Biogeography of Wild *Arachis*: Assessing Conservation Status and Setting Future Priorities

Andy Jarvis,* Morag E. Ferguson, David E. Williams, Luigi Guarino, Peter G. Jones, H. Tom Stalker, Jose F. M. Valls, Roy N. Pittman, Charles E. Simpson, and Paula Bramel

ABSTRACT

The conservation status of wild *Arachis* spp. is not well characterized for its maintenance and possible future exploitation for the improvement of cultivated peanut, *Arachis hypogaea* L. Our objectives were to use 2175 georeferenced observations of wild peanut (*Arachis* spp.) to assess the conservation status of the genus and to prioritize biologically and geographically future conservation actions. Species distribution predictions were made on the basis of 36 climate variables, and these data were synthesized with land-use data to map the potential distribution of each species, and hence the species richness of the whole genus, excluding *A. hypogea*. hotspots of species richness were found in Mato Grosso around Cuiabá and Campo Grande in Brazil and around the Serra Geral de Goias, northeast of Brasilia. The current state of in situ conservation areas poorly represents wild peanut, with only 48 of the 2175 observations from National Parks. Several species were identified as being under threat of extinction. These included *A. archeri* Krapov. & W.C. Gregory, *A. setinervosa* Krapov. & W.C. Gregory, *A. marginata* Gardner, *A. hatschbachii* Krapov. & W.C. Gregory, *A. appressipila* Krapov. & W.C. Gregory, *A. villosa* Benth., *A. cryptopotamica* Krapov. & W.C. Gregory, *A. helodes* Martius ex Krapov. & Rigoni, *A. magna* W.C. Gregory & C.E. Simpson, and *A. gracilis* Krapov. & W.C. Gregory (identification based on highly restricted ranges and land-use pressures); and *A. ipaënsis* Krapov. & W.C. Gregory, *A. cruziana* Krapov., W.C. Gregory & C.E. Simpson, *A. williamsii* Krapov. & W.C. Gregory, *A. martii* Handro, *A. pietrarellii* Krapov. & W.C. Gregory, *A. vallsii* Krapov. & W.C. Gregory, and *A. monticola* Krapov. & Rigoni (identification based on insufficient observations and land-use pressures). It is suggested that ex situ conservation efforts should focus on the area around Pedro Gomes (300 km southeast of Cuiabá), 170 km south along the planned road from Cuiabá to Corumbá, and around San José de Chiquitos in Bolivia, where some of the species adapted to lower temperatures may be found.

IT IS GENERALLY AGREED that a rapid loss of plant diversity is occurring: ecosystem, species, gene, and allelic diversities are being lost forever, and the accelerating processes of habitat destruction and genetic erosion show no sign of abating. For example, the World Conservation Union's (IUCN's) Red List of Threatened Plants (Walter and Gillett, 1998) suggests that 34 000 plant species are threatened globally, equivalent to some 12.5% of the estimated world flora. Other estimates suggest that 25 to 35% of plant genetic diversity could be lost in the next 20 yr. Those taxa that include crop species and their wild relatives (crop gene pools) are of particular concern from a conservation perspective. The economic and social consequences of such an irredeemable loss of plant diversity, combined with rapid human population growth, could be potentially disastrous. The conservation of plant diversity, particularly of those species essential for human nutrition and crop improvement, is of critical importance. One of the most pressing challenges facing biologists today is the description of biological diversity at the ecosystem, species, and genetic level.

The accurate assessment of diversity is important to help reduce its loss. Some geographic areas show greater taxonomic and/or genetic diversity for a given gene pool than do others. Because funds for conservation are limited, the accurate spatial mapping of diversity is essential to prioritize conservation interventions. The task of measuring diversity at a location presents many difficulties, and the subsequent extrapolation from areas that are studied to other, less well-studied, regions is a problem central to biodiversity research (Colwell and Coddington, 1994). Conservationists therefore need methods for the rapid identification of action priorities, both geographically and in biological importance.

Conservation interventions may take the form of ex situ germplasm collections or the establishment of in situ protected areas. A germplasm collection in a gene bank aims to contain the maximum amount of genetic variation to respond to current and anticipated uses (Allard, 1970; Brown and Marshall, 1995). Hijmans et al. (2000) analyzed wild gene bank collections for bias in their geographic representativeness and detected strong over collecting along roads and within areas previously identified as hotspots for the gene pool. Herbarium collections focus on diversity at the species level, with a strong taxonomic bias reflecting the specialization of botanists. These biases must be acknowledged in any analysis of such point data.

Anderson et al. (2002) state that shaded outline maps of species ranging between and beyond known localities are likely to overestimate species distribution, while dot maps of known localities portray species distribution

---

A. Jarvis, D.E. Williams, and L. Guarino, International Plant Genetic Resources Institute (IPGRI), AA 6713, Cali, Colombia; M.E. Ferguson and P. Bramel, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Andhra Pradesh, India; P.G. Jones, and A. Jarvis, Centro Internacional de Agricultura Tropical (CIAT), AA6713, Cali, Colombia; H.T. Stalker, North Carolina State University, Raleigh, NC, USA; J.F.M. Valls, EMBRAPA Genetic Resources and Biotechnology, Brasilia, Brazil, CNPq Fellowship; R.N. Pittman, USDA, Agricultural Research Service, Griffin, GA, USA; C.E. Simpson, Texas Agric. Exp. Station, Texas A&M Univ., Stephenville, TX, USA. Received 25th April 2002. *Corresponding author (a.jarvis@cgiar.org).

Published in Crop Sci. 43:1100–1108 (2003).

Abbreviations: AVHRR, Advanced Very High Resolution Radiometer; CIAT, Centro Internacional de Agricultura Tropical; ESRI, Environmental Systems Research Institute; FAO, Food and Agriculture Organisation; GIS, Geographic Information System; IUCN, World Conservation Union; NOAA, National Oceanographic and Atmospheric Administration; PCA, principal components analysis.

conservatively. Geographic bias in collecting efforts creates further error in approximating species range. Species distribution modeling presents a means of extrapolating species range from point localities to a wider region while minimizing the risk of over- or underestimation (Franklin, 1995). Guisan and Zimmermann (2000) discuss some of the applications of species distribution modeling, and the various modeling algorithms that have been applied to the problem. Many of these methods use climatic variables as the principal drivers of geographic distribution (Walker and Cocks, 1991; Franklin, 1995; Guisan and Zimmermann, 2000). Jones et al. (1997) used the computer program, FloraMap, to predict the geographic distribution of wild bean (*Phaseolus vulgaris* L.) on the basis of the distribution of germplasm and herbarium specimens. The results correctly predicted areas where wild bean had not been collected, but was reported to occur in the literature. Segura et al. (1999) used the same software to map the geographic distribution of five species of *Passiflora*, and successfully guided germplasm collecting in Ecuador. David Williams (personal communication, 2002) used a similar process to collect germplasm of *Capsicum flexuosum* Sendtn. in Paraguay, finding six new populations having made a priori predictions of potential species range (Guarino et al., 2001).

