over 1000 explants were cultured representing 24 full-sib (hybrid) families. Only 6 embryogenic cultures were produced from these explants, and these are being maintained by monthly transfer to fresh medium of the same formulation as that on which they were initiated.

Testing of putative hybrid cultures for somatic embryo and somatic seedling production. In January, 2000, six of the fastest-growing hybrid *Liquidambar* cultures were selected and material was transferred to basal media for somatic embryo production. Somatic embryos were germinated on basal medium lacking casein hydrolysate, and germinants were potted in Fafard #3 mix and grown in a humidifying chamber. Over 200 somatic seedlings representing the 6 hybrid clones have been grown and are ready for transfer to the greenhouse. Leaf morphology of the hybrid somatic seedlings resembles that of hybrid trees produced by Santamour (1972). In preparation for field testing, material from all of our hybrid *Liquidambar* cultures was transferred to basal medium for somatic embryo production in May, 2000. We expect to harvest embryos and germinate them in early July, 2000, and to harden off the somatic seedlings and move them to the greenhouse in August, 2000, where they will be grown for the remainder of the season.

Suspension cultures of our few hybrid *Liriodendron* cultures have been initiated in liquid IM medium, and experiments to test various cultural treatments aimed at improving somatic embryo production are being conducted. The two variables currently under investigation are 2,4-D concentration and pH. Preliminary results indicated that lowering initial medium pH to 3.5 (from the standard 5.6) improved somatic embryo production following fractionation and plating of the embryogenic suspensions. To date, we have regenerated over 80 somatic seedlings representing 3 putative hybrid *Liriodendron* cultures. However, this level of somatic embryo production from so few hybrid *Liriodendron* cultures is probably too low to supply somatic seedlings for field testing this year.

Genotyping putative hybrid cultures using RAPD markers. DNA was extracted from expanding leaves of the three *L. formosana* parents, from the four *L. styraciflua* parents and from putative hybrid embryogenic cultures using a Qiagen DNeasy™ Plant Mini Kit. DNA yields averaged 13 µg/100 mg tissue for leaves and 11 µg/100 mg tissue for embryogenic culture material. PCR primers were obtained from Operon. We used the reaction mixture and PCR cycle parameters recommended by Operon. It was determined that 50 ng of DNA from leaves or embryogenic cultures was sufficient to obtain repeatable banding patterns. To date, 4 primers have been tested using 5 putative hybrid lines and the parents of each. One of these primers revealed

consistent polymorphisms between the *L. formosana* and *L. styraciflua* parents, including a band that was always present in *L. formosana* (male) parents and never present in the *L. styraciflua* (female) parents. This band was present in all of the putative hybrids, indicating the that genotypes of all four clones for which banding patterns have been obtained are indeed hybrids (one of the five hybrids tested to date has not given any banding pattern).

Cryostorage of hybrid cultures. As detailed in our December, 1999 Progress Report, we demonstrated that a cryopreservation protocol, modified from one published by Hargreaves and Smith (1992) for cyropreserving embryogenic Pinus radiata cultures, provided 100% recovery of *L. tulipifera* and *L. styraciflua* cultures following storage in liquid nitrogen (-196 C). Briefly, the osmotic and cryoprotectant treatments that gave full recovery were 0.4 M sorbitol with 5% DMSO and 0.4 M sorbitol with 10% DMSO. We have prepared and submitted for publication a manuscript reporting these results (Vendrame et al., submitted). Currently, we are testing the same protocol with the new hybrid *Liquidambar* and hybrid *Liriodendron* cultures. Once we have demonstrated recovery of the tested hybrid cultures using the protocol, we will cryostore at least 10 samples of each hybrid embryogenic culture.

Literature Cited

Hargreaves, C., and D.R. Smith. 1992. Cryopreservation of *Pinus radiata* embryogenic tissue. Comb. Proc. Int. Plant Prop. Soc. 42:327-333.

Lloyd G., and B. McCown. 1980. Commercially feasible micropropagation of mountain laurel, Kalmia latifolia, by use of shoot-tip culture. Comb. Proc. Int. Plant Prop. Soc. 30:421-427.

Merkle, S.A. and H.E. Sommer. 1986. Somatic embryogenesis in tissue cultures *of Liriodendron tulipifera* L. Can. J. For. Res. 16:420-422.

Santamour, F.S., Jr. 1972a. Interspecific hybridization in *Liquidambar*. For. Sci. 18(1): 23-26.

Vendrame, W.A., C.P. Holliday, P.M. Montello, D.R. Smith and S.A. Merkle. Cryopreservation of yellow-poplar (*Liriodendron tulipifera*) and sweetgum (*Liquidambar* spp.) embryogenic cultures (submitted to New Forests).

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Vendrame, W.A., C.P. Holliday, D.R. Smith and S.A. Merkle. 2000. Optimization of a cryopreservation protocol for embryogenic cultures of yellow-poplar (*Liriodendron tulipifera*). Proceedings of the 25th Biennial Southern Forest Tree Improvement Conference, July 11-14, 1999, New Orleans, LA, pp. 172-173.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

None

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

None

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

We have submitted the following grant proposal to request funding for the field test phase of the project, but have not received word on an award:

Merkle, S.A., and W.A. Vendrame. Field Performance of In Vitro-Derived Clones of Hybrid Southern Hardwoods. Submitted to 2000 National Hardwood Lumber Association Research Grant Program. July 1, 2000 - June 30, 2001. $60,000.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Send completed report to :

The Consortium for Plant Biotechnology Research, Inc.
P.O. Box 20634
(Express Delivery address: 10 Sylvan Drive, Suite 21)
St. Simons Island, GA 31522
Phone: 912.638.4900 Fax: 912.638.7788

Or sent as an email attachement to: cpbr@gate.net

RECEIVED SEP 1 1 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Basil Nikolau	
University:	Iowa State University	
Agreement Number:	OR22072-73	
Project Title:	*How Do Plants Generate Acetyl-CoA?*	
Reporting Period and Report Type:	**From:** **1/1/00** **To:** **6/31/00**	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

Genetically reduce the accumulation of specific enzymes required for acetyl-CoA generation

Determine the effect of each genetic alteration on:

the accumulation of acetyl-CoA-derived phytochemicals
the expression of acetyl-CoA-generation genes
***in vivo* measurements of the generation of different acetyl-CoA pools**

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

Our specific goals are aimed at identifying the mechanisms by which two physically distinct pools of acetyl-CoA are synthesized in plants. These two pools, one in the plastids and the other in the cytosol, are the precursors for the synthesis of a large number of plant chemicals that have utility in nutritional and industrial applications. For example, from the plastidic acetyl-CoA pool plants produce fatty acids, and hence seed oils; from the cytosolic acetyl-CoA pool plants

produce pigments, waxes and many compounds that plants use for defense from biotic and abiotic stresses.

We are ascertaining the role of four enzymes in generating the plastidic acetyl-CoA pool. These enzymes are: pyruvate decarboxylase (PDC), acetaldhyde dehydrogenase (ALDH), acetyl-CoA synthetase (ACS) and plastidic pyruvate dehydrogenase complex (pPDHC). In addition, we are studying the role of ATP-citrate lyase (ACL) in generating the cytosolic acetyl-CoA pool.

We have isolated genes and cDNAs that code for each of these enzymes; these are essential reagents for achieving our goals. By using cDNA probes in *in situ* hybridization experiments that allow us to measure the level of gene expression in individual cells or tissues, we were able to show that one of the possible enzymes for acetyl CoA synthesis, pPDHC, was more important for fatty acid formation in seeds than the other potential enzyme, ACS. We demonstrated that pPDHC gene expression was greatest at the same time in seed development when maximum fatty acid synthesis occurred. ACS gene expression was optimal in seeds at a time when seed oil accumulation was not significant. From these experiments we now know, for the first time, which enzyme is most important in fatty acid synthesis in seeds and have a target for engineering plants with increased capacity for making lipids.

Similar experiments indicate that ACS may be important for the synthesis of acetyl-CoA in other parts of the plant (e.g., leaves, flowers). If ACS is a physiological mechanism for the synthesis of an acetyl-CoA pool, it probably does so in coordination with PDC and ALDH.

ACL on the other hand is expressed at different times in the development of the plant, than pPDHC and ACS. In fact ACL expression mirrors the metabolic demands for the cytosolic acetyl-CoA pool.

To directly test the hypotheses generated by the above findings we are generating transgenic plants that show reduced PDC, ALDH, ACS, pPDHC and ACL expression. These genetic manipulations are being undertaken by the expression of specific antisense RNAs.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

The specific scientific achievements are described below in reference to each of the acetyl-CoA-generating enzyme:

ATP-citrate lyase: Our previous molecular research indicated that the plant ACL is composed of two distinct subunits, which we term ACL-A and ACL-B. Based upon additional cDNA cloning and sequencing experiments it has become clear that each subunit is coded by a small gene family of at least two members each. We now have nucleotide sequences for two members of the *ACL-A* gene family, which we term *ACL-A1* and *ACL-A2*, and two members of the *ACL-B* gene family, which we term *ACL-B1* and *ACL-B2*. The proteins coded by each member of each gene family are very similar (ACL-A1 and ACL-A2 are over 95% identical; and ACL-B1 and ACL-B2 are over 95% identical), but the UTRs of each of these genes are distinct.

We have produced recombinant ACL-A1 and ACL-B1 proteins in *E. coli* for the purposes of generating antibodies for each subunit. Optimal conditions for using each of the resulting antibodies have been devised. Using these ACL-A- and ACL-B-specific antibodies we have determined that the ACL enzyme is localized in the cytosol of plant cells. In addition, we have directly shown that the plant ACL enzyme is a heteromeric enzyme, composed of ACL-A and ACL-B subunits. This was demonstrated by the fact that ACL activity co-purifies with both the ACL-A and ACL-B subunits.

We have generated transgenic Arabidopsis plants in which ACL accumulation is altered by the expression of ACL antisense RNAs. We made transgene constructs in which *ACL-A1* or *ACL-B2* cDNAs are fused to the CaMV 35S promoter in opposite orientation from normal. The resulting transgenes were introduced into the *Arabidopsis* genome by *Agrobacterium*-mediated transformation, and transformants were selected on the basis of resistance to kanamycin, a trait physically linked to each transgene. We obtained over 200 independent transgenic lines, as indicated by kanamycin resistance in the T1 generation and substantiated by further testing in the T2 generation. Over 50 independent lines were selfed, and selected T2 lines were propagated to the T3 and T4 generations.

The focus of the past six months was the characterization of these transgenic lines. In particular, the phenotypes of the plant lines at multiple developmental stages has been characterized . A subset of the transgenic lines either do not germinate or germinate but grow very slowly (final rosette diameter ranging from about 2mm to 0.75 cm) and die prior to bolting. Among the plants that survive and bare progeny, plants that present the most severe phenotype show a montage of morphological alterations, including, much reduced plant body size, shorter and much thinner inflorescence stalks, smaller (or non-opening) flowers, reduced and occasionally early senescing petals, and anthers appear to have a problem with timing of or mechanisms of dehiscence. If siliques are present on these plants, they are much reduced in size, partially filled with seed or empty. The seeds within siliques are often smaller, and appear desiccated. Dehiscence within filled siliques is often premature (i.e., siliques will dehisce when half of the seed is mature and the other half is green). Anthers have a deficiency in timing or mechanisms of dehiscence. Leaves are reduced in size. Seed flavonoids are decreased, as indicated by the reduction of phlobaphens in the seed coat. The accumulation of cuticular waxes on the stems is greatly reduced or not detectable.

Plants with less severe phenotypes are smaller, often with a milder version of the above alterations. A common occurrence is curled siliques. These siliques often contain shriveled 'dry' seed. Silique dehiscence is delayed, even when completely dry and the seeds are very difficult to remove from silique chambers. The curled siliques are thinner than wild type and seed outlines are more easily differentiated from the outside of the silique.

ACL activity has been determined in T2-generation antisense ACL-plants that show a range of morphological phenotypes. There is an indirect relationship between the severity of the morphological phenotype and the ACL-activity that is retained in these plants. Interestingly, as little as a 50% reduction in ACL activity gives a severe phenotype, indicating the level of *ACL* expression is critical to normal growth and development.

In addition to the above transgenic alteration of *ACL* expression, we have begun to screen T-DNA tagged Arabidopsis populations for individual knock-outs of the *ACL-A1*, *ACL-A2*, *ACL-*

B1 and *ACL-B2* genes. This work is being conducted in collaboration with the University of Wisconsin Arabidopsis Knockout Facility (http://www.biotech.wisc.edu/NewServicesAndResearch/Arabidopsis/default.htm)

Finally, to investigate the expression patterns of each of the ACL subunit genes, promoter-*GUS* transgenes have been prepared, and these have been transformed into Arabidopsis. Currently, we are propagating genetic stocks that carry *ACL-A1-GUS*, *ACL-A2-GUS* and *ACL-B1-GUS* transgenes.

Plastidic pyruvate dehydrogenase complex: The pPDHC is composed of four different subunits (E1_, E1_, E2, and E3). In a previous progress report we outlined our development of molecular reagents for analyzing these subunits. These reagents include cDNA clones and polyclonal antisera for the plastidic E1_ subunit and the E2 subunit.

