Diversity and evolutionary history of mygalomorph spiders in the Western Mediterranean and the Canary Islands

La diversidad e historia evolutiva de las arañas migalomorfas en el Mediterráneo Occidental y las Islas Canarias

Vera Opatova

Diversity and evolutionary history of mygalomorph spiders in the Western Mediterranean and the Canary Islands



Vera Opatova

Doctoral Thesis

Universitat de Barcelona

DIVERSITY AND EVOLUTIONARY HISTORY OF MYGALOMORPH SPIDERS IN THE WESTERN MEDITERRANEAN AND THE CANARY ISLANDS

DIVERSIDAD E HISTORIA EVOLUTIVA DE LAS ARAÑAS MIGALOMORFAS EN EL MEDITERRÁNEO OCCIDENTAL Y LAS ISLAS CANARIAS

Vera Opatova

Doctoral Thesis

2014

Cover
Right to left: *Cteniza* sp. burrow, *Ummidia* sp. burrow, *Ummidia* sp. adult female
Photo credit and design: Vera Opatova (VO)

Tesis Doctoral

Facultad de Biología - Departamento de Biología Animal
Programa de Doctorado : en Biodiversidad

DIVERSITY AND EVOLUTIONARY HISTORY OF MYGALOMORPH SPIDERS IN THE WESTERN MEDITERRANEAN AND THE CANARY ISLANDS

LA DIVERSIDAD E HISTORIA EVOLUTIVA DE LAS ARAÑAS MIGALOMORFAS EN EL MEDITERRÁNEO OCCIDENTAL Y LAS ISLAS CANARIAS

Memoria presentada por

Vera Opatova

Para optar al título de

Doctora por la Universidad de Barcelona

Barcelona, Abril 2014

DOCTORANDA

Vera Opatova

DIRECTOR Y TUTOR DE TESIS

Dr. Miquel A. Arnedo Lombarte
Profesor agregado
Departamento de Biología Animal
Universidad de Barcelona

ACKNOWLEDGEMETS

After nearly six years spent in Barcelona, first working on my master's degree and then on the PhD, I find it surprisingly difficult to put all the memories and thoughts in these lines. It has been a long time and many people helped me or accompanied me in this process.

Esta tesis no sería posible sin mi director y tutor Miquel Arnedo, me aceptó en su grupo cuando vine de Erasmus para hacer el máster y luego no se asustó con la burocracia que implicaba la convalidación de mis estudios y mi aceptación en el programa de doctorado en la UB. Agradezco mucho la oportunidad que tuve para poder venir a Barcelona y trabajar con los migalomorfos que era una cosa que quería hacer desde que empecé considerar a estudiar biología como carrera universitaria. Muchas gracias Miquel, por guiarme a lo largo de la tesis, he aprendido mucho contigo.

Mis gracias van también a todos los miembros presentes y pasados del "Arnedo Lab". Leti, gracias por ayudarme mucho sobretodo en principio enseñándome todos los secretos de la PCR y por ser una buena compañera que siempre estaba cerca. Paola, no sé que haría sin ti! A parte de que nunca me dejas dormir bajo el puente cuando no tengo casa (que es muy a menudo desgraciadamente), me ayudaste un montón con el laboratorio y siempre eras una gran amiga. Coucou Eli, em van encantar les nostres sortides del camp, quines risses per tot el Mediterrani, ¿no?. Gracies a tu tindré molt per recordar per tota la vida, al lab o fora sempre va ser divertit. Muchas gracias también al resto del lab, a los miembros presentes, pasados, o visitantes que siempre traían ideas nuevas i energía positiva: Pedro, Najet, Vanina, Cristina, Núria, Edu, David, Ageliki.

Muchas gracias a mis compañeros del departamento pasados y presentes, Berta, Gema, Oriol, Enric, Lucía, Blanca, Laura :), Nayra, Mercè, Raquel Vasconcelos, Ana, Sergi, Juan, Jessica por compartir risas e ideas. Muchas gracias también a Carles, Marina y Marta que eran mi comisión del seguimiento y cada año tenían que invertir tiempo y energía evaluando mi progreso. Muchas gracias a María José por explicarme e ayudarme con todos los procesos administrativos que implicaba entregar esta tesis!

During my PhD I was also very fortunate to be able to visit another labs. I am indeed very grateful to Kanchon Dasmahapatra for accepting me in his former lab at University College of London and for teaching me the AFLPs. I had a great time in your lab Kanchon, thank you, and thank you Fraser and Neil for all the funny stories you were telling me every day. Now I know what the British humour is like and I must say I love it.

I would like to thank to Jason Bond whose lab at Auburn University I visited repeatedly, I learned a lot there and I was always looking forward to my next visit. I am grateful to the Bond family for housing me for a week during my first stay and making me feel at home. Thanks to all members of the Bond lab: Niccole, Rebecca, Chis and Charles for all the fun, inspiring conversations about spiders, phylogenies and science in general. Thanks Niccole and Rebecca for driving me around a LOT, especially during my first visit.

When I started this thesis I had literally no material at all, in total, I spend about four months in the field with a lot of different people collecting mygalomorphs across the whole Mediterranean and Eastern Canary Islands. I would like to thank to all the people who were involved in the field trips, who helped me collecting or who gave me material, thanks to them I was able to conduct the research presented in this thesis. Thank you Miquel, Eli, Enric, Pedro, Raquel, Cristina, Sara, Leti, Nuria, Saša, Jason, Niccole and Rebecca for helping me in the field. I would also like to thank to Miguel-Angel Ferrández, Jiří Král, Siegfried Huber, Arthur Decae, Nuria Macías-Hernández, Vláďa Hula, Standa Macík, Honza Korba and Martin Forman for sending me material.

I would also like to thank to Prof. Jan Buchar and Lenka Kubcová, my first advisors in Prague, who introduced me into the fascinating world of spiders and always motivated me with the research, and to Arthur Decae who guided my first steps with the mygals.

Most importantly, I would like to thank to my family and the closest friends. My mom and sister always supported me unconditionally in everything I did and often offered me their valuable perspectives without trying to influence my decisions and opinions. I indeed feel very fortunate for the relationship we have I and would need several pages to express my gratitude for all the things you did for me in my life, thank you.

Ely, also with you I would need so many pages to thank you for all the things you did for me in these years. Thank you for being always so great, so understanding and supportive, for being always close when I most needed you and for trying to teach me patience with the “unimportant” things.

I would also like to thank to all my dear friends from Prague, after all these years abroad, with many of them I keep in touch as often as I used to when we all lived “just around the corner” from each other. Thanks for the moments worth remembering spent in Společenstvo and later on Erik, Haďák, Helena, Štěpán, Saša, for the friendship lasting more that 15 years Zuzana & Zuzana, for all these many years and great times Jana, Tomáš, Radka, Martina.

Y tampoco puedo olvidarme de mis queridos amigos que conocí viviendo en Barcelona! Compartimos muchas risas y momentos divertidos tomando cervezas, paseando, en la playa, disfrutando de una buena comida. Muchos de ellos participaron también en mis mudanzas constantes o me ofrecieron su casa temporalmente, muchas gracias chicos, os lo agradezco un montón!!! Rui, eras el mejor compañero de piso EVER, cada vez que paso por el Disneyland Casp me acuerdo de ti y de Marta, os quiero mucho! Paola, eres una gran amiga, muchas gracias por siempre buscar el lado positivo de las cosas, por ofrecerme tu casa tantas veces y por todas las risas que compartimos, grazie mille! Muchas gracias a Marc, Marc, Nico, Edu, Fernando, Ángela Cristina, Hossein, Claudio, Sergi, Ibai, Jacobo por los momentos divertidos con una copa de cerveza (o muchas), la amistad y también por vuestra ayuda en algún momento de mis mudanzas constantes!

CONTENTS

General Introduction

GENERAL INTRODUCTION

Unravelling cryptic diversity and complex biogeographic patterns in species-poor and morphologically conservative lineages

Earth's Biodiversity is currently being threatened by human activity (Eldredge 2001). It is estimated that approximately 30,000 species go extinct every year, a significant proportion of them are unknown to science (Wilson 1992). Therefore, efficient biodiversity conservation plans require rapid assessment of those areas that are presumably richest in species diversity and present unique evolutionary and ecological features (Myers *et al.* 2000).

The Mediterranean Basin is listed among the twenty-five global biodiversity hotspots deserving special conservation attention (Myers *et al.* 2000). The region harbours approximately 10% of Worlds' diversity of vascular plants, 50% of which are local endemics, 4% of terrestrial vertebrates and 75% of European insect species (Medail & Quezel 1997; Blondel 2010). Within the Mediterranean region the Canary Islands stand out because of their high level of endemism; about 50 % of know invertebrates and 27 % of vascular plants inhabiting the archipelago are local endemics (Juan *et al.* 2000).

The ability to recognize species is fundamental to understanding the origin and diversification of the biodiversity. However, species recognition may be particularly challenging in morphologically uniform groups. The spider infraorder Mygalomorphae, comprised of tarantulas, funnel-web spiders and the trap-door spiders, is one of the three main evolutionary lineages recognized within spiders (Hedin & Bond 2006; Platnick & Gertsch 1976). Family level phylogenetic relationships are with a few exceptions relatively well-solved (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin & Bond 2006), but overall species diversity is probably underestimated.

Seven mygalomorph families inhabit the Mediterranean region, including the Canary Islands (Table 1, Platnick 2014). With the exception of Nemesiidae, all the families show very low levels of species diversity in the area. However, some species are remarkably widespread. For instance, *Cyrtauchenius*

walckenaeri (fam. Cyrtaucheniidae) has a circum Western Mediterranean distribution and *Atypus affinis* (fam. Atypidae) ranges from the Southern Iberian Peninsula to the United Kingdom and Sweden (Platnick 2014). Because of its extremely low dispersal capacity, mygalomorph spiders show high levels of local endemism and deep genetic population structure, as revealed by the pioneering work of Bond and collaborators (2001) on trap-door spider phylogeography. On the other hand, mygalomorphs are morphologically conservative. Therefore, it seems likely that a large portion of mygalomorph diversity in the Mediterranean to date may have been overlooked.

Family	Genera number	Species number
Atypidae	2	2
Hexathelidae	1	2
Cyrtaucheniidae	1	15
Ctenizidae	3	9
Idiopidae	2	3
Nemesiidae	4	68
Theraphosidae	3	15

Table 1. Mygalomorph spider diversity in the Mediterranean region (Platnick 2014)

Factors contributing to the current poor taxonomic knowledge of the group include morphological similarity, biased sampling methodologies and secluded habits. Many groups construct underground burrows relatively difficult to detect in the field. Additionally, direct capture of specimens from burrows usually results in females or immature individuals, because adult males abandon their burrows in search for mates. Males are seasonally active, and are more effectively captured using pitfall traps. Oftentimes, the use of a single collection strategy results in the description of species with a single sex. For example, approximately half of the species in the most species-rich Mediterranean mygalomorph genus, *Nemesia* (fam. Nemesiidae), were described from only females (Platnick 2014). Because of low vagility and burrow fidelity, mygalomorph spider distribution ranges are more likely to reflect

geological processes such as continental breakups or tectonic plate rearrangements (Bauza-Ribot *et al* 2012; Hedin *et al.* 2013; Stock 1993), making them a perfect model system for biogeographic studies. The distribution patterns of some Mediterranean mygalomorph genera triggered the discussion about their actual origins. The bulk of *Ummidia* diversity is found in the New World, which could be explained by both a former Laurasian distribution or a recent human mediated introduction (Decae 2010). The putative African or Asian origin of the European *Macrothele* species (fam. Hexathelidae) is also a matter of debate (Arnedo & Ferrández 2007; Ferrández *et al.* 1998; Haupt 2007; Jimenez-Valverde & Lobo 2007; Van Helsdingen & Decae 1992). Finally, because of the rarity of mygalomorph spiders on oceanic archipelagos, the presence of *Titanidiops canariensis* (fam. Idiopidae) on the volcanic Eastern Canary Islands has also challenged classic vicariant wisdom.

The present Ph.D. thesis provides insights into the cryptic diversity, distribution patterns and phylogenetic relationships of some mygalomorph families that inhabit the Western Mediterranean and the Canary Islands. By using a multi-locus approach and state of the art phylogenetic inference methods, the pattern and time frame of the diversification of targeted groups is inferred. Phylogenetic and temporal information sheds new light on biogeographic scenarios and diversification drivers of selected families. Molecular based species delimitation approaches are further used to identify species boundaries in morphologically conservative groups.

Methodology framework

Gene tree versus species tree approaches

Concatenation of genes evolving at different rates, widely used to resolve phylogenetic relationships, is based on the assumption that individual gene trees are congruent among them and further mirror the species tree. This approach has been the paradigm in phylogenetic studies until recently. Nevertheless, it is widely recognized that numerous evolutionary and demographic processes may lead to discrepancies among gene trees.

The acceleration of substitution rates in genes of some species with

respect to others included in the analysis or the use of highly saturated markers may affect branch lengths and tree topology (Felsenstein 1978; Degnan & Rosenberg 2006; Kubatko & Degnan 2007). Incomplete lineage sorting (ILS), also referred to as deep coalescence, especially pervasive in recently evolved species with large population sizes, may also be responsible for conflictive topologies among individual gene trees (Maddison & Knowles 2006). Other frequent sources of gene incongruence are hybridization, recombination, duplication or horizontal gene transfer (Maddison 1997; Degnan & Rosenberg 2006; Kubatko & Degnan 2007). New phylogenetic methods to infer species tree that explicitly incorporate gene incongruence due to coalescent error have been recently developed (Heled & Drummond 2010; Kubatko *et al.* 2009; Liu & Pearl 2007).

One of the most widely used methods for species tree and divergence time inference is the Bayesian based multiple species coalescence approach (Heled & Drummond 2010), implemented in the BEAST programme package (Drummond *et al.* 2012). This approach estimates the effective population sizes and divergence times for all individual gene trees contained in the species tree. The programme assumes no gene flow and explains the possible discrepancies among individual gene trees as a result of incomplete lineage sorting (Heled & Drummond 2010).

Divergence time estimation

Molecular dating of divergence times is a widely used tool in phylogenetic and phylogeographic studies. The inference of the diversification time frame facilitates contrasting alternative biogeographic hypothesis (e.g. (Noonan 2000; Paulo *et al.* 2008; Rowe *et al.* 2010), for example by discriminating between evolutionary colonisations or human mediated introductions (Chapman *et al.* 2007; Cunningham 2008).

Divergence time estimate methods are based on the molecular clock hypothesis (Zuckerkandl & Pauling 1965). It was empirically demonstrated that proteins gradually accumulate changes in nucleotide sequence through time and the rate of accumulation of the mutations differ among proteins (Zuckerkandl & Pauling 1965).

The strict molecular clock assumes a constant mutation rate trough time

for all the organisms included in the analysis, which has been shown to be usually an unrealistic assumption. Alternatively, the relaxed molecular methods allow the mutation rate to vary among branches. Relaxed clock methods may be further divided between those that assume a temporal autocorrelation of the rate between ancestral and descendent lineages e.g. R8s (Sanderson 2003), and those that do not e.g. BEAST (Drummond & Rambaut 2007; Heled & Drummond 2010).

The time divergence method implemented in the computer package BEAST (Drummond & Rambaut 2007; Drummond *et al.* 2012) operates in a Bayesian framework and implements a MCMC approach in order to obtain the posterior distribution for divergence time estimates and substitution rates. In addition, it allows the use of different evolutionary models for nucleotide substitution. BEAST provides algorithms for divergence time estimates using both concatenated genes and multispecies coalescence approaches *BEAST: (Heled & Drummond 2010) (see the gene tree versus species tree section). It has been reported in the literature that the two approaches may infer different speciation time estimates. Time estimates based on concatenated genes frequently overestimate divergence times because of the gene divergence time may precede the actual population/species split (for review see Edwards & Beerli 2000; Arbogast *et al.* 2002). Conversely, coalescent-based methods may lead to underestimation of the speciation times in the presence of unaccounted gene flow (McCormack et al 2010). Simulation studies have demonstrated the impact of gene flow or horizontal transfer on time estimates, specially when non-sister taxa are involved in the genetic exchange (Leaché *et al.* 2014).

The incorporation of fossil or biogeographic calibration points to the analysis enables the inference of absolute diversification times. The fossil record provides the primary source of information to infer absolute times and is usually incorporated as a minimum bound for the stem group to which the extinct organism belonged. Biogeographic data provides an additional source of information to calibrate phylogenetic trees. Biogeographic calibration points can be related to some well documented vicariant events. For example, the time of the opening of the Strait of Gibraltar was used to date the divergence time between sister taxa inhabiting the Iberian and African sides (Bidegaray-Batista *et al.* 2011), which are usually incorporated as nearly-fixed times. The

emergence of volcanic islands can also be used to date the divergence of sister taxa inhabiting younger and older islands, respectively such as the Canary Islands (Macías-Hernández *et al.* 2010), which are usually incorporated as maximum bounds for the time of split of sister taxa. The use of biogeography-based calibration points, however, has been criticized (Heads 2011; Kodandaramaiah 2011) due to the difficulty in demonstrating the assumptions involved.

In the absence of informative biogeographic events or fossil data that may provide calibration points, absolute divergence times can be inferred with the use of substitution rates estimated for other groups reported in the literature. For instance, in arthropods one of the most widely used substitution rates is the 2.3% pairwise sequence divergence per million years for the mtDNA (Brower 1994). More recently, a spider specific mitochondrial substitution rate has also been estimated for the ground-dwelling genus *Parachtes* (Bidegaray-Batista *et al.* 2011). It should be borne in mind, however, that the use of a substitution rates, either universal or inferred for closely related groups, is based on the assumption that the substitution rates in the group of interest are the same, which may not be true. For example, an accelerated rate of mitochondrial genes has been reported in mygalomorph spiders (Bond *et al.* 2001) and scorpions (Gantenbein 2004).

Species Delimitation

Species are the most fundamental units of biodiversity and the main outcomes of the evolutionary process (Ereshefsky 1992; Wiley 1981), but very few topics in biology have raised as much debate (Claridge *et al.* 1997; Coyne & Orr 2004; Cracraft 1989; de Queiroz 2005; Ereshefsky 1992; Hey 2001; Mayr 1982; Ridley 2004). There are more than 20 definitions of the species (i.e. species concepts) that emphasize different aspects and attributes of the species and the speciation process (de Queiroz 2007; Ereshefsky 1992; Mallet 2001; Mayden & Wood 1994). However, independence of evolutionary lineages is a common underlining feature and alternative concepts are better viewed as operational criteria for species recognition (Mayden 1997; de Queiroz 1998). Following this line of reasoning, de Queiroz (2007) defined a species as a separately evolving metapopulation lineage. The term lineage would here refer

to successive ancestor – descendant metapopulations extended through time (de Queiroz 2007). Other species concepts may be relevant in the species delimitation processes and their combination corroborates the robustness of the delimited species (de Queiroz 1998; Padial *et al.* 2010).

Species boundaries delimitation may be particularly challenging when dealing with morphologically uniform (Bond & Stockman 2008; Stockman & Bond 2007; Hamilton *et al.* 2014; Hendrixson *et al.* 2013) or recently evolved taxa (Pons *et al.* 2006; Shaffer & Thomson 2007). The integration of different lines of evidence (e.g. morphology, sequence data, genotyping, species distribution modeling) maximizes the objectivity and robustness of the species delimitation process (Bond & Stockman 2008; Dasmahapatra *et al.* 2009; Raxworthy *et al.* 2007, Rissler & Apodaca 2007; Hendrixson *et al.* 2013; Wiens & Penkrot 2002; Edwards & Knowles 2014).

While the integrative approach used for fine-tuning species delimitation is usually time consuming, the use of single molecular markers (i.e. the DNA barcoding, currently the *cox1* mtDNA gene for animals) can be used to identify species candidate lineages for fast biodiversity assessment. Several methods have been devised to identify species boundaries based on a single locus: the General Mixed Yule Coalescent model (GMYC) (Pons *et al.* 2006), the Automatic Gap Barcode Discovery (AGBD) (Puillandre *et al.* 2012) or the Barcode Index Number System (BIN) (Ratnasingham & Hebert 2013), among others.

The GMYC method is maximum likelihood sequence based approach that clusters individuals into putative species (i.e. GMYC lineages) by detecting the transition point between coalescence and speciation processes. The AGBD and the BIN, on the other hand, are distance-based methods. The GMYC approach tends to over split the data (Hamilton *et al.* 2014; Esselstyn *et al.* 2012; Kekkonen & Hebert 2014; Talavera *et al.* 2013) and thus is better interpreted as an objective first step into the assignment of the individuals into putative species for subsequent analyses (Edwards & Knowles 2014). The final decision on what constitutes an actual species will further require the integration of additional sources of evidence (Kekkonen & Hebert 2014).

The ease of generating genomic data for non-model organisms and increases in computing power have prompted the development of the multilocus

coalescent approach to tackle DNA based species delimitation (Leaché & Fujita 2010, Brown *et al.* 2012; Burbrink *et al.* 2011; Paez-Moscoso *et al.* 2011, Sistrom *et al.* 2013).

The Bayesian Phylogenetic and Phylogeography (BPP) program (Yang & Rannala 2010) uses the reversible-jump Markov Chain Monte Carlo (rjMCMC) algorithm to estimate the posterior probability of species clusters from multilocus data given a guide species tree topology. This method offers the possibility to define alternative evolutionary scenarios concerning the effective population sizes (θ) and root ages (τ), which may lead to favor a more or less conservative number of delimited species (Leavitt *et al.* 2011; Yang & Rannala 2010). The BPP approach assumes no gene flow after the speciation event, but efficient performance has been proven even under scenarios with low levels of gene flow (e.g. 0.1 migrant per generation) (Zhang *et al.* 2011).

Amplified fragment length polymorphism (AFLP)

The amplified fragment length polymorphism (AFLP) (Vos *et al.* 1995) is a popular technique to address population genetics and phylogeographic issues, especially in non-model organisms. The method does not require any prior knowledge about the genome of the studied organism and with a relatively simple protocol that involves digestion of the genome with restriction enzymes and subsequent specific amplification of adaptor tagged fragments, hundreds of loci can be generated in a cost-efficient short time (Bensch & Akesson 2005; Meudt & Clarke 2007; Mueller & Wolfenbarger 1999).

The reproducibility of the AFLP loci and the objectiveness of the scoring process was a matter of debate, but a laboratory workflow for evaluating the reproducibility through replicates has recently been proposed (Crawford *et al.* 2012; Pompanon *et al.* 2005) and the implementation of automatic scoring methods (Arrigo *et al.* 2009; Bonin *et al.* 2007; Herrmann *et al.* 2013; Whitlock *et al.* 2008) has greatly improved scoring objectivity.

The performance of AFLPs compared to microsatellites or allozymes has been investigated on several occasions and the results indicate that AFLPs are highly efficient for individual-based population assignment methods (Campbell *et al.* 2003; Mariette *et al.* 2001; Nybom 2004).

Species distribution modelling (SDM)

Species distribution modelling (SDM) is based on the relationship between ecological and environmental properties of the landscape and the known occurrence of species (Guisan & Zimmermann 2000; Elith *et al.* 2006; Peterson 2006; Soberon & Peterson 2005) and aims to reconstruct ecological requirements of a species and predict its geographic distributions in the past, present or future. This procedure first sets up the environmental space of a species and geographically projects it. Then SDM detects additional areas where a suitable environment is represented. However, it is important to bear in mind that potentially suitable habitats do not necessarily need to be inhabited by the species of study due to potential dispersal barriers (Peterson 2006). Thanks to the increasing availability of environmental data layers (WorldClim database, Hijmans *et al.* 2005) and methodological advances, SDM has become a widely used tool in ecological, phylogeographic and conservation studies (Alvarado-Serrano & Knowles 2014).

The geological and geographic setting

The Mediterranean basin

The origin of the Mediterranean basin traces back to the collision between the African and the Eurasian plates, which precipitated the closure of the Thetis Ocean about 40 million years ago (Ma) (Blondel & Aronson 1999). The impact of the collision catalysed the Alpine orogeny (i.e. the formation of mountain ranges along the Mediterranean Basin), starting at the upper Eocene and continuing throughout the Miocene (Schellart 2002).

In the middle Oligocene, a block of landmass known as the Hercynian belt, split and drifted off from the north-eastern Iberia and southern France as a result of former plate subduction rollback (Rosenbaum *et al.* 2002a; Rosenbaum *et al.* 2002b; Rosenbaum *et al.* 2002c) (Fig 1). The breakup of the Hercynian belt, comprising the Balearic Island, Corsica, Sardinia, Calabria, Betic-Rif and the Kabylies microplates, started with the opening of the Gulf of Lyon around 30 Ma (Séranne 1999), when the Corsica, Sardinia and Calabria block started rifting to their present day position. The second phase of the breakup, which separated the Betic-Rif, Balearic Island and Kabylies

microplates from the continuous landmass and induced the opening of Valencia Trough traces back to about 25 Ma (Roca *et al.* 1999).

