

RESEARCH ARTICLE

Open Access

Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders

Leticia Bidegaray-Batista* and Miquel A Arnedo

Abstract

Background: The major islands of the Western Mediterranean—Corsica, Sardinia, and the Balearic Islands—are continental terrenees that drifted towards their present day location following a retreat from their original position on the eastern Iberian Peninsula about 30 million years ago. Several studies have taken advantage of this well-dated geological scenario to calibrate molecular rates in species for which distributions seemed to match this tectonic event. Nevertheless, the use of external calibration points has revealed that most of the present-day fauna on these islands post-dated the opening of the western Mediterranean basin. In this study, we use sequence information of the *cox1*, *nad1*, *16S*, *L1*, and *12S* mitochondrial genes and the *18S*, *28S*, and *h3* nuclear genes, along with relaxed clock models and a combination of biogeographic and fossil external calibration points, to test alternative historical scenarios of the evolutionary history of the ground-dweller spider genus *Parachtes* (Dysderidae), which is endemic to the region.

Results: We analyse 49 specimens representing populations of most *Parachtes* species and close relatives. Our results reveal that both the sequence of species formation in *Parachtes* and the estimated divergence times match the geochronological sequence of separation of the main islands, suggesting that the diversification of the group was driven by Tertiary plate tectonics. In addition, the confirmation that *Parachtes* diversification matches well-dated geological events provides a model framework to infer substitution rates of molecular markers. Divergence rates estimates ranged from 3.5% My⁻¹ (*nad1*) to 0.12% My⁻¹ (*28S*), and the average divergence rate for the mitochondrial genes was 2.25% My⁻¹, very close to the “standard” arthropod mitochondrial rate (2.3% My⁻¹).

Conclusions: Our study provides the first unequivocal evidence of terrestrial endemic fauna of the major western Mediterranean islands, whose origin can be traced back to the Oligocene separation of these islands from the continent. Moreover, our study provides useful information on the divergence rate estimates of the most commonly used genes for phylogenetic inference in non-model arthropods.

Background

The estimation of the timing of evolutionary events from DNA sequence information has become a major research topic in evolutionary biology. Although the use of molecular data to estimate divergence times goes back to the mid 60's of the past century [1], the number of studies that include time estimation has increased rapidly over the last decade due to the ever increasing amount of DNA sequence data and the development of new algorithms that relax the limiting assumptions of

the molecular clock (see reviews of [2-4]). Information on timescales has shed light not only on the origin of taxonomic groups but has also allowed for the testing of biogeographic and climatic hypotheses, the estimation of rates of species diversification, and the investigation of rates of molecular evolution, among other topics [5].

Genetic distances are transformed into absolute divergence times by incorporating calibration points. The choice of a specific calibration date is therefore potentially crucial to accurately infer molecular dates (e.g., [6]). Biogeographic and paleoecological data are among the main sources of information for dating phylogenetic nodes, either as a complement to the fossil record, or as

* Correspondence: letigaray@yahoo.com

Institut de Recerca de la Biodiversitat & Departament de Biologia Animal, Universitat de Barcelona, Av. Diagonal 643, 08020, Barcelona, Spain

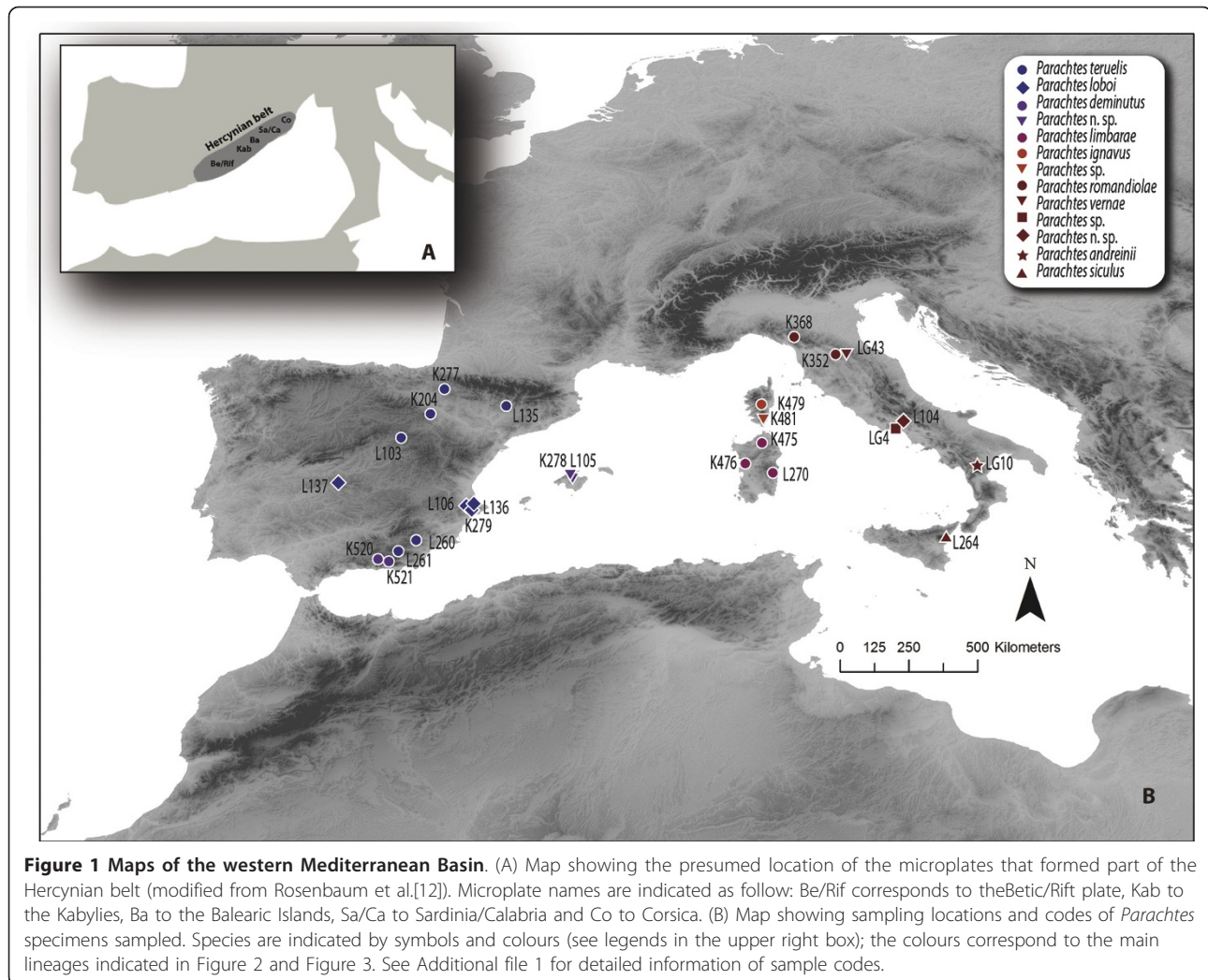
the only available evidence in poorly preserved organisms. The use of geologic and paleoclimatic evidence, however, has been criticised on the basis of their limited accuracy and unwarranted assumptions (e.g., [7-9]). Therefore, the appropriate use of biogeographic events to calibrate phylogenies requires a well-documented geochronology and the demonstration that the assumed barriers constitute a true obstacle to dispersal for the focal group.

The complex geological evolution of the western Mediterranean region, located at the edge of the converging African and Eurasian plates, is relatively well understood [10-17] and has been extensively used for molecular dating (e.g., [18-24]) as well as to explain present-day species distributions (e.g., [25]). The opening of the western Mediterranean basin started during the Alpine orogeny, when an eastward tectonic extension occurred between the African and Eurasian plates due to a subduction rollback between the oceanic and the continental slab (back arc extension). As a result, at the beginning of the Oligocene (~30-25 million years ago, Ma), several continental microplates that formed part of the Hercynian belt (Corsica, Sardinia, Balearic Islands, Calabro-Peloritan massif, including northern Sicily, the Kabylies and the Beatic-Rift Cordillera) broke off and started drifting from the eastern Iberian Peninsula and southern France to their present-day location (Figure 1(A)). During the first stage of the back arc extension, the northern microplate assemblage (Sardinia, Corsica and Calabro-Peloritan massif) started drifting counter-clockwise with respect to the Eurasian plate, whereas the southern assemblage (the Balearic Islands and Great Kabylie) drifted clockwise relative to Iberia. The movement of the northern microplates resulted in the opening of the Ligurian Sea, the Valencia Trough and the Gulf of Lyon. Corsica, Sardinia and the Calabro-Peloritan massif collided with the Apulian plate around 20-18 Ma, completing the formation of the Gulf of Lyon. There is no general consensus on the time of separation of Corsica and Sardinia, although it almost certainly occurred during the Corsica-Sardinia rotation. This rotation has been dated to approximately 21-15 Ma [13,26] based on paleomagnetic data and is further corroborated by the presence of early Burdigalian marine sediments in southern Corsica [27]. Likewise, there are doubts about the time when the exotic blocks of the Calabro-Peloritan massif drifted off from Sardinia to their present-day positions, although it is generally accepted that it occurred during the formation of the Tyrrhenian Sea [10,17]. According to Rosenbaum & Lister [28], the Tyrrhenian Sea started opening 10-9 Ma in the northern and western part, finishing around 5-4 Ma in the southern part. The same geological processes that generated the Tyrrhenian Sea also broke off the land connection

between the Maghrebides and Sicily, opening the Strait of Sicily [12]. The Provençal and the Algerian basins are both older than the Tyrrhenian Sea, and originated during the Early-Middle Miocene (21-15 Ma), when the Kabylies block broke off from the Balearic Islands (~21 Ma) and started drifting southward until they collided with the African margin (18-15 Ma). The formation of the Alboran basin is a matter of debate because the original position of the Rif-Betic mountain range is uncertain [12,15]. Some authors suggest that during the Oligocene, the Rif-Betic mountain range formed a continuous orogenic belt together with the western Alps, Calabria, Corsica and the Kabylies blocks [12]. At the time of the main subduction rollback, the Rif-Betic block drifted southwest up to its present location (~10 Ma), completing the formation of the Alboran basin.

