From *In Situ* to *Ex Situ* and Back: The Importance of Characterizing Germplasm Collections

James R. McFerson
Washington Tree Fruit Research Commission, 1719 Springwater Ave., Wenatchee, WA 98801

Biodiversity is widely perceived as desirable, although controversy still revolves around how it should be measured, conserved, and valued. While sizable collections of genetic resources have existed for centuries as zoos, botanical and estate gardens, hunting and nature preserves, etc., the development of extensive, formal, organized *ex situ* genetic resource collections of crop plants (gene banks) is recent. Such collections have never been more important, as agricultural development is fundamentally dependent on germplasm improvement, and the level of genetic diversity in crops has inevitably been reduced during domestication (Tanksley and McCouch, 1997).

As clearly described by Harlan (1992), traditional agricultural systems were based on variable, integrated, adapted populations called landraces. In these systems, genetic diversity was narrowed within landraces via artificial and natural selection, yet overall diversity was retained within geographic areas, first termed centers of origin by Vavilov (1926). With increasing agricultural productivity, expanded human activities and industrial development have severely eroded genetic resources. This process has been greatly accelerated over the past century with the continued development of scientific plant breeding and the globalization of agricultural technologies. While this genetic erosion has continued, largely unrecognized not only by the general public but by many agricultural scientists as well, certain farsighted and significant efforts have been undertaken worldwide to collect and preserve genetic resources of crop plants.

Unfortunately, most early collectors of germplasm were narrowly exploitative or haphazardly opportunistic. Further, they often ignored the rights of countries or peoples where the collections were made (Solleiro, 1998). However, acquisition strategies and techniques have evolved considerably, with greater attention to conceptual, technical, and ethical issues (Frankel, 1984; Frankel and Brown, 1984; Frankel and Soulé, 1981; Frankel et al., 1995; Guarino et al., 1995).

As awareness of the importance of collections has grown, so have the numbers of accessions held in germplasm repositories, even though funding, trained scientists, and adequate storage facilities have always been limiting. Currently, nearly three million accessions are estimated to be held worldwide in *ex situ* collections, with nearly half of these accessions being major grain or food legume crops (Plucknett et al., 1987). Considerable genetic resources also exist in crop improvement programs and as commercially-important cultivars (Duvick, 1984). Indeed, collections have grown large enough for some to argue that the value of continued germplasm acquisition is dubious and that resources should instead be allocated to evaluation of collections (Marshall, 1989). In a related vein, Kresovich and McFerson (1992) contended that the quality of a collection, rather than its sheer quantity, should be the curator's goal. Though easily stated, this approach is more difficult to articulate than is numbers-counting. It requires development and understanding of appropriate tools and interpretation of results.

CRITICAL CONSIDERATIONS IN GERMPLASM COLLECTIONS

The first step in evaluating an *ex situ* germplasm collection involves clarifying its nature and purpose. Fundamentally, germplasm collections should contain materials that offer both short-term benefits and long-term insurance. A collection of a selected taxon should be well characterized and represent as much variation as possible within a reasonable number of accessions. It should include genotypes or populations at all levels of development, including wild and weedy relatives, landraces, cultivars (obsolete and current), and genetic stocks. This approach, built on the gene pool concept first articulated by Harlan and deWet (1971), is easily applied to the needs of plant breeders, who have constituted the primary users of *ex situ* collections. Obviously, crop improvement programs will preferentially utilize germplasm within Harlan and deWet's Gene Pool 1, where sexual recombination is easy. Further, breeders will prefer accessions with genotypes that most closely approximate those of current commercially desirable cultivars. However, plant breeders have long been aware that unique and useful alleles are often found in exotic materials (Gene Pools 2 and 3). Although such alleles, particularly those governing quantitative traits, are often completely obscured by unprepossessing phenotypes of exotic, wild, and weedy germplasm, emerging techniques allow their expeditious identification and introgression, as recently highlighted by Tanksley and McCouch (1997). Thus, the maintenance of genes in Gene Pools 2 and 3 is insurance for continued crop improvement and novel biotic and abiotic challenges to existing cultivars.

In addition, curators should strive to extend their customer base beyond plant breeders. While most *ex situ* collections of plant germplasm are crop-oriented, they can offer ideal experimental materials across disciplines in basic and applied sciences. Such research can lead to tangible benefits like identification of characters valuable in production agriculture, as well as intangible ones, such as increased knowledge of a range of biological systems and phenomena.

In all cases, whether the user is exploring issues of plant ecology or gene action and evolution, the quality of a collection is determined by how well it represents its target taxon. Resolving this issue requires genetic characterization, which simultaneously and directly improves collection utility. Ideally, such characterization is equivalent to the comprehensive understanding of relevant genetic parameters for each accession. Since most *ex situ* collections of plant germplasm emphasize a given crop group, most work is conducted at the species or population (accession) level. However, a comprehensive understanding of other organizational levels is also needed, from ecosystems through cellular and molecular levels. The genetic issues of most interest to the curator can be summarized as follows:

identity—the correct cataloguing of accessions;
relatedness—the degree of similarity among accessions in a collection;
structure—the amount of genetic variation present and how it is partitioned among accessions and genotypes;
location—existence of desirable genes in individual accessions and the physical and genetic location of these DNA sequences.