The Corsica, Sardinia and Calabria block started rifting eastward and eventually collided with the Apennines between 20-18 Ma. The separation of Corsica and Sardinia has been dated between 21-15 Ma, a time of eastward migration while the plate was undergoing a series of counter clockwise rotations (Speranza *et al.* 2002; Gattacceca *et al.* 2007). Corsica and Sardinia subsequently detached from the Apennines in the mid Miocene around 9 Ma and started drifting off towards their present day position (Rosenbaum *et al.* 2002).

Around 21 Ma a continental break-up occurred in the region between the Balearic Islands and the Kabylies microplates and initiated the subsequent Kabylies southward rifting (Rehault *et al.* 1984). The separation of the two Kabylies blocks likely started in the same time period as the southward rifting and terminated before 18 – 15 Ma when the blocks collided with Northern Africa reaching thus their present day position (Frizon de Lamotte *et al.* 2000).

The Betic-Rif plate started rifting westward from its original location around the Though of Valencia around 23 Ma and 15 Ma and eventually started fragmenting into the Betic and Rif blocks, which reached their present day location on both sides of the Strait of Gibraltar by mid Miocene, approximately 10 Ma, completing the opening of the Alboran Sea (Lonergan & White 1997).

Alternatively, some authors suggest that the connection between the southernmost part of the Betic, corresponding to the Gibraltar region of the Iberian Peninsula, with the Rif block persisted until Upper Tortorian/Lower Messinian (~8 – 7.2 Ma), while the water exchange between the Mediterranean and the Atlantic Ocean was facilitated through the Guadalquivir Basin (for review see Braga *et al.* 2003) (Fig 1).

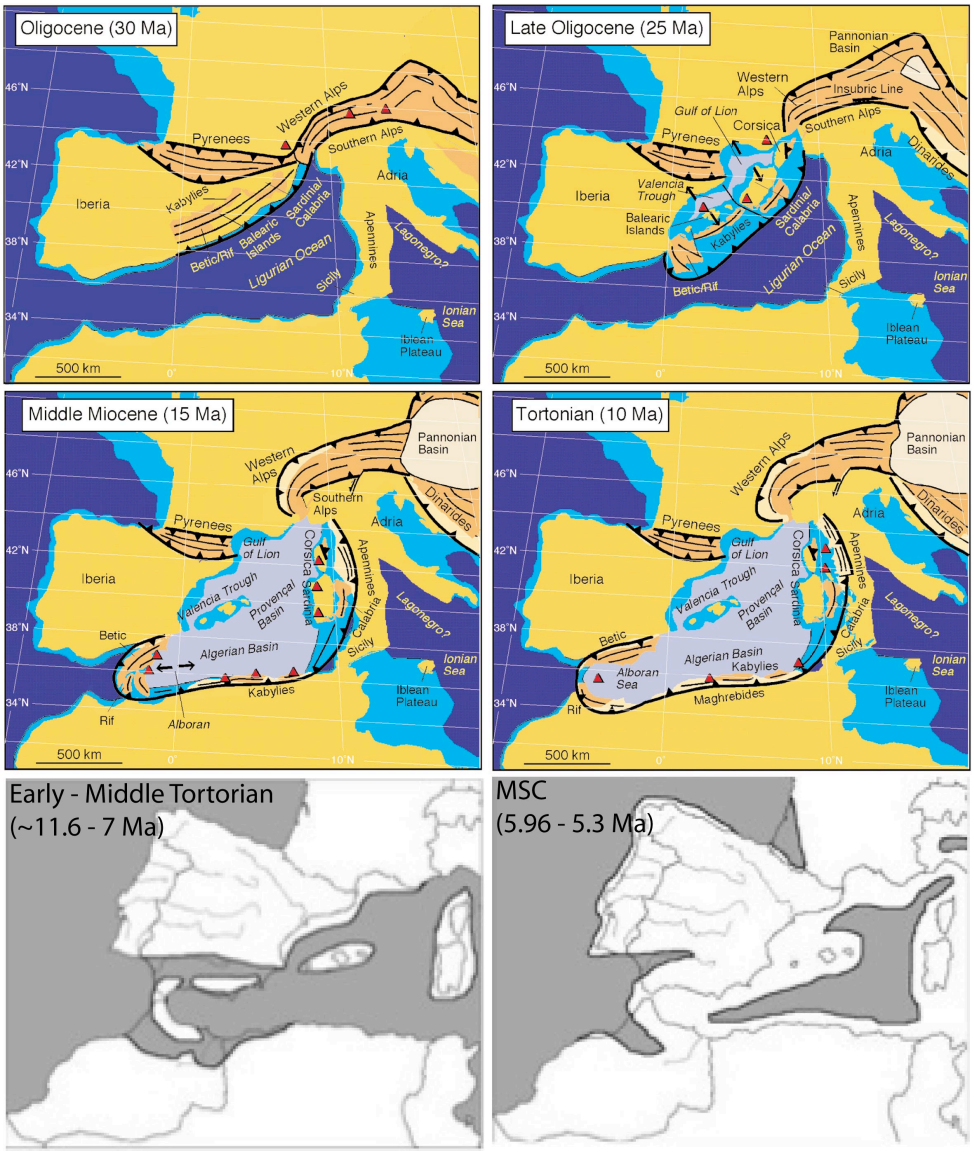


Figure 1. Geological history of the Western Mediterranean; Oligocene 30 Ma, Late Oligocene 25 Ma, Middle Miocene 15 Ma, Tortorian 10 Ma Rosenbaum et al 2002. Alternative hypotheses: Early – Middle Tortorian, Messinian Salinity Crisis (MSC) from Paulo et al 2008.

Geological studies conducted on the marine sediments in the internal zone of the central–eastern part of the Betic zone also documented that the uplifting of Betic Cordillera was a continuous process that took place from the lower Tortonian throughout the upper Messinian (~8-6 Ma), during which the region had a reticulate character, where emerged areas were partly separated by shallow basins (Braga *et al.* 2003).

The closure of the Strait of Gibraltar in the late Miocene (about 5.96 Ma) (but see Braga *et al.* 2003) drove the nearly complete desiccation of the Mediterranean Basin (i.e., the Messinian salinity crisis). The dramatic sea level drop connected formerly isolated areas through multiple land bridges (e.g., North Africa with Iberia, Sicily and Italy and Corsica with Sardinia). The previous conditions were re-established by the reopening of the Strait of Gibraltar at 5.3 Ma (Jolivet *et al.* 2006; Krijgsman *et al.* 1999).

In the Eastern Mediterranean a continuous landmass containing the present day mainland Greece, Crete, the small Aegean islands and parts of Anatolia (i.e. Agaïs) started to break up after the formation of the mid-Aegean trench in the Upper Miocene (12–9 Ma), which resulted in the opening of Anatolian sea and the disjunction of Crete about 8 Ma (Creutzburg 1963; Dermitzakis 1990). The entire region has underwent complex rearrangements following the Aegean plate breakup (8 Ma) to the present day; many of the small Aegean islands were repeatedly connected and disconnected with Peloponnesus or Anatolia during the Messinian salinity crisis and the Plio-Pleistocene eustatic sea level oscillations, while Crete remained isolated during this time (Creutzburg 1963; Dermitzakis 1990).

The Eastern Canary Islands

The Canary Islands archipelago lies on a volcanic ridge, approximately 110 km off the northwest coast of Africa and is comprised of seven main islands and several smaller islets. The archipelago was created by cyclic volcanic activity, which started in the Miocene and has continued with varying levels of intensity until today (Carracedo *et al.* 1998b; Carracedo *et al.* 2003; Ibáñez *et al.* 2012).

The Eastern Canaries were the first islands to emerge and were subsequently followed by the rest of the islands, following a East to West trend (Carracedo *et al.* 1998a). The Jandía peninsula, south of present-day Fuerteventura was the first to emerge about 20 Ma and the volcanism progressed in a SSW to NNE spatial polarity. Lanzarote emerged as two independent proto islands: the Ajaches massif, in the southwest, and Famara, in the northeast, which formed between 15.5 – 13.5 Ma and 10.2 – 3.8 Ma, respectively. After the initial intensive volcanic phase, the proto islands were

eventually connected, following a long quiescent, erosive period (Coello *et al.* 1992; Carracedo & Rodríguez-Badiola 1993). The second pulse of volcanic activity started in Lanzarote around 1.6 Ma and continued until the present, the last eruption was documented in historic times (Carracedo & Rodríguez-Badiola 1993; Carracedo *et al.* 2003). The Eastern Canary Islands, and the northern islets, were repeatedly connected via land bridges during the quaternary climatic oscillations (Carracedo *et al.* 2003).

Climatic changes in the region

Paralleling the dramatic geological events, the Mediterranean also experienced important climatic changes. At the beginning of the Eocene (~ 55 Ma) the climate was warmer and wetter than today. Following the Eocene climatic optimum, a gradual long-term trend towards lower temperatures and dryer conditions began, interrupted by a short period of temperature increase during the Miocene (17 – 15 Ma) (Zachos *et al.* 2001).

The cooling process eventually lead to the establishment of the double seasonality pattern that characterizes present-day Mediterranean climate, beginning about 3.4 Ma (Jiménez-Moreno *et al.* 2010; Suc 1984). The growth of the first ice sheets in the Northern Hemisphere marked the start of the Quaternary glacial cycles, which in the Mediterranean region began between 2.8 and 2.5 Ma (Gibbard *et al.* 2010).

During these 100,000 year long cycles, several regions of the Mediterranean Basin played an essential role as a refugia to thermophilic species (Blondel & Aronson 1999; Hewitt 1996; Hewitt 2000). The cyclic growth of the Northern Hemisphere ice sheets also altered the sea level in the region. Land bridges connected previously isolated areas in both the Mediterranean (Creutzburg 1963; Dermitzakis 1990) and the Eastern Canary Islands. The last glacial period (i.e. Würm) traces back to 115,000 – 110,790 years before present (bp), with the glacial maximum dated about 21,000 to 18,000 bp (Bowen *et al.* 1986).

The model system: the mygalomorph spiders

The infraorder Mygalomorphae, comprising tarantulas, purse-web spiders, funnel-web spiders, trap-door spiders and their kin, is one of the three main lineages recognized within spiders (Hedin & Bond 2006; Platnick & Gertsch 1976). The group is of ancient origin (Penney & Selden 2011) and it is often described as 'primitive' because of the lack of specialization in the spinning structures and the retention of characters regarded as plesiomorphic such as the presence of four book lungs and chelicerae bearing longitudinal fangs with unsynchronized movement (Raven 1985). Although mygalomorphs inhabit all the continents except Antarctica, they are significantly less diverse both in number of genera and species than their sister group, the Araneomorphae: there are less than 3,000 described mygalomorph species whereas over 41,000 araneomorphs are known (Platnick 2014).

Females of this group are usually long-lived and show very strong burrow or nest fidelity, while adult males searching for mates mediate the gene flow among populations. Mygalomorphs tend to show high levels of local endemism, generally attributed to low vagility of the species (Arnedo & Ferrández 2007; Bond *et al.* 2001; Bond & Stockman 2008; Cooper *et al.* 2011; Hendrixson *et al.* 2013; Satler *et al.* 2011). The limited dispersal ability and sedentary nature makes mygalomorphs an excellent model system for biogeographic studies (Hedin *et al.* 2013; Raven 1980).

Dispersal through ballooning, which is very common in araneomorph spiders (Duffey 1956, 1997) has been reported only in three mygalomorph families, namely Atypidae, Ctenizidae and Actinopodidae (Coyle *et al.* 1985; Eberhard 2006; Coyle 1983; Cutler & Guarisco 1995; Ferretti *et al.* 2013). The efficiency of long distance dispersal through ballooning may also differ among groups. The presence of *Ummidia* (Ctenizidae) species in some Caribbean Islands of volcanic origin without previous connection to any landmass (Platnick 2014), indicates long distance dispersal capability, but deep genetic structuring detected among *Atypus* populations (Atypidae) (Pedersen & Loeschcke 2001) suggest dispersal over short distances.

Because of their low vagility, mygalomorph spiders are extremely rare in oceanic islands. Few cases of non-ballooning mygalomorphs are known from

the Australian region, although colonization via land bridges from the adjacent mainland could not be ruled out for most of these taxa (Raven 1980). In this regard, the presence of a trap door spider *Titanidiops canariensis* (Idiopidae) on the Canary Islands is of great biogeographic interest, as the archipelago has never been connected to the mainland and the family Idiopidae has no reported cases of aerial dispersal.

Mygalomorph taxonomy is challenging, most of the closely related taxa are morphologically homogenous (Bond & Hedin 2006) and the majority of diagnostic characters are based on reproductive organs of adult males, usually available in the field only during a short period of time. The conservative morphology and poor dispersal ability causing extreme genetic structuring even among geographically close populations, makes mygalomorphs a good candidate to investigate species boundaries (Bond & Stockman 2008; Hamilton *et al.* 2011; Hendrixson *et al.* 2013; Satler *et al.* 2011; Stockman & Bond 2007).

There are seven mygalomorph families present in the Mediterranean Region and the Canaries (Table 1) (Platnick 2014). In the present study, I have used genera of the trap-door spider families Ctenizidae and Idiopidae and the funnel-web spider family Hexathelidae as model systems (Fig 3).

Family Ctenizidae

The family Ctenizidae is widespread across North and South America, the Mediterranean region, South Africa, South-East Asia and Australia (Fig 4). In spite of its worldwide distribution, ctenizids, comprising 128 species and nine genera, rank among the less diverse spider families (Platnick 2014). All genera construct underground, silk-lined burrows that open to the surface with a trapdoor (Decae 2010; Decae 1996). The monophyly of Ctenizidae was not recovered in recent molecular phylogenetic studies of the entire infraorder Mygalomorphae, but no alternative placement of the genera has received substantial support (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin & Bond 2006).

In the Mediterranean region the family Ctenizidae is represented by three genera with non-overlapping distributions (Fig. 4), *Cteniza* Latreille, 1829, *Cyrtocarenum* Ausserer, 1871 and *Ummidia* Thorell, 1875. The genus *Cteniza* and *Cyrtocarenum* are mostly endemic to the region, whereas the bulk of *Ummidia* diversity lays in the New World (Platnick 2014). Unlike the vast

majority of mygalomorphs, the spiders of the genus *Ummidia* are capable of air-borne dispersal (Coyle *et al.* 1985; Cutler & Guarisco 1995; Eberhard 2006; Ferretti *et al.* 2013).

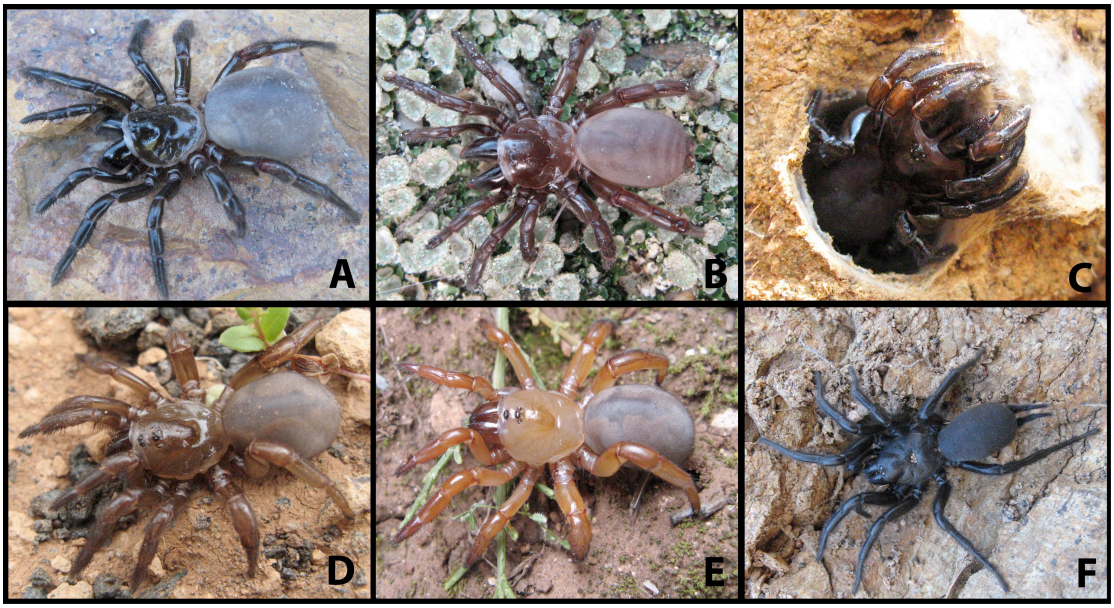


Figure 2. Mediterranean mygalomorph spiders. Family Ctenizidae: A: *Cyrtocarenum* sp., Crete B: *Cteniza* sp., Corsica C: *Ummidia* sp., Spain, Family Idiopidae: D: *Titanidiops canariensis*, Betancuria, Fuerteventura, E: *Titanidiops* sp. Morocco, Family Hexathelidae: E: *Macrothele calpeiana*, Sierra de Aracena, Spain, Photo credit: VO

Currently, there are four *Ummidia* species described from the Mediterranean region, namely *U. aedificatoria* (Westwood, 1840), *U. algarve* (Decae, 2010), *U. algeriana* (Lucas, 1846) and *U. picea* (Thorell, 1875). The amphi-Atlantic distribution of the genus and the fact that the bulk of its diversity is found in the New World, led to suggest that the presence of *Ummidia* in the Mediterranean is a product of a human-mediated introduction (Simon 1864, 1910). An alternative hypothesis has been offered recently by Decae (Decae 2010), suggesting that the presence of the genus on both sides of the Atlantic Ocean could be a result of a former Laurasian distribution.

There is virtually no information about the phylogenetic position of *Cteniza* and *Cyrtocarenum* and even the validity of both genera has been brought into question (Decae 1996; Raven 1985).

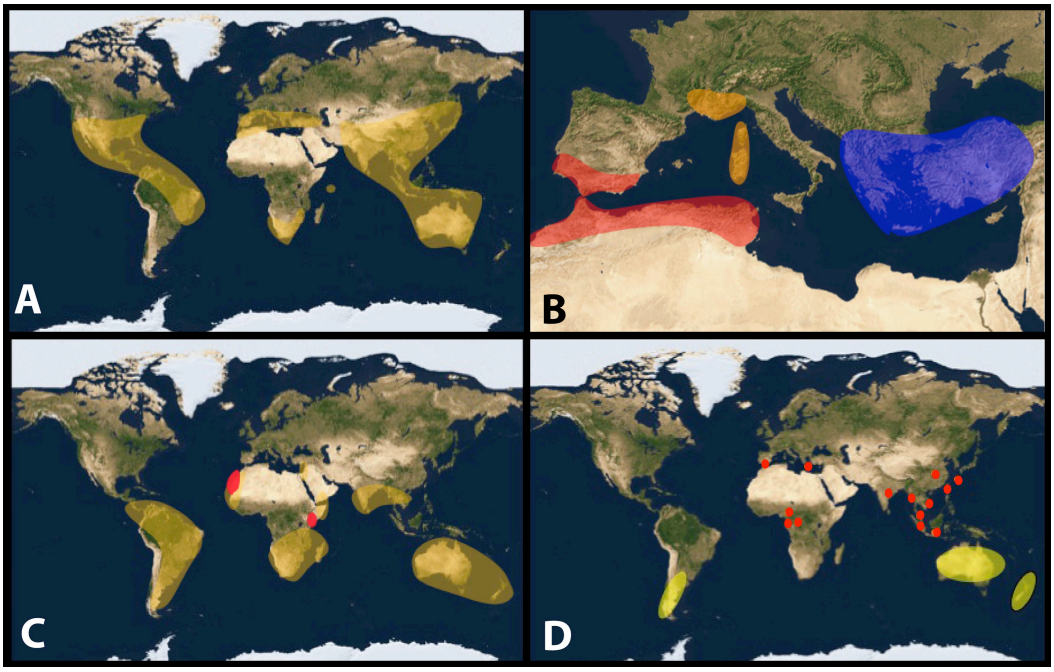


Fig. 3 A: Ctenizidae distribution, B: Distribution of the Mediterranean Ctenizidae genera: *Cyrtocarenum*: blue, *Cteniza*: yellow, *Ummidia*: red, C: Idiopidae distribution: yellow, genus *Titanidiops*: red, D: Hexathelidae distribution: yellow, genus *Macrothele*: red dots.

Family Idiopidae

The trap-door spider family Idiopidae is widespread across Australia, New Zealand, South-East Asia, Sub-Saharan Africa, Madagascar and South America, but has also few remotely located species in North Africa, the Canary Islands and the Middle East (Platnick 2014). This wide distribution is often explained by the Gondwanan origin of the family (Ayoub *et al.* 2007; Hedin & Bond 2006) (Fig 4).

With 318 species placed in 22 genera the family ranks among the most species-rich mygalomorph groups (Platnick 2014). The family is well characterized on morphological grounds (Raven 1985) and its monophyly has been corroborated in recent molecular phylogenetic studies (Ayoub *et al.* 2007; Bond 2012; Hedin & Bond 2006).

The genus *Titanidiops* Simon, 1903 currently comprises only three described species; *T. canariensis* Wunderlich, 1992, endemic to Canary Islands, *T. maroccanus* Simon, 1909 from Morocco and *T. compactus* (Gerstäcker, 1873), which occurs in East Africa. *Titanidiops canariensis* inhabits Fuerteventura, Lanzarote and the islet of La Graciosa, in the Canary Islands, which is of great biogeographic interest because it constitutes one of the few examples of non-ballooning mygalomorphs colonising oceanic islands.

Family Hexathelidae

The family Hexathelidae includes 113 described species and is one of the less diverse mygalomorph families. The greatest numbers of species are found in Australia and New Zealand, but some genera are also present in South America and South-East Asia (Fig 4). Both morphology based analyses (Goloboff 1993) and molecular phylogenetic studies (Ayoub *et al.* 2007; Hedin & Bond 2006), or the combination of both data types (Bond *et al.* 2012), have failed to recover the monophyly of the family.

The genus *Macrothele* Ausserer, 1871 is the only hexathelid distributed outside the Australasian region and the one with the most widespread distribution (Platnick 2014). The bulk of the *Macrothele* diversity is found in Southeast Asia, few species are known from central Africa and only two species with disjunct distributions inhabit Europe: *Macrothele calpeiana* (Walckenaer, 1805) is found in the southern Iberian Peninsula while *Macrothele cretica* Kulczynski, 1903 is endemic to Crete (Snazell & Allison 1989).

The disjunct distributions of the European species promoted discussion about their origins and colonization pathways into the Mediterranean region. Some authors proposed an African origin of the Iberian *M. calpeiana* (Arnedo & Ferrández 2007; Ferrández *et al.* 1998; Van Helsdingen & Decae 1992), while others argue for an Asian origin (Jimenez-Valverde& Lobo 2007). Similarly, some authors have questioned the natural colonization of the Iberian peninsula by *M. calpeiana* and have proposed the species is a result of a recent introduction either from northern Africa (Van Helsdingen & Decae 1992) or Asia (Haupt 2008).

Both European *Macrothele* species are of conservation concern. *Macrothele calpeiana* is the only spider protected by European Union legislation

because of its presumable affinity to the diminishing primary cork oak forest (Collins & Wells 1987; Snazell 1986; Snazell & Allison 1989). However, its status as a bioincator was challenged by subsequent field observations (Ferrández & Ferrández de Cespedes 2001; Van Helsdingen 1993; Van Helsdingen & Decae 1992) and species distribution modelling studies (Jimenez-Valverde & Lobo 2007; Jiménez-Valverde & Lobo 2006). The present day distribution of the species is highly fragmented, with limited gene flow among populations, and some of the marginal populations could qualify as vulnerable (Arnedo & Ferrández 2007).

There is virtually no information about the biology or life cycle of *M. cretica*. Given the fact that this species is known from a few localities in Western Crete, where the natural habitat might be negatively influenced by human activity, the species has been included in the IUCN red list under the data deficient category.

Aim & specific Objectives

AIM

The main goal of this study was to unravel the phylogenetic relationships and phylogeographic patterns in Western Mediterranean and Canarian mygalomorph spiders by means of a multi-locus approach and state-of-the-art phylogenetic inference methods to identify the main drivers of diversification, to test alternative biogeographic scenarios for the origin and to uncover overlooked diversity of this so-far neglected group.

Specific objectives

1. To provide novel insights into the phylogeny of the family Ctenizidae and to infer a timeframe for its diversification with the aim of unravelling the origins of the Mediterranean taxa and shedding light on their taxonomic status.

2. To unravel the biogeography patterns of the genus *Ummidia* in the context of the dynamic geologic history of the Mediterranean region, to provide a temporal framework for its diversification and to assess the ecological preferences of three *Ummidia* species in order to evaluate their ecological exchangeability.

3. To uncover the origins and colonization pathways of the genus *Macrothele* in Europe and to investigate the limits of the family Hexathelidae in the context of a wider mygalomorph phylogeny framework provided by previous studies.

4. To uncover the phylogenetic origins of *T. canariensis*, one of the few examples of trap-door spiders endemic to an oceanic archipelago, to infer the temporal framework for the colonisation of the islands and to assess the possible existence of cryptic species by means of new multi-species coalescent methods.

