

## The relationship between geographic range size and life history traits: is biogeographic history uncovered? A test using the Iberian butterflies

Enrique Garcia-Barros and Helena Romo Benito

*E. Garcia-Barros (garcia.barros@uam.es) and H. Romo Benito, Dept of Biology, Univ. Autónoma de Madrid, ES-28049 Madrid, Spain.*

The geographic range of a species is influenced by past phylogenetic and biogeographic patterns. However, other historical interactions, including the interplay between life history and geography, are also likely involved. Therefore, the range size of a species can be explained on the basis of niche-breadth or dispersal related hypotheses, and previous work on European butterflies suggests that both, under the respective guise of ecological specialisation and colonising ability may apply. In the present study, data from 205 species of butterflies from the Iberian peninsula were processed through multiple regression analyses to test for correlations between geographic range size, life history traits and geographic features of the species distribution types. In addition, the percentage of variance explained by the subsets of variables analyzed in the study, with and without control for phylogenetic effects was tested. Despite a complex pattern of bivariate correlations, we found that larval polyphagy was the single best correlate of range size, followed by dispersal. Models that combined both life history traits and geographic characteristics performed better than models generated independently. The combined variables explained at least 39% of the variance. Bivariate correlations between range size and body size, migratory habits or egg size primarily reflected taxonomic patterning and reciprocal correlations with larval diet breadth and adult phenology. Therefore, aspects of niche breadth i.e. potential larval diet breadth emerged as the most influential determinants of range size. However, the relationships between these types of ecological traits and biogeographic history must still be considered when associations between life history and range size are of interest.

Range size not only represents a proxy for species rarity (IUCN 2001), but also a “parameter” for patterns of diversity (Rosenzweig 1975, 2003, Gaston and Blackburn 1999, He and Legendre 2002, Gaston 2003, Storch et al. 2007). The geographic range of a species is determined by the interaction between habitat suitability and connectivity, species niche requirements and population dynamics (Gregory and Gaston 2000). Regardless of the complexity of the processes involved (Brown 1984, Gaston et al. 1997, 2000, Cowley et al. 2001b, Gilbert and Lechowicz 2004), these interactions operate via the biological properties of individual species. Hence, the relationships should be identifiable as correlations between species range size and ecological features (Brown 1984, 1995, Gaston and Blackburn 1994, Gaston et al. 2000, Diniz-Filho et al. 2005).

Large-scale inter-specific analyses of present species ranges are often hampered by three factors: 1) phylogenetic patterning; 2) geographic turnover in species life histories; and 3) the biogeographic history underlying present ranges. Taxonomic variability and phylogenetic relationships across a species range are routinely addressed (Carrascal et al. 2008, Gove et al. 2009, Calosi et al. 2010), however geographic turnover and biogeographic history remain

largely unresolved. First, species life histories exhibit geographic variability across a species range. Consequently, analyses of “mean” trait values measured across geographic gradients may suggest a misleading niche breadth-based explanation (Gaston et al. 2007). A provisional solution is to focus on intermediate scale patterns. Finally, evidence for hidden historical causes in observable ranges can be tested by comparing apparent relationships between range size and ecological factors with simple biogeographic features of the species range, such as the chorotype and range position (as shown by recent work on plant distributions; Weiser et al. 2007, Gove et al. 2009).

Butterflies are highly sensitive to environmental changes, particularly changes impacting vegetation because butterfly larvae are specialised herbivores. Several Lepidopteran life history traits display notable interspecific variation, as well as intraspecific differences along geographic gradients (Nylén 2009). Studies assessing butterfly range size and ecology across species have been reported in temperate areas, most notably the British Isles (Hodgson 1993, Quinn et al. 1998, Dennis et al. 2000, 2004, 2005, Cowley et al. 2001a, b), among others (Hughes 2000, Komonen et al. 2004). Dennis et al. (2004) provides evidence that

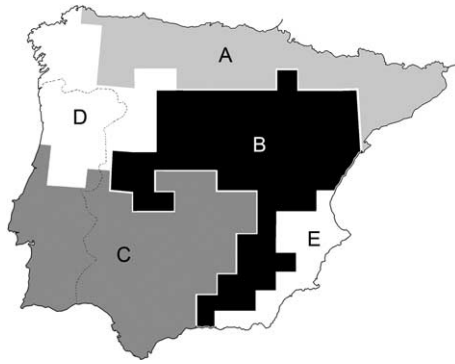


Figure 1. Main Iberian faunal regions based on butterfly distributions (adapted from Romo 2008).

correlates range size with population density (depending on the geographic scale; Hughes 2000), development time, number of broods per year, adult mobility, and resource use specificity. Furthermore, host plant type is correlated with phenology (Cizek et al. 2006) and adult mobility with larval polyphagy, resource availability and range position (Komonen et al. 2004). However, Dennis et al. (2005) points out that range size and larval host range are associated because of a reciprocal dependence on other

life history and resource variables. Therefore, the available evidence suggests a complex pattern of interrelated factors, and those compatible with a dispersal- or niche- related explanation of range size are most relevant.

The role of history in the evolution of the ecological links within the land biotas remains untested, and evidence suggests that the geographic structure of the West-Palaeartic butterfly fauna has been strongly modelled by postglacial events (Dennis et al. 1991, 1998, Schmitt 2007). The present study served to determine how the integration of data on species ranges and ecology might modify former explanations of range size in these insects (Dennis et al. 2004, 2005). In addition, to what degree the correlations among variables were supported from data derived from faunal regions of different climate and physiography. To answer these questions, we tested the relationships between life histories and geographic ranges of Iberian butterflies. Obscure historical patterns, which might be identifiable on a topological basis, were addressed by analyzing relationships between range size and variables that describe species range (with and without controlling for phylogenetic relatedness). We subsequently compared the explanatory power of each of the two subsets of variables (life history and geography), and tested for a relationship between subsets in terms of their shared variance.

Table 1. Bivariate correlations (R) between occupancy (AREA) and the independent variables, based on the raw species data (columns a, b) and on the standardised independent contrasts (columns c, d), and from the full data set (a, c) or the subset of species with known egg sizes (b, d). Sample sizes are a=205, b=155, c=146, d=119. ns =  $p > 0.05$ , \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ .

