J. L. Sánchez, G. Domina & J. Caujapé-Castells

The Sicilian circumscription of the genus *Matthiola* (*Brassicaceae*): population genetic insights from isozymes

Abstract

Sánchez, J. L., Domina, G., Cujapé-Catells, J.: The Sicilian circumscription of the genus *Matthiola* (*Brassicaceae*): population genetic insights from isozymes. — Bocconea 19: 243-252. 2006. — ISSN 1120-4060.

We examined the genetic variation at 12 isozyme loci at 14 populations of *Matthiola* (*Brassicaceae*) representing the distribution of the species *Matthiola incana, M. fruticulosa* and *M. tricuspidata* in the Sicilian insular system and adjacent mainland areas to gain insight into the levels of variation. While the low polymorphism indicators for *M. incana* subsp. *incana* are consistent with its presumptive origin through cultivation and the associated lack of gene flow, *M. incana* subsp. *pulchella* and *M. incana* subsp. *rupestris* have only slightly higher indices of polymorphism and they are probably neo-endemics. By contrast, the Sicilian populations *M. tricuspidata* and *M. fruticulosa* subsp. *fruticulosa* have maintained a moderate-to-high polymorphism through an abundant genetic interchange. These two latter species are also closer genetically than each of them to *M. incana*.

Introduction

The study of plant diversification on islands has proven of outstanding value to understand population genetic processes in space and time (Givnish 1998). In contrast to mainland areas, islands represent closed systems where the influx of immigrants is not constant, and they usually have different habitats concentrated in a narrow area where we can link the study of the mode and tempo of biological diversification (in most cases, the approximate ages of the islands are known).

However, insular systems do not constitute a uniform testing ground for biological diversification, as they show differences that stem from (among many other factors) their degree of connection to the landmasses. Based on this factor, a sharp distinction is made between oceanic and continental islands depending, respectively, on whether they have been attached to the continent in the past, or not (Moore 1979).

Up to now, studies that analyse comparatively the genetic structure of plant populations in continental vs. oceanic islands are scarce or inexistent. Therefore, we are at a loss of

insight into the relative importance of intrinsic and extrinsic factors on plant diversification in islands of different nature.

The genus *Matthiola* R. Br. (*Brassicaceae*) is mainly Eurasian, and contains ca. 50 species of annual and biannual herbs, and perennial sub-shrubs characterised by a silicle that features horns or swellings in its apical end (Appel & Al-Shehbaz 2003). Because of its distribution, encompassing mainland and island areas from Macaronesia to Mediterranean Northeast Africa, *Matthiola* represents an adequate model system to compare the different patterns of diversification in continental vs. oceanic islands.

In this contribution, we report the first data on the levels of molecular population genetic (isozyme) variation of these three species: *Matthiola incana*, *M. tricuspidata* and *M. fruticulosa* from the continental island of Sicily, and the surrounding oceanic islands of Salina and Pantelleria and several mainland areas from peninsular Italy. These results will be integrated into a broader body of comparative molecular genetic and morphological data corresponding to other species of *Matthiola* with a view to understand the differences in plant diversification processes between Macaronesian and Mediterranean islands.

Material and Methods

We sampled 14 populations of *Matthiola* (Table 1) representing the distribution of *M. incana*, *M. fruticulosa* and *M. tricuspidata* in the Sicilian insular system and in peninsular Italy. Although there is one known population of *M. sinuata* in Sicily, we did not sample it because of its extreme exiguity. All sampled populations of *M. incana* subsp. *incana* are in all likelihood ornamentals that have escaped cultivation and naturalised, whereas the remaining populations included in this investigation are wild. We collected 2Kg of soil at each of the sampled populations (except for Salina) for edaphic analyses on 14 variables (Table 2).

Horizontal starch-gel electrophoreses proceeded according to the specifications in Table 3. Figure 1 illustrates the average resolution of the gels through the patterns obtained for enzymes PGM and PGI, as well as the interpretation of the two pictured gels in the computer program Transformer (Caujapé-Castells 2001), that was used to format the arising data according to the specifications of the different population genetic programs used for statistical analysis.