Evaluations of species distribution models typically use presence–absence of data to test how well the prediction fits with reality (Fielding and Bell, 1997). Manel et al. (2001) conclude that Cohen's kappa provides the most appropriate statistical evaluation. This form of evaluation is problematic in germplasm collections because only species presence is available, and it is often impossible to complement this with confirmed points of species absence.

Arachis hypogaea is the most widely cultivated grain legume in the world and is one of the five most important oilseeds. Total world production in 2001 was estimated at 34 395 951 Gg (FAO, 2001). Peanut has important nutritional qualities, containing approximately 50% high quality unsaturated fats and 30% digestible protein. The center of origin of the cultivated peanut is thought to lie in northern Argentina and southern Bolivia (Stalker and Simpson, 1995), but various questions remain unanswered regarding its domestication and evolutionary history. Recent evidence indicates northwest Peru may be another possible site for the origin of the cultivated peanut (Simpson et al., 2002).

Some wild *Arachis* species have proved useful in peanut breeding. A recent example of this is a new cultivar released by Simpson and Starr (2001), which incorporates germplasm from three wild relatives (*A. cardenasii*, *A. diogoi*, and *A. batizocoi*) to reduce infestation of root-knot nematodes by >90% over nonresistant cultivars. It is important to note that there are just 17 conserved accessions of *A. cardenasii*, five of *A. diogoi*, and 12 of *A. batizocoi*. This is a clear demonstration of the importance of conserving wild relatives to respond to the ever-changing attacks of pests and diseases and to the needs of farmers and consumers. It is important that germplasm collections are sufficiently complete to anticipate these needs.

Despite over 60 yr of systematic effort to collect germplasm of wild *Arachis* species throughout the genus' natural range (Krapovickas and Gregory, 1994), a significant amount of the genetic diversity present in the peanut's secondary and tertiary gene pools remains unrepresented in any gene bank. Outstanding gaps include additional accessions of several species for which only a very few specimens exist. Ironically, many of these underrepresented wild species are among those believed to have been most closely involved in the origin and domestication of the cultivated peanut and are therefore of primary importance from a plant breeding perspective. Other significant gaps include germplasm from a handful of entirely new, but still undiscovered, *Arachis* species that are believed to survive in some of the more remote unexplored areas of western Brazil, southeastern Bolivia, and northwestern Paraguay (Williams, 2001).

Our objectives were to use 2175 georeferenced observations of wild peanut (68 species) to assess the conservation status of the genus, and to prioritize biologically and geographically future conservation actions.

MATERIALS AND METHODS

The data presented in this paper were derived from the "Catalog of *Arachis* Germplasm Collections" compiled by Stalker et al. (2000), available at http://www.icrisat.org/text/research/grep/homepage/groundnut/arachis/start.htm; verified 6 December 2002. The data are a compilation of collections of 2175 unique observations of wild *Arachis* germplasm accessions, herbarium specimens, and citations from Krapovickas and Gregory (1994).

Basic statistics of the distribution of these point observations were calculated to assess the geographic biases in the data. The distance of each point from the nearest road was calculated using the Digital Chart of the World roads coverage for Latin America (ESRI, 1992). The average distance from the nearest road for a set of 20 000 random points within the study region was taken to provide the control.

This paper uses three methods to identify conservation priorities, each tackling the problems associated with point sampling. First, species distribution mapping and climatic characterization was made and combined to give a map of potential species richness, identifying regions not visited, but with a high potential for finding various species of wild *Arachis*. Second, an analysis of agricultural land use in the region was made to examine where genetic erosion has already taken place. Finally, complementarity analysis (Rebelo, 1994) was used to identify the fewest number of protected areas needed to effectively conserve all 68 species.

Species Distributions

A computer program called FloraMap (Jones and Gladkov, 1999) was used to develop climatic models for predicting the diversity of *Arachis* spp. in the study area. FloraMap was developed at the International Center for Tropical Agriculture (CIAT, the Spanish acronym) for predicting the distribution of organisms in the wild when little or nothing is known of the physiology of the species involved. It is assumed that the climate at the points of observation and/or collection of a species is representative of the environmental range of the organism. The climate at these points is used as a calibration set to compute a climate probability model.

FloraMap uses climatic data from a 10-min grid (corresponding to 18 by 18 km at the equator) derived from observations from over 10 000 meteorological stations in Latin America. A simple interpolation algorithm based on the inverse square of the distance between stations and the interpolated point is used. For each interpolated pixel, the five nearest stations are used in the inverse distance equation. The climatic variables included are the monthly averages for temperature, rainfall, and diurnal temperature range. Mean temperature is standardized with elevation by means of the NOAA TGP-006 (NOAA, 1984) digital elevation model and a lapse rate model (Jones, 1991). Rainfall and diurnal temperature range remain independent of elevation. A 12-point Fourier transform is applied to each variable to adjust for geographic differences in the timing of major seasons. For further information the reader is referred to Jones et al. (1997, 2002).

For each accession, the 36 climate variables (comprised of 12 monthly means for temperature, rainfall, and diurnal temperature range) are extracted for the pixel in which the accession is located, and a principal components analysis (PCA) is applied to identify a smaller number of variables that account for the bulk of the variance in climates among the accession locations. The PCA is performed on the variance-covariance matrix since the Fourier analysis has transformed the variables to comparable scales. A multivariate-Normal distribution is fitted to the principal component scores so that a probability of belonging to the distribution can then be calculated for all pixels. The result is a probability surface for all of Latin America. It should be noted that this merely maps the potential climatic envelope where an organism could exist, and does not account for factors such as dispersal mechanism.

