Biogeography of wild *Arachis* (*Leguminosae*): distribution and environmental characterisation

MORAG E. FERGUSON[1,2,*], ANDREW JARVIS[3],
H. TOM STALKER[4], DAVID E. WILLIAMS[3], LUIGI GUARINO[3],
JOSE F.M. VALLS[5], ROY N. PITTMAN[6], CHARLES E. SIMPSON[7]
and PAULA J. BRAMEL[1,8]

[1]*International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Andhra Pradesh, India;* [2]*Present address: International Institute for Tropical Agriculture(IITA), c/o ILRI, P.O. Box 30709, Nairobi;* [3]*International Plant Genetic Resources Institute (IPGRI), Regional Office for the Americas, Cali, Colombia;* [4]*North Carolina State University, Raleigh, NC, USA;* [5]*Centro Nacional de Recusos Geneticos EMBRAPA, Brazil;* [6]*United States Department of Agriculture, Agricultural Research Service, Griffin, GA, USA;* [7]*Texas Agric. Exp. Station, Texas (A & M) University, Stephenville, TX, USA;* [8]*Present address: c/o Dr. John Peacock, KISR, P.O. Box 24885,13109 Safat, Kuwait;* **Author for correspondence (e-mail: m.ferguson@cgiar.org; fax: +254-20-631499)*

Abstract. Geographic Information System (GIS) tools are applied to a comprehensive database of 3514 records of wild *Arachis* species to assist in the conservation and utilisation of the species by: (a) determining the distributional range of species and their abundance; (b) characterising species environments; (c) determining the geographical distribution of species richness; and (d) determining the extent to which species are associated with river basins. Distributional ranges, climatic variables and indices of endemism for each species are tabulated. *A. duranensis* Krapov. & W.C. Gregory, the most probable donor of the A genome to the cultivated peanut, is distributed in close proximity to both the proposed donor of the B genome, *A. ipaënsis*, and the closest wild relative of the cultigen, *A. monticola* Krapov. & Rigoni. This region in the eastern foothills of the Andes and the adjoining chaco regions of Argentina, Bolivia and Paraguay, is a key area for further exploration for wild *Arachis*. An area of particularly high species richness occurs in the State of Mato Grosso, close to the Gran Pantanal in southwest Brazil. Seventy-one percent of the species were found to have some degree of association with water catchment areas, although in most cases it was difficult to determine whether this was due to climatic adaptation reasons, restricted dispersal due to geocarpic habit, or the role of watercourses as a principal dispersal agent. In only two cases could climatic adaptation be eliminated as the reason for species distribution.

Introduction

The importance of the wild relatives of the cultivated peanut (*Arachis hypogaea* L.) to the improvement of the crop has long been recognized. Wild species are endemic to South America and have been documented, collected and conserved since the times of the first European plant explorers. Today there are numerous records of populations, herbarium samples and gene bank accessions that are housed at a number of institutes around the world.

Unfortunately, each institute has assigned a different unique identifier to an accession according to its own numbering system. To bring clarity to past conservation, research and plant breeding efforts, as well as enhance the future efficiency of both conservation and utilisation of wild *Arachis* germplasm, a comprehensive and extensive database of 3514 records, comprising and cross-referencing accessions held in the major genebanks, specimens held in herbaria and citations made in publications were compiled by Stalker et al. (2000). This comprehensive database should greatly facilitate both conservation and utilisation efforts, and has already been used as the basis to determine conservation priorities (Jarvis et al. 2003).

The genus *Arachis* currently consists of 69 described species divided into 9 sections, with Section *Arachis* containing the cultivated peanut and its closest wild relatives. All *Arachis* species are diploid ($2n = 2x = 20$ or $2n = 2x = 18$), except the cultigen (*A. hypogaea*) and *A. monticola* (wild progenitor or weedy form) as well as *A. glabrata* Benth in Section *Rhizomatosae* Krapov. & W.C. Gregory, that is tetraploid ($2n = 4x = 40$). *Arachis monticola* is indistinguishable from the cultivated peanut on the basis of DNA markers (Halward et al. 1991; Kochert et al. 1991) and closely related morphologically.

Cultivated peanut is an allotetraploid, having an A and a B genome (AABB), and is thought to have evolved relatively recently from a single hybridisation event, either between the unreduced gametes of two diploid species pertaining to different genomes (AA + BB) or two haploid gametes (A + B) which subsequently underwent spontaneous doubling, thereby restoring fertility. Either way, this natural hybridisation event reproductively isolated *A. monticola* and *A. hypogaea* from their original genome donors and other wild species (Kochert et al. 1996). Kochert et al. (1996) concluded as a result of genomic DNA and chloroplast DNA RFLP evidence, supported by cytological (Fernandez and Krapovickas 1994) and simple sequence repeat marker evidence (Ferguson, unpublished data; Moretzsohn 2001), that the most likely donors, from currently known species were the diploid species *A. duranensis* contributing the A genome and *A. ipaënsis* Krapov. & W.C. Gregory contributing the B genome. The possibility however does exist that one or both of the original donor species is yet undiscovered, or has been contributed by another known species.

This reproductive isolation imposed a genetic bottleneck on *A. hypogaea* leaving a relative scarcity of genetic variability in the cultivated peanut, while extensive variability remained in the wild species. This is evident from genetic marker studies using RFLPs (Kochert et al. 1991; Paik-Ro et al. 1992), SSRs (Hopkins et al. 1999; Moretzohn 2001), RAPDs (Halward et al. 1992), isozymes (Grieshammer and Wynne 1990; Stalker 1990; Lacks and Stalker 1993; Stalker et al. 1994) and seed storage proteins (Tombs 1963; Bianchi-Hall et al. 1991).

The narrow genetic base of the cultivated peanut has obliged plant breeders to tap the genetic pool of the wild species where high levels of resistance to many major pests and diseases have been found (Stalker and Moss 1987;

Stalker 1992). Fertility barriers in the form of different ploidy levels and irregular chromosomal pairing has made introgression of desirable traits from wild species a time consuming and difficult process. Several interspecific crosses are available, however, and have been used in breeding programmes (Stalker and Moss 1987; Simpson et al. 1993; Simpson and Starr 2001; Stalker and Lynch 2002; Stalker et al. 2002a, b).

The genus *Arachis* is distinguished from other closely related genera by its geocarpic pods. This characteristic has a major effect on seed dispersal and thus rates of migration and species distribution. Gregory et al. (1973) and Smartt and Stalker (1982) make a plausible suggestion that one of the major dispersal agents is moving water, capable of dislodging both soil and fruits. They argue that this is supported by the geographical distribution of taxa, which tend to be closely associated with specific drainage basins of both ancient and recent times, and that this has also played a major role in the isolation and evolutionary divergence of the major sub-generic groups.

Thus, the objectives of this study were to analyse data in the above mentioned database to assist in the conservation and utilisation of wild *Arachis* genetic diversity by: (a) determining the distribution ranges of species and their abundance; (b) characterising species environments; (c) determining the geographical distribution of species richness; and (d) determining the extent to which taxa are associated with particular river basins.

