

# Quantitative biogeography in the South America highlands—recognizing the Altoandina, Puna and Prepuna through the study of Poaceae

Lone Aagesen<sup>a,\*</sup>, Claudia A. Szumik<sup>b</sup>, Fernando O. Zuloaga<sup>a</sup> and Osvaldo Morrone<sup>a</sup>

<sup>a</sup>*Instituto de Botánica Darwinion, Labardén 200, CC 22, San Isidro B1642HYD, Argentina;* <sup>b</sup>*Instituto Superior de Entomología, Instituto Miguel Lillo, Miguel Lillo 205, 4000 San Miguel de Tucumán, Argentina*

Accepted 19 November 2008

## Abstract

The distribution data of 340 grass species sampled in a region of 53,219 km<sup>2</sup> in the northwestern corner of Argentina (between ~21°S and ~24°S) were analyzed to search for concordance in species distributions by using the program NDM/VNDM. Here, the traditional biogeographic hypothesis proposed for the region is evaluated for the first time by using a quantitative method and an optimal criterion specifically developed within the context of areas of endemism. Three different grid sizes (0.5° × 0.5°, 0.35° × 0.35° and 0.2° × 0.2°) were used to analyze three nested data sets: species found in the Andes of Argentina, Bolivia and/or Chile; Andean distributed species; and all grass species found in the study region. The main areas supported by the analyses correspond generally to the traditional biogeographic hypothesis proposed for the region. Local distribution patterns defined by species restricted to the study region were best supported under the small grid sizes, while the bigger grid sizes recovered areas defined by species with a broader distribution. The local distribution patterns emerged in all the analyses even when widespread species were added to the data set.

© The Willi Hennig Society 2009.

When developing their grid-based method to establish areas of endemism, Szumik and Goloboff (2004) stated that concordant species distribution could be due to several causes, such as historical, physical or ecological barriers, considering that their method could be seen as an ecological method as much as a historical one. This point of view encouraged an exploration of the approach in small-scale analyses of regions where species distribution may be determined by ecological factors rather than by historical ones.

Szumik et al. (2002; Szumik and Goloboff, 2004) developed their method as an alternative to commonly used methods such as parsimony and clustering (e.g. Morrone, 1994; Linder, 2001), which have an inability to recover partial endemic areas or to interpret non-endemic species as if they were overlapping (Szumik et al., 2002). An area of endemism was defined as any

area (group of cells in a grid) in which two or more species, not found anywhere else, co-occur. An optimality criterion, the area score, was simultaneously developed in order to compare areas and give the best scores to those defined by several well fitting species and a lower score to areas defined by fewer and/or loosely fitting species. Species appearing outside the area were initially considered widespread and did not add to the area score (Szumik et al., 2002). Several relaxed criteria were later developed to account for partially matching species distributions, by accepting that species may also occur in cells neighbouring the endemic area (Szumik and Goloboff, 2004).

Hitherto, this purely pattern-based optimality criterion has been employed only within the context of areas of endemism for which it was developed (e.g. Szumik and Roig-Juñent, 2005; Domínguez et al., 2006; Lizarralde and Szumik, 2007; Díaz Gómez, 2007). Here, we use this quantitative method to analyze species distribution in a local context while relaxing the criterion of endemism; hence, most species used in the

\*Corresponding author:

E-mail address: laagesen@darwin.edu.ar

present analysis are also found outside the study region. Rather than searching for areas of endemism, we searched for concordant species distribution patterns within the region. This paper is focused on prominent patterns shared by several species in order to compare them to previous published phytogeographic studies of the same region.

The northwestern province of Argentina, Jujuy, the complex topography of which harbours a wide array of classic biogeographic units, was chosen as the study region. The three Andean highland units, the Altoandina, the Puna and the Prepuna, were the main concern of this study (Cabrera and Willink, 1973; Cabrera, 1976). A new analysis of the data set compiled by Morrone et al. (2005) that includes all grass species found in the study region, was made by searching for conspicuous distribution patterns among grass species. The resulting areas were compared with pre-established commonly used biogeographic units, and analytical aspects such as robustness to parameter choice and sample density are discussed.

### The study region

The Andean Province of Jujuy, located in the extreme northwestern corner of Argentina between  $\sim 21^\circ\text{S}$  and  $\sim 24^\circ\text{S}$ , was selected as the study region (Figs 1 and 2). Jujuy includes all three high Andean units evaluated in this study as well as other biogeographic units such as Yungas and Chaco that may interfere with the chosen units (Fig. 1). The topography of the area is complex with high peaks in the lowlands and deep valleys

intersecting the highlands. This rugged landscape creates a framework for the analysis where several of the cells defined by the chosen grid include very uneven altitudes (see transects in Fig. 1). The variable topography implies that the delimitation of the distribution patterns is not straightforward, as several cells may include more than one distribution pattern; hence, overlapping patterns are to be expected.

The grass flora in the chosen region is rich and diverse, with approximately 30% of the grass species and 50% of the grass genera known in Argentina represented here, in only 0.02% of the Argentinean territory (Zuloaga et al., 1994, 1999). The flora of this region has been well sampled through several past and recent major collections (e.g. Ruthsatz and Movia, 1975; Cabrera, 1957; Morrone et al., 2005).

### Traditional phytogeographic division of the study region

Within Argentina, the most commonly used phytogeographic scheme was developed by Cabrera (1957, 1976). Five of Cabrera's phytogeographic units are found within the study region, roughly divided along the altitude gradient. The highest slopes of the Andean mountains (above approximately 4400 m) are occupied by Altoandina vegetation, where grass steppes are the principal vegetation type. The grass steppes are followed by an open bush steppe of Puna vegetation mainly formed by the families Solanaceae, Fabaceae and Asteraceae (3400–4500 m). Below 3400 m, the Puna changes into Prepuna vegetation that is also dominated by bushes, but with a prominent presence of the families

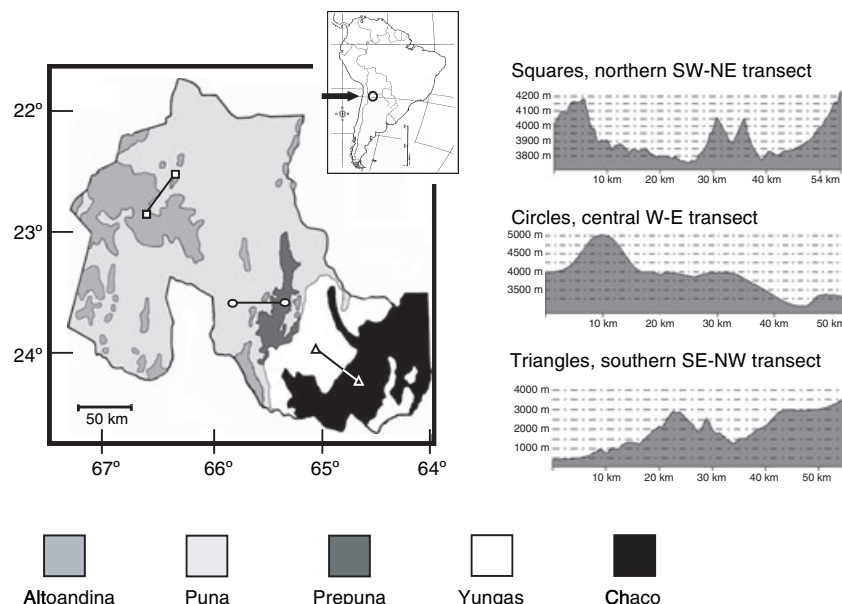


Fig. 1. The study region showing the traditional biogeographic units according to Cabrera and Willink (1973), redrawn after Morrone et al. (2005). Three 50 km transects show the variable topography in the study region.

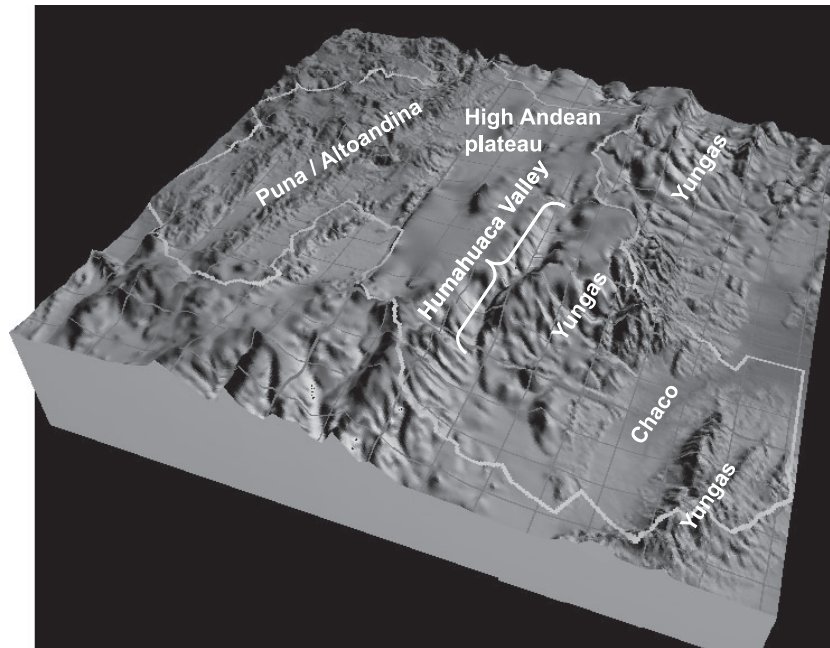


Fig. 2. 3D topographic image of the study region, showing several biogeographic units *sensu* Cabrera and Willink (1973) as well as the Humahuaca valley with Prepuna vegetation and the High Andean plateau occupied by Puna vegetation.

