

# Using phylogenies to trace the geographical signal of diversification

Pedro Abellán<sup>1</sup>  | Ignacio Ribera<sup>2</sup> 

<sup>1</sup>Department of Biology, Queens College, City University of New York, New York, NY, USA

<sup>2</sup>Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Barcelona, Spain

## Correspondence

Pedro Abellán, Department of Biology, Queens College, City University of New York, New York, NY, USA.  
Email: pabellan@qc.cuny.edu

Editor: François Guilhaumon

## Abstract

**Aim:** Current geographical distributions in conjunction with species-level phylogenies have been viewed as offering an enormous potential for investigating the causes of speciation. However, many authors have concluded that species ranges are too dynamic to retain the historical signal of the speciation processes. We develop an approach to explore the extent to which range changes have erased the geographical signal of speciation in the current distribution of species lineages, and whether phylogenetic relationships among species and their current distributions can provide information regarding the geography of diversification. As case studies we used different lineages of aquatic Coleoptera.

**Location:** Western Palaearctic.

**Methods:** We first assessed the persistence of a geographical signal in the current distribution of a clade from the association between evolutionary relationships and geography using Mantel tests. We then tested two potential scenarios of diversification using bivariate plots of the geographical distance between the centroids of the species ranges and phylogenetic distances, and an assessment of the match between the observed geographical distribution and the phylogenetic topology. We used as a study system a set of 10 monophyletic lineages of water beetles with different ecological characteristics (either standing water, running water or a mixture of both habitat types).

**Results:** Our results point to a common pattern of geographically conserved ranges where current species seem to have originated through range fragmentation of formerly more widespread species. Exclusive standing water clades showed a much weaker signal, as expected from their higher mobility, which erases the geographical signal at much shorter temporal scales.

**Main conclusions:** Overall, our findings show that for at least some lineages it is possible to obtain strong evidence of stasis of the geographical ranges of species, as well as information regarding the diversification process, through the study of their current distributions and phylogenetic relationships.

## KEYWORDS

allopatric speciation, aquatic Coleoptera, biogeography, geographical distributions, phylogenies, species ranges

## 1 | INTRODUCTION

The analyses of species' ranges provided the original evidence of the key role of geography in the study of speciation (see e.g. the geography of speciation in Mayr, 1963). Since then, current distributions in conjunction with species-level phylogenies have been considered to offer an enormous potential for investigating the causes and modes of speciation (e.g. Barraclough & Nee, 2001; Barraclough & Vogler, 2000; Brooks & McLennan, 1991; Johnson & Cicero, 2002). Most previous research in this area has focused on inferring the geographical mode of speciation (allopatric, sympatric or parapatric) and assessing how frequent are these different speciation modes (Losos & Glor, 2003). Early approaches examined the relative range size and range overlap of extant sister species under the assumption that spatial isolation of sister taxa is a signature of allopatric speciation (e.g. Anderson & Evensen, 1978; Chesser & Zink, 1994; Lynch, 1989). More recently, the relation between the range overlap of sister lineages and the age of their separation has been studied using comparative methods. In this approach, the observed patterns between sympatry and node age are compared with the expected patterns of range overlap under alternative modes of speciation (Barraclough & Vogler, 2000; Fitzpatrick & Turelli, 2006; Perret, Chautems, Spichiger, Barraclough & Savolainen, 2007). The assumption is that the amount of overlap between the geographical ranges of sister species or clades depends initially on the geography of speciation, but becomes randomized over time due to subsequent range shifts (Barraclough & Vogler, 2000; Berlocher & Feder, 2002; Lynch, 1989; Phillimore et al., 2008; Pigot, Owens & Orme, 2012).

It is thought that species ranges are not constant over time but can vary through the species' lifetime in both size and geographical position (Gaston, 2003), often over time scales significantly faster than the cladogenetic processes (Davis & Shaw, 2001; Lessa, Cook & Patton, 2003). This has led many authors to conclude that rapid changes in species geographical ranges may mask any potential relationship between their current distribution and the geography of speciation. As a consequence, evolutionary inferences concerning the geography of species in the past is often considered to be unreliable (Bolnick & Fitzpatrick, 2007; Chesser & Zink, 1994; Fitzpatrick & Turelli, 2006; Gaston, 1998; Losos & Glor, 2003). However, it would be equally wrong to assume that all species have suffered modifications in their ranges large enough to erase any geographical signal from the past, as related groups of species often tend to be found in the same parts of the world, at all geographical and phylogenetic scales. In some cases there is strong evidence to support the stasis of geographical ranges, either through the fossil record (e.g. Jablonski, 1987) or indirectly from ecological and phylogenetic data (e.g. Carranza & Wade, 2004; Martínez-Solano, Gonçalves, Arntzen & García-París, 2004). An important question to be investigated is then whether such high rates of range evolution are typical of different groups, and to which extent range changes have obscured the information regarding the geography of speciation.

It is increasingly recognized that the classification of speciation as allopatric, parapatric or sympatric is an artificial subdivision of a continuum and misses many potentially important nuances (Fitzpatrick, Fordyce & Gavrilets, 2009). With independence of the degree of range overlap during the speciation process, and the particular mode of speciation, a relevant issue that has not been as thoroughly explored is the persistence of a biogeographical signal of the cladogenetic process, and whether it is possible to infer the diversification mechanism from this biogeographical signal. To test the persistence of a geographical signal in the current distribution of the species of a clade, Ribera et al. (2011) compared the spatial relationship between the centroids of the range of the species and their evolutionary relationships (see also Cardillo, 2015). While distribution boundaries can change through range expansion or contraction, affecting the degree of sympatry among species and obscuring the diversification process, the relative position of the distribution of each species in relation to the others (as estimated from the range centroids) may still remain constant, leaving a geographical signal of where and how the species originated. As shown by Ribera et al. (2011), congruence between phylogenetic and spatial relationships can provide evidence on the persistence of a geographical signal in the current distribution of species within a lineage. The comparison of observed phylogenetic and spatial relationships with a random null model (Barraclough & Nee, 2001) allowed testing different scenarios of speciation.

Given the diverse factors that can potentially affect post-speciation range shifts, and the different modes of speciation across the allopatric to sympatric continuum, a useful strategy to test the persistence of biogeographical signal of diversification are comparative studies in closely related groups of taxa differing in life history or ecological traits potentially affecting their dispersal capability, and in consequence their mode of speciation. Here, we develop the approach introduced in Ribera et al. (2011) to explore the extent to which range changes have erased the phylogenetic signal in the current distribution of species, and whether phylogenetic relationships among species and their current distributions can provide information regarding the biogeography of diversification. As case studies we used several lineages of Western Palearctic aquatic Coleoptera. Water beetles are a rich and well-known insect group in both Europe and the Mediterranean Basin, exhibiting a high level of endemism but also with species widely distributed across the Palearctic and Holarctic regions, with subsets of taxa with overlapping ranges but others geographically isolated (Ribera, Foster & Vogler, 2003). Most species are specialized to either standing (lentic) or running (lotic) waters, with differences in habitat preference occurring at all phylogenetic levels (Ribera, 2008). Due to the short-term geological persistence of small lentic water bodies, species living in them show higher dispersal capacity as well as more dynamic and larger geographical ranges (Ribera, 2008). It could then be hypothesized that in lineages of lentic species range changes are more likely to have obscured the information regarding the geography of speciation than in lineages of lotic species.

## 2 | MATERIALS AND METHODS

### 2.1 | Geography of diversification: two potential scenarios

As a demonstration of the use of species ranges to infer the geography of diversification we test here two potential scenarios of speciation by geographical isolation, differing in the resulting relationship between geographical and phylogenetic distances (Ribera et al., 2011). Under such scenarios, the persistence of a geographical signal in the current distribution of a clade can be evaluated through the comparison of observed phylogenetic and spatial relationships against a random null model (Barraclough & Nee, 2001).

#### 2.1.1 | Scenario 1

Speciation by stepping-stone colonization (Figure 1a). Range expansion occurs through successive bouts of dispersal with subsequent speciation ("island hopping", MacArthur & Wilson, 1967). The starting situation is a small range to which new areas are added sequentially, to be eventually removed again owing to speciation. Typical examples would be the colonization of archipelagos (Gillespie & Rodrick, 2002; Keppel, Lowe & Possingham, 2009) or of new available areas, for example by deglaciation (Hewitt, 2000). Under this scenario, there should be a general positive relationship between phylogenetic and geographical distances, with more distant species having on average older divergences. However, this relationship will be asymmetrical (triangular in a bivariate plot; Figure 1a, "Scenario 1"): while there cannot be species that are geographically distant but phylogenetically close (unless there was long-range dispersal), there can be species that are geographically close but phylogenetically distant (e.g. species resulting from the initial, most ancient splits). The age of the species will generally increase towards the geographical origin of the range expansion (as postulated by Hennig, 1966 progression rule).

#### 2.1.2 | Scenario 2

Speciation by increased reduction of gene flow between "islands" of favourable conditions. In this case, speciation occurs owing to a fragmentation of the initial range. The starting situation could be a species that extends its range in a short time due to a window of favourable conditions, either climatic (e.g. Rizzo, Comas, Fadrigue, Fresneda & Ribera, 2013) or ecological (e.g. after deglaciation, García-Vázquez & Ribera, 2016). If there is a progressive and more or less uniform degradation of the general conditions that allowed the initial range expansion, populations of the species could become increasingly isolated in refuges of still suitable habitat (Figure 1a, "Scenario 2"). The fragmentation resulting from suboptimal conditions should most likely affect first the area of minimum gene flow, which would correspond to the largest area with the lowest population density. If the degradation of the suitable conditions continues there would be successive fragmentations, each of them sequentially

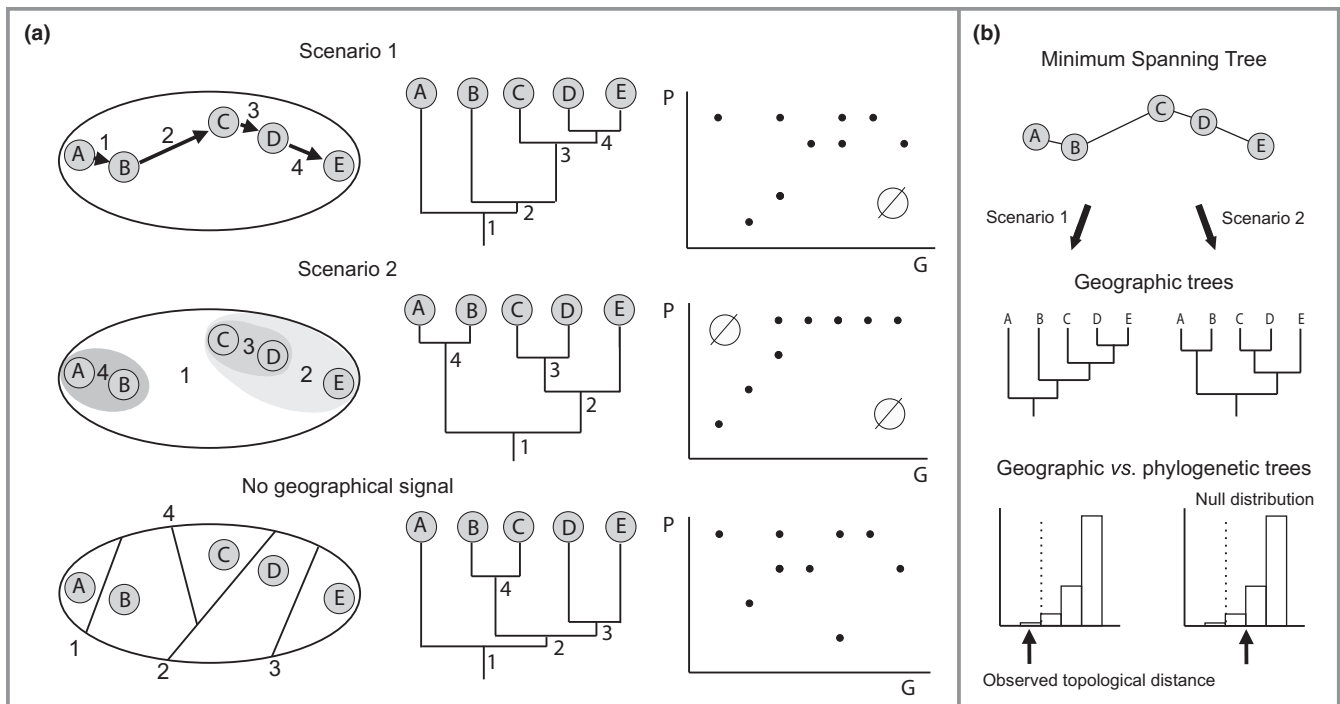
affecting the next largest area with the lowest population density. This scenario would be equivalent to the refuge speciation of Moritz, Patton, Schneider and Smith (2000) or to the vicariance by niche conservatism of Wiens (2004). In this case, the general relationship between geographical and phylogenetic distances should increase monotonically. Under a perfect scenario, the match between geographical and phylogenetic distances should be optimal; that is, the topology of the observed phylogeny should be identical to the one that minimizes the geographical distances between species.