Materials and methods

Data for the analysis were derived from the 'Catalog of *Arachis* Germplasm Collections' compiled by Stalker et al. (2000) and available for querying at http://www.icrisat.org/text/research/grep/homepage/groundnut/arachis/start.htm. This database cross-references accessions, based on collector name and number, in the databases of EMBRAPA, United States Department of Agriculture (USDA), Centro Internacional de Agricultura Tropical (CIAT), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Texas A & M University, and North Carolina State University, as well as in herbaria and publications. The 2175 records that have been identified to the species level and geo-referenced were included in the analysis. This includes both herbarium specimens and germplasm accessions. Observations were not included in the analysis if they were known to have come from an *ex situ* field collection, with corresponding geographical co-ordinates, or if they were indicated as being cultivated. Three species are commonly cultivated as well as being found in the wild, these are *A. glabrata*, *A. repens* Handro and *A. pintoi* Krapov. & W.C. Gregory. *Arachis villosulicarpa* Hoehne has only been cultivated by indigenous people in the northwestern part of the Brazilian state of Mato Grosso (Gregory et al. 1973) and *A. stenosperma* Krapov. & W.C. Gregory was apparently cultivated for its seed and distributed by early Europeans from central Brazil to the eastern coast of Brazil. Two species have sub-specific classifications, namely

A. glabrata and *A. paraguariensis* Chodat & Hassl. The majority of observations, however, have not been determined to that level and these two taxa were therefore analysed at the species level only.

To verify co-ordinate data, we first checked for gross errors by plotting all species on a dot-map using DIVA (Hijmans et al. 2001). DIVA-GIS is available at no cost from http://gis.cip.cgiar.org, and the reader will find further information about its use in a plant genetic resources context in Hijmans et al. (2002). Dot-maps of individual species were then produced in FloraMap (Jones and Gladkov 1999) and the co-ordinates of any geographical outliers verified using the Gazetteer available at http://164.214.2.59/gns/html/index.html.

The number of observations (records) of each species and their distribution by country were tabulated. The distributional range of species was described on a country basis by tabulating the number of observations, number of species and number of sections per country. The distribution of sections was plotted on dot-maps using FloraMap (Jones and Gladkov 1999).

To quantify the area over which each species is distributed, given the available data, the maximum distance (MaxD) and the circular area (CAr) over which observations were distributed was calculated following the methods of Hijmans et al. (2001). MaxD is the longest distance between any pair of observations of one species. CAr is calculated by assigning a circle of radius (r), in this case $r = 50$ km, to each observation. The area over which the species is distributed is then calculated with overlapping areas being included only once. Area is expressed relative to the area of one circle. The CA_{50} statistic was plotted against the number of observations of a species to explore differences in abundance among species. This was quantified using a relative CA_{50} (R CA_{50}) calculated as CA_{50}/number of observations. The number of observations recorded in protected areas was also determined using the Global Protected Areas Dataset held by the World Conservation Monitoring Centre (2000).

To characterise species environments, highlight the factors that may influence the geographic distribution of species, and provide an indication of the abiotic stress tolerances that exist within species variation, FloraMap was used to determine the distributional ranges of wild peanuts over 36 climatic variables and elevation. Climatic variables include the mean maximum and minimum temperatures for each month and mean monthly rainfall. FloraMap includes climatic data from a 10-min grid (corresponding to 18 km at the Equator). The grids were derived by interpolation of data from over 10,000 meteorological stations (Jones 1991). Rainfall and diurnal temperature range remain independent of elevation. Altitude was inferred for all records whether or not elevation data was available in the original data. Here reported are the mean maximum and minimum temperatures, minimum and maximum monthly rainfall, and mean annual rainfall, together with the altitudinal range of each species. In an effort to compare the climatic adaptations of different species, the means of each of the 36 variables, for 68 species, were subjected to Principal Components Analysis and to cluster analysis (Euclidian distances, Ward's method) using STATISTICA software.

Areas of high species richness were located by determining and displaying the number of species occurring in each cell of a 50×50 km grid using DIVA (Hijmans et al. 2001). Species richness is used as a measure of taxonomic diversity because it is a simple, useful, widely used and understood parameter (Gaston 1996). It is also less sensitive to the problems of unsystematic sampling intensities and procedures than other diversity indices (Hijmans et al. 2000).

The hypothesis that species distributions are associated with watersheds was tested by overlaying species dot-map distributions with watershed patterns using ESRI's ArcView. Watershed information was obtained from the USGS Hydro1k Basins dataset (2000) at http://edcdaac.usgs.gov/gtopo30/hydro/samerica.html that provides six levels of watersheds with increasing resolution (i.e., level 1 has very broadly defined catchment areas (low resolution), whereas level 6 has a large number of small catchment areas (high resolution)). Level 1 catchments include the entire Amazon basin, the Paraná basin (including the Rio Paraguay), and the Tocantins basin among others. The level at which each section and species is distributed was tabulated, and the number of watersheds at the next level across which each species is distributed was also recorded. The distribution across watersheds was compared with the potential species distribution based on climatic similarities, generated using FloraMap to determine whether species were likely to be restricted to the watershed by their seed dispersal mechanism and/or by climatic limitations. A minimum of 10 geographically unique accessions were defined in order to determine the potential distribution of a species based on climatic parameters. This is an established methodology with fewer locations resulting in unreliable climatic ranges (Jarvis et al. 2003). If there is significant climatic potential distribution outside of the catchment within which the species is restricted, this provides evidence that watersheds are a limiting barrier to species distributions.

Results and discussion

Ecogeographic distribution

The genus *Arachis* is distributed across 5 countries (Argentina, Bolivia, Brazil, Paraguay and Uruaguay), from the highlands of Ceara, Piauí, and Maranhão in northeastern Brazil, just south of the equator, to the northern bank of the Rio de la Plata in Uruguay (35°S), and from the Atlantic coast to the Parana and the eastern foothills of the Andes. Seventy-six percent of all observations were from Brazil (Table 1), representing 54 species and all 9 sections. Two sections are endemic to Brazil, these are *Heteranthae* Krapov. & W.C. Gregory and *Triseminatae* Krapov. & W.C. Gregory. The geographical distribution of each section is illustrated in Figure 1. Maps of all species ranges and locations can be generated at http://www.icrisat.org/text/research/grep/homepage/groundnut/arachis/start.htm and details of climatic conditions and elevations in Table 2. Mean annual rainfall ranges from 706 mm for *A. batizocoi* Krapov.

Table 1. Number of species and number of observations per country.

