American Journal of Botany 88(11): 2101–2112. 2001.

Geographic distribution of wild potato species[1]

Robert J. Hijmans[2,4] and David M. Spooner[3]

[2]International Potato Center, Apartado 1558, Lima 12, Peru; and
[3]USDA, Agricultural Research Service, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, Wisconsin 53706-1590 USA

The geographic distribution of wild potatoes (Solanaceae sect. *Petota*) was analyzed using a database of 6073 georeferenced observations. Wild potatoes occur in 16 countries, but 88% of the observations are from Argentina, Bolivia, Mexico, and Peru. Most species are rare and narrowly endemic: for 77 species the largest distance between two observations of the same species is <100 km. Peru has the highest number of species (93), followed by Bolivia (39). A grid of 50×50 km cells and a circular neighborhood with a radius of 50 km to assign points to grid cells was used to map species richness. High species richness occurs in northern Argentina, central Bolivia, central Ecuador, central Mexico, and south and north-central Peru. The highest number of species in a grid cell (22) occurs in southern Peru. To include all species at least once, 59 grid cells need to be selected (out of 1317 cells with observations). Wild potatoes occur between 38° N and 41° S, with more species in the southern hemisphere. Species richness is highest between 8° and 20° S and around 20° N. Wild potatoes typically occur between 2000 and 4000 m altitude.

Key words: GIS; geographic distribution; geographic information systems; potato; section *Petota*; *Solanum*; Solanaceae; species richness.

The wild tuber-bearing *Solanum* L. species (Solanaceae sect. *Petota* Dumort.) and outgroup relatives in *Solanum* sect. *Etuberosum* (Bukasov and Kameraz) A. Child, hereafter called wild potatoes, are relatives of the cultivated potato (Hawkes, 1990; Contreras and Spooner, 1999). The cultivated potato (*Solanum tuberosum* L. and six other cultivated species that are only grown in the Andes) is one of the world's principal food crops (Walker, Schmiediche, and Hijmans, 1999).

In addition to these seven cultivated potato species, there are 199 wild potato species (Spooner and Hijmans, in press). These wild species all grow in the Americas, from the southwestern United States to central Argentina and Chile. They have been collected intensively to contribute to potato genebanks, and they have been used in breeding programs to improve the cultivated potato (Ross, 1986; Hawkes, 1990). There is no evidence that wild potatoes are currently being domesticated or traded or that their geographic distribution is deliberately altered by man, except for the case of *S. chacoense,* a species that is grown as an ornamental plant in Lima, Peru (Ochoa, 1962). However, in some parts of the Andes there is ongoing gene flow between wild and cultivated potato species (Hawkes, 1979; Watanabe and Peloquin, 1989; Rabinowitz et al., 1990).

Despite a purported great variability in wild potatoes (Hawkes, 1990), many of the species have a general appearance similar to the cultivated potato (Spooner and Van den Berg, 1992). A typical wild potato has pinnately dissected leaves (a few species have entire leaves), corollas in various shades of purple (colors vary from white to blue to purple and pink), corollas pentagonal to rotate (some are stellate), and fruits spherical to ovoid (some are conical). This similarity has led to widely conflicting taxonomic treatments of wild potatoes. Hawkes (1990) and Spooner and Hijmans (2001) provide the most recent taxonomic overviews of the entire group.

Wild potatoes have been the object of intensive study. Spooner and Hijmans (2001) recently reviewed wild potato systematics and germplasm collection. Jansky (2000) reviewed the value of wild and cultivated potatoes in breeding for disease resistance. However, there has been no comprehensive analysis of the geographic distribution of wild potatoes, which we provide here for the first time. We used geographic information systems (GIS) to analyze a large georeferenced database of locations where wild potatoes were observed.

We computed country- and species-level statistics. For each species, we estimated the geographic area over which it occurs and mapped the number of observations and species richness, using grid cells. We determined the minimum number of grid cells needed to include all species and related species richness to latitude and altitude. Species richness is used because it is a simple, widely used, well-understood, and useful measure of taxonomic diversity (Gaston, 1996) and because it is less sensitive than diversity indices to the problems of unsystematic sampling intensities and procedures (Hijmans et al., 2000).

This study explores the use of geographic information systems (GIS) to describe the geographic distribution of wild crop relatives. This type of study can provide baseline data for further GIS analysis for exploration, conservation, and use of germplasm of wild crop relatives (Guarino et al., in press), as well as for studies of the factors that explain the geographic distribution of these species.

MATERIALS AND METHODS

Wild potato distribution data—Our sources of wild potato distribution data were the following: (1) the Inter-genebank Potato Database, which has data from seven genebanks in the USA, Peru, The Netherlands, Germany, Argentina, the UK, and Russia (in order of size of contribution) (Huamán, Hoekstra, and Bamberg, 2001). (2) Data from 16 collecting expeditions in 12 countries by D. M. Spooner and coworkers (Spooner and Hijmans, 2001). These included records that are not in genebanks because they were collected as herbarium species or were lost as living specimens after collection. (3) A database of herbarium records developed by J. G. Hawkes (Hawkes, 1997). (4) Hawkes and Hjerting (1969; Argentina, Brazil, Paraguay and Uruguay). (5) Hawkes and Hjerting (1989; Bolivia). (6) Ochoa (1990; Bolivia). (7) Ochoa (1999, Peru). (8) Spooner et al. (1998; Guatemala). (9) Spooner et al. (1999;

[1] Manuscript received 12 December 2000; revision accepted 19 April 2001.
The authors thank Mariana Cruz, Jorge de la Cruz, Alberto Salas, and Monica Túesta for helping improve the coordinate data in the wild potato database, and Luigi Guarino, Sandra Knapp, John Stares, Karl Zimmerer, and two unidentified reviewers for review.
[4] Author for reprint requests.

Peru). (10) Spooner, Hoekstra, and Vilchez (2001; Costa Rica). (11) Spooner et al. (in press; Mexico).

We used all records from sources 1–3 that included a species name and passport data. Passport data include a description of the location of origin (such as locality name), administrative units (such as departments and districts), and geographic coordinates. However, coordinate data were absent for many records. For the genebank databases, coordinates were assigned using the locality description where possible. For Hawkes' (1997) database this was only attempted for species for which we had fewer than five observations with coordinate data. Sources 5–7 were used to verify and improve the geographic coverage of our species distribution data (see below). Additional herbarium records were taken from sources 8–11.

The presence of coordinate data allowed the analysis of the database with GIS; we used ArcView-GIS (Environmental Systems Research Institute, 1999) and DIVA-GIS software (Hijmans et al., 2001). Coordinate data in genebank databases often lack precision and were checked and modified following procedures described by Hijmans et al. (1999). First, we checked for gross errors, such as accessions located in the oceans. Then, we made overlays (simultaneous spatial queries) of the collection sites and administrative boundary databases (first level subdivision for Mexico and Central America; first and second level for the United States and South America; and first, second, and third level subdivision for Peru). We compared the names of the administrative units according to the wild potato distribution database with those of the administrative boundary database. In case of discrepancies between the two databases, the coordinates were checked against the locality description and new coordinates were assigned where needed.

Dot maps of the distribution of all species were compared with published species distribution maps from the floristic sources 5–7. When general areas of occurrence were already represented on our maps, we did not include additional points, because of a possible lack of precision of many of these maps, and the risk of duplicating records. However, if it appeared that our distribution maps did not include all major areas where a species was reported to occur, we did copy additional observations for these areas. Species names follow Spooner and Hijmans (in press), who list 196 wild species in sect. *Petota* and three outgroup species in sect. *Etuberosum*. Taxonomic groups below the species level (subspecies, varieties, and forms) are all treated within their component species.

Country- and species-level distribution—The number of observations and species in the database were tabulated by country. This was done separately for rare species, here defined as species for which we had fewer than five observations. The number of observations per species was calculated and plotted. The average number of observations per species was calculated to assess intensity of collection by country, given the species richness it harbors.