Informe del director de tesis

Informe del Director de tesis sobre el factor de impacto de los artículos publicados

Como director de la tesis doctoral de Vera Opatova, presento el siguiente informe sobre el factor de impacto de las publicaciones presentadas en la tesis:

1. Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae) (2013) Vera Opatova, Jason E. Bond, & Miquel A. Arnedo *Molecular Phylogenetics and Evolution*, **69**: 1135-1145

Molecular Phylogenetics and Evolution tiene en la última edición disponible de los *Journal Citation Reports* (2012) un índice de impacto de 4,066. Esta revista figura en el segundo cuartil (15 de 47) del área "*Evolutionary Biology*". *Molecular Phylogenetics and Evolution* es un referente en el campo de la sistemática molecular y el uso de filogenias moleculares para el estudio de la evolución.

2. From Gondwana to Europe: inferring the origins of Mediterranean *Macrothele* spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae (in press) Vera Opatova & Miquel A Arnedo *Invertebrate systematics*.

Invertebrate systematics tiene en la última edición disponible de los Journal Citation Reports (2012) un índice de impacto de 1,983. Esta revista figura en el primer cuartil (31 de 151) del área "*Zoology*". *Invertebrate systematics* publica artículos relevantes en el campo de la sistemática y evolución de invertebrados, con especial interés en la región de Australasia y en islas. Ha publicado un buen numero de estudios sobre arañas por lo que es un referente en el mundo de la aracnología.

3. Spiders on a hot volcanic roof: Colonization pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae) (en revisión) Vera Opatova & Miquel A Arnedo. *Plos ONE*.

Este capítulo se encuentra actualmente en fase de revisión en la revista *Plos ONE.*
Plos ONE tiene en la última edición disponible de los Journal Citation Reports (2012) un índice de impacto de 3.73. Esta revista figura en el primer cuartil (7 de 56) del área "*Multidisciplinary Sciences*". *Plos ONE* es una revista *Open Access* que en poco tiempo se ha convertido en un referente entre las revistas de este tipo que publican sobre temáticas evolutivas.

4. Loosening the belt: How the Hercynian Belt breakup shaped the diversity and distribution of the trap-door spider genus *Ummidia* (Araneae, Ctenizidae) in the Western Mediterranean. Vera Opatova, Jason E Bond & Miquel A Arnedo

Se ha finalizado la redacción del artículo correspondiente a este capítulo que está actualmente en fase de *friendly review*. Se prevé enviarlo para su publicación a la revista *Journal of Biogeography* (Índice de impacto en 2012: 4,863, primer cuartil de "*Ecology*", posición 18 de 136), un referente en el campo de la biogeografía y el estudio de los factores que han modelado las distribuciones actuales de los organismos.

Finalmente, se ha adjuntado un quinto trabajo a la tesis en forma de apéndice y que se prevé finalizar en los próximos meses. Los datos filogenéticos y filogeográficos presentados en dicho apéndice se combinarán con información generada por modelos de distribuciones de especies, con proyecciones en el pasado y el futuro, así como modelos ABC para investigar la demografía histórica de la especie objeto de estudio (*Macrothele calpeiana*). El artículo final será probablemente enviado para su consideración a la revista *Molecular Ecology*.

Informe del Director de tesis sobre la participación de la doctoranda en cada uno de los artículos publicados

Como director de la tesis doctoral de Vera Opatova, presento el siguiente informe sobre la contribución de la doctoranda en las publicaciones en coautoría presentadas en la tesis:

1. Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae) (2013) Vera Opatova, Jason E. Bond, & Miquel A. Arnedo *Molecular Phylogenetics and Evolution*, **69**: 1135-1145

Contribución de la doctoranda: Participación en el diseño del trabajo, recogida de las muestras, trabajo de secuenciación, realización de los análisis y redacción de la primera versión del manuscrito y revisiones posteriores.
Contribución de los otros autores: MA, diseño y supervisión, recogida de las muestras, participación en los análisis y en la redacción. JB, supervisión, recogida de las muestras, participación en la redacción.

2. From Gondwana to Europe: inferring the origins of Mediterranean *Macrothele* spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae (in press) Vera Opatova & Miquel A Arnedo *Invertebrate systematics*.

Contribución de la doctoranda: Participación en el diseño del trabajo, recogida de las muestras, trabajo de secuenciación, realización de los análisis y redacción de la primera versión del manuscrito y revisiones posteriores.
Contribución de los otros autores: MA, diseño y supervisión, recogida de las muestras, participación en los análisis y en la redacción.

3. Spiders on a hot volcanic roof: Colonization pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae) (en revisión) Vera Opatova & Miquel A Arnedo. *Plos ONE*.

Contribución de la doctoranda: Participación en el diseño del trabajo, recogida de las muestras, trabajo de secuenciación, realización de los análisis y redacción de la primera versión del manuscrito y revisiones posteriores.
Contribución de los otros autores: MA, diseño y supervisión, participación en los análisis y en la redacción.

4. Loosening the belt: How the Hercynian Belt breakup shaped the distribution of the trap-door spider genus *Ummidia* (Araneae, Ctenizidae) in the Western Mediterranean. Vera Opatova, Jason E Bond & Miquel A Arnedo

Contribución de la doctoranda: Participación en el diseño del trabajo, recogida de las muestras, trabajo de secuenciación, realización de los análisis y redacción de la primera versión del manuscrito y revisiones posteriores.
Contribución de los otros autores: MA, diseño y supervisión, recogida de las muestras, participación en los análisis y en la redacción. JB, supervisión de los análisis de modelación de nichos, participación en la redacción.

5. Apéndice - Threatened or Threatening: Inferring the population structure of Iberian endangered funnel-web spider *Macrothele calpeiana* (Araneae, Hexathelidae). Vera Opatova, Kanchon K Dasmahapatra, Miguel-Ángel Ferrández, Miquel A. Arnedo

Contribución de la doctoranda: Participación en el diseño del trabajo, recogida de las muestras, trabajo de secuenciación, realización de los análisis y redacción del apéndice.
Contribución de los otros autores: MA, diseño y supervisión, participación en los análisis y en la redacción. KD, supervisión de los análisis de los AFLPs. MAF, recogida de las muestras.

Por otro lado, dejo constancia que ningún material de estas publicaciones será utilizado en ninguna otra tesis doctoral.

Barcelona, 28 de Abril 2014 Miquel A. Arnedo Lombarte

Chapter 1

Vera Opatova, Jason E. Bond, & Miquel A. Arnedo (2013) Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae)

Orígenes antiguos de las arañas migalomorfas Mediterráneas de la familia Ctenizidae (Araneae, Mygalomorphae)

Resumen

La familia Ctenizidae es un grupo de arañas migalomorfas con distribución mundial, un modesto número de géneros y especies y una biogeografía interesante. Su monofilia ha sido cuestionada des de la base de evidencia morfológica hasta la molecular. La familia está representada en la Cuenca Mediterránea por tres géneros y nueve especies: *Cteniza* y *Cyrtocarenum*, en su mayoría endémicos de la región; y *Ummidia*, un género considerado durante mucho tiempo una introducción antropogénica al Mediterráneo debido a que la mayoría de su diversidad se encuentra en el Nuevo Mundo. El estatus taxonómico de algunas de las especies y géneros (por ejemplo, especies Mediterráneas de *Ummidia* o géneros *Cteniza* y *Cyrtocarenum*) ha sido puesto en duda debido a sus similitudes morfológicas. En este estudio, utilizamos un enfoque multilocus que combina los datos de secuencias de ADN de tres genes nucleares, *rRNA 28S*, *EF1γ* y *H3* para investigar los orígenes y la posición filogenética de los taxones Mediterráneos en el contexto de la diversidad al nivel genérico de los ctenízidos. Por primera vez, *todos* los géneros conocidos la familia Ctenizidae se han incluido en análisis filogenéticos. Los métodos bayesianos de estima de edades implementando el reloj molecular relajado y las tasas de sustitución específicas se utilizan para inferir el marco temporal de la diversificación del grupo. Nuestros resultados no están de acuerdo con la división tradicional de la familia Ctenizidae y establecen que las dos subfamilias son polifiléticas, destacando la necesidad de una re-evaluación de los caracteres morfológicos que se han utilizado en la clasificación del grupo. Las estimas de tiempos de divergencia indican un origen antiguo y una larga historia evolutiva de los ctenízidos mediterráneos. La actual distribución discontínua de *Ummidia* se originó probablemente debido a la formación del Océano Atlántico, sugiriendo una distribución previa por Laurasia. Esta hipótesis está apoyada también por la presencia fósiles del dicho género en el ámbar del Báltico. Del mismo modo, la formación de la

cuenca mediterránea occidental probablemente ha jugado un papel clave en la diversificación de *Ummidia* y *Cteniza*, mientras que el origen de las especies del género *Cyrtocarenum* precedió a la desintegración de la antigua masa de tierra continua que formaba la región del Egeo. Tiempos de divergencia profundos y una monofilia recíproca entre *Cteniza* y *Cyrtocarenum* apoyan el estatus de linajes evolutivos independientes. Alternativamente, el estatus taxonómico de *Ummidia* con respecto al género relacionado *Conothele* sigue sin resolver. Se necesita una muestra más completa de estos últimos géneros para evaluar la posible sinonimia de los dos.

Molecular Phylogenetics and Evolution 69 (2013) 1135–1145

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae)

Vera Opatova [a,b,*], Jason E. Bond [b], Miquel A. Arnedo [a]

[a] Institut de Recerca de la Biodiversitat (IRBio), Departament de Biologia Animal, Universitat de Barcelona (UB), Av. Diagonal 643, 08028 Barcelona, Spain
[b] Department of Biological Sciences and Auburn University Museum of Natural History, Auburn University, Auburn, AL 36849, USA

ARTICLE INFO

Article history:
Received 15 February 2013
Revised 31 July 2013
Accepted 2 August 2013
Available online 13 August 2013

Keywords:
Phylogeny
Biogeography
Laurasia breakup
Mediterranean tectonics
Bayesian time estimates

ABSTRACT

The family Ctenizidae is a worldwide-distributed trapdoor spider group, with a modest number of genera and species but interesting biogeography. Its monophyly has been questioned on the basis of both morphological and molecular evidence. The family is represented in the Mediterranean Basin by three genera and nine species: *Cteniza* and *Cyrtocarenum*, mostly endemic to the region, and *Ummidia*, long considered an anthropogenic introduction to the Mediterranean because the bulk of its diversity is in the New World. The taxonomic status of some of the species and genera (e.g. Mediterranean *Ummidia* species or *Cteniza* and *Cyrtocarenum*) has been called into question due to their close morphological affinities. Here, we use a multilocus approach that combines DNA sequence data from three nuclear genes *28S rRNA*, *EF1*γ and *H3* to investigate the origins and phylogenetic position of the Mediterranean taxa within the context of ctenizid generic-level diversity. For the first time, *all* known ctenizid genera are included in a phylogenetic analysis. Additionally, Bayesian relaxed clock methods and specific substitution rates are used to infer the timing of the group's diversification. Our results disagree with the traditional division of the family Ctenizidae into two subfamilies finding them polyphyletic and stress the need for re-evaluating the morphological characters that have been used in the group's classification. Time estimates indicate an ancient origin and long history of Mediterranean ctenizids. The present day disjunct distribution of *Ummidia* seems to be the result of the opening of the Atlantic Ocean, suggesting a former Laurasian distribution of the genus that is further supported by Baltic amber fossils. Similarly, the opening of the western Mediterranean Basin has likely played a key role in the diversification of both *Ummidia* and *Cteniza*, whereas the origin of *Cyrtocarenum* species preceded the breakup of the former continuous landmass that formed the Aegean region. Deep divergence times and reciprocal monophyly support the status of *Cteniza* and *Cyrtocarenum* as independent evolutionary lineages. Alternatively, the taxonomic status of *Ummidia* with regard to the closely related genus *Conothele* remains unclear; a more thorough sampling of the latter is needed to evaluate whether the synonymy of the two genera is necessary.

1. Introduction

Geological processes such as continental level vicariance coincident with long-range dispersal events have greatly contributed to the present day distribution of many extant organismal groups (e.g. McCarthy, 2003; Michalak et al., 2010; Sanmartin et al., 2001). Ancient lineages with narrow habitat preferences and low vagility are more likely to mirror the fingerprint of geological history (Bauzà-Ribot et al., 2012; Stock, 1993), although rare long distance dispersal events have been reported even in unlikely cases (Rowe et al., 2010).

The present day richness of the Mediterranean Basin, defined as one of twenty-five global biodiversity hotspots, is likely the consequence of the region's geological and climatic history (Myers et al., 2000). The formation of the Mediterranean traces back to the upper Eocene, about 40 million years ago (Ma), when the movement of the African plate towards the Eurasian plate precipitated the closure of the Thetis Ocean (Blondel and Aronson, 1999) (Fig. 1). The collision between the two tectonic plates catalyzed the Alpine orogenic process, which continued through to the Miocene (Schellart, 2002), and drove the formation of new basins that separated formerly continuous land masses (Rosenbaum et al., 2002a, 2002b, 2002c). The drop in sea level following the closure of the Strait of Gibraltar (about 5.96 Ma) and subsequent nearly complete desiccation of the Mediterranean Basin (i.e., the Messinian salinity crisis) created land bridges that connected formerly isolated areas (e.g., North Africa with Iberia, Sicily and Italy, and Corsica with

* Corresponding author at: Institut de Recerca de la Biodiversitat (IRBio), Departament de Biologia Animal, Universitat de Barcelona (UB), Av. Diagonal 643, 08028 Barcelona, Spain. Fax: +34 934035740.
E-mail address: vera.opatova@gmail.com (V. Opatova).

1055-7903/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.08.002

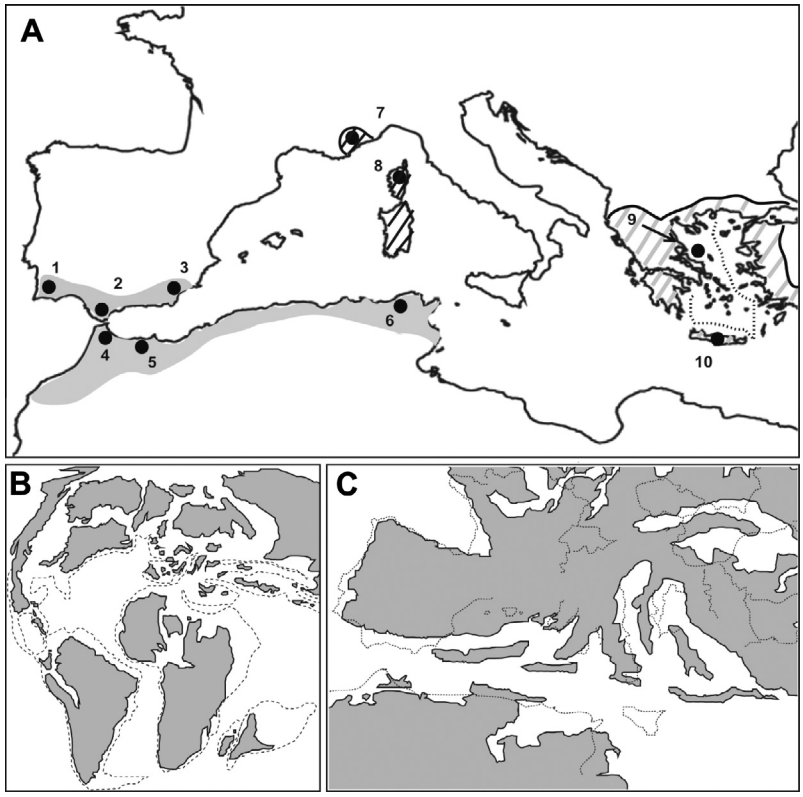


Fig. 1. Map of the Mediterranean region (A) with the location of the sampled specimens and distribution ranges of the Mediterranean ctenizid genera (adapted from Decae (2011)). Distribution of *Cteniza*: black striped, *Cyrtocarenum*: grey striped, *Ummidia*: solid grey. Locality numbers correspond to the following species, 1: *Ummidia algarve*, 2: *Ummidia* sp. Tarifa, 3: *Ummidia picea*, 4: *Ummidia aedificatoria*, 5: *Ummidia* sp. Morocco, 6: *Ummidia algeriana*, 7: *Cteniza moggridgei*, 8: *Cteniza sauvagesi*, 9: *Cyrtocarenum cunicularium*, 10: *Cyrtocarenum grajum*. Dotted lines indicate the location of the Mid-Aegean Trench (MAT), after Lymberakis and Poulakakis (2010). Paleogeographic reconstructions of geographic areas discussed in the text: (B) a Mollewide globe corresponding to the Late Cretaceous (90 Ma), redrawn from Blakey (2008) and (C) Eocene (50 Ma) reconstruction of the Mediterranean region, redrawn from map available at http://cpgeosystems.com (credits to Ron Blakey, Colorado Plateau Geosystems, Inc.).

Sardinia). The reopening of the Strait of Gibraltar 5.3 Ma (Jolivet et al., 2006; Krijgsman et al., 1999) re-established previous conditions.

Paralleling the geological events described above, the Mediterranean also experienced important climatic shifts. The climate at the beginning of Eocene (~55 Ma) was warmer and wetter than today. Following the optimal Eocene climatic maximum, a long-term trend towards lower temperatures and dryer conditions began (Zachos et al., 2001), eventually leading to the establishment of the double seasonality pattern that characterizes present-day Mediterranean climate beginning about 3.4 Ma (Jiménez-Moreno et al., 2010; Suc, 1984). Changes in climatic conditions resulted in an increase of seasonality and aridity (Bruch et al., 2011; Eronen et al., 2009) and was accompanied by a wave of extinctions of terrestrial and aquatic life forms (Böhme, 2003). The growth of the first ice sheets in the Northern Hemisphere marked the start of the Quaternary glacial cycles, which in the Mediterranean region began between 2.8 and 2.5 Ma (Gibbard et al., 2010). During these 100,000 – yearlong cycles, the Mediterranean Basin provided refugia for thermophilic species (Blondel and Aronson, 1999).

The family Ctenizidae belongs to the spider infraorder Mygalomorphae, one of the three main lineages recognized within spiders (Hedin and Bond, 2006). The mygalomorph spiders retain some characters considered primitive among spiders, such as four book lungs, unsynchronized movement of the chelicerae with longitudinal fangs and the lack of specialization in spinning structures (Raven, 1985).

The family Ctenizidae is widespread across North and South America, the Mediterranean region, the South of Africa, South-East Asia and Australia. In spite of its worldwide distribution, ctenizids comprising 128 species placed among nine genera rank among the less diverse spider families (Platnick, 2012). These mid-size, ground-dwellers construct underground silk lined burrows that open to the surface with a trapdoor. The trapdoor is occasionally covered with leaf litter layer, which makes it very difficult to detect (Bond and Coyle, 1995; Gertsch and Wallace, 1936; Gertsch and Platnick, 1975; Hunt, 1976). In some cases, a second inner trapdoor has been reported (Decae, 2010; Saunders, 1842; Siliwal et al., 2009). Females are long-lived and sedentary while males leave their burrows in search of females after the adult moult and hence mediate gene flow between populations. Low vagility has commonly been cited as the main reason for the high level of local endemism in mygalomorph populations and species (Bond et al., 2001; Bond and Stockman, 2008). In *Cyrtocarenum cunicularium* (Olivier, 1811), for instance, offspring usually construct their burrows only a short distance away from their mother's, forming aggregates where up to seven different generations may overlap (Decae et al., 1982). Dispersal through ballooning, which is very common in araneomorph spiders (Duffey, 1956, 1997) but almost unknown in mygalomorphs, has been observed repeatedly in the ctenizid genus *Ummidia* (Coyle et al., 1985; Coyle, 1985; Eberhard, 2006). The efficiency of ballooning dispersal even for relatively long distances is demonstrated by the presence of *Ummidia* species in some Caribbean Islands of volcanic origin without previous connection to any landmass (Platnick, 2012).

The family Ctenizidae was traditionally divided into two subfamilies on the basis of morphological characters: Ctenizinae and Pachylomerinae (Raven, 1985), the latter name subsequently replaced by Ummidiinae (Ortiz, 2007). The Ummidiinae comprises three genera: *Conothele*, *Hebestatis* and *Ummidia*, and is mainly

defined by the presence of a saddle depression on the tibia III that presumably has an anchoring function in the burrow (Coyle, 1981; Gertsch, 1979). Recently, the genus *Hebestatis* has been removed from the group due to some morphological differences (Decae, 2010). The second subfamily Ctenizinae, which contains the remaining six genera: *Cteniza*, *Cyrtocarenum*, *Cyclocosmia*, *Bothriocyrtum*, *Latouchia*, *Stasimopus*, lacks any synapomorphic character and thus its monophyly is doubtful (Raven, 1985). The monophyly of Ctenizidae was not recovered in recent molecular phylogenetic studies of the entire infraorder Mygalomorphae, but no alternative placement of the genera in the phylogeny received substantial support either (Ayoub et al., 2007; Bond et al., 2012; Hedin and Bond, 2006).

The family Ctenizidae is represented by three genera with non-overlapping distributions in the Mediterranean region (Fig. 1). *Ummidia* (Fig. 2) is restricted to the southwestern Mediterranean, occurring on the Iberian Peninsula and in North Africa; *Cteniza* is distributed in the Maritime Alps near the border between France and Italy, and in Corsica and Sardinia; *Cyrtocarenum* extends through the Peloponnesus, smaller Aegean Islands, Crete and Turkey. The genus *Cteniza* and *Cyrtocarenum* are mostly endemic to the region, whereas the bulk of *Ummidia* diversity is in the New World (Platnick, 2012).

Despite the description of three different species of the genus *Ummidia* in the Mediterranean region, namely: *U. aedificatoria* (Westwood, 1840), *U. algeriana* (Lucas, 1846) from North Africa and *U. picea* from southern Spain (Thorell, 1875), the presence of the genus *Ummidia* in Europe was considered a human mediated introduction and the former species synonymized under *U. aedificatoria* (Simon, 1864, 1910). In a recent revision the Mediterranean region's species, all three of the species were re-erected and one additional new species described (Decae, 2010). This new taxonomic arrangement casts doubt on the hypothesis of a recent anthropogenic introduction and provides an alternative explanation for the presence of *Ummidia* in the Mediterranean as a possible relict of a former Laurasian distribution.

The poor morphological diagnosis of some of the other Mediterranean ctenizid genera brings their validity into question. Close morphological affinities between *Cteniza* and *Cyrtocarenum* were already highlighted in the original description of *Cyrtocarenum* (Ausserer, 1871), and subsequent authors have questioned their division (Decae, 1996; Raven, 1985). Similarly, *Aepycephalus*, once considered a distinct genus by several authors (Ausserer, 1871, 1875; Decae, 1996; Raven, 1985; Simon, 1864), was synonymized with *Cteniza* (Wunderlich, 1995).

Fig. 2. *Ummidia algeriana* adult female from Tunisia: PN El Feija, N 36.50705 E 8.32080, 829 m.

Although there is extensive information on the natural history, behaviour, and hunting habits, described both from direct field observation and captivity (Bond and Coyle, 1995; Coyle, 1981; Decae et al., 1982; Moggridge, 1873), as well as the morphology of the burrow (Bond and Coyle, 1995; Buchli, 1962; Coyle, 1981; Decae, 2010) in *Ummidia*, (Decae et al., 1982; Saunders, 1842) *Cyrtocarenum* and (Moggridge, 1873) *Cteniza*, phylogenetic relationships among these genera and their position in the ctenizid phylogeny remains unknown.

The main goal of this study is to provide novel insights into the phylogeny of the family Ctenizidae and infer a timeframe for its diversification with the aim of unravelling the origins of the Mediterranean taxa and shedding some light on their taxonomic status. To accomplish these objectives, we have sampled, for the first time reported, all the described genera in the family and have obtained sequence data from three nuclear genes to infer the group's phylogeny.

2. Materials and methods

2.1. Taxonomic sampling

Thirty-four specimens representing all known ctenizid genera were included in this study. Twenty-two specimens sampled from across the eastern and southwestern United States and Mexico were used to represent the distribution of the New World *Ummidia* ranging from Brazil to United States (Platnick, 2012).

Six specimens representing all described Mediterranean *Ummidia* species and two potentially undescribed species (unpublished data) were also included in the analyses. All known species of *Cyrtocarenum* and *Cteniza*, with the exception of Mediterranean *Cteniza brevidens* (Doleschal, 1871) and the central Asian *Cteniza ferghanensis* Kroneberg, 1875, were analyzed (Fig. 1). Two North American specimens represented the genus *Cyclocosmia* and single specimens from the remaining genera (*Bothriocyrtum*, *Conothele*, *Hebestatis* and *Stasimopus*) were sampled. Four outgroups were chosen on the basis of a recently published mygalomorph phylogeny (Bond et al., 2012). Three of them are placed in the Domiothelina clade that includes Ctenizidae, Migidae (*Migas*), Idiopidae (*Segregara*), and Euctenizidae (*Myrmekiaphila*). A representative of the family Nemesiidae (*Brachythele*), placed in the clade Crassitarsae, was used as outgroup based on the sister-group relationship of this clade to Domiothelina.

