

Editor's Choice Series on Sharing Data and Materials

The Preservation of Plant Genetic Resources. Experiences with *Arabidopsis*¹

David Meinke and Randy Scholl*

Community Coordinator of Information on *Arabidopsis* Mutants and Genetic Maps, Department of Botany, Oklahoma State University, Stillwater, Oklahoma 74078 (D.M.), and *Arabidopsis* Biological Resource Center, Department of Plant Biology, Ohio State University, Columbus, Ohio 43210 (R.S.)

The preservation of genetic resources used to support multinational research programs in plant biology requires extensive sharing and curation of information and materials. Each sector of the community has an important role to play, from individual investigators and volunteer coordinators to curators of central databases and public stock centers. When empowered with a sense of responsibility and goodwill, stable sources of funding, and a standardized system of genetic nomenclature, plant biologists can build a lasting collection of genetic resources that aid their research programs. In this article, we reflect upon our personal experience with genetic resources and nomenclature of *Arabidopsis*, consider how this example might benefit those working with other plant species, and discuss the opportunities and challenges that lie ahead.

HISTORICAL PERSPECTIVE

The value of preserving natural and induced variation in *Arabidopsis* was recognized by the earliest investigators (discussed in Somerville and Koornneef, 2002). Initial efforts to curate genetic collections were undertaken by individual researchers as an avocation not funded by external sources. This work began with Friedrich Laibach, Gerhard Roebbelen, and Albert Kranz, whose *Arabidopsis* Information Service collection formed the cornerstone of genetic variation used by the explosion of *Arabidopsis* investigators in the 1980s. Information about the *Arabidopsis* Information Service collection was published and samples were distributed worldwide for modest fees. George Redei isolated a number of mutants during this period, produced the first partial linkage maps, and freely shared his collection. Maarten Koornneef generated additional mutants and completed development of a comprehensive genetic map. Demand for these seed resources increased just as Drs. Kranz and Redei approached retirement. Requests for molecular resources such as libraries and

clones used in physical mapping also increased during this period. By 1990, the *Arabidopsis* community recognized the need for long-term curation and efficient distribution of stock collections. The Multinational *Arabidopsis* Steering Committee (MASC) articulated a vision that included two stock centers, one in the U.S. and one in Europe. These centers were designed to represent convenient distribution points for the two largest concentrations of *Arabidopsis* scientists and to hold mirrored collections to ensure preservation of stocks. As a result of this vision, the Nottingham *Arabidopsis* Stock Centre (NASC, Nottingham University, Nottingham, UK) was established in 1990 with British public funding, and the *Arabidopsis* Biological Resource Center (ABRC, Ohio State University, Columbus) was founded in 1991 with support from the National Science Foundation (NSF).

The plan outlined by the MASC was unique in another aspect: that resources of a molecular nature should be curated by the stock centers. Important molecular resources at the time included restriction fragment-length polymorphism mapping clones from the laboratories of Elliot Meyerowitz and Howard Goodman. The dual role of seed and clone curation was adopted by the ABRC in the U.S., whereas these functions were delegated to two distinct centers in Europe: NASC for seeds and the European DNA Resource Center in Germany for clones. NASC was founded by Mary Anderson and Bernard Mulligan and is presently directed by Sean May. ABRC's DNA facility was directed first by Keith Davis and later by Doreen Ware. Jeffrey Dangl headed the European DNA Center. After a few years of operation, public distribution of clones from the European center was discontinued. ABRC has subsequently provided molecular resources for the entire community. The usage of ABRC and NASC resources has increased continuously since their inception. The number of seed and DNA samples sent annually by the two centers combined now approaches 100,000, a rate that substantially exceeds anything imagined in the beginning. This magnitude of curation and distribution activities needs to be considered when planning for similar resources to support research with other organisms. Because the number and diversity of resource types to be distributed often exceeds expectations, individual

¹ The *Arabidopsis* Biological Resource Center is supported by the National Science Foundation.

* Corresponding author; e-mail scholl.1@osu.edu; fax 614-292-0603.

www.plantphysiol.org/cgi/doi/10.1104/pp.103.024877.

stock centers must remain flexible in their plans to allow for expansion of facilities and personnel.