Cactaceae and Bromeliaceae. The eastern slopes of the Andes mountains (500–3500 m) are covered by moist forest of Yungas vegetation, whereas the lowest part of the study region is covered by xerophytic Chaco vegetation. Each vegetation type is described in more detail in Appendix 1.

## Material and methods

### *Specimen list*

We used the specimen list published by Morrone et al. (2005), adding species from the literature of *Aristida* (Sulekic, 2003), *Poa* (Negritto and Anton, 2000), *Nassella* (Torres, 1997a) and *Jarava* (Torres, 1997b). Subdivisions below species level were ignored. Introduced species were deleted from the list and redundant localities were removed. The final list includes a total of 340 species and approximately 1960 specimens. All locations were geo-referenced using GPS or maps and satellite photographs processed in the program Global Mapper ver. 6.06 (<http://www.globalmapper.com>).

The number of records per square kilometer ( $0.037 \text{ records/km}^2$ ) roughly corresponded to the sampling density used in other published studies (for example  $\sim 0.016 \text{ records/km}^2$  in Crisp et al. (2001);  $\sim 0.002 \text{ records/km}^2$  in Domínguez et al. (2006)). One may be concerned that this sampling density was inadequate for a biogeographic study. However, our data set represents the grass collections of three major

herbaria of Argentina with extra data extracted from several other herbaria as well as from the literature in the case of critical species. The sampling reflects present day knowledge of grass species distribution within the study region. Grasses have been collected selectively in the study region for decades by grass specialists. Thus, the absence of a specific grass species in a given cell is most likely to imply real absence and not just poor sampling. Furthermore, the present study is explorative, and focused on analytic abilities with no intention of modifying pre-established biogeographic units.

The sampling density is shown in Fig. 3, where the number of species per cells is shown for the smallest cell size used in the present study ( $0.2^\circ \sim 20 \text{ km}$ ). Sampling density was improved by the fill option (see Area analysis and Fig. 4). Coordinates were recorded in degrees and minutes and transformed into plane coordinates, which were used as input in the program NDM/VNDM.

All information concerning species distribution discussed in the text, for example whether a species is restricted to Argentina or found in other countries as well, was obtained from the “Catálogo de las Plantas Vasculares de la República Argentina I” (Zuloaga and Morrone, 1996; Zuloaga et al., 2008).

### *Area analysis*

The final list was divided into three nested data sets. In order to recognize both local patterns and

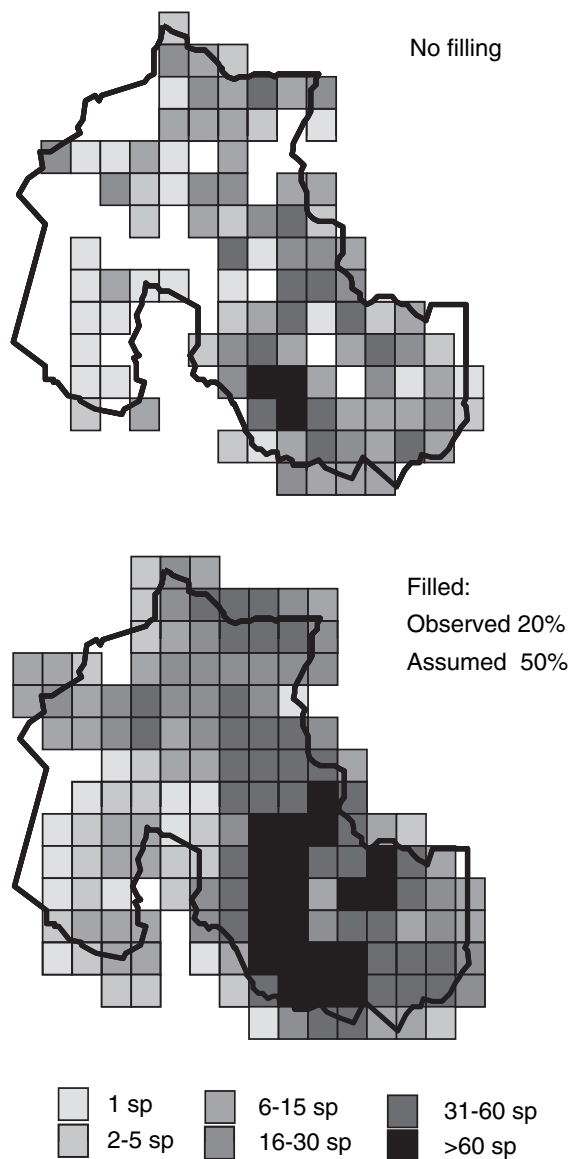


Fig. 3. Sampling density is shown as number of species per cell for the smallest cell size used in the present study ( $0.2^\circ \approx 20$  km). Sampling density was improved by the fill option.

widespread patterns simultaneously, both local and widespread species must be included in the analysis. However, widespread species have commonly been considered a problem for biogeographic analyses of areas, as these species may blur the pattern established by more locally distributed species. Initially, we included only local Andean grass species found exclusively in Argentina, Bolivia, and Chile (the ABC data set of 104 species). In the next stage, we added widespread Andean species (the AND data set including 176 species). Finally, we analyzed all grass species found in the study region, including species with an eastern South American distribution, broad American

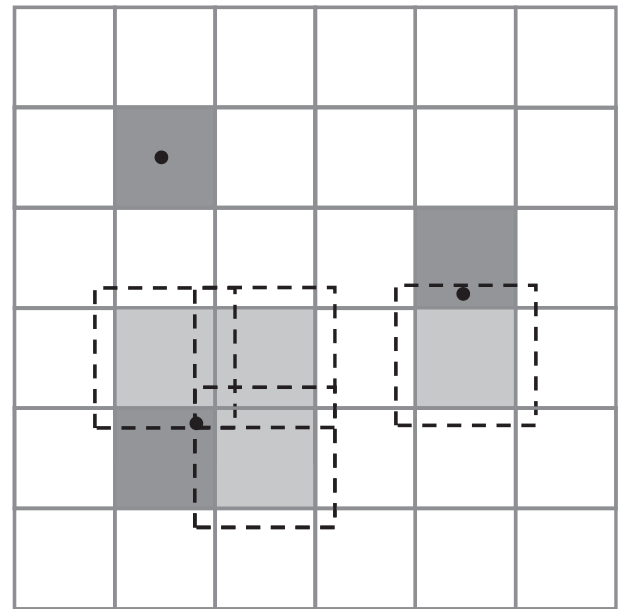


Fig. 4. Under the fill option 20% presence of a given species applies to all cells in which the species was directly observed as well as to neighbouring cells in which the species was observed within a distance of 20% of the cell length. Black dots—observed species. Dark grey—cells in which the species is directly observed. Light grey—cells in which the species is recorded as observed according to the fill option. Dashed line—the 20% fill zone.

distribution and species found in other continents as well (the TOT data set of 340 species). We explore how the added species affect the pattern obtained, by either appearing in several areas or by occurring in the transition zones between the areas.

When analyzing the data set, NDM/VNDM converts the plane coordinates into grids of species presence/absence (Szumik et al., 2002). Analyses applying parsimony or UPGMA normally use only one grid cell size (the most frequent being  $1^\circ \times 1^\circ$  or  $2^\circ \times 2^\circ$ ), although there is no formal argument to use only one. Here we analyze the data matrix of 340 species and 1960 specimens with three grid cell sizes:  $0.5^\circ \times 0.5^\circ$ ,  $0.35^\circ \times 0.35^\circ$  and  $0.2^\circ \times 0.2^\circ$  (approximately  $50 \times 50$ – $20 \times 20$  km). All the grids have the same latitudinal and longitudinal origins. It is important to note that when using different grid sizes but the same latitudinal and longitudinal origins on the grid, the individual sample locations are re-allocated in each analysis; thus, two locations previously included in the same cell may fall into distinct cells when the grid size is changed. Areas that do not survive a change in grid size may therefore be a simple artefact of a specific grid. Conversely, resistance to grid size changes becomes an important factor when evaluating a specific distribution pattern.

The optimality criterion (Szumik et al., 2002; Szumik and Goloboff, 2004) that is used here to evaluate areas (set of cells) takes into account the classical concept of



areas of endemism, i.e. an area that has two or more taxa found there and nowhere else (see for example Platnick, 1991). The concept of areas of endemism is not associated with a specific type of causal factor, but only with the existence of a common one; if a single factor affects the distribution of numerous groups of organisms at the same time, the distributions of those organisms are expected to show similar patterns, regardless of whether the causal factor is historical or ecological (Szumik and Goloboff, 2004). The criterion (implemented in NDM/VNDM, Goloboff, 2004) assigns a score to a species based on how well its distribution appears as “endemic” in a given area (set of cells) and this depends on how the limits of the area are defined (Szumik and Goloboff, 2004). The endemicity value ( $V$ ) for the species  $Z$  for a given area ( $A$ ) applied by NDM is:

$$V_z = \frac{p + (i \times f_i) + (a \times f_a)}{\text{size} + \frac{o}{f_o} + \frac{d}{f_d} + \frac{n}{f_n}}$$

where  $p$  is the number of cells of  $A$  where the species  $Z$  is actually present;  $i$  is the number of cells of  $A$  where the species  $Z$  is inferred;  $a$  is the number of cells of  $A$  where the species  $Z$  is assumed (defined by the user); size number of cells that conforms  $A$ ;  $o$  number of cells adjacent to  $A$  where the species  $Z$  is actually present;  $d$  number of cells adjacent to  $A$  where the species  $Z$  is assumed;  $n$  number of cells non-adjacent to  $A$  where the species  $Z$  is assumed;  $f_i$ ,  $f_a$ ,  $f_o$ ,  $f_d$  and  $f_n$  are the factors which determine how influential the corresponding terms are. Species then add different scores to a given area according to their distribution inside and outside it. The maximum value is given to a species which occurs in each cell of the area, and in no cells outside the area. If, on the contrary, it occurs outside the area in non-adjacent cells, it is considered widespread and does not add to the score (Szumik and Goloboff, 2004).