FloraMap was used to map a probability distribution for each of the 68 wild species in genus *Arachis* across a geographical range spanning all of central South America. Seventeen species with fewer than 10 accessions were omitted from the analysis. These were *A. brevipetiolata, A. chiquitana, A. cruziana, A. giacomettii, A. herzogii, A. ipaënsis, A. martii, A. monticola, A. microsperma, A. pietrarellii, A. praecox, A. rigonii, A. trinitensis, A. valida, A. vallsii, A. villosulicarpa,* and *A. williamsii*. Some of these species have been identified as possible progenitor species of the cultigen (Kochert et al., 1991), underlining the need for further collecting and conservation. For each species, displays the number of accessions and number of components used in the PCA in FloraMap, and the percentage variance that was explained.

While the climatic potential for a species may be geographically very large (e.g., Cuba is climatically suitable for many of the species), in many cases the actual distribution is much more limited. This is likely to be predominantly a result of the geocarpic habit of the wild peanut, reducing migration rates to no more than 1 m per year, given no fluvial transport of seeds (Gregory et al., 1973) or human interference. Other factors, such as historical environmental and anthropogenic change, may be responsible for confining a species distribution to a smaller range than its climatic potential. For these reasons, the climate-based potential distribution must be limited to a feasible area. Each distribution map was therefore limited to a 300-km buffer around the existing observations of the species. This value was chosen on the basis of an analysis of the geographic gaps in the collections and of the system of road access in central South America. Areas in the Bolivian and Paraguayan Chaco identified as particularly undercollected regions (Williams, 2001) are sufficiently lacking in infrastructure that areas accessible for collection lie as much as 300 km apart. This is an indication of the inaccuracy that the existing collections might represent in defining the species distributions.

Additionally, the probability distribution was subjected to an analysis of land cover to identify areas where wild habitats have already been converted to cropland. A dataset of agricultural extent was used that was derived from Advanced Very High Resolution Radiometer (AVHRR) satellite data with a resolution of 1 km (Wood et al., 2000). The dataset was reclassified into two variables, wild habitat and agriculture, where agricultural land cover was defined as having 30% or greater cropland cover. These individual species distribution maps were then combined to give a map of species diversity. If the probability of finding a species in an individual grid square was 0.5 or greater, then the species was assumed to be present.

Complementarity Analysis

In the database, 48 *Arachis* observations occur within currently recognized protected areas, but these only account for nine of the 68 wild species in the genus (Ferguson, personal communication, 2002). To determine optimal locations for in situ reserves to conserve maximum species diversity, a study based on species complementarity was undertaken using DIVA-GIS software (http://www.cipotato.org/gis/; verified 6 December 2002; Hijmans et al., 2001, 2002). The species complementarity procedure is based on the algorithm described by Rebelo (1994) and Rebelo and Sigfried (1992). The aim is to identify grid cells with a defined size, which complement each other in terms of species composition, although any biological characteristic may be used whether taxonomic, morphological, or genetic. The process is iterative, whereby the first cell is the most species rich. The second iteration locates a grid cell that is richest in species not already represented in the first iteration. This iterative process continues until all species have been represented. We computed the minimum number of grid cells needed to capture all 68 wild *Arachis* species. The grid cell size was defined as 50 by 50 km.

RESULTS

Basic statistics on the distribution of observations indicate a strong bias in collecting along roads. The average distance from each observation to the nearest road was found to be 3.31 km, while the average distance for a set of random points in the study region was 9.92 km. This is more exaggerated in some areas than in others, depending on the density of the road network and the intensity of collecting. This provides a strong case for the use of spatially extrapolative modeling to fill in the geographic gaps in collecting.

Wild peanut species potentially cover an area of nearly 5 000 000 km^2, with 364 000 km^2 harboring five or more species sympatrically (Fig. 1). These values represent the potential distributions, and do not take into account potential anthropogenic effects that may have destroyed wild peanut habitats. Forty-one percent of the potential habitat of all *Arachis* species is currently under agricultural land use (Table 1). This limits the potential climatic distribution to a more restricted range (Fig. 2).

To examine the validity of the model used, the predicted species richness is compared with the actual species richness encountered from the field collections and observations (Fig. 3). Modeling species richness in 18- by 18-km grid cells against the observed species richness

Table 1. Number of accessions used in the creation of each species probability distribution, with the associated number of principal components used in the analysis and accounting for the percentage variance. The total area of predicted distribution is shown and the percentage of this area that is now under agricultural land use (defined as >30% agricultural).