Although the formation of the main western Mediterranean islands has been used to date evolutionary divergence times and to explain species ranges, most studies using independent calibration points or substitution rates estimated from related organisms have concluded that the origin of present-day fauna in the region greatly post-dates the opening of the basin [29-33]. In particular, the Messinian Salinity Crisis (MSC) has been pointed out as one of the main drivers of local diversification. The late Miocene (8 Ma) marine corridors in southern Iberia and northern Morocco closed about 5.96 Ma, isolating the Mediterranean Sea from the Atlantic Ocean and causing a large reduction in Mediterranean water levels and the emergence of land connections between North Africa, Corsica, Sardinia and Eurasia and between the Balearic Islands and Iberia [34,35,16]. The opening of the Strait of Gibraltar about 5.3 Ma restored the water exchange between the Atlantic and Mediterranean waters, reestablishing effective isolation of the island ecosystems [36,37]. To date, the plant family Araceae constitutes the only case reported of Mediterranean organisms in which diversification has been shown to correspond to the Oligocene break-off of the Hercynian belt [38].

The ground-dweller spider genus *Parachtes* Alicata, 1964 (Araneae, Dysderidae) is a promising candidate for a terrestrial animal whose diversification may have been shaped by western Mediterranean plate tectonics. This genus is restricted to the western Mediterranean, where it exhibits a disjunct distribution [39]. It currently includes 12 species: *P. ignavus* Simon, 1882 and *P. inaequipes* Simon 1882 from Corsica, *P. limbarae* Kraus, 1955 from Sardinia, *P. siculus* Caporiacco, 1949 from Sicily, *P. romandiola* Caporiacco, 1949, *P. vernae* Caporiacco, 1936, *P. latialis* Alicata, 1966 and *P. andreinii* Alicata, 1966 from the Italian Peninsula, and *P. cantabrarum* Simon, 1914, *P. teruelis* Kraus, 1955, *P. lobo* Jiménez-Valverde, 2006 and *P. deminutus* Denis, 1957



from the Iberian Peninsula. Two new species, one endemic to Majorca (Balearic Islands) and another to the Italian Peninsula (Lazio), have been recently discovered and are awaiting formal description (Arnedo, unpublished data). *Parachtes* shows a remarkable uniformity in its somatic morphology and most species' diagnostic features are restricted to the male and female genitalia [39]. *Parachtes* species are nocturnal wandering hunters, usually found in leaf-litter, under dead logs or stones and in dark and humid habitats at mid to high altitudes (> 500 m). These spiders do not build webs, and ballooning (i.e., aerial dispersal by means of silk threads) has been reported neither in the genus nor in the whole family. The low vagility and habitat preferences suggest limited opportunities for overseas dispersal, although transport by floating islands has been proposed to explain colonisation of the oceanic Canary Islands by the closely related genus *Dysdera* [40].

Here, we investigate the role of the major geological events associated with the opening of the Western Mediterranean basin in shaping diversification of the spider genus *Parachtes*, by inferring a molecular phylogeny based on eight mitochondrial and nuclear genes from a thorough taxonomic sampling of the genus and its close relatives. Relaxed clock models in combination with multiple, independent biogeographic and fossil calibration points are further used to reconstruct the temporal framework of species diversification.

Results

Samples and sequences analysed

The specimens and sequences analysed in the present study are summarised in Additional file 1. The alignment of the non-protein coding genes, including informative gap characters coded as absence/presence characters (see Additional file 2 for details), were

merged with the protein-coding genes (*cox1* = 1257, *nad1* = 358 and *h3* = 327 characters), resulting in a combined matrix of 4454 characters for the “default alignment”, 4462 for the “gappy alignment” and 4447 for the “compressed alignment”. Gap characters were included in the parsimony and Bayesian inference analyses but excluded in maximum likelihood and divergence time analyses.

Phylogenetic analyses

Preliminary analysis using parsimony conducted on the concatenated matrices of the default, gappy and compressed alignments resulted in almost identical topologies and similar jackknife supports, although jackknife values were slightly higher for the default alignments (Figure 2). The *I2S* and *I6S-L1* genes were the most variable and hence the most sensitive to changes in the alignment parameter values, recovering slightly different topologies. Differences, however, involved unsupported alternative positions.

Bayesian inference and maximum likelihood analyses were only conducted on the default alignment concatenated matrix, because no remarkable differences were found among the resulting topologies of each alignment. The selected substitution models for each gene and gene partition are shown in Additional file 3. Phylogenetic analyses under the three inference methods and partition strategies resulted in almost identical topologies, summarised in Figure 2.

All methods support the monophyly of *Parachtes* and its sister group relationship to the western Mediterranean species of *Harpactocrates*. As suspected [41], the eastern European species currently included in *Harpactocrates* did not form a clade with its western counterparts. The most basal split in the genus *Parachtes* separates species of the Iberian Peninsula (Iberian clade) from those of the western Mediterranean islands, Italy and the Betic species (Hercynian-Italian clade). Although all analyses support this basal split, the support for the Hercynian-Italian clade was moderate to low (parsimony jackknife 67%, Bayesian posterior probability 0.95, maximum likelihood bootstrap 51%), and the AU topology test could not reject the alternative topology where all Iberian and the Balearic species formed a clade sister to the remaining species ($p = 0.464, 0.469, 0.473$, for P1, P2 and P3, respectively). Maximum likelihood and Bayesian inference analyses conducted on P2 and P3 partition schemes, however, increased support for the Hercynian-Italian clade (72/0.96 and 84/0.98, respectively, see Figure 2). The Hercynian-Italian clade was further resolved into two clades, one clade including the Balearic and Betic species, and a second clade where the Italian species form a monophyletic group sister to Corsica, with both of them in turn

sister to Sardinia. All of the former clades were highly supported.

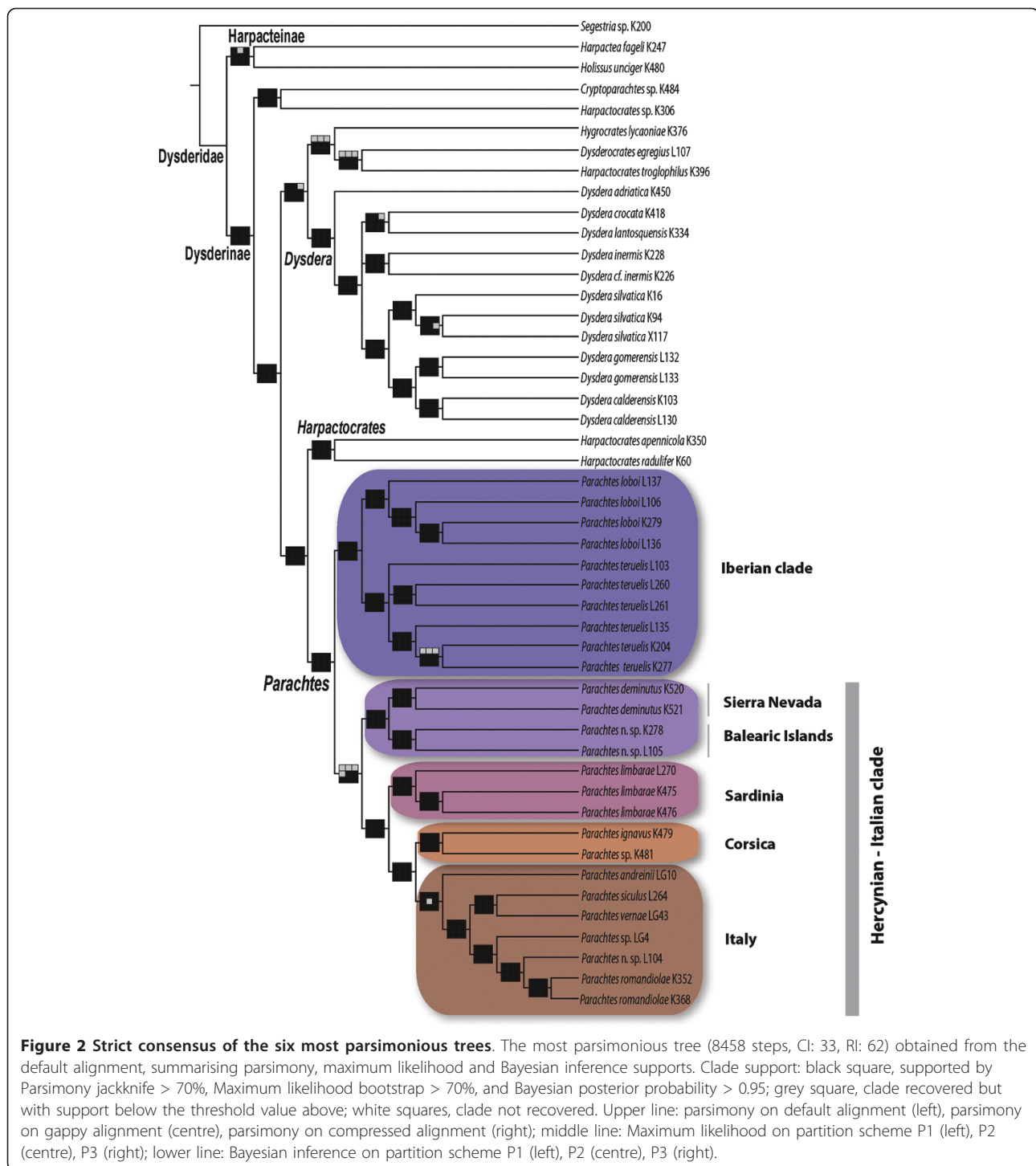
Temporal framework of *Parachtes* diversification

Bayes factor analysis provided decisive support for selecting relaxed clock models against the strict clock, but could not discriminate between the two Bayesian relaxed clocks. The lognormal model, however, yielded the highest harmonic mean likelihood and was used in subsequent analyses. Likewise, BEAST analyses using Yule as tree prior yielded the highest likelihood harmonic means, but provided no strong evidence as compared to the more complex birth-death models. Finally, Bayes factors indicated decisive evidence in favour of the P3 partition scheme over the alternative schemes (Bayes factors summarised in Additional file 4).

The chronogram with corresponding confidence intervals obtained with the preferred analysis options is shown in Figure 3. The time of the split between the Iberian clade and the Hercynian-Italian clade was estimated at 23.16 Ma (31.81-15.32 Ma), while the time of divergence between the Balearic-Betic clade and the Sardinian-Corsican-Italian clade was estimated at 19.17 Ma (26.63-12.24). The Balearic and the Betic clades split at 9.73 Ma (16.14-4.09). Sardinia split from the Corsican-Italian clade at 14.34 Ma (20.81-9.00), and the Corsican lineage split from the Italian clade at 11.27 Ma (16.28-6.65), similar to the estimated time of divergence of the two Iberian species *P. lobo* and *P. teruelis*, 11.09 Ma (17.39-5.50). The southern Italian species *P. andreinii* diverged from the remaining Italian species at 8.80 Ma (12.93-4.96). Italian *P. siculus* and *P. vernae* split from *P. romandiola* and *P. n. sp.* at 5.90 Ma (9.14-3.23). The most recent species splits correspond to the two former pairs, at 3.20 Ma (5.40-1.26) and 1.69 Ma (3.14-0.56), respectively.

