Ann. Bot. Fennici 48: 449–454 ISSN 0003-3847 (print) ISSN 1797-2442 (online)
Helsinki 30 December 2011

Identification of Sardinian species of *Astragalus* section *Melanocercis* (Fabaceae) by seed image analysis

Gianluigi Bacchetta[1,*], Giuseppe Fenu[1], Oscar Grillo[1], Efisio Mattana[1] & Gianfranco Venora[2]

[1)] *Centro Conservazione Biodiversità (CCB), Dipartimento di Scienze Botaniche, Università degli Studi di Cagliari, v.le S. Ignazio da Laconi 13, IT-09123 Cagliari, Italy (*corresponding author's e-mail: bgsar@ccb-sardegna.it)*
[2)] *Stazione Consorziale Sperimentale di Granicoltura per la Sicilia, via Bouganvillea 20, IT-95041 Caltagirone (CT), Italy*

Received 8 Jan. 2010, revised version received 4 July 2011, accepted 17 Mar. 2010

Bacchetta, G., Fenu, G., Grillo, O., Mattana, E. & Venora, G. 2011: Identification of Sardinian species of *Astragalus* section *Melanocercis* (Fabaceae) by seed image analysis. — *Ann. Bot. Fennici* 48: 449–454.

Using a computer-aided imaging system, the seed morphometric and colorimetric features of *Astragalus terraccianoi*, *A. tegulensis* and *A. thermensis* — all belonging to *Astragalus* sect. *Melanocercis* — were measured and analysed. The results confirmed differences between *A. terraccianoi* and *A. tegulensis*, as well as the validity of *A. tegulensis*. Differences between the Corsican and Sardinian populations of *A. terraccianoi* were also found.

Introduction

The genus *Astragalus* (Fabaceae) comprises at least 2300 species (largely based on currently accepted names in Podlech 1999) and perhaps up to 2500 species in some 245 sections (Lewis *et al*. 2005). The main speciation centre of this genus is in Turkey, Iran and Afghanistan, with ca. 1200 species in the Middle East, SW Asia and the Sino-Himalayan region to W China (Lewis *et al*. 2005).

The studied species *Astragalus terraccianoi* and *A. tegulensis* belong to *Astragalus* subgen. *Cercidothrix* sect. *Melanocercis* (Chater 1968, Chamberlain & Matthews 1969) because of their spiny suffruticous habit, mediafixed black-and-white hairs, imparipinnate leaves with spine-like rachis, stipules adnate to the petiole, flowers in racemes, tubular calyx, and legume exceeding the calyx. *Astragalus tragacantha* and the closely related *A. balearicus* and *A. thermensis* are also included in this section (Podlech 2008, Bacchetta & Brullo 2010). *Astragalus terraccianoi* and *A. tegulensis* are closely related, and Bacchetta and Brullo (2010) have recently identified a south-Sardinian population of *A. terraccianoi* as belonging to a new species *A. tegulensis*. Therefore, *A. terraccianoi* is endemic to the NW part of Sardinia and southern Corsica, *A. tegulensis* is known from only one population in SW Sardinia (Bacchetta & Brullo 2010), and *A. thermensis* is endemic to N Sardinia (Valsecchi 1994, Giardina *et al*. 2007) (Fig. 1).

Using seed-image analysis techniques, Bacchetta *et al*. (2008) were able to confirm the taxonomic distance between *A. terraccianoi*