Specimens were collected live in the field, immediately transferred to the 100% EtOH and stored at −20 °C for the following DNA extraction. The authors collected most of the material, however, some of the specimens were already utilized in previous phylogenetic studies of the suborder Mygalomorphae (Bond and Hedin, 2006; Bond et al., 2012; Hedin and Bond, 2006) and some were kindly donated by colleagues. Specimens are deposited at following institutions: *Centre de Recursos de Biodiversitat Animal*, University of Barcelona (CRBA, www.ub.edu/crba), Spain and Auburn University Museum of Natural History, Auburn, AL, USA (AUMNH).

2.2. DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from one or two legs per specimen depending on the size of the individual. Extractions were conducted using the SpeedTools Tissue Extraction Kit (Biotools) or DNeasy Tissue Kit (Quiagen) following the manufacturer's protocol. Partial fragments of nuclear genes 28S rDNA (*28S*), Elongation

factor-1 gamma (*EF1γ*) and Histone H3 (*H3*) were amplified using the following primer combinations. The nuclear ribosomal *28S* was amplified in two fragments with ZX1/ZR2 and ZR1/AS8, or alternatively with ZX1/rd5b and rd4.8a/rd7b1. In some cases internal sequencing primer pairs 28SO/28SB, AS3/AS6, ZR3/ZR4 and AS7/AS8 were used to obtain better overlap in the sequence (primer sequences available in Giribet et al. (1999), Hedin and Maddison (2001), Mallatt and Sullivan (1998), Schwendinger and Giribet (2005) and Winchell et al. (2002)). A fragment of the *EF1γ* was amplified using the primers described in Ayoub et al. (2007). First, a pre-amplification was carried out using a "touchdown" amplification strategy with primers EF1gF78/EF1gR1258 and then a second "nested" amplification was performed with the primer combinations EF1gF218/EF1gR1090, EF1gF179/EF1gR1168 or EF1gF179/EF1gR856. Histone H3 (*H3*) was amplified with the primer combination H3a F/H3a R (Colgan et al., 1998).

The general PCR conditions were as follows: 5 min at 94 °C followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 62–52 °C for 35–60 s and extension for 45–60 s at 72 °C (depending on the fragment size) and final extension at 72 °C for 5 min. *28S* and *H3* yielded successful amplification on higher annealing temperatures; between 62 and 58 °C for 60 s and 35 s for the *28S* and *H3*, respectively. *EF1γ* was successfully amplified at temperatures between 54 and 52 °C. For details about the "touchdown" pre-amplification see Ayoub et al. (2007). All reactions were carried out in total reaction volume of 25 μl 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 1.5 μl of DNA and the amount of *Taq* buffer recommended by the maker. The amplification of *28S* was carried out with the FailSafe™ PCR PreMix Kit and FailSafe™ PCR Enzyme Mix I (Epicentre) according to the manufacturer's instructions adding 0.2 μM of each primer and 1.5 μl of DNA.

PCR products were purified using ExoSAP-IT (USB Corporation) and sequenced in both directions using one of the respective amplification primers and BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The sequencing was performed on ABI 3700 automated sequencer at both the *Centres Científics i Tecnològics* of the University of Barcelona (CCiTUB, www.ccit.ub.edu) and Auburn University sequencing facilities. The chromatograms were assembled and edited in Geneious v. 5.3.6. (Drummond et al., 2010). The sequences were deposited in GenBank under KF471411–KF471500 accession numbers.

2.3. Phylogenetic analysis

The alignment of *H3* was trivial, given that the amplified fragment showed no insertions/deletions and hence the sequences were adjusted manually. Both *28S*, and *EF1γ* fragments had some length polymorphism due to indel mutations. Sequence alignment of *28S* was undertaken using the online version of MAFFT v. 6 (available at http://mafft.cbrc.jp/alignment/server/, Katoh and Toh, 2008) using the Q-INS-i strategy, which considers the secondary structure of RNA, with default settings (gap opening penalty, GOP set to 1.53; offset value set to 0.0). The online version of multiple alignment program TranslatorX (available at http://www.translatorx.co.uk/, Abascal et al., 2010) was used to build the nucleotide alignment of *EF1γ* using the amino acid information.

Preliminary Maximum Likelihood analyses were performed in RaxML v.7.2.8 (Stamatakis, 2006) separately on each gene fragment under the GTR + G evolutionary model. The best maximum likelihood tree per each gene was selected from 100 iterations and resampled by 1000 bootstrap replicates. The preliminary results revealed similar topologies in supported clades (results shown as Supplementary data S1–S3) suggesting no major incongruences among the genes. All three genes were concatenated in a single matrix for subsequent analyses using Geneious.

The best partitioning scheme and corresponding substitution models were selected with the help of the program PARTITIONFINDER (Lanfear et al., 2012). The best partition strategy was selected from 37 alternative combinations obtained by regrouping an original fully partitioned data set, that is, partitioned by gene and by codon (in the case of the two protein coding genes), using a greedy algorithm.

Maximum Likelihood (ML) analyses were conducted in RaxML v.7.2.8 (Stamatakis, 2006), using the six-partition scheme and defining unlinked GTR + G model for each partition. The best maximum likelihood tree was selected from 1000 iterations, each starting with distinct randomized parsimony trees. Bootstrap support was inferred from 400 replicates, as selected by the MRE-based stopping criterion (autoMRE option) (Pattengale et al., 2010).

The Approximately Unbiased (AU) topology test (Shimodaira, 2002) as implemented in the program CONSEL (Shimodaira and Hasegawa, 2001) was used to test statistical significance of alternative topologies (see below).

Bayesian inference analyses were conducted with MrBayes v. 3.1.2 (Ronquist and Huelsenbeck, 2003) and run remotely at the Bioportal computer resource of the University of Oslo (http://www.bioportal.uio.no/). The combined data matrix was partitioned and models assigned as suggested by PARTITIONFINDER. Two independent runs of 5×10^7 generations with 8 MCMC (Markov Chain Monte Carlo) chains each were conducted simultaneously, starting from random trees and resampling each every 1000 generations. The standard deviation of split frequencies between runs (<0.01) and the effective sample size (ESS) as measured by the program TRACER version 1.5 (Rambaut and Drummond, 2009), were monitored to ensure stationarity, convergence and correct mixing of the chains and to determine the correct number of generations to discard as a burn-in for the analyses (first 10%). All trees were visualized and manipulated in the program FigTree v. 1.3.1 (available at http://tree.bio.ed.ac.uk/software/figtree/).

2.4. Estimation of divergence times

Substitution rates for the *EF1γ* in spiders were estimated by Ayoub et al. (2007) to be approximately 0.0011 substitution/sites/million years. However, these estimates were obtained with the NPRS method (Sanderson, 1997), which has been shown to overfit the data (Sanderson, 2002). The estimates were also provided without confidence intervals and did not integrate phylogenetic uncertainty. Therefore, we have here reanalyzed Ayoub et al.'s data using a Bayesian framework with a relaxed clock model as implemented in the program BEAST v.1.7.2 (Drummond et al., 2012). We defined a single partition and a GTR + I + G model; the tree prior was set to follow a Yule speciation model. The same fossil calibration information as used in Ayoub et al. (2007) was incorporated into the analysis with minimum bounds of uniformly distributed priors as follows: Mygalomorphae, 240 Ma, Orbiculariae (Deinopoidea + Araneoidea 170 Ma, Araneoidea (*Argiope argentata* + *Latrodectus Hesperus*) 135 Ma, Deinopoidea (*Deinopis spinosa* + *Uloborus diversus*) 135 Ma, Atypoidina (*Aliatypus plutonis* + *Atypoides riversi* + *Atypus snetsingeri*) 96 Ma (see Ayoub et al., 2007 for a justification of the callibration points). In addition, we included the age of *Attercopus*, 392 Ma (Devonian, Givetian sediments) (Selden et al., 1991), currently classified in Uraraneida, the sister lineage of spiders (Selden et al., 2008), as a maximum bound for all calibration points plus the root. A preliminary tree including all time constraints was obtained with the program STARTTREE (Heath, 2012) and specified as starting tree in the BEAST runs. A uniform prior was assigned to the ucld.mean parameter of the lognormal relaxed clock with lower and upper bounds 0 and 0.0115 (universal substitution rate of arthropod mitochondrial DNA), respectively, and starting value 0.0011. We enforced

well-supported clades recovered in a recent multi-locus phylogenetic analysis, including *EF1γ*, of the suborder Mygalomorphae (Bond et al., 2012) to speed up analyses. Three independent runs of 5×10^7 generations were conducted. Convergence between runs and correct mixing within each run were visualized with TRACER. The resulting tree is shown in Supplementary Fig. S4.

The *EF1γ* substitution rate estimated following the protocol above was further used to infer the timeframe of Ctenizidae diversification. To speed up the computational time and make the output graphically more readable, a reduced version of the original data matrix was obtained by removing most of the New World *Ummidia* samples and retaining single representatives of the main lineages having the most complete sequences. The dataset was split into the preferred partitions and their corresponding evolutionary models. Three relaxed lognormal clocks were defined: one for the *28S* + 1st positions of the *H3*, one for the remaining positions of the *H3* and one for the *EF1γ*. The mean and standard deviation of the estimated substitution rates of *EF1γ* was set to the values obtained in the preliminary analysis and the rates of the other two clocks were included as parameters to estimate. Three independent chains of 5×10^7 generations were run independently. Convergence and correct mixing was assessed as in the former analyses. The BEAST accompanying programs LOGCOMBINER and TREEANNOTATOR were used to combine independent runs following burn-in and to estimate a consensus chronogram.

3. Results

3.1. Phylogenetic analyses

Specimens, locality data and GenBank accession codes of the sequences analyzed in this study are listed in the Table 1, a map showing the location of Mediterranean samples is provided in Fig. 1. The final concatenated matrix of three nuclear genes resulted in 3397 characters, including the nearly complete ribosomal *28S* (2119 characters, 323 variables), and fragments of the *EF1γ* (882, 339 variables) and *H3* (396, 131 variables). Gaps were treated as missing data in subsequent analyses.

The Bayesian information criterion implemented in PARTITIONFINDER selected a scheme with 6 partitions corresponding to each of the protein coding gene codon positions except for the 1st positions of *H3* that was combined with *28S*. The preferred evolutionary models for each partitions were: GTR + I + G for the *28S* + 1st positions of *H3*, JC and K81 + I + G for the 2nd and 3rd positions of *H3*, respectively, and K80 + G, HKY + I + G and K80 + G for the for 1st, 2nd and 3rd positions of the *EF1γ*, respectively.

Both maximum likelihood ($-\ln L = 14260.890357$) and Bayesian analyses recovered highly similar tree topologies. The tree obtained in Bayesian analysis with both Bayesian posterior probabilities and ML bootstrap supports, is shown in Fig. 3. Overall, Bayesian clades supports were slightly higher than ML ones.

None of the methods recovered the monophyly of family Ctenizidae. Bayesian analyses split Ctenizidae in three distinct lineages, one corresponding to *Stasimopus*, which is the sister group of the remaining lineages, a second one including *Cteniza* and *Cyrtocarenum*, more closely related to non-ctenizid lineages (the idiopid *Segregara* and the euctenizid *Myrmekiaphila*), and a third lineage including all remaining ctenizids. In general, basal relationships were poorly supported. Maximum likelihood put *Stasimopus* as sister to the larger ctenizid clade, albeit with low support. The AU topology test, however, could not reject the monophyly of Ctenizidae.

Cteniza and *Cyrtocarenum* formed a clade with high support in both analyses, as did *Conothele* and *Ummidia*, and *Bothriocyrtum* and *Hebestatis*. The genus *Ummidia* was recovered as monophyletic in both analyses, albeit only supported in Bayesian analysis, and further divided into Mediterranean and North American clades. Although Bayesian and maximum likelihood analyses fully agree in the internal resolution of *Ummidia*, support values for the main lineages were low. Within the Mediterranean clade, *U. algarve* was recovered as the sister group to the remaining species, and *U. picea* as sister to the North African lineages. The North American clade showed two deeply divergent lineages.

Bayesian analyses provided further support for the sister clade relationship of *Latouchia* to the *Conothele* + *Ummidia* clade, and the inclusion of both lineages along with *Bothriocyrtum*, *Hebestatis* and *Cyclocosmia* in the same clade. Maximum likelihood recovered the same clades albeit with low support.

3.2. Divergence time estimation

Reanalysis of the data under a Bayesian relaxed clock framework results in slightly younger ages that those previously reported by Ayoub et al. (2007). The estimated ucld.mean was 0.00117 (0.00094–0.0014) and the ucld.stdev = 0.388 (0.28–0.51). These values were incorporated into the divergence time analysis of Ctenizidae as normal priors for both the ucld.mean (mean = 0.00117, stdev = 0.00014) and ucld.stdev (mean = 0.388, stdev = 0.08) of the lognormal clock of the combined *EF1γ* partitions. The resulting chronogram is shown in Fig. 4. The estimated time of split between the Mediterranean and the North American *Ummidia* lineages was 50.12 Ma (34.04–69.72 Ma), and between *Cteniza* and *Cyrtocarenum* 74.73 Ma (45.7–109.5). The time of the most common ancestor was dated to 41.99 Ma (27.8–60) for Mediterranean *Ummidia*, 43.35 Ma (28.1–61.7) for the North American lineage, 30.61 Ma (14–52.8) for *Cteniza* and 61.81 Ma (34.2–94.6) for *Cyrtocarenum*.

The BEAST analyses provided further support for some of the phylogenetic relationships, including the monophyly of *Ummidia* and the sister group relationship of *U. algarve* to the remaining Mediterranean *Ummidia* lineages.

4. Discussion

4.1. Origins and phylogenetic relationships of Mediterranean ctenizids

The occurrence of the genus *Ummidia* in Europe was traditionally hypothesized as a human mediated introduction from the New World, most likely transported with soil (Simon, 1864, 1910). This assumption served as ground for earlier suggestions to synonymize available species of Mediterranean *Ummidia* species under the single name *U. aedificatoria* (Simon, 1910). Almost 100 years elapsed before a full revision of the Mediterranean *Ummidia* taxa was completed, which reinstated the original species and added a new one. These results suggest that the present-day distribution of *Ummidia* both in the Mediterranean and North America is either the remnant of a former Pan-Laurasian distribution or, alternatively, that *Ummidia*-like phenotypes evolved independently on both sides of the Atlantic ocean, and hence that the genus is polyphyletic (Decae, 2010).

Our results are in agreement with Decae's taxonomic revision. The diversity of lineages and their deep divergences, which trace back its origin to the Palaeogene, suggest a long-time presence of *Ummidia* in the Mediterranean region and rules out a more recent contemporary human mitigated introduction.

Concerning the origin of the genus *Ummidia*, our results recover its monophyly and basal split into a Mediterranean and a North American clade. The split between these two main lineages occurred during the Palaeogene (~50 Ma), which is roughly concordant with the final split of the Laurasian supercontinent (Fig. 1),

Table 1
Specimen locality data and Genbank accession numbers.

Family	Genus	Species	Country	Locality	Lat/long	code	28S	EF-1γ	H3
Ctenizidae	*Bothriocyrtum*	*californicum*	USA	California, Lake Mathews	N33.8266 W117.438	MY2277	KF471414	KF471444	KF471473
Ctenizidae	*Conothele*	sp.	Australia	Western Australia, SE of Nungarin	S31.29432 E118.19944	MY2070	DQ639909	JQ358734	KF471470
Ctenizidae	*Cteniza*	*moggridgei*	France	Monti, Sospel	N43.82472 E7.48162	Z627	KF471432	KF471462	KF471494
Ctenizidae	*Cteniza*	*sauvagesi*	France	Corsica, Coll de Palmarela	N42.36393 E8.64826	Z777	KF471436	KF471466	KF471498
Ctenizidae	*Cyclocosmia*	sp.	USA	Alabama, Lee Co.	N32.5786 W854543	AUMS139	KF471412	KF471440	
Ctenizidae	*Cyclocosmia*	*truncata*	USA	Alabama, Sipsey WA	N34.3409 W87.4710	MY457	DQ639903		
Ctenizidae	*Cyrtocarenum*	*cunicularium*	Greece	Crete, Gorgolaini	N35.20581 E24.9863	Z715	KF471435	KF471465	KF471497
Ctenizidae	*Cyrtocarenum*	*grajum*	Greece	Sporades, Alonnisos	N39.12824 E23.56266	Z713	KF471434	KF471464	KF471496
Ctenizidae	*Hebestatis*	*theveneti*	USA	California, N Mariposa	N37.5039 W119.9940	MY278	DQ639905	JQ358743	
Ctenizidae	*Latouchia*	sp.	Taiwan	Nantou		Z821	KF471434		KF471500
Ctenizidae	*Stasimopus*	*mandelai*	South Africa	Eastern Cape, Great Fish River Reserve	S33.12755 E26.67287	MY557	DQ639901	JQ680321	KF471486
Ctenizidae	*Ummidia*	*algarve*	Portugal	Lagos/Odiaxere junction	N37.13945 W8.67204	Z597	KF471431	KF471461	KF471493
Ctenizidae	*Ummidia*	*algeriana*	Tunisia	El Feija National Park	N36.50709 E8.3208	Z817	KF471437	KF471467	KF471499
Ctenizidae	*Ummidia*	*aedificatoria*	Morocco	M.F.B. Bellota	N34.94961 W5.52892	Z584	KF471430	KF471460	KF471492
Ctenizidae	*Ummidia*	*picea*	Spain	Almeria, Vera	N37.21327 W1.82724	Z121	KF471427	KF471458	KF471489
Ctenizidae	*Ummidia*	sp.	Mexico	El Mezquitito hills	N24.100305 W110.269416	MY2695	KF471421	KF471452	KF471481
Ctenizidae	*Ummidia*	sp.	Mexico	El Triunfo watercourse	N23.80194 W110.11833	MY2696	KF471422	KF471453	KF471481
Ctenizidae	*Ummidia*	sp.	Morocco	Bni-Hadifa	N35.00545 W4.17899	Z582	KF471429		KF471491
Ctenizidae	*Ummidia*	sp.	Spain	Tarifa, Sierra de Fates	N36.12643 W6.4418	Z668	KF471433	KF471463	KF471495
Ctenizidae	*Ummidia*	sp.	USA	Alabama, Butler Co.	N31.56635 W86.74021	MY2536	KF471418	KF471449	KF471478
Ctenizidae	*Ummidia*	sp.	USA	Alabama, Lee Co.	N32.5786 W854543	AUMS116	KF471411	KF471439	KF471468
Ctenizidae	*Ummidia*	sp.	USA	Arizona, Lake Charles State Park	N36.06459 W91.14906	MY2815	KF471425	KF471456	KF471485
Ctenizidae	*Ummidia*	sp.	USA	Florida, Santa Rosa	N30.95616 W87.31464	MY2549	KF471419	KF471450	KF471479
Ctenizidae	*Ummidia*	sp.	USA	Florida, Santa Rosa	N30.95616 W87.31464	MY2550	KF471420	KF471451	KF471480
Ctenizidae	*Ummidia*	sp.	USA	Kentucky, Olaton	N375325 W86.7296	MY2041	KF471413	KF471441	
Ctenizidae	*Ummidia*	sp.	USA	North Carolina, Durham Duke forest	N36.1 W78.96	MY653	KF471426	KF471457	KF471488
Ctenizidae	*Ummidia*	sp.	USA	North Carolina, Midland	N35.21800 W80.57670	MY2042	DQ639907	KF471442	KF471469
Ctenizidae	*Ummidia*	sp.	USA	North Carolina, Weyerhauser Cool Springs	N35.1825 W77.0754	MY2310	KF471415	KF471445	KF471474
Ctenizidae	*Ummidia*	sp.	USA	North Carolina, Weyerhauser Cool Springs	N35.1825 W77.0754	MY2311	KF471416	KF471446	KF471475
Ctenizidae	*Ummidia*	sp.	USA	North Carolina, Weyerhauser Cool Springs	N35.1825 W77.0754	MY2312	KF471417	KF471447	KF471476
Ctenizidae	*Ummidia*	sp.	USA	North Carolina, Weyerhauser Cool Springs	N35.1825 W77.0754	MY2313	DQ639908	KF471448	KF471477
Ctenizidae	*Ummidia*	sp.	USA	South Carolina, Sumter National Forest	N34.08 W82.25	MY2784	KF471424	KF471455	KF471484
Ctenizidae	*Ummidia*	sp.	USA	Tennessee, Sequatchie Co.	N35.39514 W85.39662	MY2716	KF471423	KF471454	KF471483
Cyrtaucheniidae	*Myrmekiaphila*	*fluviatilis*	USA	Virginia, Cascades Rec Area	N37.3538 W80.5999	MY2234	DQ639889	KF471443	KF471472
Idiopidae	*Segregara*	sp.	South Africa	Western Cape Province	S32.3347 E22.4747	MY604	DQ639922	DQ680329	KF471487
Migidae	*Migas*	*variapalpus*	Australia	Queensland, Lamington National Park	S28.19347 E153.18722	MY2104	DQ639895	JQ358749	KF471471
Nemesiidae	*Brachythele*	sp.	Greece	Stomio	N39.85972 E22.73722	Z33	KF471424	KF471459	KF471490

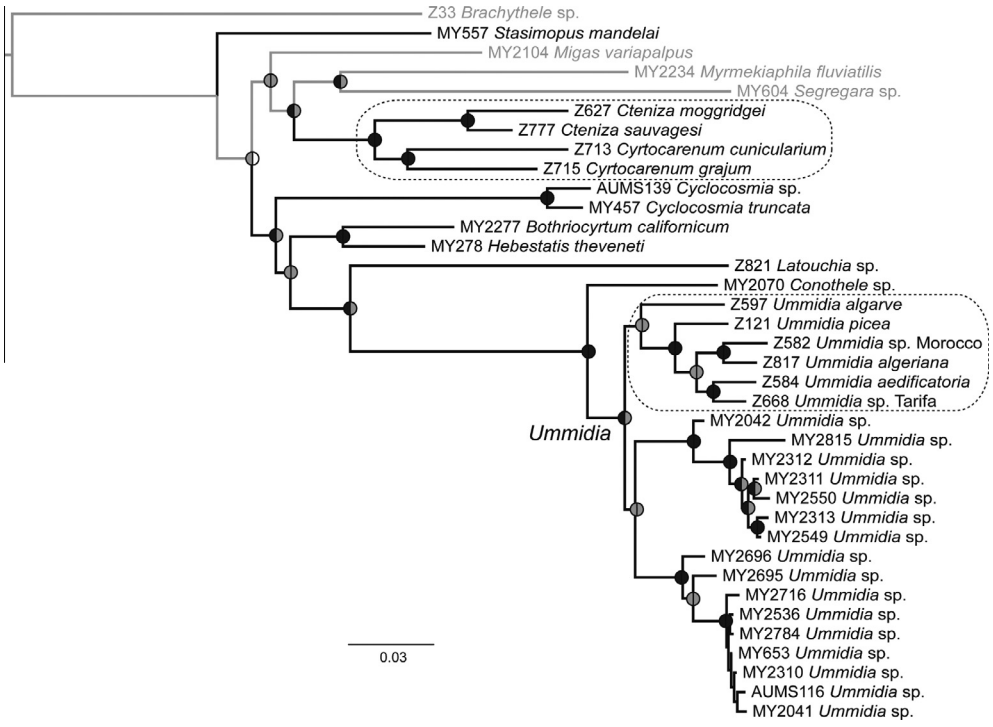


Fig. 3. Topology obtained in the Bayesian analyses. Outgroup taxa are in grey. Mediterranean ctenizids are highlighted in boxes. Dots on nodes denote support as follows: left semi-circle are Bayesian posterior probabilities (PP) and right ones maximum likelihood bootstraps, black = PP > 0.95, ML bootstrap support >80%, grey = clade recovered but with support values less than thresholds above, white = topology not recovered.