The lack of association between phylogeny and geography could be the result of independent range shifts large enough to have erased the geographical signal of diversification under either scenario, but also be the result of other speciation modes. For instance, a process of speciation by vicariance owing to the formation of random barriers to gene flow ("no geographical signal" in Figure 1a) could result in a lack of correlation between geographical and phylogenetic distances, as species that are close geographically may have large phylogenetic divergences, and vice versa. Under such model, as in Scenario 2, speciation also occurs owing to the fragmentation of the initial, maximum range, but in this case the fragmentation occurs as a result of vicariance events independently of the distribution of the species. Typical examples could be fragmentation of a range by a raise of sea level, or by tectonic fragmentation of microplates (e.g. Sanmartín, 2003).

### 2.2 | Testing for geographical signal of diversification

We first used Mantel tests to assess the relationship between phylogenetic (patristic) distances and the geographical distance between the species range centroids. The comparison of the observed relationship between phylogenetic and geographical distances with suitable null models allows discriminating between the two potential scenarios of speciation outlined above. We also used bivariate plots of the geographical distance between the centroids of the species ranges with phylogenetic distances, as, as seen above, our two alternative scenarios should result in different patterns in the relationship between phylogenetic and geographical distances (Figure 1).

Furthermore, we performed an assessment of the match between the observed geographical distribution and the phylogenetic topology. First, we obtained a minimum spanning tree (MST) connecting the centroids of the species ranges, where the weight of the edge between each pair of points is the geographical distance between those two points. The result is a graph or network connecting all points minimizing the weight of the edges (see e.g. Graham & Hell, 1985), i.e. the distances among centroids. Then, this network is converted to a geographical tree or dendrogram representing the expected cladogenetic events according to the geography of speciation under each one of the scenarios outlined above. Two alternative trees can be tested, corresponding to Scenarios 1 and 2 above. In Scenario 1, the hierarchical structure of the tree will be determined simply by the ordination of the species in the MST, so that neighbour relationships are used to determine sequential splits in the geographical tree (Figure 1a). In this tree, the geographical neighbours



**FIGURE 1** (a) Schematic representation of the various hypothesized scenarios of diversification (from Ribera et al., 2011, modified). The first column shows the geographical distribution of the species (A to E); the second column, the phylogenetic relationships among them; the third column, a bivariate plot of the 10 geographical linear distances (G) vs. phylogenetic distances (P) (approximate values; note that some of the species pairs have identical values). The two tested scenarios are shown (see text), plus a generalized representation of other speciation modes that could result in no geographical signal: (1) speciation by stepping-stone colonization; (2) speciation by increased reduction of gene flow between favourable patches; and (3) speciation by vicariance owing to the formation of random barriers to gene flow (represented by lines). The three scenarios result in different patterns in the relationship between phylogenetic and geographical distances: (1) triangular relationship, with no closely related species at large geographical distances; (2) positive relationship, with neither closely related species at large geographical distances nor distant phylogenetic species at close geographical distances; and (3) no relationship. Note that the pattern represented in “no geographical signal” is indistinguishable from that resulting from random range movements obscuring scenarios 1 or 2 (see text for details). Numbers indicate the temporal sequence of the cladogenetic processes. (b) Major steps of assessment of the match between the observed geographical distribution and the topology obtained from the phylogeny: first, a minimum spanning tree connecting the centroids of the species ranges is obtained; second, this network is converted to a geographical tree representing the expected cladogenesis events according to the geography of speciation under each one of the scenarios outlined above; and third, topological distance between phylogenetic and geographical trees are computed and compared to a null distribution of distance values between the phylogenetic tree and random trees with the same taxa

are contiguous species in the phylogeny, irrespective of the length of the edges. As the MST trees are unrooted, the order of the cladogenetic events is not relevant. On the contrary, in Scenario 2 the sequence of cladogenetic events (and as consequence the resulting topology) will be determined by the length of the edges connecting the centroids (Figure 1a): the first split will be between the taxa at either extreme of the longest edge, the second will be between taxa at either end of the second longest, and so on until the two geographically closest species are separated. As noted above, a tree with random geographical relationships among species could be interpreted either as the result of other speciation modes (such as e.g. speciation by vicariance owing to the formation of random barriers to gene flow) (Figure 1a, “no geographical signal”), or to Scenarios 1 or 2 with subsequent range shifts erasing any geographical signal.

The match between the observed phylogenetic topology and the two topologies obtained with this procedure can be computed as a topological distance between two (unrooted) trees, measured as

twice the number of internal branches defining different bipartitions of the tips (Penny & Henny, 1985). It can then be tested if the observed topological distance is significantly smaller than what would be expected by chance by obtaining a null distribution of distance values between the phylogenetic tree and random trees with the same taxa (Figure 1b). This procedure differs from the one used by Ribera et al. (2011), who only assessed the probability that the observed phylogenetic topology was identical to that obtained from the MST build from the geographical location of species.

Under the framework outlined above, a significant concordance between geographical and phylogenetic patterns would be a compelling evidence of a non-random geographical arrangement of the species. A potential issue to be considered when assessing the association between geographical and phylogenetic patterns is the probability of false positives, i.e. the random generation of significant patterns (type I error). To estimate this source of error, we simulated random range centroids and phylogenetic relationships for sets of  $n$

species. We then applied our methodology to those simulations multiple times, and recorded the proportion of significant cases as an assessment of the probability of false positives (see Appendix S1 in Supporting Information for details). The results of these simulations showed that type I error was comparable to the standard 5% for both Mantel tests (mean 5.25%; range 3.7%–6.7%) and the procedure assessing the match between the phylogenetic and geographical topologies (mean 2.53%; range 1.3%–7.7%; see Table S1 in Appendix S1).

### 2.3 | Study cases: Western Palearctic water beetles

We used a phylogenetically heterogeneous set of 10 monophyletic lineages of water beetles (Table 1) occurring in the western Palearctic, some with both lotic and lentic species, and others encompassing exclusively either lotic or lentic species. The lineages used here belong to three families of two suborders of Coleoptera (Adephaga and Polyphaga), representing several independent invasions of the aquatic medium (Hunt et al., 2007). The full list of species and data used in this study are provided in Appendix S1. Within family Dytiscidae we used four lineages: (1) the *Ilybius subaeneus* group (genus *Ilybius*); (2) the western Mediterranean clade of the genus *Deronectes*; (3) genus *Graptodytes*; and (4) the *Hydroporus fuscipennis* group (genus *Hydroporus*). Within family Hydraenidae we included five lineages: (5) subgenus *Enicocerus* (genus *Ochthebius*); (6) the *Limnebius nitidus* complex (genus *Limnebius*); (7) the *Hydraena gracilis* and (8) the *H. dentipes* clade within the “*Haenydra*” lineage (genus *Hydraena*); and (9) subgenus *Phothydraena* (genus *Hydraena*). Lineage (10) was the west Mediterranean species of genus *Hydrochus* (family Hydrochidae).

It is possible that a relationship exists between phylogenetic and geographical signal within a subclade of the phylogeny, but that this relationship is masked when different subclades are considered in a single analysis. We thus also analysed monophyletic subclades at different hierarchical levels within the main lineages studied (see Appendix S1 for the delimitation of subclades).

The phylogenetic relationships within each lineage of water beetles and their relative age of divergence, reconstructed from different combinations of mitochondrial and nuclear genes, were taken from previous works (Abellán & Ribera, 2011; Hidalgo-Galiana & Ribera, 2011; Ribera & Faille, 2010; Ribera, Castro & Hernando, 2010; Ribera et al., 2011; see Appendix S1 for the final trees used). Details about how phylogenies were obtained are provided in Abellán and Ribera (2011). Geographical range data were taken from Abellán and Ribera (2011).

Multiple Mantel tests for association between the matrix of phylogenetic distances and the matrix of geographical distances between centroids were done using the function “mantel” in the R (R Development Core Team, 2011) library “ecodist” (Goslee & Urban, 2007) with 10,000 randomizations to generate a null distribution and assess the significance of the results. Great-circle distances between range centroids were calculated using the function “dism”

in the R library “geosphere” (Hijmans, 2015). Although it could be possible to use weighted distances to account for different types of barriers distorting the landscape (e.g. Lawson, 2013), in all analyses we assume a “flat”, Euclidean landscape, in which the only alteration of the relationship between geographical and phylogenetic distances are range movements. This reduces the power of our tests, which are already very conservative (see Discussion), as a genuine relationship between geography and phylogeny may be masked and considered to be the result of random range movements.

For comparison, we also tested the association between the matrix of phylogenetic distances and the matrices of minimum distance between ranges, estimated as the minimum distance between the polygons representing the geographical range of two species, and percentage of range overlap calculated between every taxon pair, measured as the area of overlap divided by the area of the taxon with the smaller area (i.e., 100% overlap if a species range is enclosed in the range of the another species) as in previous studies (e.g. Barraclough & Vogler, 2000; Chesser & Zink, 1994).

The “mst” function in the R package “ape” (Paradis, Claude & Strimmer, 2004) was used to produce the MST, which were manually transformed into “geographical trees” according to the two tested scenarios. The topological distance between geographical and phylogenetic trees was computed with the function “dist.topo” of the “ape” package, which was compared to that obtained between the phylogenetic tree and 1,000 random trees with the same taxa. All statistical analyses were performed using R.

Phylogenetic uncertainty can potentially have an important impact when assessing the association between evolutionary relationships and geography. A simple method for controlling for uncertainty in the phylogenetic reconstruction involves repeating the analyses on each (or a subset of) alternative phylogenetic trees (Donoghue & Ackerly, 1996). As an example, we additionally accounted for phylogenetic uncertainty in the analyses of the *Limnebius nitidus* complex, which had low support for most of its internal nodes (see Appendix S1). Thus, we conducted the Mantel test and the procedure assessing the match between geographical and phylogenetic topologies for each of the last 1,000 trees obtained in the BEAST runs (Drummond & Rambaut, 2007) from the phylogeny of Hydraenidae of Abellán et al. (2013).