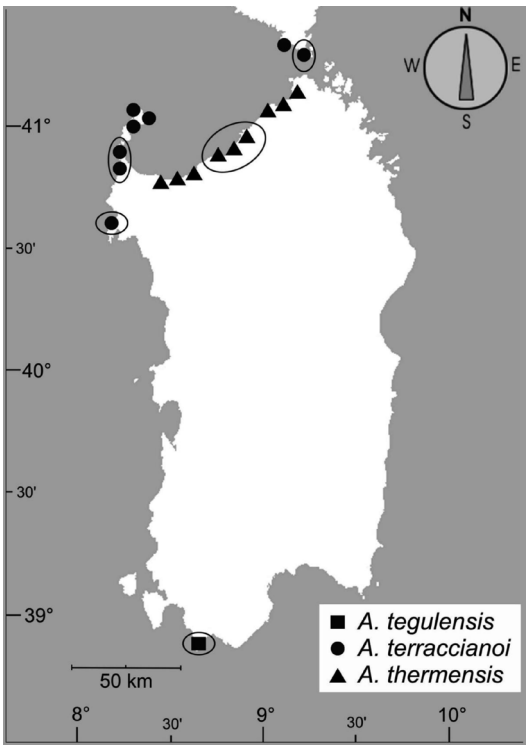

Fig. 1. Distribution of *Astragalus tegulensis*, *A. terraccianoi* and *A. thermensis* populations in Sardinia and Corsica. The study populations are encircled.

and *A. thermensis* (86.8% of correct classification). With the same techniques, Mattana *et al.* (2008) achieved 100% correct classification for *Astragalus verrucosus* and *A. maritimus*, subgen. *Trimeniaeus* endemic to Sulcis-Iglesiente (SW Sardinia). Hence, this approach seems appropriate for comparative taxonomic studies within the genus *Astragalus*.

The aims of this work were: (1) to analyze taxonomic distances among *A. thermensis*, *A. terraccianoi* and *A. tegulensis*, representing all the Sardinian species of *Astragalus* sect. *Melanocercis*, (2) to verify the validity of the new species proposed by Bacchetta and Brullo (2010), and (3) to analyze variability among populations of *A. terraccianoi*.

Material and methods

The images of six seed samples obtained from the five Sardinian populations and one Corsican population (a total of 539 seeds) (Fig. 1 and Table 1) were acquired following Bacchetta *et al.* (2008). The seed samples were scanned using a flatbed scanner (Epson GT-15000) with a resolution set at 200 dpi and a scanning area not exceeding 1024 × 1024 pixels. Following Venora *et al.* (2009a), before the seed images were acquired, the scanner was colour-calibrated using the Shahin and Symons (2003) protocol i.e., after scanning a Q60 Kodak Input Colour Target a Look Up Table was created that allowed to match a colour profile of any new scanned image to the standardized reference.

To prevent any artificial variation in seed size, shape or colour, the seeds were scanned before drying (at 15 °C to 15% RH) for long term conservation. When the original accession consisted of more than 100 seeds, subsamples of 100 seeds were used for the digital image analysis, otherwise, the analysis was caried out on the whole batch. The digital images were captured and analysed using the software package KS-400 ver. 3.0 (Carl Zeiss, Vision, Oberkochen, Germany). A computer macro specifically developed for characterization of wild seeds (Bacchetta *et al.* 2008) and later modified by adding 20 new

Table 1. Studied accessions of *Astragalus* species from Corsica and Sardinia.

Species	Accession code	Locality	Altitude (m a.s.l.)	Coordinates
A. terraccianoi	179/06	Capo Pertusato, Bonifacio, Ajaccio (Corsica)	73	41°22.678´N, 9°10.731´E
	259/06	Cala della Barca, Alghero, Sassari (Sardinia)	45	40°36.355´N, 8°09.062´E
	268/06	Punta Falcone, Stintino, Sassari (Sardinia)	37	40°57.969´N, 8°12.189´E
A. thermensis	197/06	Badesi mare, Badesi, Sassari (Sardinia)	5	40°58.448´N, 8°51.245´E
	199/06	Foci del Coghinas, Badesi, Sassari (Sardinia)	11	40°56.336´N, 8°49.040´E
A. tegulensis	138/06	Capo Teulada, Teulada, Cagliari (Sardinia)	3	38°53.750´N, 8°38.607´E

seed morpho-colorimetric features (Mattana *et al*. 2008), was used to measure the seed size (area, perimeter and diameter of the seed projection), shape (shape, roundness and circular geometrical descriptors of the seed projection), colour (red, green, blue, hue, lightness and saturation values of the seed surface pixels and their relative dispersion values) and density. The relative dispersion values of the colour and density parameters were calculated for each seed in the batch. In addition, the mean seed weight of each accession was determined by weighing ten times 20 seeds (precision 0.0001 g) and included in the analysis to further increase its discriminant power (for the list of all variables *see* Appendix).