Average number of alleles per locus (A), percentage of polymorphic loci (P) and expected and observed heterozygosity (He and Ho) were calculated from genotype data using the computer program BIOSYS-1 version 1.7 (Swofford & Selander 1988). Genestat (Lewis 1993) was used to obtain the values of Gst (Nei 1973), that estimate the proportion of genetic variation that resides among populations. The matrix of allele frequencies was implemented in NTSYS-PC (Rohlf 1993) to carry out a Principal Component Analysis (PCA) based on the variance-covariance matrix. Gene flow was estimated according to the values of Nm (Wright 1931) using the computer program Popgene (Yeh & al. 1997).

The relationship among geographic location, genetic makeup and edaphic composition of the five areas sampled was further evaluated through Mantel tests (Mantel 1967). This is a non-parametric statistical procedure that calculates the correlation between pairs of distance matrices and tests its significance. Mantel analyses involved matrices of geo-

Table 1. Sampling details of the 14 populations sampled (including the life form of each species) and basic indicators of population genetic polymorphism as estimated from the isozyme analyses, **A:** average number of alleles per locus; **P:** percentage of polymorphic loci; **Ho, He:** observed and expected heterozygosity; **N:** the population numerical code (it corresponds with those in Fig. 1); **n:** he number of samples analysed per population; ***** labels the mainland populations. Letters in parenthesis after the population names are the population codes assigned.

N	Species / Life form / Locality	n	A	P	Ho	He
	***M. incana* ssp. *incana* / Perennial subshrub**					
1.	Salina (IIPS)	28	1.3	8.3	0.003	0.050
2.	Palmi* (IIPA)	21	1.0	0.0	0.000	0.000
3.	Balestrate (IIBA)	25	1.0	0.0	0.000	0.000
4.	Monte Pellegrino (IIMP)	30	1.0	0.0	0.000	0.000
5.	Pizzo Calabro* (IIPC)	27	1.0	0.0	0.000	0.000
	***M. incana* ssp. *rupestris* / Perennial subshrub**					
6.	Capo Zafferano (IRCZ)	22	1.0	0.0	0.000	0.000
	M. incana ssp. *pulchella* / Perennial subshrub					
7.	Pantelleria (MIPP)	21	1.1	8.3	0.017	0.015
	***M. tricuspidata* / Annual-biannual**					
8.	Isola delle femine (TRIF)	12	1.3	8.3	0.063	0.052
9.	Campofelice di Roccella (TRCR)	16	1.3	16.7	0.063	0.074
10.	Castel di Tusa (TRCT)	27	1.3	25.0	0.093	0.113
11.	San Ferdinando Calabro* (TRSF)	32	1.7	33.3	0.050	0.144
12.	San Isidoro* (TRSI)	30	1.6	41.7	0.090	0.134
	***M. fruticulosa* ssp. *fruticulosa* / Perennial herb**					
13.	Vallone Madonna degli Angeli (FFMA)	25	2.0	50.0	0.159	0.190
14.	San Martino delle Scale (FFSM)	27	1.7	33.3	0.192	0.173

graphic (in km), genetic (Nei 1978) and edaphic distances between all pairwise combinations of zones. Edaphic distances were calculated using the option 'taxonomic distance' in the computer program NTSYS-pc (Rohlf 1993) after standardizing the values of the 14 variables derived from physico-chemical soil analyses (Table 2).

Results

The higher levels of polymorphism were detected in *M. fruticulosa* subsp. *fruticulosa*, with values of average number of alleles per locus (*A*) and expected heterozygosity (*He*) of A = 2.4 and H_e= 0.201 (Table 1, Fig. 2) this is also testified by the large number of subspecific taxa reported from Sicily (cfr. Lojacono 1888). The lowest levels of polymorphism were found in *M. incana* subsp. *incana*, with population averages of A = 1.06 and H_e = 0.01. However, when all the populations of *M. incana* subsp. *incana* were merged in a single assemblage (Fig. 2), the values of polymorphism increased sharply

Table 2. Values of the 14 edaphic variables obtained after the soil analyses in the 13 populations for which we could collect soil samples (all except for Salina). CTY: conductivity; OM: Organic matter. OM, C, and the physical analyses are expressed in %. Amounts of K, Na, Ca and Mg are given in meq/100g. Amounts of P an N are given in parts per million. Population codes are those in Table 1.

