LETTER TO THE EDITOR

Bensen, R.J., Johal, G.S., Crane, V.C., Tossberg, J.T., Schnable, P.S., Meeley, R.B., and Briggs, S.P. (1995). Cloning and characterization of the maize *An1* gene. Plant Cell **7,** 75–84.

Braun, D.M., and Walker, J.C. (1996). Plant transmembrane receptors: New pieces in the signaling puzzle. Trends Biotechnol. **21,** 70–73.

Brown, R.C., Lemmon, B.E., and Olsen, O.-A. (1994). Endosperm development in barley: Microtubule involvement in the morphogenetic pathway. Plant Cell **6,** 1241–1252.

Brown, R.C., Lemmon, B.E., and Olsen, O.-A. (1996a). Development of the endosperm in rice (*Oryza sativa* L.): Cellularization. J. Plant. Res. **109,** 301–313.

Brown, R.C., Lemmon, B.E., and Olsen, O.-A. (1996b). Polarization predicts the pattern of cellularization in cereal endosperm. Protoplasma **192,** 168–177.

Clark, S.E. (1996). Plant cell communication: The world outside the plasma membrane. Trends Plant Sci. **1,** 406–407.

Doan, D.N.P., Linnestad, C., and Olsen, O.-A. (1996). Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus cell layers. Plant Mol. Biol. **31,** 877–886.

Drews, G.N., Lee, D., and Christensen, C.A. (1998). Genetic analysis of female gametophyte development and function. Plant Cell **10,** 5–17.

Engell, K. (1989). Embryology of barley: Time course and analysis of controlled fertilization based on serial sections. Nordic J. Bot. **9,** 265–280.

Friedman, W.E. (1994). Evidence of a preangiosperm origin of endosperm: Implications for the evolution of flowering plants. Science **255,** 336–339.

Johannsen, W. (1884). On the importance of the endosperm and its development in barley. Meddelelser fra Carlsberglaboratoriet, Kopenhagen, Bind **II,** 103–133.

Kranz, E., and Lörz, H. (1993). In vitro fertilization with isolated single gametes results in zygotic embryogenesis and fertile maize plants. Plant Cell **5,** 739–746.

Kranz, E., von Wiegen, P., Quader, H., and Lörz, H. (1998). Endosperm development after fusion of isolated, single maize sperm and central cells in vitro. Plant Cell **10,** 511–524.

McClintock, B. (1978). Development of the maize endosperm as revealed by clones. In The Clonal Basis of Development, S. Subtelny and I.M. Sussex, eds (New York: Academic Press), pp. 217–237.

Nusslein-Volhard, C. (1996). Gradients that organize embryo development. Sci. Am. **275,** 54–55.

Olsen, O.-A., Lemmon, B.E., and Brown, R.C. (1995). The role of cytoskeleton in barley endosperm cell wall deposition. BioEssays **17,** 803–812.

Olsen, O.-A., Brown, R.C., and Lemmon, B.E. (1998). A model for aleurone development. Trends Plant Sci., in press.

Opsahl-Ferstad, H.-G., Le Deunff, E., Dumas, C., and Rogowsky, P.M. (1997). *ZmEsr*, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. Plant J. **12,** 235–246.

Russell, S.D. (1993). The egg cell: Development and role in fertilization and early embryogenesis. Plant Cell **5,** 1349–1359.

Singh, H. (1978). Embryology of Gymnosperms. Handbuch der Pflanzenanatomie Vol. 10. (Berlin: Gebruder Borntraeger).

Stebbins, G.L. (1974). Flowering Plants: Evolution Above the Species Level. (Cambridge, MA: Harvard University Press).

van Lammeren, A.A.M., Kieft, H., Ma, F., and van Veenendaal, W.L.H. (1996). Light microscopical study of endosperm formation in *Brassica napus* L. Acta Soc. Bot. Pol. **65,** 267–272.

A Plant Genome Initiative

Over the last several months, we have attended various meetings and small conferences concerning the need for a coordinated and comprehensive plant genome initiative (PGI). From these meetings, and from our extensive discussions with plant and animal scientists with interests in genetics, agriculture, and science policy, we have been able to discern a general consensus among biologists both that a comprehensive program is desperately needed and that certain components of this program are indispensible.

Of course, with such a broad set of discussants, there were some lively debates on a number of technical points and some significant disagreements. Nevertheless, it is clear that many scientists feel that a comprehensive PGI will be essential if we are to maintain the agricultural competitiveness of the United States and to develop new sources of agricultural productivity at a time when demand for increased food production worldwide threatens to exceed traditional approaches to crop improvement. The fact that maize alone generates $80 billion annually in farmgate value that is converted into over $400 billion through value-added activities is just one indication of the importance of agriculture to the economy of the United States.

