Nordic Journal of Botany 29: 001–013, 2012
doi: 10.1111/j.1756-1051.2011.01105.x,

Subject Editor: Jens Christian Svenning. Accepted 6 October 2011

Modelling the geographic distributions of endemic genera in the eastern Central Asian desert

Song-Mei Ma, Ming-Li Zhang, Jian Ni and Xi Chen

S.-M. Ma, M.-L. Zhang (zhangml@ibcas.ac.cn) and X. Chen, Key Lab of Biogeography and Bioresources in Arid Land, Xinjiang Inst. of Ecology and Geography, Chinese Academy of Sciences, PR-830011 Urumqi, PR China. SMM also at: Graduate Univ., Chinese Academy of Sciences, CN-100039 Beijing, PR China. MLZ also at: Inst. of Botany, Chinese Academy of Sciences, CN-100093 Beijing, PR China. – J. Ni, State Key Lab of Environmental Geochemistry, Inst. of Geochemistry, Chinese Academy of Sciences, CN-550002 Guiyang, PR China.

This study simulates the distributions of 13 endemic and near-endemic genera (*Ammopiptanthus*, *Sympegma*, *Iljinia*, *Elachanthemum*, *Potaninia*, *Tugarinovia*, *Kaschgaria*, *Sarcozygium*, *Timouria*, *Zollikoferia*, *Stilpnolepis*, *Synstemon* and *Tetraena*) to indicate areas of plant diversity and conservation importance within the eastern Central Asian desert, and to identify the determinant environmental variables contributing to the spatial distribution patterns. Using known distribution localities and 14 environmental variables, the Maxent and Domain species distribution models were employed to map the patterns of geographic distribution. The power of predictability of the models was tested using the receiver operating characteristic method and the jackknife validation approach, according to the different number of species localities available. The estimated richness and the superimposed potential distributions of 13 genera were used to indicate endemic patterns of distribution. The comparison of Maxent and Domain further identified previously unknown areas of endemism and described the distribution for each taxon. Both observed species occurrence and the species occurrence predicted from the Maxent indicated that the eastern Alashan of Inner Mongolia is the most noticeable endemic area, and the northwestern and northern Tarim Basin of Xinjiang is the secondary center of plant diversity. These regions were then prioritized for conservation importance. Potential evapotranspiration ratio and precipitation seasonality played important roles in driving the observed patterns of endemic distribution.

Knowledge about the geographical distribution of species is critical for addressing numerous issues in biogeography, ecology, and evolution, and is essential for conservation and management of biodiversity (Margules and Pressey 2000, Guisan and Thuiller 2005). However, in many regions detailed data on the distribution of most taxa are lacking, and collecting such data is both costly and labor intensive (Ottaviani et al. 2004). Existing data collections are always incomplete and biased, so they cannot be taken to represent accurate range limits of species (Graham et al. 2004). Obtaining such distributional knowledge, therefore, strongly depends on species distribution models (SDMs).

SDMs combine points of known occurrence with suites of environmental variables to infer the ecological niche requirements of a species, and can provide detailed predictions of distribution. Consequently, SDM is becoming an indispensable tool for providing a measure of a species' potential occupancy in areas not covered by previous botanical surveys (Guisan and Thuiller 2005), and for estimating patterns of species distribution and informing conservation strategies (Ortega-Huerta and Peterson 2004).

Conservation biologists often pay attention to endemic and/or rare taxa. Endemic taxa with geographic distributions restricted to specific geographic regions, are potentially most sensitive to habitat perturbation and thereby especially vulnerable to extinction (Linder 1995, Peterson and Watson 1998). Modelling potential distributions of the endemic taxa can detect previously unknown areas of endemism, and further establish the detailed distributions (Manrique et al. 2003). In previous works, SDMs have been employed to reveal areas of high endemism, to evaluate regional concentrators of species richness, and to set conservation priorities (Peterson et al. 2000, Manrique et al. 2003, Harris et al. 2005, Murray-Smith et al. 2009, Young et al. 2009). However, these studies did not fully identify the environmental variables determining the observed pattern of endemic distribution, which could produce information about the relationships between taxa and environments as well as the ecological requirements of taxa.

In the eastern Central Asian desert (ECAD) 13 endemic or near-endemic genera with isolated taxonomic positions and speculated ancient origins occur, and most of them are considered as descendent species of the Tethys flora (e.g. *Sympegma regelii*, *Iljinia regelii* and *Zollikoferia polydichotoma*), or of the paleo-tropic climate in the Tertiary (e.g. *Ammopiptanthus mongolicus* and *A. nanus*, *Tetraena* *mongolica* and *Potaninia mongolica*) (Fu 1992, Zhao and Zhu 2003). Previous studies of these 13 endemic genera have focussed on their taxonomic position and genetic diversity or biology (Zhao 2000, Ge et al. 2003, Zhu et al. 2003, Fang et al. 2004, Chen et al. 2009). Zhao and Zhu (2003) presented the known geographical distribution limit for each genus. However, the specific distribution areas and potential distributions are unknown. Potential distribution, indicated by the SDM as regions of potential presence, may be occupied by closely related species, or may represent suitable areas to which the species has failed to disperse or in which it has gone extinct (Anderson et al. 2003).

In this study, two SDMs are employed to determine the potential distributions of 13 genera based on known localities and 14 environmental variables. We also aim to use the derived distributions of endemic and endemic-threatened or rare species to indicate areas of plant diversity and conservation importance within the ECAD and to identify the importance of these predictive variables on the spatial distribution patterns of endemic taxa in the models.

Hence, our specific objectives are to address the following questions: 1) what is the potential geographical distribution of each endemic taxon? 2) what are the superimposed distribution patterns of the 13 genera? 3) what are the determinant environmental variables contributing to the observed endemic distribution patterns in the models? and 4) what are the implications of our results to nature conservation?

Material and methods

Study area

The ECAD (75°–115°E, 35°–50°N) (Zhao and Zhu 2003) is located in the central part of Central Asia (Fig. 1). It includes the driest northwestern part of China (the whole Xinjiang Autonomous Region, the Qaidam Basin of Qinghai Province, the Hexi corridor of Gansu Province, and the Alashan and western Ordos plateaus of Inner Mongolia Autonomous Region), the Tianshan Mountains between China and Kyrgyzstan, and the Junggar, Transaltai, Alashan, and the east Gobi of Mongolia. The specific areas mentioned above are divided into two physical geographical areas described as the central Gobi and the Alashan.

ECAD is the driest area of the temperate zone in the Northern Hemisphere, and the temporal–spatial distribution of precipitation is variable. The maximum precipitation is 100–200 mm annually in the Alashan, and 40–60 mm around the Tarim Basin, while the lowest is 30–40 mm in the central Gobi. Moreover, evaporation in this area is extremely high. The maximal potential evapotranspiration ratio is 15–25 in the central Gobi, 5–20 around the Tarim Basin, and 2–5 in the Alashan. Vegetation in such arid conditions has formed very special features (drought tolerant and avoiding) to adapt to this extreme environment (Pan et al. 2001). Xerophytic small trees (e.g. *Haloxylon* spp.), shrubs (e.g. *Calligonum*, *Ephedra*, *Nitraria* and *Tamarix*), semi-shrubs and dwarf shrubs (e.g. *Anabasis*, *Artemisia*, *Ceratoide*, *Reaumuria*, *Salsola* and *Sympegma*), and succulent halophytic dwarf semi-shrubs (e.g. *Kalidium* and *Suaeda*) are dominant in this desert region, companied with some low grasses (e.g. *Stipa gobica*) and annual or biennial ephemerals appearing only in early spring.