Country	No. of observations	No. of species	No. of endemic species	No. of sections	No. of endemic sections
Argentina	172	6	1	2	0
Bolivia	142	17	9	3	0
Brazil	1658	54	38	9	2
Paraguay	151	13	0	5	0
Uruguay	55	2	0	2	0
Total	2178		48		2

& W.C. Gregory to 1797 mm for *A. trinitensis* Krapov. & W.C. Gregory. Temperatures range from a mean monthly minimum of 7.2 °C in *A. monticola* to 21.3 °C in *A. williamsii* Krapov. & W.C. Gregory, and a mean monthly maximum of 20.3 °C for *A. monticola* to 32.6 °C in *A. williamsii*.
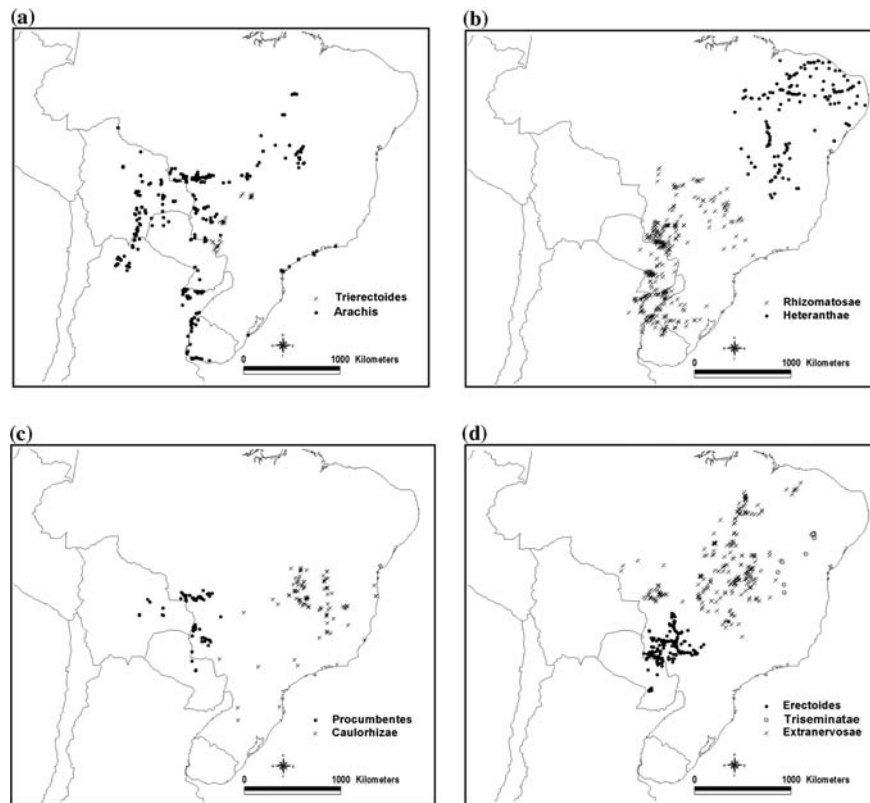

Figure 1. Geographical distribution of sections in the genus *Arachis*. (a) Sections *Arachis* and *Trierectoides*, (b) Sections *Heteranthae* and *Rhizomatosae*, (c) Sections *Procumbentes* and *Caulorhizae*, and (d) Sections *Erectoides*, *Triseminatae* and *Extranervosae*.

Table 2. The number of observations per species, the countries through which they are distributed, climatic characteristics of species environments and the indices of endemism, MaxD and RCA.

Species	Observations per country					Mean min temp (°C)	Mean max. temp (°C)	Min. mean annual rainfall (mm)	Max. mean annual rainfall (mm)	Min. alt. (m)	Max. alt. (m)	Number of observations in protected areas	Area CA50	Max D (km)	R CA50
	ARG	BOL	BRA	PAR	URU										
Section *Arachis*															
batizocoi	21				2	16.7	29.1	505	2191	1	2378		5.9	259	178
benensis		8				16.7	29.4	554	787	285	1498		2.1	451	37
cardenasii	26			2		20.7	32.1	1453	1797	120	872	7	7.6	557	107
correntina	38				4	18.6	31.5	637	1232	147	78	8	7.9	452	137
cruziana	4					15.8	27.3	1144	1600	40	305		1.8	70	207
decora			31			18.7	31	1126	1154	294	675		7.4	274	189
diogoi			11		8	14.9	30.5	1360	1568	277	305		8.0	1191	53
duranensis	39	20				18.8	27.7	1103	1510	68	1652	1	13.5	742	143
glandulifera	5			1		13.9	31.1	505	1186	187	581	4	4.6	539	67
helodes			25			18.3	30.5	1080	1456	207	194		3.6	130	220
herzogii	2					19.0	29.7	1252	1396	122	293		1.2	16	599
hoehnei		8			5	18.7	31.1	1154	1154	290	247		4.3	494	68
ipaënsis	2					18.8	29.5	1095	1473	65	689		1	0	3927
kempff-mercadoi	25					16.1	30.0	786	786	689	442		4.0	232	137
kuhlmannii			61			19.0	30.9	1037	1520	256	581		14.7	821	140
magna	3		10			18.4	30.6	1108	1467	78	577		6.2	421	115
microsperma			5			17.7	30.1	1053	1381	200	217	1	2.5	114	169
monticola	12					7.2	20.3	876	878	141	2378		1.3	23	454
palustris			7			19.5	32.5	1177	1428	1371	207		3.0	281	83
praecox	3					17.7	30.0	1567	1697	132	181		1.4	33	334
simpsonii	3		10			17.5	30.2	1279	1285	134	305	3	3.5	109	251
stenosperma			68			17.9	29.5	1246	2191	172	418		17.6	1564	89
trinitensis	2					21.0	32.3	1797	1797	155	155		1.0	0	3927

Table 2. Continued.

Species	Observations per country					Mean min. temp (°C)	Mean max. temp (°C)	Min. mean annual rainfall (mm)	Max. mean annual rainfall (mm)	Min. alt. (m)	Max. alt. (m)	Number of observations in protected areas	Area CA$_{50}$	Max D (km)	R CA$_{50}$
	ARG	BOL	BRA	PAR	URU										
valida			7			19.5	31.3	1096	1113	63	281		1.7	56	245
villosa	16		3		32	13.0	23.9	957	1357	1	75		13.6	843	127
williamsii		3				21.3	32.6	1797	1797	155	155		1.0	0	2618
Section Caulorrhizae						16.1	29.0	775	2069	1	1098				
pintoi			132			16.1	29.2	834	2069	27	1098		29.8	1219	192
repens		1	33			16.2	28.2	775	1836	1	936		19.4	3606	42
Section Erectoides						17.0	30.1	922	2009	46	695				
archeri			39			16.1	28.9	1443	1462	300	668		3.1	100	243
benthamii			46			16.8	29.5	1104	1487	65	527		14.7	584	198
brevipetiolata			2			16.3	29.2	1400	1487	552	600		2.0	450	35
cryptopotamica			17			17.9	29.9	1292	1529	122	496		3.9	160	190
douradiana			16			16.1	29.4	1455	1505	299	585		4.9	194	198
gracilis			12			16.1	28.9	1279	1483	303	527		5.9	250	185
hatschbachii			7			16.8	29.1	1324	1442	250	695		2.6	124	167
hermannii			11			17.4	29.7	1380	1467	212	643		3.6	244	115
major			45	12		17.1	29.9	1174	1526	87	585		12.2	510	188
martii			3			16.0	28.8	1446	1446	494	647		1.3	23	441
oteroi			56			16.1	28.8	1278	1471	243	689		10.9	382	224
paraguariensis			41	19		17.4	29.4	1146	1517	46	601	2	11.3	605	147
stenophylla			10	1		17.4	29.5	1190	1454	151	574		4.2	169	194
burchellii			91			18.5	31.6	1292	1797	73	639	1	36	1727	163
lutescens			68			17.7	30.1	1151	1671	122	883		21.9	1152	150
macedoi			31			16.5	31.5	1318	1643	170	800		13.6	1322	81
marginata			6			15.0	31.4	1474	1520	583	761		1.7	44	295
pietrarellii			12			18.7	31.1	1217	1643	168	612		3.4	794	33
prostrata			94			16.2	30.8	922	1816	125	1070	2	40.3	1789	177