We are altering the accumulation of the pPDHC in transgenic *Arabidopsis* plants by the expression of the E2 cDNAs in the antisense configuration. Transgenic plants are currently being propagated, and will be analyzed once we have reached the homozygous T3 generation.

Acetyl-CoA synthetase: We have generated a series of transgenic *Arabidopsis* plants with *ACS* activity that ranges from 5% to 500% of wild-type values. The amount of *in vivo* ACS activity as measured by the incorporation of ^{14}C-acetate into lipids is directly correlated with the ACS activity measured *in vitro*. We have looked at the physiological effects of altering the acetyl CoA levels. Surprisingly, despite the tremendous differences in ACS activity there is no measurable differences in size, vigor, and developmental schedules for the mutant plants. An analysis of the fatty acid content and composition showed no differences between the mutant and wild-type plants. Similarly, photosynthetic rates and the rates of total lipid and fatty acid formation from photosynthetically fixed carbon dioxide was not different between the mutant and wild-type plants.

We conclude from these studies that acetyl-CoA formation for fatty acid synthesis is not the major role of ACS. What ever the role of this enzyme it is either not responsible for the synthesis of important biomolecules under our present growth conditions or its function within the plastid can be completely usurped by pyruvate dehydrogenase.

Acetaldehyde dehydrogenase (ALDH): To prove that *Arabidopsis* cDNA clones, which we are characterizing (*ataldh1* and *ataldh2*) code for ALDH enzymes, genetic complementation tests were performed in an *E.coli* strain (JA111) that carries a mutation in an ALDH gene (*ald9*). A construct consisting of the *ataldh1* cDNA cloned into an expression vector (pET17b) without its mitochondrial targeting peptide, failed to complement the JA111. In contrast, the *ataldh2* cDNA did complemented JA111. Crude extracts of *E. coli* expressing either AtALDH1 and AtALDH2 were assayed for ALDH activity. Both extracts exhibit ALDH activity when assayed with either acetaldehyde or glycoglutaldehyde as the substrates.

Gene-specific primers were designed from the genomic sequences of *ataldh1* and *ataldh3*. These primers will be used to screen for insertion mutants among the T-DNA lines in the Arabidopsis Knockout Facility at University of Wisconsin Biotechnology Center. Both primers were extensively tested for specificity and have been sent to the Arabidopsis Knockout Facility.

Since there is a high degree of similarity and identity between the deduced AtALDH1, AtALDH2 and AtALDH3 protein sequences, gene-specific peptides were synthesized and will be used to raise antibodies for subsequent immunolocalization experiments.

Pyruvate decarboxylase (PDC): Probes derived from the 5'ends of the *pdc1* and *pdc2* genes were hybridized to Southern blots containing Landsberg and Columbia genomic DNA digested with various restriction enzymes. These DNA gel blots revealed that the *Arabidopsis* genome contains only these two *pdc* genes.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Ke, J./Behal, RH, Yunkers, S., Nikolau, B.J., Wurtele, E.S., and Oliver, D.J. 2000. The role of pyruvate dehydrogenase and acetyl-CoA synthetase in fatty acid synthesis in developing Arabidopsis seeds. Plant Physiology 123: 497-508

Nikolau, B.J., Wurtele, E.S., Oliver, D.J., P.S. Schnable. 2000. Molecular biology of acetyl-CoA metabolism. *In* The Proceedings of the 14th International Symposium on Plant Lipids, Cardiff, Wales. In press.

Fatland, B., M. Anderson, B.J. Nikolau and E.S. Wurtele. 2000. Molecular biology of cytosolic acetyl-CoA generation. *In* The Proceedings of the 14th International Symposium on Plant Lipids, Cardiff, Wales. In press.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

International Patent Application No. PCT/US99/14382. (Claiming Priority to U.S.S.N. 60/090,717) "Materials and Methods for the Alteration of Acetyl-CoA Levels in Plants"

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Research undertaken in conjunction with this grant is providing opportunities to educate and train undergraduate and graduate students in research. In addition, this grant is providing continued research training of post-doctoral research associates. This grant is providing multidisciplinary training in the areas of plant molecular biology, biochemistry, genetics with the aim of solving a complex problem in plant metabolism.

Specifically, this grant is supporting the following research based degrees:

Stephanie Back - M.S. 2000 completed - Thesis title: The Enzymatic and Functional Characterization of Acetyl Coenzyme A synthetase

Beth Fatland, Ph.D., plant physiology Iowa State University, expected completion 2002. Thesis title: The role of ATP-citrate lyase in generating acetyl-CoA in plant cells.

Li Wei Cui, candidate for a PhD in Biochemistry, Iowa State University, expected completion 2003.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

National Science Foundation, REU fellowship (summer 2000) to Lashae (Michal) Stallworth, undergraduate summer student from Jackson State University. Total amount: $4,000.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Dr. Marc Anderson, who was post-doctoral research associate under the direct supervision of Wurtele, has left the project and has taken a position of Assistant Professor at North Dakota State University.

Wei Huang, was a research assistant under the direct supervision of Nikolau, has left the project and is currently in a graduate program in Statistics at Iowa State University.

RECEIVED APR 1 9 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	John B. Ohlrogge Christoph Benning	
University:	Michigan State University	
Agreement Number:	OR22072-98	
Project Title:	DNA microarray discovery of genes and networks which control plant storage products	
Reporting Period and Report Type:	**From:** **9/30/1999** **To:** **3/31/2000**	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

1. Partial sequencing of cDNAs from an *A. thaliana* seed library.
   Previously accomplished.

2. Identification of *Arabidopsis* genes specifically expressed in developing seeds.

3. Global expression analysis in selected mutants and transgenic plants.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

To provide a broad analysis of gene expression in developing *Arabidopsis* seeds, microarrays have been produced which display approximately 2600 seed expressed genes. DNA for genes spotted on the arrays were selected from >10,000 clones partially sequenced from a cDNA library of developing seeds. Based on a series of controls, sensitivity of the arrays was estimated at 1-5 copies mRNA per cell and cross-hybridization was estimated to occur if closely related genes have >70-80% sequence identity. These arrays have been hybridized in a series of experiments with probes derived from seeds, leaves, and roots of *Arabidopsis*. Analysis of expression ratios between the different tissues has allowed the tissue-specific expression patterns of many hundreds of genes to be described for the first time. Approximately 25% of the

2600 genes were expressed at ratios ≥2 fold higher in seeds than leaves or roots and 10% at ratios ≥10. Included in this list are a large number of proteins of unknown function, and potential regulatory factors such as protein kinases, phosphatases, and transcription factors. The *Arabidopsis* arrays were also found to be useful for analysis of mRNA isolated from developing *Brassica napus* seeds and expression patterns correlated well between the two species.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

RESULTS

A Microarray from Developing *Arabidopsis* Seeds

From a cDNA library of developing Arabidopsis seeds 27,648 clones were arrayed on filters and hybridized with probes specific for highly-abundant transcripts (such as storage proteins) in *Arabidopsis* seeds. Over 10,000 clones, which showed no signal in this subtractive screening, were partially sequenced from their 5` ends. Subsequent BLASTX and contig analysis condensed the number of these expressed sequence tags (ESTs) down to about 5,000 putative unique sequences. About 30% of these sequences were not represented in the public *Arabidopsis* EST database (dbEST) as of Oct 1999, and >45% of these sequences had no significant similarity (BLAST score <100) to the entries in the Genbank protein data base. This high amount of potentially new sequences in part reflects the lack of study of *Arabidopsis* seeds by EST approaches and emphasizes the value of this cDNA set as an interesting source to discover novel gene functions. A more complete description of the generation of the seed-specific cDNA library, the sequencing project and its analysis is given in White *et al.* (2000).

For microarray fabrication a subset of 2,715 clones was selected from the 5,000 putative unique sequences. Several of these ESTs were very similar and are likely to represent the same gene. The number of unique genes represented on the arrays is therefore approximately 2600. To monitor the expression pattern from as many genes involved in glycerolipid and carbohydrate metabolism as possible, 82 additional cDNA clones were collected which complemented the seed microarrays with most of the missing sequences from these pathways. In addition, a collection of 60 control DNAs was generated. The inserts of the three clone collections were amplified by PCR with vector specific primers. PCR samples which yielded less than 0.2 µg/µl DNA or showed several DNA fragments were re-amplified or replaced with alternative clones. The PCR products were arrayed on and bound to coated microscope slides. To increase the reliability of the detected signals, each PCR sample was spotted twice in two subarrays resulting in a total array of 7680 data points. The identity of 12 randomly chosen DNA samples was confirmed by re_sequencing their PCR products used for microarray printing and comparing the obtained sequence results with the corresponding EST sequences in our database. In all 12 cases the sequences of the PCR samples matched with their original EST sequence. This sequence confirmation increases the confidence in the identity of the DNA elements on our microarrays and makes is unlikely that major errors like frame shifts in the data lists or in the sample plates occurred during sample preparation. Additional details of the microarray results from this study are available online at: http://www.bpp.msu.edu/lgc/johnweb5.htm

Data Evaluation

To evaluate the reliability of the hybridization experiments, the microarrays contained several control elements. To detect the sensitivity limit and to have an additional control for balancing the intensities of the two channels, nine non-related human cDNA fragments were arrayed on the slides and the corresponding *in vitro* transcribed poly(A)$^+$ RNA species were added to 1.0 µg of the tissue mRNA samples as internal standards in decreasing concentrations from 1.0 ng (1:1.0x10^{-3}) to 0.01 ng (1:1.0x10^{-5}) (Figure 1). The lowest RNA concentrations of 7.5x10^{-4} and 1.0x10^{-5} gave in most experiments fluorescence signal intensities (FSI units) higher than two times the local background. Similar detection limits of 1.0x10^{-5} (Ruan *et al.*, 1998) and 5.0x10^{-5} (Schena *et al.*, 1996) were detected from other groups. According to mRNA

quantifications from Okamuro and Goldberg (1989), this detection limit corresponds to approximately 1-5 mRNA copies per cell.

Cross-hybridizations between different members of gene families are an issue in cDNA based microarray experiments and many genes involved in lipid metabolism are members of gene families (Mehkedov et al. 2000). To estimate the extent of such non-specific hybridizations, the cross-hybridization threshold of each experiment was detected with several specificity controls. These controls included synthetic gene fragments and heterologous sequences from other plant species, which have decreasing sequence identities of 100-60% to three moderately expressed *Arabidopsis* genes. First, we synthesized and arrayed 365bp synthetic fragments of the *Arabidopsis FAD2* gene in three different forms of identical length and constant GC content of 48%, but decreasing nucleotide identities of 100%, 90% and 80%. As shown in Figure 1 and 2, the 100% fragment gave comparably strong signals (within 80-90%) to a ~1.1 kbp PCR fragment from *FAD2* indicating that a target length of 365bp is sufficient for an efficient probe binding in this technique. The 90% identity fragment gave 50% weaker signals compared to the 100% form, whereas the 80% form showed almost no detectable signals suggesting a cross-hybridization threshold under the conditions of these experiments between 80-90% identity. Cross-reactions with other *Arabidopsis* transcripts are unlikely, because for *Arabidopsis* no genes are known which are closely related (>60%) to *FAD2* (Okuley *et al.*, 1994). The synthetic gene fragments were designed with evenly spaced mismatches. Two other specificity control sets consisted of four ferredoxin sequences and three acyl-ACP-desaturase sequences from other organisms. These contain more variable similarity clusters to the *Arabidopsis* sequences than the synthetic *FAD2* fragments, and showed cross-hybridization thresholds between 60-70%. Based on these experiments it is clear that some closely related gene family members will not be discriminated. However, with complete availability of the *Arabidopsis* genome it is possible to assess the approximate extent of potential cross-hybridization. For example, most of the seven known *Arabidopsis* ACP genes are less than 70% identical and unlikely to cross-hybridize, whereas four of the five members of the stearoyl-ACP desaturase family are >80% identical (Mehkedov et al, 2000). Additional controls, shown in Figure 2 and as described in the legend and in "METHODS" monitored for, non-specific hybridization, carry-over during printing and for mRNA integrity/probe length.

Microarray Hybridizations

To monitor seed-specific gene expressions, mRNA samples from seeds, leaves and roots of *Arabidopsis* were isolated, and reverse transcribed with oligo-dT primers into first strand cDNA fluorescent probes. The mRNA isolated from seeds was the reference to which the samples from leaves and roots were compared. Each tissue comparison was performed at least two times using in most cases independently isolated RNA samples as starting material. For repeated experiments, the probe pairs contained the fluorochromes Cy3 and Cy5 in opposite orientation. Results of repeated experiments were only used for further analyses, if the ratios of all data points on the array showed a correlation coefficient close to one. To eliminate highly variable and therefore less reliable expression data, we used data for further analysis only if at least two experiments showed the same trend of expression. Averaging ratios across experiments was considered a less stringent strategy, because it neglects the variability between measurements. This is particularly true when low tissue mass (as with developing *Arabidopsis* seeds) is a limitation for the number of feasible experiments. For the experiments described here, over 20 hours of dissection of developing seeds from siliques was required to harvest material for a single fluorescent probe.