## 3 | RESULTS

Phylogenetic and geographical distances were significantly correlated in five of the 10 main clades examined, and in nine of the 22 tested subclades within them (Table 1). In total, 44% of the studied clades and subclades showed a significant correlation. While none of the four clades or subclades with only lentic species had a significant correlation, 10 of 18 clades with only running species had (chi-square test  $\chi^2 = 4.074$ , d.f. = 1,  $p = .044$ ). When the degree of range overlap was used, only 21.9% of the clades and subclades showed a significant correlation, and 28.1% when the distance between ranges was used (see Table S2 in Appendix S2).

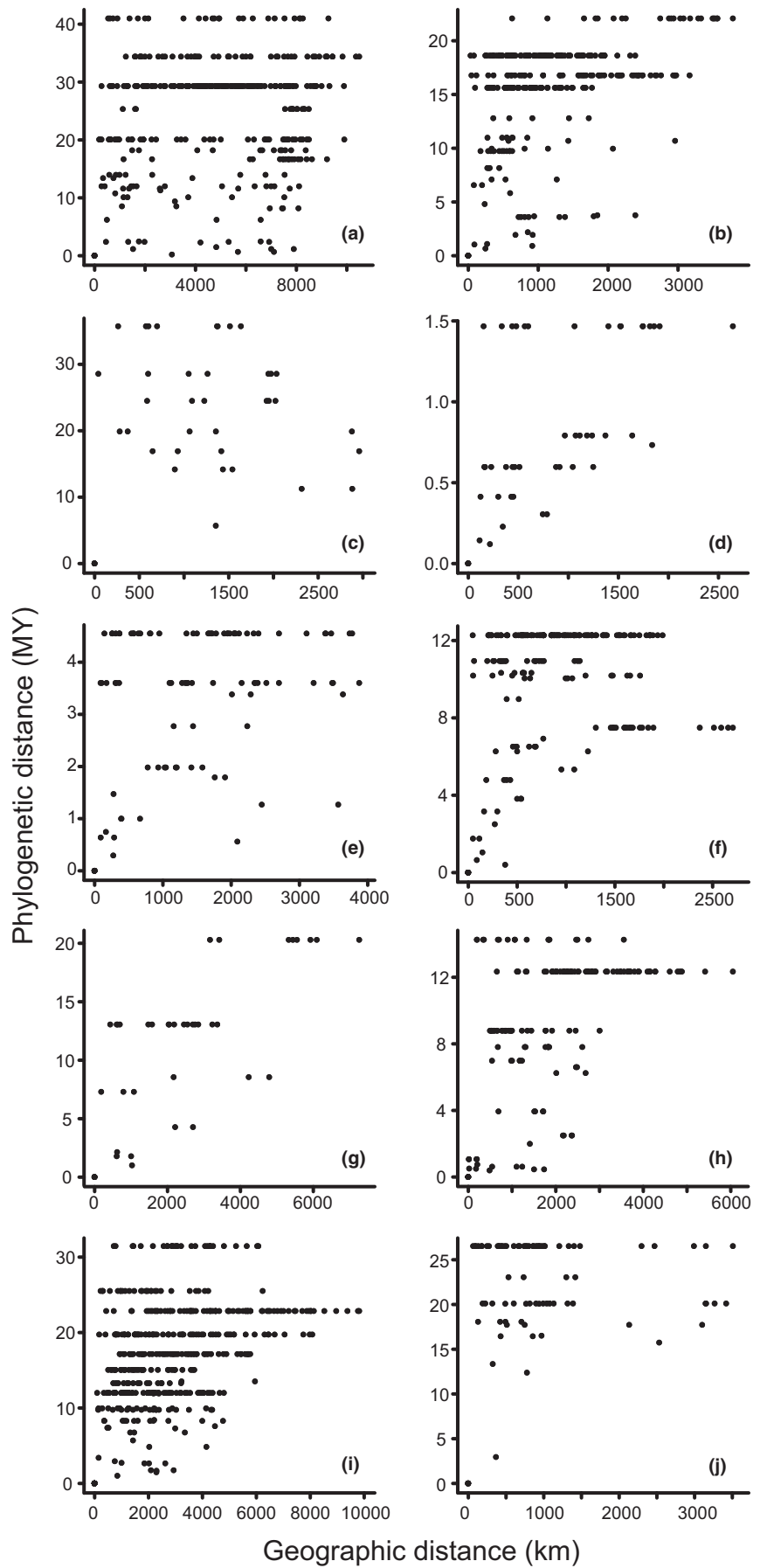
**TABLE 1** Results of the association between phylogenetic relationships and geography for the different lineages studied (and some subclades within them), as assessed by Mantel tests and the match between geographical and phylogenetic tree topologies (MST). Mantel's test analyses assess the correlation between phylogenetic and geographical distance matrices (distance between range centroids). *n* indicates the number of taxa in the lineage and *ND* indicates nodal distance between geographical and phylogenetic tree topologies. In bold, statistically significant *p* values at  $p < .05$ . The habitat type (HT) of the species in the clade is also indicated (r, running; s, standing; m, both). See Appendix S1 for the phylogenetic position of the different subclades

Lineage	Clade	<i>n</i>	Mantel test		MST scenario 1		MST scenario 2		HT
			<i>r</i>	<i>p</i>	ND	<i>p</i>	ND	<i>p</i>	
<i>Ilybius</i>	whole lineage	27	.10	.09	48	.59	46	.09	s
	<i>discedens</i>	24	.10	.08	42	.58	40	.09	s
	<i>vittiger</i>	8	.18	.11	8	.13	8	.13	s
	<i>guttiger</i>	11	.29	.07	16	.60	16	.60	s
<i>Deronectes</i>	whole lineage	24	.38	<b>&lt;.01</b>	42	.60	42	.60	r
	<i>semirufus</i>	5	.75	<b>.03</b>	2	.21	0	<b>.04</b>	r
	<i>moestus</i>	6	−.03	.45	6	.64	6	.64	r
	<i>algibensis</i>	8	.49	<b>&lt;.01</b>	8	.15	10	.63	r
<i>Enicocerus</i>	whole lineage	9	−.30	.06	12	.60	12	.60	r
	<i>exsculptus</i>	6	−.27	.27	6	.65	4	.18	r
<i>Limnebius nitidus</i>	whole lineage	10	.52	<b>.04</b>	10	<b>.03</b>	12	.16	r
	<i>maurus</i>	8	.64	<b>&lt;.01</b>	7	.05	10	.65	r
<i>Hydraena gracilis</i>	whole lineage	14	.22	.06	22	.60	18	<b>.02</b>	r
	<i>gracilis</i> s.s.	10	.31	<b>.04</b>	12	.13	10	<b>.03</b>	r
	<i>emarginata</i>	6	.90	<b>&lt;.01</b>	4	.18	2	<b>.04</b>	r
<i>Hydraena dentipes</i>	whole lineage	20	.11	.09	30	<b>&lt;.01</b>	24	<b>&lt;.01</b>	r
	<i>evanescens</i>	10	.39	<b>.02</b>	11	<b>.02</b>	12	.09	r
	<i>catalonica</i>	5	.68	.08	1	.08	2	.20	r
	<i>gaditana</i>	9	.69	<b>&lt;.01</b>	7	<b>&lt;.01</b>	4	<b>&lt;.01</b>	r
	<i>manfredjaechi</i>	5	.90	<b>.03</b>	2	.20	0	<b>.04</b>	r
<i>Phothydraena</i>	whole lineage	9	.67	<b>&lt;.01</b>	12	.60	10	.12	m
<i>Graptodytes</i>	whole lineage	18	.44	<b>&lt;.01</b>	27	<b>.01</b>	24	<b>&lt;.01</b>	m
	<i>veterator</i>	6	.46	.07	5	.29	4	.18	s
	<i>atlantis</i>	7	.24	.19	6	.21	4	<b>.04</b>	m
	<i>fractus</i>	5	.97	.10	4	.66	4	.66	r
<i>Hydroporus</i>	whole lineage	31	.34	<b>.01</b>	50	<b>.02</b>	50	<b>.02</b>	m
	<i>discretus</i>	8	−.17	.22	9	.29	10	.65	r
	<i>tessellatus</i>	8	.09	.28	9	.31	8	.19	m
	<i>nigrita</i>	17	.00	.48	28	.61	26	.12	s
<i>Hydrochus</i>	whole lineage	13	−.04	.48	20	.62	20	.62	m
	<i>flavipennis</i>	8	.21	<b>.04</b>	8	.21	10	.68	m
	<i>aljibensis</i>	5	.35	.16	3	.34	4	.67	m

The bivariate plots between geographical and phylogenetic distances displayed a triangular relationship in most of the main clades, with no closely related species at large geographical distances but with some distantly related species in close proximity (Figure 2). Exceptions were *Ilybius*, *Enicocerus* and, to a lesser extent, *Deronectes* and the *Hydraena gracilis* clade, for which the plot showed a general lack of relationship between phylogenetic and geographical distances. This lack of relationship was maintained for subclades within the *Ilybius* lineage (with only lentic species), and in some subclades within the *Deronectes* and

*Hydroporus* lineages. For other lineages, most of the subclades showed a positive relation between phylogenetic and geographical distances, with some cases showing a monotonic increase of phylogenetic distance as the distance between centroids increased, without apparent outliers and in particular without phylogenetically distant species pairs in close geographical proximity (e.g. *semirufus* subclade within *Deronectes*, *emarginata* subclade within the *Hydraena gracilis* lineage, and *catalonica* and *manfredjaechi* subclades within the *Hydraena dentipes* lineage, all exclusively lotic) (see Fig. S1 in Appendix S2).





**FIGURE 2** Bivariate plots of the geographical distance between the centroid of the species ranges (km) against the phylogenetic distance (age estimate in MY) for the studied lineages of water beetles: (a) *Ilybius*; (b) *Deronectes*; (c) *Enicocerus*; (d) *Limnebius*; (e) *Hydraena gracilis*; (f) *Hydraena dentipes*; (g) *Phothyaena*; (h) *Graptodytes*; (i) *Hydroporus*; (j) *Hydrochus*

In five of the main clades the observed geographical distribution, as represented in a geographical MST, was significantly more congruent with the topology obtained from the phylogeny that expected at random, under either Scenario 1 or 2 (*Limnebius*, *Hydraena gracilis*, *Hydraena dentipes*, *Graptodytes* and *Hydroporus*, Table 1). This was also the case for different subclades within the *Hydraena gracilis*, *Hydraena dentipes*, *Graptodytes* and *Deronectes* clades (Table 1). Overall, c. 38% of the clades (12 out of 32) showed a higher congruence than expected at random between the MST topology and the phylogeny, showing a significant association between the current distribution of the species and their evolutionary relationships. Among the significant cases, geographical topologies according to Scenario 2 were more frequent (10 of the 12), and when both scenarios were significant, the topological distance was often lower for the geographical tree obtained under Scenario 2 (Table 1). Again, none of the clades or subclades including exclusively lentic species showed significant congruence between phylogenetic and geographical trees.

Results of the comparison between geographical MSTs and phylogenetic topologies for the *Limnebius* lineage were robust to phylogenetic uncertainty: for 87% of the 1,000 examined trees the phylogenetic topology was significantly more congruent with the geographical MST topology that expected at random. Mantel tests between phylogenetic and geographical distances were significant for only 43% of these trees (mean  $\pm$  SD correlation coefficients were  $0.49 \pm 0.03$ ).

