McFerson outlines the importance of molecular markers in developing core collections of crop species, and discusses the types of decisions that are routinely faced by the director of a germplasm repository. José Luis Solleiro describes the recent legislation to restrict germplasm acquisition in Latin America, as well as the feelings of policy makers about what they view as the theft of germplasm resources.

We hope that this colloquium will stimulate further work on cataloging the diversity of crop species and developing optimal preservation strategies. As space limitations become increasingly

stringent, the repositories must shift from being general storehouses of material provided by a hodgepodge of germplasm workers to systematic collections of highly representative crop genotypes. Of course, valuable material needs to be saved that is in danger of eroding *in situ*, but most importantly, highly representative collections of the total diversity of each crop species needs to be maintained for ready deployment by breeders. Judicious selection of material is particularly critical in Andean countries where further access to native populations and land races is being restricted.

Evolution, Population Genetics, and Germplasm Preservation

Tim Holtsford

Division of Biological Sciences, University of Missouri, Columbia, MO 65211

Jim Hancock

Department of Horticulture, Michigan State University, East Lansing, MI 48824

The collection and maintenance of plant germplasm requires a basic knowledge of both evolutionary mechanisms and evolutionary history of the crop in question. Germplasm collectors need to understand the patterns and processes of evolution associated with the particular crop so that they can most effectively allocate their limited time and resources while collecting a representative sample of genetic variation. Once a collection is made, germplasm managers should seek to minimize *ex situ* evolution, preserving as much of the natural genetic variation as possible for future breeding programs.

The basic evolutionary forces of mutation, selection, gene flow, and genetic drift determine the amount and distribution of genetic variation within and among populations. An understanding of these forces, coupled with ecological and life-history information, can help in designing a sampling strategy that makes efficient use of time and resources while yielding a representative collection. Evolutionary history is also valuable information for germplasm collectors. A phylogenetic tree, which depicts the evolutionary relationships among races of the focal species and its wild relatives, can identify other lineages that may be of use to breeders.

Evolutionary considerations for germplasm preservation, including practical recommendations for sampling, have been recently treated by Schoen and Brown (1995) and Brown and Marshall (1995). Herein, we provide a brief review of the evolutionary forces and their pertinence to germplasm collection and preservation, and present an example of the use of population genetic analyses to uncover the structure of genetic variation in an endangered annual plant *Clarkia springvillensis* Vasek.

EVOLUTIONARY CONSIDERATIONS

Mutation provides the variability that is the raw material of evolution; the other evolutionary forces merely rearrange and sieve the products of mutation. Spontaneous rates of mutation are generally thought to be low (10^{-5} to 10^{-6} per locus per generation) (Kahler et al., 1984; Stadler, 1942). Most mutations, even advantageous ones, rarely persist because their initial low frequency makes them particularly susceptible to being lost through random forces (Crow and Kimura, 1970; Falconer, 1981). Therefore, the genetic variability present in most extant populations has accumulated over millions of years. Mutagenesis, the deliberate production of mutations at a rate above the normal background level, has proven to be very important in creating

genetic maps and in identifying genes responsible for particular traits, but has contributed little usable material to plant breeders. This is why the role of germplasm conservation is so critical—mutation is a slow process that human efforts can at best poorly duplicate. We must preserve extant genetic variation because it cannot be easily replaced.

Genetic drift is a change in gene frequency due to random events. Such changes are most dramatic in very small populations where rare alleles can be lost due to sampling errors. Drift can have an impact at two levels of germplasm management: during the collection process and in *ex situ* preservation. Where native population sizes are quite small, the germplasm collector should make every effort to sample a large number of populations, even at the expense of individuals per population or propagules per individual. In species with highly variable native population sizes, Schoen and Brown (1995) have suggested that the size of a germplasm collection should be proportional to population size. Germplasm managers need to minimize the effects of drift in their collections by maintaining as many of the original clones of perennial material as possible, and periodically growing large populations of annual material that is stored as seed.

