
Assessing the Geographic Representativeness of Genebank Collections: the Case of Bolivian Wild Potatoes

R. J. HIJMANS,* K. A. GARRETT, Z. HUAMÁN, D. P. ZHANG, M. SCHREUDER, AND M. BONIERBALE

International Potato Center, Apartado 1558, Lima 12, Peru

Abstract: *Genebank collection databases can be used for ecogeographical studies under the assumption that the accessions are a geographically unbiased sample. We evaluated the representativeness of a collection of wild potatoes from Bolivia and defined and assessed four types of bias: species, species-area, hotspot, and infrastructure. Species bias is the sampling of some species more often than others. Species-area bias is a sampling that is disproportionate to the total area in which a species is found. Hotspot bias is the disproportionate sampling of areas with high levels of diversity. Infrastructure bias is the disproportionate sampling of areas near roads and towns. Each of these biases is present in the Bolivian wild potato collection. The infrastructure bias was strong: 60% of all wild potato accessions were collected within 2 km of a road, as opposed to 22%, if collections had been made randomly. This analysis can serve as a guide for future collecting trips. It can also provide baseline information for the application of genebank data in studies based on geographic information systems.*

Evaluación de la Representatividad Geográfica de las Colecciones de Bancos de Genes: El Caso de las Papas Silvestres Bolivianas

Resumen: *Las bases de datos de los bancos de germoplasma pueden ser usadas para estudios ecogeográficos bajo el supuesto que las entradas constituyen una muestra geográficamente imparcial. Evaluamos la representatividad geográfica de una colección de papas silvestres de Bolivia y definimos y evaluamos cuatro tipos de sesgos: sesgos de especie, de especie-área, de áreas con gran diversidad (hotspot), y de infraestructura. El sesgo de especie implica el muestrear más algunas especies que otras. El sesgo de especie-área es un muestreo que es desproporcionado con respecto al total del área en la cual se encuentra una especie. El sesgo de "hotspot" es el muestreo desproporcionado de áreas con niveles altos de diversidad. El sesgo por infraestructura es aquel muestreo desproporcionado de áreas cercanas a carreteras y pueblos. Cada uno de estos sesgos se presenta en la colección de papas silvestres de Bolivia. El sesgo por infraestructura fue muy elevado: 60% de todas las entradas de papas silvestres fueron colectados dentro de un radio de 2 km de distancia de las carreteras, cuando se debería esperar un 22% si las colectas se hubieran hecho de manera aleatoria. Este análisis puede servir como guía para futuras exploraciones y proporciona una base para la aplicación de los datos de bancos de genes en estudios basados en sistemas de información geográfica.*

*email r.bijmans@cgiar.org

Paper submitted November 23, 1998; revised manuscript accepted March 1, 2000.

Introduction

The collection of plant genetic resources provides access to the greatest possible amount of genetic variability in a particular species and helps reveal the ecological and geographic distribution of plant species (Bennett 1970). A genebank is a collection of a particular crop and its wild relatives and, ideally, includes at least one example of each alternative allele for each locus (Chapman 1984). Because researchers can hope to sample only a fraction of the genetic variation that occurs in nature, it is important that this sample be as large as possible and contain the maximum amount of useful variation for both present and future use (Allard 1970; Brown & Marshall 1995).

The role of genebanks in the conservation and use of germplasm for crop improvement is well established. More recently, genebanks have been used as a source of data on the ecological and geographic distribution of species. For example, Jones et al. (1997) used genebank databases and data on climate surfaces to identify areas where the wild common bean (*Phaseolus vulgaris*) might occur but had not actually been recorded. Analysis of georeferenced genebank data also proves useful in identifying areas of high diversity (Nabhan 1990; Frankel et al. 1995); targeting genetic resources for breeding programs (Nabhan 1990; Guarino et al. 1999); and selecting and designing sites for in situ conservation (Guarino et al. 1999). The scientific value of these alternative uses of genebank data is questionable if the database of localities used for geographic analysis is small, skewed, or otherwise unrepresentative of the total natural distribution (Nabhan 1990).

Diversity is not uniformly distributed in space or among taxonomic groups. Some genetic heterogeneity can be attributed to selection and gene flow and reflects ecogeographic adaptation, whereas other heterogeneity is due to factors such as founder effects and genetic drift (Guarino 1995). Although ecological factors are a major determinant of genetic diversity (Bennett 1970), the amount of genetic variation that can be attributed to ecogeographic factors rarely exceeds 50% (Chapman 1989).

Because the distribution of diversity is not known prior to data analysis, successful collecting in terms of diversity may depend on the proper identification of populations closely adapted to specific environments and land-use patterns (Brown & Marshall 1995). Therefore, it is often suggested that collectors take a stratified sample, dividing the population into as many ecological and geographical subpopulations as possible and drawing a sample from each one (Cochran 1977). In doing so, it is important to include extremes of the range of a species (Allard 1970). Although these subpopulations may not present great genotypic variation, they may harbor unique traits or taxa (Von Bothmer & Seberg 1995). At present, the optimal scales for accurate sampling of genetic diversity have not been determined. Brown and

Marshall (1995) suggest sampling 50 plants from each of 50 sampling populations in each region, but they do not clearly define the size of those regions.

A genebank collection may deviate from the ideal sample for a number of reasons. Practical constraints include limits on accessibility, time, and funding. In addition, the collection may be skewed if the germplasm was collected with a specific objective such as rescuing material at risk of extinction, identifying a specific trait, or fulfilling a specific research need (Engels et al. 1995).

Identification of optimal accession sample size for a representative genebank collection is an important issue, but it will not be covered here. Instead, we analyze how genebank collections can systematically deviate from a representative sample in a relative sense: whether some species and/or areas have been sampled more intensively than others. Whether or not a sample may be considered representative depends partly on how representative sampling has been defined. Samples may be thought of as drawn from a pool of individual organisms or from a pool of geographic areas (Hayek & Buzas 1997). If drawn from a pool of individual organisms, a random sample would tend to produce collections with proportionately more accessions from the more abundant species or common varieties. Whether or not that is desirable depends on the specific research goals. To assess whether a genebank collection is representative in the relative sense, we have defined four types of bias. Three of these biases are spatial: species-area, hotspot, and infrastructure bias. The fourth, species bias, is not. We studied the presence of these four types of bias in a genebank collection of wild potatoes from Bolivia.