	Chemical analyses											Physical analyses		
Population	**pH**	**CTY**	**$CaCO_3$**	**C**	**OM**	**K**	**Na**	**Ca**	**Mg**	**P**	**N**	**Clay**	**Lime**	**Sand**
IIPA	7.60	228	0.44	0.26	0.48	0.5	0.86	13.60	4.40	11.00	35	39.20	17.90	25.95
IIBA	8.18	103	24.64	0.21	0.36	0.1	0.31	11.20	0.80	2.39	5	13.57	4.59	7.60
IIMP	7.69	155	0.66	1.61	2.77	0.39	0.28	24.60	6.75	17.40	25	8.20	10.07	15.35
IIPC	7.98	170	7.70	0.05	0.09	0.37	0.76	9.80	3.00	7.70	10	18.05	6.97	74.90
IRCZ	8.00	147	0.22	1.31	2.26	1.63	0.54	8.42	4.74	11.20	10	26.45	13.87	27.65
MIPP	7.51	3,020	0.44	0.60	1.03	0.9	11.10	6.00	2.20	20.30	375	8.52	7.02	50.55
TRIF	7.68	177	25.96	0.64	1.10	0.2	0.50	9.40	1.33	23.60	15	9.85	0.90	88.15
TRCR	7.76	90	15.40	0.45	0.78	0.1	0.13	8.82	0.80	8.00	5	2.35	3.30	80.10
TRCT	7.85	78	3.30	0.22	0.38	0.1	0.10	6.40	0.82	4.00	0	1.77	4.69	91.65
TRSF	7.55	201	0.66	0.05	0.08	0.12	0.58	1.35	0.65	9.70	20	8.27	2.90	87.45
TRSI	7.89	243	22.44	1.05	1.81	0.42	0.49	11.10	3.50	10.70	221	6.55	7.27	66.00
FFMA	8.08	42	8.14	0.22	0.38	0.1	0.10	6.30	2.75	3.00	0	16.05	2.90	64.80
FFSM	7.31	197	1.76	0.68	1.17	0.1	0.17	6.60	2.26	26.00	44	13.05	3.40	41.60

(A = 1.6 and H_e = 0.131) by virtue of the conspicuous qualitative differences in allelic composition among them.

The average value of Gst was high in *M. incana* subsp. *incana* (*Gst* = 0.759) and only moderate in *M. tricuspida0ta* (*Gst* = 0.275) and *M. fruticulosa* subsp. *fruticulosa* (*Gst* = 0.193), despite their much higher polymorphism (Fig. 2).

Estimates of gene flow through *Nm* values (Wright, 1931) between pair-wise combinations of populations were considerably low in *M. incana* subsp. *incana* (average *Nm* =

Table 3. Gel/buffer systems used and isozymes resolved in this investigation. Slice: part of the gel (from top to bottom) used to resolve the corresponding enzyme after gel slicing (any of the gels used can consist of up to four stainable slices); **EC code:** numerical sequence used by the enzyme commission to designate the enzyme; **N:** number of loci resolved per enzyme; **A:** number of alleles per locus over all populations; **Q:** quaternary structure of the enzyme (**d:** dimer, **m:** monomer).

Gel / enzyme	**Slice**	**Abbreviation**	**EC code**	**N**	**A**	**Q**
Morpholine-Citrate 6.1						
6-phosphogluconate-dehydrogenase	1st	6-PGD	1.1.1.44	2	(2,3)	d
Isocitrate dehydrogenase	2nd	IDH	1.1.1.42	2	(3,1)	d
Malate dehydrogenase	4th	MDH	1.1.1.37	4	(2,3,1,3)	d/m
Histidine 7.0						
Phosphoglucose isomerase	1st	PGI	5.3.1.9	1	6	d
Fosfoglucomutase	2nd	PGM	5.4.2.2	3	(5,7,4)	m