The presence observed ( $f_a$ ) and presence assumed ( $f_d$ ) for a given species were filled into the cells according to a given radius, which is expressed in percentage. Under the fill option, presence of a given species applies to all cells in which the species was directly observed as well as to neighbouring cells in which the species was observed within a given distance (% of the cell length and/or height) from the cell (see Fig. 4 for an example using a 20% fill option). The use of the fill option alleviates, to some extent, the problem of sparse sampling. During the process of transforming the plane coordinates into grids of species presence/absence some observed species (points) will inevitably lie close to the cell limit. The fill option presumes that the species is also present in the neighbouring cell if lying within a defined distance of this cell.

At present the fill option is user defined. An objectively defined radius could be the mean and standard

error of the georeferenced specimens following standard procedures such as the point radius method (Wieczorek et al., 2004). In the present study most records were georeferenced by the use of the textual descriptions noted on the specimens' label. When GPS data were available only degrees and minutes were generally recorded. Using standards from Wieczorek et al. (2004), the error distance of coordinates with a precision to the nearest minutes is 2.805 km. Since the data mainly consists of manually georeferenced records, they have been based on a slightly higher error distance of 4 km. We used the fill option when analyzing the data set under the smallest grid size (0.2°, approximately 20 km). Figure 3 illustrates the effect of using the fill option in the present study.

Two criteria were used to search for areas (Table 1). One very strict criterion assumes the borders are well defined ( $f_o = 0.0001$ ), which means species with records outside the borders are not considered. The other criterion, which is more realistic, assumes the limits of the area are diffuse ( $f_o = 0.5^\circ$ ) in which case species with records outside and adjacent to the area are considered endemic.

The results are shown through consensus areas that merge areas (sets of cells), which share  $x\%$  of their defining species; e.g. areas (sets of cells) which differ little in their composition will be summed. Thus, the resulting consensus area shows cells with maximum, low and minimal values of endemicity—reflecting the different scores of the added areas (see consensus areas in Fig. 5). Here, we summarize the areas that share 50% and 75% of their species respectively; obviously, the second criterion is stricter and results in more consensus areas. Two consensus rules are available in the program. The strict consensus rule adds areas (sets of cells) when each area shares  $x\%$  of the defining species with each of the other areas (sets of cells) in the consensus. The less strict consensus rule adds areas (sets of cells) when each area shares  $x\%$  of the defining species with at least one other area (set of cells) in the consensus (Fig. 5). In the present analysis, the less strict consensus rule was used unless otherwise stated.

#### *Altitudinal range*

Because the traditional phytogeographic units of the study region mainly represent altitudinal divisions, the altitudinal range of the defining species from each consensus areas was evaluated. The altitudinal ranges of the defining species were produced by the path profile/LOS tool in Global Mapper ver. 6 using a loaded elevation dataset (GTOPO30). The path profile/LOS tool produces a vertical profile along a user-specified path, which in the present case zig-zags through all the defining specimens of each area (as produced by the

Table 1  
Criteria used to search the sets and resulting sets

Data	Size	<i>M</i>	fo	Rf	Ra	Cells	Spp	Sets	50%	75%	IE
ABC	0.5°	4	0.5°	–	–	26	69	54	6	16	16.27
ABC	0.5°	4	–	–	–	26	61	33	7	13	9.25
AND	0.5°	4	0.5°	–	–	29	118	63	9	34	24.66
AND	0.5°	4	–	–	–	29	111	36	10	23	13.54
TOT	0.5°	4	0.5°	–	–	30	251	90	11	45	69.54
TOT	0.5°	4	–	–	–	30	232	79	4	17	58.63
ABC	0.35°	4	0.5°	–	–	37	57	48	9	25	10.2°
ABC	0.35°	4	–	–	–	37	39	12	2	9	7.64
AND	0.35°	4	0.5°	–	–	39	101	74	11	46	13.62
AND	0.35°	4	–	–	–	39	81	36	6	16	9.07
TOT	0.35°	4	0.5°	–	–	45	237	101	15	55	45.43
TOT	0.35°	4	–	–	–	45	214	81	7	24	24.29
ABC	0.2°	2	0.5°	20	50	133	44	20	11	19	4.93
ABC	0.2°	2	–	20	50	133	35	29	7	9	3.63
AND	0.2°	2	0.5°	20	50	140	78	42	17	36	7.67
AND	0.2°	2	–	20	50	140	45	37	9	14	7.60
TOT	0.2°	2	0.5°	20	50	153	248	203	31	134	24.53
TOT	0.2°	2	–	20	50	153	191	93	23	65	24.53

Data, data sets used; Size, size of the grid cells; *M*, minimum value of the endemism; fo, factor of presence observed in cells adjacent to the area; Rf, radius size to fill as present; Ra, radius size to fill as assume; Cells, total number of cells of the grid; Spp, total number of species that give score to some area; Sets, total number of sets obtained; 50%, 75%, total of consensus areas with a cut off of 50% and 75% of similarity; IE, maximum value of endemism obtained.

output of NDM/VNDM—see documentation and manual). The altitude of each specimen is approximate as it depends both on the precision of each georeferenciation and the underlying GTOP30 file. Instead of translating the altitudes into mean and standard deviation, the direct output from Global Mapper is shown and used in the discussion (Fig. 6).

### Terminology

To avoid tedious repetitions, the term “consensus area” has been reduced to “area” throughout the Results and Discussion section while the term “set” is used when referring to an optimal group of cells obtained from NDM/VNDM. In all other cases, we use the word “region”, for example, when referring to the study region, or to the region of the Yungas, Altoandina, etc. or when discussing pre-established phytogeographic units (e.g. Cabrera and Willink, 1973).

The term “endemic” is used exclusively when referring to species currently known to be present only within the study region, the province of Jujuy. Among the endemic species, eight are grasses: *Anatherostipa henrardiana* (Parodi) Peñailillo, *Aristida pedroensis* Henrard, *Aristida pubescens* Caro & E.A. Sánchez, *Elymus tilcarensis* (J.H. Hunz) A. Löve, *Eragrostis andicola* R.E. Fr. *Festuca superba* Parodi ex Túrpe, *Nassella yaviensis* Torres, and *Poa grisebachii* R.E. Fr. All endemic grass species were included in the analysis, except *Poa grisebachii*, which could not be georeferenced with adequate precision.

## Results

### General results

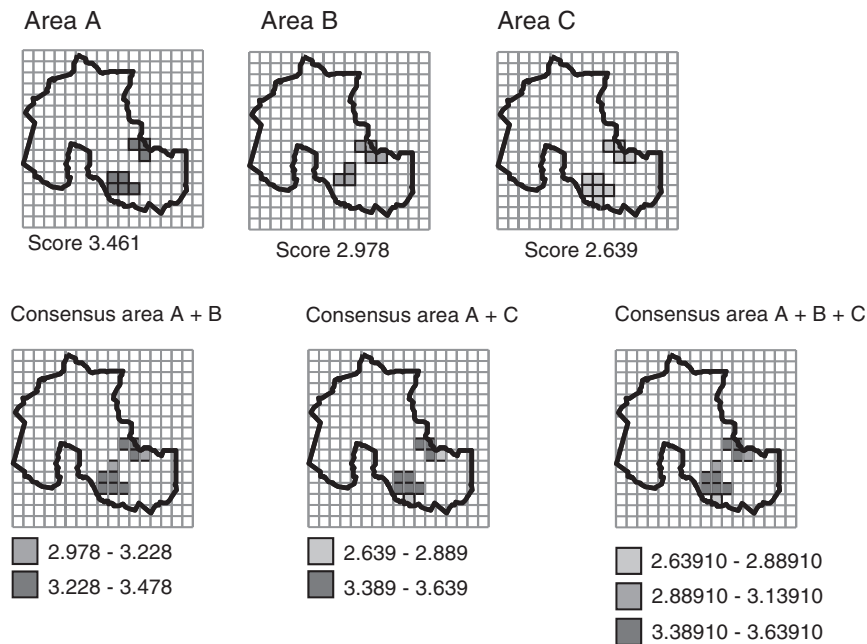
The results for all analyses are summarized in Table 1. In the unrestricted analyses using the three different grid sizes (0.5°, 0.35° and 0.2°) and the consensus criterion 50%, the three data sets yielded: ABC 6–11 areas, AND 9–17 areas and TOT 11–31 areas.

The six areas found with the ABC data set (grid size 0.5° and consensus criterion 50%, hereafter and throughout the text: 0.5°, 50%) are seen in Fig. 6, where the altitudinal range of the defining species is shown (It should be noted that the altitude is not part of the present study but an outcome—see Material and methods).