Species Bias

Species bias refers to sampling more from one species than from another. Species bias might result from differences in the probability of finding one species relative to other species because of differences in abundance or because of collectors' preferences for a particular species due to a specific expected trait or a specific research need. In terms of genetic resource applications, species bias could be justified if, for example, some species were considered more useful than others because they contained more genetic diversity or desirable agricultural traits.

Species-Area Bias

Species-area bias refers to over- or undersampling of a species in relation to the size of the area in which it occurs. Theoretically, to maximize genebank diversity, the number of accessions collected per species should increase with the size of the area in which it is found, assuming that a larger distribution area implies a proportionate in-

crease in intraspecific diversity. This is true, in general, although biological diversity is more directly dependent on ecological diversity and environmental stress factors (Nevo 1998). The increase in diversity with area tends to be logarithmic (Rosenzweig 1995; Magurran 1988).

Therefore, a species-area bias exists if a collection contains too many accessions from some species, and hence too few from others, relative to the size of the areas in which the species appear. No species-area bias exists if each species is represented proportionately according to its geographic range. Unless all species distribution areas are the same size, there will be a species bias, a species-area bias, or both.

Hotspot Bias

Hotspot bias refers to excessive or insufficient collection in certain geographic areas according to their level of genetic diversity. This is likely to occur if a genebank contains material from different expeditions. To avoid this bias, collection expeditions should cover complementary areas. Collectors may replicate parts of previous expedition routes to find particular species and to ensure that they obtain a sufficient amount of germplasm. Although it is important to consult genebank and herbarium databases and the route maps of previous collectors to design a collecting strategy (Nabhan 1990; Brown & Marshall 1995), this might lead to redundancy if collectors aim only for areas where previous studies and collections indicate that high species diversity exists. Consequently, some areas may be explored more intensively than others, creating a hotspot bias. The collection would have too many accessions from certain highly diverse areas, in a relative sense, despite the potentially high level of species diversity.

Infrastructure Bias

Infrastructure bias refers to oversampling near roads and towns. For reasons of efficiency, logistics, and convenience, collectors tend to follow the roads that connect the main towns. Hermann (1988) showed that most Andean tuber crop collection sites in Ecuador were located along the Pan-American Highway and other major roads. Based on herbarium records, Von Bothmer and Seberg (1995) produced a distribution map for *Elymus cordilleranus* in South America which shows distinct aggregation around La Paz (Bolivia), Lima (Peru), and the Pan-American Highway in Ecuador. This does not reflect species distribution accurately (Von Bothmer & Seberg 1995). Øllgaard (1995) analyzed the geographic distribution of *Huperzia* based on collections considered representative and of sufficient quality for determining priorities in selecting biodiversity reserves. He found that these collec-

tions showed a strong tendency to be located near roads and special collection sites.

Methods

Wild Potato Genebank Data

The wild tuber-bearing *Solanum* species (Solanaceae sect. *Petota*) are relatives of the cultivated potato (*Solanum tuberosum* and six other species; Hawkes 1990). They appear in the Americas from Colorado in the United States to Argentina and Chile and are most abundant in the Andes of Peru and Bolivia (Hawkes 1990). Wild potatoes have been used in modern breeding programs to improve the cultivated potato since the early 1900s when Mexican *S. demissum* was used to breed for resistance against *Phytophthora infestans*, the fungus that causes potato late blight (Ross 1986). We restricted our study to data for wild potatoes collected in Bolivia. An overview and in-depth information about wild potato species from Bolivia are provided by Hawkes and Hjerting (1989), Hawkes (1990), and Ochoa (1990). A recent wild-potato collecting expedition in Bolivia is reported by Spooner et al. (1994).

Our data source was the Intergenebank Potato Database (IPD). The IPD contains passport and evaluation data from the world's six main wild potato genebanks. Passport data include location variables such as the geographic coordinates and the locality, province, and department of origin for each accession. Each accession represents a number of potato seeds stored in one or more of the genebanks. The IPD currently has over 7128 records, of which 1440 are from Bolivia. The Bolivian data were drawn from 18 different expeditions conducted from 1953 to 1994.

Geographic Information System

Each accession has passport data that include geographic coordinates. We used these data to incorporate the IPD into a geographic information system (GIS) with IDRISI, Arc/Info, and ArcView software. This yielded a digital map of collection sites consisting of points with the IPD's records as attributes. The location of these points was checked according to procedures described by Hijmans et al. (1999). This included the production of an overlay of the collection site map and an administrative boundary map of Bolivian departments and provinces to allow comparison of the province and department names from both sources. If coordinates were wrong or missing, new coordinates were assigned when possible, based on the IPD locality description. After applying these procedures, we found that 1051 out of 1440 Bolivian accessions in the database had confirmed geographic coordinates (Fig. 1) and could thus be used in our study.