Natural selection is perhaps the best known of the evolutionary forces. In response to selection alone, those genes that confer a survival or reproductive advantage to their bearers will increase in frequency. The magnitude and direction of natural selection will vary across a species' range because the ecological challenges may differ among populations, or even among microsites within a population. Preserving the genes that allow plants to survive and reproduce in spite of various physical or biological stresses is the primary goal of germplasm preservation. Therefore, germplasm collectors should endeavor to sample populations from throughout the species' geographical and ecological ranges. Within populations, care should be taken to identify and sample all microhabitats; e.g., dry and wet, sunny and shady, various soil types, etc.

Once a collection is made, natural selection will act to adapt the germplasm to its new *ex situ* conditions (Allard, 1988). Because directional selection may decrease genetic variation, germplasm managers should try to minimize the selection acting upon their collections so that breeders will have more genes available. For perennial plants, this may mean keeping alive even the seemingly weaker plants because their traits may be beneficial in other environments or if challenged by a new pest. For annual plants, propagation by single-seed descent, or a constant number of seeds from each maternal plant, would preserve more genetic variation than harvesting in bulk.

Gene flow can act to inject new variability into a recipient population and can have a homogenizing effect on the genetic differentiation between populations. In general, if the migration rate is greater or equal to selection, then the impact of selection on population differentiation will be insignificant (Ellstrand and Marshall, 1985). Likewise,

Received for publication 10 Feb. 1998 Accepted for publication 9 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.

very small amounts of gene flow can overwhelm the effects of genetic drift. Unless populations exceed 200 individuals, migration rates as low as 0.001 diploid migrants (seeds) per generation will prevail over genetic drift (Grant, 1963). As a result, where gene flow is high between populations, the collector may want to concentrate on sampling fewer, but more widely spaced populations.

Levels of gene flow have been estimated in many species using genetic markers (Hancock et al., 1997; Kirkpatrick and Wilson, 1988; Klinger et al., 1991 and 1992; Wilson and Manhart, 1993), but where such information is not available, knowledge of a species' mating system, mode of pollination, and seed dispersal patterns can aid in predicting levels of interpopulational differentiation. In general, outcrossing species have higher levels of gene flow than self-pollinating species and lower levels of intra- and interpopulational differentiation (Hamrick and Godt, 1997). Insect-pollinated species tend to have more limited dispersal distances than wind-pollinated species and as a result their populations are generally more substructured and differentiated (Hamrick and Godt, 1990). This suggests that the distance between sampling locations for wind-pollinated or outcrossing species should be greater than insect-pollinated or selfing ones. Conversely, *ex situ* isolation distances should be greater for wind-pollinated or outcrossing species than for insect-pollinated or selfing ones. Ironically, surveys of recommended isolation distances for the production of pure seed stocks average less for wind-pollinated species (200 m) than for insect-pollinated ones (500 m) (Hancock et al., 1997; Levin and Kerster, 1974).

In general, most seeds and pollen do not move very far from their home population. Pollen movement of both insect- and wind-pollinated species is leptokurtic and the bulk of pollen travels only 5–10 m (Colwell, 1951; Handel, 1983; Hokanson et al., 1997). The distribution of seed dispersal is also leptokurtic, with most propagules landing within a few meters of the mother plant (Levin and Kerster, 1974). However, significant long distance dispersal does occur. Wind-carried pollen has been found hundreds of kilometers from its source (Ehrlich and Raven, 1969; Moore, 1976) and insects have been shown to carry pollen for several kilometers (Devlin and Ellstrand, 1990; Ellstrand and Marshall, 1985).

Gene flow among accessions in germplasm collections should be avoided by separating populations in space or time. Germplasm collections should preserve not only the individual genes but also the genetic structure of the species; the distribution of genetic variation within and among accessions should be kept the same as that among the source populations. The structure of genetic variation is an intrinsic attribute resulting from many generations of evolution that cannot be duplicated *ex situ*. Therefore, if the purpose of the collection is to preserve the genetic attributes of a species, the genetic differentiation among populations should be preserved, especially if the collection may someday be used to reintroduce rare plants into the wild.