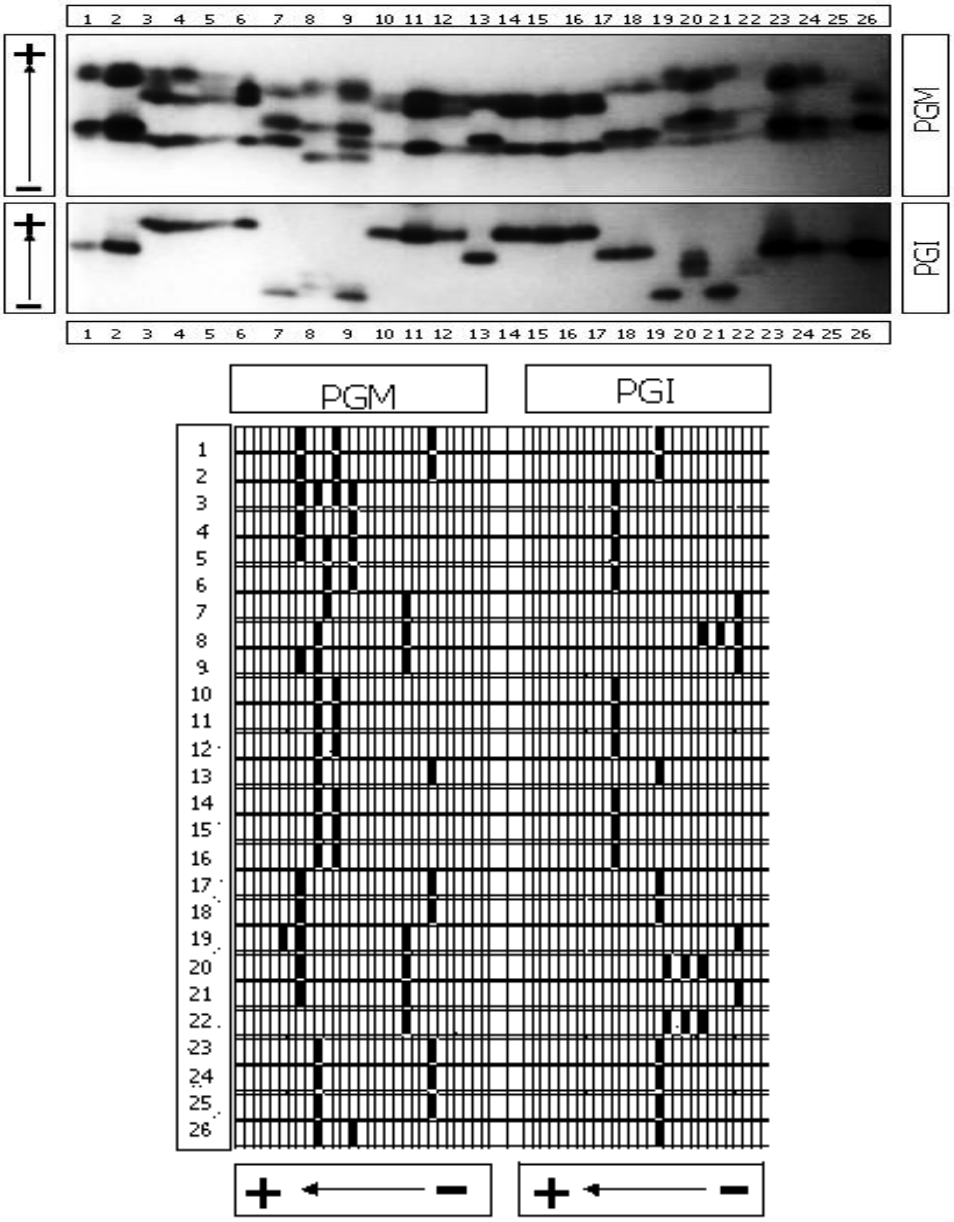

Fig. 1. Average resolution of the enzymes PGM and PGI for 26 individuals, and interpretation of their genetic patterns in the program Transformer-1 (Caujapé-Castells 2001). In the pattern of PGM, bands with different tonalities correspond to each of the three loci interpreted. The 26 individuals represented in these gels are a mixture of all the species and populations included in the paper. According to the codes in Table 1, lanes represent different individuals from the following populations: IIPS (1 and 2), TRIF (3 and 4), TRCR (5 and 6), FFMA (7 to 9), TRSF (10 to 12), IRCZ (13), TRSI (14 and 15), IIPA (16 and 17), FFSM (18 to 21), IIMP (22 to 24), and MIPP (25 and 26).