To a degree, the United States Government apparently recognizes the need for increased support in the area of plant genomics and is prepared to back that recognition up with new funding. Indeed, since we prepared our first drafts of this letter, the National Science Foundation (NSF) has announced a plant genome

LETTER TO THE EDITOR

research program (PGRP) to expand plant genomics beyond the ongoing Arabidopsis genome initiative (AGI). We welcome this development and initial indications that this support will be ongoing.

The motivation for the NSF PGRP came largely from the educational efforts of plant scientists and the National Corn Growers Association, whereas the conceptual and organizational framework came primarily from the recommendations of an interagency working group (IWG) appointed by the science advisor to the President on request of Senator Bond of Missouri. (The history of the legislative process can be viewed at http://www.inverizon.com/ncgi.) The IWG, which is chaired by Dr. Ronald Phillips, the current chief scientist at the United States Department of Agriculture, attended and/or received reports from many of the same meetings that we attended. Its final report is available via http://www.whitehouse.gov/wh/eop/ostp/html/ostp_home.html or from the National Science and Technology Council Executive Secretariat at 202-456-6120 (voice); 202-456-6026 (fax).

If properly focused and implemented, we believe that the NSF PGRP could provide the foundation for a comprehensive PGI. The AGI, which is also administered by the NSF, has been a successful program with tremendous value to all plant molecular biologists and geneticists. However, because it is based on only one (albeit important) plant species, the AGI cannot by definition be termed comprehensive.

In this Letter, we first describe the justification for a comprehensive PGI and then outline some priorities for implementation within a PGI. We go on to point out some of the areas of continuing disagreement regarding such priorities that we found within the plant science community. The hope is that in this way, we may help stimulate thinking on how to most efficiently perform, utilize, and promote plant genomics research.

Genome Technology Is Readily Available

Genomics technology can be applied immediately to agriculturally important crops. The rapid progress that is being made in the human genome project, and in similar projects for mice, nematodes, Drosophila, yeast, and Arabidopsis, indicates how quickly technologies for genome analysis are becoming more powerful and less expensive. In addition, the exceptional value of the information these projects are generating is being demonstrated in greater depth and in ever more unexpected ways on a day-by-day basis. For instance, it now has become possible to characterize the chromosomal organization of multigene families or clusters of related genes, as well as the actual gene density, on larger contiguous genomic sequences than ever before (Bevan et al., 1998). An entire chromosome has now been assembled in overlapping clones that are only separated by the centromeric region (Kotani et al., 1997).

The PGI will be able to utilize knowledge and technologies developed for these other genome projects to provide a cost-effective program. Moreover, the PGI can benefit from lessons learned by other genome projects and can make use of recent advances in plant genetics to create a uniquely valuable resource. Furthermore, the ongoing AGI can be viewed as the initial phase of a PGI, serving as an organizational guide for a more comprehensive program that includes as its main focus the genomes of important crop species.

Plant Genomes Are Directly Comparable

A major recent discovery in plant genetics is that the complement and chromosomal arrangement of genes in different plant genomes is very similar. This is particularly true for the grasses, such as maize, sorghum, wheat, and rice (Ahn and Tanksley, 1993; Bennetzen and Freeling, 1993; Moore et al., 1995). Thus, maize and rice differ more in the versions of genes (i.e., alleles) that they carry than they do in the actual types of genes. This means that the isolation and study of genes from one grass can provide information and sequences that can be used in the improvement of all other grass species. Moreover, understanding the genetic basis for the similarities and differences among crop species can help us to identify the changes that evolution and/or humans have selected to generate today's crops and, thereby, how we may enhance this selection process to create truly superior varieties.

Nevertheless, this potential synergy does not extend indefinitely, and it is clear that a full understanding of the Arabidopsis genome will not provide sufficient information for a complete understanding of the biology of monocot crops such as maize and wheat. Moreover, a PGI that pursues only one or two plant species would needlessly waste one of the great strengths of plant science—a wealth of comparable data on a broad array of plant species. On the other hand, dispersing genomics research funds across a very large number of plant species would not yield a timely or valuable resource for anyone. Therefore, most plant scientists believe that a successful PGI will require intensive investigations into a small number of key species. However, it is essential that the results and tools developed with these key species are immediately useful to all major crops. The frequent colinearity and similar gene content of plants provide the potential for making such connections.