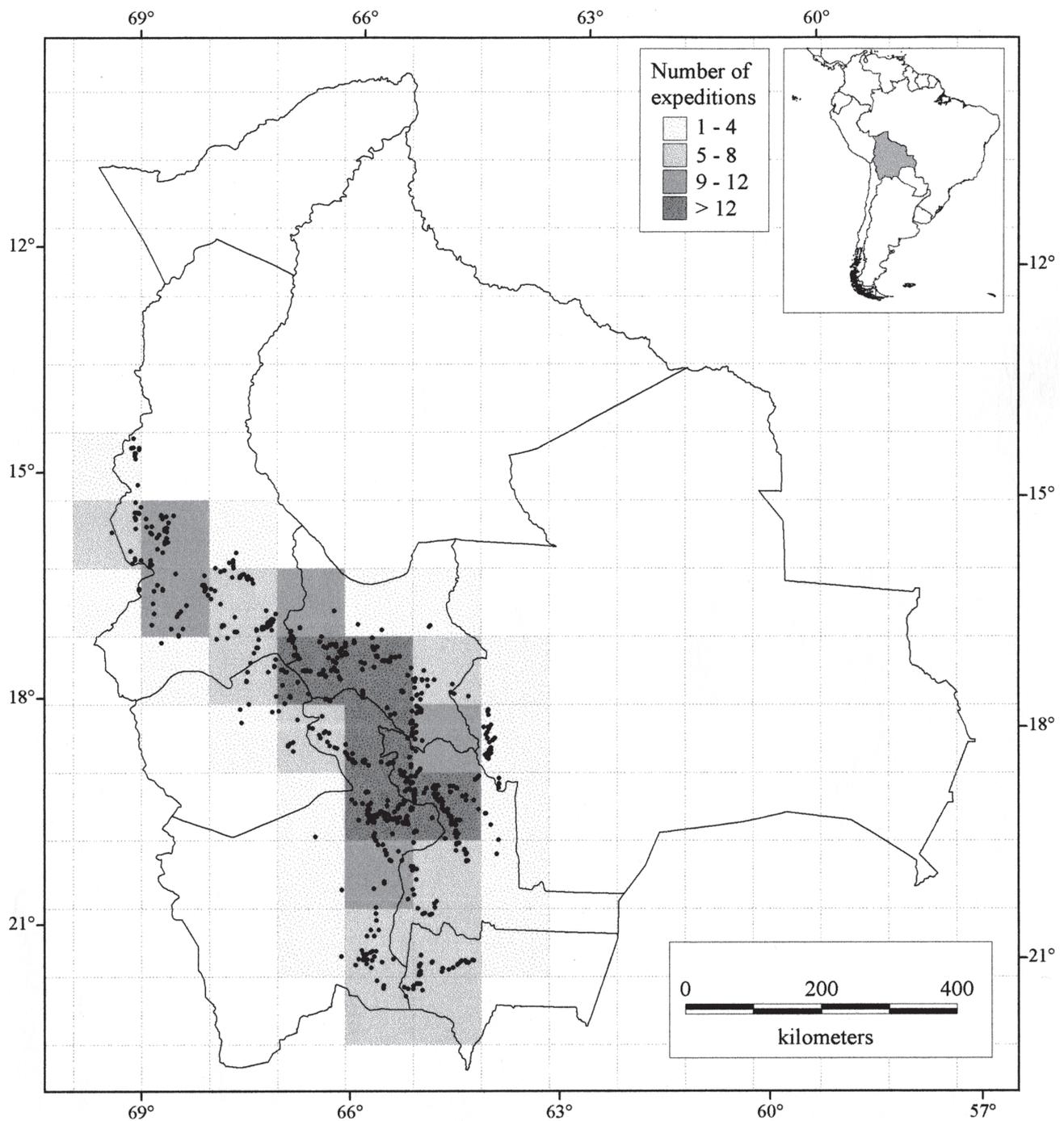


Figure 1. Bolivia departments (solid lines), locations where wild potatoes (*Solanum spp.*) were collected (dots, $n = 1051$), 100×100 km grid cells (dotted lines), and number of collecting expeditions per grid cell.

Species Bias

There is an ongoing debate about species boundaries in wild potatoes (Spooner & Castillo 1997; Van den Berg et al. 1998). We have treated all species, subspecies, and varieties included in the database as separate taxa, according to the identification criteria determined by collectors and genebank taxonomists. Using this criterion,

we found that there are 45 wild potato taxa (hereafter referred to as species) collected in Bolivia, conserved in ex situ genebanks, and documented in the IPD. We tabulated the number of accessions per species, A_S . The presence of a species bias was evaluated with a chi-square test to compare the observed number of accessions per species to the number per species that would be ob-

served if accessions were equally allocated between all species.

Species-Area Bias

We divided Bolivia into 100×100 km grid cells (Fig. 1), and for each species, S , we determined the number of grid cells in which it was collected (G_S). If there is no species-area bias, there should be a high degree of correlation between A_S and G_S . The relation between A_S and G_S was described by linear and logarithmic regression. We used a chi-square test to compare the observed number of accessions per species to the number per species that would be observed if accessions were equally allocated to each species based on the number of grid cells in which it occurs. The results of the comparison should reveal systematic deviations from the expected relationships between A_S and G_S .

Hotspot Bias

To examine the presence of a hotspot bias, we again divided Bolivia into 100×100 km grid cells (Fig. 1). For each grid cell, G , we determined the number of accessions collected, A_G ; the number of species collected, S_G ; and the number of expeditions that collected a sample (accession), E_G . The relationships between A_G , S_G , A_G/S_G and E_G were determined by linear regression. We determined a hotspot bias to be present if there was a significant (and positive) correlation between A_G/S_G and E_G . We determined the number of expeditions per area from published route maps (e.g., Spooner et al. 1994) or inferred them by plotting collection sites on a road map. Large gaps between collection sites were not included,

because the routes between them were not always clear and because they were likely to represent stretches where no active searching for wild potatoes took place.

Infrastructure Bias

To assess the presence of an infrastructure bias, we divided Bolivia into 1×1 km grid cells. For each grid cell, we determined the distance to the nearest road, nearest main road, and nearest large city (i.e., the department capital). Based on this grid, we calculated the distance between each accession area and the main road. This yielded a specific distribution, which we then compared by chi-square test with the "random" distribution, the distribution of the distance between the (main) road or city over the grid cells where wild potatoes are known to occur.

Results

Species Bias

According to our data, 45 wild potato species have been collected in Bolivia. The distribution of the number of accessions per species (A_S) was highly skewed (chi-square test; $p < 0.01$; Fig. 2). *Solanum acaule* was collected 166 times, whereas some other species appeared only once. The number of accessions of *S. acaule* was more than double that of the second most collected species, *S. sparsipilum*. Together, these two species accounted for 23% of all accessions in the genebanks. Seven species yielded 52% of the accessions, and 24 species (almost half of the total collection) yielded 90% of the accessions.