The resulting data were statistically analysed using the stepwise linear discriminant analysis (LDA) — included in the SPSS (release 15) software package (SPSS 1999) — which is commonly used to classify/identify unknown groups characterized by quantitative and qualitative variables (Fisher 1936, 1940). To find the best features for seed identification using this procedure, tolerance, *F*-to-enter and *F*-to-remove were calculated for each measured feature (*see* Appendix). Tolerance indicates the proportion of a variable variance not accounted for by other independent variables. A variable with very low tolerance contributes little to the model. *F*-to-enter and *F*-to-remove define the power of each variable in the model. Stepwise LDA starts with a model that does not include any of the variables. At each step, a variable with the *F*-to-enter value ≥ 3.84 is added to the model. The variables whose $F < 3.84$ after the last step are left out of the model. The process was automatically stopped when no remaining variables increased the discrimination ability (Bacchetta *et al*. 2011, Grillo *et al*. 2011).

A cross-validation procedure was applied to verify the performance of the obtained classifiers. This method, when applied to small data sets, analyzes the individual data points and classifies them in relation to all other points (SPSS 1999). To graphically highlight the differences among groups, a multidimensional plot was drawn using the two available discriminant functions (Fisher 1936, 1940) and the Mahalanobis distance (Mahalanobis 1936), a measure between two data points in the space defined by two or more discriminant functions; a higher value indicates that a particular case includes extreme values for one or more independent variables, and can be considered dissimilar to other cases (Bacchetta *et al*. 2008, Venora *et al*. 2009b).

Results

Using stepwise LDA, of the 34 variables 18 were found best for seed identification (Table 2). The principal discriminating features were colorimetric and densitometric, even though the first parameter selected by the model on the basis of its discriminatory power was the mean seed weight (SW), which was also the parameter with the highest *F*-to-remove value (Table 2).

A cross-validation procedure indicated that using the selected set of variables, 99.8% of the samples were correctly classified (Table 3 and Fig. 2). Only one seed of *A. terraccianoi* was misidentified as *A. thermensis*, but no *A. thermensis* and *A. tegulensis* seeds were misclassified.

To evaluate the inter-population variability in *A. terraccianoi*, the seeds of the three investigated populations (Fig. 1) were also analysed giving an identification score of 92.5%, with the Corsican population differing from the Sardinian ones (Table 4 and Fig. 3).

Conclusions

The differences between *A. terraccianoi* and *A. tegulensis* previously suggested by Bacchetta and Brullo (2010) were confirmed by the results. *Astragalus tegulensis* was supported as a distinct species from Sardinia in *Astragalus* sect. *Melanocercis*. The comparison of the seed features of these two species with those of *A. thermensis* confirmed the differences as previously suggested by Valsecchi (1994). The Corsican population of *A. terraccianoi* is well defined not only geographically but also by the morphometric and colorimetric seed characters, while the two Sardinian populations, although distinguishable, are not sufficiently isolated or geographically separated to introduce clear phenotypic differences, or perhaps they are not two genetically distinct populations.

Table 2. Mean values and standard deviations of morpho-colorimetric variables of the analysed *Astragalus* species from Corsica and Sardinia, and ranking of the 18 selected features (in boldface) after 22 cycles of the stepwise linear discriminant analysis (for the feature abbreviations and definitions *see* Appendix).