Area of distribution—For each species we estimated the area over which it occurs, using two statistics: (1) Maximum distance (MaxD) between two observations of a single species was calculated as the largest distance (in meters) between all possible pairs of observations of one species. (2) We assigned a circular area (CA_r) with a radius r to each observation and calculated the total area of all circles per species. Areas where circles of a species overlap are only included once. Area is expressed as the area relative to the area of one circle, i.e., the number of circular areas covered. We decided to use a radius of 50 km (i.e., CA_{50}).

The assumption is that each point observation represents a group of plants that covers a circular area with a 50-km radius. Expressing CA_r as the number of circles instead of the absolute area makes it more easily comparable across different studies and scales (when a radius other than 50 km is chosen).

The CA_{50} statistic was plotted against the number of observations to explore differences in abundance between species. A species with a relatively high number of observations per CA_{50} would be abundant within its area of distribution, whereas a low number would indicate that a species was more scattered over the range in which it occurs.

To describe species distributions we use the terms ''endemic'' and ''rare.'' We use ''endemic'' for species that occur in relatively small areas (have a small range size) (Rabinowitz, 1981; Gaston and Williams, 1996) and ''rare'' for species that have been observed in relatively few cases.

Grid based distribution—We compared the number of observations and species using a grid with 50 × 50 km cells and summarized the results by country. We used ArcView-GIS to transform the coordinate data to the Lambert equal-area azimuthal projection, with 80° W as the central meridian and the equator as the reference latitude (Environmental Systems Research Institute, 1999). Because the origin of a grid is arbitrary but can influence the results, it may not be accurate to assign a point to one cell only, if the point is located near one or three other cells. Therefore, the data were assigned to grid cells using a circular neighborhood (Bonham-Carter, 1994; Cressie, 1991) with a radius of 50 km, using the DIVA-GIS software. All the observations within that neighborhood were assigned to its respective grid cell, and an observation can, therefore, be assigned more than once. The result is a smoother grid, which is less biased by the origin of the grid and also less sensitive to small changes (errors) in the coordinate data. When we discuss grid cells in this paper, these refer to circles with an area of $\pi r^2 = 7854$ km^2 with their center in the middle of the grid cells with an area of 2500 km^2.

To assess the distribution of species over grid cells, we plotted the number of species per grid cell against the number of grid cells. Data on plant distributions can be biased due to spatial differences in recorder (a person who takes data) effort (Rich and Woodruff, 1992; Gaston, 1996; Hijmans et al., 2000). In extreme cases, differences in number of species between areas would reflect the amount of time spent there by recorders and not reflect actual differences in distribution. We assessed the extent to which the number of observations predicts the number of species by plotting the number of species vs. the number of observations per grid cell.

Complementarity analysis—To further analyze aspects of species distribution and endemicity we identified the smallest area (number of grid cells) needed to capture all wild potato species. This type of complementarity analysis is typically used in studies for optimal reserve selection (Csuti et al., 1997). We used the algorithm described by Rebelo (1994; see also Rebelo and Siegfried, 1992) and implemented in the DIVA-GIS software that selects grid cells so as to identify the minimum set of cells that captures a maximum amount of species. The algorithm selects the cell with most species in it and then, step by step, selects cells that contain the highest number of additional (not previously included) species. In the case of cells having the same number of additional species, a random cell is selected from such cells. Selecting these complementary cells is a nonlinear optimization problem for which Rebelo's (1994) algorithm finds a near-optimal solution. We determined the minimum number of grid cells needed to include all species and mapped the location of these grid cells.

Distribution by latitude and altitude—To summarize the species distribution data, the number of species was tabulated by latitude and altitude. First, the number of species that occur in strips of 1° latitude was determined. Then, to obtain a smoothed line, for each 1° latitude zone the moving average was calculated, using five adjacent zones (two at each side). The GTOPO30, a 30″ grid (each cell is ~0.8 km^2) with altitude data (United States Geological Survey, 1998) was used to estimate altitude for all wild potato localities. This estimate was only used for the records for which the passport data did not include altitude. Observations were then grouped in classes of 250 m altitude and the number of species per class was plotted. The five-observation moving average was calculated and plotted.

RESULTS

Country- and species-level distribution—Wild potatoes occur in 16 countries (Table 1). Four countries (Argentina, Peru, Bolivia, and Mexico) account for 88% of the records in the database. Peru has by far the highest number of species (93 species; 47% of the total). Peru also has the highest absolute and relative (over all species in a country) number of rare species (here defined as species with five or fewer observa-

TABLE 1. Wild potato distribution by country. Number of observations (obs), species, rare species (obs ≤ 5), and the ratio of observations to species.

Country	obs	Species	Rare species	obs/species
Argentina	1688	28	4	60.3
Bolivia	1303	39	5	33.4
Brazil	24	3	0	8.0
Chile	100	4	1	25.0
Colombia	144	13	7	11.1
Costa Rica	24	1	0	24.0
Ecuador	142	15	7	9.5
Guatemala	69	6	0	11.5
Honduras	1	1	0	1.0
Mexico	926	36	6	25.7
Panama	15	2	0	7.5
Paraguay	22	2	0	11.0
Peru	1420	93	42	15.3
USA	158	3	0	52.7
Uruguay	13	2	0	6.5
Venezuela	24	3	1	8.0
Total	6073	199	72	30.5

tions), and 15 Peruvian species occur only once in our database (out of 17 species that occur once).

The distribution of the number of observations by species is far from uniform (Fig. 1). The most frequently observed species are *S. acaule* (630 observations), *S. leptophyes* (337), *S. megistacrolobum* (320), *S. bukasovii* (252), and *S. chacoense* (205). These five species account for 29% of the records and *S. acaule* alone accounts for 10%. The 72 species (36% of all species) with the least number of observations (five or less) make up only 3% of the records. A similarly skewed distribution has been described for Bolivian genebank accessions by Hijmans et al. (2000) and for the interpotato genebank by Huamán, Hoekstra, and Bamberg (2001).

The ratio between the number of observations and the number of species varies strongly across countries (Table 1). The ratio is very high in species-poor USA as well as in species-rich Argentina, indicating that these two countries have been explored more intensively for wild potatoes than have other countries, relative to their species richness. The ratio is low in many countries, some of which have low wild potato species richness. However, other countries in this group have an intermediate level of species richness, such as Ecuador and Colombia. Because the number of species tends to go up with collecting effort, the countries with a low ratio between species and observations would be the most likely places to find species that have not yet been discovered.

Area of distribution—Most species are narrow endemics and only 39 species occur in two or more countries. Most of these country co-occurrences are from Bolivia and Argentina (13 species in common) and from Bolivia and Peru (10 species in common). *Solanum chacoense* is the only species that occurs in five countries, followed by *S. acaule* and *S. commersonii* which occur in four countries. The average greatest distance between two observations of the same species (MaxD) is 411 km. For 77 species, MaxD is <100 km, and for 100 species (50% of the total), MaxD is <200 km (Fig. 2). The greatest MaxD observed was for *S. acaule.* At the time of this research, this number had recently increased by 732 km to a total of 3253 km, due to the recent discovery of this species in Ecuador (Spooner, Castillo T., and López J. [1992] who identified it as *S. albicans*; but Kardolus [1998] recognized it as an anomalous new hexaploid variety of *S. acaule*). Average circular area (CA_{50}) over all species is 6.7, but its distribution is strongly skewed. Seventy-two species have a CA_{50} of <2, and 124 have a CA_{50} of <5 (Fig. 2).