The curator needs accurate, rapid, and cost-effective tools to assess the collection for these parameters. In the past, such tools were limited and consisted predominantly of characters of agricultural interest that sometimes segregated in a Mendelian fashion, but more frequently were polygenic and greatly affected by the test environment. In many instances, even morphometric data sets for a collection were incomplete and/or of questionable validity. Even when complete data sets were assembled, their utility was often hampered by the limited number of markers available, different test environments, and variation associated with genotype x environment interactions.

While such phenotypic information on agricultural traits remains critically important to the curator and the crop-oriented user community, it is still insufficient to resolve many of the curators' questions on the genetic parameters outlined above, even in the relatively well-characterized systems, such as barley, maize, pea, tomato, or wheat. In many plants—fruits, vegetables, medicinals, trees and shrubs—almost nothing is known of genetic structure. Given the declining funding for genetic resources programs, this situation is unlikely to change dramatically. Many collections already cannot fund activities in the "life-support systems" of maintenance and rejuvenation, let alone expensive, long-term evaluations.

Received for publication 10 Feb. 1998. Accepted for publication 4 Mar. 1998. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked *advertisement* solely to indicate this fact.

Fortunately, advances in technology and genetic knowledge over the past decade promise new strategies and tools of considerable potential to curators of collections. Methods using molecular markers, especially those based on DNA polymorphisms, are evolving at a breathtaking rate. Molecular markers are now routine tools in many scientific fields, used to study animal, bacterial, fungal, plant, and viral genomes at population and individual levels. Technical advances in instrumentation, protocols, and data handling capability show no signs of abating and contribute to an ever more favorable cost/unit biological information ratio.

UTILITY OF MOLECULAR MARKERS

As highlighted by Avise (1994), molecular markers are used most intelligently when they address controversial areas or when they are employed to analyze problems in biology that have proven intractable using traditional nonmolecular observation. This applies equally to conservation biology or crop genetic resources—the researcher must ask the right questions and use the right tools. Avise (1994) further highlights that molecular genetic markers are so valuable because: 1) molecular data are genetic; 2) molecular methods open the entire biological world for genetic scrutiny; 3) molecular methods access a nearly unlimited pool of genetic variability; 4) molecular data can distinguish homology from analogy; 5) molecular data provide a common measure for assessing divergence; 6) molecular approaches facilitate mechanistic appraisals of evolution; and, finally, 7) molecular approaches are challenging and exciting. These same arguments hold true for application of molecular genetic markers for more effective conservation of agricultural genetic resources.

The ultimate goal of these genetic analyses will be to determine, analyze, and store DNA sequence information for a wide variety of applications in agriculture and conservation biology. While isozymes, seed proteins, and other molecules have been, and will be, widely used to assess genetic diversity, they suffer several disadvantages compared to molecular DNA markers. Polymorphisms are limited and sometimes nonexistent in many crops, loci occur in a small number of sites in the genome, and sample throughput is low. Where isozyme or protein polymorphisms are known, they are certainly a valuable source of information, as long as the researcher is careful, as Bachman (1994) pointed out, to treat polymorphisms as "tools that provide access to relevant characters, rather than as characters themselves."

Several recent reviews have described in detail a range of molecular markers useful for assessing plant genetic diversity (Bretting and Widrlechner, 1995; Gepts, 1995; Kresovich et al., 1993; Staub et al., 1996). Numerous crops and crop relatives have been characterized using molecular markers, at first using RFLPs (restriction fragment length polymorphisms), and more commonly now with some variant of PCR-based molecular markers: randomly amplified polymorphic DNA (RAPDs), simple sequence repeats (SSRs), and amplified fragment length polymorphisms (AFLPs). These techniques have created fantastic opportunities for fine-scale genetic characterizations of germplasm collections—the only major limitation being cost. Yet, since PCR-based markers are highly polymorphic and simple to process, they generate very large amounts of information for the time and money invested.

Molecular markers can be particularly useful in avoiding unnecessary duplication and reducing redundancies within and among collections. Very often, duplicate accessions exist in active collections, because of faulty passport information, transcriptional errors, mishandling, or lack of quality control. Systematic duplication is standard operating procedure when done to provide secure backup collections; however, unrecognized duplication and redundancy in an active collection is costly, genetically needless, and confusing.