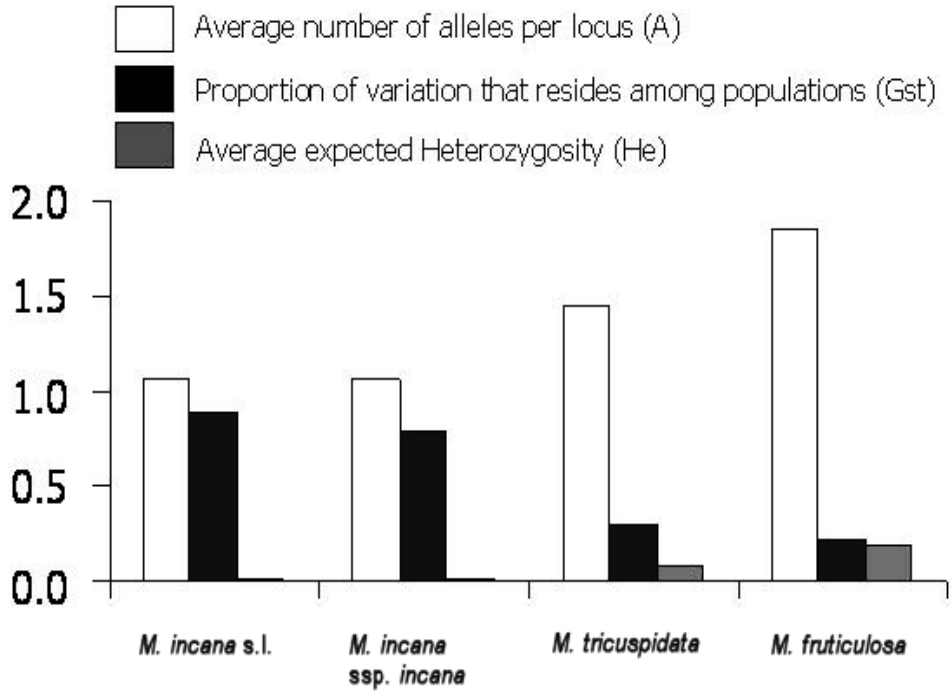

Fig. 2. Graphic representation of the basic parameters that estimate the levels and distribution of genetic variation in *Matthiola incana* s. l., *M. incana* subsp *incana*, *M. tricuspidata* and *M. fruticulosa* subsp *fruticulosa*.

0.05) but surpassed 1 in *M. fruticulosa* and *M. tricuspidata* (*Nm* = 3.37 and 5.46, respectively). *Matthiola tricuspidata* exhibits a clear (though non significant) trend to a diminished gene flow as geographic distance increases, which is less apparent in *M. incana* s.l. (Fig. 3a).

All the populations of *M. incana* clustered in the middle of the multivariate space of the PCA (Fig. 4). By contrast, populations of *M. tricuspidata* and *M. fruticulosa* showed a much higher degree of dispersal in this representation, with a trend to a closer relationship between them than with respect to *M. incana*.

Mantel tests for *M. incana* subsp. *incana* and *M. tricuspidata* did not reveal significant correlation between geographic and genetic distances (r = -0.551, P = 0.446 and r = 0.522, P = 0.498, respectively) or between genetic and edaphic distances (r = 0.299, P = 350 and r = -0.275, P = 0.399, respectively).

Discussion

Although the results presented here need be streamlined by additional data analyses that also consider a broader geographical context, they may be used to derive several preliminary insights into the diversification of *Matthiola* in Sicily.