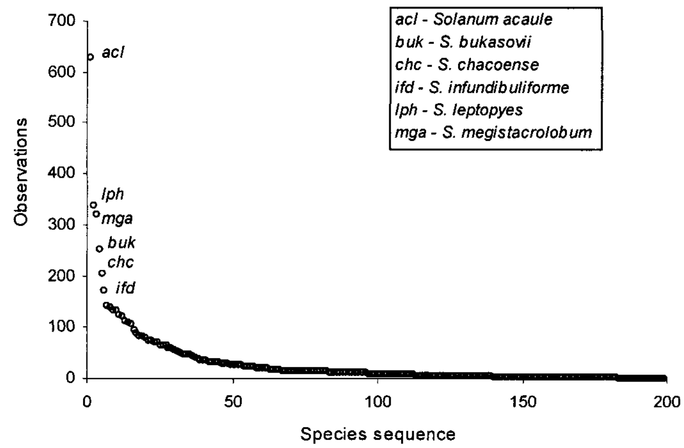

Fig. 1. Number of observations of wild potatoes by species.

Clearly, both maximum distance and circular area go up with the number of observations. On average, a species has a CA_{50} of 0.16 times its number of observations (Fig. 3, regression line). There are, however, important differences among species. For example, *S. acaule* and *S. chacoense* occur in an area of comparable size ($CA_{50} = 73$ for *S. acaule* and 74 for *S. chacoense*), but *S. acaule* has been observed about three times more often, suggesting that *S. acaule* is much more abundant. *Solanum commersonii* is third in terms of CA_{50}, but is only 17th in terms of number of observations. The CA_{50} of *S. commersonii* is 0.36 times its number of observations while for *S. acaule* it is 0.12 times its number of observations. This suggests that *S. commersonii* is less abundant within its area of distribution than *S. acaule.*

Grid based distribution—The grid based maps showing the number of observations (Fig. 4) and species richness (Fig. 5) give a much more refined picture than the country summaries

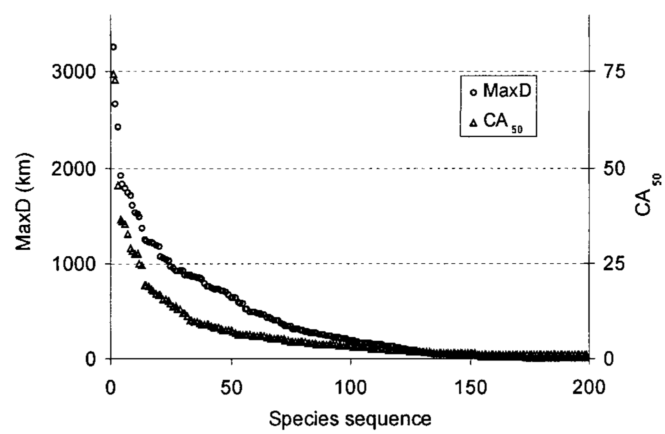

Fig. 2. Maximum distance between two observations of one wild potato species (MaxD) and circular area (CA_{50}). A circular area with a 50 km radius was assigned to each observation. Areas where circles of a species overlap were only counted once. The area is expressed relative to the area of one circle. Species sequence is not necessarily the same for MaxD and CA.

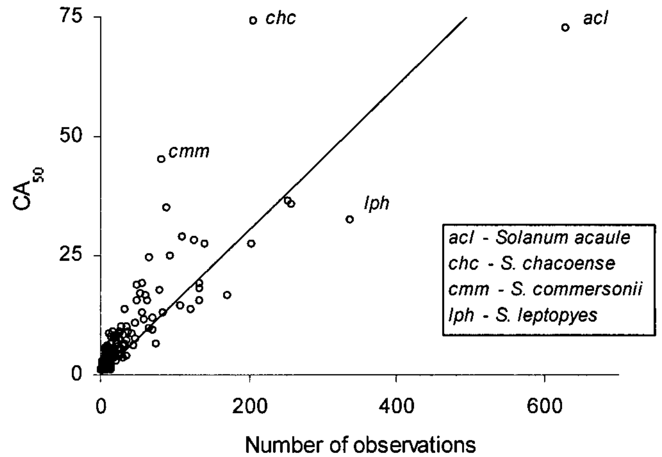

Fig. 3. Circular area (CA_{50}) vs. number of observations of wild potato species. Each dot refers to one species. A circular area with a 50 km radius was assigned to each observation. Areas where circles of a species overlap were only counted once. The area is expressed relative to the area of one circle. Regression line: $y = 0.15x$, $R^2 = 0.68$. The two dotted lines ($y = 0.1x$ and $y = 0.5x$) are included for comparison only.

presented in Table 1. Species richness is clearly not homogeneously distributed within countries. There are few areas with many species, and many areas with few species (Figs. 5, 6). The number of species follows a similar pattern to that of the number of observations. There is a strong positive correlation between the number of observations and species richness per grid cell (Fig. 7; compare Figs. 4 and 5). On average over all grid cells, there are 4.2 observations per species. Important deviations from this average are some areas in the USA and Argentina. As mentioned above, the number of observations in these two countries is high in comparison to the number of species. There is a relatively high number of samples in Argentina, particularly in the areas with high diversity. For example, there are 421 observations in the cell in Argentina with the highest number of species (17), a much higher ratio than for Peru (213 observations, 22 species) or Bolivia (141 observations, 20 species). However, in Argentina wild potatoes occur over a larger area than in any other country (295 grid cells) and the number of observations averaged out over that area has an intermediate value (17 per grid cell), lower than that of Peru (21) and of Bolivia (36; Table 2).

Species richness is particularly high in the southern and central Andes, and in central Mexico (Fig. 5). Going from north to south, the principal areas with high species richness are (1) the central Mexican highlands (México and Michoacán states); (2) a small area in central Ecuador (Chimborazo province); (3) a stretch from northern to central Peru (in Ancash, southern Cajamarca, La Libertad, and Lima departments); (4) southern Peru (in Cusco department); (5) central Bolivia (in Cochabamba, Chuquisaca, and Potosí and to a lesser extent La Paz an Tarija departments); and (6) northern Argentina (Jujuy and Salta provinces).

There are few cells with many species (Fig. 6). Cells with more than 15 species are only found in Peru, Bolivia, and Argentina; Ecuador and Mexico are the only two additional countries that have cells with nine or more species (Table 2). Only 5% of the cells have more than ten species, while 52% of the cells only have one species (Fig. 6).

The highest number of species in a single grid cell is 22 and occurs in the department of Cusco in south Peru. Two cells have 20 species, one in the Bolivian department of Potosí (on the border with Chuquisaca) and one in the Peruvian department of Ancash. There are two cells in the Peruvian department of Cusco with 19 species and one cell in the Bolivian department of Chuquisaca with 18 species. Although Peru has more species, its most species-rich areas are comparable in species richness to those of Bolivia. However, Peru has more cells with a high number of species, and its most species-rich cell only has 24% of all species present in the country. This again illustrates the high number of endemic species in Peru. In Bolivia, in contrast, the most species-rich grid cell has 51% of all Bolivian species (Table 2). There are also occurrences of relatively species-rich areas in Ecuador (60% of all species in that country), Argentina (61%), and in all countries with only a few species, but to a lesser extent in Colombia (31%) and Mexico (36%).

Complementarity analysis—Although nine grid cells are enough to capture 51% of all wild potato species, the minimum number of grid cells needed to capture all species at least once is 59 (out of 1317 total cells) (Figs. 8, 9). Twenty-three cells contribute only one additional species each (Fig. 9). The locations of the first 15 grid cells that get selected by Rebelo's (1994) algorithm follow a pattern that can only partly be inferred from Fig. 5. The early appearance of cells from Mexico and Ecuador (grid cells number 4 and 7 in Fig. 8) may seem surprising because they have only an intermediate level of species richness (Fig. 5), but the species in these countries are all different from those observed in Peru and Bolivia (the single observation of *S. acaule* mentioned above is the only exception). Areas in Argentina are selected later (numbers 6 and 13 in Fig. 8) than might be expected on the basis of its high number of species (Fig. 5). This is because some of the species in these cells were already included in area number 3 (Fig. 8) in southern Bolivia.

In northern Peru and southern Ecuador, there is a group of four nearby cells that are selected at an early stage (within the first 15 iterations). This means that in these areas, there are not only individual cells with a high level of species richness, but that there is also a high turnover of species composition between nearby grid cells. Despite its second rank (tied with Ecuador) for rare species (Table 1), Colombia is not included in the selection at an early stage because its species are not geographically clustered (see also Fig. 5; Table 2).