A scatter plot of the data for a seed vs leaf comparisons is shown in Figure 3. It is clear from this representation that the majority of genes analyzed fall near the X-axis and have less than a two fold difference in signal intensity between the leaf and seed probes. Thus, although the microarray was based on a set of ESTs primarily derived from sequencing of a seed cDNA library, the overall expression pattern shown in Fig 3 clearly indicates that a large proportion of seed expressed genes are also expressed in other tissues. These data support the general conclusion which was previously based on COT analysis of RNA complexity that 60-77% (the majority) of plant genes do not have strong tissue-specific expression (Okamura and Goldberg, 1989; Kamalay & Goldberg, 1980). Expression analyses with smaller and non-seed specific

arrays from *Arabidopsis* detected comparable amounts of tissue specific (Ruan *et al.*, 1998) or differentially expressed genes (Desprez *et al.*, 1998; Kehoe *et al.*, 1999; Richmond & Somerville, 2000).

Nevertheless, the microarrays reveal that a substantial number of genes can be considered seed-specific. In the seed versus leaf co-hybridizations, approximately 30% of the spotted cDNAs showed more than 2 fold stronger signals in seeds and approximately 12% gave a more than 10 fold higher expression in seeds than in leaves (Table I). In the corresponding seed versus root experiments similar comparisons yielded 33% and 13% of the genes, respectively. If both tissue comparisons are combined, 25% of genes showed more than 2 fold and 10% more than 10 fold stronger signals in seeds than in leaves or roots. One factor should be noted which influences these numbers. The reliability of the signals used to calculate these ratios was ensured by including only those values which showed fluorescent intensity levels in at least one channel above three times the local background. This high signal to noise ratio and the stringent limit for the ratios of more than two fold in each experiment of both tissue comparisons selects preferentially for genes which are moderate to strongly expressed in seeds and only to a very low extent in the other tissues. A disadvantage of this sorting for high confidence values is its tendency to disregard weakly expressed genes, which generally do not reach a high and stable enough signal to background ratio in several experiments to appear in this list.

Characteristics of the Seed-expressed Set

The set of highly seed-specifically expressed sequences (ratio ≥4) contains several seed storage proteins, and a number of other genes which are well known to be predominantly seed expressed. These include oleosins (Abell et al, 1997), fatty acid elongase (*FAE1*) (James et al. 1995), lipoxygenase (Fauconnier *et al.*, 1995), and other genes. Similarly, our arrays included a number of genes involved in photosynthesis and carbon fixation such as chlorophyll a/b binding protein and the small subunit of RuBisCo. These and other related photosynthetic genes were found to be expressed preferentially in leaves. Thus, the overall reliability of the microarrays was confirmed by obtaining the expected preferential seed or leaf expression patterns for dozens of well characterized genes. The tissue-expression ratios for several of these genes and the variability observed in replicated experiments is presented in Table II.

We previously classified the seed-expressed ESTs according to codes which catogorize their putative function (White et al., 2000). Table III presents a partial summary of the microarray analysis of groups of clones from several categories. Of the 113 genes included on the microarrays which are related to lipid biosynthesis, only 10 were found to occur in the subset with ≥10 fold higher seed vs leaf or root signals. These numbers reflect the fact that lipid biosynthesis is essential for growth of all tissues, and can be considered a "housekeeping" function. The 10 lipid related genes with high seed to leaf/root expression ratios include oleosin, FAE1, and lipases.

Approximately 15 cDNAs with homology to transcription factors, kinases, phosphatases and proteins involved in development were highly seed-specific. About 145 of the ≥10 fold subset, - more than 50% - show no significant homology to known sequences (BLAST score <100). Since the sequences of most structural genes are known, it is likely that this set of sequences with unknown function contains many additional regulatory genes.

Identification of New Strong Seed-Specific Promoters

Because EST abundance is in most cases related to mRNA abundance, the sequencing of >10,000 ESTs from a seed cDNA library has provided a set of data which can be used to identify highly expressed genes (White *et al.*, 2000). Microarray data as described here provides additional information on tissue specificity of gene expression. By combining these two types of data, it is possible to identify genes which are both strongly expressed, and expressed with high tissue-specificity. Of course many seed storage proteins and other genes are well known to fall into this category. In Table IV, we have identified a number of additional such candidates which have both high EST abundance and high seed-specificity based on microarrays. Most of these are proteins of unidentified function and therefore may be of particular interest in future functional genomic studies of seed metabolism and development. In

addition, the promoters from such genes may be useful to control the expression of economic traits in the production of transgenic plants and further examination may reveal that some have particularly useful timing of expression during embryogenesis.

Application of *Arabidopsis* Microarrays to *Brassica napus*

Species within the genus *Brassica* are the major vegetable oil crop grown in Northern Europe, Canada, and China and represent the third largest source of vegetable oils worldwide. Because of the close phylogenetic relationship of *Arabidopsis* to *Brassica* we examined the ability of the arrays developed for this study to provide information on gene expression in *Brassica napus*. When hybridized with seed and leaf mRNA samples, the correlation coefficients between *Arabidopsis* and *Brassica* experiments varied between 0.73_0.83 for ratios and 0.76_0.83 for intensities (Table V). Since these values are only slightly lower than those for repeated *Arabidopsis* experiments, which varied between 0.86_0.87 for ratios and 0.84_0.96 for intensities, it is clear that *Arrabidopsis* microarrays are a very useful tool to analyze related *Brassica* species. In addition, most seed-specific sequences, which we identified here with *Arabidopsis* probes (Table I and website), also gave seed-specific signals in the *Brassica* hybridization. However, the averaged signal intensities of *Brassica* experiments are almost two_fold lower than those from *Arabidopsis* experiments (Table IV), and therefore the signals from some weakly-expressed genes are likely to be lost in experiments with heterologous probes.

DISCUSSION

The dataset derived from this study provides initial characterization of the tissue expression patterns for a large number of *Arabidopsis* genes. For a substantial number (at least 1000) of the genes studied here no previously published data are available on their expression patterns in seeds or other tissues and therefore these data provide initial information useful toward their characterization. Furthermore, at least 40% of the genes on the arrays are of unknown function, and therefore these new data can guide future work in functional genomics. As just one example, knowledge that previously undescribed protein kinases (such as clones M19D10, M34C01, and M23F11) are seed-specific can direct future analysis of the phenotype of mutants or transgenic plants altered in their expression toward the seed. Similarly, a number of transcription factors are defined by these data as having seed-specific expression and further analysis of their function may provide clues regarding transcriptional control of seed metabolism and development.

The dataset also defines a large number of seed-specific genes which can be further analyzed by examination of the promoter regions for these genes. Previously, only a handful of genes have been available for such analysis which included primarily seed storage protein or other genes with highly abundant transcripts. The set described here includes a much wider range of examples, including genes with widely different expression levels. Bioinformatics analysis of several hundred such promoters with approaches similar to those described by Hughes *et al.* (2000), Tavazoie *et al.* (1999) or Zhang (1999) may therefore offer new insights on cis activation sequences responsible for control of seed expression. Moreover, these promoters can be used to clone their corresponding *trans* acting elements using yeast one-hybrid screenings or similar approaches.Many crop plants are phylogenetically related to *Arabidopsis* and we therefore explored the ability of *Arabidopsis* based arrays to provide useful information on such species. When hybridized with probes derived from mRNA isolated from *B. napus*, the *Arabidopsis* arrays provided a very useful dataset with only a minor loss in sensitivity. The microarray technique thus will enable detailed studies of gene expression in different *Brassica* cultivars. We are currently using the arrays to analyze seeds from transgenic *Brassica napus* lines. These results further suggest that other species within the Brassicacea (e.g. broccoli, cabbage, mustards, etc.) can likely be analyzed with *Arabidopsis* based arrays. This ability is a feature of the cDNA/glass slide based arrays used here which will continue to make them attractive alternatives to oligonucleotide based arrays (Lipshutz et al, 1999) for analysis of many species. The possibility to analyze related species with the same microarray also makes it feasible to compare *Arabidopsis* and *Brassica* probes directly by

simultaneous hybridization to the same microarray. In preliminary experiments with such chimeric comparisons, a number of genes are clearly expressed more highly in the heterologous sample.

Limitations to microarray analysis.

Based on spiking of our mRNA preparations with internal standards, it can be estimated that the sensitivity of the microarrays is approximately 1 mRNA species per 100,000. This roughly corresponds to 1 to 2 mRNA molecules per cell based on the estimate that a cotton embryo cell contains approximately 120,000 molecules per cell (Galau & Dure, 1981, ordered, check it). This level of sensitivity is thus sufficient to detect a large proportion of all genes expressed in the developing seeds. However, it should be recognized that there are other factors which limit the amount of data obtainable from these arrays. Most importantly, the arrays which we have produced, although containing thousands of genes, currently do not contain a high representation of rarely expressed genes. Because the arrays in this initial study are based on sequencing the first 5,000 of 10,000 ESTs from a partially subtracted cDNA library, mRNAs of abundance lower than 0.01% will be under-represented in the population of genes surveyed by these microarrays. Future generations of microarrays which include much more complete coverage of the *Arabidopsis* genome will become available and allow extension of the current data. However, it should be recognized that current microarray technology, whether cDNA or oligonucleotide based, will continue to have difficulty in reliable detection of the most rarely expressed genes. The presence of many highly abundant transcripts, as those for seed storage proteins, has a dilution effect on low abundant transcripts. Furthermore, the use of complex tissue samples for probe synthesis consisting of different and non_synchronized cell types causes an additional increase in probe complexity and can prevent the detection of transcripts, which are only expressed in a small proportion of the tissue sample. Laser capture systems for collecting specific cell types and subsequent RNA amplification methods, which have already been used with animal cells (Luo *et al.*, 1999), may circumvent some of these limitations specific to microarray analysis of multi_cellular organisms.

A further limitation to wide-scale transcription profiling based on cDNA arrays is the possibility of cross-contamination of DNA samples. Handling of many thousands of samples in high-density microtiter format through many steps of manipulations introduces the possibilities of cross-contamination via aerosols or other processes. If a 0.1% contamination were to occur between a seed storage protein which is expressed as 1% of the mRNA population and a transcription factor clone which is expressed at 0.001% then the expression profile observed for the transcription factor could artifactually appear as highly seed-specific. Such artifacts cannot be detected by re-sequencing of the clones used to spot the arrays or by many other common controls. Although the great majority of the data from a microarray are valid, this example emphasizes that users of microarray data must always consider the data to be preliminary and require independent confirmation by techniques such as northern analysis.

In conclusion, the present study provides only the initial information which can be derived from such microarrays. A more complete set of data from this study than can be provided here is downloadable at our website and undoubtedly other workers will be able to "mine" further useful insights by asking questions not considered here. Furthermore, in the future, more detailed studies of the timing of gene expression, patterns of gene expression in seed mutants such as *Wri1*, (Focks and Benning, 1998), and in transgenic plants will provide a second generation of more rich information useful for understanding the complexities of seed metabolism and its control.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

88. Joseph A. White, Jim Todd, Tom Newman, Nicole Focks, Thomas Girke Oscar Martínez de Ilárduya, Jan G. Jaworski, John Ohlrogge, Christoph Benning 2000 A New Set of Arabidopsis ESTs from Developing Seeds: The Metabolic Pathway from Carbohydrates to Seed Oil **Submitted to Plant Physiology**

89. Thomas Girke, Jim Todd, Sari Ruuska, Joe White, Christoph Benning, John Ohlrogge 2000 Microarray Analysis of Developing Arabidopsis Seeds **In Preparation**

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

A patent application is being prepared for the discovery of 10 new seed-specific promoters. License agreement is under negotiation.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

A number of new seed-specific promoters have been discovered and will be useful to provide additional control over transgene expression timing during seed development in transgenic crops.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

Thomas Girke and Joe White have taken industrial positions and have been replaced by new postdocs on the project

RECEIVED JUN 2 3 2000
(email)

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	David A. Somers, Ph.D.	
University:	University of Minnesota	
Agreement Number:	OR22072-66	
Project Title:	*Incorporation of Value-Added Traits Into Alfalfa for Biomass Energy*	
Reporting Period and Report Type:	**From:** **1/1/1999** **To:** **6/30/2000**	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

See submitted report.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

The goal of this project is to produce a biodegradable plastic polymer known as polyhydroxybutyrate (PHB) in leaves of alfalfa plants. This plastic is made naturally by many different bacteria. Because of it can be completely degraded by microorganisms, PHB is an attractive source of nonpolluting plastic that could be used in a variety of consumer products. Production of PHB in plants would be less expensive than production by bacteria in fermentors. Three enzymes are needed to produce PHB in plants. The genes for these enzymes were previously isolated from the bacterium *Ralstonia eutropha* and manipulated by our industry partners for use in plant cells. The genes were introduced into alfalfa and the amount of PHB produced in over 200 different plants was measured. The amount of PHB in leaves varied widely from 0.024-0.2% of the leaf dry weight. All plants had normal growth and vigor suggesting that there was no deleterious effects from PHB production. Further tests were done on 20 plants with the highest accumulation of PHB. The three genes were shown to be integrated into the chromosomes of the alfalfa plants and RNA was made from the genes. Antibodies to the proteins for PHB synthesis were used to determine if the proteins were made, but results were inconclusive. However, PHB granules were

visualized in leaf cells by two methods of microscopy. Only plants receiving the three bacterial genes had PHB granules. Futher biochemical analysis of leaf extracts showed that authentic PHB was produced by alfalfa leaves. Ten plants were used to pollinate plants from the alfalfa variety UMN2966., which is adapted to the midwestern U. S., and an experimental germplasm being developed for biomass energy. Seeds resulting from these crosses were planted and the plants analyzed for presence of the marker gene. These plants are currently being tested for production of PHB. The results from the second year show that significant amounts of PHB can be produced in alfalfa plants and that production has no detectable detrimental affect on the plants.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

See submitted report.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Saruul Purev, Somers, David A., and Samac, Deborah A. 2000. Synthesis of biodegradable plastics in alfalfa plants. Proceedings of the 36th North American Alfalfa Improvement Conference.