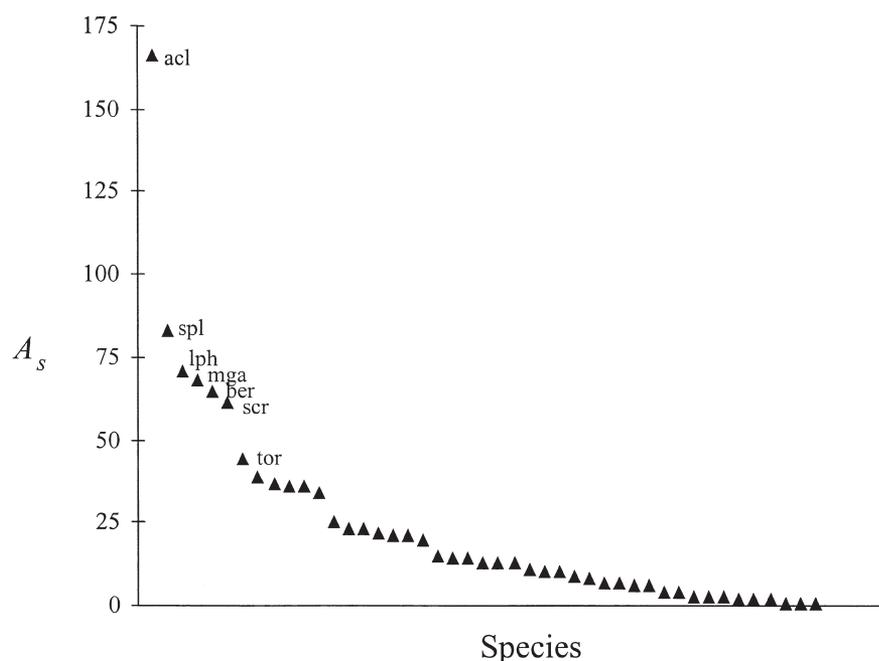


Figure 2. Number of Bolivian accessions per wild potato species (*Solanum spp.*), A_S , in the Inter-genebank Potato Database (acl, *S. acaule*; spl, *S. sparsipilum*; lph, *S. leptophyces*; mga, *S. megistacrolobum*; ber, *S. berthaultii*; scr, *S. sucrensis*; tor, *S. toralapanum*).

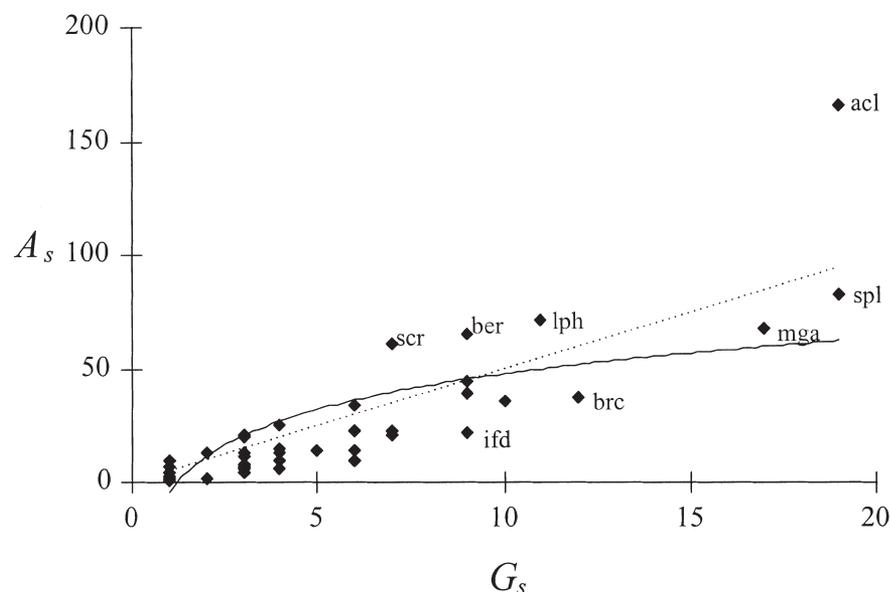


Figure 3. Number of accessions per wild potato species (*Solanum* spp.), A_s , versus number of grid cells per species, G_s (solid line, $y = 24.2 * \ln(x) - 7.3$, $r^2 = 0.54$; dashed line, $y = 5.0 * x$, $r^2 = 0.76$; *acl*, *S. acaule*; *spl*, *S. sparsipilum*; *lph*, *S. leptophyces*; *mga*, *S. megistacrolobum*; *ber*, *S. berthaultii*; *scr*, *S. sucrensis*; *brc*, *S. brevicaulis*; *ifd*, *S. infundibuliforme*).

The strong species bias evident in this case is not desirable for maximizing the diversity of a genebank collection, but this bias does not have direct implications for geographical representativeness. *S. acaule* is widespread, and the high number of accessions from this species might be a reflection of this.

Species-Area Bias

The number of accessions per species increased with increasing species distribution area (Fig. 3). The number of accessions per species (A_s) increased disproportionately with G_s , and the observed relation between G_s and A_s was significantly different from the predicted linear and logarithmic relationships (χ^2 test; $p < 0.01$). *Solanum acaule* was strongly oversampled. *Solanum sucrensis*, *S. leptophyces*, and *S. berthaultii* also were oversampled. Not all species that were collected many times are necessarily oversampled according to this criterion. The number of accessions from *S. sparsipilum* and *S. megistacrolobum*, for example, were about what would be expected on average given the area they occupy, despite their second and fourth rank on the abundance curve (Fig. 2). Other species, such as *S. brevicaulis* and *S. infundibuliforme*, were undersampled.