Distribution by latitude and altitude—Wild potatoes occur between 38° N and 41° S. The highest number of species per degree latitude (>20) occurs between 8° S and 20° S, i.e., from north-central Peru to central Bolivia, and around 20° N, in the central Mexican highlands (Fig. 10). The distribution of the number of species by latitude follows a bimodal distribution. There is a remarkably similar pattern between 20° and 40° in both hemispheres. However, in the zone between 20° N and 20° S, and particularly the zones between 8° N and 15° N and 8° S and 15° S, the number of species is rather different, with a conspicuously higher number of species in the southern hemisphere.

Wild potatoes are most common in the tropical highlands (compare Figs. 5 and 11), particularly between 2000 and 4000 m (Fig. 12). The average elevation for all species is 2770 m when weighted by species, and 2890 m when giving equal weight to all observations in the database. Ninety-one percent

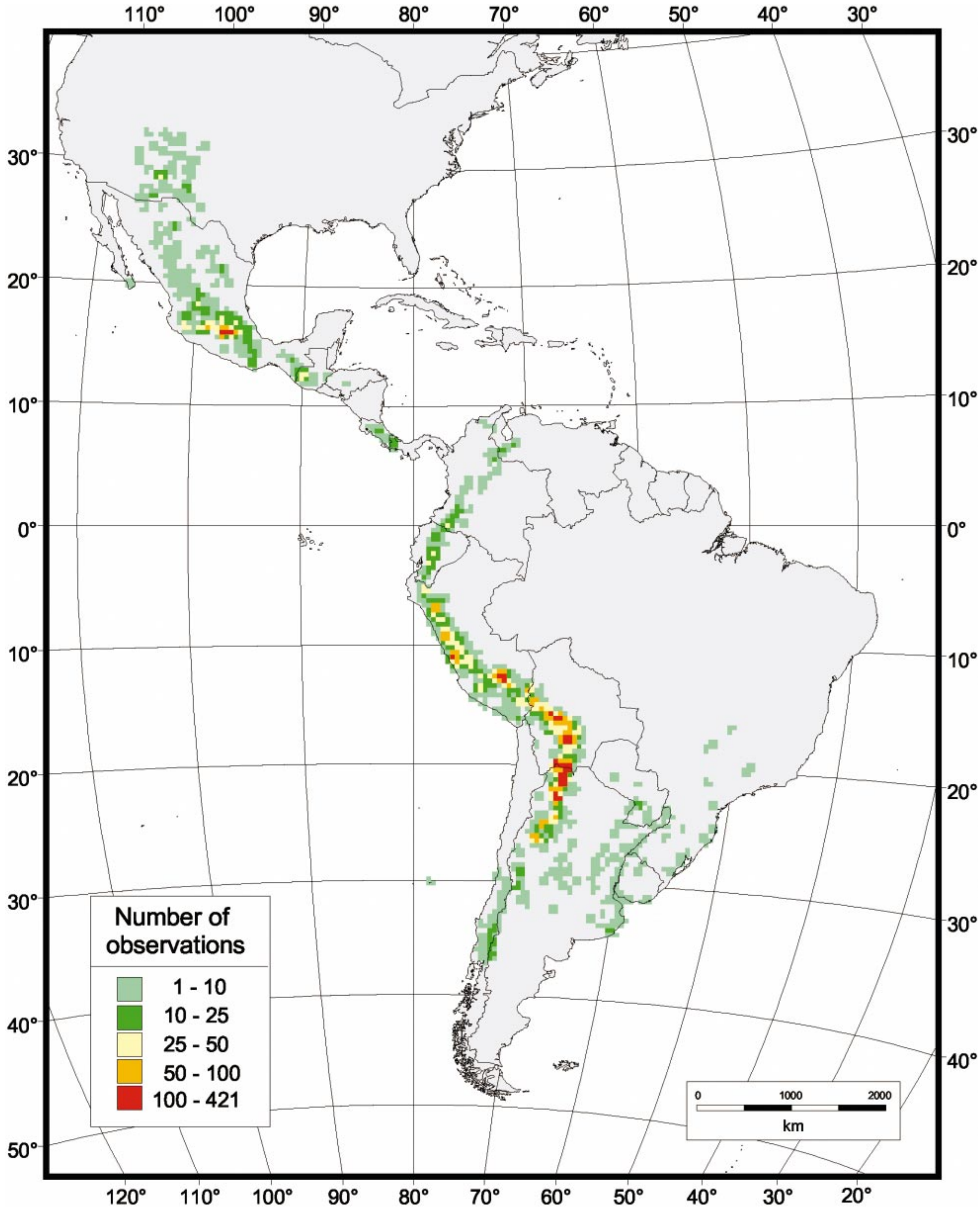

Fig. 4. Number of observations of wild potato species per 50 × 50 km grid cell. A circular neighborhood with a radius of 50 km was used to assign observations to a grid cell. There are 1317 grid cells with observations.

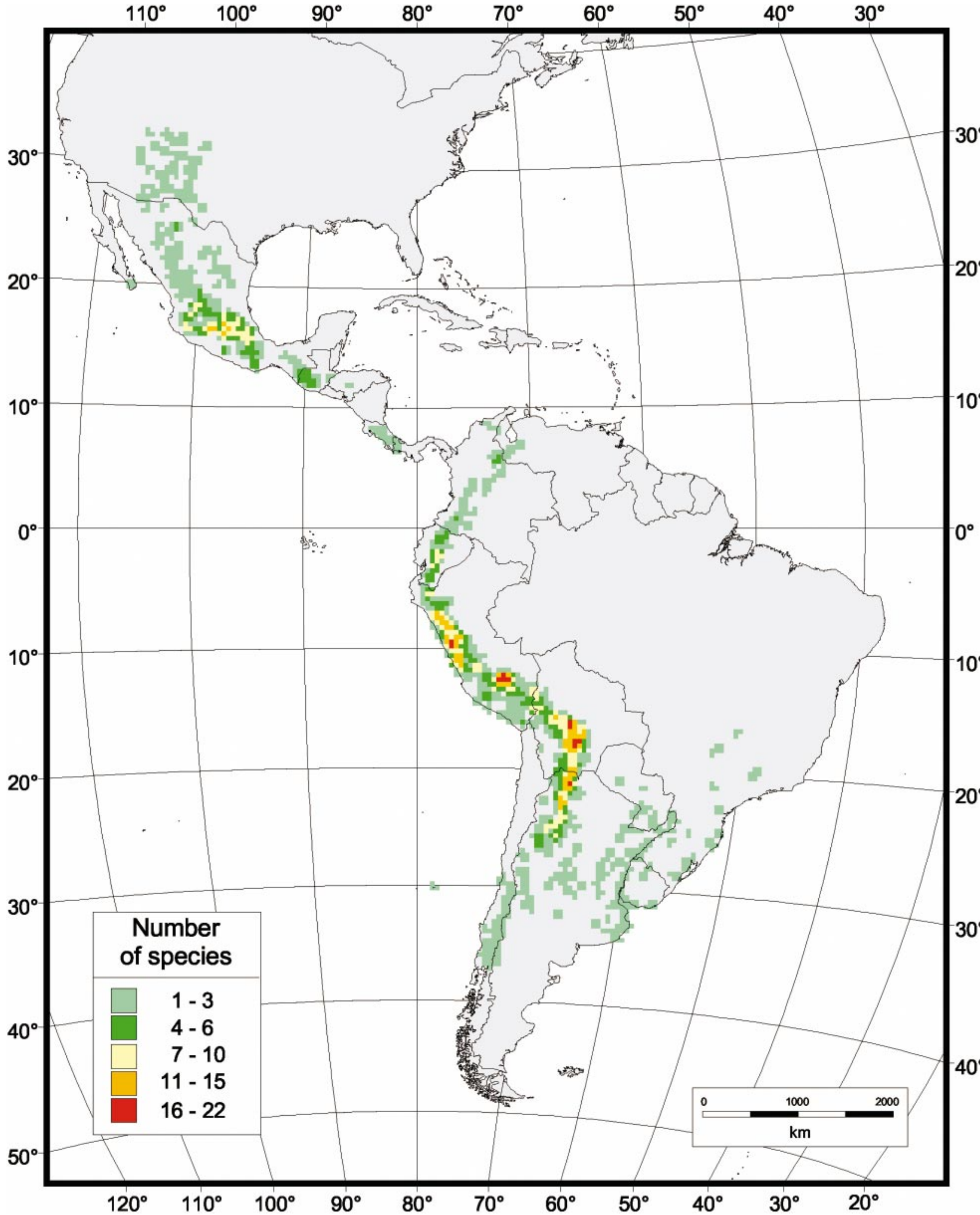

Fig. 5. Number of wild potato species per 50 × 50 km grid cell. A circular neighborhood with a radius of 50 km was used to assign observations to a grid cell. There are 1317 grid cells with observations.

