Efficient Sampling from a Collection

Edward J. Ryder
U.S. Agricultural Research Station, Agricultural Research Service, U.S. Department of Agriculture, 1636 E. Alisal Street, Salinas, CA 93905

The similar titles of Namkoong's paper and this one invite comparison, not so much for their similarity as for the differences in sampling rationale. The main difference is in the nature of the populations that are being sampled. The population in the field is usually far away. Further, it is of unknown size and its genetic makeup and complexity can only be guessed.

Geographical distance means several things. It is expensive to get to a distant place and to stay for a long period. Therefore, the opportunity to sample may not occur very frequently. In some instances, it may happen just once; on a second trip, one may find a city or a cropped field where the collection site used to be.

In the "good old days" of plant exploration, a collector might expect to stay in the field for long periods of time, collect everything of interest in several or many different species, and not worry about costs or whether the sample represented the population genetically and statistically. Now, the collector may have just one short trip in which to obtain a sample of maximum genetic usefulness. Thus, Namkoong presents a set of concerns of real interest to modern collectors regarding the structural complexity of a wild population: its allelic distribution, relative allelic desirability, and other characteristics that would influence decisions on the method of sampling and size of samples.

Sampling from existing collections is a much easier task. Existing collections are easily accessible physically. They are either geographically close or at least communicably close; i.e., questions can be asked and samples requested by telephone or letter. Physical accessibility also means that the collections can be sampled at relatively little expense.

Existing collections have a finite size. Agronomic collections, particularly of the cereal grains, are likely to be quite large and pose sampling efficiency questions that are less likely to arise with horticultural collections, which may be relatively small.

Existing collections should, theoretically, last forever. They are in no danger of being destroyed by urbanization and industrialization, or of being lost because of substitution of improved modern cultivars for land races. However, they can be lost through loss of viability, neglect, or other types of human error or administrative indifference. These are serious concerns that require constant attention.

One of the most important virtues of existing collections is the opportunity to increase the information available from them and, thus, their potential usefulness. To place this idea in proper perspective, it would be useful to consider the present state of existing collections.

It is appropriate that the word "diversity" is applicable to nearly all attributes that describe the state of the various germplasm collections in this country. Unfortunately, at this stage in the development of the national germplasm system, only those collections with some official status can be described adequately. There are many collections that are associated with research projects and are virtually unknown to others. For example, there are several lettuce breeding collections that have no official status and are essentially undescribed.

The size of collections varies tremendously, ranging from small, with a few hundred accessions, to the National Small Grains collection, with 109,000 entries of wheat, barley, rice, oats, rye, *Triticale*, and *Aegilops*. The type of material in the collection contributes to the size variation. Some collections contain only plant introduction lines, some only cultivars, and some only genetic mutant stocks. Some contain combinations of two or all three of these classes of material. The size of the collections is, in general, proportional to the economic importance of the representative crop species but is also related to personal or institutional diligence exerted towards collection and preservation of materials.

The holders of collections are themselves a diverse group. The official national curators include geneticists (tomato and barley genetic stocks), breeders (lettuce, carrot, onion), Regional Station directors (sugarbeet, tomato), and curators *per se* (small grains). They include both ARS/USDA and state experiment station people. There are 29 official curators.

The crops curated include small grains, cotton, sugarbeet, peanut, tomato and tomato genetic stocks, barley genetic stocks, pea, potato, lettuce, *Helianthus*, flax, soybean, pearl millet, *Trifolium*, pecan, tobacco, maize, carrot, onion, native grasses, wheat genetic stocks, Western range grasses, bamboo, tropical fruits, sugarcane, clover, *Tripsacum*, and various woody genera. Official curators have not yet been appointed for most species. Many species are held at the National Seed Storage Laboratory in Fort Collins, Colo., mostly in the form of cultivars. Many are at the four Regional Introduction Stations (Pullman, Wash.; Experiment, Ga: Geneva, N.Y.; and Ames, Iowa) and consist principally of wild and domesticated introductions. Most plant breeders have collections of cultivars and breeding lines and possibly some plant introduction lines and genetic stocks as well.

The diversity among those responsible for the germplasm collections is greatly responsible for the variation in the state of the collections. Many collections are in the hands of plant breeders, geneticists, or other research scientists, whose primary job is research. In those situations, the people, time, and money required for maintenance, evaluation, and distribution are taken from the research projects, with the result that one or another of those efforts suffers from lack of support. This disadvantage is partly offset by the ability of the scientist–curator to translate directly descriptor traits and their values into breeding techniques and improvement. These translations are not always obvious to the curator who is not also in research. On the other hand, I think there is a tendency for a breeder–curator to think of a collection as a personal one, for primary use in the local breeding program, and to neglect the responsibility to describe the material properly and to publicize the existence and availability of the collection.

