Composition of Committee of Visitors for the Plant Genome Research Program
Directorate for Biological Sciences

The Committee of Visitors consisted of seven members, including Dr. Burt Ensley, who represented the BIO Advisory Committee, and will report his observations to the Advisory Committee at their next meeting. Four of the COV members are female, and the members currently work in seven different states: Arizona, California, Connecticut, Kansas, Massachusetts, Minnesota, and New Hampshire. Seven of the members are from academic institutions, one is from industry, and all members hold the Ph.D., with the earliest degree awarded in 1966 and the most recent awarded in 1990.

All files presented to the committee were first scrutinized from possible conflicts with committee members. None of the sample of proposals pulled for Committee review came from any institution with which a member had a known conflict of interest. Committee members were advised about confidentiality and conflicts of interest prior to arriving at NSF and at the inception of the meeting. No conflict issues arose during the meeting.

Mary E. Clutter
Assistant Director

Plant Genome COV Members
August 11-12, 2004

BIO AC Liaison

Dr. Burt D. Ensley
PMB 1319
2675 W. Hwy 89A
Sedona, AZ 86336
(928) 203-4127
burtensley@aol.com

COV Chair

Dr. Susan Singer
Department of Biology
Carleton College
Northfield. MN 55057
(507) 646-4391
ssinger@carleton.edu

COV Members

Dr. Forrest Chumley
Dean of Research
Kansas State University
Manhattan, KS 66506
(785) 532-6148
fchumley@ksu.edu
krothwel@oznet.ksu.edu

Dr. Anita Klein
Dept of Biochemistry
University of New Hampshire
Durham, NH 03824
Anita.Klein@unh.edu

Dr. Sarah Mathews
Arnold Arboretum
Harvard University Herbaria .
22 Divinity Ave.
Cambridge, MA 02138
(617) 495-2365/2331
smathews@oeb.harvard.edu

Dr. Dong-Guk Shin
Computer Science & Engineering Dept.
University of Connecticut, Storrs
371 Fairfield Rd.
Storrs, CT 06269-2155
(860 483-2783
shin@brc.uconn.edu

Dr. Sarah Tobin
Stanford Center for Biomedical Ethics
Stanford University
701 Welch Rd., Bldg. A, Suite 1105
Palo Alto, CA 94304
(650) 725-2663
tobinsl@stanford.edu

CORE QUESTIONS and REPORT TEMPLATE
for

FY 2004 NSF BIO/PGRP COMMITTEE OF VISITOR (COV) REVIEWS

NSF relies on the judgment of external experts to maintain high standards of program management, to provide advice for continuous improvement of NSF performance, and to ensure openness to the research and education community served by the Foundation. Committee of Visitor (COV) reviews provide NSF with external expert judgments in two areas: (1) assessments of the quality and integrity of program operations and program-level technical and managerial matters pertaining to proposal decisions; and (2) comments on how the outputs and outcomes generated by awardees have contributed to the attainment of NSF's mission and strategic outcome goals.

Many of the Core Questions are derived from NSF performance goals and apply to the portfolio of activities represented in the program under review. The program under review may include several subactivities as well as NSF-wide activities. The directorate, division or program may instruct the COV to provide answers addressing a cluster or group of programs – a portfolio of activities integrated as a whole – or to provide answers specific to the subactivities of the program, with the latter requiring more time but providing more detailed information.

The Program may choose to add questions relevant to the activities under review. NSF staff should work with the COV members in advance of the meeting to provide them with the report template, organized background materials, and to identify questions/goals that apply to the program under review.

Guidance to the COV: The COV report should provide a balanced assessment of NSF's performance in two primary areas: (A) the integrity and efficiency of the ***processes*** related to proposal review; and (B) the quality of the ***results*** of NSF's investments in the form of outputs and outcomes that appear over time. The COV also explores the relationships between award decisions and program/NSF-wide goals in order to determine the likelihood that the portfolio will lead to the desired results in the future. Discussions leading to answers for Part A of the Core Questions will require study of confidential material such as declined proposals and reviewer comments. *COV reports should not contain confidential material or specific information about declined proposals.* Discussions leading to answers for Part B of the Core Questions will involve study of non-confidential material such as results of NSF-funded projects. It is important to recognize that the reports generated by COVs are used in assessing agency progress in order to meet government-wide performance reporting requirements, and are made available to the public. Since material from COV reports is used in NSF performance reports, the COV report may be subject to an audit.

We encourage COV members to provide comments to NSF on how to improve in all areas, as well as suggestions for the COV process, format, and questions.

NSF Plant Genome FY 2004 REPORT TEMPLATE FOR NSF COMMITTEES OF VISITORS (COVs)

Date of COV: August 11-13, 2004
Program: Plant Genome
Division: Biological Infrastructure
Directorate: Biological Sciences
Number of actions reviewed by COV: Awards: 36 Declinations: 45 Other: 0
Total number of actions within Program during period being reviewed by COV: Awards: 108 Declinations: 270 Other: 116 Total: 494
Manner in which reviewed actions were selected: Jackets were chosen to be representative of Standard, Continuing, and Cooperative awards, as well as ISGA, VCA, YIA, SGER and workshop proposals. Jackets were also selected to include new, renewal and resubmitted proposals.

Table A. Roster of all pulled jackets
- Table A.1. 2001 Awards and Declines
- Table A.2. 2002 Awards and Declines
- Table A.3 2003 Awards and Declines

Table B. Rosters of all proposals
- Table B.1 FY2001 PGRP Proposals
- Table B.2 FY2002 PGRP Proposals
- Table B.3 FY2003 PGRP Proposals

Table C. All awards
- Table C.1. Roster of all FY2001 PGRP Awards
- Table C.2. Roster of all FY2002 PGRP Awards
- Table C.3. Roster of all FY2003 PGRP Awards
- Table C.4. Portfolio Balance

Document 1. Review process for all programs

Table 1. Percentage of reviews addressing both review criteria

Table 2. Data obtained from pulled jackets
- Table 2.a PGRP Proposals with a Conflict of Interest (COI)
- Table 2.b PGRP Proposals Review Analysis Summary Data
- Table 2.c PGRP Proposals Panel Summary Data

Table 3. Dwell Time (See also Graph 1)

Table 4. Average number of reviews (mail + panel) per proposal

Table 5. Number of reviewers by institution type

Table 6. PGRP Reviewers from underrepresented groups

Table 7. Average award size and average award duration (See also Graph 2)

Table 8. Institution type for awards (See also Graph 3)

Table 9. Minority Involvement
- Table 9.a. Awards to Minority Serving Institutions
- Table 9.b. Minority PI Funding Rates

Table 10. Funding Rate by State
- Table 10.a FY2001 PGRP Funding by State
- Table 10.b FY2002 PGRP Funding by State
- Table 10.c FY2003 PGRP Funding by State
- Table 10.d FY2001 through FY2003 PGRP Funding by State

Table 11. Student Involvement in PGRP Projects

Table 12. Doctorates Awarded by State in Biological and Agricultural Sciences
- Table 12.a 2001 Doctorate Information
- Table 12.b 2002 Doctorate Information
- * 2003 Information will not be available until Spring 2005

Graph 1. Dwell Time

Graph 2. Average Award Size - Requested vs. Actual
- Graph 2.a. Average award size – requested vs. actual
- Graph 2.b. Average award duration – requested vs. actual

Graph 3. Award expenditure by institution type

Map 1. Geographical Distribution of PGRP Awards
- Map 1.a PGRP FY2001 Geographical Distribution of Awards
- Map 1.b PGRP FY2002 Geographical Distribution of Awards
- Map 1.c PGRP FY2003 Geographical Distribution of Awards
- Map 1.d PGRP FY2001 – FY2003 Geographical Distribution of Awards

Map 2. Geographical Distribution of Reviewers
- Map 2.a PGRP FY2001 Geographical Distribution of Reviewers
- Map 2.b PGRP FY2002 Geographical Distribution of Reviewers
- Map 2.c PGRP FY2003 Geographical Distribution of Reviewers

PART A. INTEGRITY AND EFFICIENCY OF THE PROGRAM'S PROCESSES AND MANAGEMENT

Briefly discuss and provide comments for *each* relevant aspect of the program's review process and management. Comments should be based on a review of proposal actions (awards, declinations, and withdrawals) that were *completed within the past three fiscal years*. Provide comments for *each* program being reviewed and for those questions that are relevant to the program under review. Quantitative information may be required for some questions. Constructive comments noting areas in need of improvement are encouraged. Please do not take time to answer questions if they do not apply to the program.