An Expanded PGI Should Focus First on Maize

The consensus among participants at a National Academy of Sciences Colloquium

LETTER TO THE EDITOR

entitled "Protecting Our Food Supply: The Value of Plant Genome Initiatives," which was organized by Ron Phillips and Michael Freeling and held in Irvine, California, on June 2–5, 1997, was that maize should be the primary target of a PGI. Maize is the most valuable crop in the United States and is the source of much of the feed used in the poultry and livestock industries. Equally important, maize provides an excellent genetic system for plant genome analysis, and there is a large community of productive scientists in the public, private, and Government sectors devoted to its analysis and improvement.

It is also clear that understanding the genetic composition of maize and the functions of its genes will greatly accelerate traditional approaches to maize improvement by breeding. Such advances will also provide a whole new level of potential biotechnological enhancements. For example, comparative genetic analyses among maize and the other cereals will provide new knowledge as well as genes from other species that can be used to improve maize. Genome synteny will also facilitate the reciprocal transfer of knowledge and genes from maize to other important crops, such as wheat, barley, sorghum, oats, forage species, rye, sugarcane, and rice.

Although maize should be the initial focus of an expanded PGI, that does not mean that all experiments would or should be performed with maize. Quite to the contrary. In many cases, studies of other species (particularly, but not exclusively, other grasses) will provide a more rapid route to understanding the maize genome than would experiments conducted directly on maize. Connections can be established between maize and other grasses by comparing genetic maps, physical maps, and expressed sequence tag (EST) sequence and mapping data.

In addition, some dicot species and gymnosperms are important crops as well as resources for basic plant science. Thus, connections between the AGI and genomic analyses in other dicots need to be expanded. Furthermore, crop performance is greatly limited by microbes in all agricultural environments, and so the low cost of genomics projects that focus on a small number of agriculturally important microbes would justify their inclusion in an expanded and well-funded PGI. As discussed below, the degree to which species beyond maize (and its surrogates) should be investigated in depth will depend largely on the level at which the PGI is funded.

Coordination among Plant Scientists Is Needed

Because it will differ in some ways from traditional basic research, the many facets of a PGI must be coordinated to guarantee that all necessary parts are available in a timely and appropriate manner. A steering committee composed of plant genome researchers, funding agency representatives, and end users of plant genomics data should be established to oversee this coordination. On a more immediate level, laboratories involved in genomics studies repeatedly emphasize that they are swamped by the great mass of information they are generating. Thus, the full scientific community will be needed to make the best use of genomics data.

All Data Should Be Publicly Available

All of the academic researchers whom we polled on this subject agree that the materials and information from the PGI must be made available to public databases and distribution centers as soon as it is generated. This would include (1) the immediate release of clones and related data, (2) a complete restriction of privileged access to these data, and (3) no blocks to patents.

Meeting these goals and facilitating coordination among researchers working at all levels within the PGI will require investments, as part of the PGI, in bioinformatics and in a stock center or stock management projects. The objective is to ensure that the efficient storage, distribution, and use of materials and knowledge generated during the PGI can be guaranteed at the outset. This goal may be accomplished by enhancing existing centers, creating centers de novo, and/or developing service companies. Continuing support of these aspects of the PGI will be needed for seed curation, stock management, and the development of enhanced informatics methodologies.

Funding Should Come from Multiple Agencies

Funding of PGI-related projects should be administered through the cooperation of multiple agencies, such as the United States Department of Agriculture, the NSF, and the United States Department of Energy. Projects should be awarded through a competitive grants program following peer review by a single PGI panel established by the participating agencies. Furthermore, scientists in the United States feel that the United States Congress should continue to separate fiscal support for the PGI from existing plant science funding. This is because current support for the plant sciences is at its lowest level in years, a situation so serious that many domestic graduate students do not consider this to be an attractive field in which to work. Moreover, if Congress mandates reallocations within existing plant science budgets, it will strangle postgenomic capabilities, for which the PGI will provide the enabling technologies.

On the other hand, if new funding is made available, scientists in the United States will be able to establish links and partnerships between publicly funded

LETTER TO THE EDITOR

PGI participants and the industrial and international communities, where much of the interest and many of the resources for plant genomics reside. This would optimize (i.e., leverage) the productivity of a PGI and guarantee that all participants share equally in the information generated.

All of the scientists we contacted feel that the exclusive pursuit of plant genomics in industry is not acceptable. This is because information generated by industrial concerns would not be fully available to the largest cohort of scientists, those in the public sector. The impressive recent strides in maize genomics in various companies should serve as another indication of the importance of this kind of research and should not dissuade us from providing these same resources to the broadest possible community of biologists.