Desert endemic species

There are seventeen desert xerophytic species in the 13 endemic genera belonging included in this study (*Ammopiptanthus*, *Sympegma*, *Iljinia*, *Elachanthemum*, *Potaninia*, *Tugarinovia*, *Kaschgaria*, *Sarcozygium*, *Timouria*, *Zollikoferia*, *Stilpnolepis*, *Synstemon* and *Tetraena*), and they belong to seven families (Table 1). These endemic genera

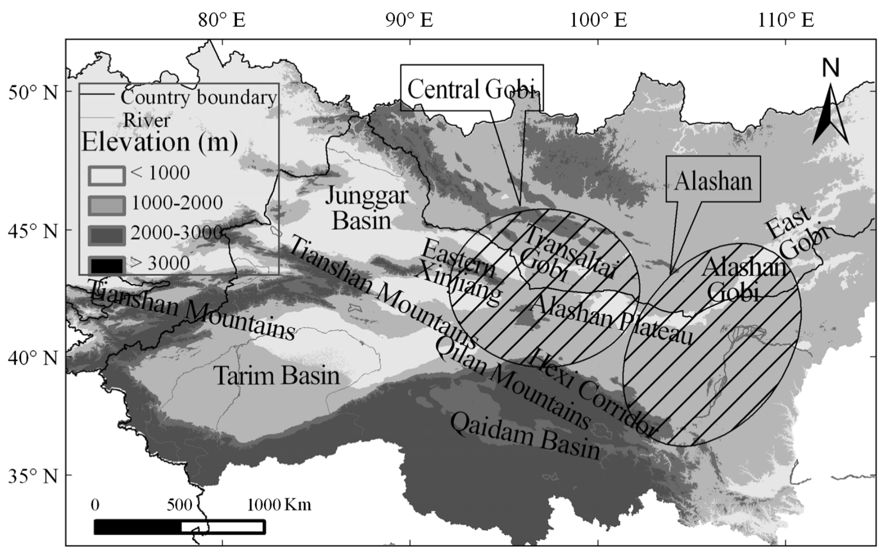

Figure 1. Map of the study area showing elevational relief (i.e. the area used for modelling) in eastern Central Asian desert. This includes the Tianshan Mountains, Junggar, Tarim and Qaidam Basins, Central Gobi (the hatched areas: the eastern Xinjiang, the western Yumen of Gansu Province, the eastern Alashan Plateau of Inner Mongolia, and the Transaltai Gobi of Mongolia), and the Alashan (the hatched areas: the middle and eastern Hexi corridor of Gansu Province, the Alashan left and right Banners and western Ordos of Inner Mongolia, and the Alashan Gobi of Mongolia).

Table 1. Simple descriptions and the number of known localities of the thirteen endemic genera. ECAD = eastern Central Asian desert.

Family	Species	Endemism	Life form	Localities
Leguminosae	*Ammopiptanthus mongolicus*	Restricted to east Alashan of ECAD	Evergreen shrub	27
	A. nanus	Restricted to the borders between China and Kyrgyzstan	Evergreen shrub	8
Chenopodiaceae	*Sympegma regelii*	Near endemic to ECAD	Deciduous semi-shrub	39
	Iljinia regelii	Near endemic to ECAD	Deciduous semi-shrub	64
Compositae	*Elachanthemum intricatum*	Near endemic to ECAD	Annual forb	109
	Kaschgaria brachanthemoides *K. komarovii*	Near endemic to ECAD	Deciduous semi-shrub	48
	Tugarinovia mongolica *T. ovatifolia*	Near endemic to ECAD	Perennial forb	39
	Zollikoferia polydichotoma	Endemic to the desert of northwestern China	Perennial forb	16
	Stilpnolepis centiflora	Restricted to south Alashan-Ordos of northwestern China	Biennial forb	13
Graminae	*Timouria saposhnikovii*	Endemic to mountain area of ECAD	Perennial grass	34
Rosaceae	*Potaninia mongolica*	Restricted to Alashan-east Gobi of ECAD	Deciduous small shrub	28
Cruciferae	*Synstemon petrovii* *S. deserticolus*	Restricted to south Alashan-Ordos of northwestern China	Annual forb	9
Zygophyllaceae	*Tetraena mongolica*	Restricted to east Alashan-west Ordos of northwestern China	Deciduous small shrub	8
	Sarcozygium xanthoxylon	Near endemic to ECAD	Deciduous shrub	42

prefer habitats with diluvial piedmont, gravel on low sloping hillsides, or desert or gravel Gobi terrain. Six genera (*Ammopiptanthus*, *Tugarinovia*, *Synstemon*, *Potaninia*, *Tetraena* and *Stilpnolepis*) are restricted to varying degrees to certain areas, whereas the remaining seven exhibit relatively widespread distributional ranges.

In the past decades, most of them have rapidly declined mainly due to free grazing, firewood collection, and especially, mining and urbanization, with no regard to their value in maintaining the desert. Seven genera have become threatened or endangered, and four have been given protected status in the Chinese red data book (Fu 1992). *Ammopiptanthus mongolicus* and *A. nanus* were defined as threatened species in the third and first conservation priority, respectively. The rare *Tugarinovia* was placed in the category of the first conservation priority while *T. mongolica* and *Potaninia mongolica* got the second conservation priority. Recently, *Elachanthemum intricatum* and *Sarcozygium xanthoxylon* were also listed as the first and second conservation priority, respectively, by the Xinjiang government, and *S. centiflora* was listed as a rare and endangered species by the government of Inner Mongolia.

Distribution data

For this research, we conducted extensive field investigations for five of the 13 genera across northwestern China, from June 2008 to September 2009. These included six species: *Ammopiptanthus mongolicus*, *A. nanus*, *T. mongolica*, *Z. polydichotoma*, *S. regelii* and *S. xanthoxylon*. The investigated localities were then used as the distribution data for the following modelling for the above six taxa.

Locality records for the remaining eight endemic genera were extracted mainly from the Chinese Virtual Herbarium (<www.cvh.cn>), including all specimen information from local herbaria in China. Additional species distribution records from Mongolia, Kazakhstan and Kyrgyzstan were obtained from the 'Plantae Asiae Centralis' (Grubov 1966) and other publications (Zhao 2000, 2002, Zhao and Zhu 2003, Zhu et al. 2003). If specific geographic coordinates were not provided for a locality, we used gazetteers and Google Earth to assign geographical coordinates to these records. All herbarium specimens were then revised and quality controlled by scientists with expertise in the species' distributions (Acknowledgements).

Environmental data sets

Nineteen climate variables at a spatial resolution of 30 arc seconds (0.93 × 0.93 km resolution) were obtained from WORLDCLIM (<www.worldclim.org>). They represent a combination of annual trends, seasonality and extreme environmental conditions (Hijmans et al. 2005). To reduce the similarity between climate layers, a principal components analysis (PCA) was employed to create new axes that summarized the variation in fewer (independent) dimensions. Collinearity was examined using a Pearson correlation matrix, where subsets of variables with a high average correlation (> 90%) were reduced to a single variable to represent this group, and the other collinear variables were removed (Ward 2007). Three PCA axes consistently explained ~90% of the variation within climate data, which always represented temperature variables in the first PCA axis and precipitation variables in the second. Finally, a total of 12 variables (Table 2) were retained for following modelling for each SDM and taxon.

Further, we derived the thirteenth variable, potential evapotranspiration (PET) ratio (PER), following Anderson et al. (2002). The PER is estimated as the mean annual biotemperature (°C) divided by total annual precipitation (mm) and multiplied by an empirically derived constant of approximately 60 (Holdridge et al. 1971). Biotemperature is the mean unit temperature range from 0–30°C, and all unit-period values outside this range are given a substitution zero to reflect that extreme high and low temperatures inhibit the physiological activity of plants (Holdridge et al. 1971). At a PER of 1.0, potential evapotranspiration approximately equals total precipitation for the long-term

Table 2. Fourteen environmental layers and their descriptions or data sources.