A. 1 Questions about the quality and effectiveness of the program's use of merit review procedures.

A. 2 Questions concerning the implementations of the NSF Merit Review Criteria (intellectual merit and broader impacts) by reviewers and program officers. Provide comments in the space below the question. Discuss issues or concerns in the space provided.

QUALITY, EFFECTIVENESS and IMPLEMENTATION OF MERIT REVIEW PROCEDURES	**YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE**

Is the review mechanism appropriate? (panels, ad hoc reviews, site visits) (Document 1) The review mechanism is rigorous and fair. A reading of several jackets leads to the conclusion that the process is followed consistently. Proposers are given an opportunity to correct deficiencies that may be detected after an initial screen for responsiveness to the Program Solicitation. The ad hoc review process is managed in a way that ensures timely receipt of at least three peer reviews from qualified, impartial reviewers. Ad hoc reviewers are clearly requested to evaluate proposals in regard to both merit review criteria. Requesting multiple panel reviews ensures that contrasting opinions may be heard at the panel meeting. POs prepare award recommendations based on panel recommendations, portfolio considerations and broader impacts. The COV notes that it has been necessary to convene more than one panel to accommodate the large number of proposals received. A possible concern arises concerning the melding of recommendations from two or possibly more panels. Internal panel dynamics can be distinctive. It could be valuable for panel chairs to meet jointly with the POs to provide an opportunity for cross-checking during formulation of the award recommendation. PGR is to be commended for working closely with Proposers to refine and refocus high-quality proposals at various stages, including pre- and post-panel questions and site visits. The criteria for determining when a site visit is required are well-conceived, and a review of several jackets shows that site visits are rigorous and well-documented. The process for reviewing SGER and Workshop Proposals is consistent with NSF guidelines. PGR is to be commended for seeking external peer reviews for supplemental requests in excess of $100,000, something that is not required by NSF guidelines. This adds rigor and should be applied in all such instances.	Yes with some recommendations
Are reviews consistent with priorities and criteria stated in the program's solicitations, announcements, and guidelines? (program solicitations and Jackets) The priorities and criteria in the solicitations is fairly broad; support of research on plant genomics, and accelerate utilization of new knowledge and innovative approaches to elucidate biological processes in plants. Emphasis was on functional genomics, defined as the identification of functions of a pathway or a cluster of genes at a genomic scale, and new informatics tools. All of the reviews considered were consistent with these guidelines.	Yes
Have the individual reviews (either mail or panel) addressed whether the proposal contributes to both merit review criteria? (Table 1) With respect to the evaluation of both review criteria by individual reviewers, the COV found that, as indicated in Table 1, most reviewers commented on criterion 2. However, a look at the jackets revealed that the quality and depth of comments on criterion 2 is variable. Some reviewers did not comment on criterion 2 at all; others made only cursory comments; others provided thoughtful critiques. For example, among the reviews of one proposal, 5 reviewers explicitly considered criterion 2 while two reviewers did not comment	Not always

at all. A similar split occurred among reviewers of three other proposals. A three-way split characterized the reviews of yet another proposal: two reviewers evaluated criterion 2, one provided a cursory evaluation, and one did not comment at all.	
Have the panel summary reviews addressed whether the proposal contributes to both merit review criteria? (Table 2) The panel summary reviews consistently addressed both merit review criteria. In a sample of 12 proposal jackets, all addressed criterion 1 and criterion 2 in depth and with detailed examples.	Yes
Do the individual reviews (either mail or panel) provide sufficient information for the principal investigator(s) to understand the basis for the reviewer's recommendation? (Jacket Review) Overall the individual reviews are sufficiently detailed in terms of both criterion 1 and 2 to understand the basis for the reviewer's recommendation. The reviewers attend to past performance for renewals, the quality of the proposed science and its feasibility, and broader impacts. For example, reviewers for one proposal acknowledged the tremendous contributions the PI had already made to the field, critically analyzed the proposed work in terms of the science (e.g. concerns about strategies), integration of personnel involved in the project, and considerations related to agricultural improvement and involving a K-12 education expert in the outreach component. Not all reviews are helpful. In particular brief reviews that rate the proposal as excellent provide minimal information. Another proposal included a review that contained a four -line summary and no other information. This was balanced by other reviews in the jacket that were more detailed. The reviewers often have a range of expertise, and each person focused on aspects of the proposal they were qualified to review.	Yes with some exceptions
Do the panel summaries provide sufficient information for the principal investigator(s) to understand the basis for the panel recommendation? (Jacket Review) Panel summaries do indeed sum up the reviews. They appear to succinctly point out issues that emerged during the panel and that could be accidentally buried in individual reviews. It is also noteworthy that in 2001 COV a suggestion was made that the PO include a summary justifying the funding and such summary could be shared with future COV. All of the reviewed jackets included such a summary and that significantly helped understanding PO's funding decision.	Yes
Have the *review analyses* (Form 7s) addressed whether the proposal contributes to both merit review criteria? (Jacket Review and Table 2.b) The COV notes that there has been improvement since the 2001 report, in the response of reviewers to addressing both scientific merit, and broader impact (criterion 2) in their review analyses. Tabular materials confirm that ~85% of the	Yes

review analyses discuss both merit review criteria. While the bulk of the review analyses focus on merit criteria one, concise comments are included about criterion 2 (re: impact and outreach). In one instance, the PO directed follow-up questions to the PI's requesting more explicit training and outreach plans. The PI, in a direct response, addressed these issues. However in annual reports, the PI tended to overlook the Merit Criterion 2. One proposal has a particularly well thought out plan for internships for both undergraduate and high school students. A part-time outreach coordinator will be hired to work with advisors to facilitate recruitment. The reviews comment on the plan and the review analyses summarize this.	
Is the documentation for recommendations complete, and does the program officer provide sufficient information and justification for her/his recommendation? (Jacket Review) PO award recommendations are based on PGR panel recommendations, portfolio considerations and broader impacts. The PO's award recommendation is presented on Form 7, which draws on primary peer reviews, panel summaries, context statements, and other relevant factors. Documentation was thorough and complete in a number of jackets that were reviewed, including an equal number of "accepted" and "declined" proposals. The Form 7 documents that were reviewed included no references to portfolio considerations. If portfolio fit was a consideration in any of the award recommendations, it was not documented.	Yes
Is the time to decision appropriate? (Table 3 and Graph 1) The time to decision was laudable. In 2001, and 2002, 99% of the proposals had been completely reviewed and a decision made within 9 months, and 60% within 6 months. In 2003, 96% were complete within 6 months.	Yes
Is the review process efficient and effective? Initial processing appears to be accomplished in a smooth and timely fashion. The COV commends the PGR program officers for maintaining the efficiency of the process, as indicated by the decrease in dwell time from FY01 (5.6 mo aver) to FY03 (4.4 mo aver), while the number of submitted proposals increased from FY01 (143) to FY03 (196). During this time there were an increasing number of proposals with a conflict of interest from 56% to 63%, decreasing the potential pool of both ad hoc reviewers and panel members. Nonetheless, during this time the average number of reviews/proposal was maintained, or increased slightly from 4.9 (FY01) to 5.3 (FY03). Overall, this suggests that the review process is indeed efficient and effective. Of the 30 proposals recommended for funding in FY03, 5 were recommended for site visits. This represents a considerable investment of time and resources. However, given the complexity of the proposals targeted for site visits and the size of the award ($3.6 to 10.9 million; average of $8 million), this additional investment seems well warranted.	Yes

A I.

Comments on the quality and effectiveness of the program's use of merit review procedures:
The quality of merit review is excellent and highly effective. Reviewers, when they can be persuaded to respond, are conscientious, informed and thorough. Proposers are responsive to reviewer's

comments and this process improves the quality and effectiveness of the funded proposals. This rigorous review does make it more difficult for less experienced investigators to obtain funding, but the Program is taking steps to address this issue.

Recommendations:

External reviewers need additional training and instruction in the range of issues that can be included in the "broader impacts" category. Particularly in the plant genome program, community contributions are appropriate and very easy to address under this category.

A 2.
Comments on NSF's merit review system:

If it reflects what we saw in the Plant Genome Program, comments in A1 are applicable.

Recommendations:

If it reflects what we saw in the Plant Genome Program, comments in A1 are applicable.

A.3 Questions concerning the selection of reviewers. Provide comments in the space below the question. Discuss areas of concern in the space provided.