Each of the six areas is defined by species that correspond to specific altitudinal ranges that differ from area to area. These six areas represent the main distribution patterns that were recovered throughout all the analyses. The score changes of each area when analyzing the ABC, AND or TOT data set are shown in Table 3 (0.5°, 50%). The scores generally improve as more data is added.

### Prepuna

Figure 7 shows one of the areas recovered under all grid sizes in all three data sets. The area is one of the best supported distribution patterns found by NDM/VNDM (Table 1). The geographical location,



Defining species	Area A	Area B	Area C
<i>Anatherostipa brevis</i>	x	x	
<i>Calamagrostis malamalensis</i>		x	
<i>Calamagrostis polygama</i>	x	x	x
<i>Chusquea lorentziana</i>	x	x	
<i>Festuca superba</i>	x	x	x
<i>Paspalum remotum</i>			x
<i>Poa hieronymi</i>	x		x

Fig. 5. Consensus rules. Areas A, B, and C combine under a 50% criterion. Areas B and C do not combine as they do not share 50% of their combined defining species. The consensus area A + B + C is not formed under the strict consensus rule where each area must share 50% (or more) of the defining species with each of the other areas in the consensus. Consensus area A + B + C is formed under the less strict consensus rule where each area must share 50% (or more) of the defining species with at least one other area in the consensus.

combined with the altitudinal range of the defining species indicates that this area corresponds to the Prepuna as defined by Cabrera and Willink (1973); Cabrera, 1976). The Prepuna areas in Fig. 7 are supported by 15–69 species with a score of 10.20–28.10; between 10 and 21 species support the area in all three data sets (Fig. 8, black bars).

To explore the Prepuna in detail, we applied an 80% consensus criterion, which split the original area into northern, southern and central sub-areas (Fig. 9).

#### Puna

A patch of the Puna is found in area C6 (Fig. 6), which appears in all analyses except under grid size 0.2° in the ABC data set (Fig. 10). The altitudinal range of the defining species (Fig. 6) is concordant with the Puna as defined by Cabrera.

The Puna areas are defined by 3–29 species with an endemism score of 2.04–8.24, and three to seven species defining the area in all three data sets. The Puna area is not separated into subunits when a smaller grid size is used.

#### Altoandina

Only a single area was defined by species appearing exclusively above 4000 m (Fig. 11, AND, 0.5°, 60%). Whenever this area appears (AND, 0.5°, 60%; TOT, 0.5°, 50%), it is defined by seven species with an endemism score of 4.21–4.46.

#### Other areas

Two additional distribution patterns emerge from the analyses in all three data sets: the area combining

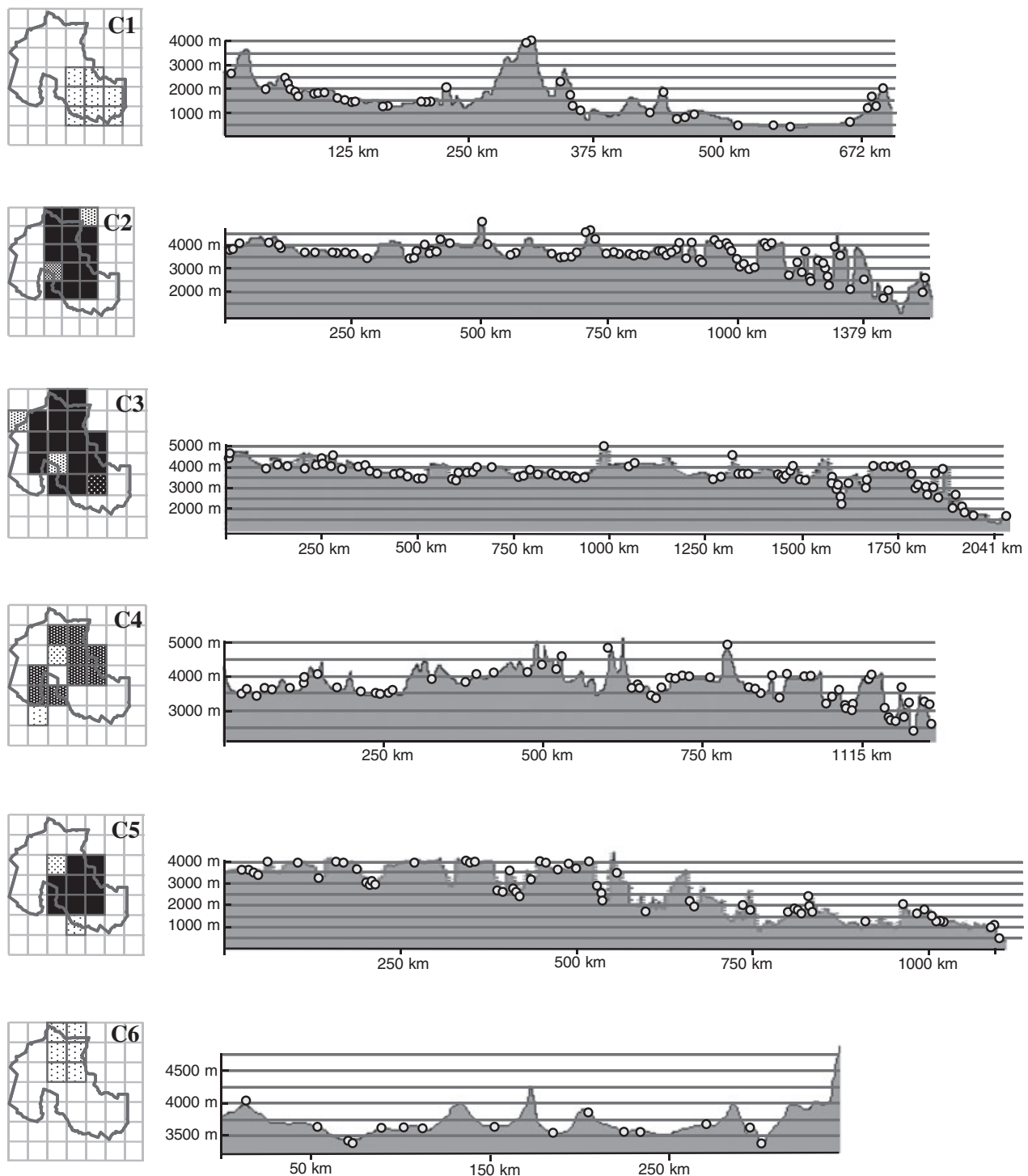


Fig. 6. The six areas found when analyzing the ABC data set, grid size 0.5° consensus criterion 50%. The altitudinal range of the defining species is shown for each area. Each circular dot represents the location of a defining specimen.

Yungas–Chaco vegetation (Fig. 6, C1), and three areas combining Puna–Prepuna vegetation (Fig. 6, C2, C3 and C4).

*Puna–Prepuna vegetation.* The areas in Fig. 6 C2, C3, and C4 are mainly defined by species that occur within the range of the Puna (Fig. 6, altitudinal



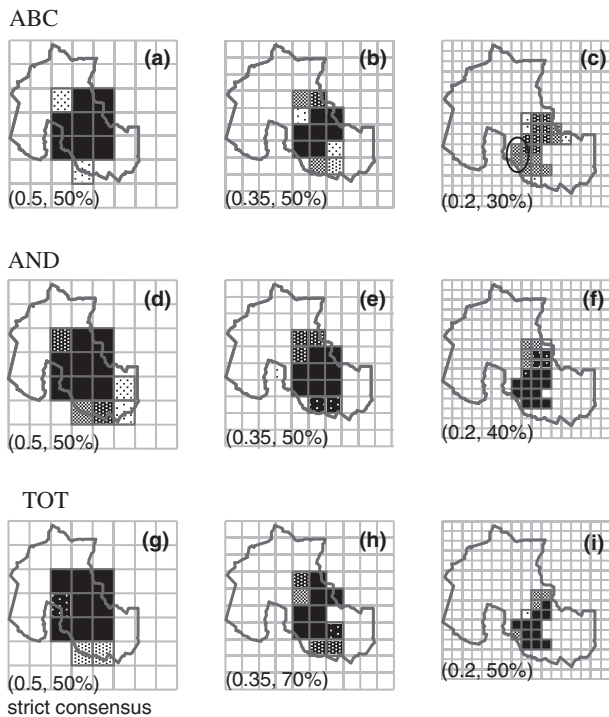


Fig. 7. The Prepuna is found under all grid sizes in all three data set. The highlighted cells in (C) fall outside the Prepuna *sensu* Cabrera (1976) but are persistently part of the Prepuna area in the present analysis.

range). The best supported of the three areas, C2 appears in all analyses, and is principally defined by species that are distributed through the Prepuna and the high plateau of the Puna area. Although part of the area defragments into the Puna and Prepuna areas, the core of the area appears in all three data sets, even under grid size  $0.2^\circ$  when a loose consensus criterion is used to recover the area (Fig. 10e, h). The areas are defined by 8, 17, or 30 species (ABC, AND, TOT,  $0.2^\circ$ , 10%), with all species that define the area in one subset of the data also defining the area in the subsequent more inclusive data sets.