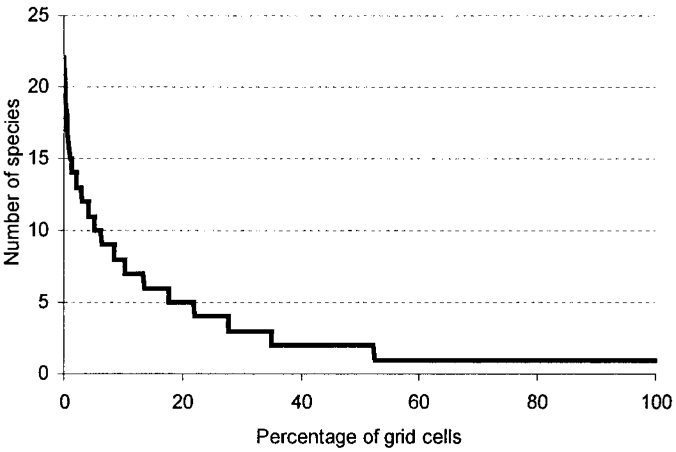

Fig. 6. Frequency distribution of the number of wild potato species per 50 × 50 km grid cell. A circular neighborhood with a radius of 50 km was used to assign observations to a grid cell.

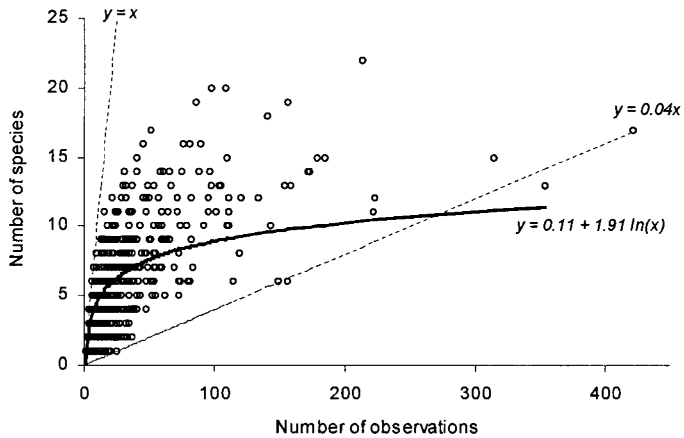

Fig. 7. Ratio of the number of wild potato species to number of observations for each grid cell (obs. = 1247). Correlation coefficient = 0.74. Regression line: $y = 0.11 + 1.91\ln(x)$, $R^2 = 0.65$.

of the wild potato species occur, on average, above 1750 m. Of all observations, 75% appear in areas above 2300 m. Almost all of the lower elevation species and observations are from the plains and hills in Argentina, Brazil, Mexico, Paraguay, Uruguay, and USA, i.e., from high latitudes.

DISCUSSION

Although distribution maps of wild crop relatives are common (e.g., Zeven and Zhukovsky, 1975), this is the first study in which a group of closely related wild crop relatives is systematically analyzed using GIS. This study is also unique in its use of a very large number of georeferenced observations for a single group of closely related wild species.

Wild potatoes occur between 38° N and 41° S. Species richness of wild potatoes is particularly high in the Central and South American tropical highlands, with clear peaks between 8° S and 20° S and around 20° N, i.e., areas in the Andes of northern Argentina, Bolivia, Ecuador, and Peru, and in central Mexico. Peru stands out for the high number of wild potato species as well as for the high number of rare wild potato species.

Many wild potatoes species are narrowly endemic, and yet a selection of only nine grid cells was needed to include 51% of the species, which emphasizes the presence of areas of high species richness. In a number of countries, the most species-rich grid cell has a high percentage all species. This might facilitate the design of in situ conservation reserves to protect these species (as called for by Huamán, Hoekstra, and Bamberg, 2000). However, the clustering of species on the scale used in this study is not directly meaningful for conservation programs, which typically would operate in considerably smaller areas.

The lower species richness around the equator, particularly in the northern hemisphere, as compared with higher tropical latitudes, contrasts with the general pattern of increasing species richness (of all flora and fauna) towards the equator (e.g., Blackburn and Gaston, 1996; Gaston and Williams, 1996). The absence of cool tropical highlands appears to be an important factor that explains the paucity of wild potato species around the equator, particularly in the northern hemisphere. The climate in these equatorial areas is also more humid and less seasonal. The absence of a clear dry (or cold) season could diminish the relative fitness of tuber-bearing perennials such as wild potatoes. At higher latitudes, where the data are more similar for both hemispheres, there is a considerable stretch of high mountains in central Mexico (the Mexican transvolcanic belt). At even higher latitudes, wild potatoes mainly occur below 1500 m altitude.

Our data were gathered from several sources (expeditions) and this may have led to some redundancies. Particularly, type localities of rare species were visited by different expeditions, as these species may not be found elsewhere. Hence, some of these species are even more rare and endemic than appears from our data. Overall, however, it may make our data more reliable given the timing dependency of the results of wild potato exploration: there are differences within and among years in the likelihood of finding certain species in certain locations.

Some of our records are recent, but many date back many years. In some cases, the habitat in which the species occurred has now disappeared, and the species may no longer occur there. For example, Spooner et al. (1998) describe a rapid rate

TABLE 2. Grid-based species richness statistics by country.

Country	No. of grid cells with one or more obs.	Mean no. of spp. per grid cell	Mean no. of obs. per grid cell	Highest no. of spp. in one cell	Total no. of spp. in the cell with highest no. of spp. (%)
Argentina	295	2.4	16.6	17	61
Bolivia	124	6.2	36.4	20	51
Brazil	56	1.0	1.5	2	67
Chile	49	1.2	5.2	2	50
Colombia	65	1.9	6.9	4	31
Costa Rica	11	1.0	6.2	1	100
Ecuador	37	4.1	12.3	9	60
Guatemala	19	3.5	11.6	6	100
Honduras	2	1.0	1.0	1	100
Mexico	269	3.2	10.7	13	36
Panama	5	1.8	11.6	2	100
Paraguay	21	1.4	3.2	2	100
Peru	211	5.3	21.2	22	24
USA	119	1.2	4.1	2	67
Uruguay	22	1.2	1.7	2	100
Venezuela	12	2.1	6.5	3	100