SELECTION OF REVIEWERS	YES , NO, DATA NOT AVAILABLE, or NOT APPLICABLE

Did the program make use of an adequate number of reviewers for a balanced review? (Table 4) POs do an excellent job of recruiting an adequate number of reviewers with appropriate expertise in the face of complex relationships and conflicts of interest among investigators.	Yes
Did the program make use of reviewers having appropriate expertise and/or qualifications? (Jacket Review) The highly integrative and collaborative nature of many of the proposals requires expertise in a broad range of areas. Reviewers were selected with a range of expertise to provide adequate coverage. This conclusion was reached by looking at the expertise of the individual reviewers in comparison to the proposal objectives and by assessing the sophistication of the technical analysis of the reviews. In general, comments are specific, knowledgeable, and detailed. Examples of this can be seen in the comments by a reviewer for one proposal (especially 2D gels) and a reviewer for another proposal (particularly with respect to QTLs). Bioinformatics expertise is sought where appropriate and complemented by reviewers with other essential expertise.	Yes
Did the program make appropriate use of reviewers to reflect balance among characteristics such as geography, type of institution, and underrepresented groups? (Map 2, Table 5 and Table 6) FY 2001 data suggests quite even geographical distribution of reviewers over 39 states, with CA having 24 most reviewers and NY being the second most with 13 reviewers. The average number of reviewers per state is 4.1. CA appears to be somewhat unusually higher number (24). There appears to be no gender gap in reviewer selection with approximately a 50-50 split in PGRP. There is a slight disparity in DBI that shows slightly higher percentage of male panelists. Both PGRP and DBI show steady increase of reviewers from underrepresented groups.	Yes
Did the program recognize and resolve conflicts of interest when appropriate? (Jacket Review) PGRP appears to carefully monitor and appropriate actions were taken to prevent conflicts of interest from interfering with reviews in most instances. Checklists in each jacket indicate when conflicts of interest occurred (either in review or in the panel). Summaries of panel discussions indicate that a panel member left the room when there was a conflict before discussion of a proposal began. A rather extensive review of potential COI's were included in one proposal, because subsequent to the panel meeting, a panelist, was added as a Co-PI to this project. The Co-PI and her home institution were chosen based on a competitive bidding process, hence it was concluded there were no COI influencing the Program award. The documentation for these discussions is extensive.	Yes

Comments on selection of reviewers:

The COV commends efforts of the PGR POs to recruit reviewers for a rapidly growing number of proposals. A few states provided a larger number of reviewers (CA, NY, TX) but this observation is balanced by the larger number of proposals originating from these states. Modest numbers of reviewers were selected from 4 year institutions and master's granting institutions, but this seems reasonable given the relatively small number proposals from these institutions. Otherwise the selection of reviewers was well balanced with respect to geography, and the notable for gender balance. The POs have been very professional in their treatments of potential COI.

Recommendations:

COV recommends proactive measures to increase the number of underrepresented groups amongst reviewers. This may in time contribute to an increase in proposals from this group. At the same time, the gender balance among reviewers and underrepresented groups probably as an undesired outcome. These reviewers may have an undo burden of "community service," particularly younger reviewers who are at a critical point in developing their careers. Some service in the form of reviews is helpful in career development, especially improving grantsmanship, but too much service can be a detriment to developing productive research careers.

A.4 Questions concerning the resulting portfolio of awards under review. Provide comments in the space below the question. Discuss areas of concern in the space provided.

RESULTING PORTFOLIO OF AWARDS	APPROPRIATE, NOT APPROPRIATE, OR DATA NOT AVAILABLE
Overall quality of the research and/or education projects supported by the program. The quality of research and education projects supported by PGR continued at a very high level during this period. PGR aims to fund projects that contribute significantly to our understanding of plant genome structure and function, with the goal of impacting economically important plant systems. Awards have consistently targeted organisms and research objectives well-suited to reaching this goal. PGR projects have contributed many publications to the primary literature, and many databases and other sophisticated research resources have been created for broad use by the research community. There are many examples of fundamental scientific advances that have been achieved through PGR-funded programs. Outstanding achievements during 2001-03 include a detailed picture of maize genome architecture (Messing DBI-0211851; Coe DBI-9872655; Schubert DBI-0221536; and others), which will enable new efforts including complete genome sequencing of maize. Other important high-quality awards were aimed at completion and annotation of the rice genome sequence (McCombie DBI-0321683; Buell DBI-03210321538; and others). The overall quality and synergy between PGR programs is reflected in Dubcovsky's success in cloning the VRN1 and VRN2 genes of wheat. Some plants, including wheat, remember that they have had prolonged exposure to cold and then flower the following spring. This adaptation, known as vernalization, prevents premature flowering. The genes required for vernalization, *VRN1* and *VRN2*, were cloned by Jorge Dubcovsky with support provided in part through DBI-9975793. His cloning strategy relied significantly on colinearity among the cereals and made creative use of data generated by other PGR programs. This work will allow identification of the cold sensing mechanism and signal transduction pathway and could lead to the development of agronomically enhanced crops. Detailed genomic analysis of plant-associated microbes, such as the rice blast fungus (Dean 0211807), is a well-conceived addition to which will greatly amplify the utility and impact of plant genome sequences. PGR's approach of supporting cooperative projects has succeeded in generating results that have broader use and greater impact, in large part due to the emphasis that has been placed on developing high quality management plans backed up by regular monitoring of progress. The quality, utility and long-term availability of databases and bioinformatics resources remains a concern, even though PGR programs have made significant progress on this front.	Appropriate

PGR aims to foster integration of research and public education, and to promote the inclusion of under-represented minorities in scientific pursuits. The strong emphasis that PGR has placed on achieving broader impacts has gradually improved the quality of education and outreach components written into investigator's proposals. An outstanding example is Himmelblau's genomics curriculum development now being supported through a supplemental grant to Osbourne (DBI-0350544). There is still room for improvement in the quality and impact of outreach programs associated with PGR awards, and the COV notes that investigators could use professional support and mentoring that goes beyond availability of the PGR Outreach Portal (Brendel DBI-0110254), which nevertheless represents an important advance. In particular, it appears there is a need for fresh ideas for PGR public education programs, together with ways to measure their impact. PGR might be well served by drawing on sources of professional expertise in development of outreach programs. Some investigators and reviewers seem confused about whether routine training of graduate students and postdoctoral fellows is relevant to criterion 2. The POs have worked hard to emphasize that routine training does not fit into criterion 2, but additional emphasis may be necessary.	
Are awards appropriate in size and duration for the scope of the projects? (Table 7 and Graph 2) The size and duration of the projects was generally appropriate in the sense that these parameters generally matched the proposer's requests. The duration of the awards closely matched that requested, the size of the awards, taken together, ranged from 95% (2002) to 77% (2003) of the amounts requested.	Appropriate

Does the program portfolio have an appropriate balance of: • High Risk Proposals? (Table C.4) While there is no obvious "right number" of High Risk Proposals to fund, the COV commends the PGRP for supporting high-risk proposals in order to promote the development of tools and to expand the scope of the program. For example, PGRP has supported proposals that seek to extend the capabilities of gene targeting (e.g., Drews, DBI-0110035, University of Utah), identify expressed genes (e.g., Meyers, DBI-0110528, UC Davis), characterize the evolutionary genetics of weeds and crops species e.g., Weinig, DBI-0227103, University of Minnesota), discover conserved noncoding sequences in grasses and develop them as tools for gene discovery (e.g., Freeling, DBI-0337083, UC-Berkeley), automate identification of orthologs (e.g., Coruzzi, DBI-0346336, NYU), development of computational and statistical methods needed to compare genetic/physical maps/whole genomes of different species, extend the toolkit for functional genomics (e.g., Jorgensen, DBI-0331626, University of Arizona), and more efficiently sequence genespace (e.g, Schnable, DBI-0121417, Iowa State University). The COV also commends the flexibility of the program officers in their support of at least one high-risk approach put forth originally in a noncompetitive proposal, after it was submitted on a smaller scale with a smaller budget (e.g., Meyers, DBI-0110528, UC Davis). Recommendation: It would be helpful to have a better understanding of what characteristics of a proposal place it in the high risk category.	Yes
Does the program portfolio have an appropriate balance of: • Multidisciplinary Proposals? (Table C.4) The program balance of multidisciplinary projects is highly appropriate, and the proportion of multidisciplinary projects may increase as access to new research data make additional innovative approaches possible. It is critical for the program to continue to emphasize multidisciplinary proposals, despite the challenges posed by obtaining reviewers who can comment on both integrative science and multiple scientific approaches without conflicts of interest. The bioinformatics component of appropriate multidisciplinary projects should be integrated as a creative and scholarly contribution, rather than as a service.	Yes