**Yungas–Chaco vegetation.** The area C1 (Fig. 6) recovers a combination of two well described biogeographic units, Yungas and Chaco. The area (ABC,  $0.5^\circ$ , 50) is defined by species distributed between 500 m and 2500 m with the exception of two Mountain Grassland species (*Deyeuxia hieronymi* (Hack.) Türpe and *Deyeuxia colorata* Beetle) found at 4000 m. The endemism score of the Yungas–Chaco area improves slightly when the Andean data set is analyzed while, not surprisingly, there is an outstanding increase when the total data is used adding species with a non-Andean distribution (Table 3). Like the Prepuna it is possible to analyze the

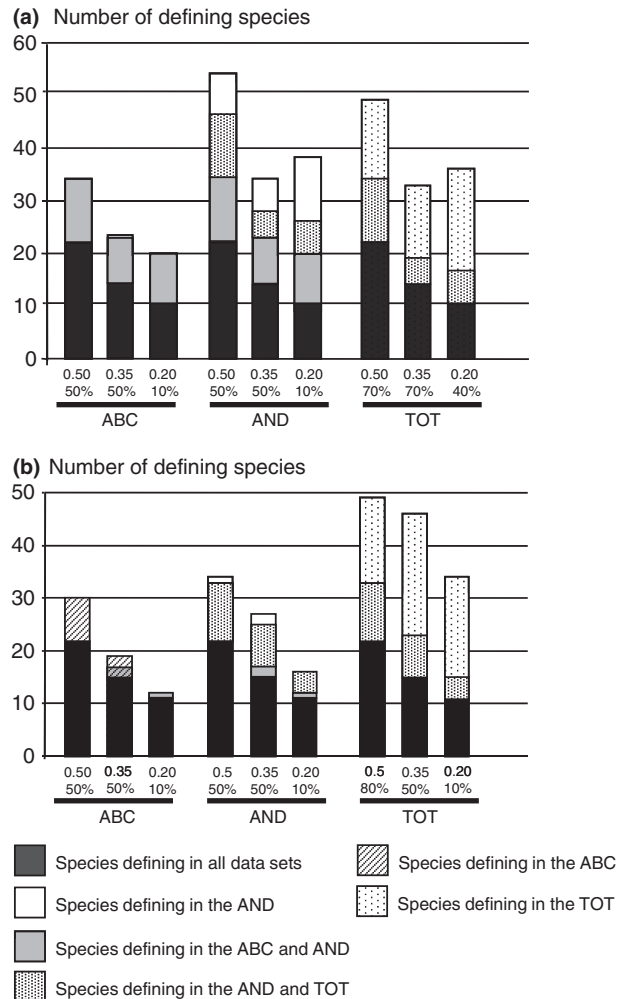


Fig. 8. Number of species that define the Prepuna area under grid size  $0.5^\circ$ ,  $0.35^\circ$ , and  $0.2^\circ$  in data sets ABC, AND and TOT. (a) Data analyzed under unrestricted consensus rule. (b) Data analyzed under strict consensus rule.

Yungas–Chaco area in detail by adjusting grid size and consensus criteria (not shown).

## Discussion

The main distribution patterns that were recovered throughout all the analyses (Fig. 6) and the altitudinal range of the defining species appeared to be concordant with the pre-established phytogeographic scheme proposed by Cabrera (Cabrera and Willink, 1973; Cabrera, 1976). However, under grid size  $0.5^\circ$  most areas combine two of the proposed units. For example, C1 (Fig. 6) combines Yungas–Chaco vegetation, while C2, C3 and C4 (Fig. 6) combine Puna–Prepuna vegetation. Note that, as the topography of each cell in the  $0.5^\circ$  grid is very variable, the specific altitudinal ranges of the

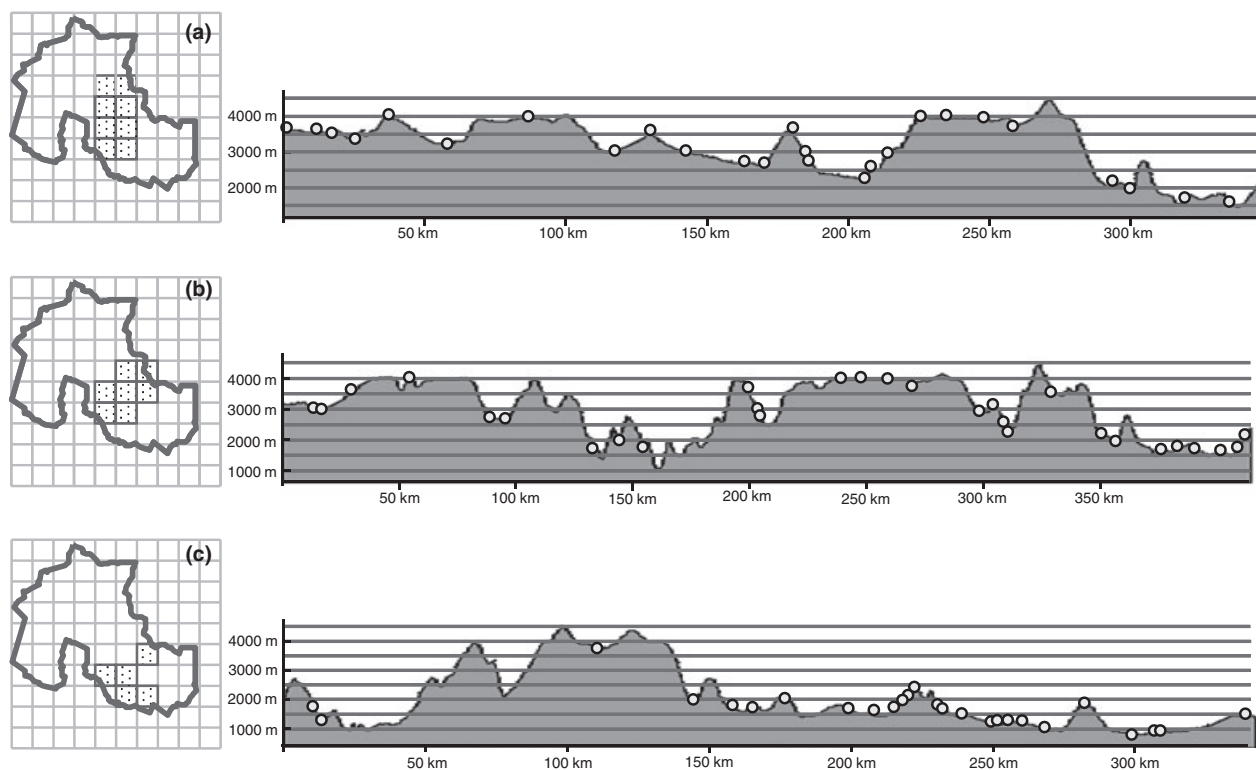


Fig. 9. ABC data set the Prepuna under an 80% consensus criterion. (a) Northern Prepuna and altitudinal range of the defining species. (b) Central Prepuna and altitudinal range of the defining species. (c) Southern Prepuna and altitudinal range of the defining species.

defining species are not a simple product of the region each area covers (compare Figs 2 and 6).

In general, adding widespread species did not decrease the number of defining species with more local distribution patterns (Table 2 and Fig. 12). Among the 104 species in the ABC data set, 90 species (~87%) defined areas in one or more analyses changing grid sizes and/or settings (Table 2 and Fig. 12). When the broad Andean species were added (AND data set) 93 ABC species (~89%), 52 AND species (~93%) were defining areas, under two or more grid sizes/settings (Table 2 and Fig. 12). When all species were analyzed (TOT data set) 96 ABC (~92%) and 52 AND species continued to define areas in addition to 163 (~91%) defining species with a non-Andean distribution (Table 2). In the TOT data set, the majority of the defining species supported areas under six different grid sizes/settings (Fig. 12).

In the following section, the areas that appear to include the three highland distribution patterns, the Altoandina, the Puna and the Prepuna, are explored. We also discuss how these three biogeographic units emerge and/or transform with changing grid size and the addition of new data.

### *Prepuna*

**Stability.** The Prepuna distribution pattern (Fig. 7) appears readily in the ABC and AND data set while stricter consensus criteria are needed to visualize clearly the same distribution pattern when analyzing the TOT data set. The species that are added in the TOT data set—those with a wide American (non-Andean) or extra American distribution—mainly reinforce the strongly supported southern pattern seen in Fig. 6 C1 (see also Table 3), which overlaps with the Prepuna distribution pattern. Either relatively stricter consensus criteria (Fig. 7h, i) or the strict consensus rule (sets of cells combined when all share 50% of the defining species) in the case of grid size  $0.5^\circ$  (Fig. 7g) results in two separate patterns.