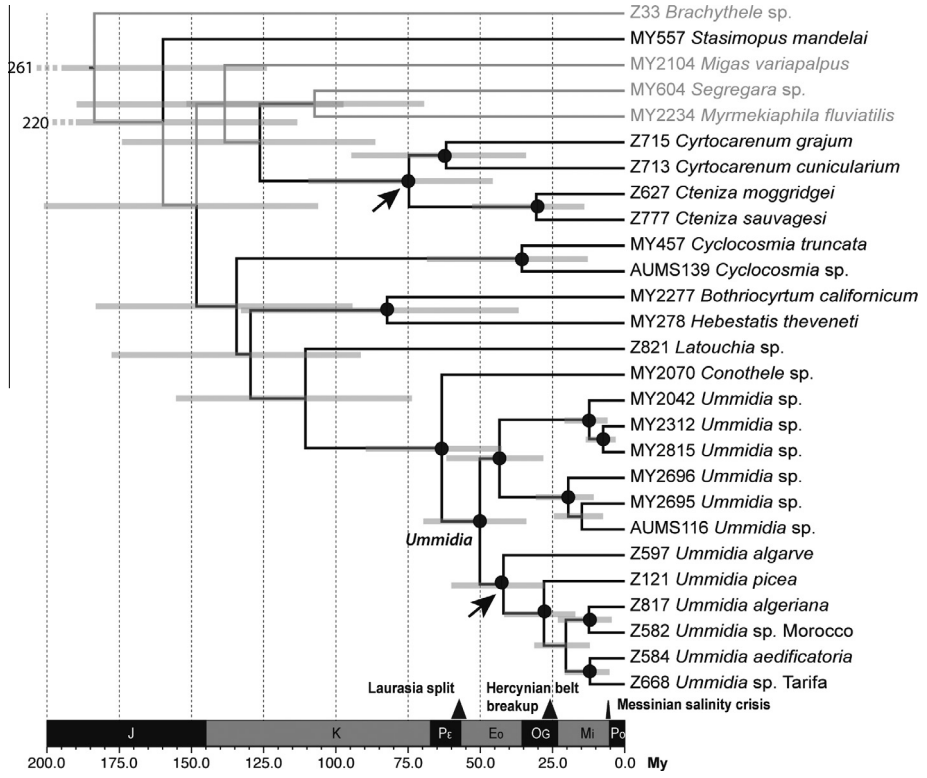


Fig. 4. Chronogram obtained with BEAST. Dots on nodes denote Bayesian posterior probabilities above 0.95. Node bars indicate the 95% HPD confidence intervals of the divergence time. The *x*-axis is time in million years (My). Triangles on the *x*-axis mark relevant geological events that may have affected diversification of Mediterranean ctenizid lineages. J: Jurassic, K: Cretaceous, P_G: Paleogene, E_O: Eocene, O_G: Oligocene, M_I: Miocene, P_O: Plio-Pleistocene. Arrows highlight the divergence times in Mediterranean Ctenizidae. Upper: divergence between the genera *Cteniza* and *Cyrtocarenum*, lower: divergence of Mediterranean clade of the genus *Ummidia.*

dated at the end of the Paleocene, about 60–55 Ma (Sanmartin et al., 2001); deep-water conditions between Iberia and North America may have already existed about 95–110 Ma (Jones et al., 1995). Therefore, according to our data, the present day disjunct distribution of *Ummidia* was the result of the break up of Laurasia and the opening of the Atlantic Ocean. A similar explanation has been put forward to explain the presence of phylogenetic sister lineages of stygobiont amphipods on opposite sites of the Atlantic Ocean (Bauzà-Ribot et al., 2012; Stock, 1993). Divergence time estimates in stygobiont amphipod American and European sister linages are, however, older (77 Ma, 57–101 Ma). Alternatively, *Ummidia* could have colonized one of the two sides of the Atlantic after the opening of the ocean. In this regard, it is interesting to note that *Ummidia* is one of the very few mygalomorph spider genera in which aerial dispersal by ballooning has been reported (Coyle et al., 1985). *Ummidia* may have used ballooning to disperse between the continents once they parted but when the distance between them was still close. Although colonization of new habitats through ballooning surely plays a role in *Ummidia* distribution, the deep population structure and haplotype divergence found among nearby localities in the Mediterranean *Ummidia* species (Opatova and Arnedo, unpublished), indicates that long distance dispersal events as required for the colonization across the opening Atlantic Ocean even through stepping stones or land bridges were probably rare. Moreover, land connections between the Eastern American coast and Europe existed via the North American land bridge until 25 Ma. In spite of the oscillating suitability for dispersal, this land bridge is hypothesized to have played an important role in plant dispersal (for review see Milne, 2006; Tiffney and Manchester, 2001; Tiffney, 1985a, 1985b).

The existence of fossil *Ummidia* taxa in Baltic amber (Wunderlich, 2000, 2004, 2011), which has been dated at about the Upper Eocene (44–47 Ma) (Ritzkowski, 1997), provides further support for the long history of *Ummidia* in Europe.

Our divergence time estimates among the main lineages of Mediterranean *Ummidia* overlap with a phase of geological turmoil in the region – the opening of the western Mediterranean Basin as a result of the breaking up of the Hercynian orogenic belt starting about 30 Ma (Rosenbaum et al., 2002a) (Fig. 1). Interestingly, north African *Ummidia* species are restricted to the Rif of Morocco and the Tell Atlas, including the Kabylia, of Algeria and Tunisia. Both regions formed part of the Hercynian block that broke off at the beginning of the Oligocene, and subsequently drifted to their present day positions. The Iberian *U. picea* and the group comprising the North African representatives (*U. algeriana*, *Ummidia* sp. Morocco and *U. aedificatoria*) split about 28 Ma (17.1–41.7), suggesting that regional tectonic vicariance played a role in the diversification of Mediterranean *Ummidia*. The more recent relationship between the *Ummidia* sp. Tarifa, from southern Spain, and the Moroccan *U. aedificatoria* could be explained as reverse colonization of the Iberian Peninsula from Northern Africa. Given the fact that the distance between the two regions is only of 15 km at its narrowest point, the colonization via ballooning is plausible. The closing of the Strait of Gibraltar during the Messinian Salinity Crisis (5.3–5.96 Ma) may have facilitated the back colonization as indicated by the marginal time overlap with the estimated split between *Ummidia* sp. Tarifa and *U. aedificatoria* inhabiting both sides of the Strait (5.3–20.8 Ma). Interestingly, comparing Mediterranean *Ummidia* species with other organisms with similar distributions across the Iberian Peninsula and North Africa, the Pleistocene climatic oscillations leading to species range shifts and subsequent dispersal are identified as the most important factors driving present day distributions (e.g. Paulo et al., 2008; Pinho et al., 2007; Santos et al., 2012).

Phylogenetic relationships and diversity of North American *Ummidia* are beyond the scope of the present study, although several samples representing the known range of the group were included to test relationships with Mediterranean taxa. It is understood, however, that the real diversity of *Ummidia* in the New World is severely underestimated (Bond and Hedin, 2006) and a thorough revision is needed. Our results support this observation, because we identified several lineages whose origin can be traced back at least to the Neogene.

Our findings concerning the sister group relationship between the genera *Cteniza* and *Cyrtocarenum* are in general agreement with their morphological similarity, already recognized by Ausserer (1871) in his original description of the genus *Cyrtocarenum*. In fact, the close morphological resemblance has led some later authors to put their status as separate genera into question (Decae, 1996; Raven, 1985). Surprisingly, our results indicate that the two genera are both ancient and long established in the Mediterranean region, their origins dating back to 75 Ma (45.7–109.5). Here, we argue that the long independent evolutionary history of the two genera, which predated the origin of the amphi-Atlantic disjunct *Ummidia*, provides enough evidence to treat them as independent taxa and hence their remarkable morphological similarities are better treated as a case of morphological stasis (not uncommon among mygalomorph taxa, see Hendrixson and Bond, 2009).

Our estimates trace back the split of the two extant species of *Cyrtocarenum* to the Paleogene (34.2–94.6 Ma), which would rival some of the best known examples of long-lived, single-species lineages, such as horseshoe crabs dated about 45 Ma (30–68) (Obst et al., 2012). Interestingly, *Cyrtocarenum* species had already been described as living fossils based on their distribution pattern (Decae, 1983). Present-day distribution of *Cyrtocarenum* includes the Peloponnese and Attica in Greece, western Turkey and surrounding islands. The two extant species occur syntopically at several localities (e.g., the Ionian Islands, Kythera and Attica). *Cyrtocarenum grajum* is dominant in most of the Peloponnesus and the Sporades, but it is absent from the Cyclades. On the other hand, all specimens hitherto reported from Crete (with one exception), the Cyclades, Anatolia and Rhodes are *C. cunicularium* (Decae, 2011). The region has a complicated geological history frequently invoked to explain phylogenetic relationships and biogeographic patterns of local organisms (e.g. Poulakakis et al., 2008 and references herein). Specifically, disjunct distributions of several organisms across the western and eastern Aegean archipelago have been shown to trace back to the formation of the Mid-Aegean Trench originated in the Upper Miocene (12–9 Ma) (Creutzburg, 1963; Dermitzakis, 1990). Species of *Cyrtocarenum*, however, are found on both sides of the trench, suggesting that they were already present during the early and middle Miocene (23–12 Ma), before the continuous landmass that included the Aegean (Agaïs) broke off (Fig. 1.).

The distribution of the genus *Cteniza* is restricted to the Maritime Alps close to the border between France and Italy and two of the largest Mediterranean islands, Corsica and Sardinia. This interesting biogeographic link between the mainland and the islands and the relationships among *Cteniza* species inhabiting this region have beckoned the attention of naturalists for over a century (Moggridge, 1873). Similar distribution patterns comprising these two islands and adjacent mainland have also been described for other groups with low dispersal capabilities. In most cases, the present day distributions have been attributed to former land connections, either by dispersal through land bridges formed during the Messinian Salinity Crisis (5.3–5.96 Ma) or the Quaternary glacial cycles (2.8–0.02 Ma), or as vicariant events driven by the opening of the Western Mediterranean Basin (e.g. Allegrucci et al., 2005; Caccone and Sbordoni, 2001; Ketmaier et al., 2003, 2006).

According to our findings, *Cteniza sauvagesi* diverged from *C. moggridgei* around 30 Ma (14–53). This time frame matches the

period when the continental microplates forming the Hercynian belt broke up and drifted towards their present-day position (i.e. the opening of the Western Mediterranean Basin) (Fig. 1). Therefore, it could be hypothesized that the ancestor of the two extant *Cteniza* species already inhabited the region by the time the different Hercynian blocks split and that the present day distribution is the result of these vicariant events. The contiguous position between south-eastern France and the blocks that gave rise to Corsica and Sardinia in geological reconstructions (e.g. Cavazza and Wezel, 2003) provide further support for this hypothesis. A similar explanation has been put forward to explain the distribution and diversification of the ground-dwelling spiders of the genus *Parachtes* (Bidegaray-Batista and Arnedo, 2011). However, the reason for the absence of *Cteniza* locality data from other terrains that were once adjacent to the former blocks, such as the Iberian peninsula, the Calabro–Perotinal massif, the Balearic Islands, the Rif or the Kabylies remains unanswered. Further studies combining population level analyses and species distribution modelling techniques will likely shed more light on the causes of the current distribution of this remarkable genus (V. Opatova in prep).

4.2. Phylogenetic structure of the mygalomorph family Ctenizidae

The results reported herein document the first ever molecular phylogeny that includes representatives of all ctenizid genera. Previous studies conducted to resolve the phylogenetic structure of mygalomorph spiders, using some or all of the markers used in the present study, have included a subset of ctenizid representatives (Ayoub et al., 2007; Bond et al., 2012; Hedin and Bond, 2006). None of these studies recovered Ctenizidae as a clade, with a single exception (Bond et al., 2012), where morphological characters were also taken into account. In these studies, non-monophyly generally involved the genus *Stasimopus*. In our study *Stasimopus* was found to form a clade with the bulk of ctenizid diversity only in the ML analysis. On the other hand, the lineage *Cteniza* + *Cyrtocarenum*, which had not been included in former analyses, never clustered with other ctenizids. Unfortunately, topology tests could not reject the single origin of the family and hence the poly- or diphyletic status of Ctenizidae remains to be tested with the addition of more data.

Our results further confirmed the sister-group relationship between *Bothriocyrtum* and *Hebestatis* (Bond et al., 2012; Hedin and Bond, 2006). These findings disagree with the traditional taxonomic division of the family Ctenizidae: *Hebestatis* belonged to subfamily Ummidiinae, while *Bothriocyrtum*, was a Ctenizinae (Raven, 1985), and confirm that the dorsal, glabrous saddle on tibia III of *Hebestatis* is not homologous to that found in other Ummidiinae, as already advanced by Decae (2010). Similarly, *Latouchia*, formerly considered a Ctenizinae, was recovered as sister group of the remaining Ummidiinae (*Ummidia* and *Conothele*).

The close affinity between *Ummidia* and *Conothele* has already been reported in previous studies (Bond et al., 2012; Hedin and Bond, 2006) and the synonymization of the two genera has been suggested on the basis of morphology (Main, 1997). Our taxonomic sampling precludes testing the reciprocal monophyly of *Conothele* and *Ummidia*. However, if lineage relative ages are used as a proxy to define genera, as recent proposals suggest (Talavera et al., 2012), then *Conothele* and *Ummidia* could be considered the same genus because their split is about 15 My younger that the next youngest genus but marginally older than the deepest species split (*Cyrtocarenum*). Regardless of the above speculations, it is obvious that more exhaustive species sampling of *Conothele* across its known distribution range will be required to fully settle the issue of the evolutionary independence of these two genera.

5. Conclusions

The trap-door spiders of the family Ctenizidae have an ancient history in the Mediterranean region and their present day distributions and diversity have been mostly shaped by the dynamic geological history of the region.

The North American and Mediterranean species of the genus *Ummidia* form reciprocal monophyletic groups that probably originated from a former Laurasian ancestor as a result of the opening of the Atlantic Ocean. Similarly, the opening of the western Mediterranean Basin was most likely involved in the diversification of the Mediterranean *Ummidia* and the genus *Cteniza*. On the other hand, speciation in *Cyrtocarenum* seems to have predated the breakup of the former continuous landmass between Greece and Turkey. Despite the morphological similarity between *Cteniza* and *Cyrtocarenum* that led some authors to suggest their synonymization, the species of the two genera were found to form exclusive sister clades with deep divergence time, which supports their status as independent evolutionary lineages. Conversely, the status of *Conothele* remains unclear; more material collected throughout its distributional range would be required to evaluate correctly a possible synonymy with *Ummidia*.

Our results emphasize the need for re-evaluating the morphological characters used in the current taxonomy of the family Ctenizidae. Phylogenetic reconstructions are in direct disagreement with some higher-level groupings but also uncover unexpected amounts of genetic diversity within some of the genera, which may hint to the existence of cryptic species and warrants further investigation.

Acknowledgments

We are thankful to E. Mora, E. Planas, R. Alvarez, L. Bidegaray-Batista, M. Chatzaki, P. Sousa, C. Ribera, N. Garrison and R. Godwin for the assistance during the field collections. We also thank S. Huber for providing us with *Brachythele* sp., M. Metallinou and E. Planas for *C. grajum*, I-Min Tso for *Latouchia* specimens and M. Hedin and B. Hendrixson for *Ummidia* specimens. We would like to thank to A. Decae for the initial guidance of VO with Mediterranean trapdoor spider taxonomy. This project was funded by Spanish Ministry of Science and Innovation (MICINN) Grant CGL 2009-07639 (MA) and additional financial support was provided by Project 2009SGR1462 from the Generalitat de Catalunya; partial support to JB by National Science Foundation Grant DEB 0315160, VO was supported by a graduate Grant (AP 2008-01841) and graduate travelling Grant FPU-2012, both provided by MICINN. MA was further supported by an ICREA Academia award for excellence in research from the Generalitat de Catalunya.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2013.08.002.

References

Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Research 38, W7–13.

Allegrucci, G., Todisco, V., Sbordoni, V., 2005. Molecular phylogeography of *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae): a scenario suggested by mitochondrial DNA. Molecular Phylogenetics and Evolution 37, 153–164.

Ausserer, A., 1871. Beiträge zur Kenntniss der Arachniden-Familie der Territelariae Thorell (Mygalidae Autor). Verhandlungen der Kaiserlich-Koniglichen Zoologisch-Botanischen Gesellschaft in Wien 21, 117–224.

Ausserer, A., 1875. Zweiter Beitrag zur Kenntniss der Arachniden-Familie der Territelariae Thorell (Mygalidae Autor). Verhandlungen der Kaiserlich-Koniglichen Zoologisch-Botanischen Gesellschaft in Wien 25, 125–206.
Ayoub, N.A., Garb, J.E., Hedin, M., Hayashi, C.Y., 2007. Utility of the nuclear protein-coding gene, elongation factor-1 gamma (*EF-1γ*), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). Molecular Phylogenetics and Evolution 42, 394–409.
Bauzà-Ribot, M.M., Juan, C., Nardi, F., Oromí, P., Pons, J., Jaume, D., 2012. Mitogenomic phylogenetic analysis supports continental-scale vicariance in Subterranean Thalassoid Crustaceans. Current Biology 22, 2069–2074.
Bidegaray-Batista, L., Arnedo, M., 2011. Gone with the plate: the opening of the Western Mediterranean Basin drove the diversification of ground-dweller spiders. BMC Evolutionary Biology 11, 317.
Blakey, R.C., 2008. Gondwana paleogeography from assembly to breakup—A 500 m.y. odyssey. The Geological Society of America Special Paper 441, 1–28.
Blondel, J., Aronson, J., 1999. Biology and Wildlife of the Mediterranean Region. Oxford University Press, Oxford, New York.
Böhme, M., 2003. The Miocene Climatic Optimum: evidence from ectothermic vertebrates of Central Europe. Palaeogeography, Palaeoclimatology, Palaeoecology 195, 389–401.
Bond, J.E., Coyle, F.A., 1995. Observations on the natural history of an *Ummidia* trapdoor spider from Costa Rica (Araneae, Ctenizidae). Journal of Arachnology 23, 157–164.
Bond, J.E., Hedin, M., 2006. A total evidence assessment of the phylogeny of North American euctenizine trapdoor spiders (Araneae, Mygalomorphae, Cyrtaucheniidae) using Bayesian inference. Molecular Phylogenetics and Evolution 41, 70–85.
Bond, J.E., Hedin, M.C., Ramirez, M.G., Opell, B.D., 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Molecular Ecology 10, 899–910.
Bond, J.E., Hendrixson, B.E., Hamilton, C.A., Hedin, M., 2012. A reconsideration of the classification of the spider Infraorder Mygalomorphae (Arachnida: Araneae) based on three nuclear genes and morphology. PLoS ONE 7, e38753.
Bond, J.E., Stockman, A.K., 2008. An integrative method for delimiting cohesion species: Finding the population–species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. Systematic Biology 57, 628–646.
Bruch, A.A., Utescher, T., Mosbrugger, V., NECLIME members, 2011. Precipitation patterns in the Miocene of Central Europe and the development of continentality. Palaeogeography, Palaeoclimatology, Palaeoecology 304, 202–211.
Buchli, H.H.R., 1962. Note préliminaire sur l' accouplement des araignées mygalomorphes Nemesia caementaria, *Nemesia dubia* et *Pachylomerus piceus* (Ctenizidae). Vie et Milieu 13, 177–178.
Caccone, A., Sbordoni, V., 2001. Molecular biogeography of cave life: a study using mitochondrial DNA from bathysciine beetles. Evolution 55, 122–130.
Cavazza, W., Wezel, F.C., 2003. The Mediterranean region – a geological primer. Episodes 26, 160–168.
Colgan, D.J., McLauchlan, A., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G., Gray, M.R., 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Australian Journal of Zoology 46, 419–437.
Coyle, E.A., Greenstone, M.H., Hultsch, A.L., Morgan, C.E., 1985. Ballooning mygalomorphs: estimates of the masses of *Sphodros* and *Ummidia* ballooners (Araneae: Atypidae, Ctenizidae). Journal of Arachnology 13, 291–296.
Coyle, F.A., 1981. Notes on the behaviour of *Ummidia* trapdoor spiders (Araneae, Ctenizidae): burrow construction, prey capture, and the functional morphology of the peculiar third tibia. Bulletin of British Arachnological Society 5, 159–165.
Coyle, F.A., 1985. Ballooning behavior of *Ummidia* spiderlings (Araneae, Ctenizidae). Journal of Arachnology 13, 137–138.
Creutzburg, N., 1963. Paleogeographic evolution of Crete from Miocene till our days. Cretan Annals 15 (16), 336–342.
Decae, A., 2010. The genus *Ummidia* Thorell 1875 in the western Mediterranean, a review (Araneae: Mygalomorphae: Ctenizidae). Journal of Arachnology 38, 328–340.
Decae, A.E., 1983. *Cyrtocarenum* Ausserer, 1871, a living fossil? In: Eberhard, W.G., Lubin, Y.D., Robinson, B. (Eds.), Proceedings of the Ninth International Congress of Arachnology. Smithsonian Tropical Research Institute, Panama.
Decae, A.E., 1996. Systematics of the trapdoor spider genus *Cyrtocarenum* Ausserer, 1871 (Araneae, Ctenizidae). Bulletin of British Arachnological Society 10, 161–170.
Decae, A.E., 2011. Diversity and Distribution of Mygalomorph Spiders in the Mediterranean Region. Terrestrial Ecology Unit (TEREC), Ghent University, Ghent, p. 163.
Decae, A.E., Caranhac, G., Thomas, G., 1982. The supposedly unique case of *Cyrtocarenum cunicularium* (Olivier, 1811) (Araneae, Ctenizidae). Bulletin of British Arachnological Society 5, 410–419.
Dermitzakis, D.M., 1990. Paleogeography, geodynamic processes and event stratigraphy during the late Cenozoic of the Aegean area. Accademia Nazionale de Lincei 85, 263–288.
Drummond, A., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2010. Geneious v.5.3. <http://www.geneious.com>.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29, 1969–1973.
Duffey, E., 1956. Aerial dispersal in a known spider population. Journal of Animal Ecology 25, 85–111.
Duffey, E., 1997. Aerial dispersal in spiders. In: Selden, P.A. (Ed.), Proceedings of the 17th European Colloquium of Arachnology. Edinburgh.
Eberhard, W.G., 2006. Dispersal by *Ummidia* spiderlings (Araneae, Ctenizidae): ancient roots of aerial webs and orientation? Journal of Arachnology 34, 254–257.
Eronen, J.T., Mirzaie, M., Karme, A., Micheels, A., Bernor, R.L., Fortelius, M., 2009. Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. Proceedings of the National Academy of Sciences 106, 11867–11871.
Gertsch, W.J., 1979. American Spiders. Van Nostrand Reinhold, New York.
Gertsch, W.J., Wallace, H.K., 1936. Notes on new and rare American Mygalomorph spider. American Museum Novitates 884, 1–25.
Gertsch, W.J., Platnick, N.I., 1975. A Revision of the trapdoor sider genus *Cyclocosmia* (Arenae, Ctenizidae). American Museum Novitates 2580, 1–20.
Gibbard, P.L., Head, M.J., Walker, M.J.C., 2010. Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. Journal of Quaternary Science 25, 96–102.
Giribet, G., Rambla, M., Carranza, S., Baguñà, J., Riutort, M., Ribera, C., 1999. Phylogeny of the arachnid order Opiliones (Arthropoda) inferred from a combined approach of complete 18S and partial 28S ribosomal DNA sequences and morphology. Molecular Phylogenetics and Evolution 11, 296–307.
Heath, T.A., 2012. Starttree. <http://bodegaphylo.wikispot.org/starttree_program> (accessed 19.10.12).
Hedin, M., Bond, J.E., 2006. Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): conflict and agreement with the current system of classification. Molecular Phylogenetics and Evolution 41, 454–471.
Hedin, M.C., Maddison, W.P., 2001. A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). Molecular Phylogenetics and Evolution 18, 386–403.
Hendrixson, B.E., Bond, J.E., 2009. Evaluating the efficacy of continuous quantitative characters for reconstructing the phylogeny of a morphologically homogeneous spider taxon (Araneae, Mygalomorphae, Antrodiaetidae, Antrodiaetus). Molecular Phylogenetics and Evolution 53, 300–313.
Hunt, R.H., 1976. Notes on the ecology of *Cyclocosmia truncata* (Araneae, Ctenizidae) in Georgia. Journal of Arachnology 3, 83–86.
Jiménez-Moreno, G., Fauquette, S., Suc, J.P., 2010. Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. Review of Palaeobotany and Palynology 162, 403–415.
Jolivet, L., Augier, R., Robin, C., Suc, J., 2006. Lithospheric-scale geodynamic context of the Messinian salinity crisis. Sedimentary Geology 188–189, 9–33.
Jones, E.J.W., Cande, S.C., Spathopoulos, F., 1995. Evolution of a major oceanographic pathway: the equatorial Atlantic. Journal of the Geological Society of London 90, 199–213.
Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings in Bioinformatics 9, 286–298.
Ketmaier, V., Argano, R., Caccone, A., 2003. Phylogeography and molecular rates of subterranean aquatic Stenasellid Isopods with a peri-Tyrrhenian distribution. Molecular Ecology 12, 547.
Ketmaier, V., Giusti, F., Caccone, A., 2006. Molecular phylogeny and historical biogeography of the land snail genus *Solatopupa* (Pulmonata) in the peri-Tyrrhenian area. Molecular Phylogenetics and Evolution 39, 439–451.
Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652–655.
Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29, 1695–1701.
Lymberakis, P., Poulakakis, N., 2010. Three continents claiming an Archipelago: the evolution of Aegean's herpetofaunal diversity. Diversity 2, 233–255.
Main, B., 1997. Hogg's phantom spider from Central Australia: a century-old mystery solved. In: Selden, P.A. (Ed.), Proceedings of the 17th European Colloquium of Arachnology. British Arachnological Society, Burnham Beeches, Edinburgh, pp. 83–90.
Mallatt, J., Sullivan, J., 1998. 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. Molecular Biology and Evolution 15, 1706–1718.
McCarthy, D., 2003. The trans-Pacific zipper effect: disjunct sister taxa and matching geological outlines that link the Pacific margins. Journal of Biogeography 30, 1545–1561.
Michalak, I., Zhang, L.B., Renner, S.S., 2010. Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. Journal of Biogeography 37, 1214–1226.
Milne, R.I., 2006. Northern hemisphere plant disjunctions: a window on tertiary land bridges and climate change? Annals of Botany 98, 465–472.
Moggridge, J.T., 1873. Harvesting Ants and Trap-Door Spiders. L. Reeve & Co, London.
Myers, N., Mittermeier, R.A., Mittermeier, C.C., da Fonseca, G.A., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Obst, M., Faurby, S.r., Bussarawit, S., Funch, P., 2012. Molecular phylogeny of extant horseshoe crabs (Xiphosura, Limulidae) indicates Paleogene diversification of Asian species. Molecular Phylogenetics and Evolution 62, 21–26.
Ortiz, D., 2007. Ummidiinae, a new replacement name for Pachylomerinae Simon 1889 (Araneae: Ctenizidae). Boletín de la Sociedad Entomológica Aragonesa 40, 395–396.