Section	Species	Accessions	Unique geographically	Principal components	Variance explained	Total area of climatic distribution	Distribution under agricultural use
		Number			%	km^2	%
Arachis	*batizocoi*	23	16	4	94.4	24 800	28.1
	benensis	8	5	1	94.2	442 000	7.5
	cardenasii	28	17	3	93.5	373 200	23.8
	correntina	42	31	3	95.6	92 000	60.6
	cruziana	4	NA	NA	NA	NA	NA
	decora	31	21	3	93.9	102 800	61.6
	diogoi	19	10	3	93.8	338 000	32.4
	duranensis	60	47	4	95.0	347 600	50.0
	glandulifera	6	6	2	93.5	85 200	30.0
	helodes	25	16	3	98.5	28 800	67.4
	herzogii	2	NA	NA	NA	NA	NA
	hoehnei	13	8	2	97.7	180 000	43.3
	ipaënsis	2	NA	NA	NA	NA	NA
	kempff-mercadoi	25	17	2	94.8	54 400	20.2
	kuhlmannii	61	41	3	94.1	186 800	39.2
	magna	13	10	3	93.9	12 800	23.5
	microsperma	5	5	1	96.4	199 200	23.5
	monticola	12	NA	NA	NA	NA	NA
	palustris	7	6	1	95.3	178 400	54.2
	praecox	3	NA	NA	NA	NA	NA
	simpsonii	13	10	2	94.8	36 000	37.7
	stenosperma	68	41	4	96.7	380 800	65.4
	trinitensis	2	NA	NA	NA	NA	NA
	valida	7	NA	NA	NA	NA	NA
	villosa	51	34	4	95.4	186 800	95.1
	williamsii	3	NA	NA	NA	NA	NA
Caulorrhizae	*pintoi*	132	85	4	94.8	588 000	67.8
	repens	34	27	4	93.9	1 330 000	71.2
Erectoides	*archeri*	39	25	4	94.3	4 400	75.3
	benthamii	46	39	4	95.6	153 600	60.1
	brevipetiolata	2	NA	NA	NA	NA	NA
	cryptopotamica	17	15	3	96.4	24 400	62.0
	douradiana	16	13	2	91.0	44 400	54.0
	gracilis	12	10	4	95.7	35 200	71.9
	hatschbachii	7	7	3	95.5	23 600	74.6
	hermannii	11	5	2	95.2	79 200	57.8
	major	57	45	3	95.5	174 000	43.3
	martii	3	NA	NA	NA	NA	NA
	oteroi	56	43	4	95.1	46 000	71.0
	paraguariensis	60	50	4	97.7	98 000	21.4
	stenophylla	11	11	2	92.1	99 600	22.1
Extranervosae	*burchellii*	91	78	5	95.6	568 400	46.7
	lutescens	68	59	3	93.9	411 200	57.9
	macedoi	31	24	3	94.7	663 200	60.4
	marginata	6	5	2	90.2	8 800	73.8
	pietrarellii	12	NA	NA	NA	NA	NA
	prostrata	94	76	5	94.9	473 200	73.3
	retusa	15	14	2	94.7	231 600	66.5
	setinervosa	6	5	2	90.9	5 200	71.8
	villosulicarpa	6	NA	NA	NA	NA	NA
Heteranthae	*dardani*	70	64	5	96.2	404 000	75.9
	giacomettii	3	NA	NA	NA	NA	NA
	pusilla	33	28	4	94.4	217 600	43.2
	sylvestris	89	71	5	94.1	728 800	51.0
Procumbentes	*appressipila*	22	14	3	97.8	21 600	57.9
	chiquitana	4	NA	NA	NA	NA	NA
	kretschmeri	14	13	3	98.2	90 400	34.9
	lignosa	12	5	2	99.7	19 600	8.1
	matiensis	41	31	3	96.1	43 200	35.5
	rigonii	3	NA	NA	NA	NA	NA
	subcoriacea	19	13	3	98.0	153 600	39.0
	vallsii	8	NA	NA	NA	NA	NA
Rhizomatosae	*burkartii*	100	81	7	94.5	162 800	93.0
	glabrata	301	241	4	95.0	639 600	54.7
	pseudovillosa	43	31	4	96.8	30 800	42.8
Triseminatae	*triseminata*	21	15	2	95.0	116 400	82.3
Trierectoides	*Guaranitica*	13	10	3	95.0	24 800	40.4
	tuberosa	17	15	2	98.5	93 200	82.1

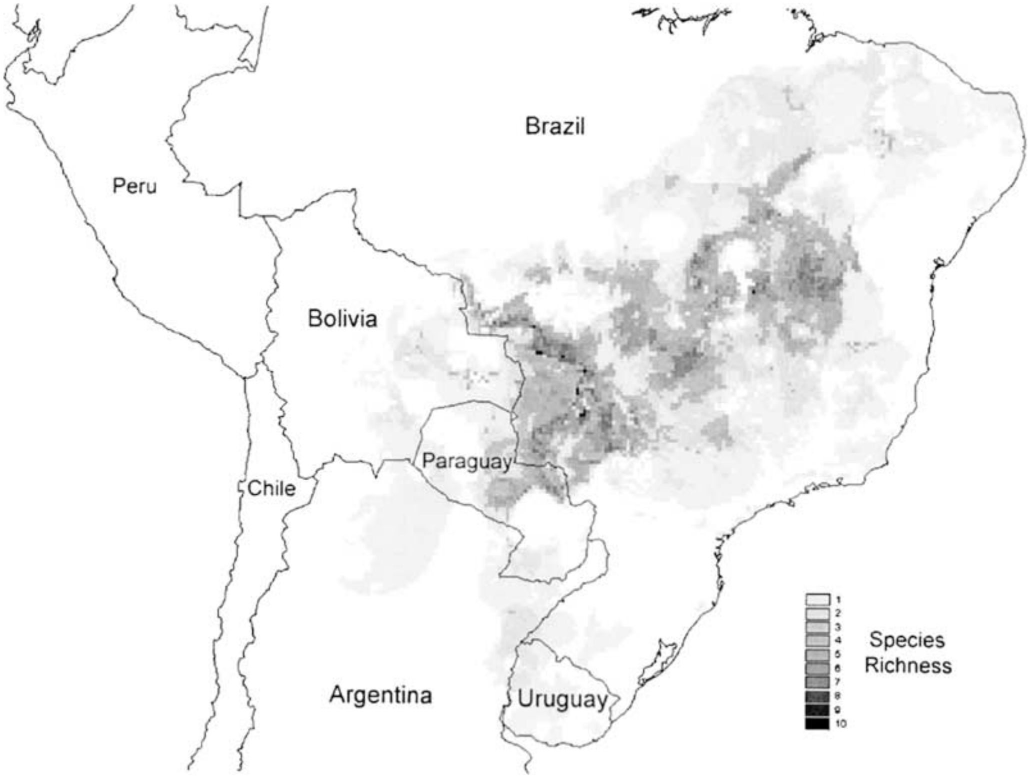

Fig. 1. Predicted distribution of species richness of *Arachis* spp. across South America based on climatic analysis and a 300-km buffer zone around known collections. Shading represents the potential number of species per 18- by 18-km grid cell.