## 4 | DISCUSSION

Our results demonstrate the frequent persistence of a geographical signal of the mode of diversification among the studied lineages, which depending on the method was still traceable in between a third and half of those tested. The finding that in a particular clade there is a strong correlation between geographical and phylogenetic distance is evidence of the non-randomness of the spatial distribution of its species. A significant similarity between the observed phylogenetic topology and that obtained from the geographical distribution of the species can also be taken as compelling evidence of a non-random geographical arrangement of the species. Hence, although there are undoubtedly cases in which range movements have erased the geographical signal of speciation, we should not throw the baby out with the bath water: in many others there may be still a signal, and this signal could give essential information to understand the evolution of many lineages and the general patterns of speciation and diversification. Widespread, highly mobile species tend to attract disproportionate attention, as they are usually common and easily accessible. This obscures the fact that in most—if not all—major lineages the general pattern is a predominance of rare, localized species (Brown, Stevens & Kaufman, 1996; Gaston, 2003; Schoener, 1987), which are the most likely to have maintained a geographical signal of speciation. This may have been the case with range movements due to past climatic changes, of which a quantitative analyses show that overall there is not much evidence

to support the assumption of major range shifts in the majority of species, especially in southern areas (Abellán, Benetti, Angus & Ribera, 2011).

In most clades studied here, diversification was better described by a model in which species have originated through the fragmentation of a more widely distributed ancestor. This fragmentation could be the consequence of a progressive and uniform degradation of the general conditions that allowed the initial range expansion (Ribera et al., 2011). This diversification mode may have been common in the Pleistocene for different groups of aquatic Coleoptera, leaving a recognizable signal in the phylogeny, but only for the more recent clades. New range expansions may originate clusters of species overlapping their geographical ranges with the previous cycles of expansion-speciation, thus erasing the signal at deep phylogenetic levels.

Exclusive lentic clades showed a much weaker geographical signal, as expected from their average higher mobility. The contrasting geological permanence and stability of lotic and lentic waters has been related to the evolutionary and biogeographical differences observed between species in the two habitat types (see Ribera, 2008 for a review). The higher mobility and colonization ability of populations of lentic species has been shown to lead to more dynamic and larger geographical ranges, as well as faster colonization of new areas (e.g. after deglaciation), a pattern observed across a range of taxa, regions and systems (e.g. Abellán, et al., 2011; Hof, Brändle, & Brandl, 2006; Ribera & Vogler, 2000; Ribera et al., 2001; Ribera et al., 2003; Sánchez-Fernández, Lobo, Millán & Ribera, 2012). The fact that our approach recovered these well-established differences suggests it produces biologically meaningful results.

Our results contrast with some previous studies, which generally found a lower prevalence of lineages retaining a strong geographical signal of speciation (e.g. Barraclough & Vogler, 2000; Fitzpatrick & Turelli, 2006; Lynch, 1989). A major difference of our approach is that it does not rely on the assessment of the degree of range overlap between species or clades, but on the relative position of the range centroids. Our results depend only on the relative position of the centroid of the distribution of each species in relation to the others, not on possible range expansions or contractions around this centroid. The relative position of the centroid seems to be less evolutionary labile than extensions of the range (e.g. Nabout, Terribile, Bini & Diniz-Filho, 2010), which will typically change at rates faster than that of speciation (Coyne & Orr, 2004). In a previous study involving the same water beetle lineages (Abellán & Ribera, 2011), we found that the phylogenetic signal of range size was relatively weak when compared with other range properties such as the position of range centroid. It is interesting to note that with the use of centroids, Cardillo (2015) also reports a high degree of conservatism in the geographical range of mammals.

Our approach overcomes other limitations of previous analyses. Thus, it does not require the detailed knowledge of the range limits of all species of a clade, or the reconstruction of ancestral ranges for internal nodes. This is usually done by pooling the ranges of the extant species for calculating their overlap at different levels in the phylogeny (e.g. Barraclough & Vogler, 2000), and has been criticized



as unrealistic (Chesser & Zink, 1994; Losos & Glor, 2003). As noted by Berlocher and Feder (2002), reconstructed range areas are necessarily larger at deeper nodes, and in consequence there is a higher probability of sympatry in comparisons deeper in the tree topology. The increase in degree of sympatry through time may thus be partly generated as an artefact of the way the method treats sister clades. Although Fitzpatrick and Turelli (2006) proposed an alternative method which uses nested averages of the pairwise overlaps between species in each clade, which does not require complete taxon sampling and does not attempt to reconstruct ancestral ranges, it is still more sensitive to range changes.

Despite the high frequency of clades retaining geographical signal of diversification, it should be stressed that our approach is strongly conservative. First, the statistical power tends to be low, as the number of taxa may be too small to detect non-random patterns in the correlations and the tree comparisons. Although Mantel tests have recently been criticized as having low statistical power in some phylogenetic comparative analyses (Harmon & Glor, 2010), this problem may be unavoidable when analysing the type of data Mantel tests were originally designed to deal with—namely, data that can only be expressed as pairwise distances (e.g. geographical distances between populations) (Harmon & Glor, 2010). Of the two statistical tests we applied, Mantel tests are the least stringent, and may still be significant when the position of the centroids of two close species swap, or when (as a result of, for example, a geographically biased range expansion) the current centroid changes its relative position with respect to the nearest neighbours. By contrast, the test comparing tree topologies is more restrictive, in that any change of the relative position of even nearest-neighbour species would result in a suboptimal topology.

A more important limitation of our approach is that diversification may simply not follow the tested models. While a positive result is a clear indication that species have maintained their relative geographical positions, when there is no significant correlation, or the topology does not match geographical distances, it is not possible to affirm that there have been range movements. As noted above, other modes of speciation may result in a lack of correlation even if the species remain in place, what would be considered a “random” pattern under our assumptions. The lack of association between phylogeny and geography could thus indicate either alternative modes of speciation or the breakdown of the expected pattern owing to subsequent changes in the range of some of the species. Additional practical limitations could be incomplete taxon sampling or topological uncertainties in the phylogenies, as missing or extinct taxa would result in the overestimation of the phylogenetic age of the sampled species and can also influence the topology of both geographical and phylogenetic trees, leading to the incorrect acceptance of scenario 1.

#### 4.1 | Concluding remarks

We have shown that it is possible to obtain strong evidence of stasis of the geographical ranges of species, as well as information regarding the diversification mechanism, through the study of their

phylogenetic relationships and current distributions. Our results for Western Palaearctic water beetles often point to a common pattern of geographically conserved ranges where current species seem to have originated through the fragmentation of an ancestor with a wider distribution. Overall, our findings contradict the view that current ranges are not informative, and that evolutionary inferences concerning the geography of species in the recent past will often not be reliable.

#### ACKNOWLEDGEMENTS

We thank all the collectors who have provided us with specimens for the phylogenies, as well as Ana Izquierdo and Rocio Alonso for lab work. We also thank three anonymous referees for useful comments to earlier versions of the manuscript. This research was supported by projects CGL2007-61665 and CGL2010-15755 to I.R.

#### REFERENCES

- Abellán, P., Benetti, C. J., Angus, R. B., & Ribera, I. (2011). A review of Quaternary range shifts in European aquatic Coleoptera. *Global Ecology and Biogeography*, 20, 87–100.
- Abellán, P., & Ribera, I. (2011). Geographic location and phylogeny are the main determinants of the size of the geographical range in aquatic beetles. *BMC Evolutionary Biology*, 11, 344.
- Abellán, P., Sánchez-Fernández, D., Picazo, F., Millán, A., Lobo, J. M., & Ribera, I. (2013). Preserving the evolutionary history of freshwater biota in Iberian National Parks. *Biological Conservation*, 162, 116–126.
- Anderson, S., & Evensen, M. K. (1978). Randomness in allopatric speciation. *Systematic Zoology*, 27, 421–430.
- Barracough, T. G., & Nee, S. (2001). Phylogenetics and speciation. *Trends in Ecology and Evolution*, 16, 391–399.
- Barracough, T. G., & Vogler, A. P. (2000). Detecting the geographical pattern of speciation from species level phylogenies. *The American Naturalist*, 155, 419–434.
- Berlocher, S. H., & Feder, J. L. (2002). Sympatric speciation in phytophagous insects: Moving beyond controversy. *Annual Reviews of Entomology*, 47, 773–815.
- Bolnick, D., & Fitzpatrick, B. (2007). Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, 38, 459–487.
- Brooks, D. R., & McLennan, D. A. (1991). *Phylogeny, Ecology, and Behaviour: A Research Program in Comparative Biology*. Chicago: Chicago University Press.
- Brown, J. H., Stevens, G. C., & Kaufman, D. M. (1996). The geographic range: Size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics*, 27, 597–623.
- Cardillo, M. (2015). Geographic range shifts do not erase the historic signal of speciation in mammals. *The American Naturalist*, 185, 343–353.
- Carranza, S., & Wade, E. (2004). Taxonomic revision of Algero-Tunisian *Pleurodeles* (Caudata: Salamandridae) using molecular and morphological data. Revalidation of the taxon *Pleurodeles nebulosus* (Guichenot, 1850). *Zootaxa*, 488, 1–24.
- Chesser, R. T., & Zink, R. M. (1994). Modes of speciation in birds: A test of Lynch's method. *Evolution*, 48, 490–497.
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.
- Davis, M. B., & Shaw, R. G. (2001). Range shifts and adaptive responses to Quaternary climate change. *Science*, 292, 673–679.