The lettuce and leafy vegetable collection at the U.S. Agricultural Research Station, Salinas, Calif., is a case in point with regard to some of the previous remarks. It is useful to describe that collection in a little detail. It is a collection I inherited from R.C. Thompson in 1959 and have kept and expanded. In the *Lactuca* group, there are 961 cultivars of lettuce, mostly from the United States, the United Kingdom, and the Netherlands, but also from many other countries. There are 998 plant introduction lines, mostly from countries near the Mediterranean Sea. There are several hundred genetic stocks, consisting of lines with one or more single-gene mutations, as well as early generations from the crosses made for inheritance studies. There are about 2000 breeding lines. Finally, there are miscellaneous accessions consisting of unnumbered introductions, wild material collected from various places in the United States, and other lines. In the *Cichorium* group, there are 277 cultivars of endive and escarole, and various chicory types (witloof, radicchio, and green leafed cultivars). There are also 19 plant introduction lines, nearly all from Western Europe.

The *Lactuca* spp. plant introduction lines consist of 787 entries of *L. sativa*, 138 of *L. serriola*, 13 of *L. saligna*, six of *L. virosa*, and 17 from other species, many of which do not cross with cultivated lettuce. The *Cichorium* spp. plant introduction lines consist of 18 entries of *C. endivia* and one of *C. intybus*.

The only economic character descriptors for *Lactuca* spp. culti-

Table 1. Descriptors for cultivars and plant introduction lines of *Lactuca* spp.

Cutivar descriptors	Plant introduction descriptors
Name	P.I. number (or other)
Date (received or recent increase)	Date of recent increase
Type (crisphead, butterhead, cos, leaf, stem, "Latin")	Species
	Type (*L. sativa* only)
Source—country, organization	Country of origin
Seed color	Seed weight
Leaf color—greenness, anthocyanin	Seed color
Resistances to disease, insect[z]	Leaf dimensions
Synonym	Leaf color
	Leaf shape
	Anthocyanin
	Flowering time
	Flower diameter
	Plant height
	Resistances
	Notes

[z]Lettuce mosaic, big vein, tipburn, sclerotinia, corky root rot; lettuce root aphid.

vars are disease and insect resistance traits (Table 1). Of the diseases and insects listed, only lettuce mosaic virus (LMV) resistance has been evaluated for most cultivars; LMV resistance has been evaluated in our screening program since 1959. There is extensive information published on downy mildew resistance in cultivars, particularly of the butterhead type, but little of this information is yet in our descriptor listing. There is published information on resistance to root aphid, corky root rot, cabbage looper, sclerotinia, and tipburn, but very little has yet been entered into our lists.

Other economic characters—bolting resistance, head development characters, maturity time, yielding ability, and head size—have never been systematically or comprehensively rated either objectively or subjectively and, consequently, little information is available for listing. A newly formed Leafy Vegetable Crop Advisory Committee is beginning to prepare a list with this information. The *Lactuca* spp. plant introduction lines have more information listed for both objective and measurement traits (Table 1). The *Cichorium* spp. collection is new, and little information is listed.

The collections are held in household freezers, which provide satisfactory storage and greater longevity than the more common storage of 0 to 10°C and 40% to 50% RH. However, the containers are crowded and not readily accessible. Most of the seed is in paper envelopes in plastic freezer containers.

Lettuce and *Lactuca* spp. collections are found in other U.S. and European locations (Table 2). Conditions of storage vary. The largest collections listed are those at Wageningen, Wellesbourne, and Leningrad; most others are quite small.

In response to questions about the state of their collections, curators of several other species have indicated that there is a great deal of variability in collection size, types of materials in the collections, and the number of entries evaluated for each of the descriptors listed. It is apparent that those who wish to sample from the collections make very specific requests when they know what they want, and vague requests when they do not. In the latter case, the requestor is usually dealing with a newly arisen problem or is unfamiliar with the species requested.

Table 2. *Lactuca* spp. collections at other locations in the United States and other countries.