The substitution rates inferred for each gene partition were (mean pairwise divergence My^{-1}): 3.5% (*nad1*), 2.5% (*cox1*), 1.7% (*I2S*), 1.3% (*I6S-L1*), 0.22% (*h3*) and 0.12% (*28S*) (Table 1). The *I8S* substitution rate could not be estimated due to low genetic variability and limited phylogenetic resolution. All genes evolved in a non-clocklike manner, as indicated by *ucl.d.stdev* values and the high coefficient of variation, which reflects heterogeneity in substitution rates along tree branches. The nuclear gene *28S* showed higher deviation from the strict clock, as *ucl.d.stev* and the coefficient of variation were greater than 1. Moreover, the 95% HPD interval of the covariance values span zero in all genes, which means that there is not evidence of autocorrelation of rates in the phylogeny (see BEAST user manual).

The lineage age estimates obtained in *r8s* under the preferred method selected by cross-validation analyses for the different partition schemes are summarised in



Additional file 5. Differences among divergence time estimation methods and partition strategies at different phylogenetic levels were assessed by monitoring the following clades: the Dysderidae node, the *Parachtes* node, and the Corsican-Italian node. The analyses performed with r8s and BEAST yielded similar estimates for the *Parachtes* and the Corsica-Italian nodes, but

differed greatly for the deeper Dysderidae node (see Figure 4).

Discussion

Timing of *Parachtes* diversification

Although our discussion and conclusions are based on the time estimates obtained using Bayesian methods, we

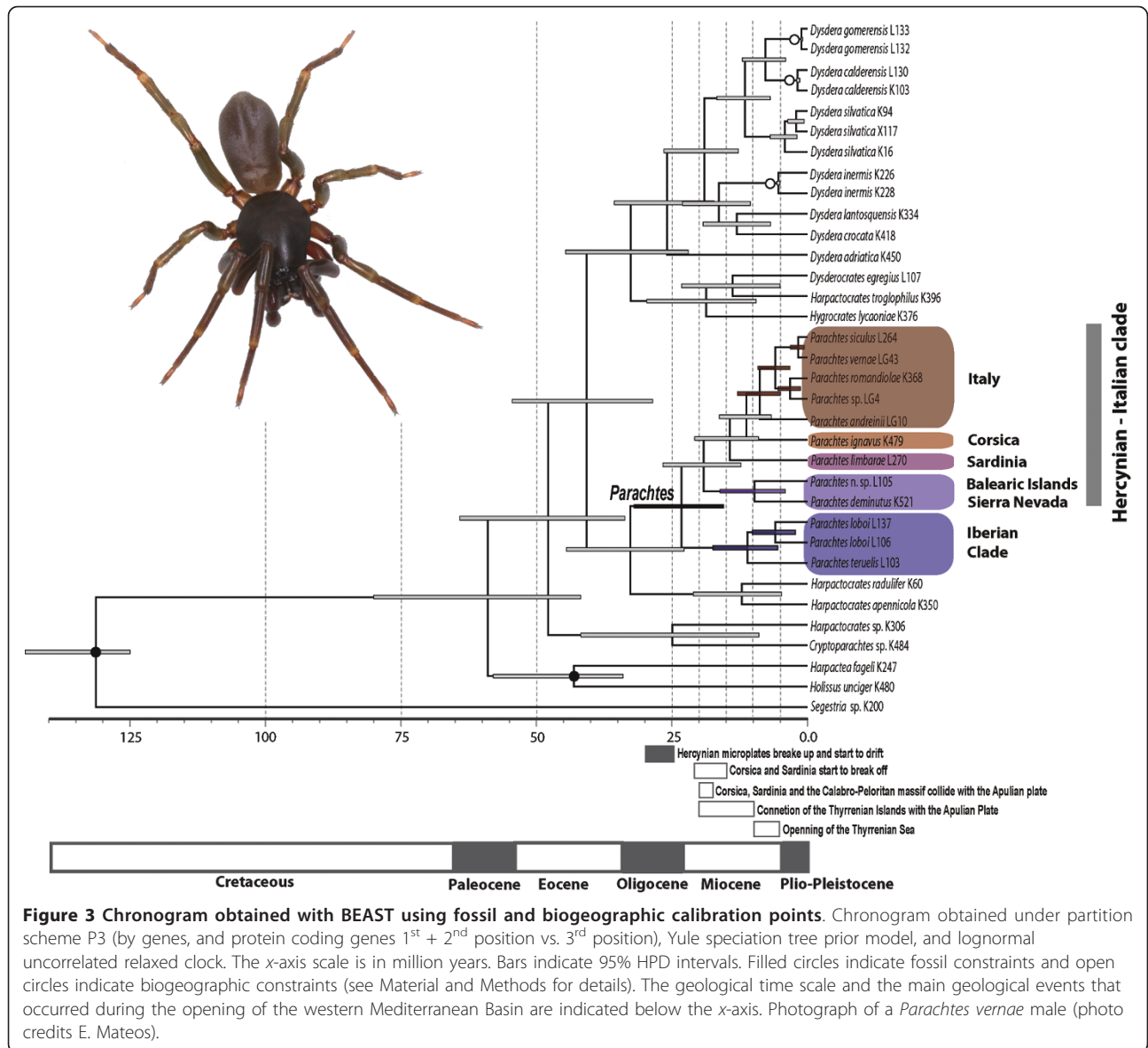
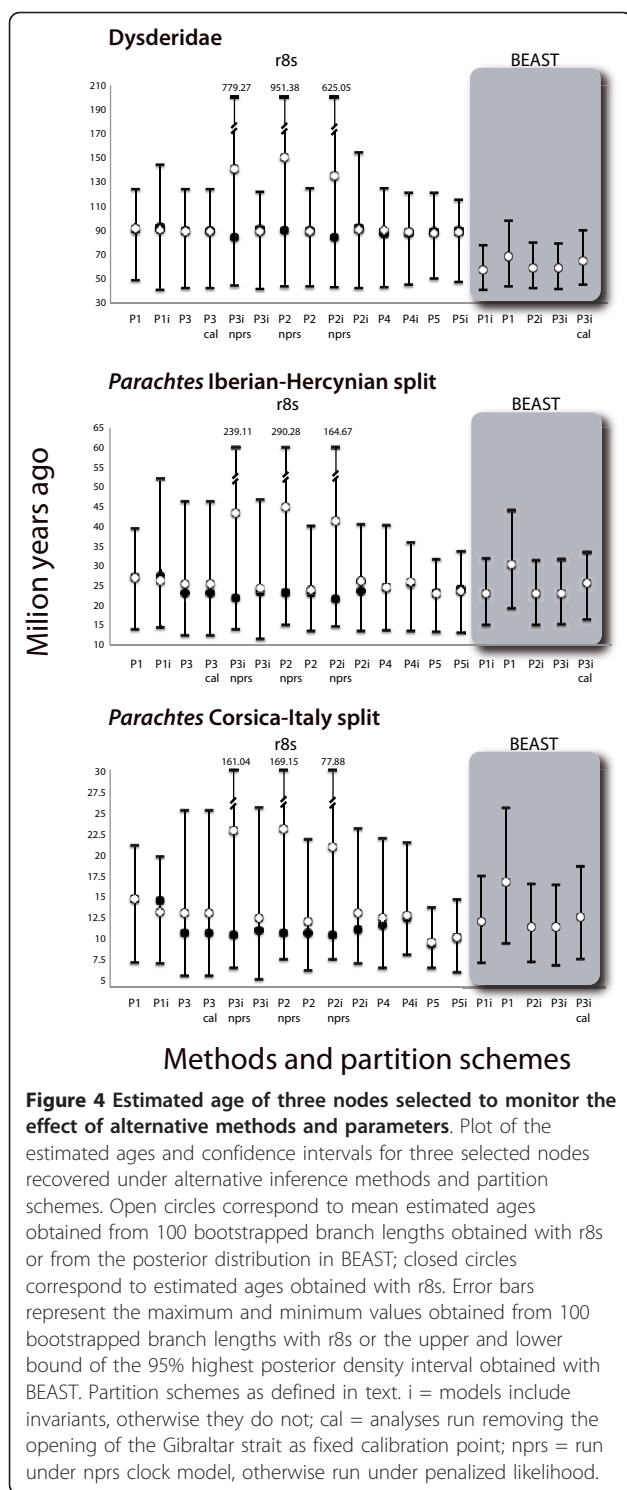


Figure 3 Chronogram obtained with BEAST using fossil and biogeographic calibration points. Chronogram obtained under partition scheme P3 (by genes, and protein coding genes 1st + 2nd position vs. 3rd position), Yule speciation tree prior model, and lognormal uncorrelated relaxed clock. The x-axis scale is in million years. Bars indicate 95% HPD intervals. Filled circles indicate fossil constraints and open circles indicate biogeographic constraints (see Material and Methods for details). The geological time scale and the main geological events that occurred during the opening of the western Mediterranean Basin are indicated below the x-axis. Photograph of a *Parachtes vernae* male (photo credits E. Mateos).

Table 1 Estimated substitution rates per gene obtained with BEAST based on fossil and biogeographic calibration points

Gene	Mean Rate	uclid.mean	uclid.stdev	Coefficient Variance	Covariance
<i>cox1</i>	0.0125 (0.0089-0.0165)	0.0199 (0.0136-0.0270)	0.6668 (0.4780-0.8728)	0.7087 (0.4823-0.9635)	0.0254 (-0.2254-0.2960)
<i>nad1</i>	0.0177 (0.0115-0.0240)	0.0258 (0.0163-0.0362)	0.6390 (0.4395-0.8792)	0.6400 (0.4088-0.8978)	0.0003 (-0.2474-0.2699)
<i>12S</i>	0.0087 (0.0057-0.0119)	0.0128 (0.0080-0.0181)	0.6950 (0.4080-1.0218)	0.7272 (0.3649-1.1102)	-0.0115 (-0.2637-0.2773)
<i>16-L1</i>	0.0063 (0.0046-0.0082)	0.0091 (0.0063-0.0124)	0.5731 (0.4032-0.7500)	0.5576 (0.3818-0.7309)	-0.0183 (-0.02596-0.2543)
<i>h3</i>	0.00108 (0.0006-0.0016)	0.0013 (0.0007-0.0019)	0.6450 (0.3241-1.0006)	0.6332 (0.2918-0.9781)	-0.0125 (-0.2216-0.2089)
<i>28S</i>	0.0006 (0.0003-0.0008)	0.0011 (0.0005-0.0019)	1.2490 (0.8895-1.6582)	1.4518 (0.9805-1.9531)	0.1892 (-0.0572-0.4451)

Estimations obtained under uncorrelated lognormal relaxed clock. Lower and upper bounds of the 95% HPD are indicated between brackets. Mean Rate: number of substitutions per site across the tree divided by the tree length. Uclid.mean: mean of the branches rates.