Saruul Purev, Somers, David A., and Samac, Deborah A. 2000. Synthesis of biodegradable plastics in alfalfa plants for biomass energy. Proceedings of the Molecular Breeding of Forage Crops.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

None during the reporting period.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

None during the reporting period.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

A post-doctoral scientist has been carrying out research for the project. This scientist has become familiar with new techniques and procedures including: plant tissue culture, plant molecular biology, protein immuno-blotting, spectrofluorometric enzyme analysis, gas chromatography, mass spectroscopy, epi-fluorescence microscopy, immunogold transmission electron microscopy, and plant breeding techniques. Two undergraduate students have worked with this scientist and have been exposed to most of the techniques listed and have been actively involved in plant propagation, nucleic acid extraction, chemical extraction of PHB, and gas chromatography data analysis. The research was presented to several small groups of non-scientists (high school students and undergraduate students) during tours of the laboratory.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

None during the reporting period.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

A post-doctoral scientist, Dr. Saruul Purev, was hired in October 1998 to carry out the project research and continued with the project during the second year.

Incorporation of Value-added Traits into Alfalfa for Biomass Energy.

Year-Two Progress Report.

In the first year of the project we demonstrated the feasibility of production of the biodegradable polyhydroxybutyrate (PHB) polymer in alfalfa. During the second year of this project, we carried out characterization of transgenic plants for (1) expression of genes for biosynthesis of PHB (phaA, phaB, and phaC) and PHBV (bktB, phbB, phbC), (2) integration of the PHB biosynthesis genes into the alfalfa genome, (3) protein levels of enzymes for PHB production, (4) analysis of the specific polymer composition of PHB produced, (5) cellular localization of PHB granules as well as (6) initiated the evaluation of F1 progeny from crosses of primary transformants with elite germplasm developed for biomass production.

Engineering expression of genes for PHB synthesis in alfalfa. Alfalfa was engineered to produce biodegradable polymers ñ polyhydroxybutyrate (PHB) and polyhydroxybutyrate-co-hydroxyvalerate (PHBV). The homopolymer PHB is somewhat brittle, while PHBV copolymers have more flexibility and therefore are suitable for more commercial applications. Three genes *from Ralstonia euthropha* encoding the enzymes for synthesis of PHB (phbA, phbB, phbC) and PHBV (bktB, phbB, phbC) engineered for plastid targeting were introduced into alfalfa by *Agrobacterium*-mediated transformation. About 128 PHB and 98 PHBV alfalfa plants were regenerated with transformation efficiencies of 88% and 89%, respectively. Presence of transgenes was assayed by PCR assays using NPT marker gene-specific primers. DNA and RNA blot analyses confirmed the integration and expression of PHB biosynthetic pathway genes both in PHB and PHBV alfalfa plants. Analysis of proteins levels of enzymes for PHB synthesis using protein-specific antibodies generated previously for *Arabidopsis* plants resulted in non-specific cross-hybridization showing common signals both in transgenic and wild-type alfalfa. Additional analysis of proteins and enzymes for PHB biosynthesis are currently under investigation.

Analysis of PHB content in transgenic alfalfa. The accumulation of PHB in leaves of alfalfa was analyzed using gas chromatography (GC). Extracts of 128 PHB plants showed that synthesis of PHB in the leaves of transgenic plants ranged from 8 to 350 mg/g of fresh weight representing 0.024% - 0.2% of the plant dry weight in 83 plants. However, only one out of the 98 transformants for PHBV actually contained hydroxyvalerate monomers. Production of PHB in alfalfa leaves was confirmed by nuclear magnetic resonance (^{1}H-NMR). Comparative analysis of the peaks in plant extracts and the bacterial PHB standard showed the characteristic NMR peaks for PHB at 1.2, 2.25 and 5.2 ppm. Analysis of leaf extracts by GC-mass spectra analysis revealed peaks that were undetectable in extracts from wild-type plants. Identification of novel peaks found by mass spectra analysis is in progress.

Appearance of PHB granules. Transmission electron microscopy of plants expressing all three genes revealed that PHB accumulated as agglomerations of electron-lucent granules 0.2 - 0.5 mm in diameter. PHB granules were correctly located in the chloroplasts as the PHB-genes were engineered to be

RECEIVED JUL 0 5 2000
(email)

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Dr. Steven H. Strauss	
University:	Oregon State University	
Agreement Number:	OR22072-78	
Project Title:	*Genes Controlling the Transition Between Vegetative and Reproductive Phases in Forest Trees*	
Reporting Period and Report Type:	**From:** 1/1/00 **To:** 6/1/00	**Check one:** [x] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

Objective 1: Isolate partial poplar homologs to known flowering time genes.

In prior and associated work, we isolated poplar orthologs to the *Arabidopsis* genes *APETALA1* (*AP1*), *AGAMOUS* (*AG*), and *LEAFY* (*LFY*). Through this project to date, we have isolated probable poplar orthologs to the flowering time genes: *ADP GLUCOSE PYROPHOPHORYLASE1* (*ADG1*), *Arabidopsis thaliana CENTRORADIALIS* (*ATC*), *CONSTANS* (*CO*), *CRYPTOCHROME1* (*CRY1*), *INDETERMINATE1* (*ID1*), *LATE ELONGATED HYPOCOTYL* (*LHY*), *MOTHER OF FT* (*MFT*), *PHOSPHOGLUCOMUTASE* (*PGM*), *PHYTOCHROME A* (*PHYA*), *PHYTOCHROME B* (*PHYB*), *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1*), *SQUA promoter binding protein-like3* (*SPL3*), *SHORT VEGETATIVE PHASE* (*SVP*), and *VERNALIZATION2* (*VRN2*). Poplar paralogs to the following gene families have also been identified: *APETALA2* (*AP2*), *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*), *CO, CRY1* and *2*, *GaMYB*, *GIBBERELLIN INSENSITIVE* (*GAI*), *SPL3*, *SPINDLY* (*SPY*), and *UNUSUAL FLORAL ORGANS* (*UFO*). A detailed summary and the number of isolated homologs to each gene is provided in the Scientific Accomplishments.

Objective 2: Study the sequences of the isolated homologs to assure they are evolutionary homologs, and analyze their genome structure via Southern blots to see if they are parts of gene families.

BLAST searches, sequence alignments, and hypothetical translation of the poplar *ADG1*, *AG*, *AP1*, *AP2*, *ATC*, *CCA1*, *CO*, *CRY*, *CRY1*, *GAI*, *GaMYB*, *ID1*, *LFY*, *LHY*, MADS-box, *MFT*, *PGM*, *PHYA*, *PHYB*, *SOC1*, *SPL3*, *SPY*, *SVP*, *UFO*, and *VRN2* homologs have been conducted to verify their evolutionary relationship to corresponding gene families. Based on Southern blot

analysis, two genes encode orthologs to *AP1* (*PTAP1-1* and *PTAP1-2*), two genes encode orthologs to *AG* (*PTAG1* and *PTAG2*), and one gene encodes a ortholog to *LFY* (*PTLF*). Three to four genes encoding homologs of *ATC* and *MFT* (members of the *TFL/FT* family) have been observed on Southern blots; cloning of these additional members is underway. Southern analysis is currently underway on the remaining genes to determine each gene's family size in poplar.

Objective 3: Study expression in different tissues over developmental and maturation gradients.

All tissues for expression analysis have been collected from a clonally-propagated male and female poplar (age range 1-6 years) over a single growing season. Tissue types include pre-dormancy vegetative buds, post-dormancy vegetative buds, new shoots, new vegetative buds, new inflorescence buds, summer shoots, and seeds (6 year old female only). RNA for expression studies has been prepared from all collected tissues, except for seedlings which is currently underway. Expression analysis is underway for *PTLF*, *PTAG2*, and *PTAP1-2*.

Objective 4: Isolate a full length cDNA from 1-2 genes if they show a promising pattern of expression.

Not Initiated Yet.

Objective 5: Begin transformation of altered cDNAs for 1-2 genes to study the effects of overexpression and inhibition in *Arabidopsis* and poplar.

Not Initiated Yet.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

The delay between vegetative growth and initiation of flowering remains a major hurdle to progress in breeding, marker aided selection, and genetic engineering of forest trees. Utilizing information from studies in model plant species, we have isolated partial gene fragments from forest trees to gain new insights into the control of flowering and juvenility. We have isolated *Populus trichocarpa* (poplar) gene fragments which are homologs of seventeen distinct plant genes (*ADG1*, *AG*, *AP1*, *ATC*, *CO*, *CRY1*, *ID1*, *LFY*, *LHY*, *MFT*, *PGM*, *PHYA*, *PHYB*, *SOC1*, *SPL3*, *SVP*, and *VRN2*) shown to affect or act as indicators of floral timing. We have also identified poplar fragments that correspond to eight additional gene families which contain members involved in or acting as indicators of floral timing. We have established a collaboration with the Swedish Agricultural University's Poplar EST Project that has resulted in the generation of ~5,300 poplar floral tissue ESTs, which are currently being annotated. Expression studies have been initiated and are currently underway on a subset of the identified poplar floral timing gene families.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Isolation of poplar flowering time gene family homologs.

Based on homology to known floral timing and indicator genes, candidate poplar genes have been identified to 25 gene families to date. Identification and isolation was conducted using a combination of degenerate primer-based PCR, library screening, and screening of a poplar EST database (PopulusDB; http://www.biochem.kth.se/PopulusDB/). A summary of the poplar orthologs and paralogs identified to each gene family follows. These twenty-five families conclude the goals of Objective 1. Identification of additional timing gene families and members will come from periodic searching of the poplar EST database and analysis of the poplar floral ESTs (following their annotation).

ADG1: *ADG1* encodes an ADP glucose pyrophosphorylase and affects flowering time. Searches were conducted of the poplar EST database and a likely poplar ortholog to the *Arabidopsis ADG1* large subunit gene was identified from mature poplar leaf ESTs. This clone has been requested from our Swedish collaborators and we are currently awaiting its arrival.

API, ***AG***, and ***LFY***: In a related project, we have isolated poplar homologs to these three genes through screening of cDNA libraries. These genes are involved in aspects of floral timing, floral meristem identity, and/or floral organ identity. Two poplar homologs to *API* (*PTAP1-1* and *PTAP1-2*) have been isolated. Two alleles to *PTAP1-1* have also been identified, as well as full-length cDNAs and partial genomic DNAs to each *PTAP1* gene. Two poplar homologs to *AG* (*PTAG1* and *PTAG2*) have been isolated. Full-length cDNAs and genomic clones to *PTAG1* and *PTAG2* have also been isolated. Both *API* and *AG* are members of the large MADS-box gene family. One poplar homolog to *LFY* (*PTLF*) has been isolated. A genomic clone and full-length cDNA to *PTLF* have been isolated. Southern analysis has been conducted on each gene and these isolates represent the entire orthologous gene family for each class. Phylogenetic analysis has confirmed their relationship to the appropriate families. Suitable regions of sequence are present in each gene which have been used to design gene-specific probes for expression studies (below).

AP2: *AP2* functions as a meristem and floral organ identity gene and is the founding member of the AP2 domain-containing *AP2/EREBP* gene family. Searches of the poplar EST database failed to detect a poplar *AP2* ortholog, but four poplar ESTs containing an AP2 domain protein motif were identified. Three ESTs correspond to members of the ethylene responsive *EREBP* subfamily and the fourth EST is orthologous to *AINTEGUMENTA*, a member of the *AP2* subfamily. All four clones have been obtained.

CO: *CO* encodes a zinc-finger type transcription factor and functions to control flowering time. Utilizing degenerate primers, we have isolated multiple near full-length *CO* cDNA and genomic DNA products, corresponding to at least 2-3 separate *CO* genes (designated *PtCOL1* and *PtCOL2*). We have also identified a putative expressed *CO* pseudogene. Utilizing searches of the poplar EST database, we have identified additional *CO*-related family members that encode probable paralogous members, and requested five from our Swedish collaborators.