The Prepuna was originally defined as a strictly Argentinean biogeographic unit (Cabrera, 1976). However, 15 of the 21 species that define the area in all data sets under grid size  $0.5^\circ$  (50%; Fig. 8a) are also found in neighbouring countries, while only six species are restricted to Argentina. Consequently, several of the species that define the Prepuna under grid size  $0.5^\circ$  must be part of a wider extended pattern. This result is in agreement with López and Beck (2002) and López

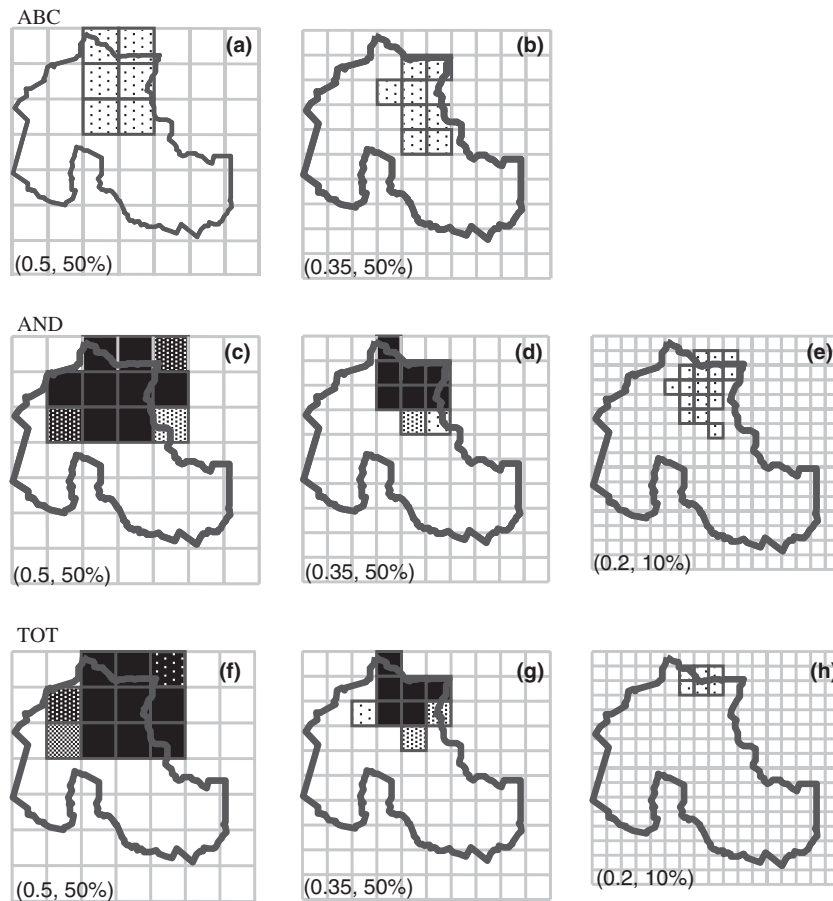


Fig. 10. The Puna is found under all grid sizes in all three data sets except ABC (0.2°).

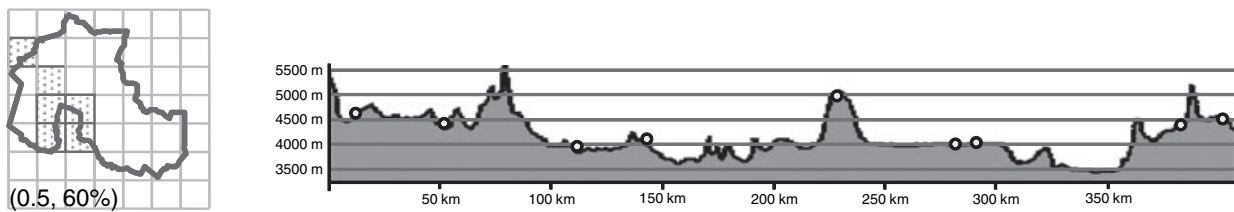


Fig. 11. The Altoandina area found with the AND data set (0.5°, 60%). The altitudinal range of the defining species is shown with each circular dot representing the location of a defining specimen.

Table 2

Number of species adding score to areas with a distribution in either Argentina, Bolivia and/or Chile, in the Andes in general or in non-Andean habitats

Distribution	No. species	Defining in ABC	Defining in AND	Defining in TOT
Argentina, Bolivia, Chile	104	90	93	96
Broad Andean	56	—	52	52
Other	180	—	—	163

(2003), who discussed the existence of Prepuna vegetation (sensu Cabrera, 1976) in southern Bolivia.

The small grid size seems, however, to be capturing a more local distribution pattern, which is essentially concordant with the extension of the Prepuna, as defined by Cabrera. Under grid size 0.2° the 10 shared species in Fig. 7A are either endemic to the defined area (four species) or restricted to the study region (four species) or to Argentina (three species) while only three species are found in neighbouring countries. In general,

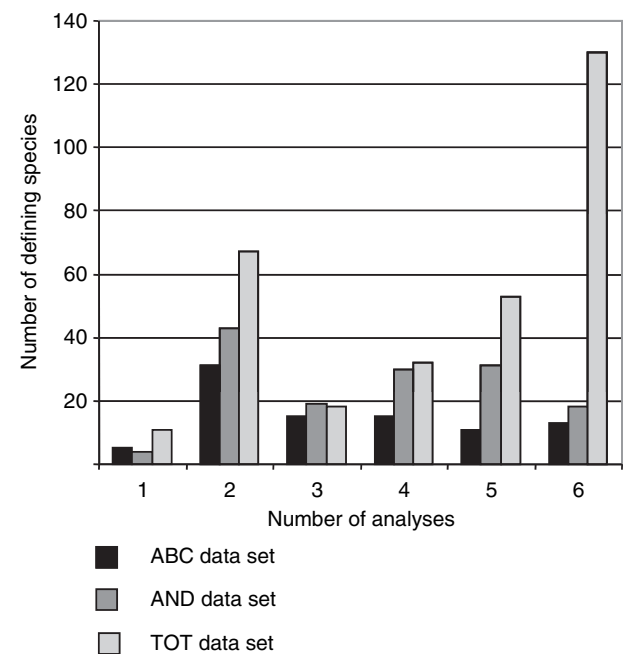


Fig. 12. Number of species in the ABC, AND, and TOT data set that support areas in one, two... or, six analyses using different grid sizes and settings.

Table 3  
Change of score in areas from ABC, AND, and TOT

Area	ABC	AND	TOT
C1	6.02	16.81*	69.53
C2	14.61	24.66	31.62
C3	9.23	15.60*	5.96
C4	6.72	6.82	7.87
C5	16.27	23.76	27.38*
C6	4.32	7.21	8.63

Grid size 0.5°, consensus criterion 50%.  
\*Found only in 60% consensus, hence the score is not directly comparable with the areas found under a 50% consensus rule.

using a small grid size combined with strict consensus rules proved useful when searching for local distribution patterns (see Fig. 13).

**Fragmentation of the Prepuna.** To explore the Prepuna in detail, we applied an 80% consensus criterion, which separated the original area into northern, southern and central sub-areas (Fig. 9). The three sub-areas differed notably in the altitudinal range of the defining species. Each area lies within the range of the Prepuna *sensu* Cabrera, with the northern area including most of the higher altitude locations and the southern area including most of the low altitude locations (Fig. 9). The central area combines the pattern found in the northern and southern areas, and corresponds to the cells of the highest endemicity value in the Prepuna area above (see

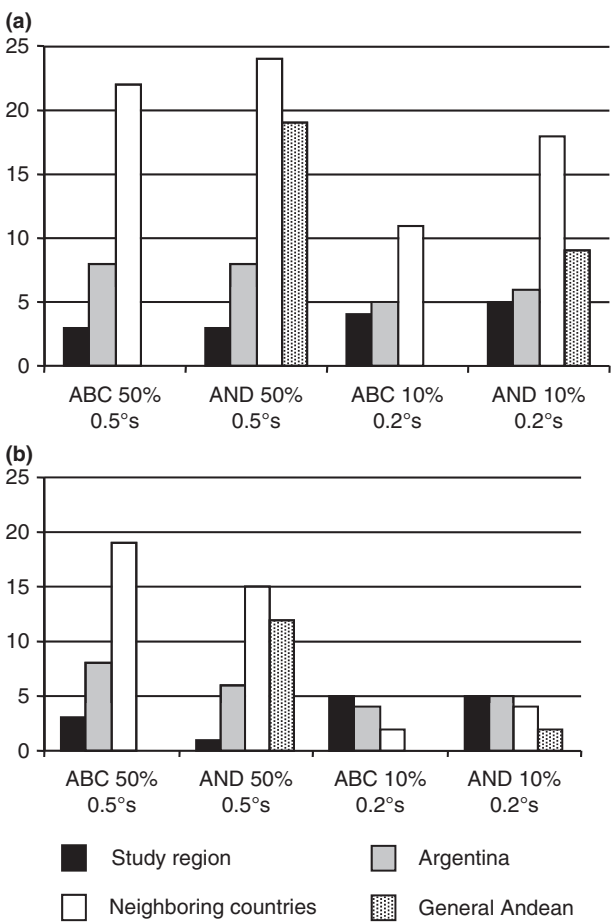


Fig. 13. Geographic distribution of the species that define the Prepuna area under grid size 0.5° and 0.2° in data sets ABC and AND. (a) Data analyzed under unrestricted consensus rule. (b) Data analyzed under strict consensus rule.

dark shaded area in Fig. 7b). The species that define the central area are distributed both in the Prepuna and in the neighbouring Mountain Grassland. Following Cabrera (1976), the Mountain Grasslands have traditionally been defined as the upper strata of the Yungas. However, this interpretation has been disputed (Brown, 1995). The Prepuna and Mountain Grasslands have not been previously been related to each other, but according to grass species distribution they are clearly part of the same pattern.

**Identity.** None of the Prepuna species mentioned by Cabrera (1976) support the area (see Appendix 1). *Digitaria californica* (Benth.) Henrard and *Munroa argentina* Griseb. both fit the area well but have been found at single locations outside the area in non-adjacent cells. A third species mentioned by Cabrera, *Jarava leptostachya* (Griseb.) F. Rojas is also found in the area but reaches far into the Puna and Altoandina provinces. It is part of the species that define the

general east–west running area C3 (Fig. 6 and below), and, as such, it is not a good indicator of the Prepuna.