United States
Regional Plant Introduction Station, Pullman, Wash.
National Seed Storage Laboratory, Fort Collins, Colo.
Breeding collections
Univ. of California, Davis.
Cornell Univ., Geneva, N.Y.
Univ. of Florida, Belle Glade.
Outside the United States
Research Institute of Plant Protection, Prague, Czechoslovakia
Central Inst. for Genetics and Crop Plant Breeding, Gatersleben, G.D.R.
Institute for Horticultural Plant Breeding, Wageningen, Netherlands
Vavilov All-Union Inst. of Plant Industry, Leningrad, USSR
National Vegetable Research Station, Wellesbourne, U.K.
National Institute for Agricultural Variety Testing, Tapioszeli, Hungary
Germplasm Institute, Bari, Italy

[z]From ref. 1.

How, then, do we make our germplasm collections most informative and most useful? The first and most obvious step is to make the collections themselves grow. The more material there is in captivity, the greater is the potential for its usefulness. The capacity of a collection to provide information and raw material for crop improvement is limited to the extent that the size of the collection itself is limited. Very simply, then, we must collect as many materials as possible.

The traditional means is to collect materials in the crop species and in related species by traveling to centers of origin and development. This is expensive, time-consuming, and often physically difficult, but necessary. Many populations are in transition, i.e., they are often here today and gone tomorrow, and it is imperative that we maintain and even increase the pace of collection.

I have what may be an unusual suggestion for increasing our germplasm base in another way. Plant breeders throw away 90% to 99% of the material they observe because most recombinants from a cross are unsuitable in terms of the goals of the breeding program. On the other hand, it is possible that some of the genotypes may have future economic value. Consumer tastes change. Plant architecture changes may become necessary as tillage methods change. Early or late maturity may become more desirable than current needs dictate. Other characteristics may have to be altered in response either to increased knowledge of physiologic–genetic responses, or to economically or socially inspired outside requirements. Therefore, it would seem useful for plant breeders to save a random or selected sample of segregating populations in addition to the material selected for continued breeding. Seeds of the potentially useful material then could be kept in storage or forwarded to the appropriate curator or other collection holder.

It is necessary to stress the importance of domesticated material, including cultivars and breeding lines, as germplasm. Germplasm often is taken to mean only plant introduction material, and there is a strong implication that its usefulness is in supplying single genes for resistance to save a crop species from a certain disaster brought on by some disease or insect epidemic. This is the biotechnologist's germplasm; the plant breeder's germplasm also includes a vast array of genetic combinations that must be assessed for their usefulness in an equally vast array of environments.

Efficiency in sampling the collections is dependent on several factors. First, the collections must be permanent, i.e., they must be stored in an adequate facility, they must be secure, and they must be monitored for viability and renewed at appropriate times.

Second, collections must be evaluated for both noneconomic and economic traits, a massive job that requires support independent of the associated breeding or other research programs. A number of crop advisory committees delegate responsibility to appropriate researchers to perform specific evaluations: screen for resistance to a disease or insect, measure response to a particular stress condition, etc. This practice should be encouraged. It is relatively easy to do this for crops such as wheat and cotton, where there is a fairly large group of researchers to whom tasks may be assigned. It is more difficult with many horticultural crops, where the number of people associated with the crops may be relatively small.

Third, collections must be publicized so that researchers are aware of their existence, the kinds of information available about the entries, and the fact that the materials are obtainable. I believe that collections should be described and published by the holder as well as being placed in the Germplasm Resources Information Network (GRIN) system. This adds a measure of redundancy of information available, which I consider to be desirable for maximizing awareness of the existence and usefulness of the collection.

Finally, redundancy in the collections themselves is desirable. Any single entry for a species should exist in more than one location. Losses of single entries or sometimes catastrophic losses of entire collections or parts of collections can occur. Redundancy can prevent irretrievable losses.

Present practices in sampling are based upon the knowledge about and accessibility of the collections, the information available about the entries, and the purpose for which samples are needed. According to the curators, most or nearly all requests are specific, i.e., for a single item or for a group known to have a desired trait or traits. A substantial number of requests are somewhat vague; for example, the requestor may ask the curator to choose material that might be resistant to a disease. There are also many requests for a representative geographic distribution of materials. Rarely are requests for an entire collection or for a whole section of a collection.

One may hope that, in the future, efficiency of sampling will be improved. As the number of traits described increases and as the completeness of description of each trait increases, the greater will be the opportunity for a researcher to request specific items and the lesser the need for vague or shotgun requests. Along with the increases in information will be the greater accessibility of that information by inclusion in the GRIN system. This sort of information development will be most useful for known traits, and will enable the breeders to incorporate into their programs additional useful genes and gene systems for yield, resistance, earliness, etc. It will also provide the geneticist, the physiologist, the plant pathologist, the entomologist, etc. specified, identified materials for appropriate experiments.