An important mission is to achieve the scientific objectives of the PGI at the lowest possible cost. It may be appropriate to use private contracts for some activities and to form public–private and international coalitions to facilitate exchange of information, learning, and resource development. Given the broad nature of the initiative, it is anticipated that some activities would be most efficiently accomplished at concentrated centers, whereas other objectives would be best pursued at a local level by smaller groups or individuals.

EST and Physical Maps Are High Priorities

Both the nature and order of the exact steps that should be pursued in a PGI will depend somewhat on the size of the program and its funding. Nevertheless, one necessary emphasis should be on the identification, sequencing, and mapping of ESTs which, although proving very valuable in other genome projects, are notably deficient in plants. The biggest EST effort should focus on maize, although the size of this component would depend on whether some data can be acquired (without strings attached) from existing private databases. Other grasses (i.e., barley, oats, rice, sorghum, and wheat) should also be subjected to comprehensive EST analyses.

One objective of these experiments is to provide allelic variation for improvement of maize and other crops. However, ESTs will also help establish informational contacts to the physiological, developmental, and genetic data on these other crops. Expanded or new EST projects on a set of important and model species other than grasses (e.g., soybean, tomato, Arabidopsis, a conifer) are also justified in a first PGI both because of the low cost of these activities and because these species may contain valuable genes that are not represented in the target monocot species.

The physical mapping of several grass species (initially rice, sorghum, and maize) should also be initiated. These maps will serve both as the templates for future genomic sequencing and to provide an efficient framework in which to place ESTs and other genes. Cross-species comparisons of these physical maps and more detailed recombinational mapping across a wide array of grass species will be necessary to determine the degree of colinearity and common gene content between different grass genomes and to determine the frequency and nature of exceptions. As with EST projects, low-cost and highly valuable physical maps should be generated for a few important dicots with small genomes, such as tomato.

Whole-Genome Sequencing Should Be Pursued in Both a Dicot and a Monocot

Although Arabidopsis is not a crop, the many advantages of this organism as a model plant species indicate that its genome can and should be fully sequenced in the next few years. The sequence of the Arabidopsis genome will represent the most cost-effective genomic sequencing project for the purpose of gene discovery and will provide a valuable resource for comparison to the genomic sequences of other plants. Therefore, cost-effective enhancement of the current Arabidopsis sequencing capacity should be explored. However, the scientists we talked to (with the exception of several Arabidopsis researchers) feel that it would not be appropriate to accelerate Arabidopsis genomics at the expense of delays in other plant genome programs.

The scientists we consulted also agree that obtaining the complete sequence of a small cereal genome, such as that of rice, would be a tremendously valuable accomplishment and a landmark goal for the PGI. Other nations are planning to begin sequencing the rice genome, and the United States could gain access to this information by producing some portion of the rice genome sequence data. With the identification of homologies between rice ESTs and those of other grasses and the comparison of their map positions to the detailed physical and genetic maps of other important cereals, such as maize and wheat, a completed rice genome sequence would accomplish one of the major goals of the PGI—determining the DNA sequence of all of the genes in an important crop plant. It would also provide the first opportunity to compare comprehensively the genome of a dicot species (i.e., Arabidopsis) with that of a monocot species (i.e., rice).

The value of whole genome sequencing, particularly when it is performed in concert with EST projects, was noted by a large majority of our discussants. However, this topic was also the source of animated discussions regarding the priority that should be placed on this activity, the suitability of current and

LETTER TO THE EDITOR

soon-to-be-available technologies, and the identity of the target species (i.e., rice, sorghum, maize, or a combination thereof). Moreover, because the value of sequencing a grass genome may depend on acquiring the entire sequence, many scientists are worried that the cost could compromise other valuable experiments. These unknowns led some researchers to believe that it is best to approach genomic sequencing with a staged/phased strategy, including preliminary feasibility tests in several possible target species. Others feel that the United States will become a weaker partner in the international research community if other countries press ahead with rice genome sequencing projects without a major contribution from the United States.

Therefore, although all agree that genomic sequencing would be very valuable, the degree and nature of the commitment to this approach will be best determined by consultation among participating nations, agencies, and investigators. For example, the Rockefeller Foundation supports a rice genome working group that was formed at the Fifth International Congress of Plant Molecular Biology, which was held in Singapore last September. This group has outlined a position paper for an international collaboration to sequence the rice genome that has been posted for input from the scientific community at http://www.staff.or.jp.

A Portion of the PGI Budget Should Be Devoted to Ethics and Education

A portion of the funds committed to the PGI should be used to explore ethical considerations of the direct and indirect effects of the initiative. It will also be important to conduct workshops on the skills that are needed to utilize effectively genomics information and the materials generated during the PGI.