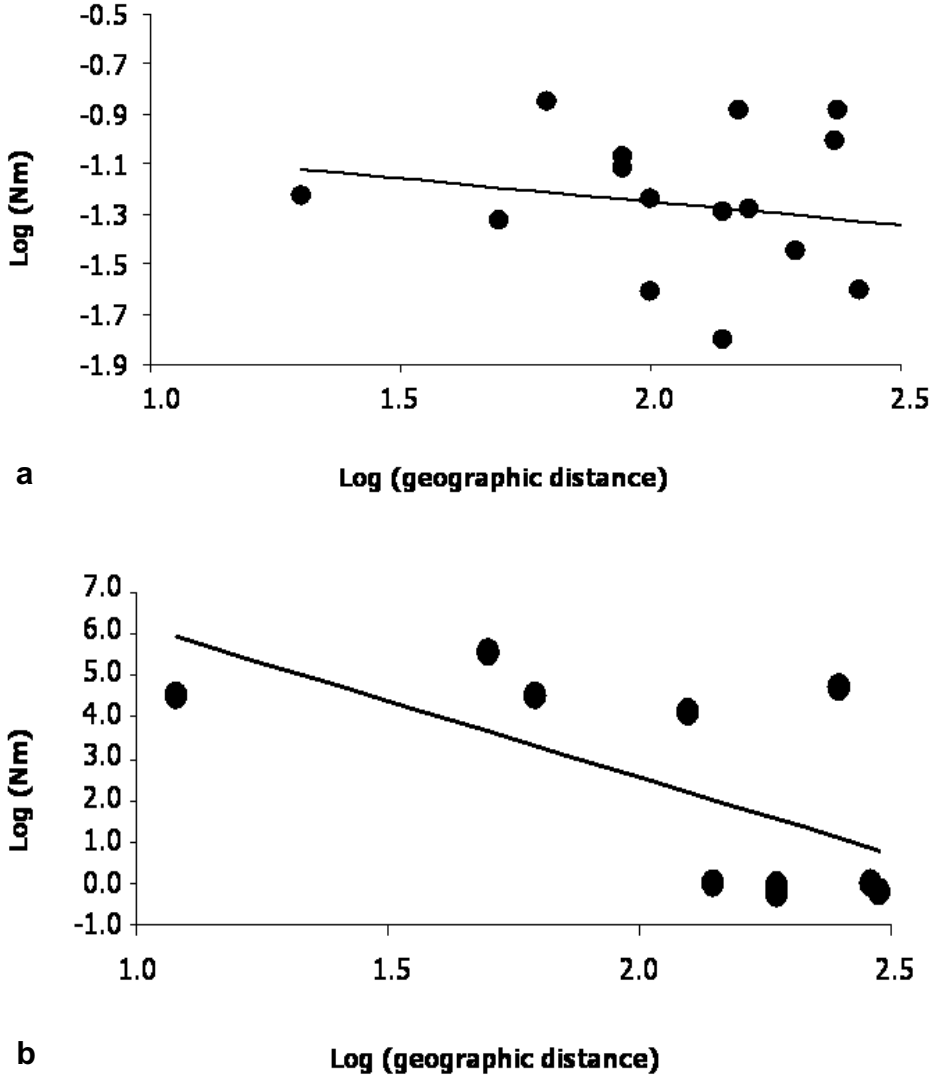

Fig. 3. Relationship between geographic distance and gene flow (as measured by Nm) in pair-wise combinations of populations of *M. incana* s. l. (a) and *M. tricuspidata* (b).

The striking monomorphism of almost all populations of *M. incana* analysed hints either at a very recent origin through founder events or to a high incidence of inbreeding. Although we are not in the position of favouring any of these hypotheses at the present stage of knowledge, it is noteworthy that even the continental populations manifest low levels of genetic variation, a fact that probably bears relationship with their origin in cultivation and a diminished level of gene flow, as revealed by the high Gst value in this taxon