CRY1* and *2: *CRY1* and *CRY2* encode related blue light photoreceptors that regulate multiple aspects of plant development, including flowering time. Three separate poplar *CRY* gene fragments have been isolated by degenerate PCR from a small insert genomic DNA library. The three genes have been tentatively designated *PtCRYa*, *PtCRYb*, and *PtCRYd*. Based on the available sequence, the genes likely encode poplar *CRY1* orthologs. *PtCRYb* and *PtCRYd* encode alleles of the same *CRY1* gene, while *PtCRYa* encodes a second poplar *CRY1* gene. In searches of the poplar EST database, we identified a poplar homolog to the *Arabidopsis* gene *PHR2*, which encodes a DNA photolyase and is a member of the extended *CRY* gene family. This EST has been requested from our Swedish collaborators.

FLC: The *FLC* gene (also known as *FLF*) plays a central role in vernalization-based regulation of flowering time. *FLC* is a member of the well conserved MADS-box gene family. Isolation of *FLC* is currently underway. We have screened a poplar genomic library using the C-terminal domain region of the *FLC* cDNA. We have identified seven cross-hybridizing plaques, one of which cross-hybridized strongly. Cross-hybridizing fragments from the strongly reacting plaque are currently being subcloned and will be subsequently analyzed by sequence analysis.

GAI: *GAI* is closely related to *RGA* and *RGAL* and affects flowering time in short days. *GAI*, *RGA*, and *RGAL* are members of the GRAS-type gene family and likely encode transcription factors. The GRAS gene family includes the defining member *SCARECROW* and a number of related genes, designated *SCARECROW*-like (*SCL*) genes. Searching of the poplar EST database failed to detect a poplar *GAI* (or *RGA* or *RGAL*) ortholog, but a poplar EST belonging to the GRAS gene family was identified that is closely related to *SCL3*, which we have obtained.

GaMYB: *GaMYB* is a member of the large R2R3 MYB-domain plant gene family; *Arabidopsis* encodes over 100 MYB-domain proteins. *GaMYB* responds to gibberellic acid and affects flowering time. Searching of the poplar EST database failed to detect a poplar *GaMYB* ortholog, but at least nine poplar ESTs containing MYB domains were identified. Four of the MYB domain poplar ESTs which are phylogenetically close to *GaMYB* have been obtained.

ID1. *ID1* is a member of a medium sized gene family and affects flowering time. Using degenerate primers, we have obtained nine *ID1* clones that encode five individual genes. The isolated fragments appear to encode very closely related *ID1* family members; however, due to the highly conserved nature of the amplified region, additional flanking sequence will be needed to distinguish between true orthologs vs. paralogs. Searches of the poplar EST database failed to detect *ID1*-related sequences.

LHY* and *CCA1: *LHY* and *CCA1* are closely related but distinct *Arabidopsis* genes, encoding single MYB domain-based transcription factors, involved in regulating circadian rhythms and flowering time. Utilizing PCR, we have isolated a partial poplar *LHY* ortholog. Through searches of the poplar EST database, we have identified two *LHY/CCA1*-related ESTs. One EST encodes the same *LHY* gene we have isolated, but appears to contain additional 5' sequence relative to the fragment we isolated. Based on BLAST search results, the second EST is more closely related to *CCA1* than *LHY*, but likely encodes a *CCA1* paralog, not ortholog, and is most closely related to a hypothetical *Arabidopsis* protein. This EST has been requested.

MADS-box family: The MADS box gene family in plants is large and members are involved in many aspects of development, including floral timing and organ identity. Based on limited searches (cambial ESTs only) of the poplar EST database, we have identified four additional ESTs that contain the highly conserved MADS domain. Searches of the remaining poplar EST tissue databases are currently underway.

PGM: *PGM* encodes a phosphoglucomutase that affects starch accumulation and flowering time. Utilizing PCR, we have isolated a putative full-length poplar *PGM* ortholog. Through searches of the poplar EST database, we have identified two *PGM* ESTs. Both appear to encode portions of the same gene we have cloned.

PHYA and ***PHYB***: *PHYA* and *PHYB* encode related but distinct red light photoreceptors that regulate multiple aspects of plant development, including flowering time. Degenerate primers to *PHYA* and *PHYB* genes were utilized to obtain fragments to each gene. We have isolated one *PHYA* (*PtPHYA1*) and two *PHYB* (*PtPHYB1* and *PtPHYB2*) genes, which correspond with the reported family size for these two genes in poplar. Screening for additional *PHY* genes has not yielded additional members. Searching of the poplar EST database failed to identify ESTs corresponding to the photo-labile *PHYA* gene, but did turn up four *PHYB*-related ESTs. Three of these ESTs contain sequence outside the region we have cloned and have been requested.

SOC1. *SOC1* (previously designated *AGL20*) is a member of the MADS-box gene family and assists *CO* in promoting flowering. Searches of the poplar cambial EST database identified a MADS-box containing EST that is highly similar to *SOC1*. The encoded product is either an *SOC1* ortholog, or an extremely close subfamily member. The EST containing this clone has been obtained and additional sequencing is currently underway to help discriminate if a true ortholog is present.

SPL3: *SPL3* is a member of the SBP-box containing gene family and affects flowering time. Using degenerated primers, we have isolated a poplar *SPL3* ortholog representing approximately 50% of the *Arabidopsis SPL3* coding region. Searching of the poplar EST database failed to detect a poplar *SPL3* ortholog, but two poplar ESTs were identified that displayed similarity to other SBP-box family members outside the SBP-box. One EST was similar to *SPL1* and the other to *SPL11*; both have been obtained.

SPY: *SPY* is a member of the tetratricopeptide protein repeat (TPR) family and affects flowering time. PCR was used to isolate a poplar *SPY*-like family member that contains similarity in the catalytic C-terminal domain, outside the TPRs, designated *PtSPY*-L1 (for *SPY*-Like1). *PtSPY*-L1 is a paralog of *SPY*, and likely encodes a plant O-GlcNAc transferase. Searching of the poplar EST database identified four poplar ESTs with similarity to *SPY* in the catalytic domain, though none encode a *SPY* ortholog. Three ESTs display similarity to *SPY* in the catalytic C-terminal region while the fourth shows similarity to the last two TPRs. Comparison of the four ESTs suggests two to three additional *SPY*-like genes are represented. We have obtained one additional *SPY*-like EST and requested the EST with similarity to the TPRs.

SVP. *SVP* is a member of the MADS-box gene family and delays flowering. Searches of the poplar cambial EST database identified a MADS-box containing EST that encoded either an *SVP* ortholog, or an extremely close subfamily member. The EST containing this clone has been obtained and additional sequencing is currently underway to help discriminate if a true ortholog is represented.

TFL* and *FT: *TFL* and *FT* are related members of a small gene family that regulates flowering time. While *TFL* delays flowering, *FT* has the opposite effect, promoting flowering. Degenerate primers to the *TFL/FT* family were used to isolate two related family members from poplar floral cDNA. One poplar clone (*PtATC*) groups most closely with *ATC*, the *Arabidopsis CENTRORADIALIS* ortholog, which phylogenetically branches with *TFL* (Figure 1). The other poplar clone (*PtMFT*) groups closest with *MFT*, which phylogenetically branches with *FT* (Figure 1). Southern blot analysis has indicated 2-4 distinct cross-hybridizing family members each for *PtATC* and *PtMFT*, suggesting a possible family size of at least 4-8 members.

UFO: *UFO* affects floral meristem and organ identity and is useful as an early indicator of flowering. Searches were conducted of the poplar EST database. A poplar ortholog was not identified, however a related paralog was observed that displays high similarity to a predicted *Arabidopsis* gene. This clone has been requested.

VRN2: *VRN2* is an unpublished gene that affects flowering time in conjunction with *FCA*. In collaboration with Caroline Dean, unpublished information on *VRN2* was shared allowing the identification of two overlapping poplar ESTs that encode a likely *VRN2* ortholog. We have obtained both these ESTs.

Poplar EST collaboration and poplar floral EST generation.

We have established a collaboration with Dr. Ove Nilsson and the Swedish Agricultural University Poplar EST Project. A floral cDNA library from a female *Populus trichocarpa* tree was supplied, from which approximately 5300 ESTs are being sequenced; ESTs are currently being annotated. These ESTs should yield a large number of gene sequences involved in tree flowering. Microarray expression studies will be conducted on interesting ESTs (and our timing gene clones) in the future using our RNAs from the tissue collection studies (below). Additionally, as a result of our collaboration, we are able to search other poplar EST collections that have not been made publicly accessible yet. Current EST sequence tissue sources include cambial and wood forming tissues, mature leaf, vegetative shoot tips, and floral buds.

Tissue collections and RNA isolations for expression analysis.

For expression studies, tissue has been collected from clonally-propagated trees over a 1-6 year age range. The male tissue source is hybrid poplar (*Populus trichocarpa* x *deltoides*) clone 24-305 and the female tissue source is hybrid poplar (*Populus trichocarpa* x *deltoides*) clone 15-29. Trees are growing at Fort James plantations at Claskanie, OR, and Scapoose, OR (Lower Columbia River Basin). In general, only trees of ages 5-6 had initiated flowering. Tissue samples collected were pre-dormancy vegetative buds, post-dormancy vegetative buds, new shoots, new

vegetative buds, new inflorescence buds, summer shoots, and seeds (6 year old female only). All RNA samples have been isolated from each collected tissue except for seedlings germinated from seeds of the six year old female; germination's are currently underway.

Gene expression studies.

Gene expression analysis is currently underway on the meristem identity genes *PTLF*, *PTAP1-2*, and *PTAG2* in vegetative buds of 1, 2, 5, and 6 year old trees. In post-dormancy vegetative buds (ages 1, 2, 5, 6) of the female 15-29 clone, *PTLF* expression levels were constant across the age gradient while *PTAG2* and *PTAP1-2* showed higher expression at age 5 relative to ages 1 and 2, but dropped significantly at age 6 (Figures 2, 3). Comparing expression levels in post-dormancy, pre-dormancy, and newly differentiated vegetative buds of clone 15-29 (year five and six), *PTAG2* levels were highest in post-dormancy vegetative buds (Figure 4). In contrast, *PTAP1-2* levels were highest in newly differentiated vegetative buds (Figure 4). An increase in *PTLF* levels was observed in newly differentiated vegetative buds relative to post-dormancy vegetative buds over both years; *PTLF* levels fluctuated in pre-dormancy vegetative buds however (Figure 4). Based on the observed expression patterns, *PTLF* expression is regulated by distinct factors relative to *PTAG2* and *PTAP1-2*. Overall, a consistent pattern of gene expression vs. flowering competence was not observed for the three meristem identity genes in these tissues (Figure 5). Further expression analysis of the three meristem identity genes in additional tissues is currently underway. Expression analysis is being initiated for poplar *ID1*, *MFT*, *ATC*, and *CO* genes.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

A seminar entitled "Flowering Control" was presented at the Tree Genetic Engineering Research Cooperative (TGERC) Annual Meeting (November 17-18, 1999, Corvallis, OR). The seminar was directed towards the TGERC representatives, a mixture of scientific and nonscientific backgrounds. The talk was an update on the various flowering related projects, including the poplar flowering time gene project and directed towards a combination of scientific and business research personnel. A second seminar entitled "Floral Regulatory Genes As Tools For Control Of Reproduction In Poplars" was presented as an invited, keynote lecture during a symposium on tree biotechnology at the International Society for Plant Molecular Biology Congress (June 2000, Quebec, Canada).

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

No technology transfer arrangements have resulted from the work.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

No commercial applications have currently resulted from the work.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

New educational aspects since the last Progress Report occurred in the form of attendees at scientific conferences where aspects of the project's work were presented (see Publications and Presentations section).

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

No additional funding has been generated as a result of the project yet.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

No personnel turnover has occurred since the last Progress Report.