Climate layers	Description
Annual mean temperature Isothermality Temperature seasonality (standard deviation × 100) Min temperature of coldest month Mean temperature of driest quarter Mean temperature of coldest quarter	Isothermality means the mean diurnal range (Mean of monthly (max temp–min temp)) divided by the temperature annual range (Monthly (max temp–min temp)) and then multiplied by 100
Annual precipitation Precipitation seasonality (CV) Precipitation of wettest month Precipitation of driest month Precipitation of wettest quarter Precipitation of driest quarter Potential evapotranspiration ratio (PER)	The coefficient of variation (CV) is the standard deviation of the weekly precipitation estimates expressed as a percentage of the mean of those estimates (i.e. the annual mean)
Elevation	USGS (http://edc.usgs.gov/products/elevation/gtopo30/hydro/asia.html)

average year. Values above 1.0 indicate increasing aridity whreas values below 1.0 indicate increasing humidity.

Additionally, we used elevation as the fourteenth variable (Table 2). Continental-scale distributions of species are principally determined by climate (Pearson and Dawson 2003), therefore we did not include other environmental variables, such as soil type.

Potential distribution

Numerous SDMs have been successfully used in ecology and conservation (Guisan and Thuiller 2005). In this study, two different modelling methods were used: Maxent (Phillips et al. 2006) and Domain (Carpenter et al. 1993). The two methods determine the potential niches for species occurrence based on environmental conditions where species should occur. They differ, however, in the approach to find these conditions. Maxent estimates a target probability distribution by finding the probability distribution of maximum entropy subject to a set of constraints that represent the incomplete information about the target distribution. The output models produce predictions in the form of real numbers between 0 and 100, representing cumulative probability of occurrence. Domain uses a point-to-point similarity metric (based on the Gower distance statistic) to assess new sites in terms of their environmental similarity to the closest (most similar) sites of known species presence (Carpenter et al. 1993). The occurrence probability calculated in Domain expresses as an index of habitat suitability on a continuous scale (0–100).

To estimate the potential distribution of each endemic taxon, Maxent ver. 3.2.19 was used with recommended default parameters for all runs (Phillips et al. 2006). Domain was implemented in the DIVA-GIS ver. 5.4 software. The two species of the genus *Ammopiptanthus* (*A. mongolicus* and *A. nanus*), were modelled separately because of their obviously disjunct distributions. The remaining three bi-species genera: *Synstemon*, *Kaschgaria* and *Tugarinovia*, were simulated at the genus level because the two species in each genus are closely related. Thus, the 13 genera were modelled as 14 taxa (Table 1).

Data for the nine taxa (ECAD endemic) with more than 25 locality records (the rule according to Pearson et al. 2007), was randomly sampled to obtain a dataset of roughly 75% of the localities for training and the remaining 25% for model evaluation. The data were divided in this way 10 times for each of the nine taxa. Ten partitions were made to assess the variability of model results (Phillips et al. 2006). The presence locality data set aside for evaluation was merged with equivalent amounts of 'pseudo-absence' data randomly selected from the background pixels, and the subsequent data entered into a receiver operating characteristic (ROC) plot to calculate the area under the ROC curve (AUC) (Fielding and Bell 1997). The theoretically perfect result is AUC = 1, whereas a model no better than random yields AUC = 0.5. To aid the presentation of model results, we apply the threshold value at which the sum of sensitivity and specificity is maximized (Cantor et al. 1999, Manel et al. 2001, Liu et al. 2005).

For the remaining four taxa (northwestern China endemic) and *A. nanus*, with available localities less than 25, the jackknife procedure (Pearson et al. 2007) was used to evaluate models. Using this approach, one locality point was removed from the total observed localities (n), and the model was built using the remaining n − 1 localities. Thus, for a species with n localities, n individual models were built for testing. Model accuracy and significance were evaluated based on the ability of each model to predict the one excluded test locality as present. The 'lowest presence threshold' (LPT) was selected to distinguish 'suitable' from 'unsuitable' areas, i.e. identifying the pixels that represented areas of habitat that are at least as suitable as those where the species is known to occur (Pearson et al. 2007).

Determinant environmental variables analysis

To identify the variables determining the distribution patterns of each endemic taxon in ECAD, we examined the contribution of each environmental variable to the models generated by Maxent. This value is provided by Maxent, which quantifies and sums the contribution of every environmental layer in each model of its training runs. We examined minimum, maximum, mean and standard deviation of the contribution of these determinant environmental variables (which provided summed contributions over 70%) to the 10 partitions or n individual models (differing by species depending on the number of localities available).

The zonal statistics routine (ArcMap 9.2) was used to extract from the potential map the values of determinant variables in the localities where the taxon is known or predicted to be present and in 1000 random locations generated with a random point generator in ArcMap. Furthermore, these values were also used in determining the ecological requirements of each taxon.

Endemic distribution patterns analysis

To reveal the observed distribution patterns, the richness (the total number of endemic genera occurring together in 1 × 1 km grid cells, regardless of frequency) of the 13 endemic taxa was estimated using DIVA-GIS (estimators of richness) based on all known occurrence records. Similarly, according to Anderson et al. (2002), the superimposed potential distributions were used to explore the modelled patterns of endemic distribution.

Results

Geographic distributions predicted by Maxent and Domain

Model evaluation for both Maxent and Domain shows high scores of performance (AUCs > 0.94) (Table 3). Maxent was the better performing method for five of the nine taxa (ECAD endemic), with higher AUC values. The predicted geographic distributions are shown in Fig. 2a–i (Maxent) and Fig. 3a–i (Domain). There are important differences in the outputs of the two models. For example, the best model for the widespread species *S. regelii*, following the AUC values, is Maxent, which predicted a wider potential distribution than Domain did. For *E. intricatum*, both methods indicated a concentric distribution around the arid Alashan Plateau (Fig. 1), but the Maxent predictions gave more reliable results, as indicated by the higher AUC and by specialist validation based on field knowledge (the produced distribution in the grassland area in east Mongolia and Inner Mongolia by Domain is unlikely for the species). Moreover, *P. mongolica* also displayed a remarkable difference between the methods, with Maxent producing more reasonable distributions than Domain did. Maxent generated suitable habitats that were either closely matched to or surrounding the observed distribution sites. However, for *Kaschgaria* and *I. regelii*, Domain predicted more observed localities as present and obtained higher AUC values than Maxent, and also correctly removed most of the potential distributions of the latter in the northwestern Tarim Basin.

Table 3. AUC score and the standard deviation (SD) of the ten partition models for each of the nine taxa.

	AUC (SD)	
Taxon	Maxent	Domain
A. mongolicus	0.997 (± 0.001)	0.997 (± 0.001)
P. mongolica	0.989 (± 0.004)	0.975 (± 0.003)
Tugarinovia	0.989 (± 0.002)	0.995 (± 0.006)
I. regelii	0.942 (± 0.007)	0.963 (± 0.012)
S. regelii	0.963 (± 0.010)	0.959 (± 0.014)
S. xanthoxylon	0.971 (± 0.007)	0.966 (± 0.012)
Kaschgaria	0.967 (± 0.008)	0.982 (± 0.018)
E. intricatum	0.952 (± 0.006)	0.949 (± 0.006)
T. saposhnikovii	0.966 (± 0.007)	0.943 (± 0.034)

Jackknife tests demonstrated high and significant number of successes (localities correctly predicted as present, Table 4) for each of the remaining five taxa. These taxa also present considerable differences between method outputs (Fig. 2j–n, 3j–n). According to our field experience, *Z. polydichotoma* is mainly distributed around the Tarim Basin of Xinjiang and Hexi corridor in the Gansu Province. Maxent reproduced this trend, but Domain produced some unlikely distributions in the western Inner Mongolia and southern Mongolia. Additionally, Maxent produced wider and/or more significant distributions for the other four taxa than Domain did. For example, the produced potential distribution for *A. nanus*, restricted to the border region between China and Kyrgyzstan, and also dispersed toward the north and south around the Tianshan and Kunlun Mountains.