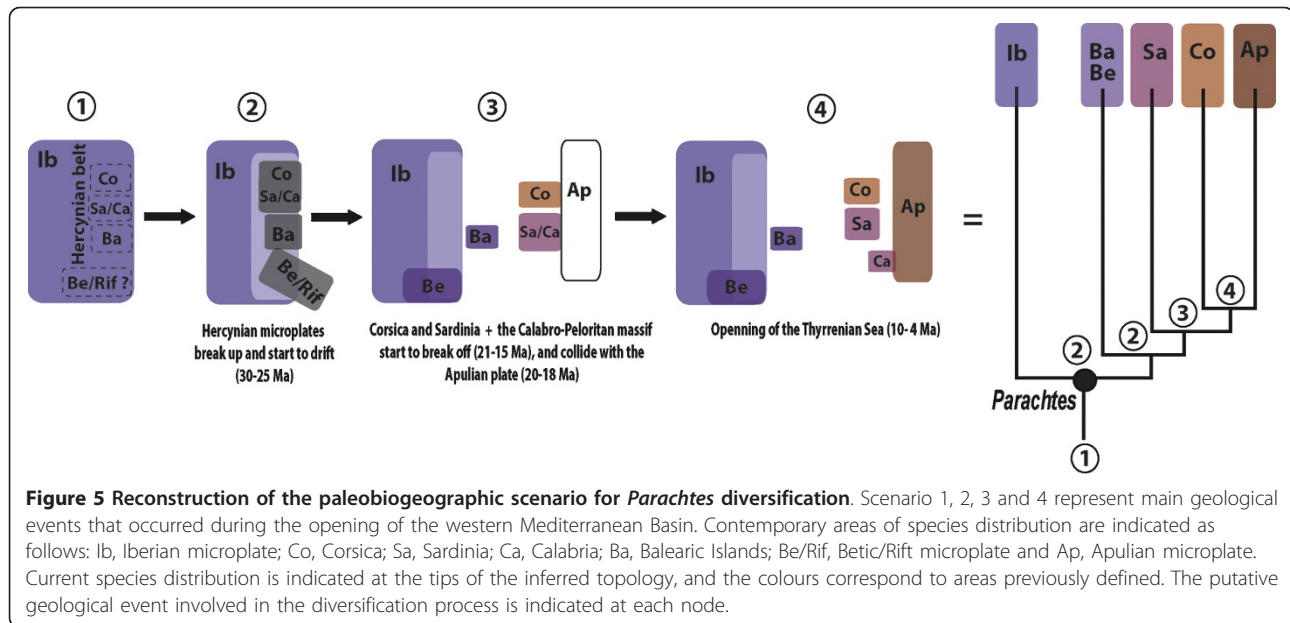


also investigated estimates based on alternative, widely-used smoothing methods. The results showed that time estimates of *Parachtes* diversification are consistent across methods and data partition schemes. The two time estimation methods, however, greatly differed at older time splits, suggesting that there could be

incorrect assumptions under certain analytical conditions. Smoothing methods differ from Bayesian methods, as implemented in BEAST, by the use of fixed topologies and roots and by assuming rate autocorrelation between ancestral and descendant branches [2-4,42]. To date, there has been a limited advance in evaluating the strengths and weaknesses of both methods, because direct comparisons between them are hampered by their numerous differences [3]. Finding out the causes of the reported discrepancies are beyond the scope of this study, but our results recommend exerting caution when trying to estimate ages close to the root.

The tree topology and estimated divergences identify the opening of the Western Mediterranean basin as the main driver for the diversification of *Parachtes* species (Figure 2, 3 and 5). The estimated age for the split of the *Parachtes* stem group closely matches the geochronology of the Hercynian belt break-off, and these node age estimates are robust to alternative analytical procedures and to partition schemes (Figure 4). Therefore, our results reject previous suggestions for a Quaternary origin of the genus [39], mostly based on the absence of these species from the Rift and Kabylies regions in northern Africa. Indeed, the absence of *Parachtes* in northern African exotic terrains poses a challenge to the Miocene origin hypothesis. It may be argued that biotic surveys of the region are far from complete, and that, hence, the presence of the species cannot be completely ruling out. Alternative explanations, such as species extinction due to unfavourable climatic conditions seem unlikely given that other humid fauna with northern Mediterranean affinities, such as the *Salamandra* newts [43] or *Alytes* toads [44], are still found in northern Africa. Evidence of absence is always hard to prove, and further exhaustive, systematic biotic surveys of north-western Africa will be necessary to discard the presence of *Parachtes* in the region.

Most current evidence indicates that Corsica, Sardinia, northern Sicily, the Balearic Islands, and the Kabylies microplates were part of a single block by the time of the onset of the back arc extension (~30-25 Ma), when the microplates broke and drifted off the Iberian Peninsula [10,12,14,15,28,17]. Nevertheless, the actual location of the Rift-Betic block is still controversial. It has been suggested that the internal part of the Rift-Betic belt was laid near the Balearic Hercynian microplates (see Figure seven of [15]). Some authors go further and propose that the Balearic microplate, together with the Betic belt, formed a Betic-Balearic domain, which suffered a WNW progressive thrust sheet stacking during the back arc extension, ending approximately at the time of the Middle-Miocene (16.4-11.2 Ma) [45,46]. The existence of a Betic-Balearic corridor during the Langhian-Serravallian marine regression (Middle Miocene,



c. 14.2 Ma) has been proposed to explain the origin of some Balearic mammalian fossil remains [47]. Alternatively, the Betic-Balearic clade could be the result of a later colonisation facilitated by the land bridge connections between the Balearic Islands and the continent established during the Messinian Salinity Crisis [21,48]. However, our much older time estimates disagree with this scenario.

The sister group relationship of the Sardinian lineage with the remaining Tyrrhenian and the Italian species together with the divergence time of these two lineages (14.34 Ma, 20.81-9 Ma) are consistent with the time frame of the separation of Corsica from Sardinia, dated at 21-15 Ma [12]. A similar phylogenetic pattern, i.e., Sardinia as a sister group of Corsica and Italy, has also been reported in the subterranean aquatic stenasellid isopods [20] and the terrestrial isopod *Helleria brevicornis* [32], although in the latter case time estimates suggest an Early Pleistocene split of Corsica and Sardinia.

The close relationships between Corsican and northern Italian species revealed in *Parachtes* have also been recovered in a diverse assemblage of arthropods, including cave crickets, wasps or isopods [20,49-51,32]. In most cases, these links originated as a result of active dispersal during the Pleistocene glacial cycles (2-0.5 Ma), when recurrent marine regressions led to the formation of land bridges between Corsica and Tuscany or when the Tuscany archipelago could have been used as stepping-stones. Our results, however, point towards a deeper split between the Corsican and Italian lineages (11.27, 16.26-6.68 Ma), which matches the onset of the opening of the Tyrrhenian Sea. About 10-9 Ma, Corsica, Sardinia and the Calabro-Peloritan massif separated

from the Apennines, after having progressively collided with the Apulian plate at around 20-18 Ma [12,28].

The basal split of Italian *P. andreinii* was estimated at about 9 Ma. Similar divergence times were found between northern and southern Italian species of the *Salmandrina* spectacled salamanders [52]. However, the finding of Miocene-age *Salmandrina* fossils in Sardinia led some authors to propose a double origin for the Italian *Salmandrina*: the ancestor of the northern species would have colonised the Peninsula from Corsica/Sardinia, while the southern one would have drifted eastward from Sardinia on the Calabro-Peloritan massif. The double-origin hypothesis put forward for *Salmandrina* does not match the topology inferred for *Parachtes*. Italian *Parachtes* are monophyletic, and thus a double colonisation scenario would require extinction events on both the islands and the continent. Indeed, the lack of direct links between the Sardinian and any Italian lineage seems to be at odds with geological expectations, given that Sardinia was connected to the Apennines until the opening of the Tyrrhenian Sea [28] and given the biogeographic links such as those reported in cave salamanders of the genus *Hydromantes* [53] and *Discoglossus* toads [21].

With the exception of *P. andreinii*, the estimated timeframe suggests that the diversification of *Parachtes* in the Italian peninsula was mostly driven by Plio-Pleistocene glacial cycles, as already predicted by Alicata [39]. The key role of geographic isolation due to Quaternary climatic oscillations seems to explain the origin of other components of the Italian fauna, such as frogs and snakes [54-58]. Further comprehensive phylogeographical studies on Italian and Iberian *Parachtes* species

are, however, needed to gain detailed insight into their evolutionary history and to fully expose the factors responsible for their diversification.

Recent studies have suggested that divergence time estimates under recently developed multilocus coalescent methods may provide younger time estimates than the standard concatenated approach conducted in the present study (see [59]). Although, this issue is still little explored in the literature, the amount of disagreement seems to be particularly significant at shallow divergence times. In fact, it has been argued that the younger estimates obtained with coalescent approaches may be the results of unaccounted recent gene flow [59], since most multilocus coalescent methods of species trees inference assume incomplete lineage sorting as the single source of gene tree incongruence (e.g., [60-62]).

Estimates of substitution rates

In the absence of fossil or biogeographic calibration points, the extrapolation of substitution rates from independent analyses provides the only information for inferring absolute timeframes. The confirmation that tectonic evolution of the Western Mediterranean shaped *Parachtes* diversification opens the possibility of using this model system to characterise substitution rates in molecular markers of interest.

The use of extrapolated substitution rates for one lineage and markers to date another may be compromised by the variation in rate of molecular evolution among lineages [63]. In the present study, the estimated average substitution rate for the mitochondrial genes estimated ($2.25\% \text{ My}^{-1}$) closely matches the “standard” arthropod mitochondrial DNA clock estimated at $2.3\% \text{ My}^{-1}$ [64], one of the most widely used universal rates. These results are surprising, given that the analytical procedures and assumptions used to estimate rates widely differ in the two studies. Nevertheless, Papadopoulou et al. [65] have reached similar conclusions by using another well-dated geological event, the formation of the Mid-Aegean Trench on the eastern Mediterranean, dated at 9-12 Ma, (from 2.39 to $2.69\% \text{ My}^{-1}$, 3.54% and 1.06% , for mtDNA, *cox1* and *16S*, respectively). These results may be viewed as a further support for the geological scenario proposed for *Parachtes* diversification.