Taxon													
retusa		15		14.4	30.7	1360	1642	277	1186		6.2	301	163
setinervosa		6		18.1	32.1	1614	1643	287	539		2.3	236	77
villosulicarpa		6		16.7	28.9	1634	2009	301	424	5	2.7	217	98
Section *Heteranthae*				19.1	21.2	536	1712	9	768				
dardani		70		20.3	31.0	536	1675	9	768		37.0	1514	192
giacomettii		3		17.2	30.3	931	950	610	610		1.0	4	2221
pusilla		33		18.3	31.2	767	1531	139	767		15.7	1557	79
sylvestris		89		18.4	31.3	767	1712	65	706		41.4	1654	197
Section *Procumbentes*				18.3	30.4	1052	1498	44	554				
appressipila		22		19.0	30.7	1052	1113	68	298		2.7	129	164
chiquitana	4			18.5	31.2	1149	1154	275	301	1	2.0	83	190
kretschmeri		14		18.1	30.5	1100	1498	77	301		5.7	349	129
lignosa		1	11	18.3	29.5	1167	1306	44	99		3.1	216	113
matiensis	10	31		18.0	30.5	1076	1396	106	554		8.7	669	102
rigonii	3			19.7	29.7	1163	1163	420	450	7	1.1	9	941
Subcoriacea		19		17.9	30.0	1186	1396	112	397		5.8	311	147
vallsii		8		19.0	31.0	1100	1246	148	281		2.1	140	120
Section *Rhizomatosae*				15.7	27.9	1098	1757	22	853				
burkartii	24	53	23	13.4	24.9	1098	1753	22	598		27.6	930	233
glabrata	43	178	80	16.5	28.7	1180	1757	48	853		74.5	1602	366
pseudovillosa		38	5	15.0	28.7	1216	1563	77	695	6	8.7	398	171
Section *Triseminatae*				18.5	31.0	706	1059	304	1053				
triseminata		21		18.5	31.0	706	1059	304	1053		7.1	751	74
Section *Trierectoides*				15.4	28.8	1443	1592	275	607				
guaranitica		12	1	15.4	29.1	1443	1563	394	607		3.8	139	215
tuberosa		17		15.4	28.6	1443	1592	275	846		5.4	463	91

Section *Arachis* contains 27 of the 69 species in the genus including *A. hypogaea*, and has the broadest geographical distributional range (Figure 1a). It is found in all 5 countries of the distributional range of the genus (Table 2), from the southern extreme of the genus along the river Uruguay to the eastern most extreme of the genus in Bolivia and Argentina and north-eastwards across the Brazilian Highlands. The section occupies a broad altitudinal range from 1 to 2378 m. The species *Arachis villosa* Benth. occurs at particularly low altitudes (1–75 m), and *A. monticola* at exceptionally high altitudes (1371–2378 m). *A. monticola* is associated with correspondingly low temperatures (7.2 to 20.3 °C mean minimum and maximum monthly temperatures).

Section *Heteranthae* contains 4 species, *A. dardani*, *A. giacomettii* Krapov., W.C. Gregory, Valls and C.E. Simpson, *A. pusilla* Benth. and *A. sylvestris* (A.Chev) A.Chev. The section is endemic to the north-east highlands of Brazil, and defines the north-east distribution of the genus, in tropical and sub-tropical dry forest and savannah environments (Figure 1b, Table 2).

Section *Trierectoides* Krapov. & W.C. Gregory contains two species *A. guaranitica* Chodat & Hassl. and *A. tuberosa* Bongard ex Benth. The section has a very narrow distributional range, being almost endemic to Brazil, apart from one population of *A. guaranitica* in Paraguay (Figure 1a). Section *Trierectoides* is distributed from Sierra de Amambay northwards, at high elevation, across Mato Grosso do Sul to Goias.

Section *Caulorhizae*, including *A. pintoi* and *A. repens*, is endemic to Brazil and is centred in the eastern Brazilian Highlands with scattered populations found towards the highlands of Mato Grosso do Sul. Both species in this section occupy a wide altitudinal range from around sea level to 1098 m and may be found in weedy and cultivated situations as well as in the wild (Figure 1c).

Section *Procumbentes* Krapov. & W.C. Gregory, consisting of 8 species, is distributed where the borders of Paraguay, Bolivia and Brazil come together, near an area known as the Pantanal (Figure 1c). *Arachis appressipila* Krapov. & W.C. Gregory, *A. kretschmeri* Krapov. & W.C. Gregory, *A. subcoriacea* Krapov. & W.C. Gregory, and *A. vallsii* Krapov. & W.C. Gregory are endemic to Brazil, with *A. chiquitana* Krapov. & W.C. Gregory & C.E. Simpson and *A. rigonii* Krapov. & W.C. Gregory being endemic to Bolivia (Table 2). Apart from a single population of *A. lignosa* (Chodat & Hassl.) Krapov. & W.C. Gregory in Brazil, all other eleven populations occur in Paraguay. *Arachis lignosa* tends to occur at low altitudes, with a maximum recorded altitude of 100 m.

Section *Erectoides* Krapov. & W.C. Gregory, consisting of 13 species has a restricted distribution, largely in the Brazilian province of Mato Grosso do Sul, stretching southwards into Paraguay (Figure 1d).

Section *Extranervosae* Krapov. & W.C. Gregory, consisting of 9 species, is also endemic to Brazil, inhabiting the Brazilian Highlands north and west of Mato Grosso do Sul, spreading across the Brazilian plateau, as far as 5°S (Figure 1d).

Section *Triseminatae* is represented by a single species endemic to the northeastern Brazilian Highlands (Figure 1d).

Section *Rhizomatosae*, comprised of three species, is found in 4 countries, inhabiting areas surrounding the Parana basin, and southwards through Paraguay, Argentina and into Uruguay, following the Rio Paraguay, and meeting with the Rio Uruguay (Figure 1b). *Arachis burkartii* Handro is the only species in this section to be distributed in Uruguay. *Arachis glabrata* is represented by a particularly large number of populations (301). Each species in this section occupies a wide range of elevations.

Ecogeographic distribution of putative wild progenitors

Twelve collections of A. *monticola* have been made from possible different locations, however location data are scanty, and it is likely that some of these collections are of the same population. All collections are from around Yala, in Jujuy province, Argentina, in the eastern foothills of the Andes, and at the most westerly extreme of the distributional range of the genus.