Variable	Raw species values		Independent contrasts	
	(a)	(b)	(c)	(d)
AWL	0.221***	0.188*	0.004 ns	-0.053 ns
ASXD	0.148*	0.125 ns	0.104 ns	0.076 ns
EGGS	-	-0.179*	-	-0.246**
MIGR	0.345***	0.354***	0.178*	0.255**
VOLT	0.357****	0.420****	0.445***	0.308***
NMON	0.782****	0.781****	0.647****	0.743****
MMON	-0.239***	-0.266***	-0.178*	-0.287**
SDMN	0.724****	0.697****	0.565****	0.663****
OWSA	0.173*	0.144 ns	-0.010 ns	0.063 ns
OWSE	-0.050 ns	-0.116 ns	-0.089 ns	-0.007 ns
OWSL	-0.169*	-0.140 ns	0.024 ns	0.118 ns
OWSP	0.246***	0.289***	0.159 ns	0.212*
LPDB	0.624****	0.651****	0.681****	0.596****
LHPL	0.139*	0.182*	0.045 ns	0.152 ns
LHPF	-0.158*	0.091 ns	0.256 ns	0.084 ns
LHTH	-0.032 ns	0.033 ns	0.014 ns	0.040 ns
LHTW	0.158*	0.091 ns	0.146 ns	0.127 ns
LHPM	0.029 ns	-0.011 ns	0.002 ns	0.007 ns
EGGB	0.070 ns	0.065 ns	0.096 ns	-0.081 ns
MIRW	-0.138*	-0.215**	-0.093 ns	-0.063 ns
MIRS	0.026 ns	0.072 ns	0.075 ns	0.054 ns
MIRO	-0.073 ns	-0.117 ns	0.089 ns	0.029 ns
MIRM	-0.133 ns	-0.203*	0.015 ns	-0.109 ns
MIRX	-0.107 ns	-0.161*	0.096 ns	-0.040 ns
LATM	-0.241***	-0.316***	-0.309***	-0.412****
DISP	0.761****	0.499****	0.634****	0.724****
ALTM	-0.466****	-0.508****	-0.406***	-0.461****
ENDM	-0.495****	-0.444****	-0.419***	-0.328****
REGA	0.315***	0.281***	0.162 ns	0.166 ns
REGB	0.671****	0.679****	0.535****	0.544****
REGC	0.479****	0.469****	0.407***	0.501****
REGD	0.592****	0.604****	0.538****	0.542****
REGE	0.641****	0.648****	0.555****	0.541****

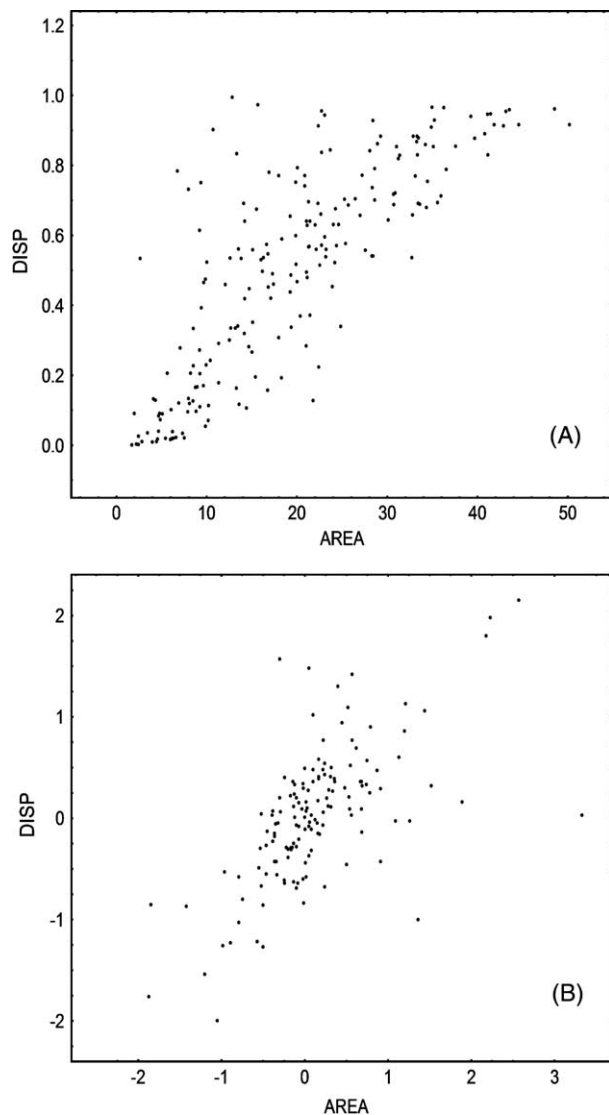


Figure 2. Relationship between dispersion (DISP) and the species size range (AREA). (A) after the species values:  $DISP = -0.033 - 0.0004 \times AREA + 0.038 \times AREA^{1/2}$  ( $R = 0.779$ ,  $p < 0.00001$ ). (B) measured on the independent contrasts:  $DISP = 0.575 \times AREA$ ,  $R = 0.634$ ,  $p < 0.00001$ .

## Methods

### Species and range sizes

The data matrix was comprised of 205 butterfly species (superfamilies Hesperioidea and Papilionoidea) from the Iberian Peninsula (Iberia, the continental territories of Portugal and Spain). Non-native, long-distance migrant species (such as monarch butterflies, *Danaus* spp.), and species with missing life history data were excluded. Each species range (AREA, the dependent variable) was measured as the number of 10 km square cells (UTM military grid) occupied (data from García-Barros et al. 2004, updated for this study by the authors). Area coverage at this grid size does not adequately represent the region and sampling was concentrated on the most diverse areas.

Consequently, species occupancies may be underestimated for widespread and overestimated for rare species (Romo and García-Barros 2005, Romo et al. 2006).

### Life history variables

Variable selection and measurements were chosen based on the available study material and literature. Care was taken to obtain only information specific to the study area (sources in García-Barros et al. 2004, with information from the French Pyrenees from Lafranchis 2000 incorporated when required), with the exception of relative egg size (as detailed below). Life history variables were divided into three subsets: adult features (size, migratory habits and fecundity); phenology; and larval features (larval host specificity, gregariousness and ant relationships). Multistate categorical factors were recoded as dummy binary variables to facilitate the identification of the precise nature of any significant effects. The variables and abbreviations used throughout the remaining text are as follows.

#### Adult features

AWL: adult size (forewing length in mm, log-transformed), estimated as the median of male and female mean wing lengths from sample cabinet specimens ( $n > 20\,000$ : García-Barros unpubl.). When  $< 10$  individuals of each sex were available, measurements from the literature were obtained and included in the data set (Manley and Allcard 1970, Fernández-Rubio 1991, Maravalhas 2003). ASXD: sexual dimorphism in adult size range, residuals from a linear regression of female size onto male size, both log-transformed ( $\text{female AWL} = -0.022 + 1.031 \times \text{male AWL}$ ;  $R = 0.988$ ;  $p < 0.00001$ ). MIGR: migratory habits, where the species were classified as “sedentary” or “migratory” (including eumigrants, intra-regional migrants and long-distance dispersers; Templado 1976, Eitschberger et al. 1991). EGGS: relative egg size, estimated as the residuals from a regression of log-transformed egg volume (from García-Barros 2000b) on AWL ( $R = 0.628$ ,  $p < 0.00001$  using raw data; and  $R = 0.428$ ,  $p < 0.00001$  using independent contrasts, see below). Estimated egg size does not always include Iberian samples (consequently, a combination with local adult sizes resulted in some level of measurement error), and was available from only 154 species. Therefore, it was necessary to duplicate some analyses, as detailed below.