For breeders, minimizing gene flow between accessions is important for at least two reasons. First, gene flow, combined with selection for adaptation to the *ex situ* environment can diminish the genetic variation in the collection. For example, genes from a xerically-adapted population may be lost through the combined action of gene flow from a more mesically-adapted population and selection for adaptation to a mesic *ex situ* location. Second, if breeders want to increase heterozygosity or combine traits from populations that evolved under different conditions, they may waste time and resources if the accessions have already been homogenized by gene flow at the germplasm repository.

Gene flow is defined as the movement of genes *between* populations, but the dispersal of seeds and pollen *within* populations is also important to germplasm conservation programs. Low dispersal distances result in the substructuring of populations into "neighborhoods." The differences in gene frequencies among neighborhoods will cause the frequency of heterozygotes of the total population to be lower than the average of the subpopulations, even in the absence of inbreeding, a phenomenon known as the Wahlund Effect (Wahlund, 1928). The size of neighborhoods is greatly affected by a species' breeding system, seed dispersal mechanism, and selection for microsite adaptation.

In highly substructured populations, the germplasm explorer should

make an effort to sample as many neighborhoods as possible. Where there is limited seed and pollen dispersal leading to substructure within populations, levels of differentiation between populations may also be high. To sample this diversity, the collector may have to either increase the total number of propagules collected or reduce the number of propagules collected per plant. The mode of pollination can further complicate this issue, as seeds from single fruits of wind-pollinated species tend to have multiple male parents, while seeds from single fruits of animal-pollinated types tend to have fewer male parents (Brown and Marshall, 1995). As a result, more fruits per plant should be collected in animal- than in wind-pollinated species.

IMPORTANCE OF MARKER LOCI

Genetic information from isozymes or other marker genes is extremely useful to the germplasm collector because the degree of genetic differentiation among populations or subpopulations can be estimated directly (Bretting and Widrechner, 1995; Gepts, 1995; Schoen and Brown, 1995). If genetic information is not available, the life history, growth habit, and seed and pollen dispersal mechanisms of the focal species should be assessed, because these traits are often associated with the degree of interpopulation differentiation (Hamrick and Godt, 1990). Genetic information from close relatives may also give the collector an idea as to how differentiated populations or subpopulations might be. Marker genes are useful for discerning the spatial scale over which genetic differentiation has occurred, and can therefore save time and money by reducing redundancy in collections. However, marker genes are not necessarily the most important genes to consider. Many isozymes and DNA markers, such as restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNAs (RAPDs) and microsatellite loci are fairly neutral with respect to plant fitness and so do not necessarily reflect the degree of genetic differentiation at genes affecting vigor or yield (Hamrick and Godt, 1990; Harrison et al., 1997). Therefore, efforts should be made to identify all possible habitats that occur across the range for the species, and if a species occurs in a range of environmental conditions, collections should be made in each microclimate, regardless of whether or not marker gene surveys indicate differentiation. Conversely, the neutrality of marker genes makes them better indicators of evolutionary history than are phenotypic traits. Therefore, neutral marker genes are well suited for identifying unique lineages which should be included in a collection.

Clarkia springvillensis Vasek (Onagraceae) provides a good example of the utility of marker genes for assessing spatial and temporal population structure. McCue et al. (1996) assayed seven isozyme-encoding loci at two spatial scales in each of three populations of this narrowly-distributed endangered species from central California. The greatest genetic differentiation was found among patches of plants separated by tens of meters; very little differentiation was found within these patches. There were no obvious microclimatic differences among the patches, nor was there any relationship between genetic distance and geographic distance among patches. Interestingly, two populations, which were composed of several patches each, were on average very little differentiated from each other even though they were 800 m apart. Therefore, a representative sample of the two populations could be obtained by collecting from only one population, provided all the patches within that population were included. The fact that the patches were more differentiated from each other than the more widely separated populations suggests that the genetic structure was not in equilibrium. A third population, 8 km distant, merited a separate sample as it was significantly differentiated from the other two and contained unique genes at the marker loci.