Environmental variables determining the potential distribution

The determinant variables with the summed contributions to the potential distribution over 70% were demonstrated for each taxon (Table 5). The distribution patters of different endemic taxa were determined by different environmental variables. However, precipitation seasonality (CV) (bio_15) and PER had the greatest contribution in identifying areas of occurrence for most endemic genera in ECAD. In addition, precipitation of the wettest and driest quarter (bio_16, 17), temperature seasonality (standard deviation × 100) (bio_4), mean temperature of the coldest month (bio_6), mean temperature of the coldest quarter (bio_11), elevation, and annual precipitation (bio_12) also played important roles in driving the observed patterns of distribution.

Specifically, temperature seasonality had considerable effects on the observed distribution patterns of *Tugarinovia* and *T. saposhnikovii*. Precipitation seasonality gave almost half of the contribution (45.31%) to *Kaschgaria* distribution. The determinant variables for *S. centiflora* were PER and mean temperature of the coldest quarter, for *T. mongolica* they were annual mean temperature (bio_1) and PER, and for *P. mongolica* it was PER and elevation.

Ecological niche of each endemic taxon

A digital display of distribution influenced by concrete and tangible environmental parameters forms a basis for gaining insights about the ecological requirements of each taxon (Table 5). Within the regions where species are known to occur, or have a potential to occur, the majority of habitats such as for *A. mongolicus* had a min temperature for the coldest month between −17.5 and −16.1°C, and precipitation of the wettest quarter between 104–116 mm. Most suitable habitats for *T. mongolica* had an annual mean temperature range between 7.2–8.2°C, and a PER range between 3–4. The same variables for *Systemon* are an annual precipitation between 26–98 mm, and min temperature of coldest month between −19 and −3.1°C. The variation range of determinant variables of the above taxa, with relatively restricted range, is low.

In contrast, the widespread taxa could survive a broad combination of temperature and precipitation conditions. For example, the majority of suitable habitats for *I. regelii* had a PER range between 3 and 20, and mean temperature of the coldest quarter between −18.2 and −7.3°C, and for *S. regelii* mean temperature of the coldest quarter ranged from −15.2 to −2.6°C, and precipitation of the driest quarter ranged from 3 to 12 mm.

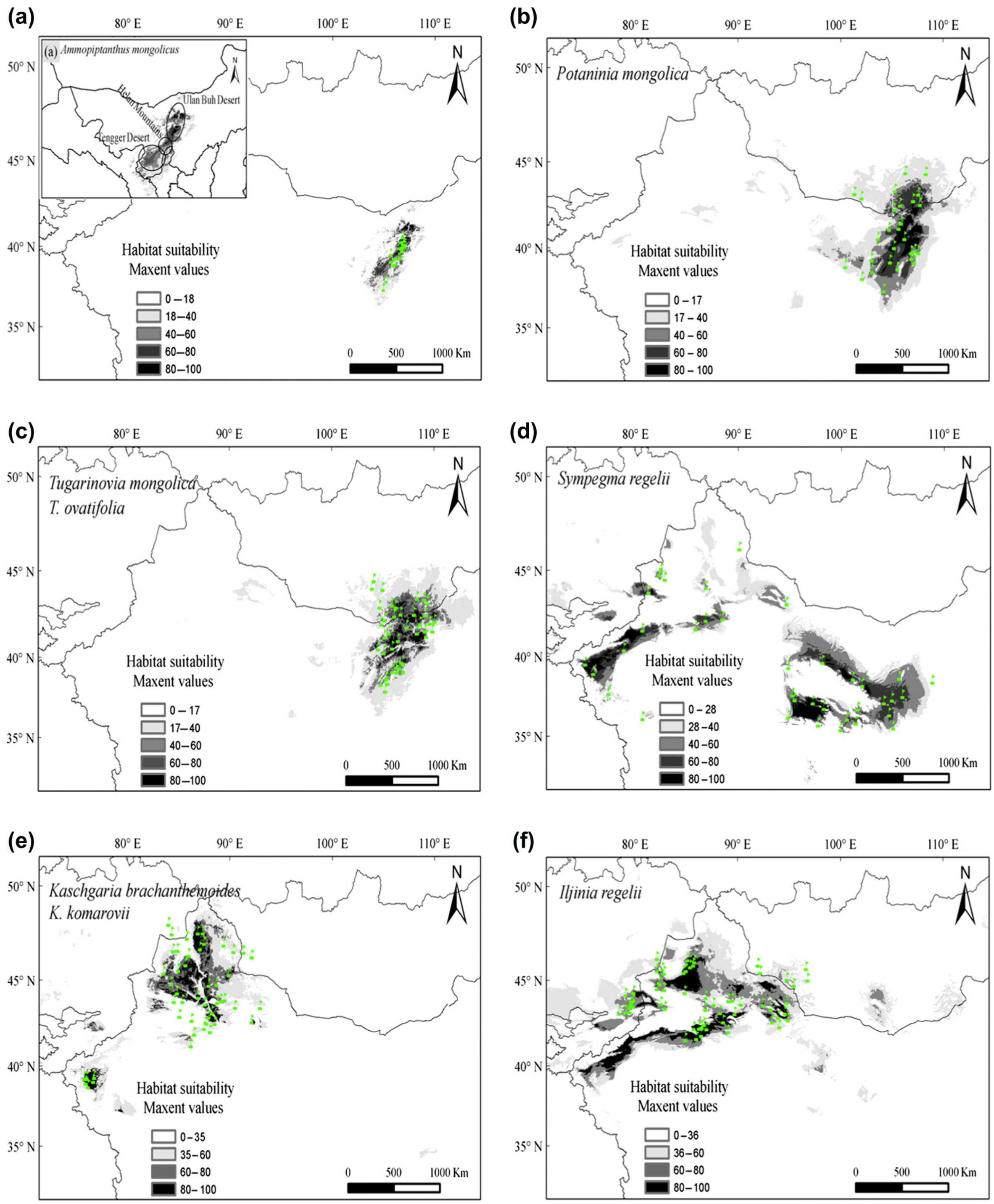

Figure 2. Final potential distributions of the 14 taxa according to Maxent, display in geographic space by the application of threshold at which the sum of sensitivity and specificity is maximized (a)–(i) or the lowest presence threshold (LPT, see "Potential distribution" section) (j)–(n), and these for *Ammopiptanthus mongolicus* were magnified and shown in a (a). The flags indicate the available occurrence records (range from 8 to 109) for each taxon.