Four grass species are almost perfectly adjusted to the Prepuna area in the present analysis (ABC, 0.35°, 50%, Fig. 7e). Two of these may prove to be confined to the Argentinean Prepuna: *Eragrostis andicola* R.E. Fr. (endemic to the study region) and *Jarava media* (Speg.) Peñailillo (only found in the Argentinean Prepuna habitats, according to Torres, 1997b). The two other species *Deyeuxia polygama* (Griseb.) Parodi and *Poa hieronymi* Hack. are found in the Mountain Grasslands of Argentina (Rúgolo de Agrasar, 2006).

The Prepuna-Mountain Grassland distribution pattern includes four of the eight grass species endemic to the study region: *Anatherostipa brevis* (Torres) Peñailillo, *Eragrostis andicola*, *Elymus tilcarensis* (J.H. Hunz.) A. Löve, and *Festuca superba*, all found in the Prepuna except for *Festuca superba* Parodi ex Türpe, which is found in the Mountain Grasslands. This suggests that the Prepuna-Mountain Grassland is not only a well supported pattern but might also be an area of endemism. Further analyses including species from other Angiosperm families are, however, needed in order to confirm these findings.

Finally, the analysis showed a case of predictive power. According to traditional biogeographic judgments, the Prepuna does not extend into the most southwestern cells of the Prepuna area (highlighted in Fig. 7c), while these cells are persistently part of the Prepuna area shown in Fig. 7. Both Prepuna indicators found in the present analysis, *Eragrostis andicola* and *Jarava media*, have been collected in this region. In a recent visit to the location, we found a vegetation mainly composed of Puna elements (*sensu* Cabrera, 1976) but with several Prepuna indicators appearing at sheltered rocky sites, among which the most emblematic was the columnar cactus *Trichocereus atacamensis* (Phil.) W.T. Marshall.

### Puna

**Stability.** Above the Prepuna lies the Puna (approximately 3400–4500 m—*sensu* Cabrera). Compared to the Prepuna, the Puna is supported by very few species, perhaps reflecting a general impoverishment of herbarium collections at high altitudes. In the ABC data set only seven species define the area under grid size 0.5° (50%—Fig. 10a). In the AND data set (0.5°, 50%—Fig. 10f) 22 species mainly distributed from Argentina to Peru define the area. Unlike the Prepuna, the distribution pattern that defines the Puna area is not affected by the species added in the TOT data set. The Puna area is recovered by same settings and grid size as in the ABC and AND data sets (Fig. 10f). The Andean aspect of the area also persists. Of the 29

species defining the area in the TOT data set (0.5°, 50%) only three species are not restricted to the Andes.

As opposed to the Prepuna, it was not possible to separate the Puna area into subunits, nor does the distribution pattern that defines the area become more local when a smaller grid size is used.

**Identity.** Among the seven species that define the area under grid size 0.5° (ABC, 50%) five are found in neighbouring countries, while one species is limited to Argentina, and one species is endemic to the study region. The species *Muhlenbergia fastigiata* (J. Presl) Henrard, that adds highest score to the area (AND, 0.35°, 50%) is found in the high Andes of Argentina, Bolivia and Peru, and appears in the list of Puna species mentioned by Cabrera (1976).

However, the patch of the Puna found in this study also includes local flora that reaches Bolivia but not Peru. *Nassella cabreræ* Torres, which adds the second highest score to the area (AND, 0.35°, 50%), is one of the species that is restricted to the Puna of Argentina and Bolivia while *Nassella yaviensis* is endemic to the study region (Torres, 1997a).

### Stability.

Species that are distributed through the Prepuna and the high plateau of the Puna consistently defined an area which appeared in all analyses (Fig. 6, C2). In the complete analysis (TOT, 0.35°, 50%) 44% of the defining species are confined to Argentina, Bolivia and Chile, while 24% are general Andean species (14% widely distributed), indicating that the pattern is a patch of more general distribution that continues into the neighbouring countries, with part of the pattern reaching the northern part of the Andes mountains. The Puna–Prepuna distribution pattern was also found by Morrone et al. (2005) but not mentioned by Cabrera who listed several of these species as part of the Puna vegetation.

### Altoandina

**Stability.** Above the Puna lies the Altoandina biogeographic province, which forms the top of the vegetation line that limits with the perpetual snow. Grasses are the principal vegetation type in this biogeographic province, but Altoandina areas were difficult to recognize in the present study. When areas with Altoandina vegetation do appear, they are only weakly supported or mixed with species from the Puna.

Samples of localities above 3500 m are not frequently found in collections (Fig. 3) and above 4400 m they are concentrated into a few long-established localities, where access has been available, for example by mining



activities. This implies that not only are high Andean collections rare but are also reduced to a few points unevenly spaced across the study region. There is an acute need for more and evenly spaced collections above 4000 m in order to analyze species distribution at these altitudes.

However, under a coarse grained grid (0.5°) a single area was defined by species appearing exclusively above 4000 m. (Fig. 11, AND, 0.5°, 60%). The defining species are exclusively Andean, one of them endemic to the study region, four found in neighbouring countries, and two reaching Peru.

**Identity.** Cabrera listed six of the seven defining species as Altoandina; among these, the two species with greatest fit: *Anatherostipa bomanii* (Hauman) Peñailillo, distributed from Argentina to Peru, and *Anatherostipa henrardiana* endemic to the study region. Other Altoandina species *sensu* Cabrera, for example *Festuca orthophylla* Pilg. and *Poa gymnantha* Pilg. are more widespread and are among the species that define a broad Puna area in Fig. 6 (C2).

Compared to the Puna area, the patch of the Altoandina found here appears to be of a more local distribution pattern, partly endemic, partly continuing into neighbouring countries with only few species reaching Peru.

#### General considerations

In the present analysis of grass species distributions, concordant distribution patterns in the northwestern Argentina mainly conformed to altitudinally divided units as has already been suggested in traditional phytogeographic schemes (Cabrera and Willink, 1973; Cabrera, 1976).

However, even the most robust distribution pattern overlapped with other areas. This should come as no surprise given that transition zones may be found between two biogeographic units. Therefore, one of our main conclusions is that methods designed to analyze species distribution must take this simple fact into consideration. In order to separate and explore overlapping patterns in the present study, we depended strongly on flexibility in variable grid size and adjustable consensus criteria/rules.

We found no reason for excluding species from the analysis. Several areas were reinforced when more data were added even if the new data belonged to an overall different distribution (e.g. Andes versus non-Andes distributions). The fact that NDM/VNDM does not consider support from a species when it also occurs outside the area in a non-adjacent cell prevents widespread species from blurring distribution patterns when using a small grid size. Such species may still support areas in a more coarse grained analysis of the same

region. In some cases, strongly supported distribution patterns like the Yungas–Chaco area appeared to absorb overlapping but less strongly supported patterns such as the Prepuna area. Those areas were, however, easily separated by increasing the percentage of shared species required to combine two sets of cells.

The sampling problem that is found in biogeographic data sets was somewhat solved under the small grid size (0.2°) by using the fill option available in the program. The fill option compensates for the fact that each species is represented by points of collected specimens which are translated into present absence in the cells of a grid. Uncertainties rise when a point falls very close to the limit between two cells, especially if the specimen is georeferenced manually by the textual description of the collection site. The fill option allows points to have some extension by adding an uncertainty radius that should ideally reflect uncertainty in the georeferenciation.

At the present time huge amounts of species distribution data are found in herbaria and museums. An enormous effort must be undertaken to make this data ready for quantitative distribution analyses. Gathering the information in data bases and providing reliable georeferences for the collected specimens is an extremely time-consuming task. However, once overcome, it is possible to analyze this distribution data with quantitative tools such as NDM/VNDM. We suggest here that the outcomes of such analyses are suitable for defining new areas as well as in evaluating pre-established ones, in both local and large scale analyses. It has been noted that species distributional patterns are the one common element in both historical and ecological biogeography, and cannot be fully understood without an integration of both disciplines (Crisci et al., 2006). Given this fundamental role of distributional patterns, we believe that the common and concordant distribution patterns that are the objects of biogeographic studies should themselves be defined on the basis of quantitative analyses.

#### Acknowledgements

We thank editors James Carpenter and Dennis Stevenson, Peter Hovenkamp, Lusía Montivero and anonymous reviewers for helpful comments and criticism. Pablo Goloboff added new options to VNDM during the analytical part of this study. L.A.A. thanks María Negritto for providing material of *Poa* ssp. C.A.S. also thanks Instituto Argentino de Estudios Filogenéticos. This work was supported by ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Argentina) grants 11793, 13374 and 12605 and CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), grants 2131 and 2567.