The geographical range of the progenitor and/or wild or weedy form of A. *hypogaea*, A. *monticola*, is towards the upper extreme of elevation of the proposed donor of the A genome, A. *duranensis* (Figure 2). This explains
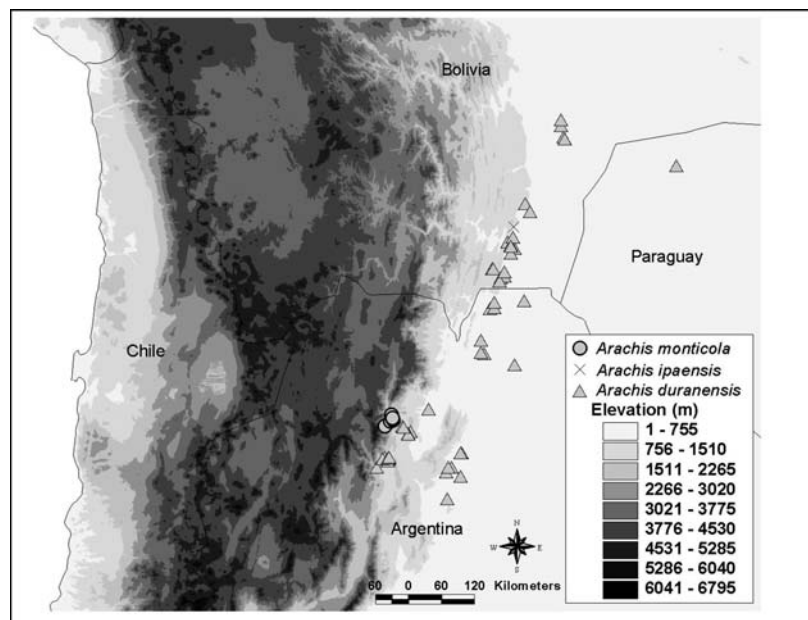

Figure 2. Geographical distribution of putative wild progenitor and genome donors of *A. hypogaea.*

differences in mean monthly maximum and minimum temperatures of the 2 species. *A. duranensis* has been collected in the Argentinian provinces of Yala and Salta, and the Bolivian provinces of Tarija, Santa Cruz and Chuquisaca as well as Alto Paraguay, which are all adjacent to one another (two outliers from Tocantins were removed from the analysis on the grounds of being cultivated).

A. ipaënsis, the most likely donor of the cultivated peanut's B genome is only represented by two accessions that are derived from the same population located in the village of Ipa, Tarija, Bolivia. The geographical distribution of *A. ipaënsis* overlaps with that of *A. duranensis* (Figure 2) in the lower altitude range (689 m). *Arachis villosa* has been proposed as a potential A genome donor (Raina and Mukai 1999a, b); however, its distribution in Corrientes, Argentina, and in Uruguay makes this unlikely.

Climatic adaptation

PCA analysis of the means of 36 climatic data derived from FloraMap reveals some variation among wild peanut species in their climatic adaptation. The first two principal components (PCs) accounted for 34 and 31% of the total variation in the climate data, respectively (Figure 3). PC1 had strongly negative loadings for June–August rainfall and strongly positive loadings for December–February rainfall and May–October diurnal temperature range.
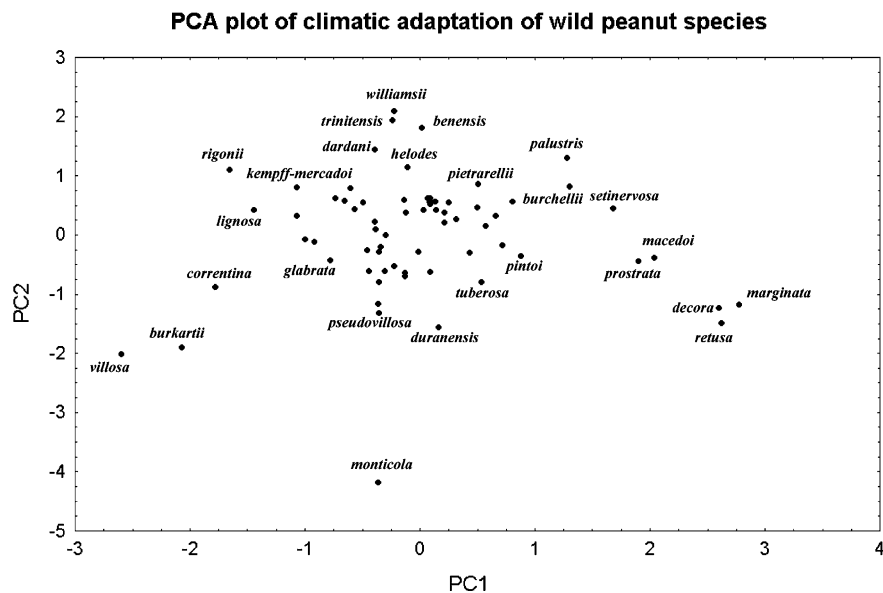

Figure 3. Scatter plot of PC1 against PC2 derived from a PCA analysis of the mean of 36 climatic variables for 68 species.
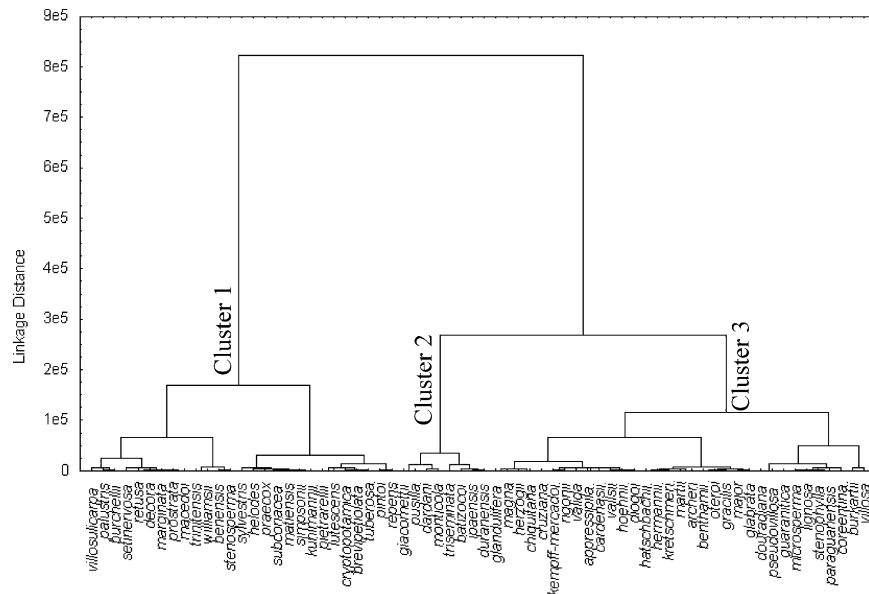
Figure 4. Species relationships according to Euclidean distance calculated from the mean of each of 36 climatic variables. The cluster of species at the extreme left of the dendrogram are those with positive PC1 scores, while the species in the cluster at the extreme right have negative PC1 scores.

PC2 had strongly positive loadings for mean temperature in all months and negative loadings for December–January diurnal temperature range. According to both the PCA analysis and the Cluster analysis (Figure 4), most species fall in a central group with very similar adaptations. However, *A. marginata*, *A. retusa* and *A. decora* are detached from this group and displaced in the positive direction of PC1, indicating adaptation to a relatively high winter rainfall and low summer rainfall. *Arachis villosa* and *A. burkartii* are displaced in the negative direction, indicating an adaptation to low winter rainfall and high summer rainfall. No species are particularly detached from the main group in the positive direction of PC2 indicating that many species are adapted to high temperatures, but *A. monticola* is isolated away from all other species in the negative direction indicating adaptation to lower temperatures.

The Cluster analysis shows that some taxonomic sections tend to cluster together (Figure 4), for example 100% of section Rhizomatosae and 75% of Section *Procumbentes* populations fall within Cluster 3 (Table 3). This reflects the tendency of sections to predominate in geographical areas (Figure 1) with different climates, and may reflect specific adaptation of these clusters to particular climatic conditions. Section *Erectoides* only occurs in Clusters 1 and 3, but section *Arachis* is fairly evenly distributed across all three clusters. Cluster 2 does, however, contain all three potential progenitor or genome donor species (*A. duranensis*, *A. monticola*, and *A. ipaënsis*), indicating that these species are similarly adapted. Cluster 1 is most distantly related to the

Table 3. Percentage of each section occurring in each of the three predominant clusters derived from Euclidean distance and UPGMA based on climatic variables (Figure 3).