Does the program portfolio have an appropriate balance of: • Innovative Proposals? (Table C.4) The COV found this term to be a bit confusing. Given the nature of the PG program we would expect most proposals to have innovative aspects to them. That only 5 of 107 proposals are considered to be innovative by the program is of concern if the definition of innovative is broad and includes tools, ideas, and people. We agree that the 5 proposals were innovative in terms of the problems they addressed (e.g. water efficiency from a genomic perspective). However, we found other proposals in the portfolio, beyond the high risk proposals, to have innovative approaches to problems. Again, the terminology is the issue, but it would be inappropriate if 102 of 107 proposals were considered to be banal rather than innovative. The jackets we reviewed during our visit generally reflected a high degree of innovation and creativity. Plant Genome Program PO 'rescued' one innovative proposal. The proposal had had tough reviews in two panels, partly because the PI, not in a regular tenure-track position, had requested a large grant. By funding the proposal initially for one year, with a renewal based on progress, a new robust method for identifying and quantifying rare messages has been validated for Arabidopsis. The resulting data base for this project was cited as excellent in the November (1) 2003 issue of Genetic Engineering news in a column on "Surfing for the Best Biotech information".	Not appropriate
Does the program portfolio have an appropriate balance of: • Funding for centers, groups and awards to individuals? (Table C.4) In terms of award dollars the distribution between Center and Group proposals and individual proposals appear to be reasonable by having 46% for the former and 54% for the latter. The portfolio also includes 12 workshop/symposium/conference proposals providing an excellent evidence for the PGRP community's proactive outreach activities. It is unfortunate, however, that only one workshop (Rocheford, 3/7-11, 2004) includes the bioinformatics component. This issue in conjunction with training and education should be addressed in a broader context.	Appropriate
Does the program portfolio have an appropriate balance of: • Awards to new investigators? (Table C.4) Eight out of 107 [7.5%] funded proposals were awarded to new investigators. Two additional high risk proposals were awarded to new investigators (Meyers DBI 0110528; Weinig DBI 0227103). A larger number of new investigators may be Co-PI's on some of the big awards. It would be worthwhile to track this type of information regarding Co-PI's for the next COV. We recognize that the program has tried to address this through the YIAs that were not ultimately a good mechanism and by including CAREER proposals in the program. Continued efforts are encouraged.	Needs improvement

Does the program portfolio have an appropriate balance of: • Geographical distribution of Principal Investigators? (Map 1 and Table 10) During the period 2001-03, the geographical distribution of PI awards was clearly non-random. Five states received 50 of the 108 total awards. Twenty-seven states plus D.C. received 58 awards. Eighteen states received no awards. However, this distribution roughly mirrors the geographical distribution of reviewers, and to a large extent it probably reflects the geographical distribution of science professionals who are qualified to serve as leaders on PGR cooperative projects. This non-randomness is not necessarily undesirable, nor is it necessarily a reflection of regional bias. The top states (CA, IA and NY) span the nation, and they are home to a number of top plant science research institutes. It is not surprising that they have garnered a large share of the awards. Some states stand out as having strong plant science research capabilities, but low success rates at PGR. For example, it appears that Texas succeeded with only one of 21 proposals. The COV wonders why Texas experienced such a poor success rate and encourages the POs to reach out to Texas investigators with guidance on writing more competitive PGR proposals. PGR awards extended to all the country's major geographic regions. As PGR progresses with the objective of including smaller and more diverse institutions, the geographical distribution should become less skewed. The COV notes the geographical distribution of co-PIs may differ significantly from the distribution of lead PIs. Geographical distribution may significantly affect allocation of overhead funds, and PGR should take care to ensure that indirect costs are covered appropriately at all participating institutions.	Appropriate
Does the program portfolio have an appropriate balance of: • Institutional types? (Table 8 and Graph 3) There is an appropriate mix of institutional types for projects of this scope and complexity. However, more opportunities should be taken to promote educational aspects of this program and support research activities at smaller schools. The committee noted the 4-year institutions received none of the funding in 01 and 02, and $125K in 03.	Generally appropriate

Does the program portfolio have an appropriate balance of: • Projects that integrate research and education? (Jacket Review) There are proposals that are exemplary in their integration of research and education. For example, Stapleton's project (DBI-0234250, University of North Carolina, Wilmington) developed a data mining approach and tested the prototype system in a course at NCA&T. This required significant input from the undergrad assistant who was key to the development of the system and it directly involved students in an undergrad course in the development of bioinformatics software. These students were "excited about using bioinformatics tools and felt they learned a lot about what kinds of questions could be asked of a database". The undergrad assistant was offered a job before her graduation date. Another notable example is the project of Carpita (DBI-0077719, Purdue University), who is developing a collection of plant cell wall mutants for the whole research community. High school students joined the group to characterize the cell wall sugar composition of a pectin mutant, the first to be identified directly from this type of screening analysis. Through this project students participated in cutting edge plant research using approaches that are still under development.	Data not fully available. We could not look at all 107 jackets.
Does the program portfolio have an appropriate balance: • Across disciplines and subdisciplines of the activity and of emerging opportunities? (Jacket Review) NSF has done an excellent job of creating a plant genome research pipeline that involves communities in establishing priorities, identifying model organisms, developing research tools such as databases and bioinformatics tools, and fostering workable approaches to improvement of crops of economic importance. Crop-specific research involving ESTs, microarrays, and expression data approaches will energize communities and provide needed impetus to the translational process. It will always be important to leverage research programs from discoveries and tools forged from work with model organisms. The portfolio is balanced for a program with a mission of undergoing portfolio change as it achieves milestones. The COV encourages continuing attention to creative bioinformatics projects designed to increase broad access to data. In addition, projects to enhance cross-talk among databases through development of common standards will also be an important priority.	Yes

Does the program portfolio have appropriate participation of underrepresented groups? (Table 9.a and 9.b) There have been minimal proposal submissions from HBCU, tribal colleges, and Hispanic-serving institutions (1 proposal, that was declined, in the 3 year period). Currently, individual PIs are asked to self-identify as members of underrepresented groups, so there are limitations to that data set. From 2001-2003 9 PIs from underrepresented groups applied for a grant and 3 were funded (2 in 2001, 1 in 2002, and 0 in 2003). The success rate is excellent, but an increase in proposal submission would be highly desirable. The COV noted that most of these institutions are not PhD granting institutions and that proposals may be of a different nature, with a limit to the number of large-scale Plant Genome proposals that might be expected from PIs from these institutions. The COV is not cognizant of the total number of researchers from underrepresented groups that would have the expertise to apply to this program. The POs might consider offering a grants workshop at the SACNAS meetings to provide information to Hispanic and Native American researchers. The POs might consider working closely with the ARD, an association of 1890s institution research deans with a focus on agriculture. Increasing involvement of underrepresented groups is an issue far broader than just this program and the POs have a sincere commitment to increasing representation from all groups. Their open meetings at site visits and interactions with participants at meetings are excellent venues to inform and possibly connect individuals from underrepresented groups with funded projects that are synergistic. Funding outreach efforts like Barbara Baker's (0218166) work with Makah tribe students in Washington State on the origins of native potatoes, not only introduces an underrepresented group to a relevant and exciting scientific question, but provides them with the opportunity to visit TIGR and the NSF, seeing science in a broader context. This is a forwardlooking approach to increasing the long-term participation of underrepresented groups in plant genomics.	Contextually appropriate