Another requirement for proper maintenance of genetic resources in model organisms involves the use of a standard system of nomenclature. Important advances in this regard were made at the Third International Arabidopsis Conference at Michigan State University, where initial guidelines for naming genes with mutant phenotypes were established. Additional details were published ten years later (Meinke and Koornneef, 1997) in a document that was designed to establish community standards for Arabidopsis genetics. The need for coordination of Arabidopsis nomenclature became apparent in the 1980s, when large numbers of mutants started to be isolated and assigned gene symbols in one laboratory without knowledge of related activities in other laboratories. This resulted in the occasional use of a single gene symbol to describe mutants with different phenotypes and gene identities. This was the original incentive for the establishment of a centralized list of mutant gene symbols that was made available to the community via the Internet, initially through the Meinke laboratory web page and later through The Arabidopsis Information Resource (TAIR) at arabidopsis.org. Responsibility for curation of genetic maps in Arabidopsis has also progressed over the years from individual research laboratories to a combination of community organizers and database curators (Meinke et al., 2003). The result has been the establishment of an impressive collection of mutants with standard identifiers indicative of disruptions in all aspects of plant growth and development. Rather than implementing a rigid system of controlled vocabulary and establishing a formal governing body for approval of acceptable nomenclature, the Arabidopsis community has emphasized instead the widespread advantages of maintaining and adhering to standardized guidelines. This flexible but persistent approach has, for the most part, been quite successful.

THE CENTRAL ROLE OF STOCK CENTERS

The principal tasks envisioned for the Arabidopsis stock centers, to collect, preserve, and distribute genetic resources, have proven to be substantial. In addition, the number and diversity of materials involved have increased far beyond anything initially imagined. Many of these resources have been handled by the stock centers, whereas some have been shared in other ways. Stock centers offer several advantages over alternative providers: they cannot be considered a competitor, they have no interest in or claims to intellectual property, and they maintain large numbers of stocks according to common standards. Although stock centers should not be responsible for managing or maintaining the intellectual property of others, they cannot impinge on the exist-

ing rights of individuals. The evolution of genetic resources in Arabidopsis has resulted in many surprises over the years, reshaping resource collections, challenging oversight policies, and affecting how research is conducted. Many community resources have become available in recent years, including public and private Arabidopsis materials and stocks of other plants and model genetic organisms. The Arabidopsis Information Resource database maintains current lists of these resources linked directly to its home page (<http://arabidopsis.org>).

The collaboration between ABRC and NASC has functioned well. There is close communication on all matters, including division of labor in growing of seed stocks and reciprocal participation in advisory board meetings. The ABRC Advisory Committee includes the Director of NASC, in addition to scientists chosen mainly from the North American Arabidopsis Steering Committee. Members of the Advisory Committee are selected for their expertise in genetics, genomics, and molecular biology. Linkages to the North American Arabidopsis Steering Committee, the MASC, funding agencies, and laboratories in different countries help to ensure that stock centers respond to the current and future needs of the science community. In retrospect, the necessity to reconcile disparate, nationality-based views on policy issues such as intellectual property claims on stocks has resulted in more equitable policies throughout the community. Given the implementation of stringent new import restrictions by many countries, the availability of dual distribution points for seed stocks may prove even more useful in the future.

ABRC and NASC decided from their inception to maximize the use of electronic media, specifically the Internet, for dissemination of stock information and ordering. Paper catalogs and ordering by mail and fax were also offered, but researchers made full use of stock information in the databases. Significantly, this took place three years before the advent of the World Wide Web. ABRC maintained a relational database in collaboration with Sakti Pramanik at Michigan State University, and NASC processed orders via e-mail and maintained electronic catalog versions that could be downloaded via ftp. Because all of the stock data were electronic, images of stocks were incorporated into the databases so that researchers could observe phenotypes of seed accessions and selected molecular data. This feature has been well received. Although it is not possible to display all critical stock data, photographs of actual plants are useful to those first working with a particular mutant or ecotype because they provide valuable information on subtle and transitory phenotypes and reassurance that the correct stock has been obtained.

Two early donations influenced the future direction of the stock centers. The first was in 1992 when Kenneth Feldmann donated more than 4,900 T-DNA insertion lines to ABRC. Although this would be

considered a modest donation by current standards, it appeared mammoth at the time. Refusal of such a valuable resource would have relegated the stock centers to a peripheral role, particularly in the management of genomic resources. Hence, these lines were quickly prepared for public release, and 37 sets were shipped in the first two days of distribution—a forced introduction to the challenges of mass distribution. This donation also established an important precedent for distribution of large populations with no legal strings attached (i.e. without a requirement for a material transfer agreement [MTA]). Although this example has not been followed in every case, similar donations have been made on a regular basis, and each has reinforced the spirit of sharing that is prevalent in the community.

A second important donation occurred a year later with clones from the expressed sequence tag project at Michigan State University headed by Christopher Somerville and Thomas Newman. This collection, when augmented by similar materials from Institut National de la Recherche Agronomique (Versailles, France) and the Christoph Benning laboratory (Michigan State University, East Lansing), ultimately grew to 40,000 clones representing over 7,000 genes. An unusual condition accompanied the initial donation: that the names of individuals receiving clones from the stock center should be public information. For uniformity of policy, the same approach was soon applied to all ABRC stocks. Although the best policy for Arabidopsis may not be right for other groups, our community has largely supported this approach. It cannot be denied that openness has been promoted and that duplication of effort among laboratories has been minimized.