The few spider rates available in the literature are significantly higher than the ones obtained here. Bond and collaborators [66] estimated a substitution rate for *16S* of $4\% \text{ My}^{-1}$ in the trapdoor spider *Aptostichus simus*, while phylogeographic studies in woodlouse hunter spiders endemic to the Canary Island have reported *cox1* rates as high as $9.8\% \text{ My}^{-1}$ [67] or $10.2\% \text{ My}^{-1}$ [68]. Interestingly, the former rates were derived from population-level comparisons, which have been shown to produce higher estimates of mutation rates as compared

to substitution rates inferred in phylogenetic (species-level) studies (Ho et al. [69]; however, see Emerson [7]). The slower estimates found in *cox1* ($1.75\% \text{ My}^{-1}$) in a species-level phylogeny of eastern Canarian *Dysdera* [70] supports this last contention.

Our study provides valuable information on the substitution rates of the handful of markers commonly used to infer phylogenies in non-model arthropods. Next-generation sequencing approaches promise to expand the range of molecular markers amenable to phylogenetic inference for organisms such as spiders. Until then, however, rigorous and thorough sampling of phylogeny of non-model organisms may rely on the set of genes used in the present study. In this context, our estimates could be an alternative to obtain absolute time estimates in the absence of calibration points when using such standard sets of molecular markers. Additionally, *Parachtes* may serve as a model system to gain further insights into the rate of molecular, morphological, and ecological trait evolution through time.

Conclusions

The pattern and timing of species formation in the spider genus *Parachtes* match the geochronological sequence of the opening of the western Mediterranean Basin. Therefore, *Parachtes* spiders provide one of the first and unequivocal pieces of evidence for endemic terrestrial taxa whose origins can be traced back to the separation of the major western Mediterranean islands from the continent (Figure 5). Our results confirm that early Oligocene tectonics played a key role in generating the diversity of the Earth's most populated biodiversity hot-spot [34,71].

Methods

Taxonomic sampling

A total of 49 specimens were analysed in the present study (Additional file 1). We included 9 of the 12 currently known *Parachtes* species, along with two new species awaiting formal description: one from Majorca and one from the Italian Peninsula (Figure 1 (B)). The missing species included the Iberian *P. cantabrorum*, the Corsican *P. inaequipes* and the Italian *P. latialis*. *P. cantabrorum* from an undetermined locality in the Pyrenees [72] has not been collected since its original description (some females assigned to *P. cantabrorum* by Denis are most likely a misidentification, [73]). *P. inaequipes* is known from a single locality in southern Corsica (Bonifacio). One of the specimens included in the present analysis is an unidentified juvenile collected near the *P. inaequipes* type locality (Porto Vecchio, 23 km). Finally, *P. latialis* is known from the Italian district of Lazio. After visiting this area, all collected specimens turned out to constitute a new species.

Half of the *Parachtes* species were represented by single individuals, two by specimens sampled from 2 populations, 2 from 3 populations and one, *P. teruelis*, from 6 populations (Figure 1(B)). We sampled 14 additional species from 5 genera of the same subfamily (Dysderinae) and two genera of the subfamily Harpacteinae, which is also in the family Dysderidae. The genus *Segestria*, a member of the family Segestriidae closely related to Dysderidae and part of the same superfamily Dysderoidea, was included as an outgroup to root trees. The length of the most basal branch in a rooted tree with molecular clock not enforced is trivial and therefore has to be removed before estimating the divergence times using smoothing methods (see below). The most recent common ancestor of Dysderoids, however, provided a calibration point and, therefore, a specimen belonging to Caponiidae, the putative sister-family to Dysderoidea [74], was included to root analyses and preserve the Dysderoid node.

Most specimens were collected in the field by the authors; some were kindly provided by colleagues. Specimens were preserved in 95% ethanol and stored at -20°C in the Department of Animal Biology of the University of Barcelona. Some specimens had been preserved in suboptimal 70% ethanol, but extraction yielded reasonable amounts of DNA for specimens collected within the past 10 years.

DNA extraction, PCR amplification and sequencing

Genomic DNA was extracted from specimens using the DNeasy Tissue Kit (Qiagen) following the manufacturer's guidelines. Partial fragments of the mitochondrial genes cytochrome c oxidase subunit I (*cox1*), the 12S rRNA (*12S*), a fragment spanning the 3' half of the 16S rRNA ribosomal subunit (*16S*), the complete tRNA leu (*L1*) and the 5' half of the NADH dehydrogenase subunit I (*nad1*), and the nuclear genes 28S rRNA (*28S*), 18S rRNA (*18S*) and Histone H3 (*h3*), were amplified using the following primer pairs: [**cox1**] C1-J-1490 [75] and C1-N-2776 [76], alternatively as two overlapping fragments using primer pairs C1-J-1490 or C1-J-1718 [77] with C1-N-2198 [75] or C1-N-2191 [77] and C1-J-2183 [77] with C1-N-2776; [**16S**, **L1**, **nad1**] LR-N-13398 [77] and N1-J-12350 (Crates: 5'-CCTARTTGRCTAR-ARTTRGCRSATCARCCAATTG-3') or N1-J-12373 [70], or as two overlapping fragments using primers pairs LR-N-13398 with LR-J-12864 [40] and LR-N-12945 [78] with either N1-J-12350 or N1-J-12373; [**12S**] 12SR-J-14199 and 12SR-N-14594 [79] or 12SR-J-14215 (Viera: 5'-AGGGTGACGGGCGATATG TGCAC-3') and 12SR-N-14522 (Forlan: 5'-AAATTATA-TACTTTGGCGGC-3'); [**28S**] 28S-B [80] and 28S-O [76], or as two overlapping fragments using primers pairs 28S-A with 28S-B [80] and 28S-O with 28S-C [76];

[**18S**] 5F and 9R [80]; [**h3**] H3a F and H3a R [81]. PCR conditions were as follows: 2 min. at 94°C followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 42-52°C for 35-45 s (depending on the primers, see below), and extension at 72°C for 30-60 s (depending on the length of the fragment), with a final single extension step at 72°C for 5 min. For the *cox1*, *16S-nad1* and *12S* gene fragments, a successful amplification was achieved with an annealing temperature of 42°C or 45°C for 45 s. For the *18S* and *28S*, a single annealing temperature of 52°C for 35 s was optimum. Finally, *h3* was amplified using a "touchdown" strategy consistent in beginning the annealing at 60°C for 35 s and lowering proportionally the temperature 1°C in each cycle (during 19 cycles), until reaching a constant temperature of 42°C, keeping this annealing temperature during the following 16 cycles. Amplifications were carried out in a 25 µl reaction volume for a final concentration of 1.25 U *Taq* polymerase (Promega), 2.5 mM MgCl₂ (Promega), 0.2 mM of each dNTP, 0.2 µM of each primer and about 2 µl of DNA sample and the amount of *Taq* buffer recommended by the manufacturer. PCR products were purified using MultiScreen PCR µ96 cleanup filter plates from Millipore. PCR products were cycle-sequenced in both directions using one of the PCR primers and the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem) and sequenced in an ABI 3700 automated sequencer at the Scientific and Technical Services of the University of Barcelona [82]. DNA sequences were edited using Geneious v.5.0.3 [83].

Phylogenetic Analyses

The alignment of the *cox1*, *nad1* and *h3* gene sequences was trivial because there was no evidence for insertion/deletion events and sequences were adjusted manually. Conversely, *12S*, *16S-L1*, *18S* and *28S* sequences showed length polymorphism, and gaps had to be included to retain positional homology. Sequence alignments of each variable length gene were constructed using the online version of the automatic alignment program MAFFT v. 6 [84,85]. The alignment was constructed using the manual strategy option set to Q-INS-i, with default options (Gap opening penalty (GOP) = 1.53 and Offset value, which works like a gap extension penalty (GEP) = 0.0). Two additional alignments were constructed to explore the sensitivity of the phylogenetic results to alternative alignment parameter values. The parameter options were set as follows: GOP = 3 and GEP = 0, to obtain a "compressed alignment", and GOP = 1 and GEP = 0.5, to obtain a "gappy alignment". Gaps were recoded as separate presence/absence characters following Simmons & Ochoterena [86]. Treating gaps in this manner allows the incorporation of gap information in the analyses while minimising the effect of increasing

the weight of overlapping multiple non-homologous gaps that result from scoring gaps as an additional state [87]. In addition, absence/presence gap scoring is amenable to Bayesian inference analyses. The program Gap-Coder [88] was used to facilitate the automatic recoding of the alignments based on the simple method proposed by Simmons et al. [89]. Gene matrices were concatenated using WINCLADA v.1.00.08 [90]. Non-sequenced fragments were scored as missing data.

Concatenation of genes evolving at different rates to resolve phylogenetic relationships among taxa across a wide range of divergence times has been the paradigm in phylogenetic inference for the last two decades. The use of concatenation is based upon the assumption that individual gene trees are congruent among themselves and to the species tree. However, it is well known that different genes may support incongruent topologies due to processes such as horizontal gene transfer, hybridization, gene duplication or incomplete lineage sorting [91-94]. This last problem is especially pervasive when dealing with closely related species, and has recently received a great deal of theoretical and methodological attention. A whole new generation of multispecies coalescent inference methods have been developed to deal with the effects of the stochasticity of the genealogical process. In the present study, however, we have decided to use a concatenation approach based on two main considerations: (1) we did not observe any instance of topological incongruence among individual gene trees, which lead us to assume that none processes know to cause gene incongruence had a major impact on our data, and (2) that our time window of interest (Oligocene/Miocene divergences) minimized the effect of coalescent stochastic errors (i.e. incomplete lineage sorting) in the sampled markers. Moreover, from a practical standpoint, the use of multiple coalescent approaches to infer species trees in the present study would have resulted in losing power to infer topology and branch lengths, given the taxonomic sampling (some species were represented by single individuals), and the low variability of the nuclear markers employed (slow evolving *18S*, *28S* and *h3* exon).

Parsimony analyses of the individual genes and concatenated matrices were conducted with the program TNT v. 1.0 [95]. Each heuristic search consisted of 1000 iterations of Wagner trees constructed with the random addition of taxa and subsequent TBR branch swapping, holding five trees per iteration and culminating in a final round of branch swapping holding up to 10000 trees. When the number of replicates finding optimal trees was less than 10%, the number of replicates was increased by 1000. Clade support was assessed via jack-knife resampling [96] using 1000 replicates with individual heuristic searches consisting of 20 iterations of

Wagner tree construction using the random addition of taxa, holding 5 trees per iteration up to 10000 trees.