#### Phenology

VOLT: voltinism pattern (one life-cycle per year, or more than one in at least part of the territory). NMON: number of months in which the adults were recorded (95% confidence limits). MMON: mean month of adult occurrence (with months scored as 1–12). SDMN: standard deviation of MMON. The last three variables (NMON, MMON and SDMN) were calculated from pooled historical records with reliable dating (data from the authors database, details in García-Barros et al. 2004). Four binary variables were used to code for the overwintering stage: OWSE (egg), OWSL (larva), OWSP (pupa) and OWSA (adult).

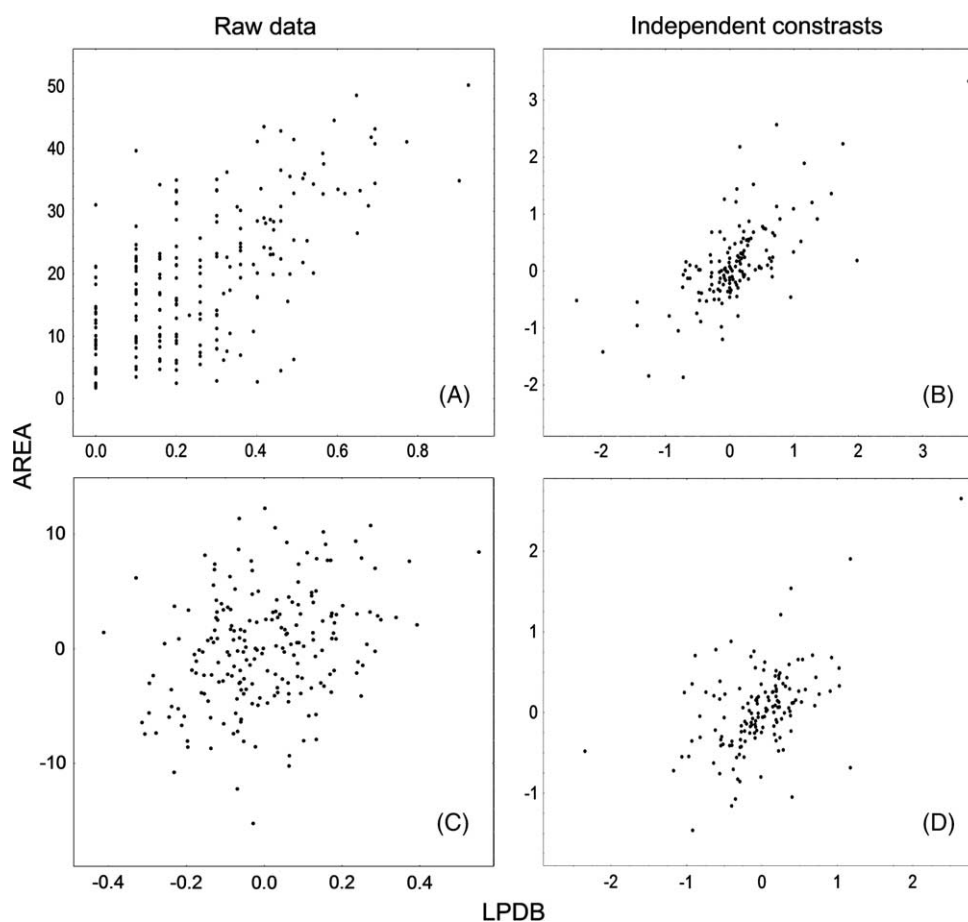


Figure 3. Relationships between range size (AREA) and larval diet breadth (LPDB) based on species raw data (A, C) and independent contrasts (B, D). A and B: bivariate plots; C and D: relationships after controlling for effects of other variables with significant effects, based on the residuals of the linear regressions of AREA and LPDB on the remaining variables in the models of Table 2 (column A) and 3 (column A).

#### Larval, and larval-host features

LPDB (larval diet breadth) was measured as the number of plant families (F), genera (G) and species (S) used as larval hosts using the formula  $LPDB = (S \times G \times F^a)^{1/3}$ , where  $a = (G + F)/2S$ . The index assumes higher values for any number of hosts when the hosts represent a greater number of genera and families (i.e. increased LPDB); LPDB was better correlated to range size ( $r = 0.62$ ) than either the number of host species, genera or families (respectively  $R = 0.589$ ,  $R = 0.567$  and  $R = 0.316$ ;  $p < 0.001$ ). LHPM: coarse taxonomy of the larval hosts (monocotyledons or dicotyledons). Plant structure was broadly described by two binary variables, LHTW (woody plants) and LHTH (herbaceous plants), and plant parts eaten were described as LHPL (leaves), and LHPF (flowers, fruits or buds). Larval gregariousness was measured by coding the egg-laying mode of each species (EGGB: eggs laid singly, or arranged in egg clusters), since the habit of laying the eggs in clusters is almost universally associated to gregarious larval habits.

The larvae of some butterflies (namely, the Lycaenidae) may show symbiotic or parasitic relationships with specific ants, whose presence may condition the success of the butterfly population. The strength of myrmecophily was

described by three binary factors (adapted from Fiedler 1991): weak (including facultative myrmecophily, MIRW), strong (MIRS) and obligate (MIRO). The three variables were also combined as a single semi-quantitative variable (MIRX, with values from 0 = no relationship with ants to 4 = obligate ant-dependence), as well as an all-or-nothing binary variable (MIRM, where “1” means any level of association and “0” no association).

#### Geographic variables (species range attributes)

The X and Y UTM coordinates of each species data subsets were used to estimate LATM (mean Y value, a surrogate of mean latitude) and DISP (dispersion of the cells occupied, calculated as the geometric average of the variances of X and Y). ALTM (mean altitude, in m a.s.l.) was derived from the original database records. ENDM: degree of endemism, quantified as the percentage of  $1^\circ \times 30''$  cells occupied by each species in Iberia over the number of European species (from Kudrna 2002). Five variables: REG (A, B, C, D, E) were adopted to qualify the species according to their presence/absence in each of five Iberian regions (A to E, Fig. 1, formerly defined by Romo 2008).