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E., Stamatakis, A., 2010. How many bootstrap replicates are necessary? Journal of Computational Biology 17, 337–354.

Paulo, O.S., Pinheiro, J., Miraldo, A., Bruford, M.W., Jordan, W.C., Nichols, R.A., 2008. The role of vicariance vs. dispersal in shaping genetic patterns in ocellated lizard species in the western Mediterranean. Molecular Ecology 17, 1535–1551.

Pinho, C., Harris, D.J., Ferrand, N., 2007. Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. Molecular Ecology 16, 1191–1205.

Platnick, N.I., 2012. The World Spider Catalog. American Museum of Natural History, New York, <http://research.amnh.org/iz/spiders/catalog> (accessed 24.12.12).

Poulakakis, N., Pakaki, V., Mylonas, M., Lymberakis, P., 2008. Molecular phylogeny of the Greek legless skink *Ophiomorus punctatissimus* (Squamata: Scincidae): the impact of the Mid-Aegean trench in its phylogeography. Molecular Phylogenetics and Evolution 47, 396–402.

Rambaut, A., Drummond, A.J., 2009. TRACER v.1.5. <http://tree.bio.ed.ac.uk/software/tracer/> (accessed 01.12.09).

Raven, R.J., 1985. The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. Bulletin of American Museum of Natural History 182, 1–180.

Ritzkowski, S., 1997. K–Ar-Altersbestimmungen der bernsteinführenden Sedimente des Samlandes (Paläogen, Bezirk Kaliningrad). In: Ganzelewski, M., Rehren, T.H., Slotta, R. (Eds.), Metalla, Sonderheft 66 zum Symposium Neue Erkenntnisse zum Bernstein. Deutsches Bergbau Museum, Bochum, Germany, pp. 19–24.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rosenbaum, G., Lister, G.S., Duboz, C., 2002a. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. In: Rosenbaum, G., Lister, G.S. (Eds.), Reconstruction of the evolution of the Alpine-Himalayan Orogen. Journal of the Virtual Explorer, pp. 107–126.

Rosenbaum, G., Lister, G.S., Duboz, C., 2002b. Relative motions of Africa, Iberia and Europe during Alpine orogeny. Tectonophysics 359, 117–129.

Rosenbaum, H.C., Weinrich, M.T., Stoleson, S.A., Gibbs, J.P., Baker, C.S., DeSalle, R., 2002c. The effect of differential reproductive success on population genetic structure: correlations of life history with matrilines in humpback whales of the gulf of maine. Journal of Heredity 93, 389–399.

Rowe, D.L., Dunn, K.A., Adkins, R.M., Honeycutt, R.L., 2010. Molecular clocks keep dispersal hypotheses afloat: evidence for trans-Atlantic rafting by rodents. Journal of Biogeography 37, 305–324.

Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14, 1218–1231.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Molecular Biology and Evolution 19, 101–109.

Sanmartin, I., Enghoff, H., Ronquist, F., 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biological Journal of the Linnean Society 73, 345–390.

Santos, X., Rato, C., Carranza, S., Carretero, M.A., Pleguezuelos, J.M., 2012. Complex phylogeography in the Southern Smooth Snake (*Coronella girondica*) supported by mtDNA sequences. Journal of Zoological Systematics and Evolutionary Research 50, 210–219.

Saunders, S.S., 1842. Description of a species of *Mygale*, from Ionica, with its nest. Transactions of the Royal Entomological Society of London 3, 160–170.

Schellart, W., 2002. Alpine deformation at the western termination of the Axial Zone, Southern Pyrenees. Journal of the Virtual Explorer 8, 35–55.

Schwendinger, P.J., Giribet, G., 2005. The systematics of the south-east Asian genus *Fangensis* Rambla (Opiliones: Cyphophthalmi: Stylocellidae). Invertebrate Systematics 19, 297–323.

Selden, P.A., Shear, W.A., Bonamo, P.M., 1991. A spider and other arachnids from the Devonian of New York, and a reinterpretation of Devonian Araneae. Palaeontology 34, 241–281.

Selden, P.A., Shear, W.A., Sutton, M.D., 2008. Fossil evidence for the origin of spider spinnerets, and a proposed arachnid order. Proceedings of the National Academy of Sciences 105, 20781–20785.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51, 492–508.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.

Siliwal, M., Nair, M.V., Molur, S., Raven, R.J., 2009. First record of the trapdoor spider genus *Conothele* (Araneae, Ctenizidae) from India, with a description of two new species. Journal of Arachnology 37, 1–9.

Simon, E., 1864. Histoire naturelle des Araignées (Aranéides). Libraire Encyclopedique de Roret, Paris.

Simon, E., 1910. Catalogue raisonné des Arachnides du Nord de l'Afrique, 1ere partie. Annales de la Societé Entomologique de France 79, 265–332.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Stock, J.H., 1993. Some remarkable distribution patterns in stygobiont Amphipoda. Journal of Natural History 27, 807–819.

Suc, J.P., 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. Nature 307, 429–438.

Talavera, G., Lukhtanov, V.A., Pierce, N.E., Vila, R., 2012. Establishing criteria for higher-level classification using molecular data: the systematics of Polyommatus blue butterflies (Lepidoptera, Lycaenidae). Cladistics 29, 166–192.

Tiffney, Manchester, 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere tertiary. International Journal of Plant Sciences 162, S3–S17.

Tiffney, B.H., 1985a. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phylogeography of the Northern Hemisphere. Journal of the Arnold Arboretum 66, 243–273.

Tiffney, B.H., 1985b. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. Journal of the Arnold Arboretum 66, 73–94.

Winchell, C.J., Sullivan, J., Cameron, C.B., Swalla, B.J., Mallatt, J., 2002. Evaluating hypotheses of deuterostome phylogeny and chordate evolution with new LSU and SSU ribosomal DNA data. Molecular Biology and Evolution 19, 762–776.

Wunderlich, J., 1995. Zur Kenntnis der Endemiten, zur Evolution und zur Biogeographie der Spinnen Korsikas und Sardiniens, mit Neubeschreibungen (Arachnida: Araneae). Beiträge zur Araneologie 4, 353–383.

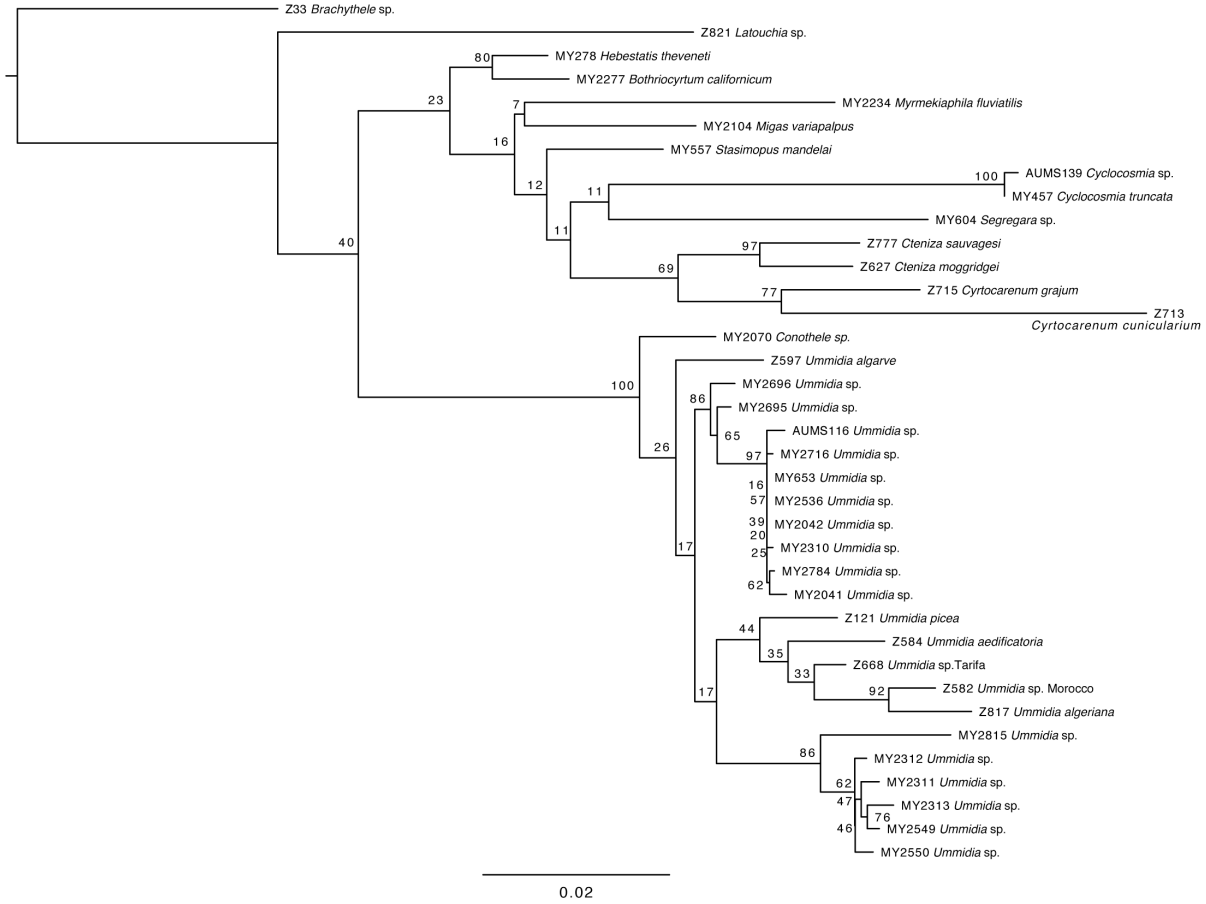
Wunderlich, J., 2000. Zwei neue Arten der Familie Falltürspinnen (Araneae: Ctenizidae) aus dem Baltischen Bernstein. Entomologische Zeitschrift 110, 345–348.

Wunderlich, J., 2004. Introduction, General Findings and conclusions. In: Wunderlich, J. (Ed.), Beiträge zur Araneologie, pp. 5–329.

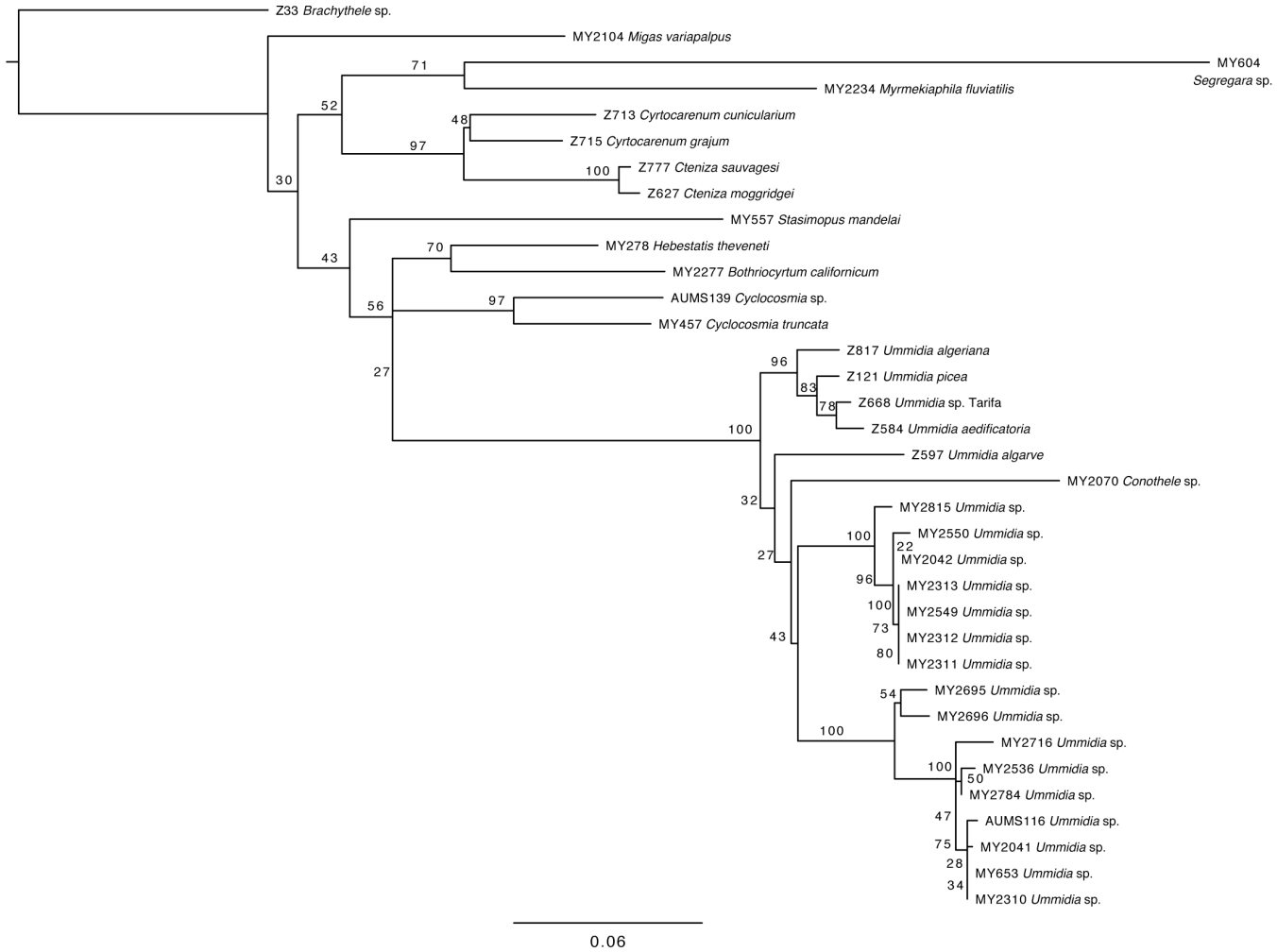
Wunderlich, J.h., 2011. Some fossil spiders (Araneae) in Eocene European ambers. In: Wunderlich, J. (Ed.), Beiträge zur Araneologie, pp. 472–538.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693.

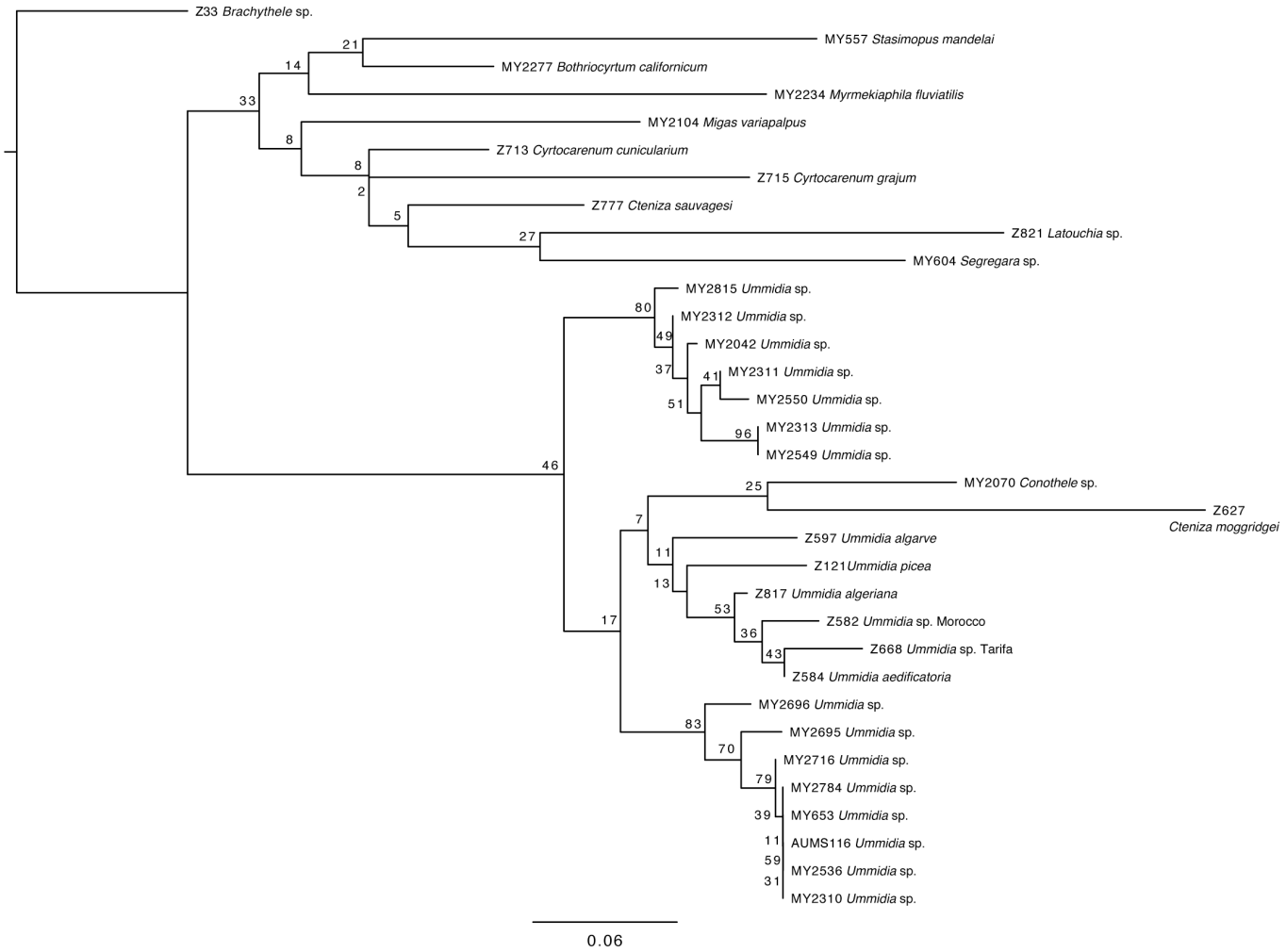
Supporting material



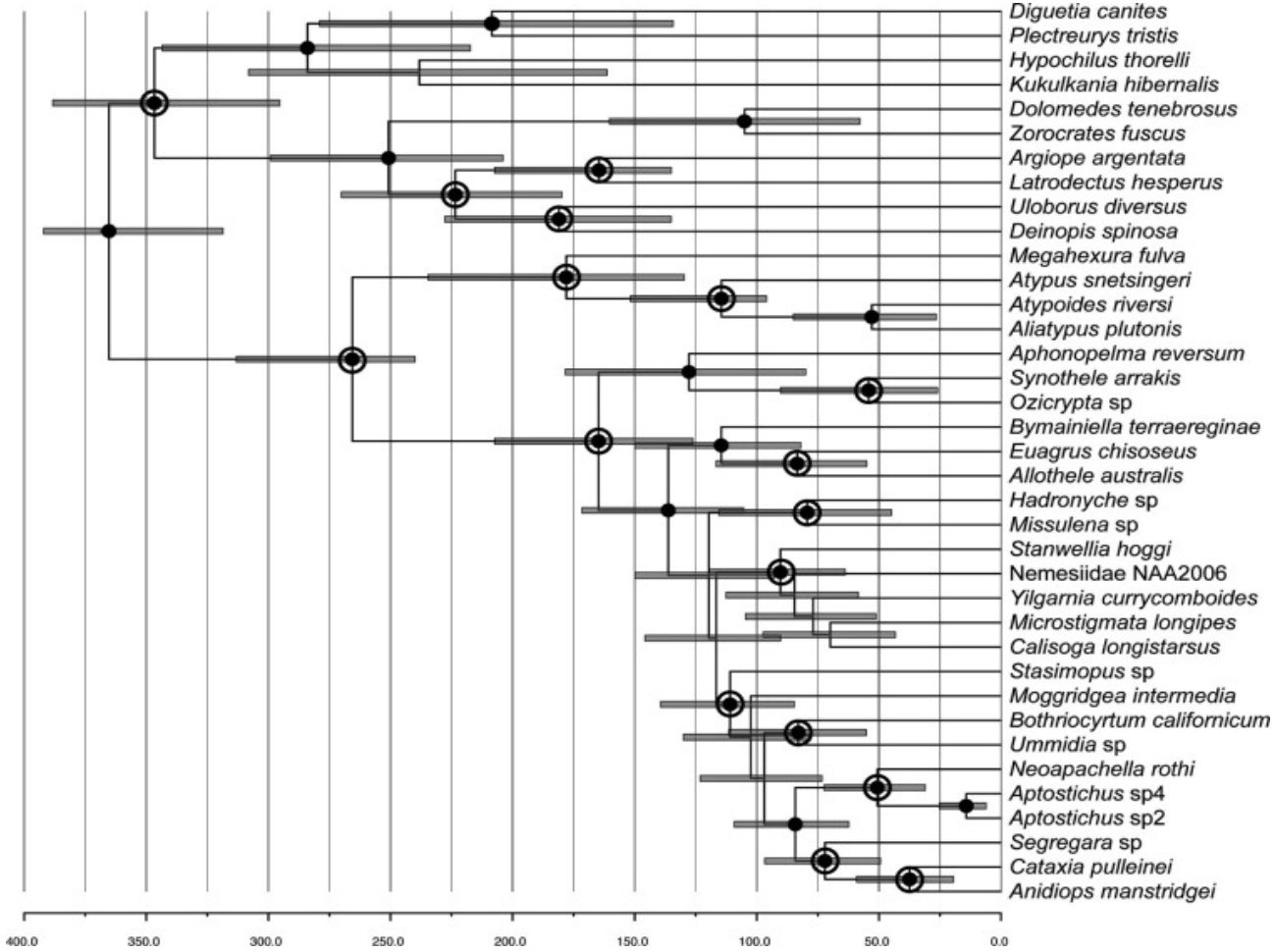
S1. – Individual ML *28S rRNA* gene tree obtained with RaxML



S2. – Individual ML *EF1γ* gene tree obtained with RaxML



S3. – Individual ML *H3* gene tree obtained with RaxML



S4.- Ultrameric *Ef1γ* tree obtained with BEAST to infer a substitution rate for the gene. Nodes constrained in the analysis labelled with a circled dot. Simple dots refer to clades supported >0.95 PP.

Chapter 2

Vera Opatova, Jason E Bond & Miquel A Arnedo. Loosening the belt: How the Hercynian Belt breakup shaped the distribution of the trap-door spider genus *Ummidia* (Araneae, Ctenizidae) in the Western Mediterranean.

Aflojarse el cinturón: La desintegración del Cinturón Herciniano modeló la diversidad y distribución de las arañas migalomorfas del género *Ummidia* (Araneae, Ctenizidae) en el Mediterráneo Occidental

Resumen

La distribución actual de las arañas migalomorfas del género *Ummidia* en el Mediterráneo Occidental ha sido previamente vinculada a los eventos geológicos pasados en la región y su ocurrencia a ambos lados del Océano Atlántico de procesos de vicariancia a nivel continental. En este estudio, utilizamos un enfoque multilocus que combina los datos de secuencias de ADN de cuatro genes nucleares y un gen mitocondrial para investigar las relaciones filogenéticas de todas las especies del género *Ummidia* descritas del Mediterráneo cubriendo la mayor parte de la distribución del grupo y proporcionar un marco temporal para su diversificación. También evaluamos las preferencias ecológicas de las tres especies ibéricas de *Ummidia* y evaluamos su intercambiabilidad ecológica. Nuestros resultados indican que la distribución actual del género *Ummidia* en el Mediterráneo Occidental se debe principalmente la migración de las placas tectónicas hasta su posición actual tras la desintegración del Cinturón Herciniano. Esta hipótesis está apoyada por el marco temporal de divergencia de los linajes principales de *Ummidia* datadas en el Mioceno. A pesar de dos casos de colonización de la Península Ibérica procedentes de África, presumiblemente durante la Crisis de Salinidad del Messiniense, los patrones filogeográficos observados y el aislamiento completo de las especies ecológicamente intercambiables y geográficamente próximas indican que la dispersión a larga distancia es probablemente muy rara y su distribución actual se atribuye principalmente a vicariancia. Los análisis filogenéticos revelan gran cantidad de linajes evolutivos independientes que pudieran corresponder a las nuevas especies putativas.