Are the programs in the program relevant to national priorities, agency mission, relevant fields and other customer needs? Include citations of relevant external reports. (background materials) The National Plant Genome Initiative (NPGI) initiated in 1988 has been dramatically reshaping the way scientists develop infrastructure for plant genetic, molecular, and applied research. NPGI is coordinated and managed by the Interagency Working Group (IWG) on Plant Genomes that include NSF, USDA, DOE, USAID, NIH, OSTP, OMB and NASA. NSF amplifies this national initiative in three aspects, develop diverse workforce, promote innovative ideas and develop research and education tools. During the five years period of PGRP, NSF has funded projects involving 487 investigators and 28 international collaborators from 23 institutes. These research projects emphasize developing the community's scientific knowledge on economically important plants. These funded projects appear to be very appropriate to the national priorities and the agency's mission. In terms of new ideas, DBI-09975866 (PI: Baker, University of California, Berkeley) reported the successful cloning of the RB locus from a wild potato relative (*Solanum bulbocastanum*). This discovery provides a potential to develop new ways of transferring durable pathogen resistance into cultivated potatoes (e.g., resistance to *Phytophthora infestans,* the causative agent of Late Potato Blight) DBI-0077709 (PI: Wessler, University of Georgia) has made spectacular discoveries while analyzing the emerging rice genome sequence for transposons. This project was a co-discoverer of the first active transposons identified in rice (http://www.nsf.gov/od/lpa/news/03/pr0303.htm). The research was published in the Jan. 9 edition of the journal Nature. This discovery may provide an explanation to rice to grow in temperate climates. DBI-0077774 (PI: Osborn, University of Wisconsin, Madison) published a significant paper in the journal Science in 2002 describing the mechanism by which plants silence genes. This discovery may offer a new way of shutting off expression of genes at will, and yet that is potentially reversible.In terms of broader impact on human resource development, DBI-0077719 (PI: Carpita, Purdue University) reports that high school students benefits from working with the investigator conducting a science fair project. DBI-0110296 (PI: Cook, University of California, Davis) uses a supplement to develop a successful teaching module available to high school teachers of plant genomics (http://www-ceprap.ucdavis.edu/). DBI-0110254 (PI: Brendel, Iowa State University) develops a web-based resource for distribution of information about Plant Genome Research Program outreach and training opportunities (http://www.plantgdb.org/pgrop/pgrop.php). DBI-0115642 (PI: Dean, North Carolina State University) allowed the six participating institutions to support students through Research Experiences for Undergraduates (REU) in which students can work at more than one site to get a broader perspective of the project. DBI-0341409 (PI:Tanksley, Cornell) supported a workshop held in November 2003 to bring together researchers from around the world to discuss how dispersed genomics efforts on members of the Solanaceae (including tomato, potato, tobacco, eggplant, and pepper) could be integrated to provide a Solanaceous genome resource. The meeting produced a detailed 10-year roadmap (SOL: http://www.sgn.cornell.edu/solanaceae-project/) that could be an excellent resource for coordinating research in this important group of plants.	Appropriate

Comments on the quality of the projects or the balance of the portfolio:

Since the last COV, the portfolio has expanded to include a number of efforts in functional genomics and metabolomics. The broader range of plants has been examined over the last three years. The COV recognizes the **continued extremely high quality** of the projects supported by the PGR.

Recommendations:

1. That the program monitor inclusive approaches to communities and continue to require documentation of substantial consensus and priority development from communities before funding new initiatives.
2. Clearly we would like to increase and enhance the integration of research and education. If not already done so in the program announcement, it would be useful to point prospective PI's to the website describing Plant Genome Research Program outreach and training opportunities (http://www.plantgdb.org/pgrop/pgrop.php).
3. While database and biocyber infrastructure are much broader than this program, an tremendous amount of data are being generated by research funded by PGR. The COV believes a proportionate amount of PGR awards portfolio should directed to support the growth of biocyber infrastructure, under the guidance of the BIO Cyber Working Group.
4. The COV recommends here, as elsewhere in this report, that efforts be made to include more agronomically important crops from Africa and other underdeveloped portions of the globe as these crops will be critical to global health and well-being.

A.5 Management of the program under review. Please comment on:

Management of the program:
The COV is impressed with level of coordination between the various PO's that contribute to this program. They balance a heavy workload and back each other up. Three PO's now contribute to PGRP, with Jane Silverthorne as the lead PO. There are also liaisons with appropriate PO's in the Research Resources cluster such as Manfred Zorn, a program director in Biological databases and informatics. If program continues to grow and if there are significant increase in applications to the program, staffing levels should be adjusted to keep pace with this growth.

There was some concern among the COV about the amount of responsibility for institutional memory that Jane Silverthorne carries. Everything possible should be done to anticipate PO openings (rotator), and allow for meaningful periods of overlap and training.

Responsiveness of the program to emerging research and education trends:

PGR Program Solicitations and awards during 2001-03 show the program has responded well to emerging research and education trends. The last COV recommended increased support for bioinformatics and proteomics. Both topics were addressed in subsequent PGR Program Solicitations, which also called for proposals in new areas such as evolutionary genomics and plant-associated microbes. Seven awards were made to projects focused on proteomics and metabolomics. Remarkably, these early projects range across 11 different plant species, which may be spreading the effort rather thin for optimal synergistic outcomes. Six bioinformatics projects were funded, with an emphasis on development of tools that will broadly serve the plant genomics community. A good example is the development of PlantGDB (Brendel DBI-0110254), a broadly useful suite of plant genome database and analysis tools that leverage EST sequences to enhance gene annotation. The Plant Ontology Consortium (Stein DBI-0321666) is working with the Gene Ontology Consortium to develop a vocabulary that will permit the development of informatics databases that span the plant kingdom. Six awards were granted for research on the functional genomics of economically important host-pathogen interactions, including diseases of rice and soybean as well as several model systems. Six "High Risk" projects were funded to explore new ideas for achieving valuable but difficult objectives such as gene targeting in plants (Drews DBI-0110035). Other interesting but risky new ideas were explored through SGER grants, providing further testimony to the innovative spirit of PGR. Educational and outreach projects have gained significantly more attention in PGR-sponsored projects during the review period. There are numerous examples of well-conceived educational and outreach plans designed to reach audiences ranging from K-12 to undergraduate and graduate students, and even to the general public (see Carpita DBI-0077719; Cook DBI-0110296; Brendel and Himmelblau DBI-0110254; Dean DBI-0115642; Schnable and Ashlock DBI-9975868). Some of these involve participation by outreach professionals, but it appears that most do not. Evidence- and Outcome-Based design of learning programs is an important trend among outreach professionals, who generally incorporate measures of success in their programs from an early stage. Proposers should be asked to provide an indication of the intended outcomes of their education and outreach programs, as well as associated indicators of success. As Criterion 2 objectives assume a larger part of the energy and resources devoted to PGR programs, it seems incumbent on PGR and the NSF to ensure that the associated projects are well-designed and that intended impacts are being achieved.

Program planning and prioritization process (internal and external) that guided the development of the portfolio under review:

The planning and prioritization process appear to have been extremely well done, based on the quality of the proposals received and accomplishments documented in the first few years of this program. The COV commends the Plant Genome program staff for their success in planning and prioritization. The initial planning process envisioned projects that addressed:

- Functional genomics with an emphasis on networks of genes
- Genomics of plant-associated microbes
- Environmental genomics
- Development of new research tools
- Informatics tools
- Genomics of economically important plants
- Novel, risky ideas

With the exception of plant-associated microbes, these targeted areas were fully addressed or are in the process of being addressed in the ensuing funded proposals. Emphasis on plants and plant processes of economic importance resulted in the exclusion or re-direction of projects involving lower plants, microbes or Arabidopsis. In 2003, the program was broadened into three categories of proposals; 1. large distributed collaborative projects, 2. individual and small group awards and 3. young investigator awards. Young investigator awards were discontinued in 2004.

The Program staff has shown admirable flexibility in responding to opportunities and adjustments in the program as they arise. Program officers have intervened to pull disparate projects together into a collaborative whole to the credit of all involved. An example is the supplementation of a Cooperative Agreement to include proposed educational activities of Dr. Edward Himmelblau, Long Island University (DBI-0350544). This kind of activity has resulted in an overall higher quality of the Program.

Results of this program effort have been excellent, particularly in view of its recent beginnings. There have been a significant number of patents filed, papers published in major research Journals, and data and research tools broadly disseminated over the Internet over the past 4 years. A number of funded projects in each of the areas of functional genomics, environmental and comparative genomics, metabolomics and proteomics, plant-pathogen interactions and genomic toolkits and what are termed "high risk projects" have all contributed to the overall Program goals.

Discuss any concerns identified that are relevant to the management of the program:

The committee notes the excellent management of PGR, but concerns were expressed that as the program continues to achieve success and expand, delegation of key functions to junior POs will be necessary. Concrete plans for future management as the program grows and diversifies are important for protecting the gains that have been made to date.

Recommendations:

The management of the Plant Genome Program has wisely emphasized development of database standards and bioinformatics tools from its very inception. However, the overwhelming success of the program is generating an avalanche of data that is increasing exponentially and catalyzing multidisciplinary projects. This has placed pressure on database development and capabilities. The COV recommends the Program Management ramp up the speed of developing commonly accepted

standards for database design and bioinformatics tools. This is an extremely important activity because database cross-talk and user-friendly bioinformatics tools will only enhance the research enterprise if they can be applied by scientists from a wide variety of fields, such as systematists, morphologists, evolutionists, and agricultural crop scientists. Databases must be able to integrate and provide easy access to a wide variety of resources (from museum specimens to germ plasm banks to seed collections to geographic and climatic information) so that the data resulting from the Plant Genome Program can be fully exploited. The results of this effort will affect all databases in the biological sciences, so it is most appropriate if consensus development is carried out at the level of the BioCyberInfrastructure Working Group. Though it will be challenging to develop common standards that will be flexible enough to accommodate future innovations in science and in bioinformatics, this process is critical to maximizing the impact of the groundbreaking successes of the Plant Genome Program and therefore is a top COV priority.