It is well to keep in mind, however, that when the research must deal with a new, previously unknown disease or situation for which descriptors do not exist, little will be different. Then, as now, the technique of sampling will depend on the nature and urgency of the problem, the resources available for research, the perception of the researcher and the curator in determining the materials most likely to be useful in solving the problem, as well as other nonobjective considerations.

It is hoped that objective criteria for sampling will be more readily defined than at present. These criteria might include protein analysis of small seed samples to gain clues associated with possible disease resistance, development of probability parameters to estimate sample sizes needed to discover genes for resistance or other traits and availability of and knowledge about genes from other species that may be transferred via new technology to the species of interest. Whatever the future holds, there will be no substitute for a well-financed system of collection, maintenance, evaluation, and distribution of germplasm in its most broadly defined sense.

Literature Cited

1. Toll, J. and D.H. van Sloten. 1982. Directory of germplasm collections, 4. Vegetables. Int'l. Board for Plant Genetic Resources. IBGPR Secretariat, Rome.

Genetic Considerations in the Collection and Maintenance of Germplasm

Paul J. Fitzgerald
Agricultural Research Service, Peoria, IL 61604

Plant germplasm has become a topic of interest and concern at the highest levels of our government and in many other parts of our society. Part of this concern no doubt relates to the increased awareness by leaders in the government and in the public at large that U.S. agricultural productivity and stability is dependent on plant resources that are not native to the United States. The United States, as compared to many agricultural producers in the world, has no centers of origin for important food and fiber crops. In fact, the United States has no major crop native to its geographical area. Of all the crops important to U.S. agriculture, only sunflower (*Helianthus* spp.), Jerusalem artichoke (*Helianthus tuberosus* L.), grape (*Vitis* spp.), pecan [*Carya illinoensis* (Wangenh.) C. Koch] and other hickories (*Carya* spp.), blueberry (*Vaccinium* spp.), cranberry (*Vaccinium macrocarpon* Aiton), and a few other miscellaneous species of minor importance are native to the United States.

Human beings have displayed an interest in plant germplasm throughout recorded history, and probably for a long time before. The involvement of Alvin Toffler led him, in a 1984 television special, to say "10,000 years ago some genius reached out and altered our lives forever . . . by planting a seed". We are beginning to appreciate the essential role plant germplasm has played in the development of all societies throughout history and, perhaps more important, in its essential role in our future.

Our earliest understanding of the importance of germplasm began with the first seed savers. These primitive agriculturists developed a practice of collecting seed from chosen plants for the next crop and, in so doing, unknowingly became the first to engage in activities later to become a part of plant breeding. The scientific basis for understanding, however, was not to come until about 1900, when Gregor Mendel's principles of heredity were rediscovered and provided evidence that characteristics of all organisms are passed from one generation to the next in an orderly and predictable manner, controlled by genetic factors, later called genes. With this new knowledge of plant behavior, improvement of important crop plants became an attainable goal of plant breeders and program leaders. These efforts to manipulate genes to produce better plants have been one long progression from observation to explanation, from art to science.

The recognition of the value of specific genes led to other conclusions. Continued improvement of plants—those for food, feed, fiber, industrial uses, and research—is dependent on a continuing and adequate supply of germplasm with useful genes. A major supply of needed genes will have to be provided by accessions (seed samples) held in storage in various crop seed collections and by plants in clonal germplasm repositories.

Adequacy of the collections to provide useful genes for current and future programs will be dependent on how successful collecting efforts have been in obtaining good representation of the existing genetic diversity within crop species and close relatives and on the quality of preservation procedures and facilities. Collectively, genes within a crop species, or a group of related species, represent genetic diversity with which scientists must work. It should be the goal of all collections to preserve the highest level of useful genetic diversity that is feasible. Further, to be most useful, the genes collected must be readily available to those who wish to use them in plant breeding and research; otherwise, progress in crop improvement may be halted. Availability is linked to identification and evaluation, and it is in this area that much additional work is needed.

Plant breeders were probably the first to recognize the need and to take action to gather, for themselves, a collection of plants containing potentially useful genes for later use. These efforts led to the awareness that special handling and facilities were needed to preserve the germplasm samples, usually seed, until needed. The periodic requirement to grow the samples in a protected nursery to multiply the seed and to make preliminary observations and evaluations of the plants also was recognized as necessary to maintain seed collections with acceptable viability and genetic reproducibility. Further, it also was recognized that some valuable crop germplasm, such as fruit and nut species, could not be suitably preserved as seeds. Consequently, a system of clonal germplasm repositories was initiated to protect and preserve, as plantlets or whole plants, the most important fruit and nut species. A national