Section	Cluster		
	1	2	3
	$n = 27$	$n = 8$	$n = 33$
Arachis ($n = 26$)	38.5 (37.0)	15.4 (50)	46.0 (37)
Caulorrhizae ($n = 2$)	100 (7.4)		
Erectoides ($n = 22$)	50.0 (40.1)		50 (33)
Heteranthae ($n = 4$)	25 (3.7)	75 (37.5)	
Procumbentes ($n = 8$)	25 (7.4)		75 (18)
Rhizomatosae ($n = 3$)			100 (9)
Triseminatae ($n = 1$)		100 (3.3)	
Trierectoides ($n = 2$)	50 (3.7)		50 (3)

The number in parentheses is the percentage of a cluster represented by a taxonomic section.

other clusters. This information could be used to target forage species for use in different environments, or to improve the climatic adaptation of the cultivated peanut for specific areas.

Species geographical range and abundance

Most species have a narrow distributional range with 15 species having a MaxD of less than or equal to 100 km and a further 12 species having a MaxD of less than 200 km (Table 2). This may be due to their geocarpic habit that severely restricts dispersal. Thirteen species have a MaxD greater than 1000 km, with *A. repens* having the greatest maxD of 3606 km. This is likely due to its spread through cultivation as a forage crop. Fifteen species have a CA_{50} of less than or equal to 2, and 41 species have a CA_{50} of less than 5, indicating very narrow geographical ranges (Table 2). Both MaxD and CA_{50} are obviously dependent upon the number of observations. An indication of abundance is given by R CA_{50}, the smaller the number, the more abundant, or more dense, the populations. Figure 5 illustrates the abundance of populations within each species geographical range, showing that population density of species varies substantially. It is important to note that it is often difficult to identify duplicates in the wild *Arachis* database, i.e., populations which have been sampled more than once by different collectors. This would bias abundance values.

Species richness

Two 2500 km^2 grids stand out for species richness (Figure 6). Both occur in Mato Grosso do Sul, Brazil. Secondary areas occur in Mato Grosso do Sul,
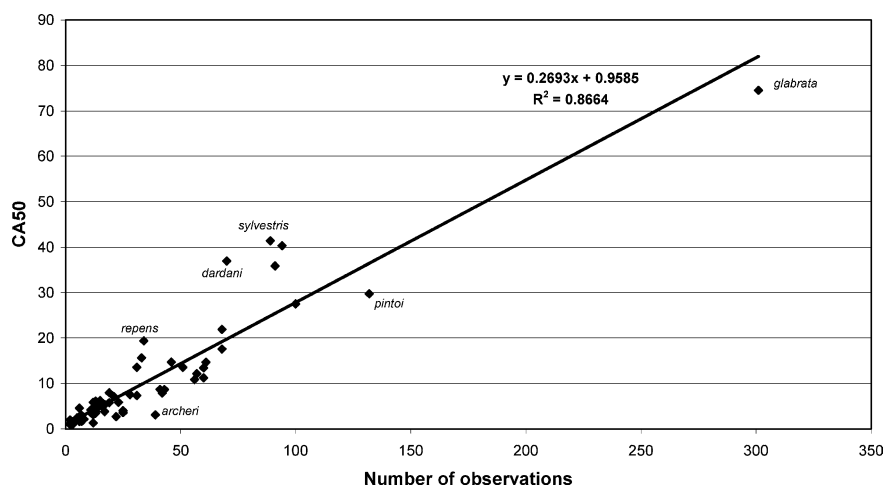
Figure 5. Graphical representation of CA_{50} versus the number of observations for each wild peanut species. A circular area with a 50 km radius was assigned to each observation, and the total area of this neighbourhood for the species was calculated. The CA_{50} area reported here is expressed relative to the area of one circle.

Matto Grosso and scattered grids in the Brazilian highlands. The high species richness in Matto Grosso do Sul is largely due to the distribution of three sections *Arachis*, *Erectoides* and *Procumbentes*. All 13 species of Section *Erectoides* are distributed in the region, 7 of 8 species of Section *Procumbentes* (excluding *A. rigonii*), and 5 of 26 wild species of Section *Arachis* are found there, with an additional 8 species found close to this centre of diversity. The largest number of species occur in the range from just below 200–300 m in altitude (Figure 7).

Catchment limitation hypothesis

The level at which each section and species is restricted to a catchment area is tabulated in Table 4. Sections *Arachis*, *Caulorrhizae* Krapov. & W.C. Gregory, *Extranervosae*, *Heteranthae*, *Procumbentes* and *Rhizomatosae* show no restrictions according to a particular river basin. Of the other sections, *Trierectoides* appears to be the most restricted to watersheds, at level 2. Sections *Erectoides* and *Triseminatae* are broadly restricted to level 1 catchments, the Paraná/Paraguay basin and northeastern Brazilian Atlantic catchments, respectively. However there is no climatic potential for the sections outside of their respective catchments, making it difficult to distinguish whether the section is restricted by the river basin itself or by climatic factors.

Twenty-four species were confined to water catchment areas at level 6, the highest level of resolution (Table 4). *Arachis marginata* Gardner was confined
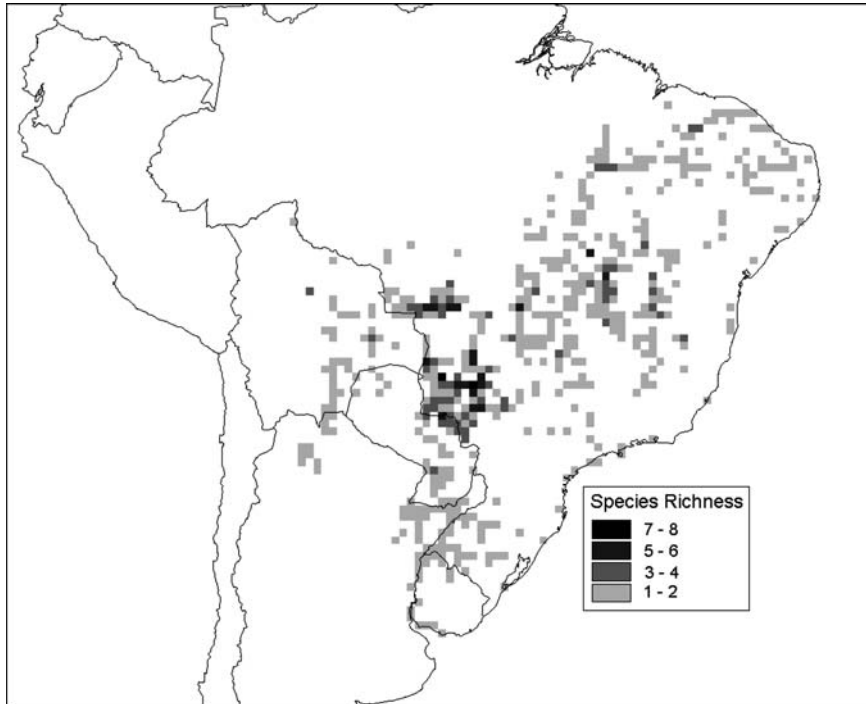
Figure 6. *Arachis* species richness in 2500 km² grid cells.