An increase in creative proposals to develop bioinformatics tools and database design should result from such a shift in priorities, and this will influence the balance in the portfolio.

The program should continue to encourage grant applicants to include partnerships with primarily undergraduate institutions. For instance, projects developing large databases or bioinformatics tools could partner with faculty at minority-serving or other undergraduate colleges to evaluate accessibility and potential adaptation to student research projects. This will in turn increase the pool of students who are excited about potential careers in bioinformatics.

PART B. RESULTS : OUTPUTS AND OUTCOMES OF NSF INVESTMENTS

NSF investments produce results that appear over time. The answers to the first three (People, Ideas and Tools) questions in this section are to be based on the COV's study of award results, which are direct and indirect accomplishments of projects supported by the program. These projects may be currently active or closed out during the previous three fiscal years. The COV review may also include consideration of significant impacts and advances that have developed since the previous COV review and are demonstrably linked to NSF investments, regardless of when the investments were made. Incremental progress made on results reported in prior fiscal years may also be considered.

The following questions are developed using the NSF outcome goals in the NSF Strategic Plan. The COV should look carefully at and comment on (1) noteworthy achievements of the year based on NSF awards; (2) the ways in which funded projects have collectively affected progress toward NSF's mission and strategic outcomes; and (3) expectations for future performance based on the current set of awards. NSF asks the COV to provide comments on the degree to which past investments in research and education have contributed to NSF's progress towards its annual strategic outcome goals and to its mission:

- To promote the progress of science.
- To advance national health, prosperity, and welfare.
- To secure the national defense.
- And for other purposes.

Excellence in managing NSF underpins all of the agency's activities. For the response to the Outcome Goal for Organizational Excellence, the COV should comment, where appropriate, on NSF providing an agile, innovative organization. Critical indicators in this area include (1) operation of a credible, efficient merit review system; (2) utilizing and sustaining broad access to new and emerging technologies for business application; (3) developing a diverse, capable, motivated staff that operates with efficiency and integrity; and (4) developing and using performance assessment tools and measures to provide an environment of continuous improvement in NSF's intellectual investments as well as its management effectiveness.

B. Please provide comments on the activity as it relates to NSF's Strategic Outcome Goals. Provide examples of outcomes (nuggets) as appropriate. Examples should reference the NSF award number, the Principal Investigator(s) names, and their institutions.

B.1 <u>OUTCOME GOAL for PEOPLE</u>: Developing "a diverse, competitive and globally engaged workforce of scientists, engineers, technologists and well-prepared citizens."

Comments:

An important outcome of the investment by NSF in proposals funded through PGRP is the enhancement of diversity in the scientific community and the enhancement of scientific literacy in the community at large. An exemplary effort is being made on both fronts by the project of Vanwijk (DBI-0211935, Cornell University). Specifically, in 2003, seven minority students from local high schools were placed in four different labs for summer internships. This project has developed an ongoing relationship with administrators at local schools that will serve as pipeline to recruit students on a regular basis. At a minimum, these students will go away from their experience with a much better

understanding of genome science, and at least some will be captivated and pursue a career in science. A focus on recruitment of minorities at the undergraduate level is a notable feature of the Keegstra project (DBI-0211797). Three minority women who are undergrads at Albany State University in Georgia or Tuskegee University were recruited for summer projects in the Keegstra and Walton (Co-PI) labs where they worked closely with postdocs. Many other projects have exemplary modules for contributing to the scientific literacy of those in their communities. The project of Wendel (DBI-0211700, Iowa State University) has led to the development of a full day, six-week summer program in 2003 for high school teachers to gain hands-on experience in plant genomics. An independent evaluation of this program gave it high marks and many teachers left the program with new lesson plans and experiments for their classroom curricula. Another project directly serving high school teachers is that of Cook (DBI-0110296, University of California, Davis). PGRP awarded Cook a supplement to expand a successful teaching module available to high school teachers of plant genomics (http://www-ceprap.ucdavis.edu/). Activities will build on a program that includes equipment loans and training workshops for local teachers. In addition, the teaching modules will be available on CD or from the web so that anyone can use them. Classrooms across the country will be able to take advantage of the tools developed and implemented locally in Davis, CA. A supplement to the project led by Patrick Schnable (DBI-9975868) at Iowa State University supported co-PI Daniel Ashlock to mount "family math night" in local schools. A "family math night" is a 2-4 hour evening event with displays, activities, and demonstrations that serve to expand both children's and parents' appreciation of mathematics. In addition to the evening's events, handouts with challenge problems are given out. Other PGRP projects are reaching a broad community of scientists and nonscientist through their development of electronic resources and databases. One example is the project of Brendel (DBI-0110254, Iowa State University), who was supplemented to allow development of a web-based resource for distribution of information about Plant Genome Research Program outreach and training opportunities (http://www.plantgdb.org/pgrop/pgrop.php). This resource allows successful programs to be replicated and implemented across multiple projects and to reach a broader group of students, teachers, and investigators than otherwise possible. The renewal project (DBI-0321600) is continuing development of this resource.

B.2 OUTCOME GOAL for IDEAS: Enabling "discovery across the frontier of science and engineering, connected to learning, innovation, and service to society."

Comments:

PIs are engaged in a range of outreach and educational activities that are now more readily shared through the PGROP web site (http://www.plantgdb.org/pgrop/pgrop.php). There is published literature supporting the positive gains of students who participate in undergraduate research and awards such as 0110189 that have use REU supplements to involve undergraduates in research.

During the period of time covered by this COV review, substantial progress has been made in a broad range of cutting-edge scientific investigations relevant to the structure and function of plant genomes. The pursuit of this diverse portfolio of projects has involved many cooperative and multidisciplinary efforts. Successes have included development of functional genomics tools that are accelerating the pace of future research and the generation of new knowledge on genome structure and function. The program is to be complimented on the substantial proportion of "high risk" projects that have achieved success. For instance, (Henikoff, DBI-0077737) developed tools that are applicable to generating mutants in any gene in a range of different organisms.
Well-designed pilot projects have paved the way for subsequent large-scale projects. In maize, two exploratory approaches to focusing on gene-rich regions were tested before moving into a full-scale sequencing effort (Schubert, DBI-0221536, Bennetzen, DBI-0236505). The portfolio of funded projects represents substantial innovation and will contribute to future development of economically

important plants that minimize environmental damage, maximize production under difficult growing conditions, and provide improved nutrition and other biological products for the benefit of society.

B.3 OUTCOME GOAL for TOOLS: Providing "broadly accessible, state-of-the-art S&E facilities, tools and other infrastructure that enable discovery, learning and innovation."

Comments:

One major deficiency pointed out during the last COV cycle was the bioinformatics area. The current COV is pleased to see the improvement in this area. DBI-0077622 (PI: Collmer, Cornell University) sequenced, assembled, and annotated the genomes of *Pseudomonas syringae* DC3000 the related *P.s phaseolicola* 1448A. DBI-0211732(PI: Henikoff, F. Hutchinson Cancer Research Center) and DBI-0321510 (PI: Henikoff, F. Hutchinson Cancer Research Center)extended Targeted Induced Local Lesions IN Genomes (TILLING), a powerful method for finding point mutations in any gene in any organism for rice and maize, respectively. DBI-0110528 (PI: Meyers, University of California, Davis and now University of Delaware) used a new and risky technology called Massively Parallel Signature Sequencing (MPSS) to isolate short sequence tags for expressed genes in Arabidopsis in a highly automated fashion. This project made 14 different libraries and in them has found evidence for more than 1,100 genes that were previously only recognized though gene prediction programs. The data are freely accessible though the project database (http://ag.udel.edu/meyers_lab/index.html). DBI-0211851 (PI: Messing, Rutgers University) has developed a detailed physical map integrating over 600,000 Bacterial Artificial Chromosome (BAC) clone fingerprints and over 450,000 BAC-end sequences. DBI-0221536 (PI: Schubert, Donald Danforth Plant Science Center) has been able to isolate and sequence genomic tags for more than 90% of the genes in maize. These genomic fragments have been assembled together with EST sequences from earlier PGRP projects and placed onto the genome map (see: http://www.tigr.org/tigr-scripts/tgi/maize2_genome.pl). The resulting resources are now enabling maize researchers to find and map genes of interest and will allow development of strategies to sequence the genome. DBI-0313887 (PI: Ronald, U. of Cal. Davis) and DBI-0321663 PI: Chandler, U. of Arizona) are developing develop and distribute community oligomer arrays for rice (http://www.ricearray.org/) and maize (http://www.maizearray.org/), respectively. Researchers will be able to either order the arrays and perform the hybridizations and analyses themselves or have them performed by the project.