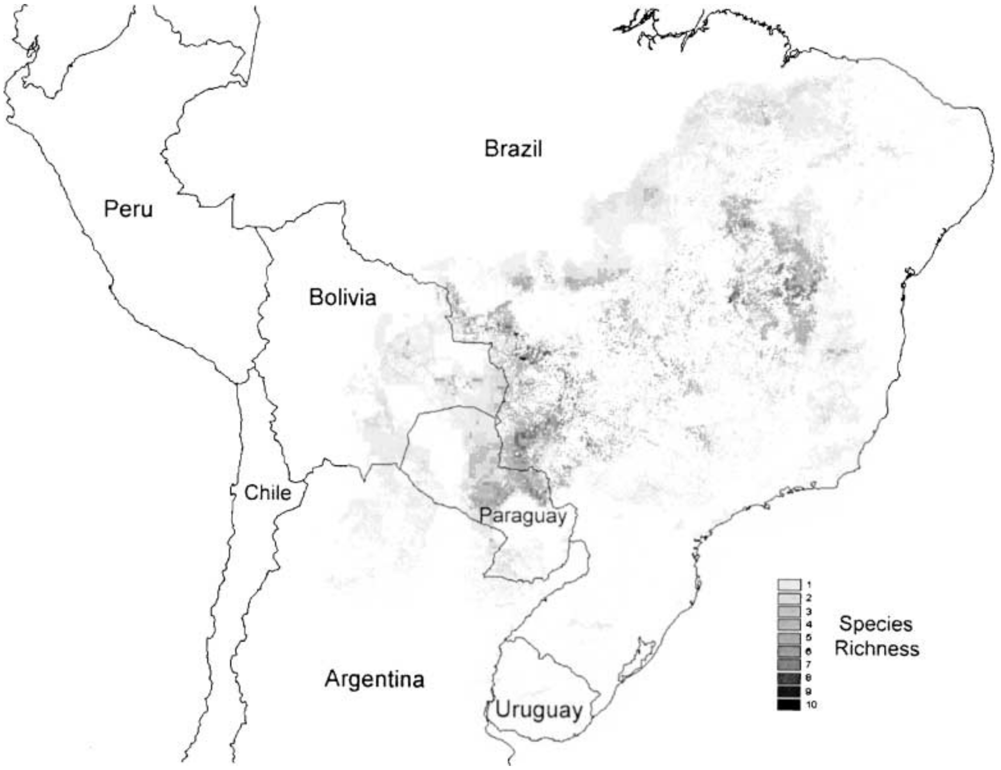

Fig. 2. Predicted distribution of *Arachis* spp. richness across South America with areas under agricultural land use excluded. Areas of 30% or greater agricultural land use are assumed not to harbor wild *Arachis* species.

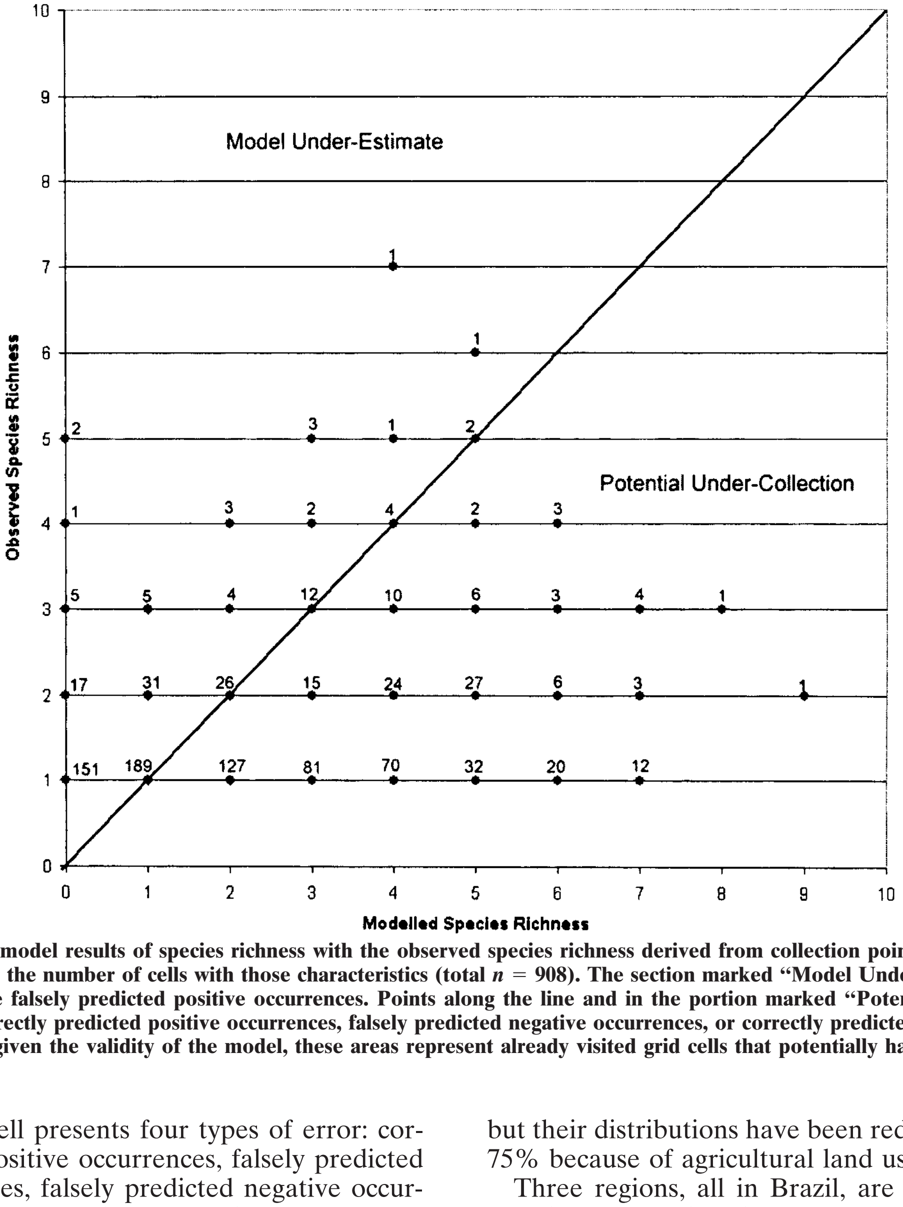

Fig. 3. Comparison of model results of species richness with the observed species richness derived from collection points. The numbers above the points represent the number of cells with those characteristics (total n = 908). The section marked "Model Under-Estimate" is the area where all points are falsely predicted positive occurrences. Points along the line and in the portion marked "Potential Under-Collection" could be any of correctly predicted positive occurrences, falsely predicted negative occurrences, or correctly predicted negative occurrences. It is assumed that, given the validity of the model, these areas represent already visited grid cells that potentially harbor more species than so far encountered.

within the same cell presents four types of error: correctly predicted positive occurrences, falsely predicted positive occurrences, falsely predicted negative occurrences, and correctly predicted negative occurrences. The latter two require absence data, which were not available for this study, as for most studies involving germplasm collections. In Fig. 3, areas where modeled species richness exceeds observed species richness (bottom right corner) indicate either undercollection, or falsely predicted negative occurrences. Areas where modeled species richness is less than the observed (top left corner) indicate cases of falsely predicted positive occurrences. Just 24% of cases fall in this category (n = 908), and 66% of these are an underestimation by only one species.

Some *Arachis* species appear to be particularly threatened by habitat loss. Those most restricted in distribution are *A. archeri*, *A. setinervosa*, *A. marginata*, *A. hatschbachii*, *A. appressipila*, *A. villosa*, *A. cryptopotamica*, *A. helodes*, *A. magna*, and *A. gracilis.* Their distribution is limited to less than 10 000 km^2 of climatically suitable wild habitat. That of *A. burkatii*, *A. triseminata*, *A. tuberosa,* and *A. dardani* remains above 10 000 km^2, but their distributions have been reduced by more than 75% because of agricultural land use.