at level 5, six species were confined at level 2 and 17 species at level 1. The remaining 20 species were not even restricted to level 1 catchments. Of the 48 species with confinement at some level (71% of all species), 20 had too few
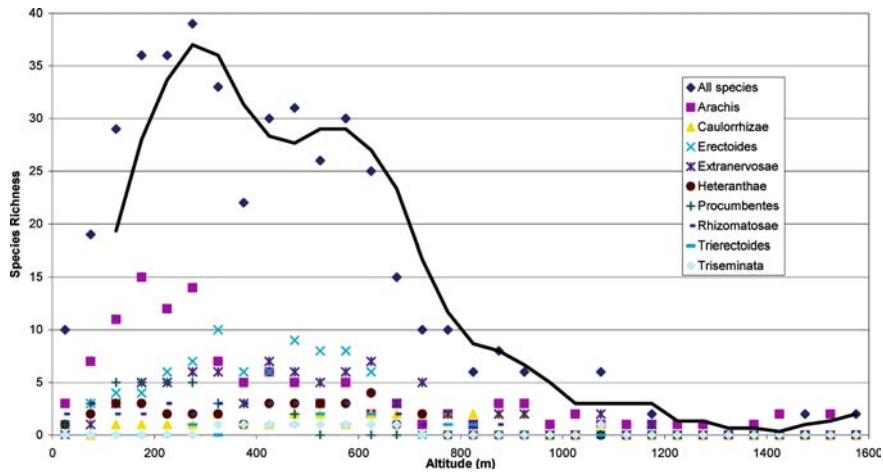

Figure 7. Number of species by section in relation to altitude.

Table 4. Catchment levels associated with each species.

	Catchment Level range	No. of catchments covered at next level up	% Climatic potential distribution lying outside catchment	Notes
Section *Arachis*	0	4	N/A	No restriction[*]
batizocoi	0	2	N/A	No restriction
benensis	6	–	N/A	Too few references
cardenasii	0	2	N/A	No restriction
correntina	1	3	0	No climatic potential outside catchment range
cruziana	6	–	N/A	Too few references
decora	2	2	39	Majority of climatic potential within catchment
diogoi	1	3	3	Majority of climatic potential within catchment
duranensis	0	2	N/A	No restriction
glandulifera	0	2	N/A	No restriction
helodes	6	–	0	No climatic potential outside catchment range
herzogii	6	–	N/A	Too few references
hoehnei	1	2	0	No climatic potential outside catchment range
ipaënsis	6	–	N/A	Too few references
kempff-mercadoi	6	–	11	Majority of climatic potential within catchment
kuhlmannii	0	2	N/A	No restriction
magna	0	2	N/A	No restriction
microsperma	6	–	N/A	Too few references
monticola	6	–	N/A	Too few references
palustris	2	4	N/A	Too few references
praecox	6	–	N/A	Too few references
simpsonii	6	–	14	Majority of climatic potential within catchment
stenosperma	0	2	N/A	No restriction
trinitensis	6	–	N/A	Too few references
valida	6	–	N/A	Too few references
villosa	0	2	N/A	No restriction
williamsii	6	–	N/A	Too few references
Section Caulorrhizae	0	3	N/A	No restriction
pintoi	0	3	N/A	No restriction
repens	0	2	N/A	No restriction
Section *Erectoides*	1	4	0	No climatic potential outside catchment range
archeri	1	2	0	No climatic potential outside catchment range
benthamii	1	2	0	No climatic potential outside catchment range
brevipetiolata	2	2	N/A	Too few references
cryptopotamica	6	–	5	Majority of climatic potential within catchment

Table 4. Continued.

	Catchment Level range	No. of catchments covered at next level up	% Climatic potential distribution lying outside catchment	Notes
douradiana	1	2	0	No climatic potential outside catchment range
gracilis	2	2	3	Majority of climatic potential within catchment
hatschbachii	6	-	N/A	Too few references
hermannii	1	2	0	No climatic potential outside catchment range
major	1	2	0	No climatic potential outside catchment range
martii	6	–	N/A	Too few references
oteroi	1	2	0	No climatic potential outside catchment range
paraguariensis	1	3	0	No climatic potential outside catchment range
stenophylla	6	–	61	Evidence for catchment limitation to distribution
Section *Extranervosae*	0	4	N/A	No restriction
burchellii	0	3	N/A	No restriction
lutescens	0	2	N/A	No restriction
macedoi	0	3	N/A	No restriction
marginata	5	2	N/A	Too few accessions
pietrarellii	0	2	N/A	No restriction
prostrata	0	3	N/A	No restriction
retusa	2	3	55	Evidence for catchment limitation to distribution
setinervosa	1	2	N/A	Too few references
villosulicarpa	1	2	N/A	Too few references
Section *Heteranthae*	0	2	N/A	No restriction
dardani	0	2	N/A	No restriction
giacomettii	6	–	N/A	Too few references
pusilla	1	4	0	No climatic potential outside catchment range
sylvestris	0	2	N/A	No restriction
Section *Procumbentes*	0	2	N/A	No restriction
appressipila	6	–	0	No climatic potential outside catchment range
chiquitana	6	–	N/A	Too few references
kretschmeri	6	–	8	Majority of climatic potential within catchment
lignosa	1	2	0	No climatic potential outside catchment range
matiensis	0	2	N/A	No restriction
rigonii	6	–	N/A	Too few references
subcoriacea	6	–	28	Majority of climatic potential within catchment
vallsii	6	–	N/A	Too few references

Table 4. Continued.

	Catchment Level range	No. of catchments covered at next level up	% Climatic potential distribution lying outside catchment	Notes
Section *Rhizomatosae*	0	3	N/A	No restriction
burkartii	0	2	N/A	No restriction
glabrata	0	3	N/A	No restriction
pseudovillosa	1	3	0	No climatic potential outside catchment
Section *Triseminatae*	1	2	0	No climatic potential outside catchment
triseminata	2	4	35	Majority climatic potential within catchment
Section *Trierectoides*	2	4	41	Majority climatic potential within catchment
guaranitica	1	2	0	No climatic potential outside catchment range
tuberosa	1	2	18	Majority climatic potential within catchment

[*] No apparent restriction of distribution by river catchment.

references (less than 10 geographically unique accessions) to determine the potential distribution based on climatic parameters, 16 had no climatic potential outside the catchment area, and 11 had the majority of climatic potential within the catchment area, making it impossible to determine whether the species was restricted by climatic factors or it was restricted by the mode of dispersal. In the case of 2 species, *A. stenophylla* Krapov. & W.C. Gregory and *A. retusa* Krapov., W.C. Gregory & Valls, there was a greater area of climatic potential outside the catchment area than inside the catchment, indicating that the observed restriction was not due to climatic parameters, and could indicate a restriction due to dispersal mechanism. However, these are just 2 species out of 48 for which the analysis was applied showing such tendencies, indicating little evidence for the importance of catchments in limiting wild peanut species distributions. The watershed data does not take into account ancient catchment areas and river basins that are likely to have changed over time, and may have affected the distribution of species.