Bayesian inference analyses were conducted with MRBAYES v.3.1.2 [97] and were run remotely at the Bioportal computer resources of the University of Oslo [98]. The combined matrix was analysed under three scheme partitions: by gene (P1), by gene and mitochondrial protein coding genes 1st + 2nd positions vs. 3rd positions (P2), and by gene and protein coding genes (both mitochondrial and nuclear) 1st + 2nd positions vs. 3rd positions (P3). Nucleotide substitution models selected by the Akaike information criterion (AIC) [99], as implemented in jMODELTEST v.0.1.1 [100,101], were specified for each partition, and a standard discrete model was defined for the gaps scored as absence/presence data [102]. The substitution parameters were allowed to vary independently between each partition. Two independent runs of 10 million generations, sampling each 1000 generations, with six simultaneous MCMC (Markov Chain Monte Carlo) chains, each starting from random trees, were carried out simultaneously. The program TRACER v. 1.5 [103] was used to ensure that the Markov chains had reached stationarity by examining the effective sample size (ESS) values and also to determine the correct number of generations to discard as a *burn-in* for the analysis (first 10%). Chain convergence was monitored by ensuring the standard deviation of the split frequencies of these two runs dropped below 0.01.

Maximum likelihood analyses of the concatenated data matrix were conducted with the software program RAxML v. 7.0.4 [104] and run remotely at the CIPRES portal [105]. The same three partition schemes described above were analysed with independent GTR +G substitution models for each partition. The best likelihood tree was selected out of 100 iterations of the random addition of taxa. Non-parametric bootstrap support values were drawn from 100 resampled matrices. Finally, confidence values were mapped onto the best topology.

Trees were visualised and manipulated with the program FigTree v. 1.1.2 [106].

The Approximately Unbiased (AU) topology test [107] implemented in the computer program CONSEL v.0.1i [108] was used to investigate whether alternative topologies could be statistically distinguished (see Results).

Estimation of divergence times and substitution rates

The use of multiple calibration points results in better and more reliable estimations, while minimising associated uncertainty [109,110]. Absolute ages were thus estimated by incorporating both fossil and biogeographic calibration points, which provided minimum and fixed or maximum ages, respectively. The oldest fossil of the family Segestriidae has been found in Lebanese amber

from the Lower Cretaceous [111]. Therefore, the minimum time of separation for the families Segestriidae and Dysderidae was set at 125 Mya. The Harpactinae fossil genus *Dasumiana* Wunderlich, 2004 from Eocene Baltic amber [111] shows close morphological affinities in the male bulb with the present day genus *Holissus* (Arnedo, pers. obs.). This information was incorporated into the analysis by placing the minimum age estimate of the common ancestor of *Harpactea* and *Holissus* at 35 Mya. The closely-related genus *Dysdera* provided several biogeographic constraints. The split of the Iberian and Moroccan populations of *Dysdera inermis* was assumed to have been caused by the opening of the Strait of Gibraltar and therefore their time of divergence was set at 5.3 Mya. The time of a volcanic island formation provides a maximum age for the lineages inhabiting the island [112]. *Dysdera* has undergone local diversification in the volcanic archipelago of the Canary Islands, and inter-island populations of two endemic Canarian *Dysdera* species were used as additional biogeographic calibration points. The age of emergence of La Palma (2 Mya) and El Hierro (1.2 Mya) [113] placed maximum age estimates for the divergence time between the populations of La Palma and La Gomera of the species *D. calderensis* and for the divergence of the populations of El Hierro and La Gomera of the species *D. gomerensis*.

Minimum and maximum constraints were included as uniform prior distributions to estimate absolute time in a Bayesian framework (see below). Uniform prior distributions for fossil calibration were preferred over more explicit distributions (e.g., lognormal) as a conservative approach, given the limited information available to decide upon the shape of the distribution [114]. For island calibration points, upper bounds were set to the island age and lower bounds to zero. For fossils, the lower bound was set to the latest fossil age and the upper bound was set to 392 Mya, which is the age of the oldest Uraraneida, the putative sister taxon to spiders [115]. The opening of the Gibraltar strait was set as a fixed calibration point following a normal distribution, with mean = 5.3 and sd = 0.15. The effect of using a fixed constraint on time estimates was explored by running additional analyses without this constraint.

Lineage ages were estimated using a Bayesian framework as implemented in BEAST v1.5.4 [116,103] and using multiple rate methods as implemented in r8s v. 1.71 [117]. Species were represented by single specimens (Additional file 1) to avoid very short or zero-length branches, which may negatively influence the performance of the algorithms in r8s [117], and to ensure that only speciation processes as tree prior were included in BEAST estimations.

A preliminary cross-validation analysis was conducted to select the best clock method and, if required, the best

smoothing parameter value [118] in r8s analyses. Analyses were conducted using the node constraints described above. Branch lengths were re-estimated with RAxML after taxon removal, enforcing the preferred topology under the alternative partitions schemes described above (P1, P2 and P3) and two additional schemes: gene partition with 3rd codon positions removed (P4) and gene partitions with mitochondrial 3rd codon positions removed (P5). Some concerns have been cast on the simultaneous use of invariants and gamma distributions in evolutionary models [119]. Therefore, branch lengths were reestimated under models with and without invariants for each partition model. Confidence intervals for the time estimates were constructed by generating 100 trees with identical topology by bootstrapping branch lengths, using RAxML.

We used an iterative strategy to select the best clock, partition and speciation model for conducting analyses in BEAST. Alternative schemes were compared using Bayes Factors, as calculated by the program TRACER v1.5. First, lineage ages were estimated under the strict and the two relaxed uncorrelated clocks (exponential and lognormal) using the partition scheme P1 (see above) and selecting the Yule speciation process as tree prior. Once the best clock model was identified, lineage ages were reestimated using the best clock and partition scheme P1, but selecting the Birth-death speciation process as a tree prior. Finally, lineage ages were estimated again by selecting the best clock and speciation process prior, under the three alternative partition schemes (P1, P2, and P3, see above). The best substitution model for each partition was assessed by AIC as implemented in jMODELTEST. An additional run was conducted without considering invariants in the nucleotide substitution models (P1_{NOINV}) under the partition scheme P1. Substitution rates were estimated for each gene under the uncorrelated lognormal relaxed clock, the Yule speciation process and the fixed topology in BEAST.

The ultrametric trees estimated with r8s were used as starting trees in BEAST analyses to ensure that time constraints were not violated. Two independent runs of 100 million generations, sampling every 10000 generations, were performed for each analysis. The convergence and mixing of each MCMC chain was assessed with TRACER. Both independent runs of each analysis were combined with LogCombiner after a 10% burn-in, and TreeAnnotator was used to summarise the information from the sampled trees.

Additional material

Additional file 1: Specimens and gene sequence information.

Specimens and sequence accession numbers included in the study with details of sample codes, voucher number, sex (f: female, m: male, juv.:

juvenile), and collection locality. Species names with asterisks were included in lineage age estimation analyses. † sequences obtained from different specimens to complete gene sampling nr: near.

Additional file 2: Length of the non-protein coding genes under different alignments. Default, compressed and gappy alignment lengths are indicated for each gene, along with numbers of gaps and informative gaps codified as absence/presence.

Additional file 3: Substitution models per gene and partitions selected by AIC. Nucleotide substitution models selected by AIC for a 48 taxon matrix used for phylogenetic analyses, and a 34 taxon matrix under different partitions used for lineage age estimations (see Material and Methods for details).

Additional file 4: Bayes factor comparisons for the selection of the clock model, the speciation process and the partition scheme. a) Bayes factor comparisons for selection between Yule and Birth-death speciation models under partition scheme P1. b) Bayes factor comparisons for selection among partition schemes P1, P2, P3 and P1_{NOINV} (see Material and Methods for details). HME: Harmonic mean of the likelihood from the posterior distribution value.

Additional file 5: Estimated age of three nodes selected to monitor the effect of alternative methods and parameters. Ages of three selected nodes under different evolutionary models and partition schemes as estimated by R8S, using the best method selected by cross-validation, and BEAST. In R8S analyses, branch lengths were first estimated with RAxML under a fixed topology in Figure 2 and GTR+I models, with (INV) or without invariants (NOINV). In BEAST, the best model was selected by jMODELTEST and, for P1 also without invariants (P1_{NOINV}). Partition schemes: by gene (P1, 4412 chars), by gene and mitochondrial protein coding genes 1st + 2nd positions vs. 3rd positions (P2, 4412 chars), by gene and protein coding genes (both mitochondrial and nuclear) 1st + 2nd positions vs. 3rd positions (P3, 4412 chars), by gene with 3rd codon positions removed (P4, 3762 chars) and by gene with mitochondrial 3rd codon positions removed (P5, 3871 chars). NPRS (Non Parametric Rate Smoothing), PL (Penalized Likelihood), log (logarithmic penalty function). S: Smoothing factor in PL. Age: estimated age. Confidence interval obtained by profiling across 100 trees of constrained topology and bootstrapped branch lengths, as estimated with RAxML: max (maximum value observed), min (minimum value observed), mean (average). 95%HPD: upper and lower bound of the 95% highest posterior density interval. †Analyses run removing the opening of the Gibraltar strait as fixed calibration point. *Preliminary cross validation analyses selected PL log as best fit but failed check. **Analyses under PL in partitions and models where cross-validation selected NPRS as best method, but obtained confidence intervals where absurdly large.

Acknowledgements

We thank H. López, N. Macías-Hernández, C. Hernando, E. de Mas, M. Mejía-Chang and V. Opatova, who assisted in the fieldwork. We would like to thank to S. Carranza, M.A. Ferrández, F. Gasparo, G. Giribet, K. Kunt, A. López, J. Pons, M. Rezac, C. Ribera, S. Taiti and N. Txasco for providing us with additional samples. We are grateful to J. Pons for assistance and support in some of the data analyses. We also thank W. Cavazza and G. Rosenbaum for their valuable comments that greatly improved the first versions of the manuscript. W. Wheeler made available sequence data collected by the ATOL Spider project (NSF grant EAR-0228699) for *Caponia* sp. This project was funded by the Spanish Ministry of Science and Innovation (MICINN) grant CGL2006-08617 (MA), and additional financial support was provided by project 2009SGR1462 from the Generalitat de Catalunya. LBB was supported by a graduate grant (FI-DGR 2009) from the Generalitat de Catalunya. MA was further supported by an ICREA Academia award for excellence in research from the Generalitat de Catalunya.