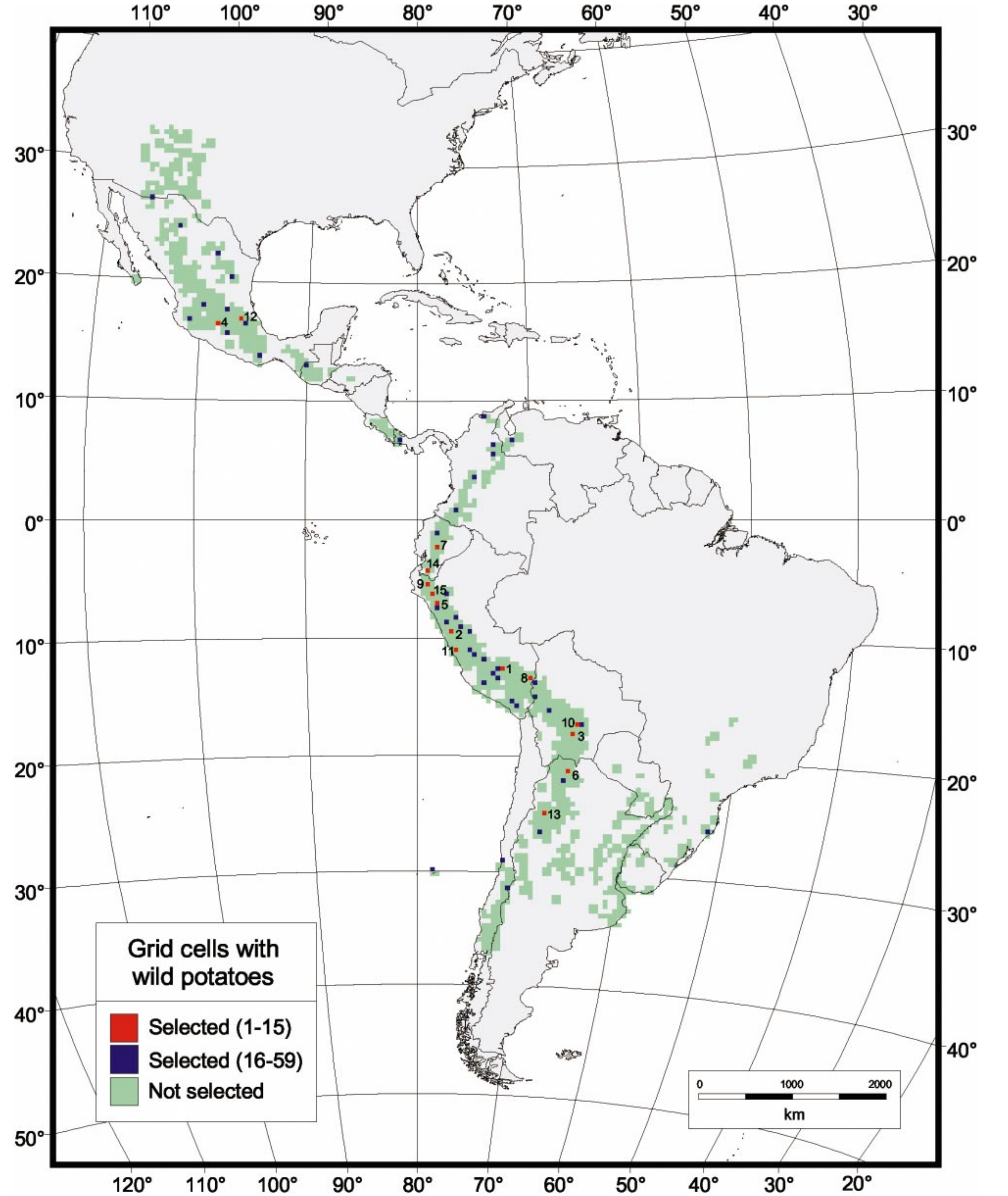

Fig. 8. The location of the first 15 grid cells selected and locations of the other 44 grid cells needed to include each wild potato species at least once.

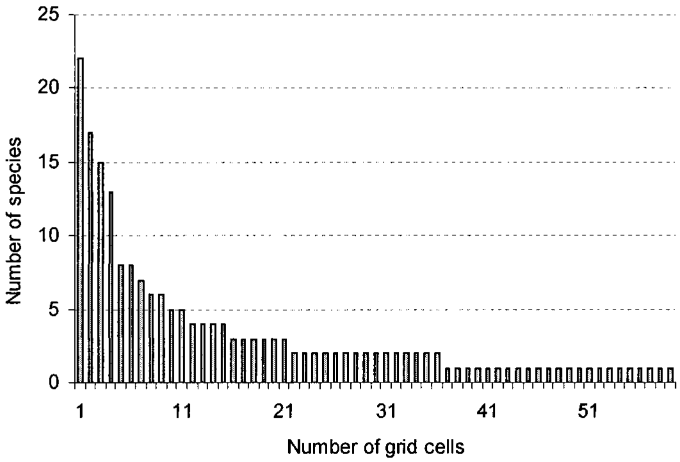

Fig. 9. Number of additional species included per grid cell, when selecting grid cells with the objective to select all species in as few grid cells as possible. The first 15 sites correspond to the numbered grid cells in Fig. 8.

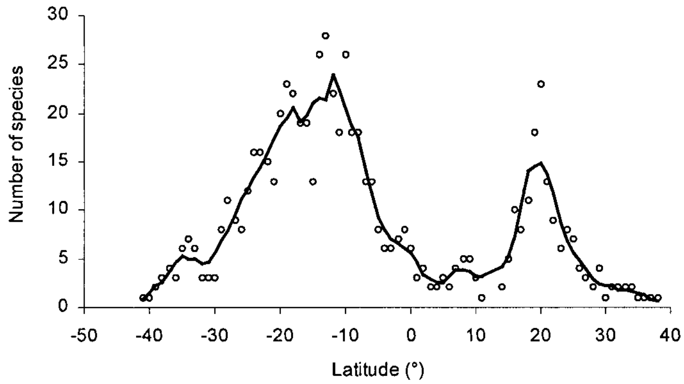

Fig. 10. Wild potato species richness by latitude. Each observation represents the number of species found in a 1° latitude wide area. The line is the five-observations moving average.

of loss of wild potato habitat in upland forests in Guatemala. However, our recent experience in Peru, and elsewhere, indicates that this is not always the case. For example, Spooner et al. (1999) and Salas et al. (2001) collected many wild potato species in Peru in the exact location, often at the type locality, where they had been collected many years before. In some cases, it was not possible to collect at documented localities, but this was often attributed to phenology, as wild potatoes often have a short growing period. In other cases, incomplete locality data hindered collections.

Recorder effort (bias) (Rich and Woodruff, 1992; Gaston, 1996; Hijmans et al., 2000) also influences the results. Nevertheless, the consistency of the results (there are no sudden gaps in the distributions) leads us to believe that we have presented a good representation of overall wild potato distribution, and this large database is one of the most comprehensive for any group of plants. Although the number of observations per grid cell is a reasonably good predictor of the number of species in that cell, we do not think that a high number of species follows causally from a high number of observations.

Wild potatoes have been the object of intensive exploration over many years (Spooner and Hijmans, 2001). While there will still be some areas where further exploration would discover additional species, a low number of observations in a cell is more likely to be the result than the cause of low species richness. Hence, the logic might be reversed: much collection has taken place in areas with a known high species diversity, referred to by Hijmans et al. (2000) as ''hotspot bias.'' Attempts could be made to correct for recorder effort through rarefaction (estimating how many species would have been observed given a constant sample size per grid cell; Sanders, 1968; Prendergast et al., 1993; Gaston, 1996). However, this would also lead to a great loss of information, because the observed number of species would be replaced by an estimate.

Distributions of wild relatives of crop plants have previously been summarized by country (e.g., Huamán, Hoekstra, and Bamberg, 2001). However, countries (or their subsidiary administrative units) have different shapes and sizes. Hence, they have only limited value in comparing geographic distributions of wild plants, despite the advantage of being familiar entities. Equal-area grid cells as used in this study are clearly to be preferred. Nevertheless, there are a number of methodological issues that need to be considered when using grids.