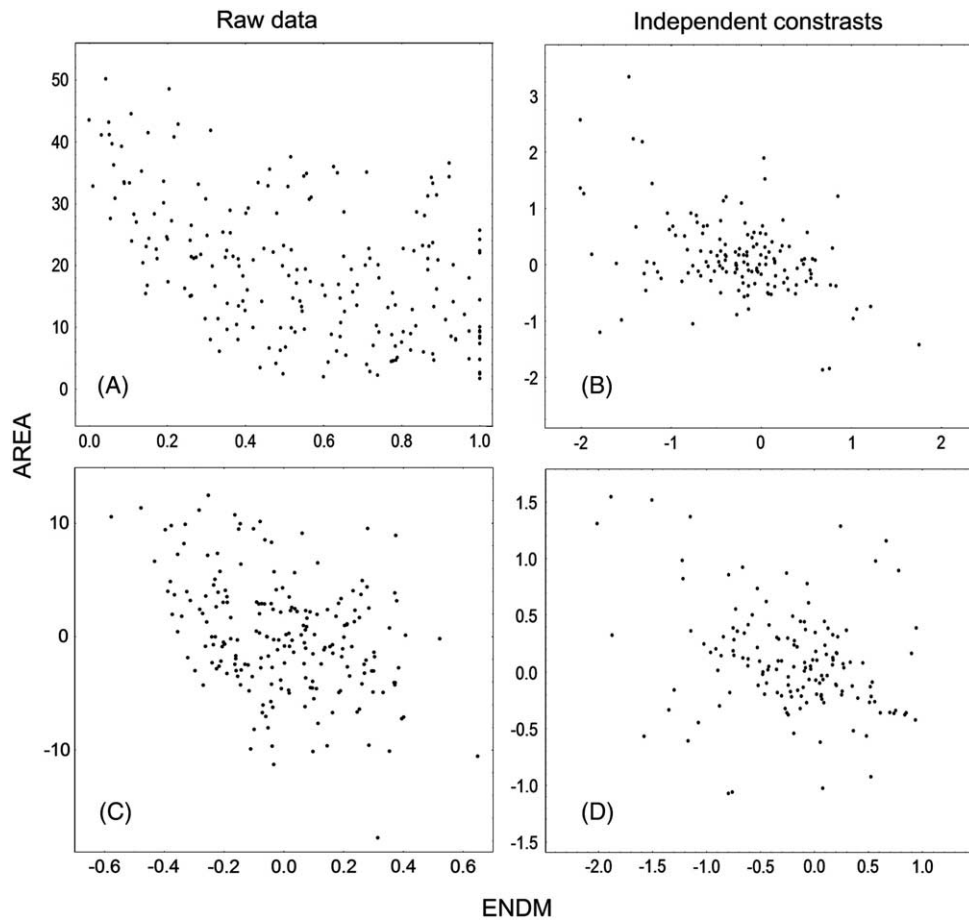


Figure 4. Relationships between range size (AREA) and the concentration of species' ranges in the study relative to Europe (ENDM) based on raw data (A, C) and independent contrasts (B, D). A and B: bivariate plots; C and D: relationships after controlling for the effects of other significant variables (residuals from the linear regressions of AREA and ENDM on the remaining variables, models in Table 2 and 3 for columns A and B, respectively).

Table 2. Explanation of AREA based on species means and three subsets of independent variables: life history plus geography (A), life-history (B) and geography (C). The figures shown are estimated coefficients and signs (Coef.), Wald statistics and significance levels (P) of the variables selected in each model. Whole model results are shown at the lower part of the table. \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ .

Model: Variable	(A) All variables		(B) Life history		(C) Geography	
	Coef.	Wald (P)	Coef.	Wald (P)	Coef.	Wald (P)
MIGR	–	–	–0.186	10.096**	–	–
NMON	–	–	0.120	252.368****	–	–
MMON	–	–	–0.001	11.912****	–	–
SDMN	0.285	89.163****	0.358	164.216****	–	–
OWSP	–	–	–0.159	10.787***	–	–
LPDB	0.039	7.274***	0.063	16.724****	–	–
ENDM	–0.346	26.068****	–	–	–0.560	77.010****
ALTM	–0.016	13.269***	–	–	–0.012	8.725**
LATM	–0.154	6.921***	–	–	–0.229	16.512****
REGA	0.840	31.840****	–	–	1.024	48.338****
REGB	0.262	21.117****	–	–	0.333	35.102****
REGC	0.393	18.935****	–	–	0.463	27.171****
REGE	0.210	16.507****	–	–	0.303	35.644****
DF		195		198		197
Deviance		193.374		320.105		307.555
Deviance/DF		0.992		1.617		1.561
Loglikelihood		–303.885		–367.251		–360.975
R <sup>2</sup>		0.820****		0.811****		0.715****

Table 3. Explanation of AREA based on the independent contrasts and variable selection. Variable subsets: life history plus geography (A), life history (B) and geography (C). The data shown are estimated coefficients and signs (Coef.), Wald statistics and significance levels (P) of the variables selected in each model. Whole model fitting results are shown at the lower part of the table. \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ .

Model: Variable	(A) All variables		(B) Life history		(C) Geography	
	Coef.	Wald (P)	Coef.	Wald (P)	Coef.	Wald (P)
SDMN	–	–	0.289	7.221**	–	–
LPDB	0.514	17.492****	0.292	7.997**	–	–
ENDM	0.381	7.209**	–	–	–	–
REGD	1.122	16.952****	–	–	1.338	28.393
DF		143		144		145
Deviance		235.083		240.065		241.998
Deviance/DF		1.644		1.667		1.669
Loglikelihood		–128.977		–137.855		–136.855
R <sup>2</sup>		0.614****		0.572****		0.268****

### Independent contrasts

Phylogenetically independent contrasts were calculated using the package COMPARE (Martins 2004). Branch lengths were set to the same length (1.00) except at polytomies, where we set a low value (0.00001; as suggested by Martins 2004). Single contrasts at these nodes were computed by initially resolving the relationships arbitrarily, then the average value per polytomy was calculated and contrasts computed. The adopted phylogenetic topology (Supplementary material) followed the sources quoted by García-Barros (2000a) and Cizek et al. (2006); and updated after Als et al. (2004), Wahlberg et al. (2005), Braby et al. (2006), Lukhtanov et al. (2006), Peña et al. (2006), Simonsen et al. (2006), Weingartner et al. (2006) and Kodandaramaiah and Wahlberg (2009).