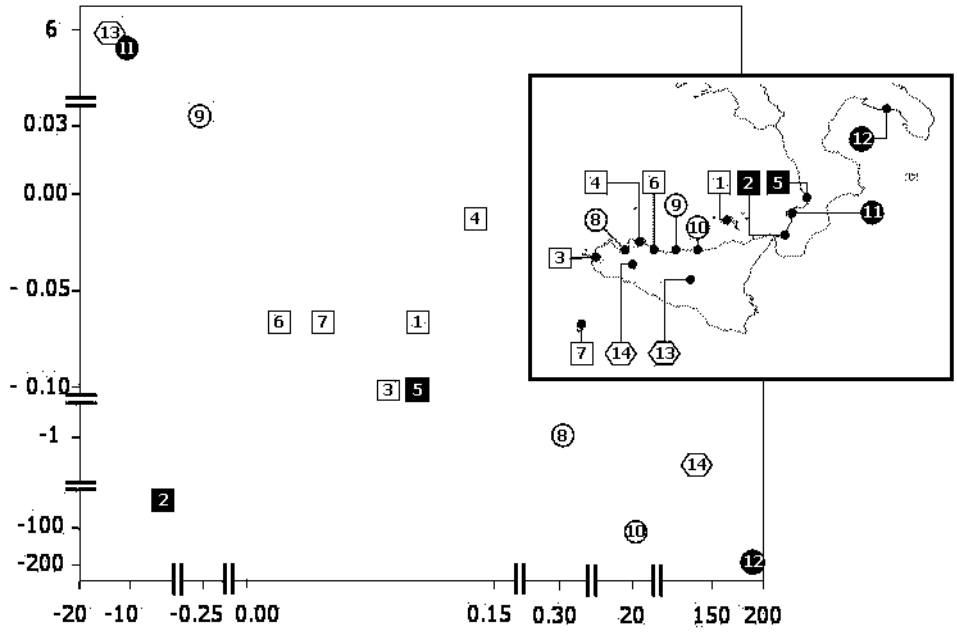

Fig. 4. Geographic distribution of the 14 populations included in this study, and Principal Component Analysis based on the allele frequencies detected. Due to the wide range of the resulting population coordinates we had to break the axes at various points to allow for representation. Population numbers correspond to Table 1. The two axes of the PCA representation explains 25% of the detected isozyme variation.

(Fig. 2). There are substantial qualitative differences among several island populations of *M. incana*, particularly (but not only) affecting subsp. *pulchella*. In the lack of evidence for the action of selection, this result entails either (a) an origin through various founder events from a genetically polymorphic mainland source that we did not sample or (b) contrasting geographic origins. Given the scarce polymorphism of the mainland populations sampled, the generalised low levels of gene flow in *M. incana* and its non-significant relationship with geographic distance (Fig. 3a), the second scenario seems better suited to construe these results.

The island populations of *M. tricuspidata* display a clear trend to having less genetic variability (e. g., in terms of the number of polymorphic loci) with increasing distance to the continent (Fig. 3b). This fact, together with the absence of population-specific exclusive alleles, suggests a diversification scenario of isolation by distance, where geographically closer populations also bear a tighter genetic relationship because of a more abundant gene flow between them.

Even though the only two sampled populations of *M. fruticulosa* subsp. *fruticulosa* are insular (Fig. 4), they hold the highest levels of isozyme variation detected (Fig. 2). This higher polymorphism, together with a low value of *Gst* in this species (*Gst* = 0.193), hints at more developed dispersal and outcrossing capabilities than its two congeners.

Nevertheless, sampling several continental populations of this taxon would be necessary to streamline this hypothesis.

Finally, although the three sampled taxa represent different sections within *Matthiola*, there seems to be a higher genetic affinity between sect. *Acinotum* (*M. fruticulosa*) and sect. *Aciloma* (*M. tricuspidata*) than between any of these two and sect. *Pachynotum* (*M. incana*), as revealed by the tighter grouping of the former two taxa (and their larger isolation with respect to *M. incana*) in the PCA (Fig. 4).

Taxonomic insights

Because the relationship between phenotype and genotype is simpler for isozyme evidence than for morphological characters (Gottlieb, 1977), our study might provide insight to assess the existing taxonomic categories within *M. incana*. However, we must exercise extreme caution in the face that the differentiation of subsp. *incana* respect to subsp. *pulchella* and subsp. *rupestris* could be lopsided toward a higher morphological divergence in the former subspecies. By and large, many genes contribute to the expression of morphological traits, and these are liable to accumulate variation more rapidly than single-gene markers as allozymes, even under the influence of drift (Barrett & Kohn 1991). In addition to this general factor, morphological divergence could have been much enhanced in subsp. *incana* because unadaptive genes have been selected against. Thus, it is likely that isozyme differentiation in this taxon (which has been kept low through repeated inbreeding) does not reflect accurately the degree of morphological divergence that determines its taxonomic differentiation.

Bearing the above in mind, the average genetic identity within *M. incana* subsp. *incana* ($I = 0.90$) falls within the range reported by Gottlieb (1977) for con-specific populations ($I = 0.95 \pm 0.2$). The genetic differentiation between the populations of subsp. *rupestris* and subsp. *pulchella* ($I = 0.756$) is higher than that of either of them with 4 *incana* ($I = 0.813$ and $I = 0.840$, respectively). However, these three values lie between Gottlieb's (1977) averages for con-specific populations ($I = 0.95 \pm 0.2$) and congeneric species ($I = 0.67 \pm 0.07$). The values reported in Gottlieb's (1977) review might change if updated; however, assuming that the eventual differences would not be substantial, our isozyme survey suggests that it is advisable to maintain the present infra-specific taxonomy within *M. incana*.