Hotspot Bias

The Bolivian wild potato collection is the result of 18 collecting expeditions. The number of accessions per cell increased with the number of expeditions to the cell. A linear approximation of this relationship is $A_G = 5.1 * E_G$ ($r^2 = 0.45$, $p < 0.01$). This relationship does not necessarily imply the presence of a hotspot bias; it may only reflect that the cells in which wild potatoes are

more diverse have been visited more often. This is supported by the fact that the number of species collected within a grid cell, S_G , also increased with the number of expeditions to the cell: $S_G = 0.96 * E_G$ ($r^2 = 0.13$, $p = 0.04$). This relationship is rather weak, however. Moreover, the ratio of S_G and A_G —the number of accessions per species in a cell—also went up with the number of expeditions to the cell: $A_G/S_G = 1.4 + 0.4 * E_G$ ($r^2 = 0.52$, $p < 0.01$) (Fig. 4). This suggests the presence of a hotspot bias. In terms of diversity, there has been excessive sampling, relatively speaking, in areas with high species diversity, unless one assumes that intraspecific diversity is higher in areas with more species.

Infrastructure Bias

Most accessions have been collected near main roads that connect major Bolivian cities and surrounding areas (Fig. 5). Sixty percent of all accessions were collected within 2 km of the nearest road, and 78% were collected within 4 km (Fig. 6). This is in sharp contrast with the distance from random points to the major road in each area: 22% within 2 km and 44% within 4 km. Forty-four percent of the accessions were found within 10 km of a main road, rather than the expected 27%. Thirty-three percent of the accessions were found within 40 km of a department capital, as opposed to the expected 21%. In all three cases, the distributions differed significantly from the expectation of a random distribution (χ^2 test; $p < 0.01$).

Discussion

The species and species-area biases we found may have occurred for several reasons. Collectors may have had a

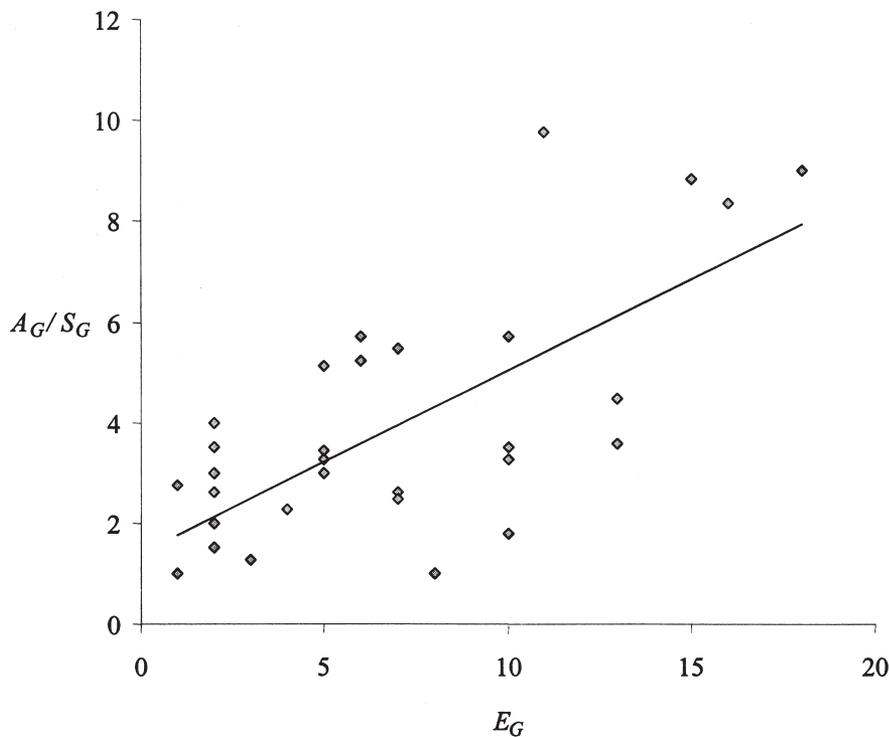


Figure 4. Ratio of the number of accessions per species (A_G/S_G) of wild potato species (*Solanum spp.*) in the Intergenebank Potato Database for Bolivia versus the number of expeditions (E_G) per 100×100 km grid cell.

preference for a particular species. Some species may be more abundant than others, irrespective of the size of their distribution area, which increases the likelihood of collection. Some species are more common in disturbed areas such as roadsides and are therefore more likely to be spotted and collected. The most frequently collected species, *S. acaule*, is an abundant inbreeder that is almost always found carrying seeds. Other species are rare, endemic, and often found without seeds; therefore, they are hard to collect. On the other hand, collection expeditions are sometimes justified by the absence of certain species in genebanks (e.g., Spooner et al. 1999), and a lot of effort may be expended to obtain samples of these species, which may be difficult to collect.

If there is no species-area bias, a stricter definition of species-area bias might be considered in which a representative collection would contain equal numbers of accessions for each species in each grid cell where that species occurs. This definition is stricter because it requires an equal allocation of accessions per grid cell within a species, as well as an allocation per species proportional to the number of grid cells a species occupies.

Species and species-area data depend on the definition of species boundaries. *S. megistacrolobum* and *S. toralapanum* (numbers 4 and 7, respectively, in the abundance ranking in Fig. 2), can be considered related varieties (Ochoa 1990) or subspecies (Giannattasio & Spooner 1994a, 1994b). If they were treated as one entity instead of two, that one species would be the second most dominant in the collection.

The presence of a hotspot bias suggests that collectors tend to collect in areas where previous expedition reports indicate the presence of many taxa. The fact that the IPD comprises several different genebank collections may have amplified this bias because all genebank curators seek some representation of the predominant species.

One weakness in analysis of hotspot bias is that it is not known whether unsuccessful attempts were made to collect wild potatoes (white areas in Fig. 5). Locations of these attempts are not recorded in genebank databases, and there are probably few collectors who systematically include in their records observations on the absence of certain species in certain areas. It would be useful if collectors would develop a more systematic documentation system that would include absence-presence and abundance data for all sites they explore, whether germplasm was collected or not. It would be useful if route maps indicated along which stretches active exploration took place, to distinguish these stretches from those that, for example, were deemed not relevant or were passed at night. In 1970, Bennett noted that "the need to formulate basic and standard procedural patterns with the capacity to absorb, integrate, preserve and make available the data of exploration missions is now urgent." Unfortunately, this statement is still valid.