The scope of molecular collections expanded in the late 1990s when the bacterial artificial chromosome clones used by the Arabidopsis Genome Initiative sequencing consortium were donated. This allowed users to analyze any localized genomic region, including those involved in chromosome walks. Arabidopsis expressed sequence tags and bacterial artificial chromosome clones have all been provided without legal restrictions. The first such clones (P1; PAC) donated in 1994 initially came with a limited MTA, but this was later discontinued with the permission of the donors. Valuable resources for functional genomics have recently become available with support from the NSF 2010 program and related initiatives worldwide. Two important donations were made in 2001. At that time, the laboratory of Joseph Ecker at the Salk Institute began to sequence T-DNA flanking regions from a population of 144,000 independent transformants. The incorporation of this population into stock center collections means that loss-of-function alleles for more than 75% of the genes are now available with no restrictions. Likewise, the 11,000 full-length cDNA clones being donated without restrictions by the Salk, Stanford, Plant

Gene Expression Center Consortium provides a valuable resource for the entire community. Many additional donations are anticipated from ongoing programs in the years ahead. Hence, public stock centers will soon become the focal point for distribution of major resources to be used in the post-genomics era. The central role of stock centers in enabling the exchange of materials has been facilitated by policies of the funding agencies, especially the NSF, which requires investigators to share the resources they produce, to budget funds to develop the necessary stocks, and to outline how this will be achieved (see article by Jane Silverthorne in this series). This model should be applicable to the sharing of additional genomic resources in the future.

The funding of stock centers has traditionally been overseen by government agencies. Grants to the host institution and public recognition associated with the presence of a stock center provide enough incentive for the institution to allocate the required space and personnel. This basic model has worked well for Arabidopsis and several other species. When the tasks confronting a stock center exceed the resources available, cost recovery through user fees may be instituted. Such a policy was adopted by ABRC and NASC in 1997. The idea is to recover some costs of operation through fees so that greater service can be provided. Such fees can also discourage frivolous requests, although these have never been a problem with the Arabidopsis community. Initially, about 25% of ABRC's budget was covered by fees. Charges have increased somewhat over the years, so that at present, approximately 40% of the ABRC budget, on a direct cost basis, is derived from user fees. Although larger fees could be charged, the research budgets of plant biologists are often modest, and substantial increases could well prove counterproductive. The present combination of user fees and government support of Arabidopsis stock centers represents a reasonable balance between community services provided and financial support required.

One approach to increasing the visibility and usefulness of a stock center is through the expansion of services. Although the NSF requires that stock center activities be focused strictly on stocks, NASC has successfully added a number of services, including curation of the recombinant inbred map, distribution of DNA microarrays, and an expanded presence in genome informatics. This has worked well for the Arabidopsis community in that ABRC specializes in offering a clone collection and NASC offers these additional services so that there is beneficial duplication of effort in selected areas in addition to appropriate specialization.

The presence of a stock center at a research university allows the center to draw upon technical expertise available at the institution. At the same time, a stock center occupies space that could be devoted to research scientists. The positive publicity and over-

head income accruing from the facility should represent enough inducement for the university to allocate the space required. The lack of motivation of faculty to supervise stock centers and curate the collections represents another potential impediment to the establishment of a stock center. In reality, whereas a stock center is strictly a service operation, the curation of stocks represents an interesting intellectual pursuit in its own right, and management of the operations represents an evolving challenge. Anyone embarking on this pursuit will be surprised if they assume that intellectual creativity is not required.

THE IMPORTANCE OF INDIVIDUAL CONTRIBUTORS AND COMMUNITY COORDINATORS

The role of stock centers in maintaining genetic resources is perhaps best described as necessary, but not sufficient, to meet the needs of the research community. Individual investigators and community coordinators must also play an important role in organizing and analyzing biological information and materials. The responsibilities of principal investigators extend from the use of appropriate genetic nomenclature to careful descriptions of mutant phenotypes and prompt donation of seeds for mutants, ecotypes, and populations. Individual laboratories may also work together to contribute important community services, as occurred several years ago when Pamela Green, Shauna Somerville, Michael Sussman, Richard Amasino, and colleagues launched the Arabidopsis Functional Genomics Consortium to benefit advances in large-scale expression analysis and screening for knockout mutants in Arabidopsis. Principal investigators can also work behind the scenes to encourage public release of extensive resources from the private sector. This approach was used to facilitate the release of valuable seed stocks and gene identities from Syngenta that provided the foundation for a 2010 project on genes with essential functions during seed development (Tzafrir et al., 2003).