Acknowledgements

This research was partly supported by a Ramón y Cajal project co-funded by the Spanish Ministerio de Ciencia y Tecnología (MCYT) and the Cabildo Insular de Gran Canaria (CIGC) to JCC, and by a Tomás de Zárate grant of the CIGC to JLS. We thank the MCYT and the CIGC for their support to carry out this research at the Jardín Botánico Canario "Viera y Clavijo" (JBCVC). We are grateful to the biologists and students at the JBCVC and at the Dipartimento di Scienze Botaniche of the University of Palermo for meaningful comments on a previous draft of this paper. We thank Pepe Cruz, Antonio Marrero, Domingo Cabrera and Nuria Sánchez, at the greenhouses of the JBCVC, for their help and expertise in the cultivation of the seeds of *Matthiola*, and Leticia Curbelo for excellent technical assistance in the lab. We also thank the Granja Agrícola Experimental of the Cabildo Insular de Gran Canaria for soil analyses.

References

Appel, O. & Al-Shehbaz., I. A.2003: *Cruciferae*. – Pp. 75-174 in Kubitzki, K. & Bayer, C. (eds.), The Families and Genera of Vascular Plants, **5**. – Berlin.

Barrett, S. C. H. & Kohn., J. R. 1991: Genetic and evolutionary consequences of small population size in plants. – Pp. 3-30 in Falk, D. A. & Holsinger K. E. (eds.), Genetics and conservation of rare plants. – Oxford.

Caujapé-Castells, J. 2001: Transformer 1.0: un programa de formateado de datos genotípicos individuales para implementación en programas de análisis molecular en genética de poblaciones. – Distributed by the author.

Givnish, T. J. 1998: Adaptive plant evolution on islands: classical patterns, molecular data, new insights. – Pp. 284-301 in Grant, P. R. (ed.), Evolution on islands. – New York.

Gottlieb, L. D. 1977: Electrophoresis evidence and plant systematics. – Ann. Missouri Bot. Gard. **64:** 161-180.

Lewis, P. O. 1993: GeneStat-PC 3.31. – North Carolina State University. Raleigh. North Carolina.

Lojacono, M. 1888: *Matthiola* R. Br. – Pp. 99-102 in Flora Sicula, **1(1).** – Palermo.

Mantel, N. 1967: The detection of disease clustering and a generalized regression approach. – Cancer Res. **27:** 209 - 220.

Moore, D.M. 1979: Origin of temperate Island Floras. – Pp. 69-86 in Bramwell D. (ed.), Plants and Islands. – London.

Nei, M. 1973: Analysis of Gene Diversity in Subdivided Populations. – Proc. Natl. Acad. Sci. U.S.A. **70:** 3321-3323.

— 1978: Estimation of average heterozygosity and genetic distance from a small number of individuals. – Genetics **89:** 583-590.

Yeh, F., Yang, R. & Boyle, T. 1997: Popgene Version 1.32. Ag/For Molecular Biology and Biotechnology Centre (University of Alberta) and Center for International Forestry Research. Alberta, Canada.

Swofford, D. L. & Selander, R. B. 1988: BIOSYS-1: a computer program for the analysis of allelic variation in genetics. – Illinois.

Rohlf, F. J. 1993: NTSYS-pc. Numerical Taxonomy and Multivariate Analysis version 1.80. – Applied Biostatistics Inc.

Wright, S. 1931: Evolution in Mendelian populations. – Genetics **16:** 97-159.

Address of the authors:
Juan Luis Sánchez & Juli Caujapé-Castells*: Laboratorio de Biodiversidad Molecular y Banco de DNA, Jardín Botánico Canario "Viera y Clavijo", Ap. de correos 14 de Tafira Alta. 35017 Las Palmas de Gran Canaria, Spain. E-mail: jlsanchez1972@terra.es, julicaujape@grancanaria.com.
Gianniantonio Domina: Laboratorio di Sistematica, Fitogeografia ed Ecologia vegetale, Dipartimento di Scienze botaniche, Universitá degli Studi di Palermo. Via Archirafi 38. - 90123 Palermo, Italy. E-mail: gdomina@unipa.it.

*Corresponding author