Species-area and hotspot biases are scale dependent. Final scale selection is usually a compromise between the conflicting objectives of a high number of grid cells and a high number of observations per cell. Because our main objective was to screen genebank data for use in

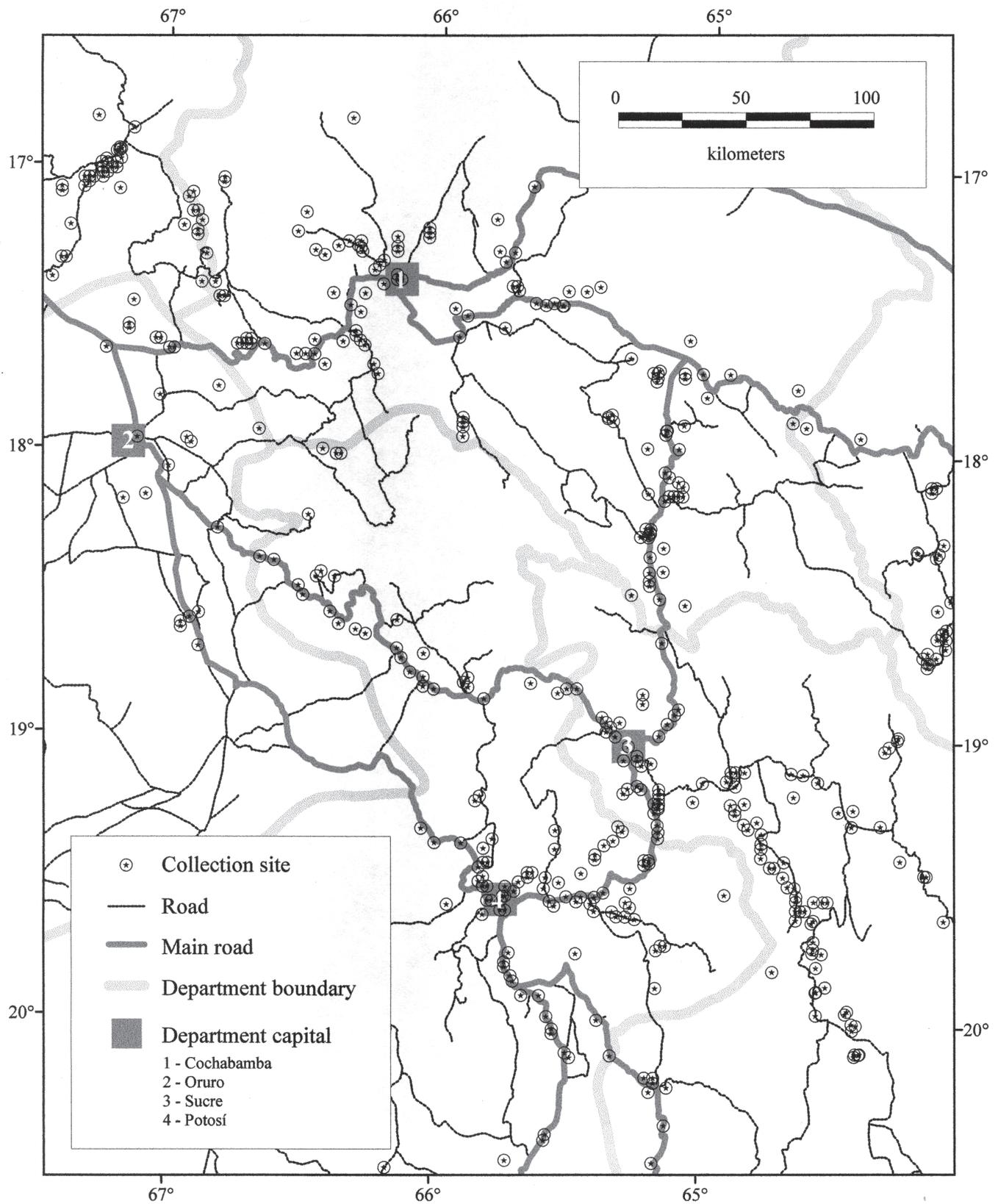


Figure 5. Location of roads, towns, and collections sites of wild potatoes (*Solanum* spp.) in central Bolivia.

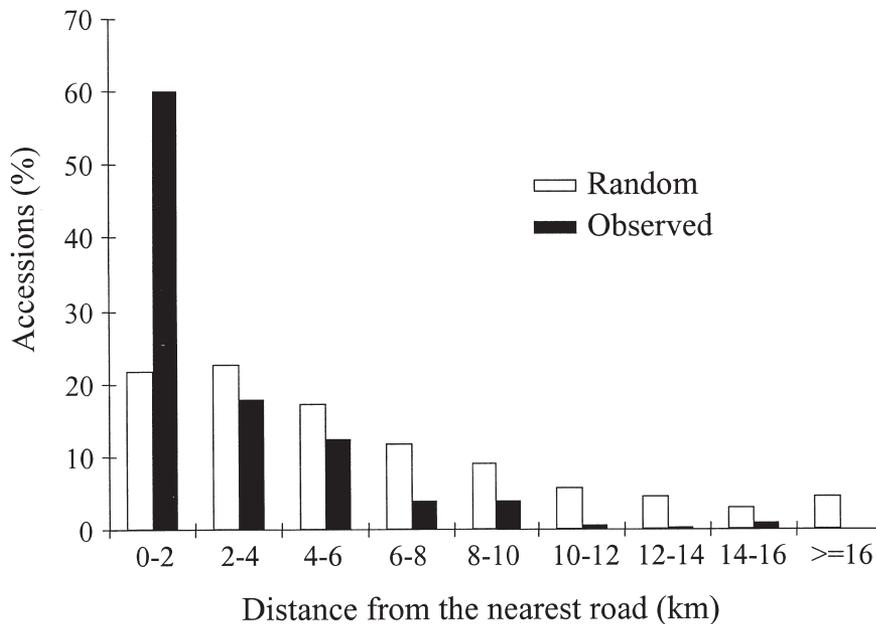


Figure 6. The distance from the nearest road in Bolivia for the accessions in the genebanks (observed) and for random points in the area where wild potatoes (*Solanum spp.*) occur.