### Procedure

The raw data values for AREA, SDMN and ALTM were square root transformed, MMON cubed, and AWL, EGGS and LPDB were log<sub>10</sub>-transformed to approach normality. Two series of computations were conducted, using species means and independent contrasts. Contrast-based regressions were forced through the origin (Felsenstein 1985).

The bivariate relationships between AREA and the independent variables were determined. Subsequently, regression models were selected using a manual stepwise procedure (GLM module of Statistica: StatSoft 2004), assuming a Poisson distribution of the residuals and a

logarithmic link function. The variables were entered in descending order according to their bivariate relationship with AREA (Table 1), and a backward selection was fit at each step to discard redundant terms. Wald's statistic was applied to measure the contribution of each variable, and model strength with the standard deviation (whole model R<sup>2</sup> values were added from GRM fitting). First, one model was fit using all variables (life-history and geographic). Two additional models were subsequently calculated with the two sets of variables independently. Finally, we applied variance partitioning (Legendre and Legendre 1998, Diniz-Filho and Bini 2008) to measure the amount of variance explained by life history, geographic features, or shared by both sets of variables (the variables included in this specific analysis were those that exhibited significant effects in any of the three formerly described models, i.e. “all”, “life history”, and “geography”).

### Preliminary analyses

Pairwise correlations are shown in Table 1. Egg size (EGGS) was significantly correlated with AREA. One additional stepwise regression restricted to data from species with known EGGS values did not select egg size (using either the log-transformed or the independent contrasts). The variables selected without controlling for phylogeny included NMON, SDMN, REGD, REGA, REGC, REGE and LPDB (deviance/DF = 0.873; log-likelihood = –232.728; R<sup>2</sup> = 0.837;  $p < 0.0001$ ), and variables from the set of contrasts included MIGR, LPDB, ALTM, LATM, REGA and REGD (deviance/DF = 1.701; log-likelihood = –74.425; R<sup>2</sup> = 0.597;  $p < 0.0001$ ). Consequently, we excluded EGGS from subsequent analyses, which resulted in a somewhat lower number of species and contrasts analyzed.

LATM displaced dispersion (DISP) (with which it was correlated;  $R = -0.427$ ,  $p < 0.0001$ ;  $R = 0.393$ ,  $p < 0.0001$ , based on the contrasts) in the analysis, probably due to heteroscedascity in the X and Y values. Therefore, to gain insights into the correlation in dispersion patterns, we tested the relationship between DISP to AREA with a polynomial function (species means) and a linear function (contrasts) (Fig. 2). We subsequently analysed the residuals following the methods described above.

Table 4. Variance partitions among the life history and geographic variables, based on the species values and on the standardised independent contrasts. The subset of variables for each column consisted of those with significant effects in any of the models described in Table 2 (raw data) and Table 3 (contrasts). The figures represent the proportion of variance explained (over 1.000).

	Species values	Contrasts
Pure life-history	0.161	0.188
Pure geographic	0.456	0.089
Shared (life history + geographic)	0.655	0.390
Residual (not explained)	0.139	0.333
Total	1.000	1.000
Model (R)	0.901	0.666

Table 5. Bivariate correlations (R) between the geographical dispersion of the species data points not explained by occupancy (residuals of the regression of DISP on AREA) and the potentially explanatory variables, derived from raw data (n=205) and independent contrasts (n=149). ns = p > 0.05, \* = p < 0.05, \*\* = p < 0.01, \*\*\* = p < 0.001, \*\*\*\* = p < 0.0001.

Variable	Species values	Contrasts
AWL	-0.097 ns	0.131 ns
ASXD	0.009 ns	0.083 ns
EGGS	-0.040 ns	0.086 ns
MIGR	-0.036 ns	0.061 ns
VOLT	0.061 ns	0.083 ns
NMON	0.161*	0.178*
MMON	-0.246**	0.062 ns
SDMN	-0.023 ns	0.159 ns
OWSA	-0.007 ns	0.028 ns
OWSE	0.016 ns	-0.064 ns
OWSL	-0.098 ns	0.031 ns
OWSP	0.183*	0.041 ns
LPDB	0.061 ns	0.125 ns
LHPL	-0.055 ns	-0.031 ns
LHPF	0.257***	-0.077 ns
LHTH	-0.284****	-0.239**
LHTW	0.270***	0.129 ns
LHPM	-0.236**	-0.201*
EGGB	0.015 ns	0.011 ns
MIRW	0.191*	-0.119 ns
MIRS	0.034 ns	0.196*
MIRO	0.046 ns	0.049 ns
MIRM	0.203*	0.181*
MIRX	0.157 ns	0.232**
LATM	-0.438	-0.280***
ALTM	-0.257***	-0.282***
ENDM	0.226**	0.229**
REGA	-0.020 ns	0.125 ns
REGB	0.115 ns	0.363****
REGC	0.177*	0.476****
REGD	0.463****	0.551****
REGE	0.329****	0.411****

## Results

Pairwise analyses found 24 variables significantly correlated with AREA (Table 1) across species means, and 14 across contrasts (Table 1, Fig. 3 and 4). The selected variables are detailed in Table 2 and 3. A maximum of six “life history” variables were selected when this subset of variables was analysed. However, in the combined analyses (life history plus geography), the geographical variables tended to displace the life history traits from the models, with the exception of LPDB (Fig. 3) and, depending on the analysis, SDMN.

Seven of the geographical variables exhibited significant effects when species means were analysed (ENDM, ALTM, LATM, REGA, REGB, REGC and REGE; Table 2). However, significant effects were not evident when contrasts were analysed, with the exception of ENDM and/or REGD, which remained significant (Table 3, Fig. 4).

Variance partitioning (Table 4) indicated that irrespective of the source of data (species mean values or contrasts), a large part of the explanation of the variance is shared by life history plus geography.

Finally, dispersion (with fixed effects for AREA) was positively correlated with geography (REGD and ENDM) and negatively with larval host type (LHTH i.e. “not

Table 6. Stepwise selection results summarising the relationships between range dispersion (with fixed AREA effects, residuals from the function in Fig. 2) and the remaining variables, based on species raw values and on independent contrasts. The values shown are estimated coefficients and signs (Coef.), Wald statistics, and significance (P) of selected variables. Whole model fitting statistics are indicated at the lower part of the table. \* = p < 0.05, \*\* = p < 0.01, \*\*\* = p < 0.001, \*\*\*\* = p < 0.0001.

Variable	Species data		Contrasts	
	Coef.	Wald (P)	Coef.	Wald (P)
REGD	0.177	52.588****	-	-
ENDM	0.164	17.754****	0.369	6.187*
LHTH	-0.099	10.433**	-	-
OWSP	0.0633	4.503*	-	-
DF		200		145
Deviance		5.2751		180.325
Deviance/DF		0.026		1.244
Loglikelihood		84.2684		-222.579
R <sup>2</sup>		0.299****		0.052**

feeding on herbaceous plants”), and with “overwintering as a pupa” (OWSP) (Table 5, 6).