Loosening the belt: How the Hercynian Belt breakup shaped the diversity and distribution of the trap-door spider genus *Ummidia* (Araneae, Ctenizidae) in the Western Mediterranean

Vera Opatova[1,2] , Jason E Bond[2], Miquel A. Arnedo[1]

[1]*Institut de Recerca de la Biodiversitat (IRBio), Departament de Biologia Animal, Universitat de Barcelona (UB), Av. Diagonal 643, 08028, Barcelona, Spain.* [2]*Department of Biological Sciences and Auburn University Museum of Natural History, Auburn University, Auburn, Alabama, 36849, USA*

Abstract

The present day amphi-Atlantic distribution of the trap-door spider genus *Ummidia* has been suggested to be a result of continental drifting. Here, we use a multi-locus approach that combines one mitochondrial and four nuclear genes to reconstruct the phylogenetic relationships within the Mediterranean *Ummidia* species and to infer a temporal framework for its diversification based on a thorough sampling across the known distribution of the genus. Species distribution modelling tools are further use to assess the ecological preferences of the three *Ummidia* species inhabiting the Iberian Peninsula and to evaluate their ecological interchangeability. Our results indicate that the present day distribution and diversity of the Mediterranean *Ummidia* was mostly shaped by the Hercynian Belt breakup and the subsequent drifting of the resulting microplates, which dates back to the Miocene. In spite of the reported ability of Ummidia for airborne dispersal the observed phylogeographic patterns and the isolation of ecologically interchangeable and geographically proximate species indicate that long distance dispersal events are rare and that the present day distribution may be mostly attributed to vicariant events drove by microplates split. Two inferred events of backward colonization of the Iberian Peninsula from northern Africa were presumably facilitated by land bridges formed during the Messinian Salinity Crisis. Finally, several additional evolutionary independent lineages that may correspond to putative new species were further detected.

Introduction

Vicariance and dispersal are the main processes responsible for the biogeographic patterns observed in the distribution of organisms. While dispersal leads to the expansion of species ranges by overcoming barriers and establishing new populations, vicariance processes involve the fragmentation of a previously continuous range by the formation of a barrier that hampers gene flow among populations. Geological processes such as the continental breakups or the tectonic plates drifting contributed to the present day distribution of many organisms (Altaba 1998; McCarthy 2003; Noonan & Chippindale 2006; Waters *et al.* 2000). This is especially true in lineages with low vagility or narrow habitat preferences (Bauzà-Ribot, 2012; Opatova *et al.* 2013; Stock 1993).

The vicariant paradigm has been subsequently challenged by the demonstration that dispersal events may have had a more relevant role in shaping organism distributions that originally considered (Michalak *et al.* 2010; Sanmartin 2003; Sanmartin *et al.* 2001). Disentangling the relative contribution of vicariance and dispersal to present day distributions is a daunting task. The use of molecular dating methods in a phylogenetic framework, however, provides the necessary information to test alternative biogeographic scenarios (e.g. Noonan 2000; Paulo *et al.* 2008), uncover rare long distance dispersal events (Rowe *et al.* 2010) and evaluate the chances of human mediated introductions (Chapman *et al.* 2007; Cunningham 2008; Opatova *et al.* 2013).

The Mediterranean Basin ranks among the twenty-five global biodiversity hotspots deserving special conservation attention. The extraordinary species richness and high level of endemism in this area is likely a consequence of the region's unique position at the crossroads between the Eurasian and Africa plates, and its complex climatic and geological history (Myers *et al.* 2000).

The formation of the Mediterranean basin dates back to the upper Eocene, about 40 million years ago (Ma), when the collision between the African and Eurasian plates precipitated the closure of the Thetis Ocean (Blondel & Aronson 1999) and catalysed the Alpine orogeny, which continued throughout to the Miocene (Schellart 2002). During the middle Oligocene, about

30 Ma, a block of landmass located between the north-eastern Iberia and southern France detached and drifted off from its previous position. This region, known as the Hercynian belt, comprised of the present day Balearic Island, Corsica, Sardinia, Calabria, the Betics, the Rif and the Kabylies, subsequently broke up into smaller blocks as a result of former plate subduction rollback, driving the opening of the Gulf of Lyon (~30 Ma) and the Valencia Through (~25 Ma) (Roca *et al.* 1999) (Fig. 1).

The Corsica, Sardinia and Calabria block collided with the Appenines between 18-20 Ma following a series of counter clockwise rotations Corsica and Sardinia subsequently detached from the Apennines in the mid Miocene around 9 Ma and started drifting off towards their present day position (Rosenbaum *et al.* 2002a).

Around 21 Ma the second block split up. The Kabylies block rifted southwards and collided with the Northern Africa margin, between 18-15Ma. The Betic-Rif block, on the other hand, reached its present day position in the southern Iberia and Northern Africa approximately 10 Ma completing the opening of the Alboran Sea (Rosenbaum *et al.* 2002a; Rosenbaum *et al.* 2002b; Rosenbaum *et al.* 2002c). The exact original position of the Betic-Rif block is still a matter of debate, but some authors suggest that during the Oligocene Betic-Rif block formed a continuous orogenic belt alongside Corsica, Calabria and the Kabylies microplates (Rosenbaum et al., 2002a).

The closure of the Strait of Gibraltar in the late Miocene (about 5.96 Ma) caused the nearly complete desiccation of the Mediterranean Basin (i.e. the Messinian salinity crisis). The dramatic sea level drop led to the emergence of land bridges between formerly isolated areas (e.g., North Africa with Iberia, Sicily and Italy and Corsica with Sardinia). The former conditions were re-established following the reopening of the Strait of Gibraltar dated to 5.3 Ma (Jolivet *et al.* 2006; Krijgsman *et al.* 1999).

The present day Mediterranean climate characterized by dry and hot summers and humid cool winters was established about 3.4 Ma (Jiménez-Moreno *et al.* 2010; Suc 1984), following a long-term cooling trend that started in the Eocene (Zachos *et al.* 2001). The shift from the warmer and wetter climate towards cooler temperatures was characterized by major extinctions and range shifts in some terrestrial organism (Böhme 2003). The Quaternary

glacial cycles began in Europe between 2.8 and 2.5 Ma and continued in 100,000 – yearlong cycles until the Holocene (Gibbard *et al.* 2010). Many European organisms suffered range shifts and the Mediterranean peninsulas served as refugia for thermophilic species (Blondel & Aronson 1999; Hewitt 1996; Hewitt 2000).

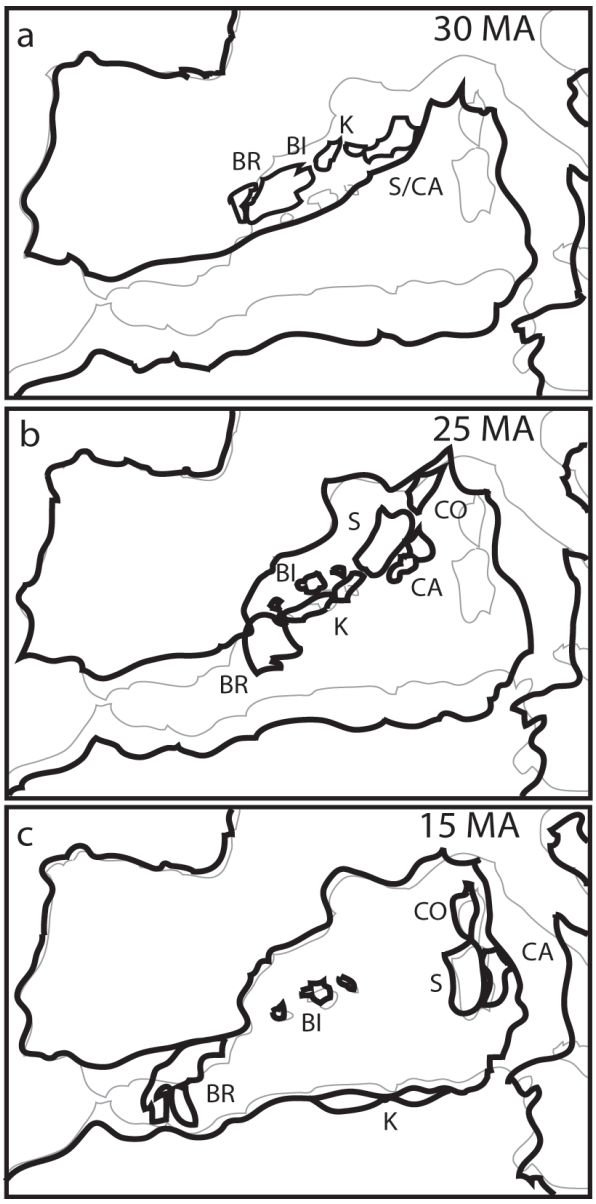


Fig. 1.- Paleogeographic reconstructions of the Mediterranean region in (a) Oligocene 30 Ma, (b) 25Ma and (c) Miocene 15 Ma adapted from Rosenbaum et al 2002. BI: Balearic Islands, BR: Betic-Rif, CA: Calabria, CO: Corsica, K: Kabylies, S: Sardinia.

The infraorder Mygalomorphae, which comprises mostly robust spiders such as tarantulas, trapdoor spiders, purse web spiders or funnel web spiders, is one of the three main evolutionary lineages recognized within spiders (Hedin & Bond 2006). The group has been frequently considered as "basal" or "primitive" because it retains morphological features that have been considered plesiomorphic in spiders, such as the presence of four book lungs or the chelicerae with longitudinal fangs and unsynchronized movement (Raven 1985). The mygalomorphs are relatively species poor, they include 2,800 species, compared to its sister lineage, the Araneomorpha (the "true spiders"), which counts more than 40,000 species, (Platnick 2014).

Mygalomorph females are usually long-lived, sedentary and show strong burrow/nest fidelity. The gene flow between populations is mostly mediated by males, wandering in search for females after reaching adulthood. Air-borne dispersal (i.e. ballooning) of juveniles or small adults which is common in araneomorph spiders (Duffey 1956, 1997) has been reported only in three mygalomorph families (Coyle *et al.* 1985; Eberhard 2006; Coyle 1983; Cutler & Guarisco 1995; Ferretti *et al.* 2013). Because of the low dispersal ability, mygalomorph spiders tend to show deep populations structure and high levels of local endemism (Arnedo & Ferrández 2007; Bond *et al.* 2001; Bond & Stockman 2008; Cooper *et al.* 2011; Hendrixson *et al.* 2013; Satler *et al.* 2011). The sedentary nature of this group provides an excellent model system for biogeography studies (Hedin *et al.* 2013; Opatova *et al.* 2013; Raven 1980).

The genus *Ummidia* (Thorell, 1875) (Fig. 2) belongs to the widespread trap-door spider family Ctenizidae, which inhabits North and South America, the Mediterranean region, the southern Africa, South-East Asia and Australia. The monophyly of Ctenizidae has been challenged by molecular phylogenetic analyses of the infraorder Mygalomorphae (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin & Bond 2006; Opatova *et al.* 2013).

The amphi-Atlantic distribution of *Ummidia*, and the fact that the bulk of *Ummidia* diversity was found in the New World, led to former authors to suggest that the presence of the genus in the Mediterranean was the result of a human mediated introduction (Simon 1864, 1910). Subsequently, Decae (2010) hypothesized that the presence of the genus both in the Mediterranean and the New World could be better explained as the result of a previous Laurasian

distribution. This hypothesis was further supported by a phylogenetic study that included molecular divergence time estimates, which dated the split between the Mediterranean and the New World *Ummidia* to the time period of the Laurasia break up (Opatova *et al.* 2013). The same study also suggested that the present-day distribution of *Ummidia* correlated with the complex geological history of the Western Mediterranean basin, although the sparse sampling prevented to make more solid conclusions (Opatova *et al.* 2013).

Fig. 2.- *Ummidia picea* adult female in the burrow.

To date, four *Ummidia* species are known from the Mediterranean region (Platnick 2014), namely: *U. aedificatoria* (Westwood, 1840) from Morocco, *U. algeriana* (Lucas, 1846) from the Tell region in Algeria and western Tunisia, *U. picea* from south-eastern Spain (Thorell, 1875) and *U. algarve* from southern Portugal (Decae 2010). *Ummidia* spiders construct underground silk lined burrows with trapdoor that opens to the surface. In *U. algarve*, a second inner trapdoor leading to a short tunnel in continuation of the burrow has also been reported (Decae 2010). Unlike the vast majority of mygalomorphs, *Ummidia* species are capable of airborne dispersal. The genus is present in some Caribbean Islands of volcanic origin, without previous connection to the mainland (Platnick 2014), which indicates an ability for overseas dispersal.

Here, we aim to unravel the biogeographic patterns of the genus *Ummidia* in the context of the dynamic geologic history of the Mediterranean region by reconstructing the phylogenetic relationships of a large sample of specimens collected throughout its known distribution using multi-locus data and inferring a temporal framework for its diversification using both gene concatenation and coalescent species-tree approaches. We further assess the ecological preferences of three *Ummidia* species to evaluate their ecological interchangeability.

Materials and Methods

Taxonomic sampling

Most of the *Ummidia* samples used in present study was obtained by the authors during extensive collecting campaigns covering most of the known *Ummidia* distribution range in the Mediterranean region (Fig. 3). Additional specimens of *Ummidia* (namely Z642 and Z1050) were kindly donated by colleagues. Detailed locality data is listed in Table S1.

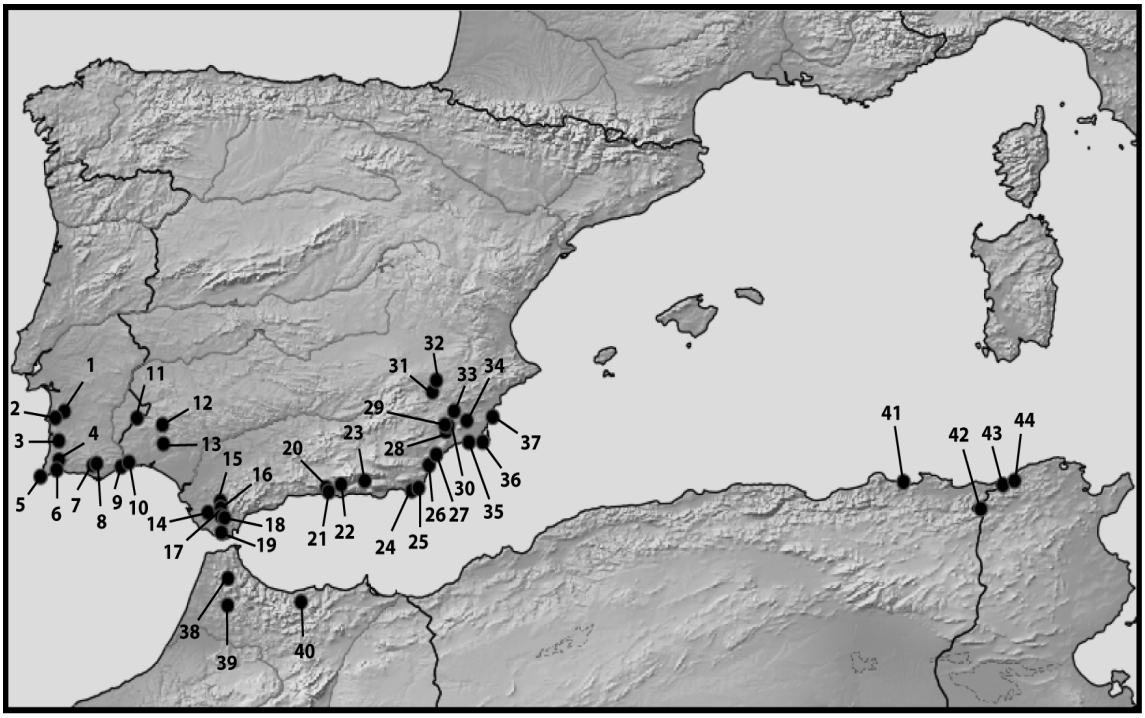


Fig. 3.- Map showing sampling locations of the genus *Ummidia*, for detailed locality information see Table S1.

DNA extraction, PCR amplification and sequencing

Whole genomic DNA was extracted from one leg per adult specimen or the whole prosoma, in case of juveniles, with the help of the SpeedTools Tissue Extraction Kit (Biotools). Partial fragments of one mitochondrial and three nuclear genes were sequenced in present study: the 5' half of the Cytochrome oxidase I (*cox1*) (the animal barcode), a fragment of the Elongation factor-1 gamma (*EF1γ*), the *wingless* (*WN*) gene and one anonymous locus 12541 (*AL12541*) obtained from an on-going restriction enzyme-based reduced representation library (RRL) sequencing project focus on mygalomorph spiders (Frias-López *et al*, unpublished).

The PCR amplifications were carried out with the following primer combinations and PCR conditions: *cox1* was successfully amplified at annealing temperature range of 43-46°C with the primer pair C1-J-1490/C1-N-2198 (Folmer *et al.* 1994). The *EF1γ* fragment was amplified with primers ER1gF78/EF1γR1258, EF1gF218/EF1gR1090 and EF1gF179/EF1gR1168 (Ayoub *et al.* 2007) or alternatively with a primer pair designed in the present study EF-gUMMIf 5´- CATAGATCAGTCGCA -3´/ EF-gUMMIr 5´- GGTCTCCGACCTGTTC -3´, at 52°C annealing temperature. Mygalomorph protein coding gene *WN* was amplified with primers MYGWNT6FOR / MYGWNT2REV (Satler *et al.* 2011) and *AL12541* with primer pair 12541FWD 5´- GAGCTTCATCATACCCTA -3´/ 12541REV 5´- GTGAGAAAAACATCCATG -3' (Frias-López *et al*, unpublished). Both primer combinations yielded amplifications at an annealing temperature between 52-54°C.

All the PCR reactions were carried out in reaction volume of 25μl, purified using ExoSAP-IT (USB Corporation) and sequenced in both directions using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on ABI 3700 automated sequencer at *Centres Científics i Tecnològics* of the University of Barcelona (CCiTUB, www.ccit.ub.edu) Spain.

The chromatograms were assembled and edited in Geneious v. 5.3.6. (Drummond *et al.* 2012a)

Sequence alignment, allele phasing and recombination testing

An online version of the alignment program TranslatorX (available at http://www.translatorx.co.uk/, Abascal *et al.* 2010), which uses amino acid translation of protein coding genes to guide the nucleotide alignment, was used to build the alignment of *EF1*γ since the fragment presented length polymorphism due to insertions/deletions. The alignment of *cox1*, *WN* and the *AL12541* was trivial since no length variation was observed in these gene fragments.

The allelic phases of the nuclear gene fragments were determined using the PHASE algorithm (Stephens *et al.* 2001; Stephens & Donnelly 2003) as implemented in DnaSP 5.10.1 (Librado & Rozas 2009).

A possible presence of recombination was tested by means of the difference of sum of squares method (DSS) as implemented in TOPALi v 2.5 (McGuire& Wright 2000; Milne *et al.* 2009). The size of sliding window was set to 500bp in all fragments except for *WN*, where a 300bp frame was used instead due to shorter length of the fragment.

Delimitation of putative evolutionary lineages

The General Mixed Yule Coalescence model (GMYC) approach (Pons *et al.* 2006) was used to delimitate coalescent groups (i.e. putative independent evolutionary lineages) within the complete *Ummidia cox1* sequence data set.

A maximum Likelihood (ML) analysis was conducted with RaxML v.7.2.8. (Stamatakis 2006) under the graphical interface raxmlGUI (Silvestro & Michalak 2011). Equal sequences were removed from the analysis. A codon partition scheme was specified, with a CAT+G model for each partition. The best maximum likelihood tree was selected from 100 random iterations. The maximum likelihood tree was subsequently transformed to ultrameric with the help of the program PATHd8 (Britton 2007). Relative branch lengths were obtained by setting the root to 1. The GMYC analysis was carried out in the R (http://www.r-project.org) environment with the help of the SPLITS package (Ezard *et al.* 2009). The outgroup was removed from the ultrameric tree prior to the analysis.

Phylogenetic analyses

The concatenation of genes evolving at different rates, widely used to resolve phylogenetic relationships, follows the assumption that individual gene trees are congruent among themselves and with the species tree. However, different genes may support incongruent topologies due to processes such as incomplete lineage sorting, hybridization, gene duplication or horizontal gene transfer (Maddison 1997; Degnan & Rosenberg 2006; Kubatko & Degnan 2007). In the present study, we conducted phylogenetic analyses using two different approaches: (1) a concatenation approach, which assumes a common underlying tree for all genes and (2) a coalescent-based species tree/gene tree approach, which allows to each gene to support different gene tree (see below). All phylogenetic analyses were rooted by including a North American *Ummidia* species (*Ummidia* sp., Alabama, Lee Co, USA) and assuming a sister group relationship between the *Ummidia* lineages at both sides of the Atlantic Ocean (Opatova *et al.* 2013).

The *cox1*, *EF1γ, WN* and *AL12541* gene fragments of one single individual representing each GMYC entity delimited were concatenated in a single matrix using Geneious v. 5.3.6. (Drummond *et al.* 2012a). Gaps inferred for the *EF1γ* were scored as presence/absence data following Simons and Ochoterena (Simmons & Ochoterena 2000), with the help of the program FastGap 1.2 (available at http://www.aubot.dk/FastGap_home.htm, Borchsenius, 2009). The best partitioning schemes and evolutionary models were selected using the greedy algorithm provided in the program PARTITIONFINDER (Lanfear *et al.* 2012).

Bayesian inference (BI) analyses were conducted in MrBayes v. 3.1.2 (Ronquist & Huelsenbeck 2003). The best partition scheme selected with their corresponding evolutionary models was defined for each gene fragment while a restriction model was assigned gaps. Two independent runs of $5x10^7$ generations with 8 MCMC (Markov Chain Monte Carlo) chains each, starting from random trees and resampling every 1000 generations were run remotely at the CIPRES server (http://www.phylo.org/ Miller *et al.* 2010). The first 20% of the generations were discarded as a *burn-in* for the analyses. Convergence of the runs and correct mixing of each chain were assessed by monitoring the

standard deviation of split frequencies (<0.01) and the correct mixing of each chain by the ESS values reported in TRACER v.1.5 (Rambaut & Drummond 2009).

Maximum Likelihood (ML) analyses were conducted with RaxML v.7.2.8. (Stamatakis 2006) under the graphical interface raxmlGUI (Silvestro& Michalak 2011). Independent GRT+G+I substitution models were assigned to each partition, and a binary model was applied to the gaps. The best maximum likelihood tree was selected from 100 random iterations and support assessed with 1000 replicates of bootstrap resampling.

All trees were visualized and manipulated with the program FigTree v. 1.3.1 (Rambaut 2009).

Estimation of divergence times

Divergence time was estimated using both concatenation and coalescent-based approaches. The analyses were conducted in a Bayesian framework with the help of the program BEAST 1.7.4. (Drummond *et al.* 2012b). In the concatenation analysis each *Ummidia* GMYC entity was represented by a single individual sampled for *cox1*, *EF1γ, WN* and *AL12541* gene fragments. In an attempt to facilitate convergence and reduce computation effort, the partition scheme was simplified by combining codon position in single gene partitions. The substitution model selected in PARTIONFINDER was assigned to each gene fragment.

Because of the lack of relevant biogeographic events and the limited fossil record, we incorporate the *cox1* substitution rates estimated for the ground-dwelling spider genus *Parachtes* (Bidegaray-Batista & Arnedo 2011) and a mygalomorph specific rate estimated for *EF1γ* gene (Opatova *et al.* 2013). A normal prior was assigned to the ucld.mean parameter of the lognormal relaxed clock for both *cox1* and *EF1γ* partitions, initial and mean value 0.0199 and standard deviation 0.001 for *cox1* and 0.00118 with standard deviation of 0.00014 for *EF1γ* respectively. Uniform priors to the ucld.mean were assigned for *WN* and *AL12541* gene fragments. Lower and upper bounds were set between 0.0001 and 0.0115, with starting value at 0.001. Upper constrain, corresponding to the universal mitochondrial substitution rate, was

selected under the assumption that the nuclear genes are about one order of magnitude slower than mitochondrial and that nuclear protein coding gene will show lower rates than the mitochondrial genes. A Yule speciation model was set as tree prior for all the genes. The only available fossil information for the genus is the presence of *Ummidia* species in Baltic amber (Wunderlich 2000; Wunderlich 2004; Wunderlich 2011). Correspondingly, a minimum bound of 44 Ma, the estimated age of the Baltic amber, was set for the tree root.

Three independent runs of $8x10^7$ generations were run remotely at the CIPRES server (http://www.phylo.org/ Miller et al. 2010) sampling every 1000 generations. Convergence between runs and correct mixing within each run were visualized with TRACER (Rambaut & Drummond 2007). Individual runs were combined and resampled at lower frequency in BEAST accompanying program LOGCOMBINER. The first 20% of the generations of each run was discarded as a burn-in. A consensus chronogram was inferred with TREEANNOTATOR.

As stated above, the species tree and divergence times were also estimated using the multi-gene coalescent approach as implemented in *BEAST (Heled & Drummond 2010). Because this approach involves the estimation of population sizes, only GMYC clusters, i.e. groups represented by at least two individuals were included in the analysis. The single representative outgroup was included by phasing heterozygous alleles, thus meeting the minimum requirement of two sequences per lineage. In order to facilitate convergence and speed up computation, the same simplified partition scheme described in the BEAST concatenated analysis was defined, and also strict molecular clock rates were assigned to each partition. Independent models, strict clocks and trees were allowed for each gene fragment. A minimum age of 44 Ma was set to the root as described above. Three independent runs of 150 millions of generations each, sampling every 2000 generations, were run remotely on the CIPRES portal (http://www.phylo.org/ Miller *et al.* 2010). The runs were combined and resampled at lower frequency in LOGCOMBINER. A consensus chronogram was inferred with TREEANNOTATOR.