Conclusions

The database used in this analysis, encompassing and cross-referencing all known collections of wild peanuts, provides an ideal basis for a comprehensive review of the ecogeographic distribution of wild *Arachis* species, and an opportunity to use the information, through the use of GIS tools for both the enhanced conservation and utilisation of genetic diversity of *Arachis*. Here we

have used the information to determine species geographic distributional ranges together with various climatic and altitudinal ranges associated with these distributions. Due to the characteristic geocarpy of the genus, it is not surprising that many of the species have a narrow distribution range.

We have found that the distributional range of *A. duranensis*, the proposed donor of the A genome to the cultivated peanut, is in close proximity to that of the closest wild relative of the cultigen, *A. monticola*, and overlaps the proposed donor of the B genome, *A. ipaënsis*. This region, in the eastern foothills of the Andes and the adjoining chaco regions of Argentina, Bolivia and Paraguay, is a key area for further explorations for wild *Arachis*. Other conservation priorities have been determined by Jarvis et al. (2003).

An area of particularly high species richness occurs in the state of Mato Grosso, close to the Gran Pantanal in southwest Brazil. The hypothesis that wild *Arachis* were dispersed from this region, via watersheds, and were confined to watersheds via their geocarpic seed dispersal mechanism, was investigated. Twenty species were found to have no catchment restriction to distribution at any catchment level. For the remaining 48 species, upon examining the climatic potential outside of their catchments, just 2 species (*A. stenophylla* Krapov. & W.C. Gregory and *A. retusa* Krapov., W.C. Gregory & Valls) were found to have significant areas of their potential climatic adaptation outside of the catchment within which they are restricted.

Acknowledgements

References

Bianchi-Hall C.M., Keys R.D. and Stalker H.T. 1991. Use of protein profiles to characterize peanut cultivars. Newsletter of the Association of the Official Seed Analysis 65: 25–26.

Fernandez A. and Krapovickas A. 1994. Chromosomas y evolucion en *Arachis* (Leguminosae). Bonpladia 8: 187–220.

Gaston K.J. 1996. Species richness: measure and measurement. In: Gaston K.J. (ed.), Biodiversity, a Biology of Numbers and Difference. Blackwell Science, London, UK, pp. 77–113.

Gregory W., Gregory M., Krapovickas A., Smith B.W. and Yarbrough J.A. 1973. Structure and genetic resources of peanuts. In: Peanuts – Culture and Uses. American Peanut Research and Education Association, Stillwater, Oklahoma, pp. 47–133.

Grieshammer U. and Wynne J.C. 1990. Isozyme variability in mature seeds of US peanut cultivars and collections. Peanut Science 18: 72–75.

Halward T.M., Stalker H.T., LaRue E. and Kochert G. 1991. Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome 34: 1013–1020.

Halward T.M., Stalker H.T., LaRue E. and Kochert G. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). Plant Molecular Biology 18: 315–325.

Hijmans R.J., Cruz M., Rojas E. and Guarino L. 2001. DIVA-GIS, Version 1.4. A Geographic Information System for the Management and Analysis of Genetic Resources Data. Manual. International Potato Center, Lima, Peru.

Hijmans R.J., Garrett K.A., Huamán Z., Zhang D.P., Schreuder M. and Bonierbale M. 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conservartional Biology 14: 1755–1765.

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

Hopkins M., Casa A., Wang T., Mitchell S.E., Dean R.E., Kochert G.D. and Kresovich S. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Science 39: 1243–1247.

Jarvis A., Ferguson M.E., Williams D., Mottram G., Guarino L. and Stalker H.T. et al. 2003. Biogeography of wild *Arachis*: Assessing conservation status and setting future priorities. Crop Science 43: 1110–1108.

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. and Gladkov A. 1999. FloraMap: A Computer Tool for the Distribution of Plants and Other Organisms in the Wild. CIAT, Cali, Colombia.

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

Kochert G., Stalker H.T., Gimenes M., Galgaro L., Romero Lopes C. and Moore K. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Americal Journal of Botany 83: 1282–1291.

Lacks G. and Stalker H.T. 1993. Isozyme analyses of *Arachis* species and interspecific hybrids. Peanut Science 20: 76–81.

Moretzsohn M.C. 2001. Microsatélites em *Arachis hypogaea* e especies associadas. In: Anais, III Simposio de Recursos Genéticos para a América latina e Caribe, 19–22 November 2001. IAPAR, Londrina, Brazil, pp. 81–83.

Paik-Ro O.G., Smith R.L. and Knauft D.A. 1992. Restriction fragment length polymorphism evaluation of six peanut species within the *Arachis* section. Theoretical and Applied Genetics 84: 201–208.

Raina S. and Mukai Y. 1999a. Genomic *in situ* hybridisation identifies the diploid wild progenitors of cultivated (*Arachis hypogaea*) and related wild (*Arachis monticola*) peanut species. Plant Systematics and Evolution 214: 251–262.

Raina S. and Mukai Y. 1999b. Detection of a variable number of 18S-5.8S-26S and 5S ribosomal DNA loci by fluorescent *in situ* hybridisation in diploid and tetraploid *Arachis* species. Genome 42: 52–59.

Simpson C.E., Nelson S., Starr J.L., Woodard K.E. and Smith O.D. 1993. Registration of TxAG-6 and TxAG-7 peanut germplasm lines. Crop Science 33: 1418.

Simpson C.E. and Starr J.L. 2001. Registration of 'COAN' Peanut. Crop Science 41: 918.

Smartt J. and Stalker H.T. 1982. Speciation and cytogenetics in *Arachis*. In: Pattee H.E. and Young C.T. (eds), Peanut Science and Technology. American Peanut Research and Education Society, Youkum, Texas, pp. 21–49.

Stalker H.T. 1990. A morphological appraisal of wild species in section *Arachis* of peanuts. Peanut Science 17: 117–122.

Stalker H.T. 1992. Utilizing *Arachis* Germplasm Resources. In: Nigam S.N. (ed.), Groundnut – A Global Perspective. Proceedings of an International Workshop. 25–29 November 1991. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India, pp. 281–295.

Stalker H.T., Beute M.K., Shew B.B. and Barker K.R. 2002a. Registration of two root-knot nematode-resistant peanut germplasm lines. Crop Science 42: 312a–313a.

Stalker H.T., Beute M.K., Shew B.B. and Isleib T.G. 2002b. Registration of five leafspot-resistant peanut germplasm lines. Crop Science 42: 314–316.

Stalker H.T., Ferguson M.E., Valls J.F.M., Pittman R.N., Simpson C.E. and Bramel P. 2000. Catalog of *Arachis* Germplasm Collections. http://www.icrisat.org/text/research/grep/homepage/groundnut/arachis/start.htm.

Stalker H.T. and Lynch R.L. 2002. Registration of four insect-resistant peanut germplasm lines. Crop Science 42: 313–314.

Stalker H.T. and Moss J.P. 1987. Speciation, cytogenetics, and utilisation of *Arachis* species. Advances in Agronomy 41: 1–40.

Stalker H., Phillips T., Murphy J.P. and Jones T.M. 1994. Variation of isozyme patterns among *Arachis* species. Theoretical and Applied Genetics 87: 746–755.

Tombs M.P. 1963. Variant forms of arachin. Nature 200: 1321–1322.

USGS Hydro1k South America 2000. Drainage Basins, http://edcdaac.usgs.gov/gtopo30/htdro/sa_basins.html

World Conservation Monitoring Centre 2000. Global Protected Areas Database V4.0, WCMC, Cambridge, UK.