Phylogenetic analysis of poplar TFL homologs

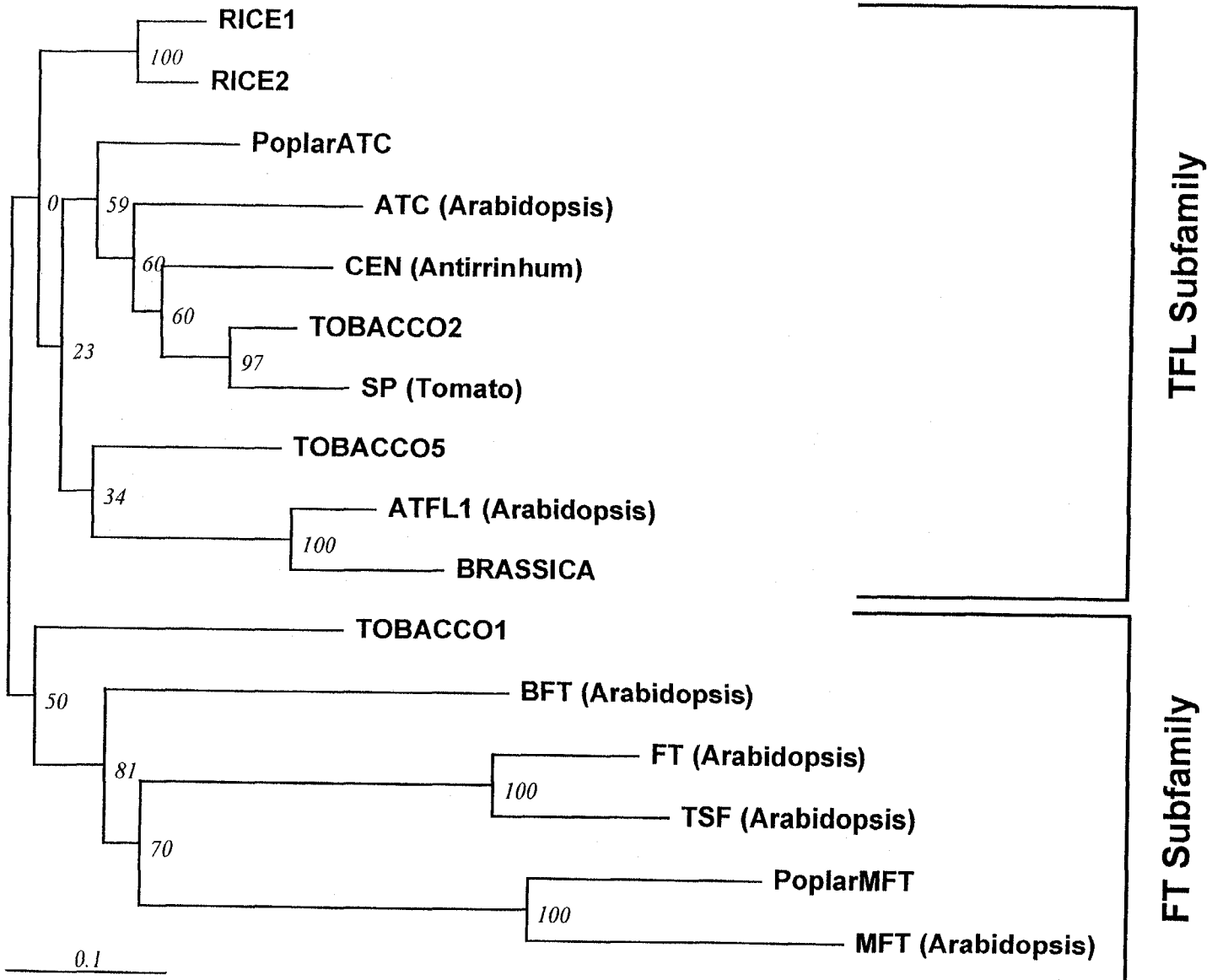


Figure 1

Age Dependent Expression of Floral Genes in *Populus trichocarpa x deltoides*
Post Dormancy Vegetative Buds of Clone 15-29

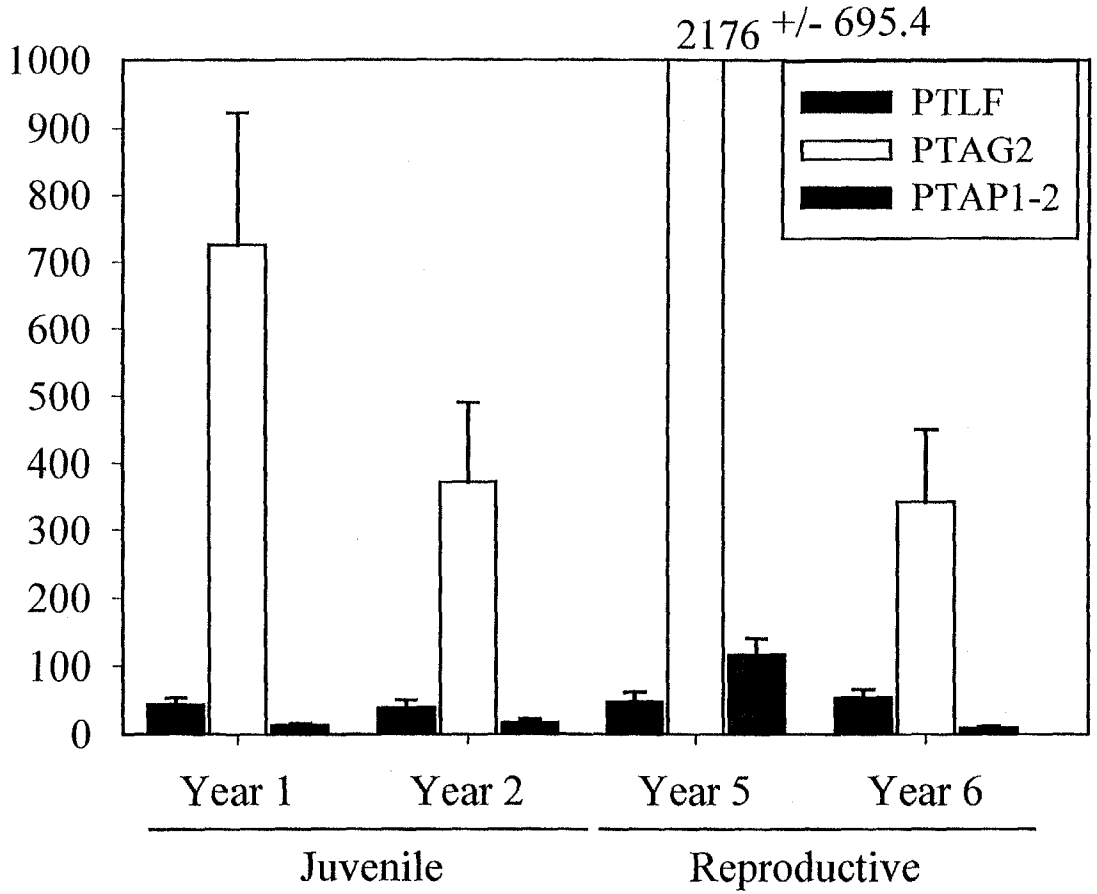


Figure 2

Age Dependent Expression of Floral Genes in *Populus trichocarpa x deltoides*
Post Dormancy Vegetative Buds of Clone 15-29

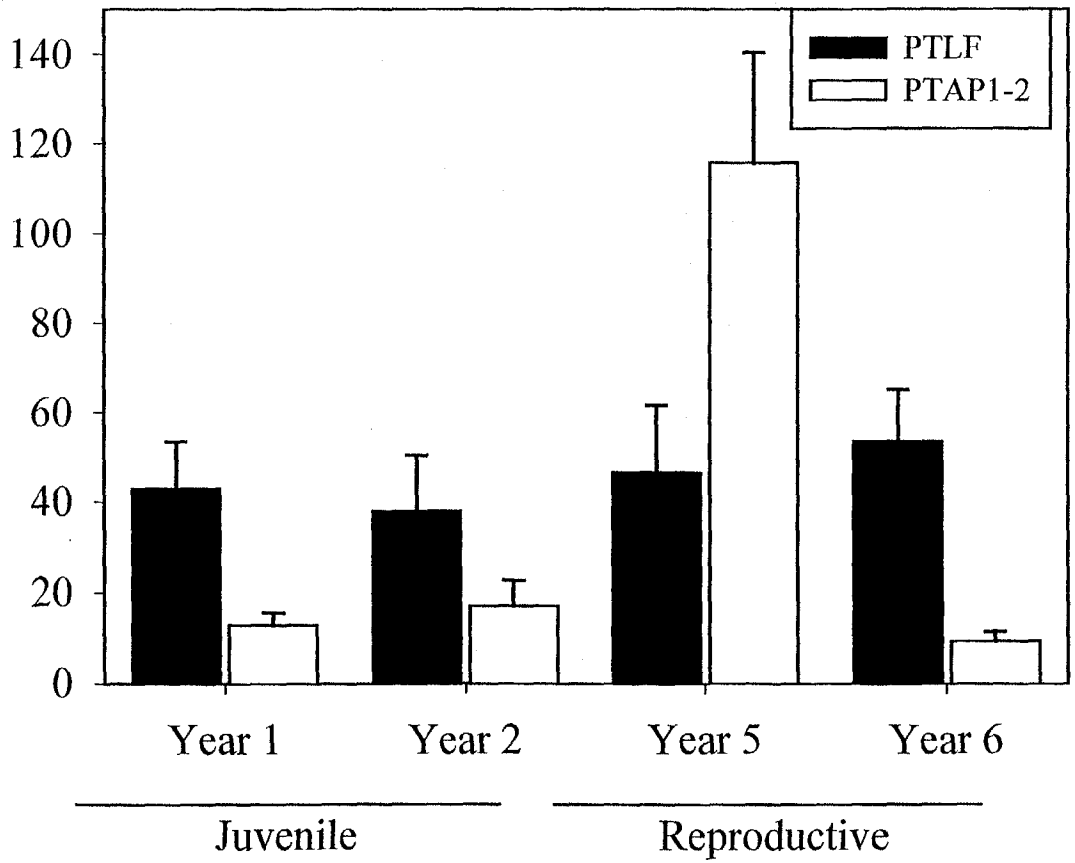


Figure 3

Seasonal Changes in Expression of rRNA Floral Genes in *Populus trichocarpa x deltoides*
Vegetative Buds of Clone 15-29

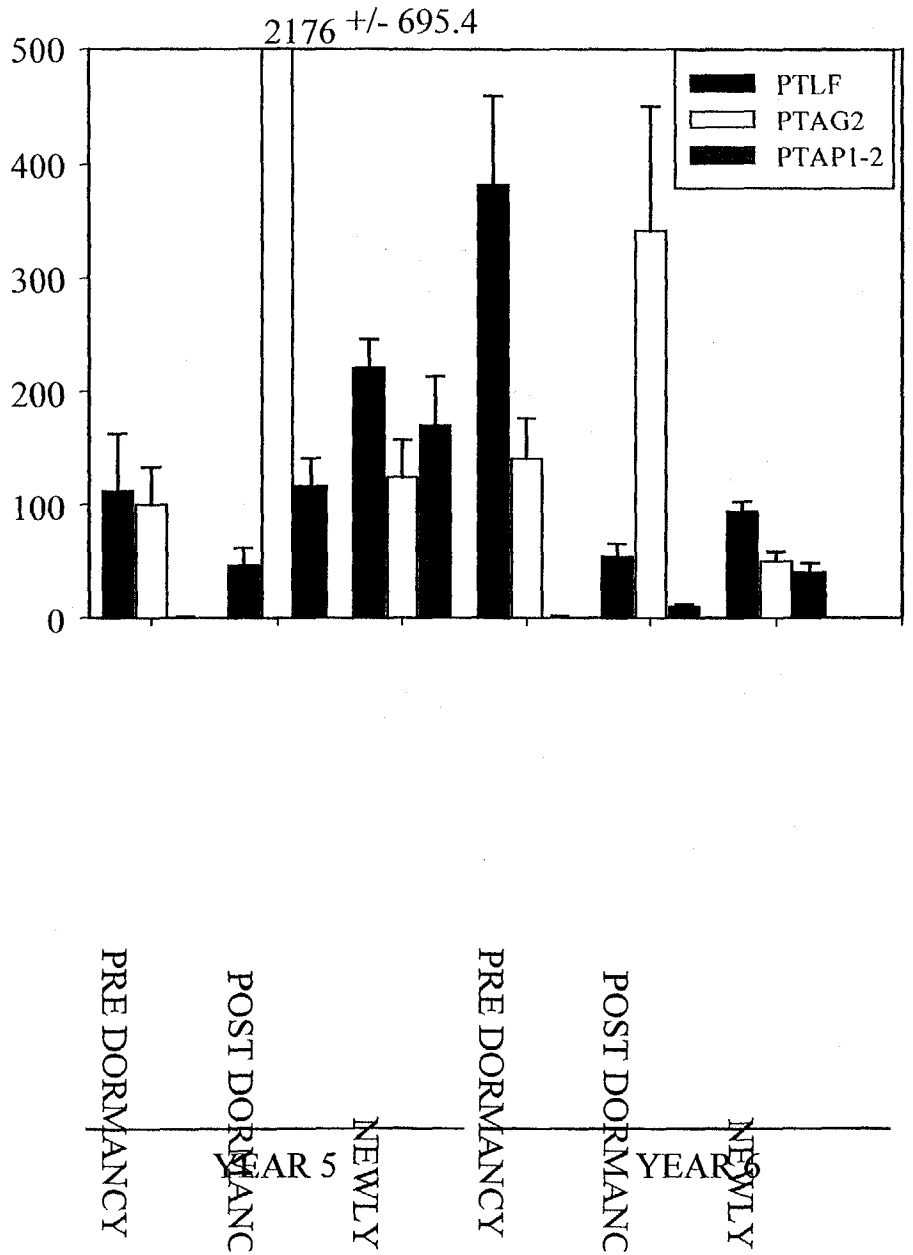


Figure 4

Seasonal Changes in Expression of Floral Genes in *Populus trichocarpa x deltoides*
Vegetative Buds of Clone 15-29, Gene-Year-rRNA Normalized