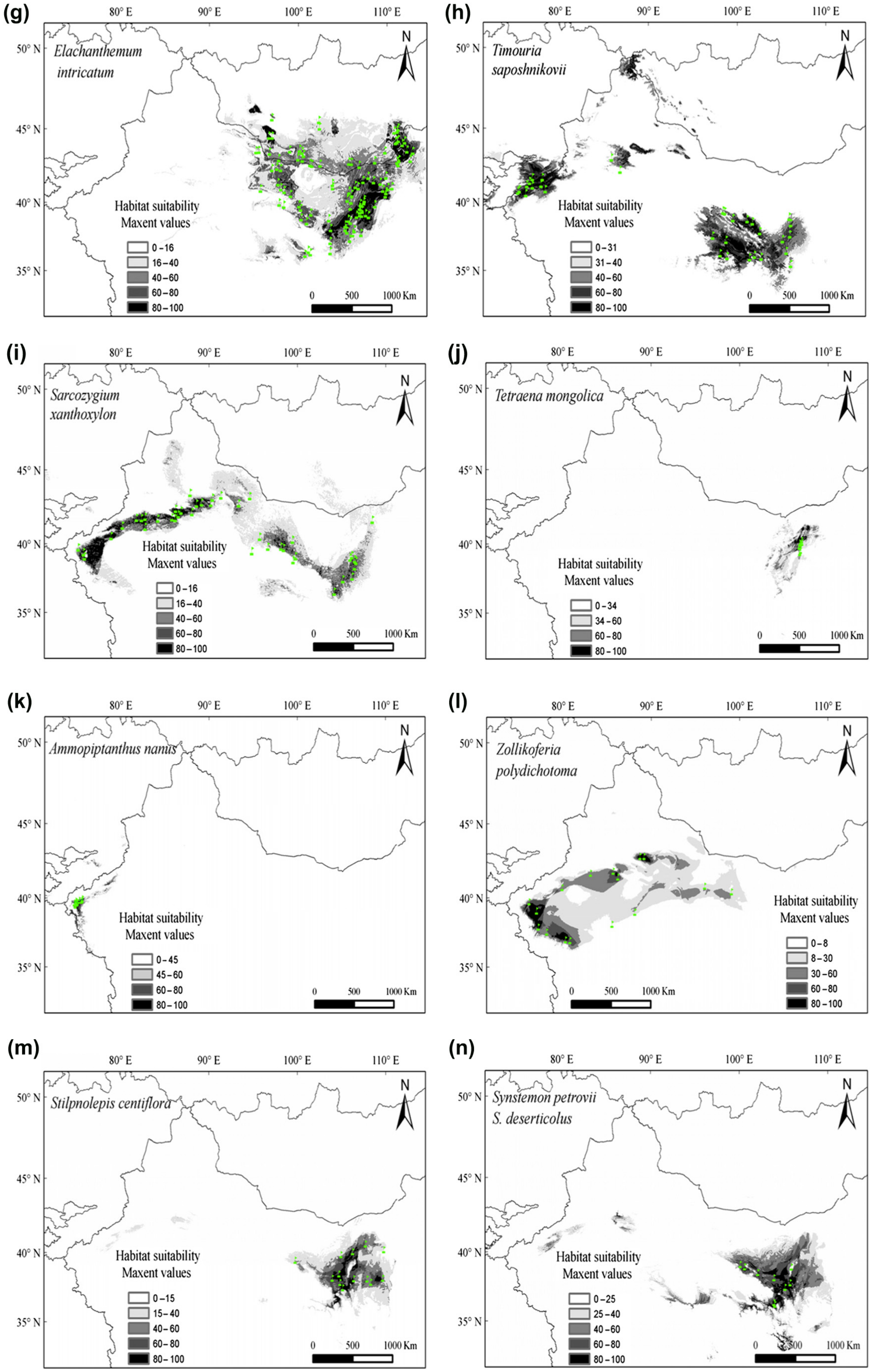

Figure 2. (Continued).

Modelled versus observed geographical distribution patterns

Observed and modelled distribution patterns according to Maxent of the endemic genera are shown in Fig. 4a–b. The richness based on species collections illustrates that notable endemism is restricted to the eastern Alashan, with four to seven genera that were indicated to focus their distributions on a total of thirteen grid cells. The isolated grids in the westernmost and northern Tarim Basin, eastern Xinjiang, and the Hexi corridor of Gansu Province possessed distributions of five to six endemic genera. Moreover, species richness presents a largely concentric distribution in the east around the Alashan Plateau, and a concentrated distribution in the west around the northwestern and northern Tarim Basin.

The superimposed potential distributions generally demonstrated similar patterns to the estimated richness.

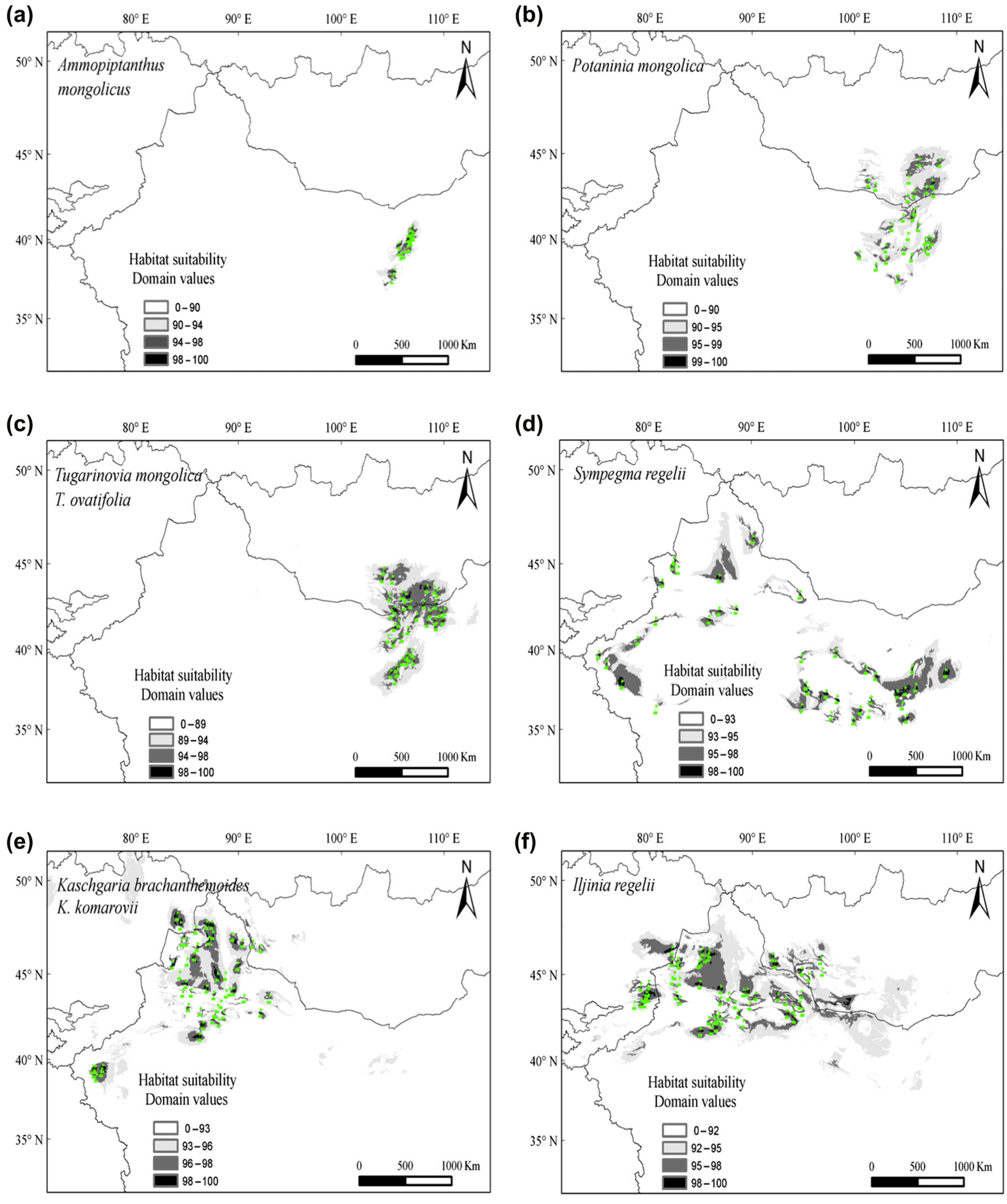

Figure 3. Final potential distributions of the 14 taxa according to Domain, display in geographic space by the application of threshold at which the sum of sensitivity and specificity is maximized (a)–(i) or the lowest presence threshold (LPT, see "Potential distribution" section) (j)–(n). The flags indicate the available occurrence records (range from 8 to 109) for each taxon.