This COV believes that the program has made a good first step but it should continuously emphasis the importance of the bioinformatics aspect of the program in particular in the area of preserving the data produced from the program. Specific recommendations are presented in C.1 and D.1.

PART C. PROGRAM LEVEL QUESTIONS

C. 1. What role should the program play with respect to data standards and databases as the community transitions from a project-specific to plant-wide model?

High throughput data production by multiple investigators makes long-term preservation of data one of the highest priorities of the program. Ideally it should be in a format that is portable to other computer platforms (e.g., XML) from the long-term preservation viewpoint. The data in electronic form should survive beyond the existence of the program. The following suggestions are made.

a. Promote use of existing standards (e.g., GO ontologies, Plant Genome Ontology, etc) in community-wide annotation
b. Promote use of standard data exchange formats (e.g., MIAME for microarray) in data submission
c. Promote and encourage submission of strictly bioinformatics proposals as long as database and tool development is relevant to the program
d. Strongly encourage participation from the bioinformatics community that develops general purpose and advanced form of bioinformatics tools as long as the relevancey to the program is demonstrated.

e. Consider creating a community wide bioinformatics core center who might play a neutral position amongst the plant genome communities. Use cooperative agreement as a way of enforcing data submission to the database core. Utilizing supplement mechanism could be another mechanism for encouraging data submission.

C. 2. How should the effectiveness of the outreach and educational activities of Plant Genome be evaluated? Are there any opportunities we are missing or communities we are not serving?

PIs are engaged in a range of outreach and educational activities that are now more readily shared through the PGROP web site (http://www.plantgdb.org/pgrop/pgrop.php). There is published literature supporting both the affective and cognitive gains of students who participate in undergraduate research. Such gains include increases in knowledge, recognition of career choices, feelings of mastery and satisfaction, and recognition of their ability to make a contribution in science. Awards such as 0110189 (Hake) that have use REU supplements to involve undergraduates in research where they directly interact with a senior scientist are exemplary, as is Barbara Baker's (0218166) program with Makah students. Other outreach and educational activities have been thoughtfully commented on by reviewers and implemented at different scales. Community buy-in to outreach and educational activities as part of their NSF awardee status has grown substantially since the last COV. The COV is pleased that the program is considering how to effectively evaluate these activities. It is curious that for an evidence-driven field like biology, there is so little data and effort devoted to evaluating the effectiveness of educational activities and outreach activities. It would be very helpful for PIs to have access to studies demonstrating the effectiveness or lack of effectiveness of different activities in achieving specific goals (e.g., retention in science, cognitive gains in specific areas, and affective changes). At the program level a continued requirement for assessment of these activities and coordination and distribution of these study results on the PGROP site would be an important first step. However, it should be kept in mind that science training does not prepare an individual for educational assessment, including randomized control groups and working with their institutional IRBs because human subjects are involved.

Fundamentally, however, this is a Foundation-wide challenge that should be approached at that level. NSF is in the best position to decide whether funding an assessment of its many educational and outreach initiatives (GK12, math science partnership programs, RET program, supplement programs, etc.) is the most effective way to learn, in an evidence-based way, what works in achieving specific goals. This is exactly the right time to approach evaluation systematically and share these findings with the scientific community. NSF provides examples of outreach and both PIs and reviewers are increasingly taking this commitment to heart. It is imperative that they have adequate information and resources to make evidence-based decisions about how to devote their precious time to outreach activities. It is inefficient and ineffective for each program to be struggling with these very important issues on their own in an institution with both the resources to address the issue and prevent the perpetual reinvention of a wheel with potentially broken spokes.

The NSF has substantial expertise in the SBE and EHR that could be used cooperatively with BIO and other directorates to address funding or contracting of studies on the effectiveness of a set of outreach and educational activities in achieving specific goals. For example, picking a goal like learning about the nature of science as stated in the NRC National Science Education Standards or AAAS Benchmarks for a K-12 and outreach (e.g. museum work) would be a place to start. Goals such as helping students like science would not be such a good choice, since it would not align with educational standards and would pose additional implementation hurdles. Carefully designed studies and analyses could then be shared with the research community via NSF to guide the development of criterion 2 activities. The COV believes that the best way to strengthen outreach and educational activities is to provide PIs in the Plant Genome program with examples that have been demonstrated to achieve specific goals successfully. This would enable them to implement outreach and educational activities about which they are passionate without hindering them or the program officers with the need to overlay this with a complex evaluation/assessment. The PIs would still report on what they had done and accomplished. The POs would continue to track these outcomes and everyone would have greater confidence that the financial and human resources devoted to these efforts were of high value.

C. 3. Is the program maintaining a proper balance between tools/resource development and biological research activities?

Eleven out of 107 funded projects could be identified from their titles as concerned with tools for wide application and resource development. It is likely a much larger number of projects would also fit under this rubric 'tools/resource development'. While tool and resource development should continue, as this discipline matures a larger proportion of the work should shift to biological research activities. This transition is noted in the number of functional genomics, metabolics, ecological and evolutionary genomics proposal funded in the 2001-2003 period.

C. 4. Are there missed opportunities in plant genomics that the Program should address?

It is clear that the knowledge gained on major crop species (rice, maize, sorghum, tomato) have made an enormous impact on plant science with these species. As this program matures, research activities need to be broadened to include other agronomically important crops, including those that are vital to populations in Africa (cassava and other indigenous crops) other places with pressing food needs. The supplemental program to bring international scientists to the US for training with PGR investigators (with USAID) is an excellent start.

PART D. OTHER TOPICS

D. 1 Please comment on any program areas in need of improvement or gaps (if any) within program areas.

The overall progress of the Plant Genome Program will benefit from a long-term vision of database evolution. It might be appropriate to institute an ongoing series of planning workshops addressing, for instance, bioinformatics and database issues that will respond to needs of researchers working on functional expression networks. Additional aspects might include working to achieve community consensus on interface formats and approaches; competitive tool shoot-outs, and/or partnerships to carry out classroom evaluations by faculty in four-year schools and undergraduates.
Fostering of creative bioinformatics programs would increase the perception of bioinformatics as an academic field (as opposed to service delivery) and assist with recruitment of more individuals into that much-needed specialty area.
Universal attention should be paid to fostering access to databases and bioinformatics tools to the larger community. Posting of data is already being handled very well, and this provides important resources to researchers. It is timely to begin a shift in emphasis from fast posting of raw data to broad tools for accessibility. This will in turn result in broadening the user base. To this end, grantees might query database users to monitor the success of such efforts.

As noted, the gains resulting from PGRP have fundamentally changed plant biology forever. One of the real gains is the creation of a broader interface for the interaction of genomics with evolutionary biology. Currently, the community of evolutionary biologists has benefited greatly from the development of tools by the genomics community. It is possible that there are more opportunities for the genomics community to gain from insights that emerge from comparative approaches. Comparative approaches are essential for the functional classification (at least at a first pass) of members of gene families and for refining the models that are built in model systems. And for many important plant processes, models inferred from one system have broad but limited applicability in other systems. In some (many?) instances, this may hinder the explicit goal of improving economically important species because a strategy designed to improve a process may not work in all species. It is important to address this limitation by sampling broadly among species in a thoughtful way. This is happening in at least some studies, such as those that focus on functional genomics, evolutionary genomics, and in certain EST studies. But the sampling is still quite skewed. Often what is needed is a sample of three species, two that are closely related and a third more divergent species. This is easier to accomplish perhaps in EST studies and functional genomics, but this should also be kept in mind during discussions of how to broaden the scope of studies that seek to sequence whole genomes.

D. 2 Please provide comments as appropriate on the program's performance in meeting program-specific goals and objectives that are not covered by the above questions.

<u>Cyberinformatics</u>

The 2001 COV stressed the importance of bioinformatics-oriented projects for managing data, ensuring broad community access and providing a framework for standards in plant genome research. To date, there have been 6 proposals funded in the area of bioinformatics. Two grants (0110254) (0321600) were awarded to Volker Brendel to develop a database of Expressed Sequence Tags and gene splicing. David Allison's proposal (0217651) is directed to microarray analysis of gene expression in plants,

Of the broader-based proposals, Lincoln Stein's project (0321666) supports the operations of the Plant Ontology Consortium to develop a common set of terms describing ontology in Arabidopsis and

cereals, later to be extend to encompass legumes and Solanaceae. This project is intended to provide a framework allowing researchers to make meaningful cross-database queries across species-specific resources. Todd Vision's grant (0227314) is directed to describing features of biological sequences and mapping across multiple organisms that are related by a known phylogeny.This research is enhanced by the development of an online plant phylogenomics database called Phytome.