- Donoghue, M. J., & Ackerly, D. D. (1996). Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351, 1241–1249.
- Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214.
- Fitzpatrick, B. M., Fordyce, J. A., & Gavrillets, S. (2009). Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology*, 22, 2342–2347.
- Fitzpatrick, B. M., & Turelli, M. (2006). The geography of mammalian speciation: Mixed signals from phylogenies and range maps. *Evolution*, 60, 601–615.
- García-Vázquez, D., & Ribera, I. (2016). The origin of widespread species in a poor dispersing lineage (diving beetle genus *Deronectes*). *PeerJ*, 4, e2514.
- Gaston, K. J. (1998). Species-range size distributions: Products of speciation, extinction and transformation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353, 219–230.
- Gaston, K. J. (2003). *The structure and dynamics of geographic ranges*. New York: Oxford University Press Inc..
- Gillespie, R. G., & Roderick, G. K. (2002). Arthropods on islands: Colonization, speciation, and conservation. *Annual Review of Entomology*, 47, 595–632.
- Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, 22, 1–19.
- Graham, R. L., & Hell, P. (1985). On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7, 43–57.
- Harmon, L. J., & Glor, R. E. (2010). Poor statistical performance of the Mantel test in phylogenetic comparative analyses. *Evolution*, 64, 2173–2178.
- Hennig, W. (1966). *Phylogenetic systematics*. Urbana: University of Illinois Press.
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.
- Hidalgo-Galiana, A., & Ribera, I. (2011). Late Miocene diversification of the genus *Hydrochus* (Coleoptera, Hydrochidae) in the west Mediterranean area. *Journal of Biogeography*, 38, 377–385.
- Hijmans, R. J. (2015) Introduction to the "geosphere" package. Technical report, geosphere vignette. Retrieved from <https://cran.r-project.org/web/packages/geosphere/vignettes/geosphere.pdf>.
- Hof, C., Brändle, M., & Brandl, R. (2006). Lentic odonates have larger and more northern ranges than lotic species. *Journal of Biogeography*, 33, 63–70.
- Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., St John, O., Wild, R., ... Vogler, A. P. (2007). A comprehensive phylogeny of beetles reveals the evolutionary origins of a super-radiation. *Science*, 318, 1913–1916.
- Jablonski, D. (1987). Heritability at the species level: Analysis of geographic ranges of Cretaceous mollusks. *Science*, 238, 360–363.
- Johnson, N. K., & Cicero, C. (2002). The role of ecologic diversification in sibling speciation of *Empidonax* flycatchers (Tyrannidae): Multigene evidence from mtDNA. *Molecular Ecology*, 11, 2065–2081.
- Keppel, G., Lowe, A. J., & Possingham, H. P. (2009). Changing perspectives on the biogeography of the tropical south Pacific: Influences of dispersal, vicariance and extinction. *Journal of Biogeography*, 36, 1035–1054.
- Lawson, L. P. (2013). Diversification in a biodiversity hot spot: Landscape correlates of phylogeographic patterns in the African spotted reed frog. *Molecular Ecology*, 22, 1947–1960.
- Lessa, E. P., Cook, J. A., & Patton, J. L. (2003). Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences USA*, 100, 10331–10334.
- Losos, J. B., & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, 18, 220–227.
- Lynch, J. D. (1989). The gauge of speciation: On the frequency of modes of speciation. In D. Otte, & J. A. Endler (Eds.), *Speciation and its consequences* (pp. 527–553). Sunderland, MA: Sinauer Associates.
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
- Martínez-Solano, I., Gonçalves, H. A., Arntzen, J. W., & García-París, M. (2004). Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography*, 31, 603–618.
- Mayr, E. (1963). *Animal species and evolution*. Cambridge, UK: Belknap.
- Moritz, C., Patton, J. L., Schneider, C. J., & Smith, T. B. (2000). Diversification of rainforest faunas: An integrated molecular approach. *Annual Review of Ecology and Systematics*, 31, 533–563.
- Nabout, J. C., Terribile, L. C., Bini, L. M., & Diniz-Filho, J. A. F. (2010). Phylogenetic autocorrelation and heritability of geographic range size, shape and position of fiddler crabs, genus *Uca* (Crustacea, Decapoda). *Journal of Zoological Systematics and Evolutionary Research*, 48, 102–108.
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20, 289–290.
- Penny, D., & Hendy, M. D. (1985). The use of tree comparison metrics. *Systematic Zoology*, 34, 75–82.
- Perret, M., Chautems, A., Spichiger, R., Barraclough, T. G., & Savolainen, V. (2007). The geographical pattern of speciation and floral diversification in the Neotropics: The tribe Sinningieae (Gesneriaceae) as a case study. *Evolution*, 61, 1641–1660.
- Phillimore, A. B., Orme, C. D. L., Thomas, G. H., Blackburn, T. M., Bennett, P. M., Gaston, K. J., & Owens, I. P. (2008). Sympatric speciation in birds is rare: Insights from range data and simulations. *The American Naturalist*, 171, 646–657.
- Pigot, A. L., Owens, I. P., & Orme, C. D. L. (2012). Speciation and extinction drive the appearance of directional range size evolution in phylogenies and the fossil record. *PLoS Biology*, 10, e1001260.
- R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ribera, I. (2008). Habitat constraints and the generation of diversity in freshwater macroinvertebrates. In J. Lancaster, & R. A. Briers (Eds.), *Aquatic insects: Challenges to populations* (pp. 289–311). Wallingford: CAB International Publishing.
- Ribera, I., Barraclough, T.G. & Vogler, A.P. (2001) The effect of habitat type on speciation rates and range movements in aquatic beetles: inferences from species-level phylogenies. *Molecular Ecology*, 10, 721–735.
- Ribera, I., Castro, A., Díaz-Pazos, J. A., Garrido, J., Izquierdo, A., Jäch, M. A., & Valladares, L. F. (2011). The geography of speciation in narrow range endemics of the "Haenydra" lineage (Coleoptera, Hydraenidae, *Hydraena*). *Journal of Biogeography*, 38, 502–516.
- Ribera, I., Castro, A., & Hernando, C. (2010). *Ochthebius (Enicocerus) aguilerai* p.n. from central Spain, with a molecular phylogeny of the Western Palaearctic species of *Enicocerus* (Coleoptera, Hydraenidae). *Zootaxa*, 2351, 1–13.
- Ribera, I., & Faille, A. (2010). A new microphthalmic stygobitic *Graptochytes* Seidlitz from Morocco, with a molecular phylogeny of the genus (Coleoptera, Dytiscidae). *Zootaxa*, 2641, 1–14.
- Ribera, I., Foster, G. N., & Vogler, A. P. (2003). Does habitat use explain large scale species richness patterns of aquatic beetles in Europe? *Ecography*, 26, 145–152.
- Ribera, I., & Vogler, A. (2000). Habitat type as a determinant of species range sizes: The example of lotic-lentic differences in aquatic Coleoptera. *Biological Journal of the Linnean Society*, 71, 35–52.
- Rizzo, V., Comas, J., Fadrique, F., Fresneda, J., & Ribera, I. (2013). Early Pliocene range expansion of a clade of subterranean Pyrenean beetles. *Journal of Biogeography*, 40, 1861–1873.
- Sánchez-Fernández, D., Lobo, J. M., Millán, A., & Ribera, I. (2012). Habitat type mediates equilibrium with climatic conditions in the distribution

- of Iberian diving beetles. *Global Ecology and Biogeography*, 21, 988–997.
- Sanmartín, I. (2003). Dispersal vs. vicariance in the Mediterranean: Historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography*, 30, 1883–1897.
- Schoener, T. W. (1987). The geographical distribution of rarity. *Oecologia*, 74, 161–173.
- Wiens, J. J. (2004). Speciation and ecology revisited: Phylogenetic niche conservatism and the origin of species. *Evolution*, 58, 193–197.

## BIOSKETCHES

**Pedro Abellán** is a post-doctoral fellow at Queens College, City University of New York. His interests include biogeography, ecology, evolution and conservation of freshwater biota. **Ignacio Ribera** is a senior researcher in the Institute of Evolutionary Biology (CSIC-UPF,

Barcelona). He is interested in the evolution, systematics and biogeography of Coleoptera.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Abellán P, Ribera I. Using phylogenies to trace the geographical signal of diversification. *J Biogeogr.* 2017;00:1–11. <https://doi.org/10.1111/jbi.13035>

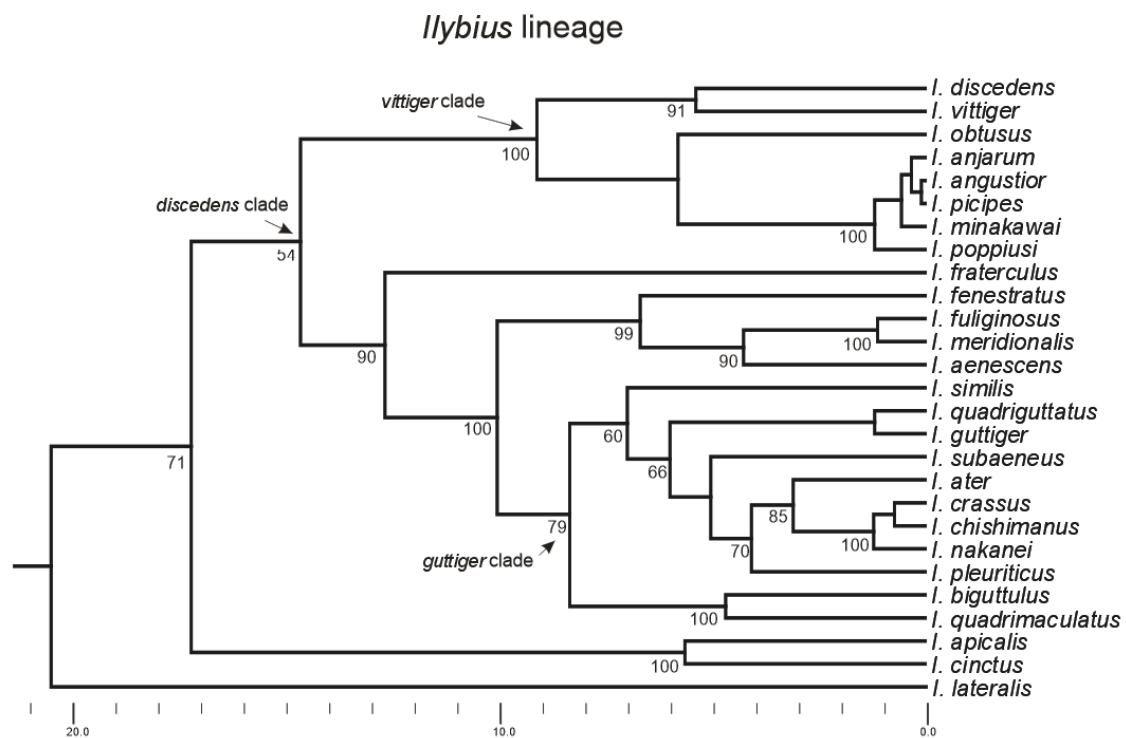


Finally, we included a lineage belonging to the family Hydrochidae: the west Mediterranean (Iberian Peninsula, Morocco and south France) species of the genus *Hydrochus* (Hidalgo-Galiana & Ribera, 2011), which forms a monophyletic group encompassing 13 species (including also *H. roberti*, so far recorded only from the Caucasus and Turkey), of which we studied 12 species.

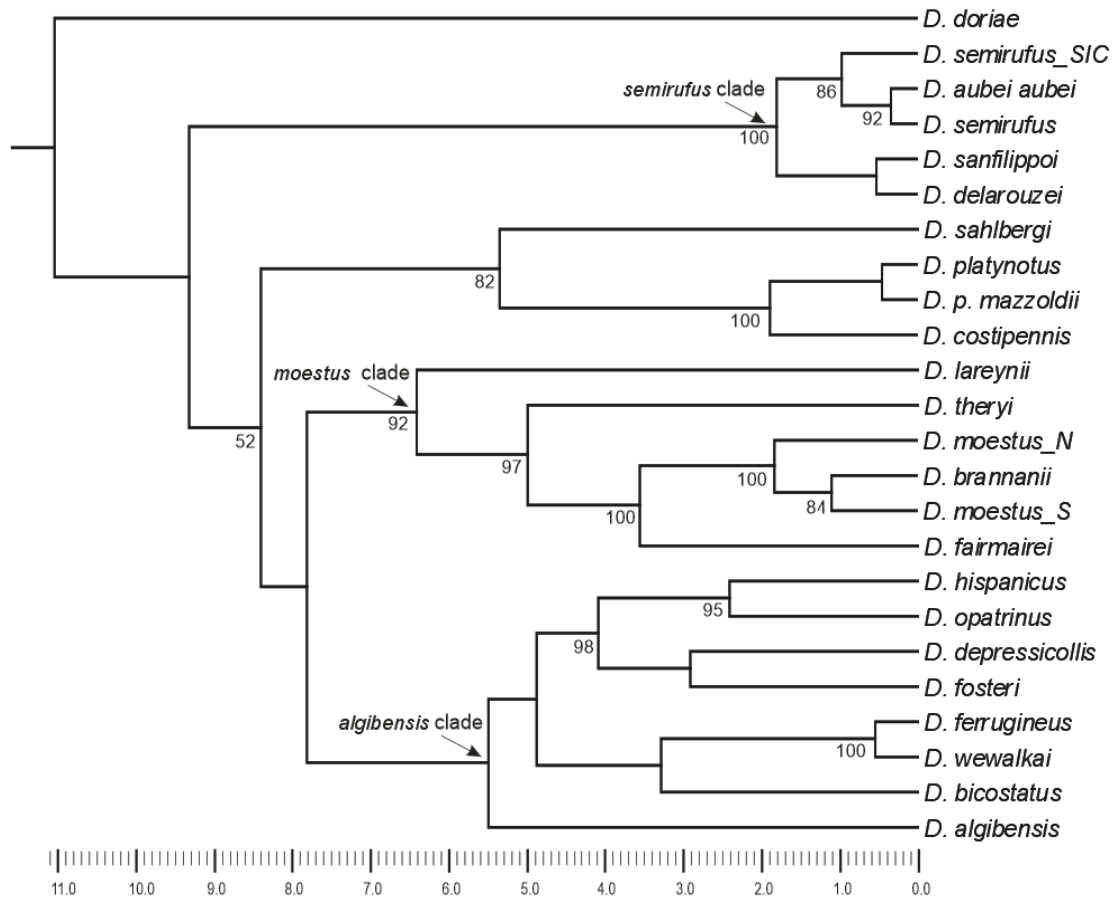
The phylogenetic relationships within each lineage of water beetles and their relative age of divergence, reconstructed from different combinations of mitochondrial and nuclear genes, were taken from recent works (Ribera & Faille, 2010; Ribera *et al.*, 2010; Ribera *et al.*, 2011; Hidalgo-Galiana & Ribera, 2011; Abellán & Ribera, 2011). Details on how phylogenies were obtained are provided in Abellán & Ribera (2011). For the *Phothydraena* lineage we included in the phylogeny two species for which no molecular data was available (Abellán & Ribera, 2011), but that are very closely related (according to the genital morphology) to one included species: *H. chersonesica* and *H. paganettii*, as sister to *H. atrata* (Jäch *et al.*, 2007)