Determining Gene Function Is Integral to the PGI

Another central goal of the PGI should be to determine the function of each of the genes in maize and related grasses. This major challenge will require both the development of new genetic tools and the participation of many laboratories with expertise in different areas of plant biology. Public collections of knockout mutants and simple screening approaches will need to be established to determine the effect of mutations in each of these genes. Known mutations and other natural genetic variation in both wild and domesticated species should be mapped onto comparative maps with anchor probes. We must develop and provide technologies for massively parallel (i.e., rapid and inexpensive) analysis of the expression patterns of all genes in a crop plant.

Seven Years, $400 million

The project we have outlined above, which is based on the consensus opinion of scientific experts in agriculture, genetics, and genomics, could be accomplished in 7 years at a cost of $400 million. At a total cost of $600 million for the same 7 years, a PGI could also sequence the gene-containing regions of maize and at least one other important grass. For $800 million, this approach could be extended to the detailed genomic characterization of at least one dicot crop and perhaps a few agriculturally important microbes.

This plan for a PGI is targeted on maize but would provide complementary information and materials for the study and improvement of other crops. Moreover, the technologies and scientific approach refined in this PGI would provide the foundation for future agricultural genomic studies that will dominate biological research in the approaching 21st century. Scientists working on dicot crops (e.g., soybean, potato, and tomato), the conifers, and livestock would all be able to use the knowledge and technologies developed in an expanded PGI.

The expected value of this maize-focused PGI is difficult to overestimate. An essentially complete set of grass genes, pertinent to the study and improvement of maize, wheat, barley, rice, oats, sorghum, and other agronomically important grasses, will be discovered and placed at the free disposal of public and private scientists. There will be great advances in our knowledge of the functions of these genes. Moreover, the PGI will develop the tools and scientific capacity in the public and private sectors to fully determine the relationships between gene function and the biology of plants.

With this knowledge in hand, researchers can use enhanced traditional approaches and novel biotechnologies to design crops that are capable of more reliable and efficient yield and that produce novel products for home and commercial use. These crops will have lower requirements for expensive and unsustainable inputs (e.g., herbicides, pesticides, and fertilizers) and will exert fewer negative impacts on the environment. Thus, the PGI would provide the foundation for a second green revolution that will be necessary to provide food for the expected increases in and improved lifestyle expectations of the world's population in the 21st century. Any nation wishing to remain competitive in agriculture will need to commit to a PGI, and the United States cannot let this opportunity pass while other nations embrace plant genomics as the great new means for crop improvement.

Perhaps even more exciting will be the unpredictable outcomes of the PGI. The technology of genomics is so powerful and plants are so underinvestigated from this perspective that we can guarantee a wealth of unique observations unmatched in any previous period

LETTER TO THE EDITOR

of plant biology research. Indeed, one cannot even guess how many novel products, industries, and public benefits will be derived from the discoveries of a PGI. Past history in discovery research is our only guide, and this history predicts an exceptional potential. It is for these reasons that the large group of United States and international scientists we have interviewed agree that the United States needs a comprehensive and well-funded PGI.

Jeffrey L. Bennetzen
Department of Biological Sciences
Purdue University
West Lafayette, Indiana 47906

Elizabeth A. Kellogg
Department of Organismic and Evolutionary Biology
Harvard University
Cambridge, Massachusetts 02138

Michael Lee
Department of Agronomy
Iowa State University
Ames, Iowa 50011

Joachim Messing
Waksman Institute
Rutgers
The State University of New Jersey
New Jersey 08855-0758
Fax: 732-445-0072
E-mail: messing@mbcl.rutgers.edu

REFERENCES

Ahn, S., and Tanksley, S.D. (1993). Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. USA **90,** 7980–7984.

Bennetzen, J.L., and Freeling, M. (1993). Grasses as a single genetic system: Genome composition, colinearity and compatibility. Trends Genet. **9,** 259–261.

Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., et al. (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana.* Nature **391,** 485–488.

Kotani, H., Sato, S., Fukami, M., Hosouchi, T., Nakazaki, N., Okumura, S., Wada, T., Liu, Y.-G., Shibata, D., and Tabata, S. (1997). A fine physical map of *Arabidopsis thaliana* chromosome 5: Construction of a sequence-ready contig map. DNA Res. **4,** 371–378.

Moore, G., Devos, K.M., Wang, Z., and Gale, M.D. (1995). Grasses, line up and form a circle. Curr. Biol. **5,** 737–739.