Another often overlooked but critical issue that must be faced by the germplasm explorer is whether there is temporal structuring of the genetic variability. Temporal structuring of a plant population arises from an important reservoir of genes that is often ignored, the seed bank. The seed bank consists of viable but quiescent seeds that remain in the soil while the above-ground cohort germinates and develops "normally." Assessments of genetic variability in the seed bank are rare, but in both known cases, the seed banks were found to contain more overall genetic diversity than mature plants (McCue and Holtsford,

1998; Tonsor et al., 1993). In *Clarkia springvillensis* the seed bank contained not only more allelic diversity (Nei's $H = 0.355$ vs. 0.260; Nei, 1973), but also a gene not found in the above-ground population (McCue and Holtsford, 1998). Crop species and their wild relatives should be evaluated for their potential to maintain seed banks in the soil. Consideration of this gene pool will be of great importance in cases where there is substantial differentiation between the seed bank and adult population, and repeat sampling visits to a site may be warranted. Because the seed bank may contain cohorts that germinate in response to different environmental conditions (Cabin 1996), germplasm collectors should strongly consider resampling populations in unusual seasons, e.g., in a dry year after several wet years. The potential for considerable genetic differentiation between the seed bank and adult population make consideration of this alternate gene source extremely important.

PLANNING A SAMPLING STRATEGY

As the considerations outlined above attest, the germplasm explorer is faced with a complex array of questions that must be answered to develop an optimized sampling strategy. A hard decision must be made at the onset about exactly how many samples can be made and stored, and then decisions must be made as to how many populations, subpopulations, individuals, and propagules per individual. In many cases, a comprehensive collection of germplasm across a species' entire range will be impossible due to space and financial constraints, so targeting certain areas or populations becomes critical. As Brown and Marshall (1995) state, "This is ensured not by collecting large quantities, but a judicious division of the target region...." The task facing the germplasm explorer is best approached with information on the genetic structure among geographic races, populations, subpopulations, individuals, and propagules per individual. Unfortunately, few germplasm collection projects will have enough resources to gather all the desired genetic information *a priori*. Additionally, genetic markers will not reveal all of the genetic structure that exists in target populations—especially in the genes that count most: unidentified genes affecting vigor and yield. The germplasm collector must therefore pay careful attention to geographic distribution and to ecological and life history attributes that are typically correlated with the degree of genetic differentiation at various levels. To optimize sampling strategies, information needs to be gathered on species': 1) geographic distribution, 2) ecological range, 3) population sizes, 4) mating system, 5) seed dispersal mechanism, 6) mode of pollination, and 7) age structure, including the dormant seed bank.