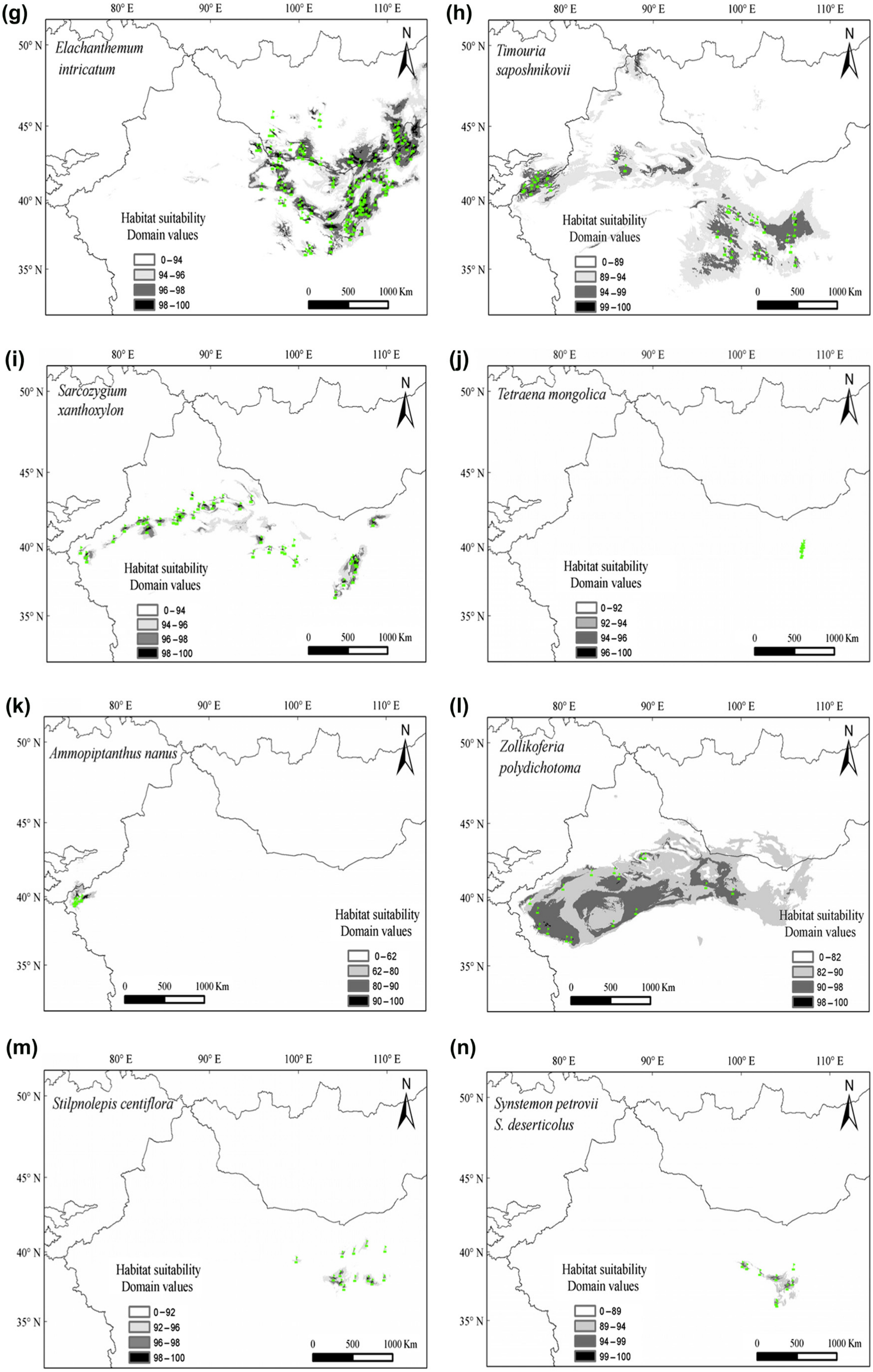

Figure 3. (Continued).

Table 4. Results of the jackknife validation method of model testing for the five taxa, showing the sample size used for modeling, and the data used to calculate the p-values. [a] = locality sample size, [b] = number of successes, [c] = mean fractional predicted area (the proportional predicted areas of northwestern China desert).

	Maxent				Domain		
Taxon	LSS[a]	NOS[b]	MFPA[c]	p-value	NOS[b]	MFPA[c]	p-value
Tetraena mongolica	8	7	0.0647	$8.49E^{-7**}$	6	$2.64E^{-06}$	0.039^{*}
Ammopiptanthus nanus	8	8	0.0263	$1.21E^{-9**}$	8	0.0032	$6.73E^{-6**}$
Synstemon	9	8	0.2321	0.0250^{*}	8	0.0015	$1.08E^{-5**}$
Stilpnolepis centiflora	13	13	0.2134	$1.18E^{-11**}$	13	$5.94E^{-03}$	0.0404^{*}
Zollikoferia polydichotoma	16	15	0.2056	$1.01E^{-10**}$	16	0.3475	$0.8\ E^{-4**}$

The eastern Alashan was indicated as having the highest environmental suitability for the tested endemic genera. Additionally, the superimposed map identified some complementary areas with the highest suitable environmental conditions, including parts of the northwestern Tarim Basin, and the easternmost portions of the Hexi corridor in Gansu Province. However, the northern areas of the Alashan Plateau which belong to the Central Gobi, was suggested to

Table 5. Determinant environmental variables (DEV) (bio = bioclim. The number corresponds to these 19 variables in WORLDCLIM) reflect the summed contributions over 70% to the potential distribution of each taxon (the relatively widespread taxa are shown in bold character), and their minimum (Min), maximum (Max) and mean values as well as standard deviation (SD). bio_1 = annual mean temperature, bio_4 = temperature seasonality (standard deviation × 100), bio_6 = min temperature of coldest month, bio_11 = mean temperature of coldest quarter, bio_12 = annual precipitation, bio_15 = precipitation seasonality (CV), bio_16 = precipitation of wettest quarter, bio_17 = precipitation of driest quarter. For full species names see Table 1 and 4.