Molecular markers will also play an important role in the development of core subsets that optimally represent the genetic diversity found in the entire collection (Gepts, 1995). Such core collections allow for the most efficient evaluation and minimize production and distribution costs (Brown, 1989 and 1995; Frankel, 1984). In the development of a core collection, not only environmental and agricultural attributes must be considered, but also the underlying patterns of genetic variability. Genetic variability is the backbone of breeding success and supplies experimental material useful in an array of biological research. Scientists utilizing a core collection often desire not only material that carries agriculturally important genes, but also genotypes that represent the maximum amount of genetic diversity. Such a collection is valuable in providing genes of interest, but also in finding useful traits that are currently unrecognized but may play an important role in the future.

For effective germplasm management, curators must have access to molecular marker systems that are inexpensive, automated, high output, user-friendly, and easily integrated into parallel processes. In this way, their needs are identical to all scientists across disciplines and biological targets. As is the case with information technologies, costs for new technologies will decrease even as their power increases by orders of magnitude. As further improvements are made in all molecular based marker systems, germplasm managers have exciting opportunities to more effectively manage their burgeoning collections.

CONCLUSIONS

To be most cost- and space-efficient, germplasm collections must be genetically well-characterized. In this way, germplasm collections can be made as lean as possible with minimal redundancy in genotypes, gene complexes, and genes. The collection also becomes most useful when information is available as to where useful genes and gene complexes are located within crop genomes. The recent emergence of relatively inexpensive and rapid molecular technologies will facilitate this process. The molecular characterization of germplasm collections should be a major priority with all germplasm managers.

Literature Cited

Avise, J.C. 1994. Molecular markers, natural history, and evolution. Chapman and Hall, New York.

Bachman, K. 1994. Molecular markers in plant ecology. New Phytol. 126:403–418.

Bretting, P.K. and M.P. Widrlechner. 1995. Genetic markers and plant genetic resource management. Plant Breeding Rev. 13:11–85.

Brown, A.H.D. 1989. Core collections: A practical approach to genetic resources management. Genome 31:818–824.

Brown, A.H.D. 1995. The core collection at the crossroads, p. 3–19. In: T. Hodgkin, A.H.D. Brown, T.J.L. van Hintum, and E.A. Morales (eds.). Core collections of plant genetic resources. Wiley, New York.

Duvick, D.N. 1984. Genetic diversity in major farm crops on the farm and in reserve. Econ. Bot. 38:161–178.

Frankel, O.H. 1984. Genetic perspectives of germplasm conservation, p. 161–180. In: W.K. Arber, K. Llimensee, W.J. Peacock, and P. Starlinger (eds.). Genetic manipulation: Impact on man and society. Cambridge Univ. Press, Cambridge, UK.

Frankel, O.H. and A.H.D. Brown. 1984. Current plant genetic resources—A critical appraisal, p. 1–11. In: Genetics New Frontiers (Vol. IV). Oxford and IBH Publ. New Delhi, India.

Frankel, O.H. and M.E. Soulé. 1981. Conservation and evolution. Cambridge Univ. Press, Cambridge, UK.

Frankel, O.H., A.H.D. Brown, and J.J. Burdon (eds.). 1995. The conservation of plant biodiversity. Cambridge Univ. Press, UK.

Gepts, P. 1995. Genetic markers and core collections, p. 127–146. In: T. Hodgkin, A.H.D. Brown, T.J.L. van Hintum, and E.A. Morales (eds.). Core collections of plant genetic resources. Wiley, New York.

Guarino, L., V.R. Rao, and R. Reid. 1995. Collecting plant genetic diversity. CAB International, Wallingford, Kent, UK.

Harlan, J.F. 1992. Crops and man. Amer. Soc. Agron., Madison, Wisc.

Harlan, J.F. and J.M.J. de Wet. 1971. Toward a rational classification of cultivated plants. Taxonomy 20:509–517.

Kresovich, S., W.F. Lamboy, A.K. Szewc-McFadden, J.R. McFerson, and P.L. Forsline. 1993. Molecular diagnostics and plant genetic resources conservation. AgBiotech News and Info. 5:555–558.

Kresovich, S. and J.R. McFerson. 1992. Assessment of genetic diversity: Considerations of intra- and inter-specific variation. Field Crops Res. 29:185–204.

Marshall, D.R. 1989. Crop genetic resources: Current and emerging issues, p. 367–388. In: A.D.H. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (eds.). Plant population genetics, breeding and genetic resources. Sinauer, Sunderland, UK.

Plucknett, D.L., N.J.H. Smith, J.T. Williams, and N.M. Anishetty. 1987. Gene Banks and The World's Food. p. 247. Princeton Univ. Press, Princeton, N.J.

Solleiro, J.L. 1998. Regulating access to genetic resources: Recent advances in Latin America. HortScience (in press).

Staub, J.E., F.C. Serguen, and M. Gupta. 1996. Genetic markers, map construction and their application in plant breeding. HortScience 31:729–741.

Tanksley, S.D. and S.R. McCouch. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277:1063–1066.

Vavilov, N.I. 1926. Studies on the origin of cultivated plants. Inst. Appl. Bot. Plant Breeding, Leningrad.