Authors' contributions

MA conceived the idea of the study, and MA and LBB designed the study. Specimens were collected by MA and LBB or kindly provided by colleagues. LBB did the molecular work. LBB and MA conducted the analyses. LBB wrote

the first draft, and MA and LBB improved successive versions. Both authors read and approved the final manuscript.

Received: 15 July 2011 Accepted: 31 October 2011

Published: 31 October 2011

References

1. Zuckerkandl E, Pauling L: **Evolutionary divergence and convergence in proteins.** In *Evolving genes and proteins*. Edited by: Bryoson V, Vogel HJ. New York: Academic Press; 1965:97-166.
2. Renner SS: **Relaxed molecular clocks for dating historical plant dispersal events.** *Trends Plant Sci* 2005, **10**:550-558.
3. Welch JJ, Bromham L: **Molecular dating when rates vary.** *Trends Ecol Evol* 2005, **20**:320-327.
4. Rutschmann F: **Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times.** *Divers Distrib* 2006, **12**:35-48.
5. Weir JT, Schluter D: **Calibrating the avian molecular clock.** *Mol Ecol* 2008, **17**:2321-2328.
6. Bromham L, Phillips MJ, Penny D: **Growing up with dinosaurs: molecular dates and the mammalian radiation.** *Trends Ecol Evol* 1999, **14**:113-117.
7. Emerson BC: **Alarm bells for the molecular clock? No support for Ho et al.'s model of time-dependent molecular rate estimates.** *Syst Biol* 2007, **56**:337-345.
8. Forest F: **Calibrating the Tree of Life: fossils, molecules and evolutionary timescales.** *Ann Bot* 2009, **104**:789.
9. Heads M: **Old taxa on young islands: a critique of the use of island age to date island-endemic clades and calibrate phylogenies.** *Syst Biol* 2011, **60**:204-218.
10. Gueguen E, Doglioni C, Fernandez M: **On the post-25 Ma geodynamic evolution of the western Mediterranean.** *Tectonophysics* 1998, **298**:259-269.
11. Roca E, Sans M, Cabrera L, Marzo M: **Oligocene to Middle Miocene evolution of the central Catalan margin (northwestern Mediterranean).** *Tectonophysics* 1999, **315**:209-229.
12. Rosenbaum G, Lister GS, Duboz C: **Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene.** *J Virtual Explorer* 2002, **8**:107-126.
13. Speranza F, Villa IM, Sagnotti L, Florindo F, Cosentino D, Cipollari P, Mattei M: **Age of the Corsica-Sardinia rotation and Liguro-Provençal Basin spreading: new paleomagnetic and Ar/Ar evidence.** *Tectonophysics* 2002, **347**:231-251.
14. Carminati E, Doglioni C: **Mediterranean tectonics.** In *Encyclopedia of Geology*. Edited by: Selley RC, Cocks LRM, Plimer IR. Amsterdam, The Netherlands: Elsevier; 2004:135-146.
15. Rosenbaum G, Lister GS: **Formation of arcuate orogenic belts in the western Mediterranean region.** In *Orogenic Curvature: Integrating Paleomagnetic and Structural Analyses*. Edited by: Sussman AJ, Weil AB. Boulder, Colorado: Geological Society of America Special Paper 383; 2004:41-56.
16. Jolivet L, Augier R, Robin C, Suc JP, Rouchy JM: **Lithospheric-scale geodynamic context of the Messinian salinity crisis.** *Sediment Geol* 2006, **188**:9-33.
17. Schettino A, Turco E: **Plate kinematics of the Western Mediterranean region during the Oligocene and Early Miocene.** *Geophys J Int* 2006, **166**:1398-1423.
18. Caccone A, Milinkovitch MC, Sbordoni V, Powell JR: **Molecular biogeography: using the Corsica-Sardinia microplate disjunction to calibrate mitochondrial rDNA evolutionary rates in mountain newts (*Euproctus*).** *J Evol Biol* 1994, **7**:227-245.
19. Caccone A, Sbordoni V: **Molecular biogeography of cave life: a study using mitochondrial DNA from Bathysciine beetles.** *Evolution* 2001, **55**:122-130.
20. Ketmaier V, Argano R, Caccone A: **Phylogeography and molecular rates of subterranean aquatic Stenaelid Isopods with a peri-Tyrrhenian distribution.** *Mol Ecol* 2003, **12**:547-555.
21. Fromhage L, Vences M, Veith M: **Testing alternative vicariance scenarios in Western Mediterranean discoglossid frogs.** *Mol Phylogeny Evol* 2004, **31**:308-322.
22. Fochetti R, Sezzi E, de Figueroa JMT, Modica MV, Oliverio M: **Molecular systematics and biogeography of the western Mediterranean stonefly**

- genus *Tyrrhenoleuctra* (Insecta, Plecoptera). *J Zool Syst Evol Res* 2009, **47**:328-336.
23. Pfenninger M, Véla E, Jesse R, Elejalde MA, Liberto F, Magnin F, Martínez-Ortí A: Temporal speciation pattern in the western Mediterranean genus *Tudorella* P. Fischer, 1885 (Gastropoda, Pomatiidae) supports the Tyrrhenian vicariance hypothesis. *Mol Phylogeny Evol* 2010, **54**:427-436.
 24. Ribera I, Fresneda J, Bucur R, Izquierdo A, Vogler AP, Salgado JM, Cieslak A: Ancient origin of a Western Mediterranean radiation of subterranean beetles. *BMC Evol Biol* 2010, **10**:29.
 25. Magri D, Fineschi S, Bellarosa R, Buonamici A, Sebastiani F, Schirone B, Simeone MC, Vendramin GG: The distribution of *Quercus suber* chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Mol Ecol* 2007, **16**:5259-5266.
 26. Gattacceca J, Deino A, Rizzo R, Jones DS, Henry B, Beauvoisin B, Vadeboin F: Miocene rotation of Sardinia: New paleomagnetic and geochronological constraints and geodynamic implications. *Earth Planet Sci Lett* 2007, **258**:359-377.
 27. Alvarez W: Fragmentation of the Alpine orogenic belt by microplate dispersal. *Nature* 1974, **248**:309-314.
 28. Rosenbaum G, Lister GS: Neogene and Quaternary rollback evolution of the Tyrrhenian Sea, the Apennines, and the Sicilian Maghrebides. *Tectonics* 2004, **23**:1-17.
 29. Carranza S, Amat F: Taxonomy, biogeography and evolution of *Euproctus* (Amphibia: Salamandridae), with the resurrection of the genus *Calotriton* and the description of a new endemic species from the Iberian Peninsula. *Zool J Linn Soc* 2005, **145**:555-582.
 30. Ketmaier V, Giusti F, Caccione A: Molecular phylogeny and historical biogeography of the land snail genus *Solatopupa* (Pulmonata) in the peri-Tyrrhenian area. *Mol Phylogeny Evol* 2006, **39**:439-451.
 31. Zangari F, Cimmaruta R, Nascetti G: Genetic relationships of the western Mediterranean painted frogs based on allozymes and mitochondrial markers: evolutionary and taxonomic inferences (Amphibia, Anura, Discoglossidae). *Biol J Linn Soc* 2006, **87**:515-536.
 32. Gentile G, Campanaro A, Carosi M, Sbordoni V, Argano R: Phylogeography of *Helleria brevicornis* Ebner 1868 (Crustacea, Oniscidea): Old and recent differentiations of an ancient lineage. *Mol Phylogeny Evol* 2010, **54**:640-646.
 33. Salvo G, Ho SYW, Rosenbaum G, Ree R, Conti E: Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the citrus family (*Ruta* L., Rutaceae). *Syst Biol* 2010, **59**:705-722.
 34. Blondel J, Aronson J: *Biology and wildlife of the Mediterranean region* Oxford, New York: Oxford University Press; 1999.
 35. Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS: Chronology, causes and progression of the Messinian salinity crisis. *Nature* 1999, **400**:652-655.
 36. Loget N, Driessche JVD, Davy P: How did the Messinian Salinity Crisis end? *Terra Nova* 2005, **17**:414-419.
 37. Loget N, Van Den Driessche J: On the origin of the Strait of Gibraltar. *Sediment Geol* 2006, **188**:341-356.
 38. Mansion G, Rosenbaum G, Schoenenberger N, Bacchetta G, Rossello JA, Conti E: Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean Basin by the angiosperm family Araceae. *Syst Biol* 2008, **57**:269-285.
 39. Alicata P: Le specie italiane di *Harpactocrates* e di *Parachtes* n. gen. (Araneae, Dysderidae). *Annuaire Ist Mus Zool Univ Napoli* 1964, **16**:1-40.
 40. Arnedo MA, Oromí P, Ribera C: Radiation of the spider genus *Dysdera* (Araneae, Dysderidae) in the Canary Islands: Cladistic assessment based on multiple data sets. *Cladistics* 2001, **17**:313-353.
 41. Deeleman-Reinhold CL, Deeleman PR: Revision des Dysderinae. *Tijdschr Entomol* 1988, **131**:141-269.
 42. San Mauro D, Agorreta A: Molecular systematics: a synthesis of the common methods and the state of knowledge. *Cell Mol Biol Lett* 2010, **15**:311-341.
 43. Escoriza D, Comas MM, Donaire D, Carranza S: Rediscovery of *Salamandra algira* Bedriaga, 1833 from the Beni Snassen massif (Morocco) and phylogenetic relationships of North African *Salamandra*. *Amphib-Reptilia* 2006, **27**:448-455.
 44. Martínez-Solano I, Gonçalves HA, Arntzen JW, García-París M: Phylogenetic relationships and biogeography of midwife toads (Discoglossidae *Alytes*). *J Biogeogr* 2004, **31**:603-618.
 45. Fontboté JM, Guimerà J, Roca E, Sàbat F, Santanach P, Fernández-Ortigosa F: The Cenozoic geodynamic evolution of the Valencia trough (western Mediterranean). *Rev Soc Geol España* 1990, **3**:249-259.
 46. Roca E: La evolución geodinámica de la Cuenca Catalano-Balear y áreas adyacentes desde el Mesozoico hasta la actualidad. *Acta Geol Hisp* 1994, **29**:3-25.
 47. Bailon S, Bover P, Quintana J, Alcover JA: First fossil record of *Vipera Laurenti* 1768. *Comptes Rendus Palevol* 2010, **9**:147-154.
 48. Bover P, Quintana J, Alcover JA: Three islands, three worlds: paleogeography and evolution of the vertebrate fauna from the Balearic Islands. *Quatern Int* 2008, **182**:135-144.
 49. Allegrucci G, Todisco V, Sbordoni V: Molecular phylogeography of *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae): A scenario suggested by mitochondrial DNA. *Mol Phylogeny Evol* 2005, **37**:153-164.
 50. Dapporto L, Wolf H, Strumia F: Recent geography determines the distribution of some flying Hymenoptera in the Tuscan Archipelago. *J Zool* 2007, **272**:37-44.
 51. Martinsen L, Venanzetti F, Bachmann L: Phylogeography and mitochondrial DNA divergence in *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae). *Hereditas* 2009, **146**:33-45.
 52. Nascetti G, Zangari F, Canestrelli D: The spectacled salamanders, *Salamandrina terdigitata* (Lacépède, 1788) and *S. perspicillata* (Savi, 1821): 1) genetic differentiation and evolutionary history. *Rend Fis Acc Lincei* 2005, **16**:159-169.
 53. Carranza S, Romano A, Arnold EN, Sotgiu G: Biogeography and evolution of European cave salamanders, *Hydromantes* (Urodela: Plethodontidae), inferred from mtDNA sequences. *J Biogeogr* 2007, **35**:724-738.
 54. Canestrelli D, Cimmaruta R, Nascetti G: Phylogeography and historical demography of the Italian treefrog, *Hyla intermedia*, reveals multiple refugia, population expansions and secondary contacts within peninsular Italy. *Mol Ecol* 2007, **16**:4808-4821.
 55. Canestrelli D, Verardi A, Nascetti G: Genetic differentiation and history of populations of the Italian treefrog *Hyla intermedia*: lack of concordance between mitochondrial and nuclear markers. *Genetica* 2007, **130**:241-255.
 56. Canestrelli D, Cimmaruta R, Nascetti G: Population genetic structure and diversity of the Apennine endemic stream frog, *Rana italica*-insights on the Pleistocene evolutionary history of the Italian peninsular biota. *Mol Ecol* 2008, **17**:3856-3872.
 57. Canestrelli D, Nascetti G: Phylogeography of the pool frog *Rana (Pelophylax) lessonae* in the Italian peninsula and Sicily: multiple refugia, glacial expansions and nuclear-mitochondrial discordance. *J Biogeogr* 2008, **19**:23-1936.
 58. Barbanera F, Zuffi MAL, Guerrini M, Gentili A, Tofanelli S, Fasola M, Dini F: Molecular phylogeography of the asp viper *Vipera aspis* (Linnaeus, 1758) in Italy: evidence for introgressive hybridization and mitochondrial DNA capture. *Mol Phylogeny Evol* 2009, **52**:103-114.
 59. McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL: Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. *Evolution* 2010, **65**:184-202.
 60. Liu L: BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 2008, **24**:2542-2543.
 61. Kubatko LS, Carstens BC, Knowles LL: STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 2009, **25**:971-973.
 62. Heled J, Drummond AJ: Bayesian inference of species trees from multilocus data. *Mol Biol Evol* 2010, **27**:570-580.
 63. Bromham L, Penny D, Rambaut A, Hendy MD: The power of relative rates tests depends on the data. *J Mol Evol* 2000, **50**:296-301.
 64. Brower AVZ: Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci USA* 1994, **91**:6491-6495.
 65. Papadopoulou A, Anastasiou I, Vogler AP: Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Mol Biol Evol* 2010, **27**:1659-1672.
 66. Bond JE, Hedin MC, Ramirez MG, Opell BD: Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Mol Ecol* 2001, **10**:899-910.
 67. Macías-Hernández N, Oromí P, Arnedo MA: Integrative taxonomy uncovers hidden species diversity in woodland hunter spiders (Araneae, Dysderidae) endemic to the Macaronesian archipelagos. *Syst Biodivers* 2010, **8**:531-553.
 68. Bidegaray-Batista L, Macías-Hernández N, Oromí P, Arnedo MA: Living on the edge: demographic and phylogeographical patterns in the