Feature	*A. terraccianoi*		*A. tegulensis*		*A. thermensis*		Tolerance	*F* to remove	Wilks' λ
	Mean	SD	Mean	SD	Mean	SD			
A	4.16	0.66	4.53	0.51	5.01	0.82	–	–	–
P	8.00	0.67	8.41	0.57	8.76	0.76	–	–	–
P_{conv}	7.80	0.62	8.19	0.52	8.55	0.71	–	–	–
P_{crof}	7.58	0.63	7.97	0.54	8.31	0.72	–	–	–
P_{conv}/P_{crof}	1.03	0.01	1.03	0.01	1.03	0.01	–	–	–
D_{max}	2.82	0.27	3.00	0.28	3.12	0.31	0.006	4.170	0.031
D_{min}	2.03	0.20	2.06	0.15	2.20	0.18	0.079	5.361	0.031
D_{min}/D_{max}	0.71	0.08	0.69	0.09	0.71	0.06	–	–	–
Sf	0.91	0.03	0.89	0.04	0.91	0.03	0.414	3.898	0.031
Rf	0.67	0.07	0.65	0.07	0.65	0.05	0.036	5.649	0.031
Ecd	2.30	0.18	2.40	0.13	2.52	0.21	–	–	–
EA_{max}	1.37	0.15	1.48	0.16	1.53	0.16	0.012	13.239	0.032
EA_{min}	0.98	0.10	0.99	0.07	1.05	0.09	–	–	–
R_{avg}	184.04	18.05	193.09	15.48	193.91	18.22	–	–	–
R_{di}	28.47	7.89	26.59	4.61	39.91	5.20	–	–	–
G_{avg}	135.21	12.76	129.51	17.38	151.97	14.35	0.046	70.712	0.039
G_{di}	27.84	4.68	33.10	4.93	33.08	4.90	0.239	9.921	0.032
B_{avg}	91.96	9.61	74.05	6.09	98.81	7.83	0.176	189.665	0.053
B_{di}	23.74	6.64	32.30	2.24	17.56	3.10	0.278	12.470	0.032
H_{avg}	21.58	5.39	25.21	9.65	24.71	2.54	–	–	–
H_{di}	15.66	15.54	24.23	26.07	17.53	9.05	0.261	31.583	0.034
L_{avg}	137.72	11.14	133.10	10.14	146.09	12.26	–	–	–
L_{di}	24.35	4.05	26.52	2.74	27.86	3.46	–	–	–
S_{avg}	114.53	27.24	144.83	16.35	132.69	32.38	–	–	–
S_{di}	29.56	8.54	25.32	6.64	47.67	15.14	0.055	31.744	0.034
D_{avg}	91.28	11.49	101.05	7.86	101.16	14.98	0.010	5.519	0.031
D_{di}	57.53	7.37	62.01	5.57	64.87	12.14	0.009	10.836	0.032
S	4.06	5.12	2.77	4.74	8.58	5.40	–	–	–
K	44.50	66.12	31.58	63.01	108.93	89.57	–	–	–
H	0.011328	0.002126	0.012006	0.001961	0.012861	0.011807	0.510	7.607	0.031
E	5.08	0.27	5.23	0.18	5.13	0.30	0.237	13.430	0.032
D_{sum}	23766.73	5092.42	28658.41	4500.50	31935.60	8378.07	0.002	22.590	0.033
$\Sigma(D_{sum})^2$	3086646.79	1003168.98	4023582.32	946694.64	4729456.16	1944426.29	0.004	25.420	0.033
SW	0.00395	0.00052	0.00442	0.00007	0.00536	0.00013	0.635	458.373	0.084

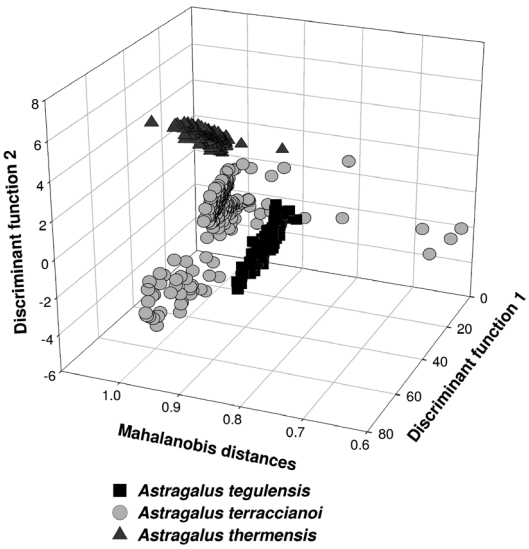

Fig. 2. Discriminant function scores for the *Astragalus* species from Corsica and Sardinia.