## Discussion

The following summarizes the results of our study: 1) significant pairwise correlations between range size and several life history variables were observed; 2) these correlations generally tend to loose strength after controlling for phylogeny; 3) the variables retained by stepwise regression selection were associated with larval host diet breadth, adult phenology, and geographic features of species distribution; and 4) an integral component of the explanation provided by the models was of a mixed (life history and geographic) nature.

Both niche theory and dispersal-related expectations are supported by the bivariate correlations across the species means results generated in this study. Comparatively, widespread species are larger, more sexually dimorphic, and exhibit increased fecundity and dispersal ability. Species with more scattered/patchy distributions tend to show longer periods of adult occurrence, their larvae typically feed on dicotyledonous woody plants, and demonstrate some degree of ant-dependence as well as a restricted Iberian range with low mean latitude and altitude.

The generally (though not universally, Table 1) lower correlations from independent contrasts do not simply reflect the reduced degrees of freedom (as argued by Dennis et al. 2004), but also the different weight of taxonomic patterning across variables (Brooks and McLennan 2002, Diniz-Filho and Torres 2002). Our comparative results were partly hindered by poor phylogenetic resolution and unknown branch lengths, two reasons why (despite the robust contrasts method under such conditions; Martins and Garland 1991, Garland et al. 1992) we concentrated on identifying the sign and strength of the relationships rather than their shape (Quader et al. 2004). Therefore, within these limitations we confidently concluded that the bivariate correlations between range size and relative egg size, dispersive adults, long period of adult occurrence, larval diet

breadth and some geographic variables were not taxonomically driven artefacts.

Among the life history variables, multiple selection on independent contrasts supports a causal interpretation of range size in terms of niche breadth (potential or realised larval polyphagy may assist the offspring to thrive in more diverse habitats) and, at least partially, dispersion in the adult “time window”. However, the second variable may actually represent a geographic turnover of species phenologies, as the Iberian lands cover a remarkable array of climate and vegetation types. Phenological variation along a geographic gradient should be higher among populations than within them, therefore the observed correlation may not denote an ability of widespread species to occupy varied habitats, but the degree to which widespread species phenologies are tuned to local conditions. Depending on the methods and variables used to study British butterflies, this explanation is supported by some reports but not others (Hodgson 1993, Dennis et al. 2004, 2005). Correlations between range size and body size or dispersal ability (formerly documented from various taxa, Gaston and Blackburn 1994, Brown 1995, Purvis et al. 2001 and Diniz-Filho et al. 2005, but cf. Hillebrand et al. 2001, Wilkinson 2001, Fernandez and Vrba 2005 or Rundle et al. 2007) are not supported by the most parsimonious interpretation of our results. The relationship between Iberian butterfly range and adult size is taxonomic, range size is weakly related to migratory status, and both relationships break down under multivariate selection protocols. However, we cannot strictly discard alternative explanations of dispersal type, namely because butterfly wing length (which we tested) might not be the best surrogate for dispersal ability (as shown for damselflies by Rundle et al. 2007). Furthermore, functional and genetic links among species life history attributes (Stearns 1977, Roff 2002) generally result in complex patterns of cross correlations (see butterfly examples in García-Barros 2000a or Cizek et al. 2006). If distribution patterns are to be evaluated in terms of the realized niche (Austin and Smith 1989, Kockemann et al. 2009) and niche is assessed based on environmental variables (Thuiller et al. 2003), “complex” variables (describing organism-habitat interactions) should perform better in multivariate tests than “proximate” variables (describing features of the organism, or of the habitat, Dennis et al. 2004, 2005). Therefore, the distinction between “causal” and “most parsimonious” solutions is of interest. However, a more exhaustive analysis of the life histories of Iberian butterflies at local or regional levels and comparable studies across different regions is warranted to provide resolution at this scale.

Results of the study suggested that the explanation for range size could be improved by including additional variables to represent processes not explicitly tested. For example, some geographic variables exhibited a high weight and proportion of variance shared with life history data. Range size might be the result of interactions between species life history patterns and geographic history. Comparable conclusions have been drawn from recent work on plant biogeography (Svenning and Skov 2005, Weiser et al. 2007, Gove et al. 2009). Resolving these processes for further analyses is not a simple task, as it requires a comparative approach to the life histories of each of the

taxa, done within a phylogeographic framework. The geographic distribution of species is a product of speciation, extinction and the temporal dynamics of its range (Gaston 1998, 2003). Consequently, the range of a species should be examined and explained in terms of the species features, its interactions with the environment (Kean and Barlow 2004) and the historical factors, which likely set limits on other interactions. However, little attention has been drawn to the relationships between different macroecological patterns (Blackburn and Gaston 2001).

Two incidental findings of our work relate to endemism and Rapoport’s pattern. The significance of endemism has a practical application in Iberia: Iberian endemics tend to be rare in Iberia (for comparable results from other regions and taxa see Gregory and Blackburn 1998 and Carrascal et al. 2008). Based on our knowledge of butterfly species and their distributions in the study area, most of the butterflies with small ranges in this region are restricted to relatively high elevations on the main mountain chains, which cover a small portion of the peninsula. These species may be particularly sensitive e.g. in terms of global warming and other environmental impacts. For the same reason, the concentration of mountain ranges in the northern half of the peninsula explains the negative pairwise correlations between range size and both latitude and altitude, incongruent with Rapoport’s pattern (where ranges should be wider at higher latitudes or altitudes; Rapoport 1975, Lomolino et al. 2006).

In summary, our results identified niche breadth (via larval polyphagy) as primarily correlated with range size, together with interactions between non-explicit historical causes (represented by chorotypes) and life histories. A more thorough “dissection” of the biological correlations of range size, as well as an integrated multi-scale protocol is required before more specific explanations for range size may be achieved.

*Acknowledgements* – We thank I. Echavarren for assistance during the early stages of the study, and C. Stefanescu and M. L. Munguira for contributing to the revision of life history data. This study was partly funded by project GL2006-10196 (M. E. C.).