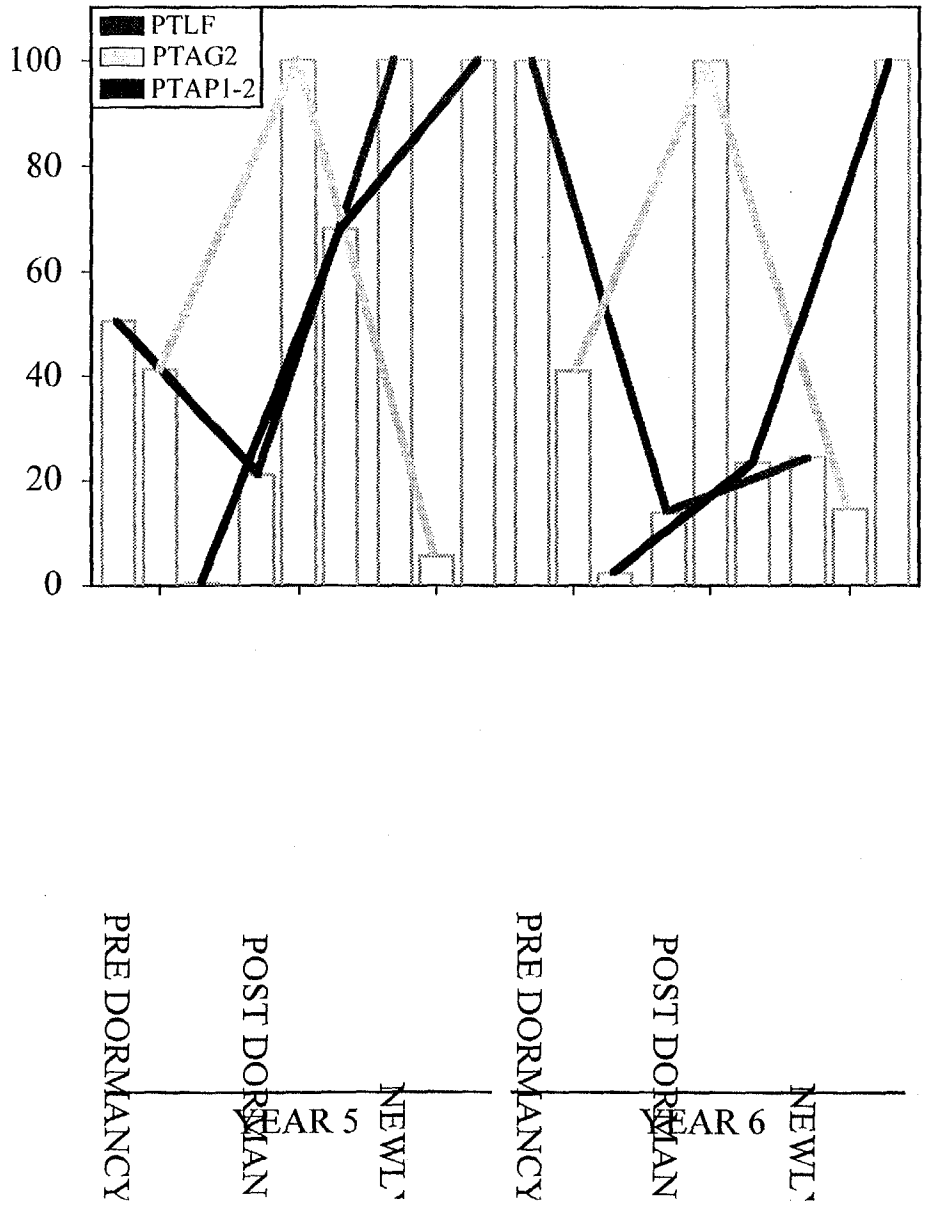


Figure 5

RECEIVED JUN 2 9 2000

Scientific Progress Report

Submitted to The Consortium for Plant Biotechnology Research, Inc., in accordance with the requirements of the Research Agreement cited.

Principal Investigator:	Jack M. Widholm	
University:	University of Illinois at Urbana-Champaign	
Agreement Number:	OR22072-79	
Project Title:	*A New Selectable Marker and Promoters of Plant Origin*	
Reporting Period and Report Type:	**From:** 7/1/98 **To:** 6/30/00	**Check one:** [X] Interim Report [ ] Final Report

Project Objectives

List each objective of the Project and the progress made toward each one during the reporting period.

The overall objectives of this project are to improve the efficiency of producing transformed plants, to produce plants that are considered to be environmentally safe by preventing expression of undesirable selectable marker genes in the plants and to express useful genes in the plants at the desired levels using plant derived sequences that are not patented by others. Our specific plans to accomplish these objectives are:

Objective 1: Develop the tobacco ASA2 gene as a selectable marker in many species.

The results thus far show that the ASA2 gene driven by the 35S CaMV promoter can be used to select transformed soybean and *Astragalus sinicus* hairy roots since the transformed roots are quite resistant to the toxic Trp analog 5-methyltryptophan. Results with tobacco leaf disks or Arabidopsis plants have not given clear-cut results but further testing will be done with plants carrying and expressing the ASA2 gene.

Objective 2: Develop the tobacco ASA2 promoter for tissue culture specific and constitutive expression in many species.

The 1.3 kb ASA2 gene promoter fragment has been found to be a strong constitutive promoter in tobacco, Arabidopsis, *Datura innoxia* and *A. sinicus* plants. This expression is being characterized in detail in these species and is being tested with tree systems by Westvaco researchers.

Layperson's Summary

Summarize in one or two paragraphs the most significant scientific accomplishments of the Project during the reporting period. Use language that a non-scientist can understand.

We have begun work demonstrating that indeed the naturally occurring ASA2 gene from tobacco can impart resistance to a toxic tryptophan analog so that this gene, which is quite different from the antibiotic or herbicide resistance usually used in gene transfer experiments, may be an effective selectable marker.

So far the results using the tobacco ASA2 gene promoter to drive gene expression show that high level expression can occur in most tissues of certain species indicating that this plant promoter may be useful for causing high level expression of useful genes in plants.

Scientific Accomplishments

Describe in two to ten pages, excluding tables and figures, the most significant scientific accomplishments of the Project during the reporting period.

Selectable marker analysis

A number of experiments have been carried out to determine if expression of the tobacco ASA2 gene (Song et al. 1998) in transgenic plant systems will lead to resistance to the toxic tryptophan (Trp) analog, 5-methytryptophan (5MT). Transformation of soybean cotyledons and *A. sinicus* seedlings with *A. rhizogenes* harboring a binary vector carrying ASA2 driven by the CaMV 35S promoter produced numerous hairy roots that would grow in the presence of 5MT concentrations that prevented hairy root growth on the controls. Molecular analysis, including Southerns, showed that all the resistant hairy root lines carried the ASA2 gene and the roots of both species and the shoots formed on the *A. sinicus* roots had greatly increased free Trp levels. The anthranilate synthase activity found in the resistant roots was also more resistant to feedback inhibition by Trp.

A. tumefaciens transformation of tobacco leaf disks and Arabidopsis plants followed by selection for 5MT-resistance has not produced clear-cut results. The work is continuing and in both cases we are also selecting for kanamycin resistance with a cotransformed *nptII* gene so lines carrying ASA2 should be obtained for direct testing.

Thus these results indicate that the use of ASA2 as a selectable may be possible with certain species.

Westvaco Corporation researchers have the ASA2 gene for testing as a selectable marker in their tree species transformation systems.

Promoter analysis

Transgenic plants of *Nicotiana tabacum* (Xanthi), Arabidopsis (Columbia), and *Datura innoxia* containing the GUS reporter gene driven by different size fragments of the ASA2 promoter,

produced by the *Agrobacterium tumefaciens* mediated transformation method, have been used for GUS histochemical assays to analyze the ASA2 promoter activity. Well known strong promoters such as CaMV 35S and actin were used as controls to evaluate ASA2 promoter activities. The 1.3 Kb ASA2 promoter was constitutive and about as strong as CaMV 35S and actin in all three different species and much stronger than the 606 bp ASA2 promoter. Interestingly, this result was different from the result of the transient expression assay using tobacco where the 606 bp ASA2 promoter was stronger than the 1.3 Kb ASA2 promoter (Song et al. 1998). These transient/stable transformation results strongly suggest the existence of a positively acting transcription factor(s) with binding site(s) between 1.3 Kbp and 606 bp upstream of the ASA2 gene and a negatively acting transcription factor(s) whose binding site(s) should be within 606 bp upstream of ASA2 promoter. The 2.3 Kb ASA2 promoter showed the tissue specific activity similar to that found in transient expression assays.

These ASA2 promoter deletion-GUS fusion constructs were also used to produce hairy roots from which shoots can develop with the forage legume, *Astragalus sinicus* (Chinese milk vetch) (Cho et al. 1999). As with the other species the 1.3 Kb promoter showed high level expression in both roots and shoots while the others, 606 and 2.3, were weak. This work is being completed and will be written up for publication soon.

These results indicate that the 1.3 Kb ASA2 promoter can be used to drive high level gene expression in several species.

References

Song, H.-S., J.E. Brotherton, R.A. Gonzales and J.M. Widholm (1998) Tissue culture specific expression of a naturally occurring tobacco feedback-insensitive anthranilate synthase. Plant Physiol. 117:533-543.

Cho, H.-J., J.M. Widholm, N. Tanaka and Y. Murooka (1998) *Agrobacterium rhizogenes*-mediated transformation and regeneration of the legume *Astragalus sinicus* (Chinese milk vetch). Plant Sci. 138:53-65.

Publications and Presentations

List all talks, posters, scientific publications, news releases, etc., about this Project during the reporting period. Provide one copy of each publication, report, or news article resulting from activities supported under the grant as well as any announcements, press releases, statements or photographs depicting Project activities.

Cho, H.-J., H.-S. Song, J.E. Brotherton and J.M. Widholm. 1998. Analysis of tobacco anthranilate synthase gene promoter using *Agrobacterium rhizogenes*-mediated *Astragalus sinicus* transformation system. In Vitro Cell. Dev. Biol. 34:63A.

Song, H.-S. and J.M. Widholm. 1998. Single-stranded DNA binding protein regulating expression of feedback-insensitive anthranilate synthase gene in *Nicotiana sylvestris*. Plant Physiol. Suppl. 117:156.

Cho, H.J., H.-S. Song, J.E. Brotherton and J.M. Widholm. 1998. Analysis of tobacco anthranilate synthase gene promoter using *Agrobacterium rhizogenes*-mediated *Astragalus sinicus* transformation system. 1998. Crown Gall Meeting, Lafayette, IN.

Brotherton, J.E. and J.M. Widholm. 1999. Metabolic consequences of overexpression of the tobacco feedback-insensitive anthranilate synthase ASA2 gene. Am. Soc. Plant Physiol. Abstracts 134.

Widholm, J.M. 2000. Tissue culture studies of anthranilate synthase the tryptophan biosynthetic control enzyme. J. Plant Biotechnol. 2:55-60.

Cho, H.-J., J.E. Brotherton, H.-S. Song and J.M. Widholm. 1999. Increasing tryptophan synthesis in a forage legume *Astragalus sinicus* by expressing the tobacco feedback-insensitive anthranilate synthase (ASA2) gene. Plant Physiol. in press.

Technology Transfer

Describe all technology transfer (inventions, disclosures, patents, licensing agreements, etc.) resulting from the Project during the reporting period.

The patent application was submitted before the project began. Submitted Invention Disclosure in Dec., 1995, entitled "Selectable marker and promoter for plant tissue culture transformation", filed Provisional Patent Application in July 1996 and Patent Application in July 1997. A portion of the claims concerning the promoter was allowed in 1998 and patent number 5,965,727 was issued Oct. 12, 1999. The selectable marker portion of the original application is still being examined by the U.S. Patent Office.

Commercial Accomplishments

Describe the most significant accomplishments resulting from the Project during the reporting period.

The plasmid pUC35SASA2 and binary vectors pBIN35SASA2 and pC2NPTASA2 in E. coli were given to Westvaco in March 2000 to test as selectable markers with their tree transformation systems. We are still discussing our plan for research interactions with RhoBio. We plan to visit the labs of Westvaco and RhoBio soon to present a seminar on our data and discuss the collaborations.

Educational Accomplishments

Describe the most significant educational accomplishments resulting from the Project during the reporting period.

Postdoctoral Research Associates are engaged in the research so are receiving training in research methods and idea development as the project develops.

Additional Funding

List any additional funding generated as a result of the Project during the reporting period.

None.

Key Personnel Hiring or Turnover

List any changes in key personnel during the reporting period.

The postdoctoral Research Associate, Junghee Kim who was hired to replace another postdoctoral Research Associate, Hee-Sook Song, who left in the fall of 1998 to take a position at Dekalb Genetics, had to leave in April due to visa problems. Dr. Niu Dong began work on the promoter expression aspects and Dr. Hyeon-Je Cho continues to work on the selectable marker aspects.