Three regions, all in Brazil, are predicted hotspots for wild peanut diversity. These are the Serra Geral de Goias northeast of Brasilia, the region west of Campo Grande in Mato Grosso do Sul, and the region 170 km south of Cuiabá in Mato Grosso. A species richness of 10 is predicted for one area 300 km southeast of the city of Cuiabá near Pedro Gomes (Fig. 4), where the species *A. cryptopotamica*, *A. diogoi*, *A. glabrata*, *A. helodes*, *A. hoehnei*, *A. kuhlmannii*, *A. lutescens*, *A. matiensis*, *A. stenosperma*, and *A. subcoriacea* are all predicted to exist sympatrically. None of these three hotspots coincide with protected areas. Ex situ collections provide a relatively better coverage of these hotspots, although some of the predicted ones remain totally unexplored. Worthy of note is the planned road running southwest from Cuiabá toward Corumbá. Locations in this region are predicted to contain as many as eight species growing sympatrically, but show no record of any ex situ collection. The state of this road for access is unclear; its northern sector was once built only from Cuiabá south to the Cuiabá River, where it reaches the

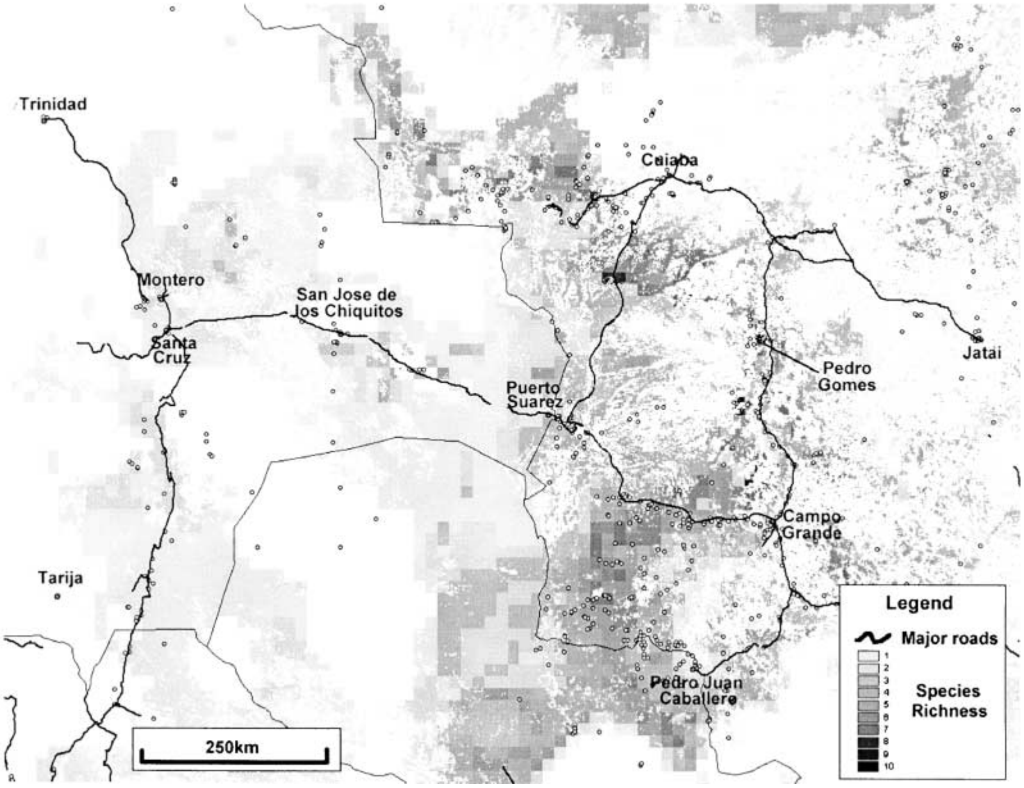

Fig. 4. Hotspots of predicted species richness for *Arachis* spp. and some of the conservation priorities in the region of Cuiabá and Campo Grande, Brazil, and in eastern Bolivia.

locality of Porto Jofre. It has since practically disappeared, with most of some 130 bridges that were originally constructed now in ruins. More recent Brazilian road maps do not mention any road from the Cuiabá River to Corumbá. Another area worthy of mention for targeting ex situ collection is the municipality of Parauna (in the state of Goias), where only three collections have been made (species *A. prostrata* and *A. glabrata*). It is predicted that as many as six different species may be found in this region, although the land in this area is predominantly agricultural. In addition to Brazil, Bolivia is highlighted as an area of interest for further collection (Fig. 4), especially on the minor road from Santa Cruz to Puerto Suarez, near the town of San José de Chiquitos in the southeast part of the country, where some five species potentially lie sympatrically.

The predictions made in this analysis are based on the data gathered from existing collections. The method attempts to fill in the climatic gaps between two climatic extremes for each species, and extrapolates this spatially using climate surfaces. Should these extremes be poorly represented in the collection data, the predicted distribution reflects this bias and may not capture the full climatic envelope to which the species may be adapted. Predictions for species that are sparsely collected, including many of the higher altitude species found on the Andean fringe in Bolivia, may have greater errors in species distributions than those that have been more exhaustively collected (such as for *A. glabrata*). This may mean that those countries where collecting activities have been less intensive (i.e., Bolivia and Paraguay) are underrepresented in terms of predicted wild species richness. It is important to note that the putative B genome progenitor species of the cultigen (*A. ipaënsis*, *A. cruziana*, and *A. williamsii*) have insufficient observations to infer their potential distribution. Of the 16 species for which data were insufficient to predict the distribution, 40 of the 76 observations now lie in areas of >30% agricultural land-use. Of special mention are *A. martii* because all three locations of previous collection are now under agricultural land use, *A. pietrarellii,* where 83% of the 12 collections are now under agricultural land use, and *A. vallsii* and *A. monticola, where* 75% of collections are now under agricultural land use).

Twenty-seven 2500-km^2 grid cells were required to include all 68 wild species, and the first ten species rich, yet complementary grid cells have been numbered to highlight the most important regions (Fig. 5). As expected, the first four grids coincide with areas of high species richness identified in Fig. 1 and 2, which indicates that each of the high diversity areas has a distinct species composition. Just five grid cells include 50% of the 68 species included in the analysis (Fig. 6).