Endemic taxa	DEV	Contributions (SD)	Min	Max	Mean	SD
A. mongolicus	bio_6°C	33.37 (± 3.96)	–17.50	–16.10	–16.20	0.43
	bio_16 mm	27.43 (± 0.99)	104.00	116.00	106.00	3.01
	bio_15	21.03 (± 0.37)	85.00	110.00	97.00	7.16
A. nanus	bio_4	33.98 (± 1.01)	9.83	10.89	10.33	0.12
	bio_15	21.74 (± 1.26)	50.00	68.00	59.56	2.11
	bio_16 mm	20.53 (± 1.48)	62.00	80.00	75.00	2.37
T. mongolica	Bio_1°C	37.62 (± 2.40)	7.20	8.20	8.00	1.00
	PER	35.92 (± 1.52)	3.00	4.00	4.00	0.01
Synstemon	bio_12 mm	39.32 (± 2.35)	26.00	98.00	57.30	8.59
	bio_6°C	23.07 (± 1.03)	–19.00	–3.10	–14.60	1.90
	PER	21.38 (± 1.00)	1.00	4.00	2.00	0.89
Z. polydichotoma	bio_4	38.78 (± 1.97)	4.20	5.39	4.41	0.11
	bio_6°C	25.76 (± 3.83)	–14.60	–9.30	–11.08	0.85
	bio_12 mm	23.80 (± 0.08)	21.00	74.00	47.93	10.99
S. centiflora	bio_11°C	32.45 (± 2.32)	–7.80	–4.60	–5.80	8.76
	PER	28.31 (± 3.05)	2.00	5.00	3.00	0.76
	bio_4	15.62 (± 1.17)	10.12	12.07	10.68	0.57
P. mongolica	PER	35.78 (± 1.97)	3.00	5.00	4.00	0.55
	Elevation m	31.16 (± 3.83)	937.00	1691.00	1216.00	186.62
	bio_12 mm	9.80 (± 0.08)	89.00	135.00	114.00	17.46
Tugarinovia	bio_4	39.69 (± 4.20)	11.29	12.94	12.15	0.38
	bio_15	17.97 (± 1.85)	98.00	115.00	105.02	4.25
	bio_16 mm	13.42 (± 2.81)	60.00	154.00	86.98	23.08
S. regelii	bio_11°C	32.45 (± 2.32)	–15.20	–2.60	–7.60	2.21
	bio_17 mm	19.37 (± 3.05)	3.00	12.00	7.00	1.89
	bio_15	15.31 (± 1.17)	35.00	94.00	73.00	14.85
I. regelii	PER	24.34 (± 5.01)	2.00	20.00	7.00	2.80
	Elevation m	21.45 (± 3.77)	–149.00	3095.00	1329.00	526.39
	bio_11°C	13.95 (± 0.88)	–18.20	–7.30	–11.00	1.63
Kaschgaria	bio_15	45.31 (± 4.05)	33.00	76.00	46.00	8.50
	bio_17 mm	18.88 (± 2.54)	3.00	23.00	12.00	4.09
	Elevation m	17.57 (± 2.09)	–56.00	1654.00	867.00	446.85
S. xanthoxylon	PER	29.37 (± 2.61)	2.00	18.00	10.00	3.84
	bio_17 mm	26.01 (± 2.44)	3.00	11.00	6.00	1.71
	bio_11°C	20.31 (± 4.38)	–10.10	–2.40	–4.80	1.84
E. intricatum	bio_1°C	31.58 (± 2.37)	2.10	10.90	4.81	1.85
	Elevation m	29.16 (± 1.01)	851.00	1883.00	1221.69	177.69
	bio_17 mm	23.38 (± 2.11)	2.00	7.00	3.60	0.87
T. saposhnikovii	bio_4	41.03 (± 4.2)	8.09	11.88	10.33	8.54
	bio_15	16.97 (± 1.85)	58.00	103.00	80.00	10.40
	bio_16 mm	13.42 (± 2.81)	53.00	269.00	128.00	29.97

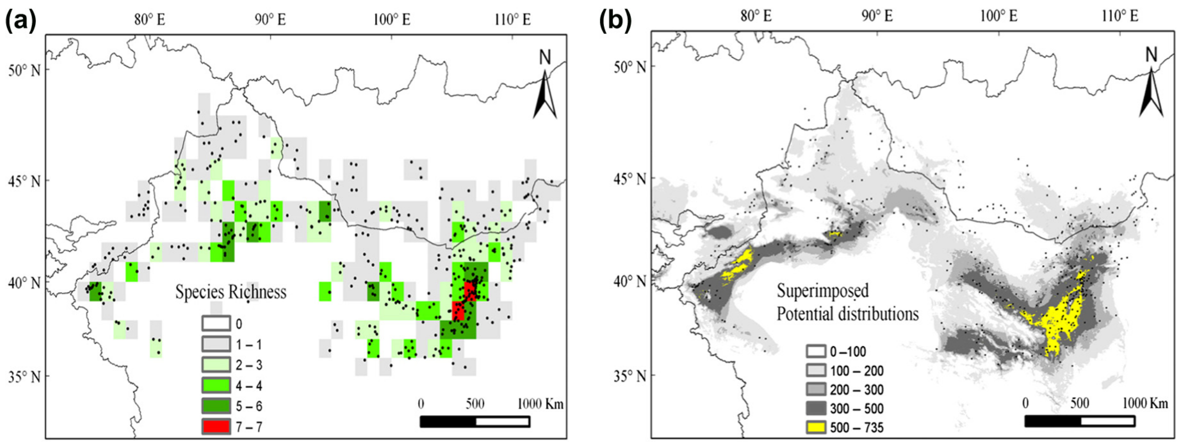

Figure 4. Grid cell (1 km × 1 km) richness (a) and superimposed potential distribution patterns of 13 endemic genera produced by Maxent (b), portrayed in summed species records or environmental suitabilities. Black dots indicate the available occurrence localities of all the taxa.

be of low suitability or unsuitable for the establishment of these endemic taxa.

Discussion

Significance of the predicted endemic distributions

Potential distribution areas with great environmental similarity to extant sites were identified for a total of 14 taxa (Fig. 2a–n, 3a–n). This based on the observed ecological niche of each taxon, and thus avoids biases and helps to compensate for the lack of available collection records. Moreover, the comparison between the Maxent and Domain helps in clarifying our present understanding of the distribution of each taxon.

Both Maxent and Domain restrict species with specific environmental affinities, and help to determine and display small areas of endemism. For example, Maxent indicated two isolated areas in the northwestern Transaltai Gobi (i.e. a part of the Mongolian Altai mountain area) of Mongolia for *E. intricatum*, and Domain identified the most suitable habitats for *Kaschgaria* in the eastern Kazakhstan. The two methods generated predictions that also further delineated the range of potential distribution for each taxon. For example, the models predicted areas of distribution in Mongolia, Kazakhstan and Kyrghizstan for the widespread taxa *S. xanthoxylon*, *T. saposhnikovii* and *S. regelii*. This is consistent with Manrique et al. (2003) suggesting that widespread species, with ill-defined distributions based on known data, can be provided with better defined distribution limits by employing SDMs.

Relationship between taxa potential distributions and environments

The potential ecological niche for each tested taxon elaborately reflects the affinities between distribution and the current environmental conditions (Fig. 2a–n, 3a–n). In the case of *A. mongolicus*, specifically, the highest suitable identified areas were restricted to the Ulan Buh Desert of Inner Mongolia, and around the Helan Mountains (Fig. 2a). However, the Tengger Desert, in the neighborhood of the areas mentioned above, was indicated as of secondary suitability (with the suitability values = 60–80). Interestingly, the revealed subtle differences of environmental conditions among the Ulan Buh Desert/Helan Mountains and the Tengger Desert are consistent with genetic divergence between two lineages specific to the two areas inhabited by *A. mongolicus* (Chen et al. 2009). Moreover, the present distribution in the Ulan Buh Desert have been speculated to be a result of rapid expansion of the species from a nearby refugium. The predictions support the above scenario in suggesting that this area harbor preferable environmental conditions for *A. mongolicus*.

Recognition of areas of endemism is dependent on scale. However, it is important to identify areas with multiple endemic taxa (Crisp et al. 2001). The species richness and superimposed potential distributions provide an interesting comparison (Fig. 4a–b). The eastern Alashan was identified as the area with most established endemic genera and possesses the highest level of environmental suitability. Six endemic taxa (*A. mongolicus*, *P. mongolica*, *T. mongolica*, *S. centiflora*, *Tugarinovia* and *Systemon*) have their actual and potential distributions principally restricted to this area. Apart from *I. regelii* and *Kaschgaria*, the remaining five genera also have extensive distributions in this area. Hence, our results support that the Alashan desert region is the distribution centre of endemic genera of Central Asian desert (Zhao and Zhu 2003).

The eastern Alashan endemism area is characterized by an annual temperature between 2.8 and 10.3°C, a total annual precipitation between 100 and 200 mm, and moderate PER ranges from 2 to 5. Therefore, it is possible that the relatively optimal combination of medium temperature and greater precipitation in this area favored the majority of the endemic taxa, especially the six ones mentioned above.

Furthermore, a small part of the northwestern Tarim Basin was identified as one of the most suitable areas for endemic distribution (Fig. 4b). This area is surrounded by the Tianshan (in the north) and Kunlun Mountains (in the south), with complex climatic–physical conditions, which thus provides diverse habitats for the tested endemic taxa.