Metrics Report for Year One Matching Company

Matching Company: Weyerhaeuser	Principal Investigator: [redacted] Scott Merkle
Company Contact (name, phone, fax, email): Nick Wheeler 360-330-1738 Ph " " -1742 Fax	Project Title: Clonal Propagation of hybrid southern hardwoods

Please complete this report and add any comments you wish. When finished, seal it in an envelope and return to the PI who sent it to you. The PI will forward the report to CPBR in partial fulfillment of reporting requirements. The information you provide will be kept confidential by CPBR and used solely for federally-required statistical reporting. Your cooperation is greatly appreciated.

A. Rate the project to date on the items below using a scale of 0 (low/poor) to 10 (high/excellent).

	low/poor 0	1	2	3	4	OK 5	6	7	8	high/exc. 9	10
Scientific accomplishments made by the project								X			
Degree to which the project's original objectives have been accomplished										X	
Degree to which the company's original expectations have been met										X	
Frequency of communication between the company and the project PI						X					
Quality of communication between the company and the project PI											X

Comments:

B. Estimate the numbers of the following that have resulted from or can reasonably be expected to result from this research project.

	to date	anticipated
Patents applied for	—	1
Patents issued	—	1
Licenses obtained	—	1
Products created	0	2
Products improved	—	—
Processes created	—	—
Processes improved	0	1
Instances of new thinking which created a framework/base for new products or processes	—	—
Dollar value of above instances of new thinking	$ —	$ —
Instances enabling commercialization of products in terms of regulatory adherence	—	—
Instances enabling the evaluation of technologies in a defensive sense (e.g., understanding the competition)	—	—

Comments:

RECEIVED SEP 18 2000

C. Estimate the future economic impact of the use and/or commercialization of the results of this project. Please list estimates in annual terms at the future times indicated.

	In 3 years	In 5 years	In 10 years
Product sales *(thousands of dollars)*	$ 0	$ 0	$ 0
Number of jobs created *(include R&D, manufacturing, distributing, marketing, etc.)*	0	0	0
Number of jobs retained *(include R&D, manufacturing, distributing, marketing, etc.)*	1	0	0
Number of jobs with increased value	1	0	0
Number of new businesses created	0	0	0

Comments:

D. Estimate the corporate investment *in addition to the cash matching funds provided* that the company has made and is likely to make to accomplish the technology transfer and commercialization of results of the project. List amounts in thousands of dollars annually at the future times indicated.

	Current/ to date	In 3 years	In 5 years	In 10 years
Personnel Time *Include R&D personnel, personnel involved in liaison with CPBR project, implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*	—	—	—	—
Facilities – space, equipment *Include R&D facilities and facilities involved in implementing the technology transfer and producing, marketing, distributing the new or improved products and/or processes, etc.*	$ —	$ —	$ —	$ —
Supplies *Include R&D supplies involved in implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*	$ —	$ —	$ —	$ —
Legal Costs *Include legal costs involved in implementing the technology transfer and in protecting and producing the new or improved products and/or processes, etc.*	$ —	$ —	$ —	$ —
Other Costs				
Item _____	$	$	$	$
Item _____	$	$	$	$
Item _____	$	$	$	$

Comments: We have elected not to pursue hardwood planting at this time.

This form was completed by:

Nicholas Wheeler — 9/13/00

Name — Signature — Date

RECEIVED SEP 15 2000

Metrics Report for Year One Matching Company

Matching Company: **Dow AgroSciences**	Principal Investigator: **Jiming Jiang**
Company Contact (name, phone, fax, email): **Dr. Michael G. Murray** **(317) 337-3982** **mmurray@dowagro.com**	Project Title: **Toward cloning a functional rice centromere**

Please complete this report and add any comments you wish. When finished, seal it in an envelope and return to the PI who sent it to you. The PI will forward the report to CPBR in partial fulfillment of reporting requirements. The information you provide will be kept confidential by CPBR and used solely for federally-required statistical reporting. Your cooperation is greatly appreciated.

A. Rate the project to date on the items below using a scale of 0 (low/poor) to 10 (high/excellent).

	low/poor 0	1	2	3	4	5	OK 6	7	8	9	high/exc. 10
Scientific accomplishments made by the project							X				
Degree to which the project's original objectives have been accomplished							X				
Degree to which the company's original expectations have been met							X				
Frequency of communication between the company and the project PI			X								
Quality of communication between the company and the project PI			X								

Comments:

B. Estimate the numbers of the following that have resulted from or can reasonably be expected to result from this research project.

	to date	anticipated
Patents applied for	?	
Patents issued		
Licenses obtained		
Products created		
Products improved		
Processes created		
Processes improved		
Instances of new thinking which created a framework/base for new products or processes		
Dollar value of above instances of new thinking	$	$
Instances enabling commercialization of products in terms of regulatory adherence		
Instances enabling the evaluation of technologies in a defensive sense (e.g., understanding the competition)		

Comments: Very early stage

C. Estimate the future economic impact of the use and/or commercialization of the results of this project. Please list estimates in annual terms at the future times indicated.

	In 3 years	In 5 years	In 10 years
Product sales *(thousands of dollars)*	$0	$0	$10 MM
Number of jobs created *(include R&D, manufacturing, distributing, marketing, etc.)*	0	0	10
Number of jobs retained *(include R&D, manufacturing, distributing, marketing, etc.)*			
Number of jobs with increased value			
Number of new businesses created			

Comments:

D. Estimate the corporate investment *in addition to the cash matching funds provided* that the company has made and is likely to make to accomplish the technology transfer and commercialization of results of the project. List amounts in thousands of dollars annually at the future times indicated.

	Current/ to date	In 3 years	In 5 years	In 10 years
Personnel Time *Include R&D personnel, personnel involved in liaison with CPBR project, implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*	0			
Facilities – space, equipment *Include R&D facilities and facilities involved in implementing the technology transfer and producing, marketing, distributing the new or improved products and/or processes, etc.*	$0	$	$	$
Supplies *Include R&D supplies involved in implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*	$0	$	$	$
Legal Costs *Include legal costs involved in implementing the technology transfer and in protecting and producing the new or improved products and/or processes, etc.*	$0	$	$	$
Other Costs				
Item_____	$	$	$	$
Item_____	$	$	$	$
Item_____	$	$	$	$

Comments: Premature to insert any numbers above. MGM

This form was completed by:

Name Michael G. Murray Signature Date5/12/00

Metrics Report for Year One Matching Company

Matching Company: **Novartis**	Principal Investigator: **Jiming Jiang**
Company Contact (name, phone, fax, email): **Dr. Danny C. Alexander** **(919) 541-8687** **danny.alexander@nabri.novartis.com**	Project Title: **Toward cloning a functional rice centromere**

Please complete this report and add any comments you wish. When finished, seal it in an envelope and return to the PI who sent it to you. The PI will forward the report to CPBR in partial fulfillment of reporting requirements. The information you provide will be kept confidential by CPBR and used solely for federally-required statistical reporting. Your cooperation is greatly appreciated.

A. Rate the project to date on the items below using a scale of 0 (low/poor) to 10 (high/excellent).

	low/poor 0	1	2	3	4	OK 5	6	7	8	high/exc. 9	10
Scientific accomplishments made by the project										X	
Degree to which the project's original objectives have been accomplished										X	
Degree to which the company's original expectations have been met										X	
Frequency of communication between the company and the project PI							X				
Quality of communication between the company and the project PI								X			

Comments:

B. Estimate the numbers of the following that have resulted from or can reasonably be expected to result from this research project.

	to date	anticipated
Patents applied for		
Patents issued		
Licenses obtained		
Products created		
Products improved		
Processes created		
Processes improved		
Instances of new thinking which created a framework/base for new products or processes		
Dollar value of above instances of new thinking	$	$
Instances enabling commercialization of products in terms of regulatory adherence		
Instances enabling the evaluation of technologies in a defensive sense (e.g., understanding the competition)		

Comments:

C. Estimate the future economic impact of the use and/or commercialization of the results of this project. Please list estimates in annual terms at the future times indicated.

	In 3 years	In 5 years	In 10 years
Product sales *(thousands of dollars)*	$	$	$
Number of jobs created *(include R&D, manufacturing, distributing, marketing, etc.)*			
Number of jobs retained *(include R&D, manufacturing, distributing, marketing, etc.)*			
Number of jobs with increased value			
Number of new businesses created			

Comments:

D. Estimate the corporate investment *in addition to the cash matching funds provided* that the company has made and is likely to make to accomplish the technology transfer and commercialization of results of the project. List amounts in thousands of dollars annually at the future times indicated.

	Current/ to date	In 3 years	In 5 years	In 10 years
Personnel Time *Include R&D personnel, personnel involved in liaison with CPBR project, implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*				
Facilities – space, equipment *Include R&D facilities and facilities involved in implementing the technology transfer and producing, marketing, distributing the new or improved products and/or processes, etc.*	$	$	$	$
Supplies *Include R&D supplies involved in implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*	$	$	$	$
Legal Costs *Include legal costs involved in implementing the technology transfer and in protecting and producing the new or improved products and/or processes, etc.*	$	$	$	$
Other Costs				
Item_____	$	$	$	$
Item_____	$	$	$	$
Item_____	$	$	$	$

Comments:

This form was completed by:

Danny C. Alexander Ph.D.

Name Signature Date

RECEIVED APR - 4 2000

Metrics Report for Year One Matching Company

Matching Company: Common Ground Group USA	Principal Investigator: Dr. Yi Li
Company Contact (name, phone, fax, email): Colie O'Donnell Telephone: 978-318-9310 Fax: 978-318-9311 Email: codonnell@cgg-usa.com	Project Title: Genetic Improvement of Seed Productivity for Biomass Crops

Please complete this report and add any comments you wish. When finished, seal it in an envelope and return to the PI who sent it to you. The PI will forward the report to CPBR in partial fulfillment of reporting requirements. The information you provide will be kept confidential by CPBR and used solely for federally-required statistical reporting. Your cooperation is greatly appreciated.

A. Rate the project to date on the items below using a scale of 0 (low/poor) to 10 (high/excellent).

	low/poor 0	1	2	3	4	OK 5	6	7	8	9	high/exc. 10
Scientific accomplishments made by the project										√	
Degree to which the project's original objectives have been accomplished									√		
Degree to which the company's original expectations have been met											√
Frequency of communication between the company and the project PI											√
Quality of communication between the company and the project PI											√

Comments:
Reduced budget impacted accomplishments but we are happy with the progress none-the-less.

B. Estimate the numbers of the following that have resulted from or can reasonably be expected to result from this research project.

	to date	anticipated
Patents applied for		1
Patents issued		
Licenses obtained		1
Products created		multiple
Products improved		
Processes created		
Processes improved		
Instances of new thinking which created a framework/base for new products or processes		
Dollar value of above instances of new thinking	$	$

Instances enabling commercialization of products in terms of regulatory adherence		
Instances enabling the evaluation of technologies in a defensive sense (e.g., understanding the competition)		

Comments:

C. Estimate the future economic impact of the use and/or commercialization of the results of this project. Please list estimates in annual terms at the future times indicated.

	In 3 years	In 5 years	In 10 years
Product sales *(thousands of dollars)*	$20,000	$50,000	$90,000
Number of jobs created *(include R&D, manufacturing, distributing, marketing, etc.)*	4	6	7
Number of jobs retained *(include R&D, manufacturing, distributing, marketing, etc.)*			
Number of jobs with increased value			
Number of new businesses created	1		

Comments: Sales projections are projected incremental value of technology at producer level so do not reflect CGG revenues. CGG will distribute the technology primarily by sublicenses so its revenue will be in the form of royalties and licensing fees negotiated as a percentage of revenues. The projections assume that there will be significant competition from alternative technologies. Amounts could be significantly higher in the unlikely event there are no suitable alternatives. Jobs created are for CGG and do not reflect new jobs likely to be created by sub licensees as a result of the technology.

We are particularly interested in the commercial potential of this technology integrated with another technology we are working on at CGG USA.

D. Estimate the corporate investment *in addition to the cash matching funds provided* that the company has made and is likely to make to accomplish the technology transfer and commercialization of results of the project. List amounts in thousands of dollars annually at the future times indicated.

	Current/ to date	In 3 years	In 5 years	In 10 years
Personnel Time *Include R&D personnel, personnel involved in liaison with CPBR project, implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or processes, etc.*	$5	$300	$500	$900
Facilities – space, equipment *Include R&D facilities and facilities involved in implementing the technology transfer and producing, marketing, distributing the new or improved products and/or processes, etc.*	$	$150	$300	$
Supplies *Include R&D supplies involved in implementing the technology transfer, and in producing, marketing, distributing the new or improved products and/or*	$	$100	$250	$450

processes, etc.				
Legal Costs *Include legal costs involved in implementing the technology transfer and in protecting and producing the new or improved products and/or processes, etc.*	$	$25	$75	$
Other Costs				
Item: Marketing Costs	$	$150	$200	$250
Item _____	$	$	$	$
Item _____	$	$	$	$

Comments: CGG will seek to sublicense this technology widely. Costs shown reflect CGG anticipated costs but do not reflect downstream costs of sub licensees.

This form was completed by:

Name: Colum P. O'Donnell Signature Date: 3/29/2000