DISCUSSION

Species distribution modeling based on climatic adaptation inferred from existing observations can be used to extrapolate from geographically biased point measurements to larger and unexplored regions. It does fail, however, in predicting the full variation within a species should the point observations poorly represent the extremes. This modeling method has proved of value in other related studies (Jones et al., 1997; Segura et al.,

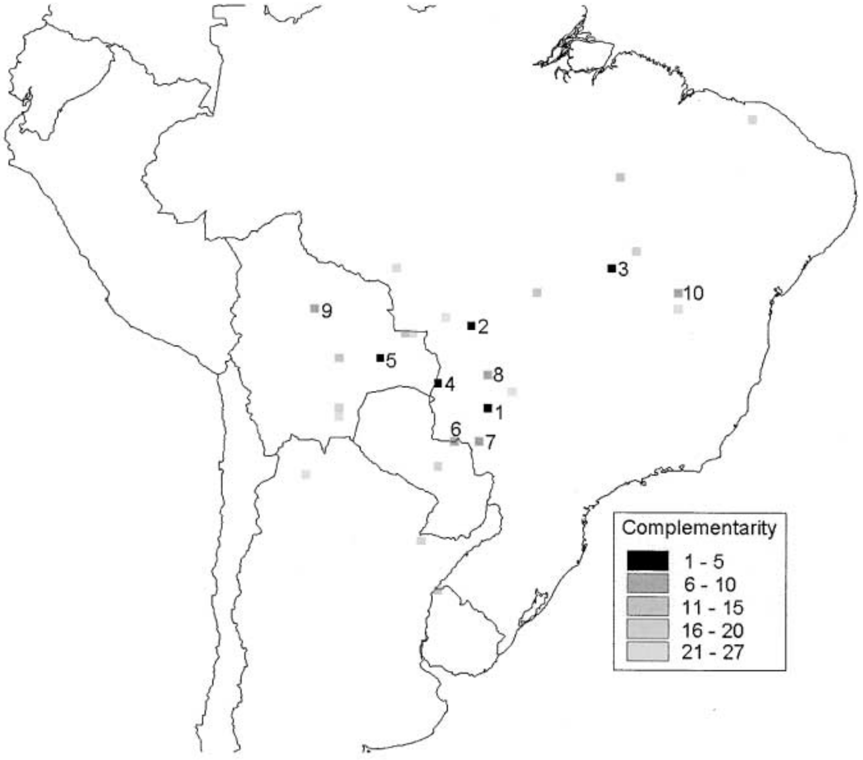

Fig. 5. Results of the complementarity analysis for prioritizing in situ conservation efforts for wild peanut across South America. Each grid is 50 by 50 km in size. The first 10 priority grids are numbered in order of importance for in situ conservation.

1999; David Williams, personal communication, 2002), but lacks a formal validation. This paper makes a preliminary comparison of predicted species richness with the observed, and finds the model performs reasonably well. However, lacking confirmed sites where the species was sought, but not found, the statistical significance of the validation remains uncertain.

Should the organism–gene pool be thoroughly collected in a given region, complementarity analyses provide a means of spatially prioritizing conservation interventions, be they in situ or ex situ. These methodologies have important implications for defining strategies for conserving gene pools, and are transferable to the intraspecific genetic level. Under rapid environmental change scenarios, any conservation action must be well focused on biologically important organisms in the geographically most vulnerable and biologically richest regions. Jarvis et al. (2002) make a regional analysis applying these methods to Bolivia, incorporating anthropogenic influences to assess the risk of *Arachis* spp. genetic erosion.

This study provides a strong case for efforts at the conservation of wild peanut in Latin America that are geographically and biologically focused. Ex situ conservation action should prioritize some of the more important species, including several of the putative B genome

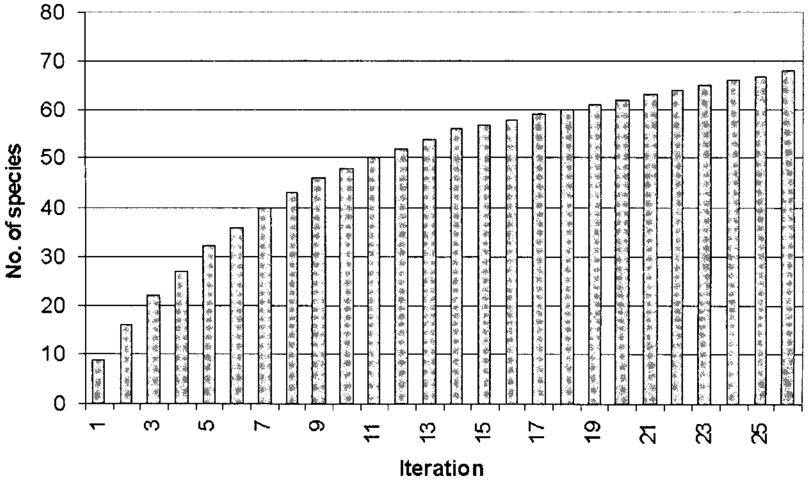

Fig. 6. Accumulation of species for each iterative addition of a 50-km grid square in the complementarity analysis. Just six of the grid cells include 50% of all 68 *Arachis* spp.

progenitors of the cultivated species, *A. williamsii*, *A. cruziana*, and *A. ipaënsis*. Also under risk of extinction are *A. martii*, *A. pietrarellii, A. vallsii*, and *A. monticola*. There are too few collections of these species to predict their distributions, thus ex situ conservation missions should focus on the remaining wild habitats in the regions where they were previously observed. Of the species with sufficient entries to make predictions of potential distribution, *A. magna* and *A. archeri* are particularly in need of further ex situ conservation. This is based on their potential importance for cultivated peanut improvement, the poor current state of collection, and the identification of potential collection gaps. Geographical areas in particular need of attention lie 40 km west of Cuiabá in Brazil, the stretch southeast out of Cuiabá, and along the minor road from Santa Cruz to Puerto Suarez around the town of San José de Chiquitos.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the Common Fund for Commodities and the World Bank under the 'Preservation of Wild Species of *Arachis*' project. The authors also thank Stanley Wood for use of agricultural extent data, German Lema for statistical guidance, and the anonymous referees for their help in improving the content of the paper.

REFERENCES

Allard, R. 1970. Population structure and sampling methods. p. 97–107 *In* Genetic resources in plants: Their exploration and conservation. F.A. Davis Company, Philadelphia, PA.