Conversely, the Central Gobi, which possess minimal and scattered distributions of some genera (e.g. *S. xanthoxylon* distributed over the eastern Xinjiang and the western Yumen of Gansu Province), was not revealed as a possible suitable area (Fig. 4b). This area is located in the centre of the Asian arid zone, the driest area of the desert region, and receives ca 30–40 mm in total annual precipitation. Moreover, the PER, ranging from 15 to 25, indicates a strong evaporation level in this area. These adverse environments might exclude most plants from the area, and only some typical Gobi components and desert species have sporadic occurrences.

Conservation implications

The distribution maps generated in the present study provide useful guidance for conservation biologists (Fig. 2a–n, 3a–n, 4a–b). A comparison between the Maxent and Domain methods pinpoint areas of greatest suitability (i.e. habitat suitability values = 80–100 or 98–100) for each taxon. These areas call for further sampling to clarify the absence or presence of each taxon, and to indicate the most likely conditions for maximizing the success of ex situ conservation. Such areas, for example for *A. mongolicus* and *T. mongolica* are all restricted to small areas of the northeastern Alashan (according to the Maxent), and for *I. regelii* are distributed over small areas of eastern Kazakhstan, northern and southern Junggar Basin, and the southeastern parts of Mongolia (according to Domain). Moreover, for in situ conservation, these suitable areas for each taxon would be subject to fencing against human disturbance, which is important to conserve extant plants and natural habitats.

Areas estimated to contain most of the endemic taxa are eastern Alashan, the western tip of the Tarim Basin, the northern Hexi corridor in Gansu Province, Middle Tianshan, and eastern Xinjiang (in two red and eleven dark green grids; Fig. 4a). Nineteen bright green grids in small areas of the eastern Alashan (spilling south into southern Mongolia), the Hexi corridor in Gansu Province, northwestern Tarim Basin and Middle Tianshan contained the second highest endemic richness. The yellow areas prioritize conservation importance by the superimposed predictions (Fig. 4b) are very similar to the above areas identified by the estimated richness. Consequently, to conserve a considerable amount of the overall endemic diversity, these above areas are the first priority for conservation.

The governments of Inner Mongolia and Ningxia have set aside three national natural reserves: the Helan Mountains, the western Ordos and the Shapotou (in Ningxia), and four endemic taxa (*A. mongolicus*, *T. mongolica*, *P. mongolica* and *Tugarinovia*) inhabit these protected areas. However, the locations of existing reserves do not cover the habitat needs of all endemic genera. Our findings demonstrate that except for the areas around the eastern Alashan, the majority of the extent of endemism areas identified (mentioned above) is unprotected. Legal rights and sufficient financial supports that complement national reserve systems will be necessary to fully protect these endemic taxa.

Acknowledgements – This study was financed by CAS Important Direction for Knowledge Innovation project (no. KZCX2-EW-305), and Xinjiang Inst. of Ecology and Geography, Chinese Acad. Sci. We thank Prof. Yizhi Zhao and Borong Pan for reviewing the distribution data. We are very grateful to Prof. Stewart C. Sanderson for his useful comments and English corrections to this manuscript.

References

Anderson, R. P. et al. 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. – Oikos 98: 3–16.
Anderson, R. P. et al. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. – Ecol. Modell. 162: 211–232.
Cantor, S. B. et al. 1999. A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. – J. Clin. Epidemol. 52: 885–892.
Carpenter, G. et al. 1993. DOMAIN: a flexible modeling procedure for mapping potential distributions of plants, animals. – Biodivers. Conserv. 2: 667–680.
Chen, G. Q. et al. 2009. Genetic structure and mating system of *Ammopiptanthus mongolicus* (Leguminosae), an endangered shrub in northwestern China. – Plant Spec. Biol. 24: 179–188.
Crisp, M. D. et al. 2001. Endemism in the Australian flora. – J. Biogeogr. 28: 183–198.
Fang, H. T. et al. 2004. A study on flower biology of endangered plant *Ammopiptanthus mongolicus*. – Guihaia 24: 478–480, in Chinese with English abstract.
Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. – Environ. Conserv. 24: 38–49.
Fu, L. G. 1992. China plant red data book. – Science Press, in Chinese.
Ge, X. J. et al. 2003. Genetic variation in the endangered Inner Mongolia endemic shrub *Tetraenamongolica* Maxim. (Zygophyllaceae). – Biol. Conserv. 111: 427–434.
Graham, C. H. et al. 2004. New developments in museum-based informatics and applications in biodiversity analysis. – Trends Ecol. Evol. 19: 497–503.
Grubov, V. I. 1966. Plantae Asiae Centralis. – Nauka Press.
Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.
Harris, G. M. et al. 2005. Refining biodiversity conservation priorities. – Conserv. Biol. 19: 1957–1968.
Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.
Holdridge, L. R. et al. 1971. Forest environments in tropical life zones: a pilot study. – Pergamon Press.
Linder, H. P. 1995. Setting conservation priorities: the importance of endemism and phylogeny in the southern African orchid genus *Herschelia*. – Conserv. Biol. 9: 585–595.
Liu, C. R. et al. 2005. Selecting thresholds of occurrence in the prediction of species distributions. – Ecography 28: 385–393.
Manel, S. et al. 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. – J. Appl. Ecol. 38: 921–931.
Manrique, C. E. et al. 2003. Phytogeographic analysis of taxa endemic to the Yucatan Peninsula using geographic information systems, the domain heuristic method and parsimony analysis of endemicity. – Divers. Distribr. 9: 313–330.
Margules, C. R. and Pressey, R. L. 2000. Systematic conservation planning. – Nature 405: 243–253.
Murray-Smith, C. et al. 2009. Plant diversity hotspots in the Atlantic coastal forests of Brazil. – Conserv. Biol. 23: 151–163.

Ortega-Huerta, M. A. and Peterson, A. T. 2004. Modeling spatial patterns of biodiversity for conservation prioritization in northeastern Mexico. – Divers. Distribr. 10: 39–54.
Ottaviani, D. et al. 2004. Two statistical methods to validate habitat suitability models using presence–only data. – Ecol. Modell. 179: 417–443.
Pan, X. L. et al. 2001. Plant flora geography and the resources utilization of the northwest arid and desert area. – Science Press, in Chinese.
Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – Global Ecol. Biogeogr. 12: 361–371.
Pearson, R. G. et al. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. – J. Biogeogr. 34: 102–117.
Peterson, A. T. and Watson, D. M. 1998. Problems with areal definitions of endemism: the effects of spatial scaling. – Divers. Distribr. 4: 189–194.
Peterson, A. T. et al. 2000. Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. – Biol. Conserv. 93: 85–94.
Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Modell. 190: 231–259.
Ward, D. F. 2007. Modelling the potential geographic distribution of invasive ant species in New Zealand. – Biol. Invas. 9: 723–735.
Young, B. F. et al. 2009. Using spatial models to predict areas of endemism and gaps in the protection of Andean slope birds. – Auk 126: 554–565.
Zhao, Y. Z. 2000. The classification and geographical distribution of *Tugarinovia*. – Acta Bot. Bor-Occid Sin. 20: 873–875, in Chinese with English abstract.
Zhao, Y. Z. 2002. The area and floristic geographic element of *Potaninia*. – Acta Bot. Bor-Occid Sin. 22: 43–45, in Chinese with English abstract.
Zhao, Y. Z. and Zhu, Z. Y. 2003. The endemic genera of desert region in Central Asia. – Acta Bot. Yunnan. 25: 113–121, in Chinese with English abstract.
Zhu, Z. Y. et al. 2003. A new species of *Elanchanthemum* and discussion on the classification and evolution of the genus. – Bull. Bot. Res. 23: 147–153, in Chinese with English abstract.