The COV views these projects as a good beginning to the essential task of managing databases in plant genetics and potentially across the spectrum of biology. However, the COV points out that the Plant Genome Project maintains an excellent cooperative spirit amongst the community scientists and presents an unique opportunity to go beyond the scope of the present funded research in developing informatics tools. Some of the specific strategies to accomplishing the goal has been suggested in Section C.1. Ultimately, these informatics tools should allow researchers to make correlations and connections between disparate databases containing information about taxonomy, physiology, ecology, predation, metabolic pathways, enzyme activity, and any other aspects of plant science, and more broadly, all scientific information that bears on the question at hand.

Programmatic efforts supporting community planning for plant genome efforts

During this period, PGR played a key role in helping communities of plant researchers come together to map the future of plant genomics within broad taxonomic areas, such as the Solanaceae, or wheat and its relatives among the Triticae. This has been accomplished by encouraging and funding workshops organized by key investigators who are natural leaders for the field, with the subsequent production of high-quality white papers. A good example is DBI-0341409, which supported a workshop held in November 2003 to bring together researchers from around the world to discuss how dispersed genomics efforts on members of the Solanaceae (including tomato, potato, tobacco, eggplant, and pepper) could be integrated to provide a Solanaceous genome resource. The meeting resulted in release of a detailed 10-year roadmap (SOL: http://www.sgn.cornell.edu/solanaceae-project/) that should lead to better international coordination of research in this important group of plants. Similarly, DBI-0344938 sponsored a workshop held in November, 2003 which brought together wheat researchers from around the world to develop a consensus and map out an international strategy for sequencing the genome of wheat. This workshop resulted in release of a white paper (http://www.ksu.edu/igrow/IGROW_workshop_report.html/) that will soon appear in a special issue of Genetics devoted to wheat genetics and genomics. These papers have established useful frameworks for development of synergistic programs. The ultimate power of this kind of integrating effort on the part of PGR is reflected in the status of today's maize genomics effort, which is poised for an efficient, effective and broadly supported genome sequencing initiative.

Response of other BIO Directorate programs to Plant Genome Program:

The COV met with program officers from other BIO divisions to more fully understand the interplay between this relatively large program and the core programs in the directorate. Four themes emerged:

1. The Plant Genome program has funded research that has benefited the plant biology community as a whole and this is reflected in the positive interactions between this program and the core programs in other divisions within the directorate. Expertise is shared in the search for reviewers and panel members. Plant Genome Venture Funds are viewed as an asset by core programs.

2. An emerging challenge for the core programs is arising from the success of the Plant Genome program. As investigators (both new and established) take advantage of the new, but more costly approaches, core programs are seeing proposals with substantially larger budgets that are appropriate for their programs, not the Plant Genome program. This is an exciting indicator of the success of the Plant Genome program, but creates some substantial funding challenges for the BIO

directorate that deserve careful consideration and planning. New investigators who did graduate work and postdoctoral work in large plant genome facilities appear to be coming to core programs with a set of expectations for how one does science that is both exciting and costly.

3. The core program officers report that some PIs are taking advantage of outreach programs established at their home institutions by Plant Genome PIs. This is an excellent use of shared resources that maximizes the impact of Plant Genome funds and the outreach efforts.

4. The need for integrated, interoperable database and bioinformatics resources was echoed by the POs we spoke with. The BIO Cyberinfrastructure Working Group represents a very important coordination of activities within the directorate. The Plant Genome program has both the experience and funding to substantially leverage these efforts in collaboration with the rest of the directorate and other relevant programs within the foundation.

D. 3 Please identify agency-wide issues that should be addressed by NSF to help improve the program's performance.

It is possible that a greater effort is needed toward communicating funding opportunities to potential young PI's from traditionally underrepresented groups. First, young PI's are generally less well informed about opportunities than are more senior PI's; second, they may be less well connected to the community of senior scientists who have a funding track record and who can identify potential mentors to assist in career in development and in the grant-writing endeavor.

Foundation-wide support of assessment of educational activities and outreach could result in a resource for PIs (see C2)

D. 4 Please provide comments on any other issues the COV feels are relevant.

Additional documentation of outreach and education is a critical issue. The current system does not provide adequate information for COV review, primarily because this information is not being requested from grantees in a systematic fashion. Every PI should be asked to gather, maintain, and report information to evaluate and track the impact of their program. This may mean that students will be tracked for several years. Privacy concerns must not be used as an excuse to avoid this responsibility because simple human subjects applications can make appropriate and ethical arrangements to gather this information. Without such data, the effectiveness of outreach and education efforts cannot be documented, and reliable information will not be available to inform future efforts. NSF must take the lead in ensuring that effective programs are documented and disseminated to the research and to the education community.

NSF might consider initiating an institution-wide tracking program to reinforce the broader impact component of a wide variety of projects (beyond tracking whether or not criterion 2 is addressed). It might also be productive to create a template for human subjects consent forms and to encourage appropriate projects to write publications documenting the success of their outreach programs and making the approaches available to the scientific community. There must be a long-term commitment to tracking so that changes in life paths will be documented.In order to prime the pipeline for young investigators, the program might consider dedicating a number of slots at the annual Awardee

Meeting to young investigators with independent faculty positions (or about to assume one) who are not yet Pis or Co-investigators on projects.

D. 5 NSF would appreciate your comments on how to improve the COV review process, format and report template.

- COV reviewers should be given access to all grant files whenever possible, rather than a selection of files. This gives COV reviewers the maximum number of options to review comprehensively or to choose selected files on a flexible basis.
- Arrangements for evening access to the building would permit COV reviewers to have additional flexibility, especially in being able to access the many individual grant files that must be reviewed.
- Several COV members had queried individuals in the plant genomics community to become informed about current issues. This information was very helpful, but was based on networking, rather than comprehensive information. We strongly suggest that grant applicants (including those funded and declined) be surveyed prior to each upcoming COV. An anonymous questionnaire might include items on the review process, on project administration, on success of particular approaches, and on perceptions of community needs that are not being met and are thus impeding research progress.
- The key information for COV members was the table of grants awarded. The usability of this table could be increased by providing information on the plant or plants utilized (as requested by this COV). In addition, a sorting mechanism would be extremely helpful so that reviewers could sort proposals according to any category. The color code was somewhat helpful, but the printer was black-and-white, requiring committee members to revisit the on-line color version to make notes. The color key was only available at the top of this multiple page document, so users had to scroll up and down several times to sort projects. Perhaps a different method (such as black-and white symbols or a code) to indicate the type of award could be instituted for the next COV.
- A spreadsheet of declined applications would have been very helpful for the COV reviewers.
- The COV members would have appreciated clarification of the "innovative" category of proposals.
- There is a great deal of very useful information and perspectives (white papers) available to the COV panel (http://nsf.gov/bio/dbi/pgcov04/GPRA%20Plans.htm), however much of the material wasn't accessible until we arrived in Arlington. Two and half day is not enough time for the COV to review that material and truly assess progress since the last COV. We realize there is a trade off with not wanting to overburden the COV but we felt could more helpful if we had two weeks in advance of the meeting with all the program announcements, reports and summaries.
- It would also be valuable to collect some breakdown information of applications or awards: for example we were asked about the portion of new investigators awards to overall awards. While some of these data were available in summary tables (i.e. Table C.4), this information is **only** provided for the PI. Without reading individual jackets it is not possible to determine if any co-PI's are new investigators. Similarly the involvement of HBC institutions and minorities in PGRP activities may be undercounted, if they participate as Co-PI's or in supplemental activities. For example Delmer DBI 0110173 had provisions for students from a HBC to work at in her lab during the summer. Improvement. This information was available only as excerpted in reports or 'nuggets'. More systematic data collection would be helpful to the COV.

OFFICIAL SIGNATURE BLOCK:

_____ _____

For the PGRP FY2004 COV
Susan Singer
Chair

Date

_____ _____

Forrest Chumley

Anita Klein

_____ _____

Sarah Mathews

Dong-Guk Shin

_____

Sally Tobin

As the designated representative to this COV and on behalf of the BIO Advisory Committee, I submit this report to the Assistant Director of the Directorate for Biological Sciences.

_____

Burt D. Ensley
Designated Representative of the BIO Advisory Committee

_____

Date