GIS-based studies, it is reasonable to check for these two biases at the same scale that will be used in subsequent studies. In terms of maximizing diversity in the actual genebanks, at least for wild potato, the optimal collection scale has not been determined. Multiscale studies on the distribution of crop genetic diversity would be useful in this respect.

A significant factor leading to infrastructure bias is time limitations, particularly for expeditions in vast, rugged areas such as the Bolivian Andes. Maximum coverage of such areas precludes roaming far from main roads. In practice, therefore, explorers cannot sample randomly, particularly for highly endemic species.

The fact that the wild potato has been oversampled near roads and cities may render the genebank less representative, but this aspect may have a positive side. Genetic erosion is caused mainly by increased human activity such as building and agriculture, which often occur near roads and cities. Therefore, it may be that the infrastructure bias has inadvertently led to a collection bias toward genepools that are the most endangered. The opposite might also be true: infrastructure bias may have favored collection of common weedy species that tend to grow near roads. For these species, an apparent infrastructure bias would be expected even if collection was done randomly throughout an area. Other species may be reduced near roads and cities because of land-use patterns such as grazing. For these species, an infrastructure bias would lead to undercollection and/or collection of only individual plants growing on the periphery of the species' ecological niche.

Because the primary objective of genebanks is to conserve diversity, sampling should be carried out accord-

ingly. By standards such as those suggested by Brown and Marshall (1995), all species in the collection are underrepresented in the absolute sense. Many wild potato species are endemic, rare, and hard to locate. For these species, theoretical sampling schemes do not have much practical relevance.

For studies of the spatial distribution of crop diversity, other sampling approaches, such as the grid-based data collection system analyzed by Prendergast et al. (1993), would be useful. It is unlikely, however, that this type of data could be obtained for many areas or species, so researchers will have to rely on data from genebanks. Checking for the four types of biases we have described can be a useful point of departure for studies based on genebank data. Depending on the objectives of the study, it may be necessary to adjust the methods and scales used. For example, if an infrastructure bias is present and could potentially skew data, the subsequent study may have to be carried out at a scale that minimizes this bias (generally a low-resolution scale). In the case of a species-area bias, one might consider eliminating some data to remove the bias.

In formal comparisons of the observed allocation of accessions to the predicted allocation, assuming no biases, the following should be noted. Because the information on the collection accessions is, in effect, a complete survey, statistical tests are not needed to determine the presence of the four types of biases we have described. Therefore, the collection can be examined to determine whether the accessions are allocated in the proportions expected in the absence of a bias. A decision on how much deviation from this expectation is needed before a bias may be considered important be-

comes a matter of judgement based on how the information will be used. The statistical tests we have used to compare observed allocations to predicted allocations in the absence of bias are indicators of inherent trends toward bias. That is, assuming that past and future processes for adding accessions to collections are the same, these tests can help indicate whether or not disproportionate allocations are likely to be maintained.

We have defined four types of biases that may be present in genebank collections—species, species-area, hotspot, and infrastructure—and have evaluated them for a collection of wild potatoes from Bolivia. Because all four types of bias were present to some extent, the genebank data may not be an adequate representation of the existing diversity or the actual geographic spread of wild potato species in Bolivia. Collecting expeditions did not use a spatially unbiased statistical sampling procedure to collect a representative sample of the wild potatoes in Bolivia. Instead, it seems that collectors often went to sites where they expected to find a particular species, perhaps based on the results of former expeditions, and that they tended to collect samples near major roads.

Because genebanks are mainly compiled to achieve maximum diversity and usefulness (e.g., crossability, presence of resistance genes), spatial biases may be acceptable. The methods and findings we have presented may help prevent redundancy and omissions in genebanks. Future collections could focus on underrepresented species and/or on those from underrepresented areas, depending on whether the genetic diversity (in useful traits) was a function of species, region, or both. This would increase the probability of conserving a more balanced collection of potentially useful diversity.

Our findings also have important implications for the use of genebank data in GIS-based exploration and conservation studies (e.g., Jones et al. 1997). An underlying assumption in such studies may be that data from collections are a representative random sample of the existing diversity of a species. We have shown that this assumption may not always be justified. Although we have analyzed data from only one genebank, the factors considered in this study may be relevant in other settings. For example, McAllister et al. (1994) mention a sampling-intensity bias that complicates the species-density mapping of coral reef fish. Therefore, analyses of spatial representativeness may be important for any biological collection being used in GIS-based studies.

Acknowledgments

We thank A. Salas and J. de la Cruz for help in improving the coordinates of the accessions; A. Hollister for editorial suggestions; and D. Spooner, L. Guarino, M. Hermann, and the anonymous reviewers for their critical review of an earlier version of this paper.