### Ultrametric trees for the different lineages showing the subclades studied.

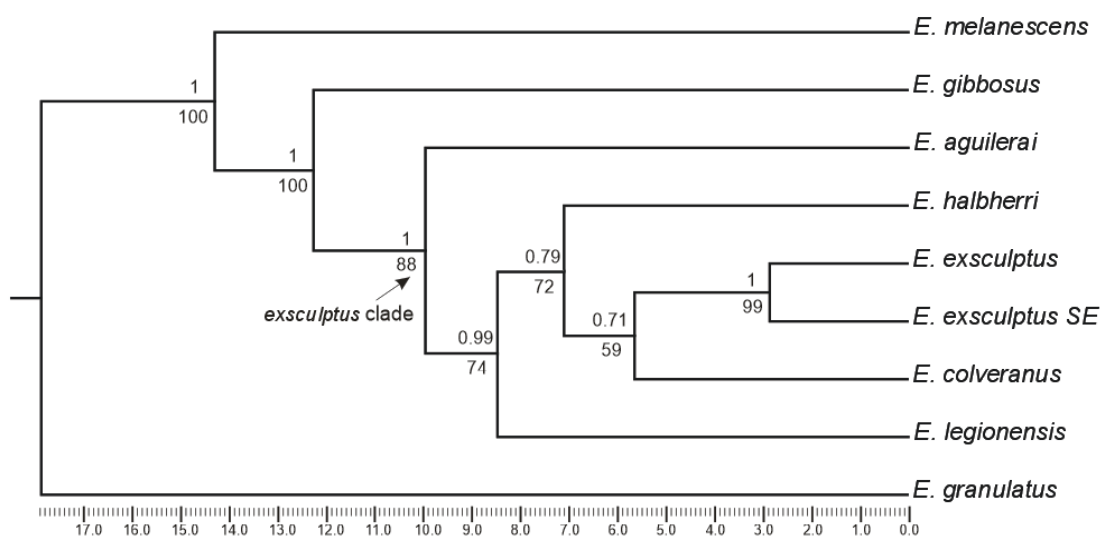
Numbers indicate node support: above nodes, Bayesian posterior probabilities (if above 0.5); below nodes, bootstrap support values from Maximum Likelihood analysis (if above 50%).