Resolution (cell size) of the grid affects the results. Number of species per grid cell will increase with the size of the grid cell, but this increase will be different among areas. We used a 50 × 50 km grid (and a circular neighborhood with a radius of 50 km) to strike a balance between the desire for high resolution and geographic sampling bias (Hijmans et al., 2000). This bias becomes less important when grid cell size increases. We used a neighborhood of 7854 km^2; most other studies of species distribution on a continental (or global) scale have used much larger grid cells. For example, Gaston and Williams (1996) review various global studies using grids of 611 000 km^2. The same was also used by Blackburn and Gaston (1996) for a study on birds in the Americas. Given the high density of observations, it was not necessary to use such large grid cells in our study. This would have amounted to serious information loss, and even smaller grid cells will be more appropriate for design of in situ reserves or for planning collecting expeditions to specific areas.

We used a circular neighborhood to assign values to grid cells because this gives a smoother result, particularly for areas with few observations (Cressie, 1991; Bonham-Carter, 1994) and is less sensitive to the origin of the grid and small errors in the locality data. However, this does include yet other factors (size, shape, and method used to define the neighborhood) that need to be considered when interpreting the results, in addition to the scale effect (effect of grid cell size). Research is needed to better understand the effects of different gridding methods and scales, in relation to data density and quality and the objectives of the study.

A complication of using species richness is the existence of conflicting taxonomic classifications (Gaston, 1996). Wild potatoes are a classic case in this respect (Harlan and De Wet, 1971; Spooner and Van den Berg, 1992). We use the list of species provided by Spooner and Hijmans (2001), which is a compilation of taxonomic names that updates Hawkes (1990). Nevertheless, future changes in species circumscription will likely change the results presented here (Spooner and Hijmans, 2001). For example, the results of ongoing research described by Van den Berg et al. (1998) and Miller and Spooner (1999) on the 30 taxa of the *S. brevicaule* complex suggest a need for reduction of the number of species in this group. These species occur in southern Peru and Bolivia and taxonomic revision would reduce species richness here.

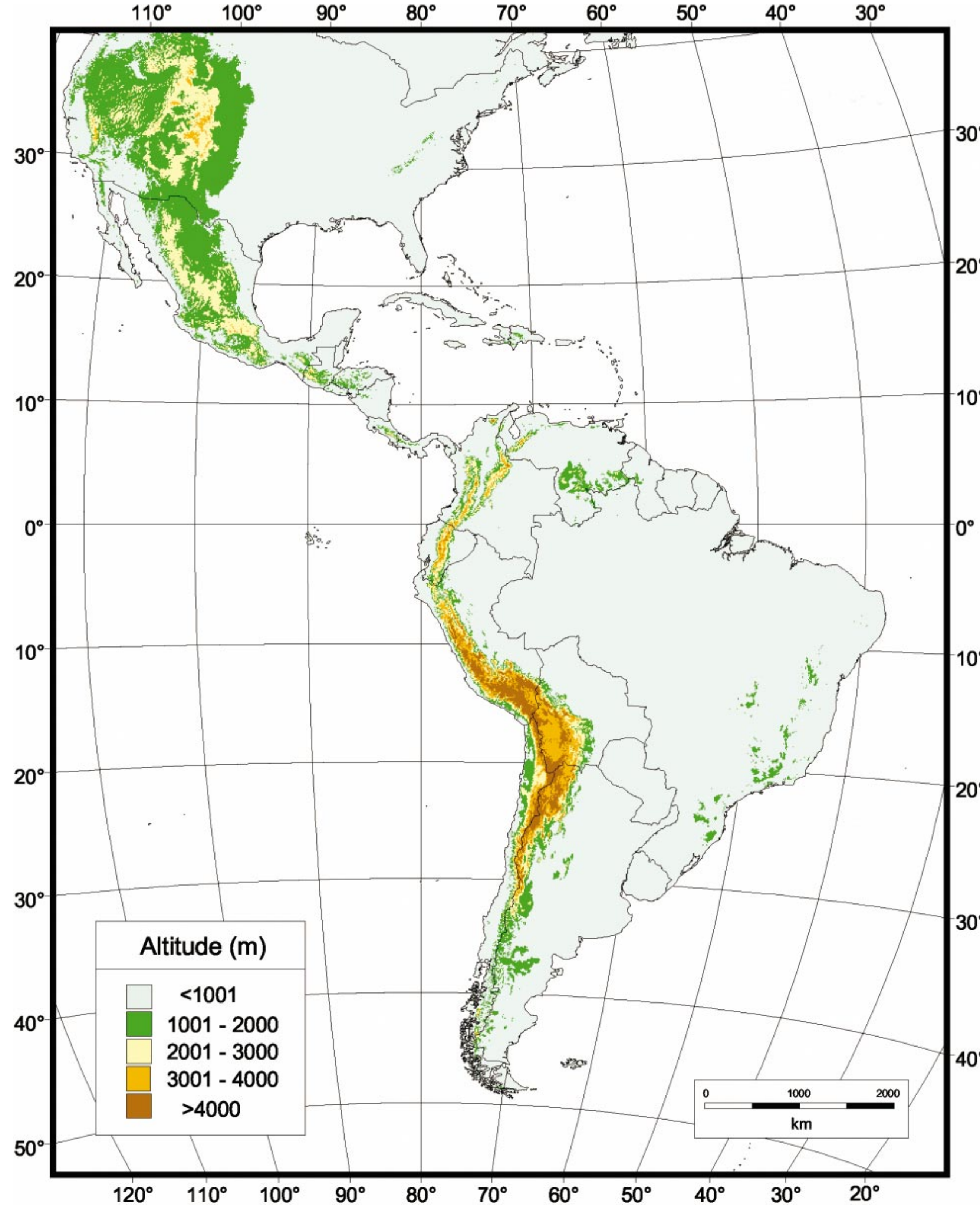

Fig. 11. Elevation in Latin America and parts of the USA. Data source: United States Geological Survey (1998).

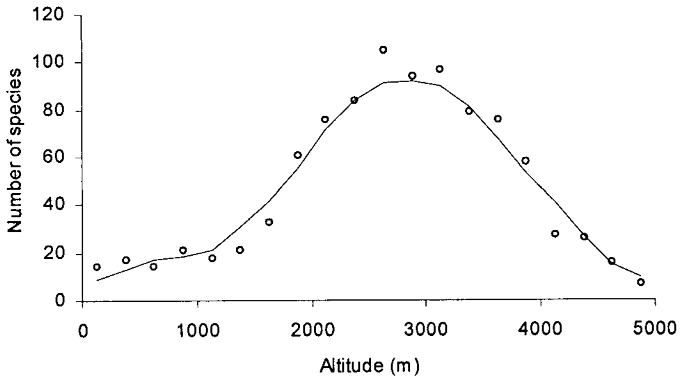

Fig. 12. Wild potato species richness by altitude. Each dot represents the number of species observed in an area covering 250 m of difference in altitude. The line is the five-observations (1250 m altitude difference) moving average.

Although Peru seems to be reasonably well explored (number of species over observations), it has an extraordinarily high number of apparently rare species. This indicates that Peru may still harbor unknown species, as illustrated by the ten new Peruvian wild potato species described by C. M. Ochoa between 1998 and 1999 (Ochoa, 1999; Spooner and Hijmans, 2001). Peruvian species are also underrepresented in genebanks, and a collecting program is currently under way to fill this important gap (Spooner et al., 1999; Salas et al., 2001).

LITERATURE CITED

BLACKBURN, T. M., AND K. J. GASTON. 1996. Spatial patterns in the species richness of birds in the New World. *Ecography* 19: 369–376.

BONHAM-CARTER, G. 1994. Geographic information systems for geoscientists. Modelling with GIS. Computer Methods in the Geosciences 13. Pergamon/Elsevier, London, UK.

CONTRERAS M. A., AND D. M. SPOONER. 1999. Revision of *Solanum* section *Etuberosum* (subgenus *Potatoe*). *In* M. Nee, D. E. Symon, R. N. Lester, and J. P. Jessop [eds.], Solanaceae IV, 227–245. Royal Botanic Gardens, Kew, UK.

CRESSIE, N. A. C. 1991. Statistics for spatial data. John Wiley & Sons, New York, New York, USA.