## References

- Als, T. D. et al. 2004. The evolution of alternative parasitic life histories in large blue butterflies. – *Nature* 432: 386–390.
- Austin, M. P. and Smith, T. M. 1989. A new model for the continuum concept. – *Vegetatio* 83: 35–47.
- Blackburn, T. M. and Gaston, K. J. 2001. Linking patterns in macroecology. – *J. Anim. Ecol.* 70: 338–352.
- Braby, M. F. et al. 2006. Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. – *Zool. J. Linn. Soc.* 147: 239–275.
- Brooks, D. R. and McLennan, D. A. 2002. The nature of diversity. An evolutionary voyage of discovery. – Chicago Univ. Press.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. – *Am. Nat.* 124: 255–279.
- Brown, J. H. 1995. *Macroecology*. – Chicago Univ. Press.
- Calosi, P. et al. 2010. What determines a species’ geographical range? Thermal biology and latitudinal range size relationships in European diving beetles (Coleoptera: Dytiscidae). – *J. Anim. Ecol.* 79: 194–204.



- Carrascal, L. M. et al. 2008. Explanations for bird species range size: ecological correlates and phylogenetic effects in the Canary Islands. – *J. Biogeogr.* 35: 2061–2073.
- Cizek, L. et al. 2006. Host plant defences and voltinism in European butterflies. – *Ecol. Entomol.* 31: 337–344.
- Cowley, M. J. R. et al. 2001a. Density–distribution relationships in British butterflies. I. The effect of mobility and spatial scale. – *J. Anim. Ecol.* 70: 410–425.
- Cowley, M. J. R. et al. 2001b. Density–distribution relationships in British butterflies. II. An assessment of mechanisms. – *J. Anim. Ecol.* 70: 426–441.
- Dennis, R. L. H. et al. 1991. A multivariate approach to the determination of faunal structures among European butterfly species (Lepidoptera: Rhopalocera). – *Zool. J. Linn. Soc.* 101: 1–49.
- Dennis, R. L. H. et al. 1998. Faunal structures among European butterflies: evolutionary implications of bias for geography, endemism and taxonomic affiliations. – *Ecography* 21: 181–203.
- Dennis, R. L. H. et al. 2000. Ecological correlates of island incidence and geographical range among British butterflies. – *Biodivers. Conserv.* 9: 343–359.
- Dennis, R. L. H. et al. 2004. Host plants and butterfly biology. Do host-plant strategies drive butterfly status? – *Ecol. Entomol.* 29: 12–26.
- Dennis, R. L. H. et al. 2005. Does diet breadth control herbivorous insect distribution size? Life history and resource outlets for specialist butterflies. – *J. Insect Conserv.* 9: 187–200.
- Diniz-Filho, J. A. F. and Torres, N. M. 2002. Phylogenetic comparative methods and the geographic range size – body size relationship in new world terrestrial Carnivora. – *Evol. Ecol.* 16: 351–367.
- Diniz-Filho, J. A. F. and Bini, L. M. 2008. Macroecology, global change and the shadow of forgotten ancestors. – *Global Ecol. Biogeogr.* 17: 11–17.
- Diniz-Filho, J. A. F. et al. 2005. Macroecology, geographic range size – body size relationships and minimum viable population analysis for new world Carnivora. – *Acta Oecol.* 27: 25–30.
- Eitschberger, U. et al. 1991. Wanderfalter in Europa (Lepidoptera). Appeal for international cooperation in the research of the migration of insects. – *Atalanta* 22: 1–67.
- Felsenstein, J. 1985. Phylogenies and the comparative method. – *Am. Nat.* 125: 1–15.
- Fernandez, M. H. and Vrba, E. S. 2005. Body size, biomic specialization and range size of African large mammals. – *J. Biogeogr.* 32: 1243–1256.
- Fernández-Rubio, F. 1991. Guía de mariposas diurnas de la Península Ibérica, Baleares, Canarias, Azores y Madeira. – Pirámide, Madrid.
- Fiedler, K. 1991. Systematic, evolutionary and ecological implications of myrmecophily within the Lycaenidae (Insecta: Lepidoptera: Papilionoidea). – *Bonner Zool. Monogr.* 31: 1–210.
- García-Barros, E. 2000a. Body size, egg size, and their interspecific relationships with ecological and life history traits in butterflies (Lepidoptera: Papilionoidea, Hesperioidea). – *Biol. J. Linn. Soc.* 70: 251–284.
- García-Barros, E. 2000b. Egg size in butterflies (Lepidoptera: Papilionoidea and Hesperioidea): a summary of data. – *J. Res. Lepid.* 35: 90–136.
- García-Barros, E. et al. 2004. Atlas of the butterflies of the Iberian Peninsula and Balearic Islands (Lepidoptera: Papilionoidea & Hesperioidea). – *Monogr. S.E.A.* 11: 1–228.
- Garland, T. J. et al. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. – *Syst. Biol.* 41: 18–32.
- Gaston, K. J. 1998. Species-range size distributions: products of speciation, extinction and transformation. – *Phil. Trans. R. Soc. B* 353: 219–230.
- Gaston, K. J. 2003. The structures and dynamics of geographic ranges. – Oxford Univ. Press.
- Gaston, K. J. and Blackburn, T. M. 1994. Are newly described bird species small-bodied? – *Biodivers. Lett.* 2: 16–20.
- Gaston, K. J. and Blackburn, T. M. 1999. A critique for macroecology. – *Oikos* 84: 353–368.
- Gaston, K. J. et al. 1997. Interspecific abundance-range size relationships: an appraisal of mechanisms. – *J. Anim. Ecol.* 66: 579–601.
- Gaston, K. J. et al. 2000. Abundance–occupancy relationships. – *J. Appl. Ecol. (Suppl. 1)* 37: 39–59.
- Gaston, K. J. et al. 2007. The scaling of spatial turnover: pruning the thicket. – In: Storch, D. et al. (eds), *Scaling diversity*. Cambridge Univ. Press, pp. 181–222.
- Gilbert, B. and Lechowicz, M. J. 2004. Neutrality, niches and dispersal in a temperate forest understorey. – *Proc. Nat. Acad. Sci. USA* 101: 7651–7656.
- Gove, A. D. et al. 2009. Dispersal traits linked to range size through range location, not dispersal ability, in Western Australian angiosperms. – *Global Ecol. Biogeogr.* 18: 596–606.
- Gregory, R. D. and Blackburn, T. M. 1998. Macroecological patterns in British breeding birds: covariation of species' geographical ranges at differing spatial scales. – *Ecography* 21: 527–534.
- Gregory, R. D. and Gaston, K. J. 2000. Explanations of commonness and rarity in British breeding birds: separating resource use and resource availability. – *Oikos* 88: 515–526.
- He, F. L. and Legendre, P. 2002. Species diversity patterns derived from species–area models. – *Ecology* 85: 1185–1198.
- Hillebrand, H. et al. 2001. Differences in species richness patterns between unicellular and multicellular organisms. – *Oecologia* 126: 114–124.
- Hodgson, J. G. 1993. Commonness and rarity in British butterflies. – *J. Appl. Ecol.* 30: 407–427.
- Hughes, J. B. 2000. The scale of resource specialization and the distribution and abundance of lycaenid butterflies. – *Oecologia* 123: 375–383.
- IUCN 2001. IUCN Red List categories and criteria: version 3.1. – IUCN Species Survival Commission, Gland and Cambridge.
- Kean, J. and Barlow, N. 2004. Exploring rarity using a general model for distribution and abundance. – *Am. Nat.* 163: 407–E416.
- Kockemann, B. et al. 2009. The relationships between abundance, range size and niche breadth in central European tree species. – *J. Biogeogr.* 36: 854–864.
- Kodandaramaiah, U. and Wahlberg, N. 2009. Phylogeny and biogeography of *Coenonympha* butterflies (Nymphalidae: Satyrinae) patterns of colonization in the Holarctic. – *Syst. Entomol.* 34: 315–323.
- Komonen, A. et al. 2004. The role of niche breadth, resource availability and range position on the life history of butterflies. – *Oikos* 105: 41–54.
- Kudrna, O. 2002. The distribution atlas of European butterflies. – *Oedippus* 20: 1–342.
- Lafranchis, T. 2000. Les papillons de jour de France, Belgique et Luxembourg et leurs chenilles. – Biotop, Mèze.
- Legendre, P. and Legendre, L. 1998. Numerical ecology, 2nd ed. – Elsevier.
- Lomolino, M. V. et al. 2006. Biogeography, 3rd ed. – Sinauer.
- Lukhtanov, V. A. et al. 2006. Rearrangement of the *Agrodiaetus dolus* species group (Lepidoptera, Lycaenidae) using a new cytological approach and molecular data. – *Insect Syst. Evol.* 37: 325–334.