*Deronectes* lineage

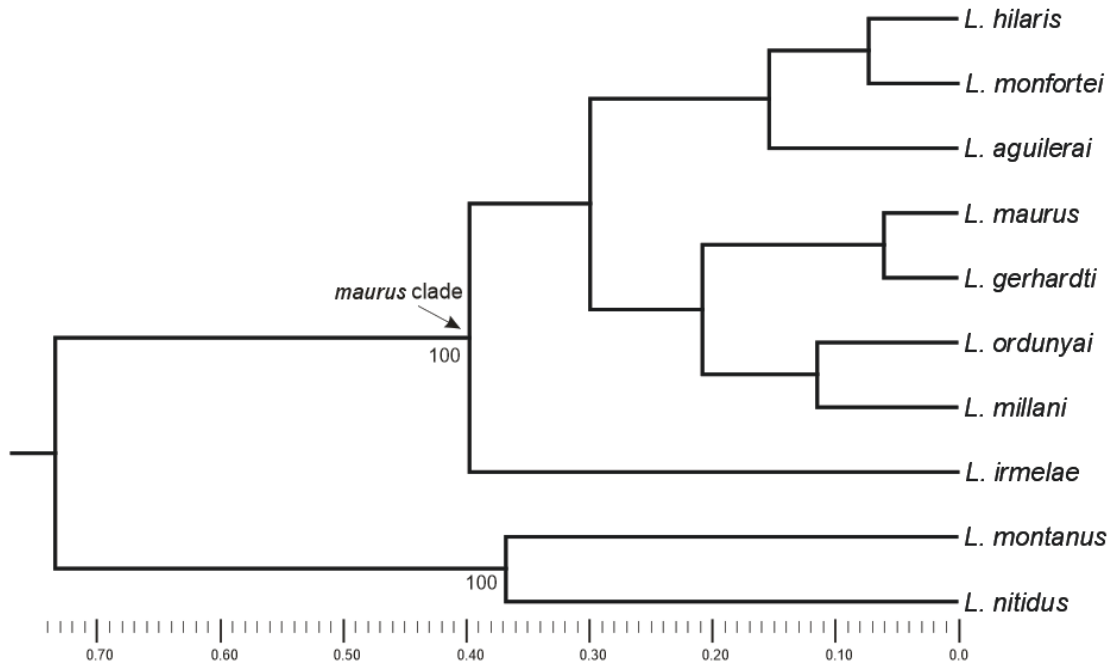


*Enicocerus* lineage

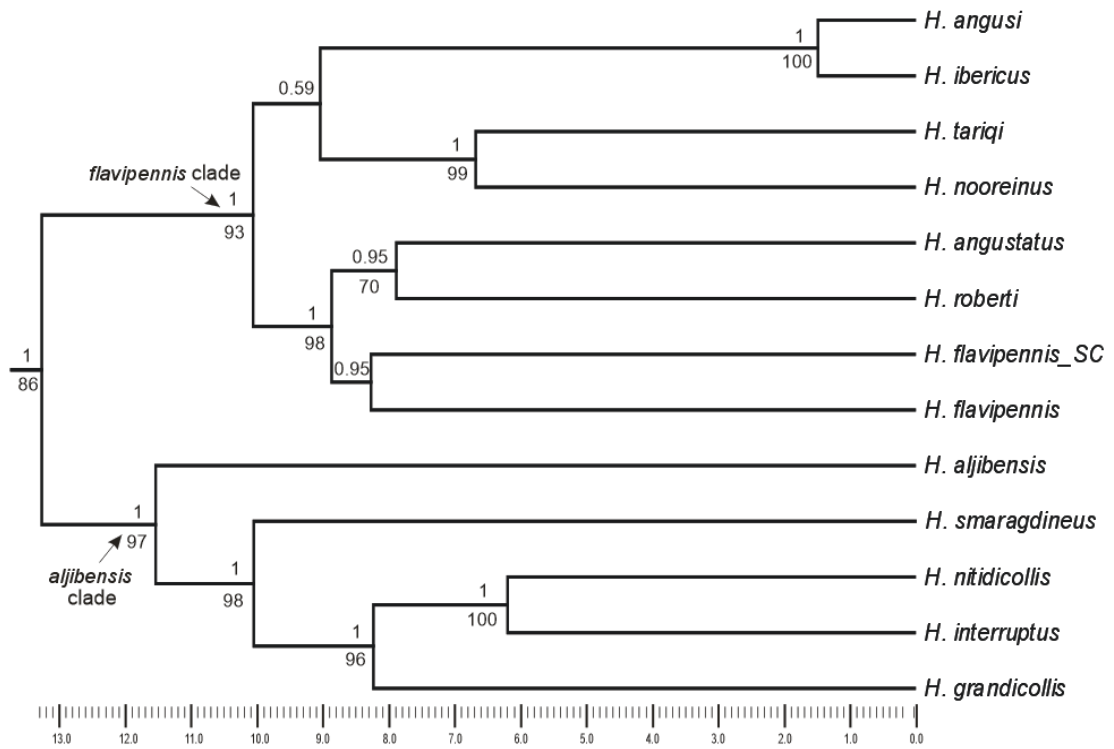




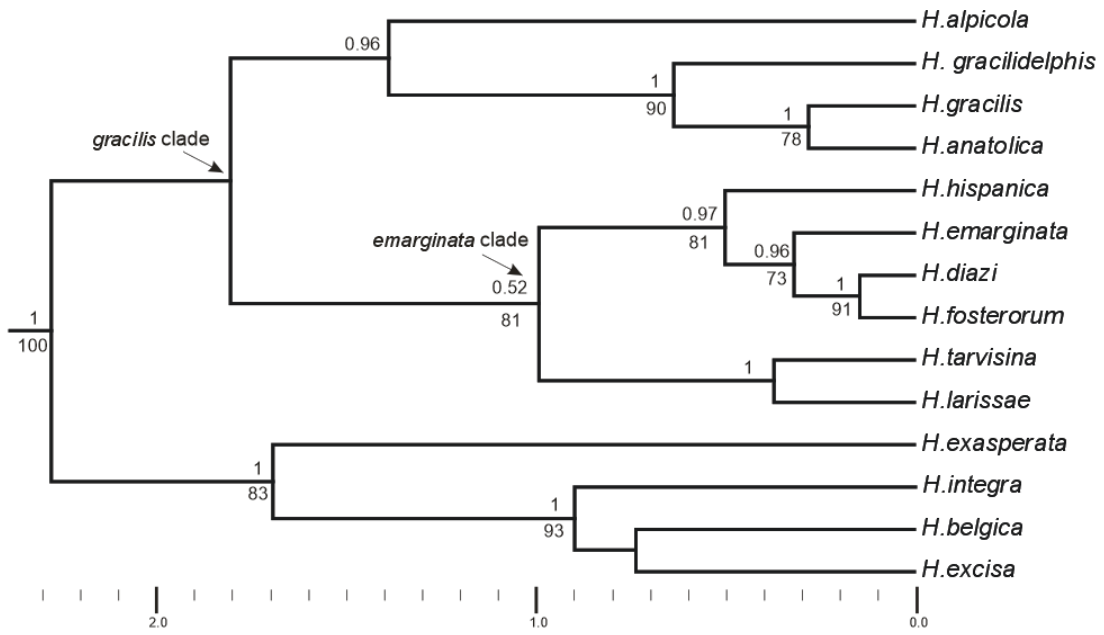
*Limnebius nitidus* subgroup lineage



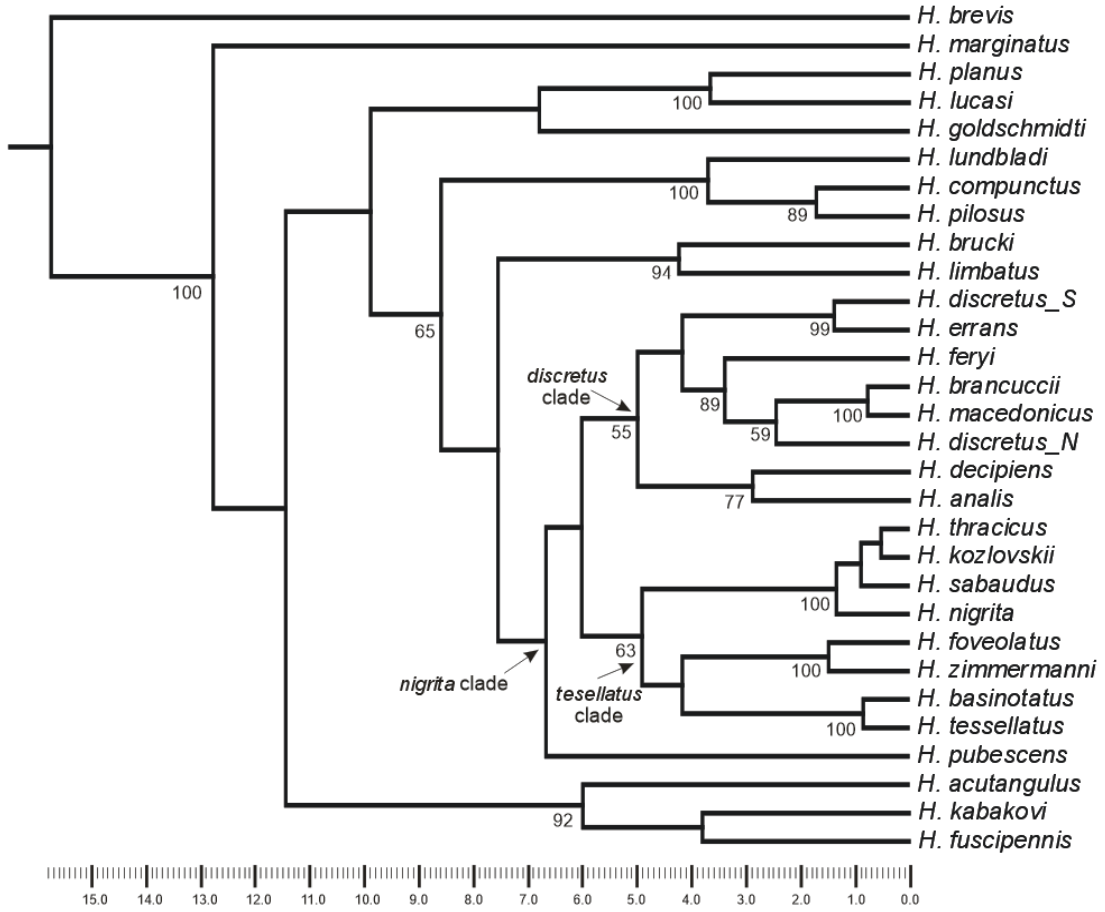
*Hydrochus* lineage



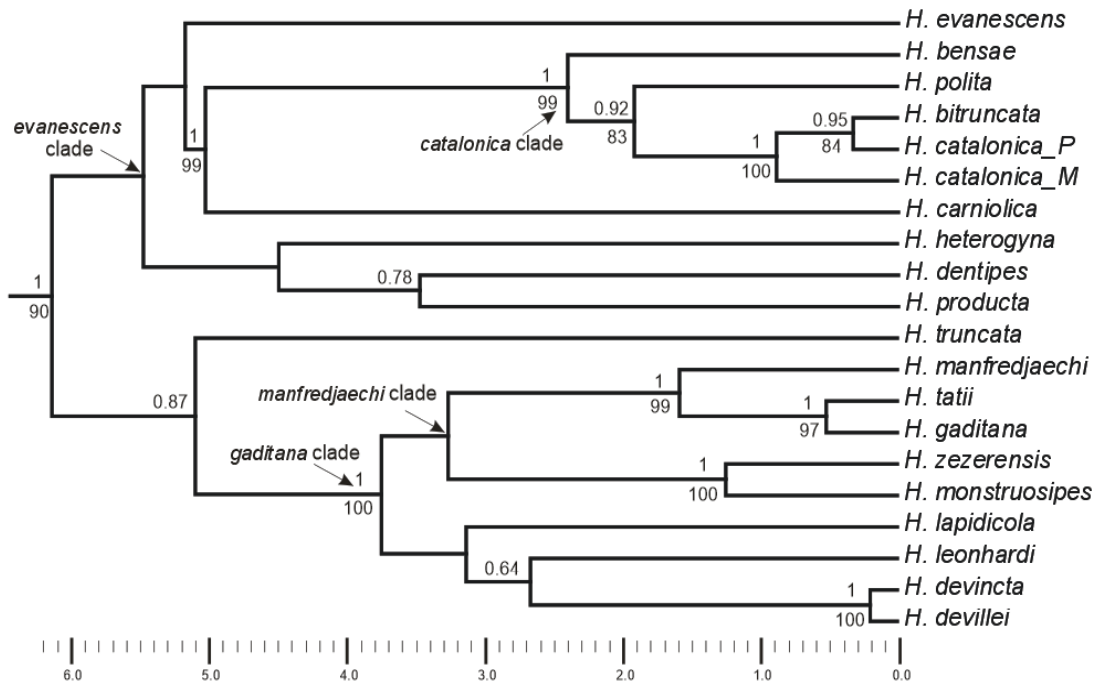
*Hydraena gracilis* lineage



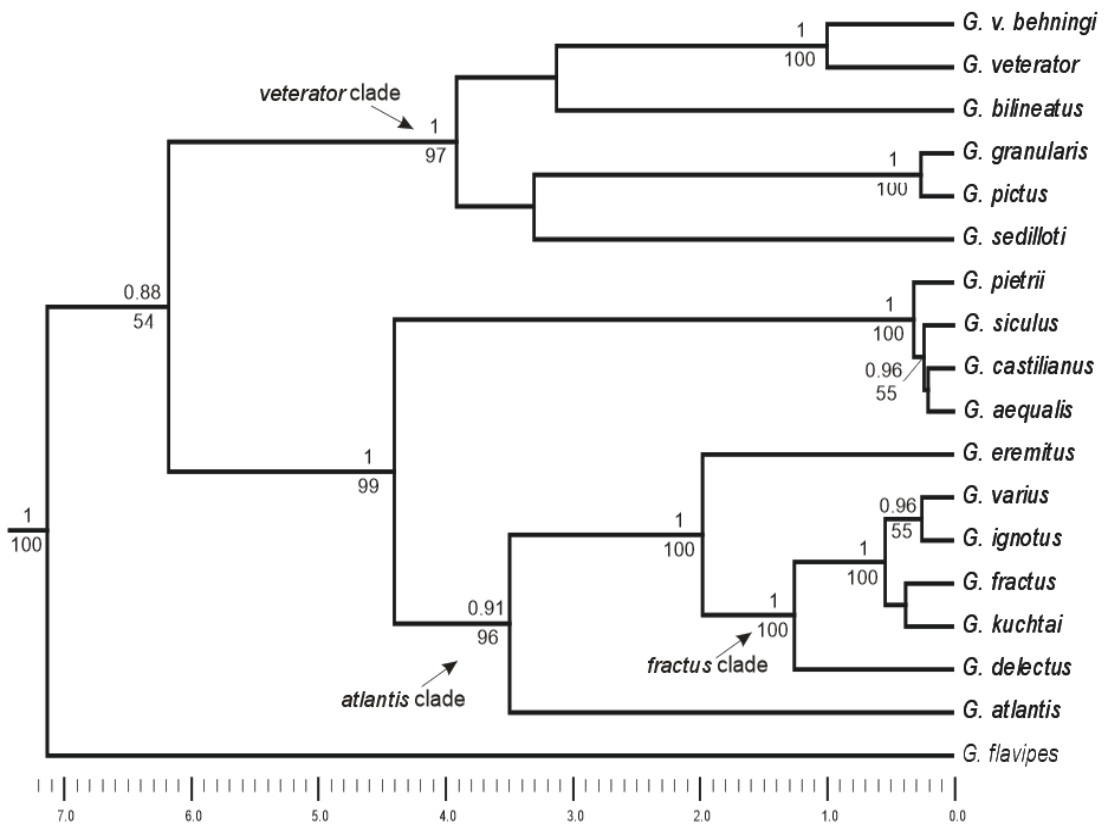
*Hydroporus* lineage



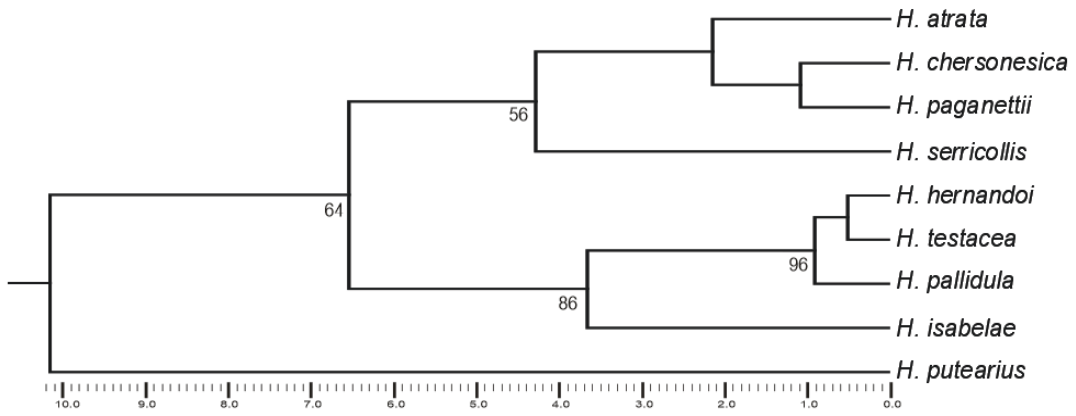
*Hydraena dentipes* lineage



*Graptodytes* lineage



*Phothydraena* lineage



## Assessing the probability of false positive results with simulations

In order to test the probability of false positive results when assessing the association between geographic and phylogenetic patterns, we performed a simulation analysis. We simulated random range centroids for sets of  $n$  species within a square continental area, and simulated also the phylogenetic relationships for these  $n$  species. Then, we applied our framework to those simulations. First, we used Mantel tests to test the relationship between phylogenetic distances for each simulated lineage of  $n$  species and the geographical distances of the corresponding random range centroids. The significance of the correlation was assessed with 10,000 randomisations. We used different species richness levels, from 5 to 100 in intervals of 5, and for each one we repeated this analysis 1000 times and recorded the proportion of significant cases as an assessment of the probability of false positives. Similarly, we assessed the match between the phylogenetic topology and the geographic topology under each one of the two diversification scenarios. The significance of the observed topological distances was evaluated obtaining a null distribution from 1000 random trees with the same taxa. Again, we repeated this analysis 1000 times for each species richness level and recorded the proportion of significant cases. Simulations and statistical analyses were performed using R and the packages ‘ape’, ‘ecodist’, ‘sp’ (Pebesma & Bivand 2005) and ‘igraph’ (Csardi & Nepusz 2006).

The results of these simulations are shown below (Table S1). The incidence of false positives was very low for both Mantel tests (mean 5.25%; range 3.7-6.7%) and the procedure assessing the match between the phylogenetic and geographic topologies (mean 2.53%; range 1.3-7.7%).

**Table S1.** Percentage of significant cases of association between phylogenetic relationships and geography for the simulation analyses, as assessed by Mantel tests and the match between geographic and phylogenetic tree topologies (MST).