CSUTI, B., S. POLASKY, P. H. WILLIAMS, R. L. PRESSEY, J. D. CAMM, M. KERSHAW, A. R. KIESTER, B. DOWNS, R. HAMILTON, M. HURST, AND K. SAHR. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80: 83–97.

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. 1999. ArcView-GIS 3.1. Environmental Systems Research Institute, Redlands, California, USA.

GASTON, K. J. 1996. Species richness: measure and measurement. *In* K. J. Gaston [ed.], Biodiversity, a biology of numbers and difference, 77–113. Blackwell Science, London, UK.

GASTON, K. J., AND P. H. WILLIAMS. 1996. Spatial patterns in taxonomic diversity. *In* K. J. Gaston [ed.], Biodiversity, a biology of numbers and difference, 202–229. Blackwell Science, London, UK.

GUARINO, L., A. JARVIS, R. J. HIJMANS, AND N. MAXTED. In press. Geographic information systems (GIS) and the conservation and use of plant genetic resources. In: Proceedings of the international conference on science and technology for managing plant genetic diversity in the 21st century (SAT21), Kuala Lumpur, Malaysia, 12–16 June 2000. International Plant Genetic Resources Institute, Rome, Italy.

HARLAN, J. R., AND J. M. J. DE WET. 1971. Towards a rational classification of cultivated plants. *Taxon* 20: 509–517.

HAWKES, J. G. 1979. Evolution and polyploidy in potato species. *In* J. G. Hawkes, R. N. Lester, and A. D. Skelding [eds.], The biology and taxonomy of the Solanaceae, 637–645. Academic Press, London, UK.

HAWKES, J. G. 1990. The potato: evolution, biodiversity and genetic resources. Belhaven Press, London, UK.

HAWKES, J. G. 1997. A database for wild and cultivated potatoes. *Euphytica* 93: 155–161.

HAWKES, J. G., AND J. P. HJERTING. 1969. The potatoes of Argentina, Brazil, Paraguay and Uruguay: a biosystematic study. Clarendon Press, Oxford, UK.

HAWKES, J. G., AND J. P. HJERTING. 1989. The potatoes of Bolivia, their breeding value and evolutionary relationships. Oxford University Press, Oxford, UK.

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 and International Plant Genetic Resources Institute, Lima, Peru.

HIJMANS, R. J., K. A. GARRETT, Z. HUAMÁN, D. P. ZHANG, M. SCHREUDER, AND M. BONIERBALE. 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. *Conservation Biology* 14: 1755–1765.

HIJMANS, R. J., M. SCHREUDER, J. DE LA CRUZ, AND L. GUARINO. 1999. Using GIS to check co-ordinates of germplasm accessions. *Genetic Resources and Crop Evolution* 46: 291–296.

HUAMÁN, Z., R. HOEKSTRA, AND J. B. BAMBERG. 2000. The Inter-genebank potato database and the dimensions of wild potato germplasm. *American Journal of Potato Research* 77: 353–362.

JANSKY, S. H. 2000. Breeding for disease resistance in potato. *Plant Breeding Reviews* 19: 69–155.

KARDOLUS, J. P. 1998. A biosystematic analysis of *Solanum acaule.* Ph.D. dissertation, Wageningen Agricultural University, Wageningen, The Netherlands.

MILLER, J. T., AND D. M. SPOONER. 1999. Collapse of species boundaries in the *Solanum brevicaule* complex (Solanaceae, Sect. *Petota*): molecular data. *Plant Systematics and Evolution* 214: 103–130.

OCHOA, C. M. 1962. Los Los *Solanum* tuberíferos silvestres del Perú (*Tuberarium* subsecc. *Hyperbasarthrum*). Privately published.

OCHOA, C. M. 1990. The potatoes of South America: Bolivia. Cambridge University Press, Cambridge, UK.

OCHOA, C. M. 1999. Las papas de sudamerica: Peru. International Potato Center, Lima, Peru.

PRENDERGAST, J. R., S. N. WOOD, J. H. LAWTON, AND B. C. EVERSHAM. 1993. Correcting for variation in recording effort in analyses of diversity hotspots. *Biodiversity Letters* 1: 39–53.

RABINOWITZ, D. 1981. Seven forms of rarity. *In* H. Syne [ed.], The biological aspects of rare plant conservation, 205–217. John Wiley & Sons, New York, New York, USA.

RABINOWITZ, D., C. R. LINDER, R. ORTEGA, D. BEGAZO, H. MURGUIA, D. S. DOUCHES, AND C. F. QUIROS. 1990. High levels of interspecific hybridization between *Solanum sparsipilum* and *S. stenotomum* in experimental plots in the Andes. *American Potato Journal* 67: 73–81.

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. SIEGFRIED. 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. *Conservation Biology* 6: 243–252.

RICH, T. C. G., AND E. R. WOODRUFF. 1992. Recording bias in botanical surveys. *Watsonia* 19: 73–75.

ROSS, H. 1986. Potato breeding: problems and perspectives. Advances in Plant Breeding 13. Paul Parey Verlag, Berlin, Germany.

SALAS A. R., D. M. SPOONER, Z. HUAMÁN, R. VINCI TORRES M., R. HOEKSTRA, K. SCHÜLER, AND R. J. HIJMANS. 2001. Taxonomy and collections of wild potato species in Central and Southern Peru in 1999. *American Journal of Potato Research* 78: 197–207.

SANDERS, H. L. 1968. Marine benthic diversity: a comparative study. *American Naturalist* 102: 243–282.

SPOONER, D. M., R. CASTILLO T., AND L. LÓPEZ J. 1992. Ecuador, 1991 potato germplasm collecting expedition: taxonomy and new germplasm resources. *Euphytica* 60: 159–169.

SPOONER, D. M., AND R. J. HIJMANS. 2001. Potato systematics and germplasm collecting, 1989–2000. *American Journal of Potato Research* 78: 237–268.

SPOONER, D. M., R. HOEKSTRA, R. G. VAN DEN BERG, AND V. MARTÍNEZ. 1998. *Solanum* sect. *Petota* in Guatemala: taxonomy and genetic resources. *American Journal of Potato Research* 75: 3–17.

SPOONER, D. M., R. HOEKSTRA, AND B. VILCHEZ. 2001. *Solanum* section *Petota* in Costa Rica: taxonomy and genetic resources. *American Journal of Potato Research* 78: 91–98.

SPOONER, D. M., A. SALAS, Z. HUAMÁN, AND R. J. HIJMANS. 1999. Wild potato 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.

SPOONER, D. M., AND R. G. VAN DEN BERG. 1992. An analysis of recent taxonomic concepts in wild potatoes (*Solanum* sect. *Petota*). *Genetic Resources and Crop Evolution* 39: 23–37.

SPOONER, D. M., R. G. VAN DEN BERG, A. RIVERA-PEÑA, P. VELGUTH, A. DEL RIO, AND A. SALAS. In press. Taxonomy of Mexican and Central American members of *Solanum* Series *Conicibaccata* (sect. *Petota*). *Systematic Botany*.

UNITED STATES GEOLOGICAL SURVEY. 1998. Database available at: http://edcwww.cr.usgs.gov/landdaac/gtopo30/gtopo30.html.

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 brevicaule* complex (Solanaceae: sect. *Petota*). *American Journal of Botany* 85: 92–109.

WALKER, T. S., P. E. SCHMIEDICHE, AND R. J. HIJMANS. 1999. World trends and patterns in the potato crop: an economic and geographic survey. *Potato Research* 42: 241–264.

WATANABE, K., AND S. J. PELOQUIN. 1989. Occurrence of 2n pollen and ps gene frequencies in cultivated groups and their related wild species in tuber-bearing solanums. *Theoretical and Applied Genetics* 78: 329–336.

ZEVEN, A. C., AND P. M. ZHUKOVSKY. 1975. Dictionary of cultivated plants and their centres of diversity. Excluding ornamentals, forest trees and lower plants. Centre for Agricultural Publishing and Documentation (PUDOC), Wageningen, The Netherlands.