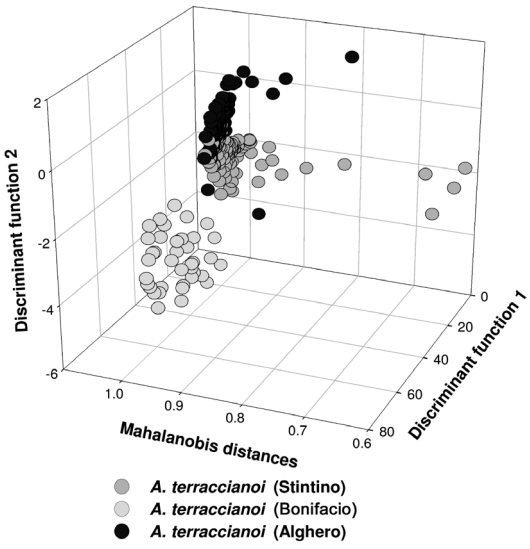

Fig. 3. Discriminant function scores for *Astragalus terraccianoi* populations from Corsica and Sardinia.

Table 3. Cross-validated percentages of correct classification for the *Astragalus* species classifier. The number of seeds in parentheses.

Species	*A. terraccianoi*	*A. thermensis*	*A. tegulensis*	Total
A. terraccianoi	99.6 (238)	0.4 (1)	0	
A. thermensis	0	100.0 (200)	0	
A. tegulensis	0	0	100.0 (100)	
Overall				99.8 (539)

Table 4. Cross-validated percentages of correct classification for the *Astragalus terraccianoi* populations classifier. The number of seeds in parentheses.

Species	*A. terraccianoi* (Stintino, Sardinia)	*A. terraccianoi* (Bonifacio, Corsica)	*A. terraccianoi* (Alghero, Sardinia)	Total
A. terraccianoi (Stintino, Sardinia)	95.0 (95)	0	5.0 (5)	
A. terraccianoi (Bonifacio, Corsica)	0	100.0 (100)	0	
A. terraccianoi (Alghero, Sardinia)	33.3 (13)	0	66.7 (26)	
Overall				92.5 (239)

Acknowledgements

The Centre for the Conservation of Biodiversity (CCB) is supported by Provincia di Cagliari — Assessorato Tutela Ambiente. This research was also supported by Ente Foreste della Sardegna. Thanks also to Regione Autonoma Sardegna for the support to the realization of this work, on the basis of Legge Regionale 7 agosto 2007, n. 7 "Promozione della ricerca scientifica e dell'innovazione tecnologica in Sardegna", pubblica selezione per il conferimento di borse di ricerca destinate a giovani ricercatori — Bando 2008.

References

Bacchetta, G. & Brullo, S. 2010: *Astragalus tegulensis* Bacch. & Brullo (Fabaceae), a new species from Sardinia. — *Candollea* 65: 5–14.

Bacchetta, G., Grillo, O., Mattana, E. & Venora, G. 2008: Morpho-colorimetric characterization by image analysis to identify diaspores of wild plant species. — *Flora* 203: 669–682.

Bacchetta, G., Escobar Garcia, P., Grillo, O., Mascia, F. &

Venora, G. 2011: Evidence on taxonomical differentiation inside the genus *Lavatera* L. sect. *Glandulosae* (Malvaceae) from germplasm image analysis. — *Flora* 206: 468–472.

Chamberlain, D. F. & Matthews, W. A. 1969: *Astragalus* L. — In: Davis, P. H. (ed.), *Flora of Turkey and the East Aegean Islands*, 3rd ed.: 49–253. Edinburgh University Press, Edinburgh.

Chater, A. O. 1968: *Astragalus* L. — In: Tutin, T. G., Heywood, V. H., Burges, N. A., Moore, D. M., Valentine, D. H., Walters, S. M. & Webb, D. A. (eds.), *Flora Europaea*, 2nd ed.: 108–124. Cambridge University Press, Cambridge.

Fisher, R. A. 1936: The use of multiple measurements in taxonomic problems. — *Annals of Eugenics* 7: 179–188.

Fisher, R. A. 1940: The precision of discriminant functions. — *Annals of Eugenics* 10: 422–429.

Giardina, G., Raimondo, F. M. & Spadaro, V. 2007: A catalogue of plants growing in Sicily. — *Bocconea* 20: 5–582.