- woodlouse-hunter spider *Dysdera lancerotensis* Simon, 1907 on the eastern volcanic ridge of the Canary Islands. *Mol Ecol* 2007, **16**:3198-3214.
69. Ho SYW, Phillips MJ, Cooper A, Drummond AJ: Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol Biol Evol* 2005, **22**:1561-1568.
70. Macías-Hernández N, Oromí P, Arnedo MA: Patterns of diversification on old volcanic islands as revealed by the woodlouse hunter spider genus *Dysdera* (Araneae, Dysderidae) in the eastern Canary Islands. *Biol J Linn Soc* 2008, **94**:589-615.
71. Myers N, Mittermeier RA, Mittermeier CC, da Fonseca GA, Kent J: Biodiversity hotspots for conservation priorities. *Nature* 2000, **403**:853-858.
72. Simon E: Synopsis général et catalogue des espèces françaises de l'ordre des Araneae: Ire partie. *Les Arachnides de France Paris* 1914, 6:1-308.
73. Ferrández MA: Notes on the Iberian Dysderidae, IV. New records, cartography and biology of *R. scutiventris* Simon, 1882 (Araneae). *Bol Asoc Esp Entomol* 1990, **14**:33-40.
74. Ramírez MJ: Respiratory system morphology and the phylogeny of haplogyne spiders (Araneae, Araneomorphae). *J Arachnol* 2000, **28**:149-157.
75. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R: DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 1994, **3**:294-299.
76. Hedin MC, Maddison WP: A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). *Mol Phylogeny Evol* 2001, **18**:386-403.
77. Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P: Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am* 1994, **87**:651-701.
78. Hedin MC: Speciation history in a diverse clade of habitat-specialized spiders (Araneae: Nesticidae: *Nesticus*): Inferences from geographic-based sampling. *Evolution* 1997, **51**:1929-1945.
79. Kambhampati S, Smith PT: PCR primers for the amplification of four insect mitochondrial gene fragments. *Insect Mol Biol* 1995, **4**:233-236.
80. Giribet G, Rambla M, Carranza S, Baguna J, Riutort M, Ribera C: Phylogeny of the arachnid order Opiliones (Arthropoda) inferred from a combined approach of complete 18S and partial 28S ribosomal DNA sequences and morphology. *Mol Phylogeny Evol* 1999, **11**:296-307.
81. Colgan DJ, McLauchlan A, Wilson GDF, Livingstone SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR: Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Aust J Zool* 1998, **46**:419-437.
82. Serveis Científicotècnics de la Universitat de Barcelona. [http://www.sct.ub.es].
83. Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A: Geneious v5.0.3. 2010 [http://www.geneious.com].
84. MAFFT version 6. [http://mafft.cbrc.jp/alignment/server/].
85. Katoh K, Toh H: Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 2008, **9**:286-298.
86. Simmons MP, Ochoterena H: Gaps as characters in sequence-based phylogenetic analyses. *Syst Biol* 2000, **49**:369-381.
87. Pons J, Vogler AP: Size, frequency, and phylogenetic signal of multiple-residue indels in sequence alignment of introns. *Cladistics* 2006, **22**:144-156.
88. Young N, Healy J: GapCoder automates the use of indel characters in phylogenetic analysis. *BMC Bioinformatics* 2003, **4**:6.
89. Simmons MP, Ochoterena H, Carr TG: Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Syst Biol* 2001, **50**:454-462.
90. Nixon KC: WinClada. Ithaca, New York: Published by the Author, v. 1.00.08 2002.
91. Maddison WP: Gene trees in species trees. *Syst Biol* 1997, **46**:523-536.
92. Degnan JH, Rosenberg NA: Discordance of species trees with their most likely gene trees. *PLoS Genet* 2006, **2**:762-768.
93. Kubatko LS, Degnan JH: Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst Biol* 2007, **56**:17-24.
94. Edwards SV: Is a new and general theory of molecular systematics emerging? *Evolution* 2009, **63**:1-19.
95. Goloboff PA, Farris JS, Nixon KC: *TNT: Tree analysis using new technologies*, v. 0.6 (Beta) 2003.
96. Farris JS, Albert VA, Källersjö M, Lipscomb D, Kluge AG: Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 1996, **12**:99-124.
97. Ronquist F, Huelsenbeck JP: MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003, **19**:1572-1574.
98. Bioportal. [http://www.bioportal.uio.no/].
99. Akaike H: Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika* 1973, **60**:255-265.
100. Guindon S, Gascuel O: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003, **52**:696-704.
101. Posada D: jModelTest: phylogenetic model averaging. *Mol Biol Evol* 2008, **25**:1253-1256.
102. Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL: Bayesian phylogenetic analysis of combined data. *Syst Biol* 2004, **53**:47-67.
103. Drummond AJ, Rambaut A: BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 2007, **7**:214.
104. Stamatakis A: RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006, **22**:2688-2690.
105. Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T: The CIPRES Portals. 2009 [http://www.phylo.org/sub_sections/portal].
106. Molecular evolution, phylogenetics and epidemiology. [http://tree.bio.ed.ac.uk/software/figtree/].
107. Shimodaira H: An approximately unbiased test of phylogenetic tree selection. *Syst Biol* 2002, **51**:492-508.
108. Shimodaira H, Hasegawa M: CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 2001, **17**:1246-1247.
109. Blair Hedges S, Kumar S: Precision of molecular time estimates. *Trends Genet* 2004, **20**:242-247.
110. Donoghue PCJ, Benton MJ: Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends Ecol Evol* 2007, **22**:424-431.
111. Wunderlich J, Milki R: Description of the extinct new subfamily Microsegestriinae (Araneae: Segestriidae) in Cretaceous Lebanese amber. *Beitr Araneologie* 2004, **3**:1867-1873.
112. Fleischer RC, McIntosh CE, Tarr CL: Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Mol Ecol* 1998, **7**:533-545.
113. Carracedo JC, Day S: *Canary Islands* Harpenden: Terra; 2002.
114. Ho SYM: Calibrating molecular estimates of substitution rates and divergence times in birds. *J Avian Biol* 2007, **38**:409-414.
115. Penney D, Selden PA: *Fossil spiders: the evolutionary history of a mega-diverse order* Manchester: Siri Scientific Press; 2011.
116. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A: Relaxed phylogenetics and dating with confidence. *PLoS Biol* 2006, **4**:699-710.
117. Sanderson MJ: r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 2003, **19**:301-302.
118. Sanderson MJ: Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol Biol Evol* 2002, **19**:101-109.
119. Yang Z: *Computational molecular evolution USA*: Oxford University Press; 2006.

doi:10.1186/1471-2148-11-317

Cite this article as: Bidegaray-Batista and Arnedo: Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology* 2011 **11**:317.