# taxa	Mantel test	MST Scenario 1	MST Scenario 2
5	3.7	7.7	6.2
10	5.6	4.9	4.1
15	6.5	4.4	2.4
20	5.9	2.1	1.7
25	5.4	3.0	1.9
30	4.6	2.9	2.4
35	5.2	2.0	2.6
40	5.1	1.9	2.2
45	5.2	2.2	1.5
50	4.4	1.5	2.5
55	5.7	2.7	1.3
60	5.4	2.8	2.5
65	6.7	2.4	2.4
70	4.8	2.2	1.8
75	5.1	2.1	1.8
80	5.6	2.2	2.2
85	4.4	1.6	2.3

90	4.9	1.7	2.3
95	4.7	1.5	1.7
100	6.1	1.7	1.9

### Literature cited only in Appendix S1

- Csardi, G., & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal, Complex Systems*, **1695**: 1-9.
- Fresneda, J., & Ribera, I. (1998) Revision of the *Limnebius nitidus* (Marsham) subgroup (Coleoptera: Hydraenidae), with description of two new species and comments on their phylogeny and biogeography. *Entomologica Scandinavica*, **29**: 395-409.
- Hansen, M. (1998) *World Catalogue of Insects. Volume 1. Hydraenidae* (Coleoptera). Apollo Books, Stenstrup.
- Jäch, M.A. (1993) Taxonomic revision of the Palearctic species of the genus *Limnebius* Leach, 1815 (Coleoptera: Hydraenidae). *Koleopterologische Rundschau*, **63**:99-187.
- Jäch, M.A. (2004) Hydraenidae. Pages 102-122 in I. Löbl and A. Smetana, eds. *Catalogue of Palaearctic Coleoptera, Volume 2: Hydrophiloidea - Histeroidea - Staphylinoidea*. Apollo Books, Stenstrup.
- Jäch, M.A., Beutel, R., Díaz, J.A., & Kodada, J. (2000) Subgeneric classification, description of head structures, and world check list of *Hydraena* Kugelann (Insecta: Coleoptera: Hydraenidae). *Annalen des Naturhistorischen Museums in Wien*, **102**:177-258.
- Jäch, M.A., Díaz, J.A., & Przewoźny, M. (2007) *Hydraena* (s.str.) *chersonesica* sp.n. (Coleoptera: Hydraenidae), a new member of the *H. testacea* species group from Crimea (Ukraine). *Annalen des Naturhistorischen Museums in Wien*, **108B**: 95-102.
- Nilsson, A.N. (2001) *World Catalogue of Insects, volume 3, Dytiscidae*. Apollo Books, Stenstrup.
- Nilsson, A.N., & Holmen, M. (1995) The aquatic Adephaga (Coleoptera) of Fennoscandia and Denmark. II. Dytiscidae. *Fauna Entomologica Scandinavica*, **32**:1-192.
- Pebesma, E.J., & Bivand, R.S. (2005) S classes and methods for spatial data: the sp package, unpublished report.
- Ribera, I., Castro, A., & Hernando, C. (2010) *Ochthebius (Enicocerus) aguilerai* sp.n. from central Spain, with a molecular phylogeny of the Western Palearctic species of *Enicocerus* (Coleoptera, Hydraenidae). *Zootaxa*, **2351**:1-13.
- Ribera, I., & Faille, A. (2010) A new microphthalmic stygobitic *Graptodytes* Seidlitz from Morocco, with a molecular phylogeny of the genus (Coleoptera, Dytiscidae). *Zootaxa*, **2641**:1-14.
- Ribera, I., Bilton, D.T., Balke, M., & Hendrich, L. (2003) Evolution, mitochondrial DNA phylogeny and systematic position of the Macaronesian endemic *Hydrotarsus* Falkenstrom (Coleoptera: Dytiscidae). *Systematic Entomology*, **28**:493-508.
- Ribera, I., Hogan J.E., & Vogler A.P. (2002) Phylogeny of hydradephagan water beetles inferred from 18S rRNA sequences. *Molecular Phylogenetics and Evolution*, **23**:43-62.



## SUPPORTING INFORMATION FOR

## Using phylogenies to trace the geographic signal of diversification

Pedro Abellán &amp; Ignacio Ribera

## APPENDIX S2. SUPPLEMENTARY RESULTS

**Table S2.** Mantel's test analyses of correlation between phylogenetic distance and geographic distance between species ranges and percentage of geographic range overlap for the different lineages studied and some subclades within them. The correlation value and its significance p-value (in parenthesis) are reported. In bold, statistically significant correlation values at  $P < 0.05$ . The habitat type (HT) of most of the species in the clade is also indicated (r, running; s, standing; m, both).

Lineage	Clade	Range overlap	Range	HT
<i>Ilybius</i>	whole lineage	-0.22 ( <b>0.003*</b> )	0.07 (0.223)	s
	<i>discedens</i>	-0.19 ( <b>0.005</b> )	0.10 (0.077)	s
	<i>vittiger</i>	-0.17 (0.303)	0.31 (0.088)	s
	<i>guttiger</i>	-0.08 (0.309)	0.35 ( <b>0.045</b> )	s
<i>Deronectes</i>	whole lineage	-0.18 ( <b>0.005</b> )	0.35 ( <b>0.000*</b> )	r
	<i>semirufus</i>	-0.80 (0.065)	0.51 (0.065)	r
	<i>moestus</i>	0.16 (0.311)	0.40 (0.134)	r
	<i>algebensis</i>	-0.12 (0.260)	0.34 ( <b>0.021</b> )	r
<i>Enicocerus</i>	whole lineage	0.35 (0.097)	-0.11 (0.303)	r
	<i>exsculptus</i>	0.06 (0.544)	0.09 (0.425)	r
<i>Limnebius nitidus</i>	whole lineage	-0.27 (0.095)	0.17 (0.247)	r
	<i>maurus</i>	-0.24 (0.151)	0.60 ( <b>0.012</b> )	r
<i>H. gracilis</i>	whole lineage	0.09 (0.188)	0.14 (0.110)	r
	<i>gracilis</i> s.s.	0.13 (0.233)	0.23 (0.063)	r
	<i>emarginata</i>	-0.46 (0.053)	0.83 ( <b>0.020</b> )	r
<i>H. dentipes</i>	whole lineage	0.11 (0.098)	0.06 (0.249)	r
	<i>evanescens</i>	-0.02 (0.428)	0.31 ( <b>0.030</b> )	r
	<i>catalonica</i>	0.19 (0.407)	0.58 (0.068)	r
	<i>gaditana</i>	-0.01 (0.246)	0.64 ( <b>0.001*</b> )	r
	<i>manfredjaechi</i>	--	0.92 ( <b>0.034</b> )	r
<i>Phothydraena</i>	whole lineage	-0.40 ( <b>0.012</b> )	0.64 ( <b>0.006</b> )	m
<i>Graptodytes</i>	whole lineage	-0.01 (0.505)	0.14 (0.102)	m
	<i>veterator</i>	-0.19 (0.124)	0.18 (0.296)	s
	<i>atlantis</i>	-0.08 (0.426)	0.13 (0.315)	m
	<i>fractus</i>	-0.64 (0.077)	0.82 (0.167)	r
<i>Hydroporus fuscipennis</i>	whole lineage	-0.08 (0.205)	0.23 (0.076)	m
	<i>discretus</i>	0.15 (0.231)	-0.18 (0.227)	r
	<i>tessellatus</i>	0.01 (0.429)	0.13 (0.246)	m
	<i>nigrita</i>	0.19( <b>0.039</b> )	-0.04(0.364)	s
<i>Hydrochus</i>	whole lineage	0.09 (0.228)	-0.06 (0.352)	m
	<i>flavipennis</i>	0.26 ( <b>0.026</b> )	0.17 ( <b>0.036</b> )	m
	<i>aljibensis</i>	-0.67 ( <b>0.018</b> )	0.32 (0.402)	m

**Figure S1.** Bivariate plots of the geographic distance between the centroid of the species ranges (G; in km) against the phylogenetic distance (P; age estimate in MY) for different subclades within the studied lineages.

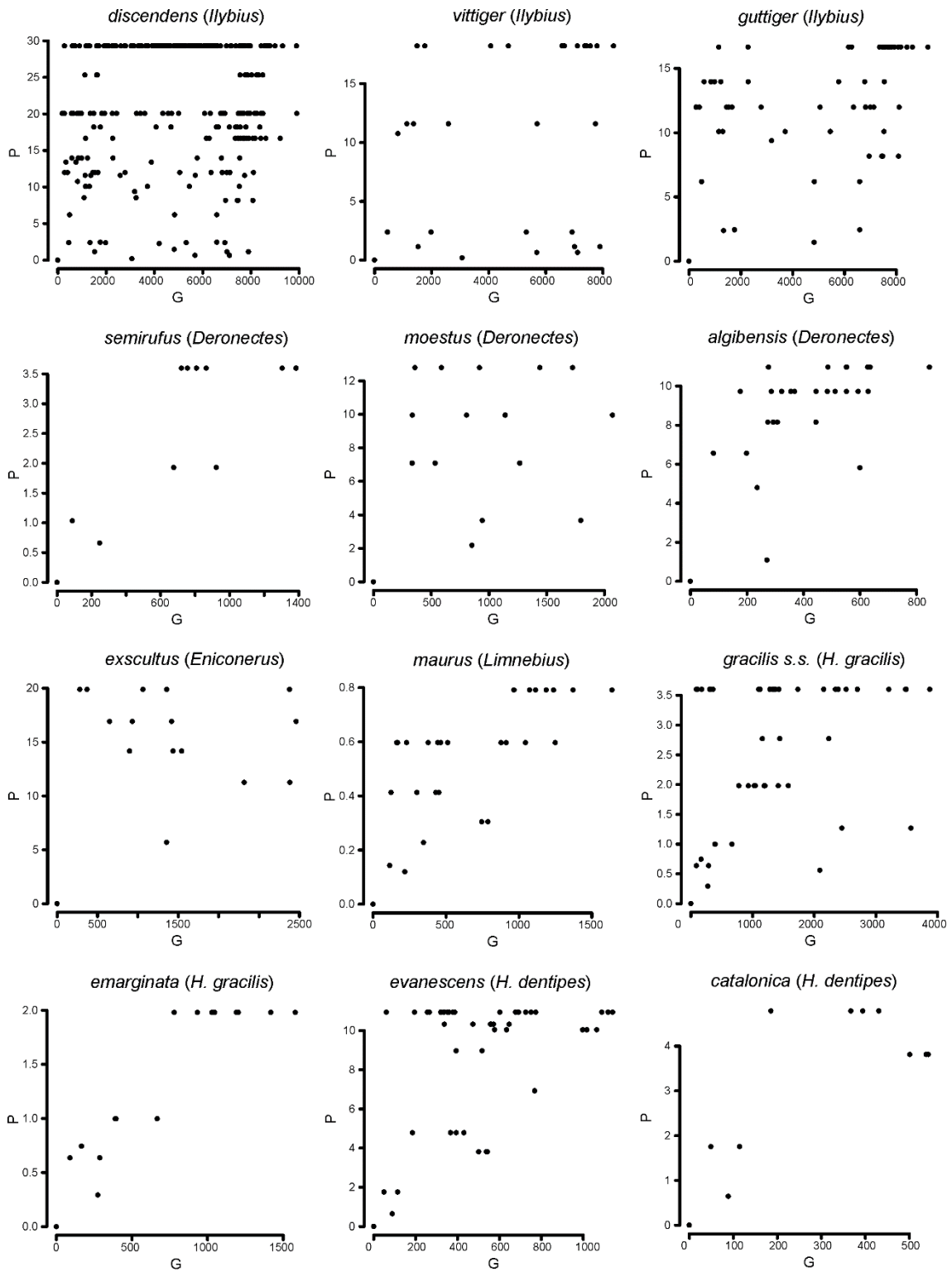


Figure S1. (Continued)