Literature Cited

- Allard, R. W. 1970. Population structure and sampling methods. Pages 97–107 in O. H. Frankel and E. Bennett, editors. Genetic resources in plants: their exploration and conservation. F.A. Davis Company, Philadelphia.
- Bennett, E. 1970. Tactics of plant exploration. Pages 157–179 in O. H. Frankel and E. Bennett, editors. Genetic resources in plants: their exploration and conservation. F.A. Davis Company, Philadelphia.
- Brown, A. H. D., and D. R. Marshall. 1995. A basic sampling strategy: theory and practice. Pages 75–91 in L. Guarino, V. Ramanatha Rao, and R. Reid, editors. Collecting plant genetic diversity. Technical guidelines. CAB International, Wallingford, United Kingdom.
- Chapman, C. D. G. 1984. On the size of a genebank and the genetic variation it contains. Pages 102–119 in J. H. W. Holden and J. T. Williams, editors. Crop genetic resources: conservation and evaluation. Allen & Unwin, London.
- Chapman, C. D. G. 1989. Collection strategies for the wild relatives of field crops. Pages 263–279 in A. H. D. Brown, O. H. Frankel, D. R. Marshall, and J. T. Williams, editors. The use of plant genetic resources. Cambridge University Press, Cambridge, United Kingdom.
- Cochran, W. G. 1977. Sampling techniques. Wiley, Chichester, United Kingdom.
- Engels, J. M. M., R. K. Arora, and L. Guarino. 1995. An introduction to plant germ plasm exploration and collecting: planning, methods and procedures, follow-up. Pages 31–63 in L. Guarino, V. Ramanatha Rao, and R. Reid, editors. Collecting plant genetic diversity. Technical guidelines. CAB International, Wallingford, United Kingdom.
- Frankel, O. H., A. H. D. Brown, and J. J. Burdon. 1995. The conservation of plant biodiversity. Cambridge University Press, Cambridge, United Kingdom.
- Giannattasio, R. B., and D. M. Spooner. 1994a. A reexamination of species boundaries between *Solanum megistacrolobum* and *S. toralapanum* (*Solanum* sect. Petota, series Megistacroloba): morphological data. Systematic Botany 19:89–105.
- Giannattasio, R. B., and D. M. Spooner. 1994b. A reexamination of species boundaries and hypotheses of hybridization concerning *Solanum megistacrolobum* and *S. toralapanum* (*Solanum* sect. Petota, series Megistacroloba): molecular data. Systematic Botany 19:106–115.
- Guarino, L. 1995. Mapping the ecogeographic distribution of biodiversity. Pages 287–328 in L. Guarino, V. Ramanatha Rao, and R. Reid, editors. Collecting plant genetic diversity. Technical guidelines. CAB International, Wallingford, United Kingdom.
- Guarino, L., N. Maxted, and M. Sawkins. 1999. Analysis of geo-referenced data and the conservation and use of plant genetic resources. Pages 1–23 in S. L. Greene and L. Guarino, editors. Linking genetic resources and geography: emerging strategies for conserving and using crop biodiversity. Special publication. American Society for Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, Wisconsin.
- Hawkes, J. G. 1990. The potato: evolution, biodiversity and genetic resources. Belhaven Press, London.
- Hawkes, J. G., and J. P. Hjerting. 1989. The potatoes of Bolivia, their breeding value and evolutionary relationships. Clarendon Press, Oxford, United Kingdom.
- Hayek, L. A. C., and M. A. Buzas. 1997. Surveying natural populations. Columbia University Press, New York.
- Hermann, M. 1988. Progress report of the IBPGR research project on the genetic resources of Andean tuber crops. AGP-IBPGR report 89/2. International Board for Plant Genetic Resources, Rome.
- Hijmans, R. J., M. Schreuder, J. De la Cruz, and L. Guarino. 1999. Using GIS to check coordinates of genebank accessions. Genetic Resources and Crop Evolution 46:291–296.
- Jones, P. G., S. E. Beebe, J. Tohme, and N. W. Galwey. 1997. The use of geographical information systems in biodiversity exploration and conservation. Biodiversity and Conservation 6:947–958.

- Magurran, A. E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey.
- McAllister, D. E., F. W. Schueler, C. M. Roberts, and J. P. Hawkins. 1994. Mapping and GIS analysis of the global distribution of coral reef fishes on an equal area-grid. Pages 155-175 in R. I. Miller, editor. Mapping the diversity of nature. Chapman & Hall, London.
- Nabhan, G. P. 1990. Wild Phaseolus ecogeography in the Sierra Madre Occidental, Mexico: aerographic techniques for targeting and conserving species diversity. Systematic and ecogeographic studies on crop gene pools 5. International Board for Plant Genetic Resources, Rome.
- Nevo, E., 1998. Molecular evolution and ecological stress at global, regional and local scales: the Israeli perspective. The Journal of Experimental Zoology **282**:95-119.
- Ochoa, C. M. 1990. The potatoes of South America: Bolivia. Cambridge University Press, Cambridge, United Kingdom.
- Øllgaard, B. 1995. Diversity of Huperzia (Lycopodiaceae) in Neotropical montane forests. Pages 349-358 in S. P. Churchill, H. Balslev, E. Forero, and J. L. Luteyn, editors. Biodiversity and conservation of Neotropical montane forests. The New York Botanical Garden, New York.
- Prendergast, J. R., R. M. Quinn, J. H. Lawton, B. C. Eversham, and D. W. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature **365**:335-337.
- Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, United Kingdom.
- Ross, H. 1986. Potato breeding: problems and perspectives. Advances in potato breeding 13. Verlag Paul Parey, Berlin.
- Spooner, D. M., and R. Castillo. 1997. Reexamination of series relationships of South American wild potatoes (Solanaceae: *Solanum* sect. *Petota*): evidence from chloroplast DNA restriction site variation. American Journal of Botany **84**:671-685.
- Spooner, D. M., R. G. van den Berg, W. Garcia, and M. L. Ugarte. 1994. Bolivia potato germplasm collecting expeditions 1993, 1994: taxonomy and new germplasm resources. Euphytica **79**:137-148.
- Spooner, D. M., A. Salas, Z. Huamán, and R. J. Hijmans. 1999. Potato germplasm collecting expedition in Southern Peru (Departments of Apurímac, Arequipa, Cusco, Moquegua, Puno, Tacna) in 1998: taxonomy and new genetic resources. American Journal of Potato Research **76**:103-119.
- Van den Berg, R. G., J. T. Miller, M. L. Ugarte, J. P. Kardolus, J. Villand, J. Nienhuis, and D. M. Spooner. 1998. Collapse of morphological species in the wild potato *Solanum brevicaulle* complex (Solanaceae: sect. *Petota*). American Journal of Botany **85**:92-109.
- Von Bothmer, R., and O. Seberg. 1995. Strategies for the collecting of wild species. Pages 93-111 in L. Guarino, V. Ramanatha Rao, and R. Reid, editors. Collecting plant genetic diversity, technical guidelines. CAB International, Wallingford, United Kingdom.