Literature Cited

- Allard, R. 1988. Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. *J. Hered.* 79:225–238.
- Bretting, P.K. and M.P. Widrechner. 1995. Genetic markers and plant genetic resource management. *Plant Breeding Rev.* 13:11–85.
- Brown, A.H.D. and D.L. Marshall. 1995. A basic sampling strategy: Theory and practice, p. 75–92. In: L. Guarino, V. Ramanatha Rao, and R. Reid (eds.). *Collecting plant diversity: Technical guidelines*. CAB Intl., Wallingford, Kent, England.
- Cabin, R. 1996. Genetic comparisons of seed bank and seedling populations of a perennial mustard, *Lesquerella fendleri*. *Evolution* 50:1830–1841.
- Colwell, R. 1951. The use of radioactive isotopes in determining spore distribution patterns. *Amer. J. Bot.* 38:511–523.
- Crow, J. and J. M. Kimura. 1970. *An introduction to population genetics theory*. Harper and Row, New York.
- Devlin, B. and N.C. Ellstrand. 1990. The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution* 44:248–259.
- Ehrlich, P.R. and P.H. Raven. 1969. Differentiation of populations. *Science* 165:1228–1232.
- Ellstrand, N.C. and D.L. Marshall. 1985. Interpopulational gene flow by pollen in wild radish, *Raphanus sativus*. *Amer. Nat.* 126:606–616.
- Falconer, D.S. 1981. *An introduction to quantitative genetics*. Longman, New York.
- 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.
- Grant, V. 1963. *The origin of adaptations*. Columbia Univ. Press, New York.
- Hamrick, J.L. and M.J.W. Godt. 1990. Allozyme diversity in plant species, p. 43–63. In: A.H.D. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (eds.). *Plant population genetics, breeding, and genetic resources*. Sinauer Assoc., Sunderland, Mass.
- Hamrick, J.L. and M.J.W. Godt. 1997. Allozyme diversity in cultivated crops. *Crop Sci.* 37:26–30.
- Hancock, J.F., R. Grumet, and S.C. Hokanson. 1997. The opportunity for escape of engineered genes from transgenic crops. *HortScience* 31: 1080–1085.
- Handel, S.N. 1983. Pollination ecology, plant population structure and gene flow, p. 163–211. In: L. Real (ed). *Pollination biology*. Academic Press, Orlando, Fla.
- Harrison, R.E., J.J. Luby, G.R. Furnier, and J.F. Hancock. 1997. Morphological and molecular variation among populations of octoploid *Fragaria virginiana* and *F. chiloensis* (Rosaceae) from North America. *Amer. J. Bot.* 84:612–620.
- Hokanson, S., R. Grumet, and J. Hancock. 1997. Effect of border rows and trap donor ratios on pollen mediated gene movement. *Ecol. Appl.* 7:1075–1081.
- Kahler, A.L., R.W. Allard, and R.D. Miller. 1984. Mutation rates for enzyme and morphological loci in barley (*Hordeum vulgare* L.). *Genetics* 106:729–734.
- Kirkpatrick, K.J. and H.D. Wilson. 1988. Interspecific gene flow in *Cucurbita*: *C. texana* versus *C. pepo*. *Amer. J. Bot.* 75:519–527.
- Klinger, T., D.R. Elam, and N.C. Ellstrand. 1991. Radish as a model system for the study of engineered gene escape via crop-weed mating. *Conserv. Biol.* 5:531–535.
- Klinger, T., D.R. Elam, and N.C. Ellstrand. 1992. Crop-weed hybridization in radish (*Raphanus sativus*): Effects of distance and population size. *Amer. J. Bot.* 79:1431–1435.
- Levin, D.A. and H.W. Kerster. 1974. Gene flow in seed plants. In: Th. Dobzhansky, M. Hecht, and W. Steere (eds.). *Evolutionary biology*. Vol. 7. Plenum Press, New York.
- McCue, K.A. and T.P. Holtsford. 1998. Seed bank influences on genetic diversity in the rare annual *Clarkia springvillensis*. *Amer. J. Bot.* 85:30–36.
- McCue, K.A., E.S. Buckler, and T.P. Holtsford. 1996. A hierarchical view of genetic structure in the rare annual plant *Clarkia springvillensis*. *Conserv. Biol.* 10:1–11.
- Moore, P. 1976. How far does pollen travel? *Nature* 260:280–281.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. (USA)* 70:3321–3323.
- Schoen, D.L. and A.H.D. Brown. 1995. Maximizing genetic diversity in core collections of wild relatives of crop species, p. 55–76. 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.
- Stadler, L.J. 1942. Some observations on gene variability and spontaneous mutation. *Spragg Mem. Lect. (Michigan State Univ.)* 3:3–15.
- Tonsor, S.J., S. Kalisz, J. Fisher, and T.P. Holtsford. 1993. A life-history based study of population genetic structure: Seed bank to adults in *Plantago lanceolata*. *Evolution* 47:833–843.
- Wahlund, S. 1928. Zusammensetzung von Populationen und Korrelation—Sercheinungen von Standpunct der Verebungslehre aus Bedrachtet. *Hereditas* 11:65–106.
- Wilson, H. and J. Manhardt. 1993. Crop/weed gene flow. *Chenopodium quinoa* Willd. and *C. berlandieri* Moq. *Theor. Appl. Genet.* 86:642–648.