Community coordinators can complement the efforts of database and stock center curators by encouraging donations, keeping track of potential duplications in genetic resources, and facilitating the release of updated information on maps and mutant collections. In the case of Arabidopsis, the Meinke laboratory has served for some time as a coordinator for information on nomenclature and maps of genes with mutant phenotypes (Meinke et al., 2003). Other individuals have contributed to the organization of information on molecular markers, physical maps, specialized techniques, and complex gene families. However, with continued growth of the community, these responsibilities can soon exceed the capacities of a single individual. The expertise of volunteer coordinators from the community must then be combined with the resources of stock centers and database curators to support continued advances.

ISSUES TO BE ADDRESSED

Despite impressive advances on a number of fronts, several critical issues remain to be addressed to realize the potential of Arabidopsis stock centers to support research. First, the curated collection of mutants, although extensive, has always been smaller than the number of mutants published. This represents a significant deficiency. ABRC and NASC have conducted a number of activities to close the gap: contacting potential donors by direct mail and e-mail, broadcasting campaigns for submissions through the Arabidopsis newsgroup, and maintaining donation pages on relevant Web sites. These measures have always produced positive results, so that about one-half of the published loci have typically been included in the collections. The addition of knockout lines for the majority of loci from the Salk (Ecker) collection will somewhat ameliorate this need, as confirmed homozygous lines are received. Nevertheless, one of the basic tenets of science is the ability to repeat experiments, and access to the exact genetic material used is vital. Hence, all published germplasm should and will be pursued by the stock centers. In this regard, several important plant journals have adopted a policy that all genetic materials resulting from genomics projects and employed in refereed publications should be donated. This represents a welcome, landmark advance in policy.

The second area of concern for ABRC and NASC involves questions of intellectual property. The original policies of both centers allowed no MTAs to be associated with stocks received. This changed in 1994 when the ABRC Advisory Committee concluded that the value of a particular stock outweighed the negatives associated with an MTA. Although this particular MTA was later discontinued, major donations have subsequently been accepted with MTAs containing a variety of clauses. Additional resources of considerable value to the community could have been acquired if MTAs with reach-through rights had been accepted. ABRC has chosen not to accept such MTAs, and has declined MTAs with other objectionable clauses. Since the year 2000, NASC and ABRC have adopted unified policies against new MTAs and have agreed that existing MTAs administered by the centers will ultimately be phased out. Such a policy is welcomed by the great majority of plant scientists for reasons being described in other articles in this series. Donations of large collections of T-DNA lines and full-length cDNAs with no strings attached played an important role in making such advances possible.

Additional issues remain to be addressed at the level of the principal investigator. Two major concerns involve prompt donation of seed stocks and appropriate recognition of existing nomenclature. Based on our experience, we believe that most people want to contribute information and seed stocks for their mutants, but have trouble prioritizing and finding the time needed to complete the task. Therefore,

a continuing challenge for stock centers and database curators is to pursue novel strategies to facilitate the periodic updating of genetic information and biological resources. One approach may involve making better use of workshops and registration materials associated with the annual *Arabidopsis* meeting. Another might be to automate linkages with funding agencies and editorial offices of major journals. Inherent in these issues is the challenge of balancing the concerns of individuals with limited resources who have chosen to focus on a particular mutant or gene of interest, and laboratories with considerable resources capable of making rapid progress with these materials once they become public. Although sharing of resources must remain a top priority for everyone involved, the standards and timetable for release of materials may need to be somewhat flexible to encourage maximal participation. A second concern involves the use of novel names and symbols to describe new alleles of established mutants (Raikhel, 2001). In the past, some of these duplications of nomenclature were unavoidable when seed stocks were not available to perform allelism tests

and gene identities were unknown. Today, these mistakes and oversights are more difficult to accept given the widespread availability of searchable databases and extensive collections of seed stocks in the public domain. By keeping such examples to a minimum, members of the community will help to maintain continuity and to minimize confusion in the preservation of genetic resources for future generations of plant biologists.

Received April 4, 2003; returned for revision May 5, 2003; accepted May 5, 2003.

LITERATURE CITED

- Meinke D, Koornneef M** (1997) Community standards for *Arabidopsis* genetics. *Plant J* **12**: 247–253
- Meinke DW, Meinke LK, Showalter TC, Schissel AM, Mueller LA, Tzafirir I** (2003) A sequence-based map of *Arabidopsis* genes with mutant phenotypes. *Plant Physiol* **131**: 409–418
- Raikhel N** (2001) Renaming genes and duplication of gene names in the literature. *Plant Physiol* **127**: 703
- Somerville C, Koornneef M** (2002) A fortunate choice: the history of *Arabidopsis* as a model plant. *Nat Rev Genet* **3**: 883–889
- Tzafirir I, Dickerman A, Brazhnik O, Nguyen Q, McElver J, Frye C, Patton D, Meinke D** (2003) The *Arabidopsis* SeedGenes project. *Nucleic Acids Res* **31**: 90–93.