## References

- Brown, A.D., 1995. Fitogeografía y conservación de las selvas de montaña del noroeste de Argentina. In: Churchill, S.P., Balslev, H., Forero, E., Luteyn, J.L. (Eds.), *Biodiversity and Conservation of Neotropical Montane Forests*. The New York Botanical Garden, New York, pp. 663–672.
- Cabrera, A.L., 1957. La vegetación de la Puna Argentina. *Revista Invest. Agríc.* 11, 316–412.
- Cabrera, A.L., 1976. Regiones fitogeográficas argentinas. *Enciclopedia Argentina de Agricultura y Jardinería* 2, 1–85.
- Cabrera, A.L., Willink, A.W., 1973. Biogeografía de América Latina Serie de Biología OEA. Monografía 13, 1–117.
- Crisci, J.V., Sala, O.E., Katinas, L., Posadas, P., 2006. Bridging historical and ecological approaches in biogeography. *Aust. Syst. Bot.* 19, 1–10.
- Crisp, M.D., Laffan, S., Linder, H.P., Monro, A., 2001. Endemism in the Australian flora. *J. Biogeogr.* 28, 183–198.
- Díaz Gómez, J., 2007. Endemism in *Liolaemus* (Iguaria: Liolaemidae) from the Argentinian Puna. *S Am J Herpetol* 2, 59–68.
- Domínguez, M.C., Roig-Juñent, S., Tassin, J.J., Ocampo, F.C., Flores, G.E., 2006. Areas of endemism of the Patagonian steppe: an approach based on insect distributional patterns using endemicity analysis. *J. Biogeogr.* 33, 1527–1537.
- Goloboff, P., 2004. NDM/VMDM, programs for identification of areas of endemism. Program and documentation. Available at <http://www.zmuc.dk/public/phylogeny/endemism>.
- Linder, H.P., 2001. On areas of endemism, with an example from the African Restionaceae. *Syst. Biol.* 50, 892–912.
- Lizarralde, M., Szumik, C., 2007. Phylogeny and Biogeography of the Genus *Pelinoidea* Cresson (Diptera-Ephydriidae). *Zootaxa* 1510, 35–50.
- López, R.P., 2003. Phytogeographical relations of the Andean dry valleys of Bolivia. *J. Biogeogr.* 30, 1659–1668.
- López, R.M., Beck, S., 2002. Phytogeographical affinities and life form composition of the Bolivian Prepuna. *Candollea* 57, 77–96.
- Morrone, J.J., 1994. On the identification of areas of endemism. *Syst. Biol.* 43, 438–441.
- Morrone, O., Aliscioni, S.S., Zuloaga, F.O., 2005. Análisis de la diversidad y distribución geográfica de la familia Poaceae en la provincia de Jujuy, Argentina. *Ann. Missouri Bot. Gard.* 92, 595–639.
- Negritto, M.A., Anton, A.M., 2000. Las especies de *Poa* (Poaceae) del noroeste argentino. *Kurtziana* 28, 95–136.
- Platnick, N.I., 1991. On areas of endemism. *Austral. Syst. Bot.* 4, xi–xii [Commentary].
- Prado, D.E., 1993. What is the Gran Chaco vegetation in South America? I. A review. Contribution to the study of flora and vegetation of the Chaco. V. *Candollea* 48, 145–172.
- Rúgolo de Agrasar, Z.E., 2006. Las especies del género *Deyeuxia* (Poaceae: Pooideae) de la Argentina y notas nomenclaturales. *Darwiniana* 44, 131–293.
- Ruthsatz, B., Movia, C.P., 1975. Relevamiento de las estepas andinas del noreste de la provincia de Jujuy. Fundación para la Educación, la Ciencia y la Cultura, Buenos Aires.
- Sulekic, A.A., 2003. Revision of species of *Aristida* (Poaceae, Aristideae) from northwestern Argentina. *Darwiniana* 4, 155–188.
- Szumik, C.A., Goloboff, P.A., 2004. Areas of endemism: an improved optimality criterion. *Syst. Biol.* 53, 973–980.
- Szumik, C.A., Roig-Juñent, S., 2005. Criterio de optimalidad para áreas de endemismo: El caso de América del Sur Austral. In: Morrone, J.J., Llorente Bousquets, J., (Eds.), *Regionalización biogeográfica en Iberoamérica y tópicos afines*. Las Prensas de Ciencias, Facultad de Ciencias, Universidad Autónoma de México, pp. 495–508.
- Szumik, C.A., Cuezco, F., Goloboff, P.A., Chalup, A., 2002. An optimality criterion to determine areas of endemism. *Syst. Biol.* 51, 806–816.
- Torres, A.M., 1997a. Nassella (Gramineae) del noroeste de la Argentina. Comisión de Investigaciones Científicas (Buenos Aires), 13, 5–46.
- Torres, A.M., 1997b. Stipa (Gramineae) del noroeste de la Argentina. Comisión de Investigaciones Científicas (Buenos Aires), 13, 47–68.
- Wieczorek, J., Guo, Q., Hijmans, R.J., 2004. The point-radius method for georeferencing locality descriptions and calculation associated uncertainty. *Int. J. Geogr. Inf. Sci.* 18, 745–767.
- Zuloaga, F.O., Morrone, O., 1996. Catálogo de las Plantas Vasculares de la República Argentina. I. Pteridophyta, Gymnospermae y Angiospermae (Monocotyledoneae). *Monogr. Syst. Bot. Missouri Bot. Gard.* 60, 1–332.
- Zuloaga, F.O., Nicora, E.G., Rúgolo de Agrasar, Z.E., Morrone, O., Pensiero, J.F., Cialdella, A.M., 1994. Catálogo de la familia Poaceae en la República Argentina. *Monogr. Syst. Bot. Missouri Bot. Gard.* 47, 1–178.
- Zuloaga, F.O., Morrone, O., Rodríguez, D., 1999. Análisis de la biodiversidad en plantas vasculares de la Argentina. *Kurtziana* 27, 17–167.
- Zuloaga, F.O., Morrone, O., Belgrano, M.J., 2008. Catálogo de las Plantas Vasculares del Cono Sur. *Monogr. Syst. Bot. Missouri Bot. Gard.* 107, 609–967.

## Appendix 1

### *Phytogeographic units (sensu Cabrera, 1976) found in the study region*

#### *Altoandina.*

The Altoandina biogeographic province lies between 4400 m and 5600 m, occupying the highest slopes of the Andean mountains (Figs 1 and 2) from Venezuela to Tierra del Fuego of southern Argentina and Chile (Cabrera, 1976). The vegetation in the region is exposed to low temperatures, sparse precipitation and strong winds. The grass steppes are the principal vegetation type, forming isolated mats mainly of *Festuca orthophylla* Pilg., *Festuca chrysophylla* Pilg., *Poa gymnantha* Phil., *Jarava frigida* Phil., *Deyeuxia cabreræ* (Parodi) Parodi, and *Anatherostipa* sp. (Cabrera, 1976).

#### *Puna.*

The biogeographic province of the Puna is found at the high Andean plateau located between the two Andean mountain chains (Figs 1 and 2), stretching from Peru to northwestern Argentina. The plateau is between approximately 3400 m and 4500 m (Cabrera and Willink, 1973). The climate is cold and dry with rain falling predominately during the summer. The dominant vegetation is open bush steppes mainly formed by the families Solanaceae, Fabaceae and Asteraceae, accompanied by grass species such as *Jarava leptostachya* (Griseb.) F. Rojas, *Pennisetum chilense* (E. Desv.) B.D. Jacks. ex R.E. Fr. *Deyeuxia rigida* Kunth, *Aristida antoniana* Steud. ex Döll and *Poa* sp.

#### *Prepuna.*

According to Cabrera (1976; Cabrera and Willink, 1973) the Prepuna is restricted to the northwestern Argentina where it occupies mountain slopes and ridges between the Yungas and the Puna (Figs 1

and 2) between approximately 2000 and 3400 m. The climate is relatively mild and dry with rainfall exclusively during the summer. Grasses are not dominant in this biogeographic province, which is mainly defined by bushes of the families Asteraceae, Fabaceae, and Caesalpinaceae and by several Cactaceae species. Especially emblematic of the Prepuna is the presence of cushion-forming Bromeliaceae species and the columnar cactus *Trichocereus atacamensis* (Phil.) W.T. Marshall (Cabrera, 1976). Few grasses are mentioned for this region, among them *Digitaria californica* (Benth.) Henrard, *Jarava leptostachya*, *Munroa argentina* Griseb., and *Eragrostis* sp.

### Yungas.

According to Cabrera (1976; Cabrera and Willink, 1973) the biogeographic province of the Yungas occupies the eastern slopes of the Andes mountains between 500 and 2500/3500 m extending from Venezuela to the north of Argentina. The climate of the Yungas is mild and humid with the mountain slopes often covered by clouds. The prevalent vegetation is cloud forest dominated by Myrtaceae and Lauraceae.

Cabrera divided the Yungas into three districts: Transition Forest (Pedemontana [Brown, 1995] or Subandean Piedmont Forest [Prado,

1993]) between ~350 and 500 m, Mountain Forest between ~550 and 1600 m, and Mountain Woodland between ~1200 and 2500 m.

Grasses become dominant only in the uppermost part of the Mountain Woodland—in the “Mountain Grasslands”—found between ~2100 and 3000 m. Cabrera (1976) mentioned eight Poaceae species as common for the grasslands: *Festuca hieronymi* Hack., *Deyeuxia colorata* Beetle, *D. polygama* (Griseb.) Parodi, *Eustachys distichophylla* (Lag.) Nees, *Lamprothyrsus hieronymi* (Kuntze) Pilg., *Paspalum candidum* (Humb. and Bonpl. ex Flügge) Kunth, *Ant-haenantiopsis fiebrigii* Parodi, and *Nassella tucumana* (Parodi) Torres.

### Chaco.

The Chaco biogeographic province is found in north and central Argentina, south-eastern Bolivia, western Paraguay, and a small portion of Mato Grosso do Sul, Brazil (Cabrera and Willink, 1973; Prado, 1993). It is occupied by lowland and low mountains, with continental climate, and rainfalls only during the summer (Figs 1 and 2). The vegetation is mainly xerophytic, with caducous forest and an understory of Poaceae, Cactaceae and Bromeliaceae (Cabrera and Willink, 1973).