- Manley, W. B. L. and Allcard, H. G. 1970. A field guide to the butterflies and burnets of Spain. – E. W. Classey, Hampton.
- Maravalhas, E. 2003. As borboletas de Portugal. The butterflies of Portugal. – Apollo Books.
- Martins, E. P. 2004. COMPARE, version 4.6b. Computer programs for the statistical analysis of comparative data. – Dept Biology, Indiana Univ., <<http://compare.bio.indiana.edu/>>.
- Martins, E. P. and Garland, T. Jr 1991. Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. – *Evolution* 45: 534–557.
- Nylin, S. 2009. Gradients in butterfly biology. – In: Settele, J. et al. (eds), *Ecology of butterflies in Europe*. Cambridge Univ. Press, pp. 198–216.
- Peña, C. et al. 2006. Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data. – *Mol. Phylogenet. Evol.* 40: 29–49.
- Purvis, A. et al. 2001. Past and future carnivore extinctions: a phylogenetic perspective. – In: Gittleman, J. L. et al. (eds), *Carnivore conservation*. Cambridge Univ. Press, pp. 11–34.
- Quader, S. et al. 2004. Nonlinear relationships and phylogenetically independent contrasts. – *J. Evol. Biol.* 17: 709–715.
- Quinn, R. M. et al. 1998. The distribution of butterflies and their foodplants. – *Ecography* 21: 279–288.
- Rapoport, E. 1975. Areografía: estrategias geográficas de las especies. – Fondo de Cultura Económica.
- Roff, D. A. 2002. Life history evolution. – Sinauer.
- Romo, H. 2008. Diversidad geográfica de las mariposas diurnas Ibero-baleares. – *Ecosistemas* 17: 106–111.
- Romo, H. and García-Barros, E. 2005. Distribución e intensidad de los estudios faunísticos sobre mariposas diurnas en la Península Ibérica e Islas Baleares (Lepidoptera, Papilionoidea y Hesperioidea). – *Graellsia* 61: 37–50.
- Romo, H. et al. 2006. Identifying recorder-induced geographic bias in an Iberian butterfly database. – *Ecography* 29: 873–885.
- Rosenzweig, M. L. 1975. On continental steady states of species diversity. – In: Cody, M. L. and Diamond, J. M. (eds), *Ecology and evolution of communities*. Harvard Univ. Press, pp. 121–140.
- Rosenzweig, M. L. 2003. How to reject the area hypothesis of latitudinal gradients. – In: Blackburn, T. M. and Gaston, K. J. (eds), *Macroecology. Concepts and consequences*. Blackwell, pp. 87–106.
- Rundle, S. D. et al. 2007. Range size in North American *Enallagma* damselflies correlates with wing size. – *Freshwater Biol.* 52: 471–477.
- Schmitt, T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. – *Front. Zool.* 4: 11.
- Simonsen, T. J. et al. 2006. Morphology, molecules and fritillaries: approaching a stable phylogeny for Argynnini (Lepidoptera: Nymphalidae). – *Insect Syst. Evol.* 37: 405–418.
- StatSoft 2004. STATISTICA (data analysis software system), version 6.1. – Statsoft, Tulsa, <[www.statsoft.com](http://www.statsoft.com)>.
- Stearns, S. C. 1977. The evolution of life history traits: a critique of the theory and a review of the data. – *Annu. Rev. Ecol. Syst.* 8: 145–171.
- Storch, D. et al. 2007. Scaling species richness and distribution. Uniting the species–area and the species–energy relationships. – In: Storch, D. et al. (eds), *Scaling biodiversity*. Cambridge Univ. Press, pp. 300–321.
- Svenning, J.-C. and Skov, F. 2005. The relative roles of environment and history as controls of tree species composition and richness in Europe. – *J. Biogeogr.* 32: 1019–1033.
- Templado, J. 1976. Sobre la variación geográfica de los Ropalóceros Ibéricos (Lepidoptera). – *Graellsia* 31: 79–92.
- Thuiller, W. et al. 2003. Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). – *Global Ecol. Biogeogr.* 12: 313–325.
- Wahlberg, N. et al. 2005. Phylogenetic relationships and historical biogeography of tribes and genera in the subfamily Nymphalinae (Lepidoptera: Nymphalidae). – *Biol. J. Linn. Soc.* 86: 227–251.
- Weingartner, E. et al. 2006. Speciation in *Pararge* (Satyrinae: Nymphalidae) butterflies – North Africa is the source of ancestral populations of all *Pararge* species. – *Syst. Entomol.* 31: 1–13.
- Weiser, M. D. et al. 2007. Latitudinal patterns of range size and species richness of New World woody plants. – *Global Ecol. Biogeogr.* 16: 679–688.
- Wilkinson, D. M. 2001. What is the upper size limit for free-living microorganisms. – *J. Biogeogr.* 28: 285–291.

Download the Supplementary material as file E6269 from <[www.oikos.ekol.lu.se/appendix](http://www.oikos.ekol.lu.se/appendix)>.