Grillo, O., Mattana, E., Venora, G. & Bacchetta, G. 2010: Statistical seed classifiers of 10 plant families representative of the Mediterranean vascular flora. — *Seed Science and Technology* 38: 455–476.

Grillo, O., Miceli, C. & Venora, G. 2011: Image analysis tool for vetch varieties identification by seeds inspection. — *Seed Science and Technology* 39: 490–500.

Lewis, G., Schrire, B., Mackinder, B. & Lock, M. 2005: *Legumes of the world*. — Royal Botanic Gardens, Kew.

Mahalanobis, P. C. 1936: On the generalised distance in statistics. — *Proceedings of National Institute of Science of India* 12: 49–55.

Mattana, E., Grillo, O., Venora, G. & Bacchetta, G. 2008: Germplasm image analysis of *Astragalus maritimus* and *A. verrucosus* of Sardinia (subgenus *Trimeniaeus*, Fabaceae). — *Anales del Jardín Botánico de Madrid* 65: 149–155.

Podlech, D. 1999: *Thesaurus Astragalorum*. — Available at http://www.sysbot.biologie.uni-muenchen.de/de/personen/podlech/thesaurus_astragalus.pdf.

Podlech, D. 2008: The genus *Astragalus* L. (Fabaceae) in Europe with exclusion of the former Soviet Union. — *Feddes Repertorium* 119: 310–387.

Shahin, M. A. & Symons, S. J. 2003: Colour calibration of scanners for scanner independent grain grading. — *Cereal Chemistry* 80: 285–289.

SPSS 1999. *Application guide*. — Prentice Hall, Englewood Cliffs, NJ.

Valsecchi, F. 1994: Sul complesso *Astragalus tragacantha* L. (Leguminosae) nel Mediterraneo. — *Webbia* 49: 31–41.

Venora, G., Grillo, O., Ravalli, C. & Cremonini, R. 2009a: Identification of Italian landraces of beans (*Phaseolus vulgaris* L.), using an image analysis system. — *Sciencia Horticulturae* 121: 410–418.

Venora, G., Grillo, O. & Saccone, R. 2009b: Durum wheat storage centers: evaluation of vitreous, starchy and shrunken kernels by image analysis system. — *Journal of Cereal Science* 49: 429–440.

Appendix. List of the morphometric and colorimetric features of seeds of *Astragalus* species from Corsica and Sardinia (Grillo *et al.* 2010).

Morphometric features	
A	Seed area (mm^2)
P	Seed perimeter (mm)
P_{conv}	Convex perimeter of the seed (mm)
P_{crof}	Crofton's perimeter of the seed (mm)
P_{conv}/P_{crof}	Ratio between convex and Crofton's perimeters
D_{max}	Maximum diameter of the seed (mm)
D_{min}	Minimum diameter of the seed (mm)
D_{min}/D_{max}	Feret ratio (ratio between minimum and maximum diameters)
Sf	Seed shape factor = $4\pi A/P^2$
Rf	Seed roundness factor = $4A/[\pi(D_{max})^2]$
Ecd	Equivalent circular diameter (diameter of a circle with equivalent area) (mm)
EA_{max}	Maximum axis of an ellipse with equivalent area (mm)
EA_{min}	Minimum axis of an ellipse with equivalent area (mm)
SW	Seed mean weight (g)

Colorimetric features of seed-image pixels	
R_{avg}	Average value of red
R_{di}	Red dispersion index
G_{avg}	Average value of green
G_{di}	Green dispersion index
B_{avg}	Average value of blue
B_{di}	Blue dispersion index
H_{avg}	Average value of hue
H_{di}	Hue dispersion index
L_{avg}	Average value of lightness
L_{di}	Lightness dispersion index
S_{avg}	Average value of saturation
S_{di}	Saturation dispersion index
D_{avg}	Average value of density
D_{di}	Density dispersion index
S	Distribution skewness
K	Distribution kurtosis
H	Intensity value of grey
E	Dispersion of the density values
D_{sum}	Sum of density values
$\Sigma(D_{sum})^2$	Sum of squared density sums