Anderson, R., M. Gomez-Laverde , and A. Peterson. 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. Global Ecol. Biogeogr. 11:131–141.

Brown, A., and D. Marshall. 1995. A basic sampling strategy: Theory and practice. p. 75–91. *In* L. Guarino et al. (ed.) Collecting plant genetic diversity. Technical guidelines. CAB International, Wallingford, UK.

Colwell, R.K., and J.A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Philos. Trans. R. Soc. Lond. B. 345:101–118.

ESRI. (Environmental Systems Research Institute). 1992. Digital chart of the world. ESRI, Redlands, CA, USA.

FAO. (Food and Agriculture Organisation). 2001. FAOSTAT Agriculture. Available at: http://apps.fao.org/page/collections?subset = agriculture; verified 6 December 2002.

Fielding, A., and J. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24:38–49.

Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geogr. 19:474–499.

Gregory, W., M. Gregory, A. Krapovickas, B.W. Smith, and J.A. Yarbrough. 1973. Structure and genetic resources of peanuts. Peanuts-culture and uses. p. 47–133. American Peanut Research and Education Association, Stillwater, OK.

Guarino, L., A. Jarvis, R.J. Hijmans, and N. Maxted. 2001. Geographic Information Systems (GIS) and the conservation and use of plant genetic resources. p. 387–404. *In* J. Engels et al. (ed.) Managing plant genetic diversity. Procs. International Conference on Science and Technology for Managing Plant Genetic Diversity in the 21st Century (SAT21). Kuala Lumpur, Malaysia. 12–16 June, 2000. CAB International, Wallingford, UK.

Guisan, A., and N. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecol. Modell. 135:147–186.

Hijmans, R., K. Garrett, Z. Huaman, D. Zhang, M. Schreuder, and M. Bonierbale. 2000. Assessing the geographic representativeness of gene bank collections: the case of Bolivian wild potatoes. Conserv. Biol. 14(6):1755–1765.

Hijmans, R.J., M. Cruz, E. Rojas, and L. Guarino. 2001. DIVA-GIS, Version 1.4. A geographic information system for the management and analysis of genetic resources data. Manual. International Potato Center, (CIP), Lima, Peru.

Hijmans, R., L. Guarino, M. Cruz, and E. Rojas. 2002. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. Plant Genet. Res. Newsl. 127:15–19.

Jarvis, A., L. Guarino, D. Williams, K. Williams, and G. Hyman. 2002. The use of GIS in the spatial analysis of wild peanut distributions and the implications for plant genetic resource conservation. Plant Genet. Res. Newsl. 131.

Jones, P.G. 1991. The CIAT climate database version 3.41, machine readable dataset. Centro Internacional de Agriculture Tropical (CIAT), Cali, Colombia.

Jones, P.G., S.E. Beebe, J. Tohme, and N.W. Galwey. 1997. The use of geographical information systems in biodiversity exploration and conservation. Biodivers. Conserv. 6:947–958.

Jones, P.G., and A. Gladkov. 1999. FloraMap: A computer tool for the distribution of plants and other organisms in the wild. CIAT, Cali, Colombia.

Jones, P., L. Guarino, and A. Jarvis. 2002. Computer tools for spatial analysis of plant genetic resources data: 2. FloraMap. Plant Genet. Res. Newsl. 130:1–6.

Kochert, G., T. Halward, W.D. Branch, and C.E. Simpson. 1991. RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor. Appl. Genet. 81:565–570.

Krapovickas, A., and W.C. Gregory. 1994. Taxonomía del género *Arachis* (Leguminosae). Bonplandia 8(1–4):1–186.

Manel, S., H. Williams, and S. Ormerod. 2001. Evaluating presence-absence models in ecology: The need to account for prevalence. J. Appl. Ecol. 38:921–931.

NOAA. (National Oceanographic and Atmospheric Administration). 1984. TGPO006 D. Computer compatible tape. Boulder, CO.

Rebelo, A.G. 1994. Iterative selection procedures: Centres of endemism and optimal placement of reserves. Strelitzia 1:231–257.

Rebelo, A.G., and W.R. Sigfried. 1992. Where should nature reserves be located in the Cape Floristic Region, South Africa? Models for the spatial configuration of a reserve network aimed at maximizing the protection of diversity. Conserv. Biol. 6:243–252.

Segura, S.D., L. Guarino, G. Coppens d'Eeckenbrugge, M. Grum, and P. Ollitrault. 1999. Mapping the distribution and regions climatically suitable for four species in *Passiflora* subgenus *Tacsonia* (*Passifloraceae*) and *P. manicata*. Poster presented at the II Simposio de Recursos Genéticos para América Latina e el Caribe–SIRGEALC, 21–26 Nov. 1999, Brasilia, Brazil.

Simpson, C., and J. Starr. 2001. Registration of'COAN' peanut. Crop Sci. 41:918.

Simpson, C.E., A. Krapovickas, and J.F.M. Valls. 2002. History of *Arachis* including evidence of *A. hypogaea* L. progenitors. Peanut Sci. 28:79–81.

Stalker, H.T., and C.E. Simpson. 1995. Germplasm resources in *Arachis.* p.14–53. *In* H.E. Pattee and H.T. Stalker (ed.) Advances in peanut science. American Peanut Research and Education Society, Inc., Stillwater, OK.

Stalker, H.T., M.E. Ferguson, J.F.M Valls, R.N. Pittman, C.E. Simpson, and P. Bramel. 2000. Catalog of *Arachis* germplasm collections. Available at: http://www.icrisat.org/text/research/grep/homepage/groundnut/arachis/start.htm; verified 6 December 2002.

Walker, P., and K. Cocks. 1991. Habitat: A procedure for modelling a disjoint environmental envelope for a plant or animal species. Global Ecol. Biogeogr. Lett. 1:108–118.

Walter, K.S., and H.J. Gillett. (ed.) 1998. 1997 IUCN red list of threatened plants. Compiled by the World Conservation Monitoring Centre (IUCN). IUCN, Gland, Switzerland.

Williams, D. 2001. New directions for collecting and conserving peanut genetic diversity. Peanut Sci. 28:135–140.

Wood, S., K. Sebastian, and S. Scherr. 2000. Pilot analysis of global ecosystems: Agroecosytems. A joint study by International Food Policy Research Institute (IFPRI) and World Resources Institute (WRI). IFPRI-WRI, Washington, WA.