Species distribution modelling (SDM)

The potential geographic distribution range of *U. picea*, *U. algarve* and *Ummidia* sp. Tarifa (see below), species for which 6 or more records were available was inferred by the maximum entropy algorithm as implemented in Maxent 3.3.3k (Elith *et al.* 2006; Phillips *et al.* 2006). All 19 bioclimatic layers plus a digital elevation model (DEM) available at the WorldClim database (www.worldclim.org, Hijmans *et al.* 2005) were downloaded and imported into ArcGIS v. 10 (ESRI) and subsequently cropped to the same extend. Additionally, correlation matrices between the individual pairs of the 20 layers were generated in order to remove highly correlated variables from the models ($r > 0.85$, Camargo *et al.* 2013). Nine bioclimatic layers (Bio_2 = Mean Diurnal Range, Bio_3 = Isothermality, Bio_4 = Temperature Seasonality, Bio_6 = Min Temperature of Coldest Month, Bio_8 = Mean Temperature of Wettest Quarter, Bio_9 = Mean Temperature of Driest Quarter, Bio_10 = Mean Temperature of Warmest Quarter, Bio_15 = Precipitation of Seasonality, Bio_16 = Precipitation of Wettest Quarter) along with the DEM were retained for the species distribution modelling.

Model performance was evaluated by using the area under the curve (AUC) of the receiver operating characteristic (ROC) plot, dividing the occurrence points into 80%/20% train/test partitions. Species-specific parameter tuning, as recommended for the enhancement of the model performance (Anderson & Gonzalez Jr 2011), was explored under a set of different values of the regularization parameter (0.1, 0.5, 1, 1.5, 2, 3, 4) and autofeatures, suggesting the default MAXENT settings as appropriate for our data. Subsequently a set of final analyses were run with 10 replicates for each *Ummidia* species, using the cross-validation testing procedure recommended for the enhancement of the performance on small datasets (Phillips *et al.* 2006), setting the test percentage to 10%. The point wise means of the 10 replicates were use to visualize the SDMs and for further hypothesis testing.

The SDM output showed a partial overlap in *U. algarve* and *Ummidia* sp. Tarifa distributions. Ecological interchangeability between species should be expected when their niches are sufficiently similar (Rader *et al.* 2005). In order to test the potential interchangeability, niche overlap was assessed with the Relative Rank (RR) metrics (Warren & Seifert 2011) by conducting pairwise

comparisons on all three *Ummidia* species. We conducted niche identity test using 100 randomized pseudo replicates to evaluate if the SDMs RR values were statistically different (one-tailed test) than expected under the null distribution. Both RR and niche identity tests were performed with ENMTools 1.3 (Warren *et al.* 2010).

Results

Sampling, sequencing and recombination testing

A total of 145 *Ummidia* specimens were used in the present study. Detailed information about locality data and GenBank sequence accession numbers are provided in Table S1. Fig. 3 shows a map of the Mediterranean basin with the sampling localities. The sequence length for each gene fragments was as follows: 675bp for the *cox1*), 893bp for the *EF1γ*, 404bp for the *WN* and 705bp for the AL*12541*. Recombination was not detected within any of the gene fragment used in this study.

GMYC lineages

The GMYC model provided a significant better fit that the single lineage null model (LR test= 9.925394e-14), and yielded 35 entities (confidence interval=, hereafter referred as lineages, 16 of which corresponded to clusters (i.e. included more than one individual). Most GMYC lineages were restricted to single localities, with few exceptions (G6, G4, G1, G12, G14, G21, G19, G27), which always corresponded to contiguous localities. Conversely, two localities (6 and 15) included more than one GMYC lineage (G9 and G30; and G14, and G33, respectively), but in both cases they were closely allied (i.e. sister taxa).

Phylogenetic analyses

All genes were concatenated in a single matrix for Bayesian and ML tree inference. The resulting data matrix consisted of 2,690 characters: 675 (271 variables) corresponding to the namely *cox1*, 903 (133 variables) to the *EF1γ*, 404 (30 variables) to *WN*, 705 (121 variables) to *AL12541* and 3 binary coded gaps from the *EF1γ* fragment. In order to speed up the analyses, each of the 35 Mediterranean *Ummidia* GMYC lineages, as well as the New World *Ummidia*

sp. outgroup, was represented by a single representatives with full gene coverage.

The partition scheme and corresponding nucleotide substitution models selected by PARTITIONFINDER were a follows: HKY+I+G to 1st and 3rd positions of *cox1* and 3rd positions of *EF1γ* and *AL12541*, F81+I+G to 2nd of *cox1*, K80+I to 1st positions of AL*12541* and *EF1γ* and 3rd position of *WN*, HKY to 2nd positions of AL*12541* and *EF1γ* and, finally, JC+I to both 1st and 2nd of *WN.*

Both the Bayesian and the maximum likelihood (-lnL 10524.476946) analyses performed on the concatenated data matrix yielded very similar tree topologies, although overall the Bayesian inference analysis yielded slightly higher clade supports (Fig. 4). The internal relationships of Mediterranean *Ummidia* lineages were well supported and highly congruent between both methods. Both analyses supported the basal split in two clades, one including lineages from the westernmost area of the distribution (Algarve in Southern Portugal and the adjacent Huelva province in Spain), hereafter referred as the Western clade and the second one including the remaining lineages. Both analyses also supported two distinct clades within the Western clade corresponding to the coastal populations and the inland populations in Sierra Aracena and Valverde del Camino (Huelva).

Similarly, the remaining lineages formed two sister clades. One of them was mostly formed by specimens from North Africa (Tunisia, Algeria, Morocco), along with a single specimen from the coast of Granada province, hereafter referred as the Maghrebian clade. The second one, hereafter referred as the Betic-Rif clade, included three well supported clades, corresponding to the nominal species *U. aedificatoria*, from Morocco, and *U. picea*, from eastern Iberian coast, and a third one composed by specimens from the Cadiz province, in the southern-most tip of the Iberian peninsula, hereafter referred as *Ummidia* sp. Tarifa. *U. aedificatoria* and *Ummidia* sp. Tarifa were supported as sister groups. Along with these three well supported clades, three additional linages of uncertain relationships were included in the same clade, two located on the coast of Granada and a third one on the south-eastern most tip of Iberia. The

Bayesian analysis support the inclusion of these last three lineages in a single clade along with *U. picea,* hereafter referred as Eastern Iberian clade.

Estimation of divergence times

The models selected for PARTITIONFINDER for the simplified, gene-based partition were assigned as follows: GTR+I+G to *cox1*, HKY+I+G to both *EF1*γ and *AL12541* and HKY+I to *WN.* Overall, the resulting tree topology was similar to the one obtained in the Bayesian and ML analyses albeit some clades were recovered with higher support (Fig. 5a).

The root age, assigned to the split of the New World and the Mediterranean *Ummidia*, was estimated to 45.66 Ma (44.5 – 36 Ma) and the TMRCA of the Mediterranean *Ummidia* was estimated at 23.87 Ma (33.02 – 16.78 Ma). The Western clade started diversifying at 11.24 Ma (15.81 – 7.51) Ma, while both Algarve and Huelva lineages diversified at 6.33 Ma (9.93 – 3.37) and 6.09 Ma (8.78 – 3.93) respectively. The clade comprising the remaining lineages split at 16.56 Ma (22.23 – 11.91) into the Maghrebian (fully resolved in this analyses) and the Betic-Rif clades. The divergence between the *Ummidia* sp. Tarifa + *U. aedificatoria* clade and the Eastern Iberian clade traces back to about 10.66 Ma (13.93 – 7.92). *Ummidia* sp. Tarifa and *U. aedificatoria* split at 6.83 Ma (9.3 – 4.69), while the Eastern Iberian clade started diversifying around 8.56 Ma (11.3 – 6.2).

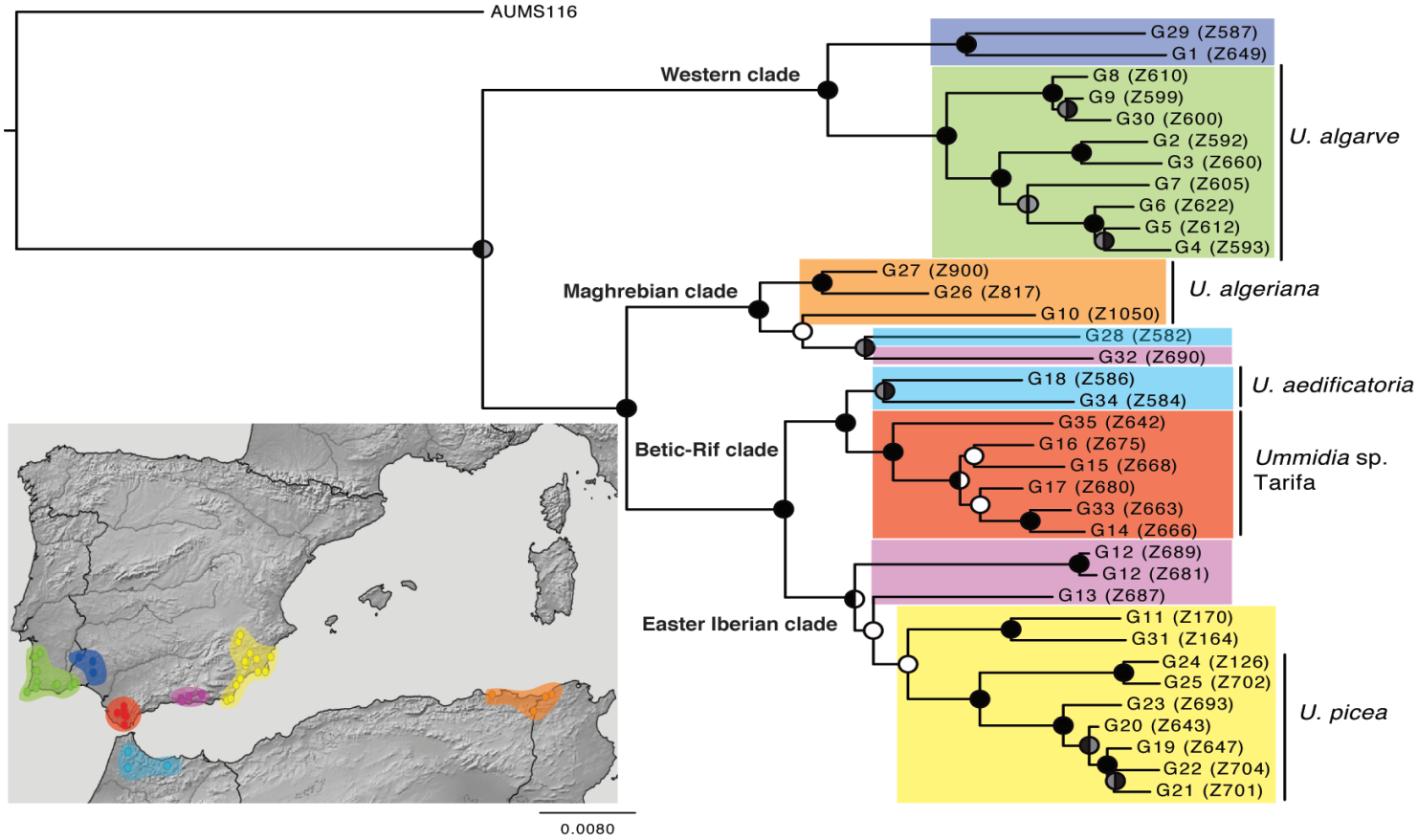


Fig. 4. - Topology obtained in the Bayesian analyses. Dots on nodes denote support as follows: right semi-circle are Bayesian posterior probabilities (PP) and left ones maximum likelihood bootstraps, black= PP> 0.95, ML bootstrap support > 70%, grey= clade recovered but with support values less than thresholds above, white= topology not recovered. Left corner: Map showing *Ummidia* sampling locations with geographically colour coded areas.

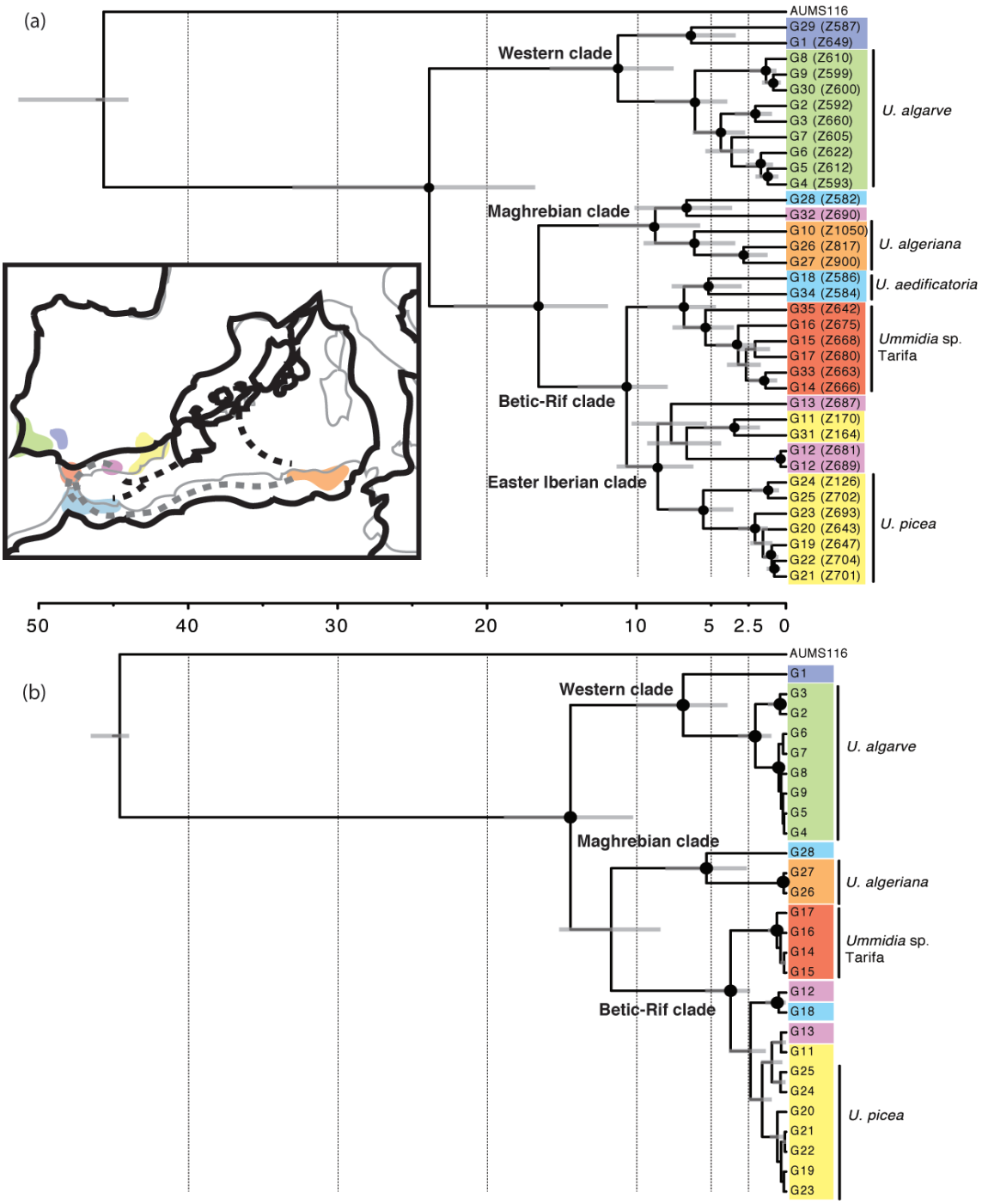


Fig. 5.- Chronograms obtained using the (a) concatenated approach and (b) the multispecies species coalescent approach, respectively. Dots on nodes denote Bayesian posterior probabilities above 0.95. Node bars indicate the 95% HPD confidence intervals of the divergence time. The x-axis is time in million years (My). Map showing Mediterranean Region in late Oligocene (25 Ma) and present and possible vicariant and dispersal events that may have contributed to present day *Ummidia* distribution. Oligocene: black lines, present: grey lines, vicariance: black dotted lines, dispersal : grey dotted lines

The coalescent-based approach

The preferred models for the simplified gene partition were: GTR+I+G to both *cox1* and *AL12541* and HKY+I+G to both *EF1γ* and *WN.*

The topology inferred by the coalescent approach was similar to the topology recovered by the concatenated analyses albeit with a lower support in some clades. The main source of incongruence is the position of *U. aedificatoria* that is shown as sister to one of the Granada lineages, instead of as a sister group of *Ummidia* sp. Tarifa (Fig. 5b), contradicting the Eastern Iberian clade supported in the concatenated analysis. Overall divergence times estimated under the coalescent approach were more recent than those reported in the concatenated analysis, especially in the younger clades. The divergence root age was estimated to 44.62 Ma (44.62 – 44.56) and the TMRCA of Mediterranean *Ummidia* to 14.4 Ma (18.09 – 10.25). The diversification of the Western clade started at about 6.9 Ma (10.04 – 3.92) and *U. algarve* split traced back to 2.05 Ma (3.22 – 0.99). The Maghrebian clade diversified at 5.34 Ma (8.08 – 2.64). The basal split within the Betic-Rif clade traces back to 3.72 Ma (5.42-2.41).

Species distribution modelling

The SDMs performed well for the three *Ummidia* species tested, *U. algarve* (AUC = 0.9860, SD = 0.011), *U. picea* (AUC = 0.97, SD = 0.04) and *Ummidia* sp. Tarifa (AUC = 0.993, SD = 0.04). Point wise means from the 10 replicate runs were use to visualize the resulting predicted ranges Fig. 6.

Results from the analysis of variable contributions indicate that the variables that have higher contribution for the model are Bio_15 (Precipitation of Seasonality), bio6 (Min Temperature of Coldest Month), bio10 (Mean Temperature of Warmest Quarter) for *U. algarve*, Bio_16 (Precipitation of Wettest Quarter), bio4 (Temperature Seasonality), DEM (altitude) for *U. picea* and Bio_15 (Precipitation of Seasonality), Bio_2 (Mean Diurnal Range), Bio_16 (Precipitation of Wettest Quarter) for *Ummidia* sp. Tarifa (Supporting information, S2). SDM predictions showed overlap between *U. algarve* and *Ummidia* sp. Tarifa distributions and identified Bio_15 (Precipitation of Seasonality) as shared climatic variable with high contribution for the model.

Pairwise comparisons of the niche overlap measured by RR metrics are

showed in the table T1. Niche identity test based on 100 randomized pseudoreplicates creating a null distribution for pair of species compared to obtained RR values did not find the SDMs statistically different (i.e. species are ecologically interchangeable) in case of *U. algarve* and *Ummidia* sp. Tarifa, alternative species pairs showed statistical difference (Fig. 7).

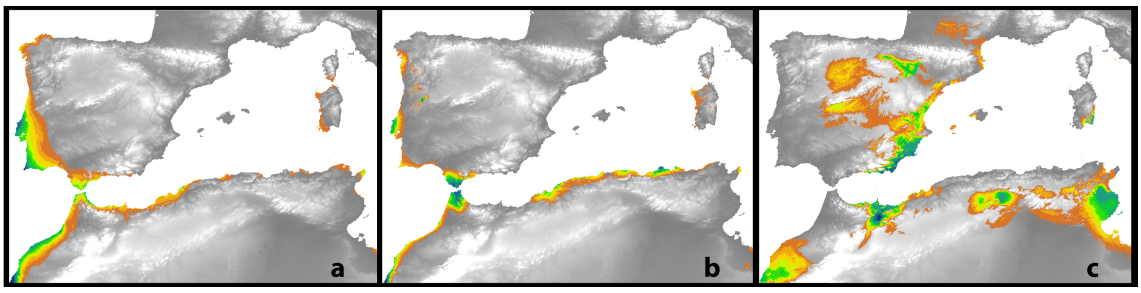


Fig. 6.- Predicted distribution for: (a) *U. algarve*, (b) *Ummidia* sp. Tarifa, (c) *U. picea*, high probability of occurrence – cooler colours, low probability of occurrence – warmer colours. Areas with less than 10% probability of occurrence – colour not shown.

Iberian *Ummidia* species pairs	RR
U. algarve vs. *Ummidia* sp. Tarifa	0.8159
Ummidia sp. Tarifa vs. *U. picea*	0.4082
U. algarve vs. *U. picea*	0.4493

Table 1. Relative ranks

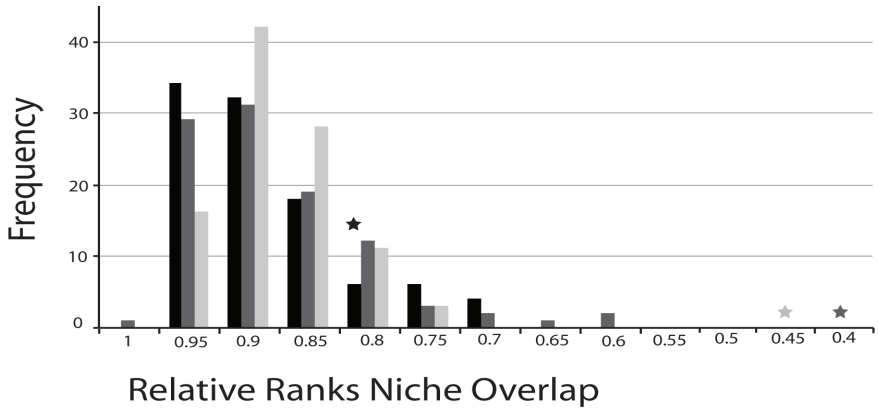


Fig. 7.- Niche simirarity results: values near 1.0 are considered identical or highly interchangeable, values near 0.0 are considered completely different or not interchangeable. Niche identity test conducted on 100 pseudoreplicates of randomized pairs of taxa (Haussler *et al.* 2009). Black: *U. algarve* and *Ummidia* sp. Tarifa, dark grey: *U. picea* and *Ummidia* sp. Tarifa, light grey: *U. algarve* and *U. picea*, stars: Relative Ranks values of niche overlap. The comparisons were not significantly different only for *U. algarve* and *Ummidia* sp. Tarifa.

Discussion

Biogeographic patterns and the timing of Ummidia diversification

The close match between the complex geology of the Western Mediterranean basin (Fig. 1) and the *Ummidia* distribution patterns was already put forward by Opatova *et al* (2013), although the limited number of samples and the low supports refrained these authors to suggest a detailed evolutionary scenario for the diversification of the lineage.

In the present study, a thorough population-level sampling across the known distribution of Mediterranean *Ummidia* and a denser sample of molecular markers resulted in a more complete and better supported phylogenetic relationships that enables us to propose a more explicit biogeographic hypothesis for the diversification of the group.

Both the topology and the inferred time frame support a close link between the diversification of *Ummidia* and the breakup of the formerly continuous land mass know as the Hercynian Belt, which stretched from the northeastern Iberian peninsula to southern France. The starting phase of the breakup resulted in the opening of the Gulf of Lyon (~30Ma) (Séranne 1999) and the Valencia Trough (~25Ma), which matches the basal split within Mediterranean *Ummidia* diversity estimated to ~ 24 Ma (16.8-33 Ma). We hypothesize that at the time of the initial fragmentation of the Hercynian Belt, *Ummidia* was probably widely distributed across Europe. The presence of *Ummidia* species in Baltic amber (Wunderlich 2000; Wunderlich 2004; Wunderlich 2011); dated to the Eocene, about 47-44 Ma, provides further support for this suggestion. The gradual world-wide cooling trend following the Eocene climatic optimum (Zachos *et al.* 2001) was accompanied by substantial extinctions in both terrestrial and marine organism (Böhme 2003), and may have contributed to the shift of *Ummidia* distribution towards southern Europe. Therefore, the break up of the Hercynian belt would have driven the divergence of Mediterranean *Ummidia* into its two main extant lineages.

Following the opening of the Valencia tough, a southward rifting of the Kabylies from the Balearic Islands and Betic-Rif blocks started in the late Miocene ~21 Ma (Rehault *et al.* 1984). The split of the Maghrebian clade was dated to 16.6 Ma (11.9–22.2 Ma), which suggests a close relationship between these two events. Under this scenario, the ancestor of *U. algeriana* and its

closely related lineages would have split from their remaining Hercynian relatives by drifting off on the Kabylies block. The westward dispersal of the Maghrebian clade towards Morocco, and eventually to Iberia, dated to ~8.7 Ma (5.7–12.5 Ma), postdates the collision of the Kabylies block with Northern Africa, ~18 to 15 Ma (Frizon de Lamotte *et al.* 2000).

The Betic-Rif plate continued drifting off westward and after a series of rotations, the plate started fragmenting around 15 Ma. The Betic and Rif blocks reached their present day location at the both sides of the present day Strait of Gibraltar by the mid-Miocene, ~10 Ma (Lonergan & White 1997). The Moroccan Rif and the southernmost part of the Iberian peninsula formed a continuous landmass isolated by marine passages from both the Betics and Morocco, respectively, as late as the Tortonian (11.6-7.2 Ma), which may explain the separation of the *U. aedificatoria* + *Ummidia* sp. Tarifa clade from its sister group, dated ~10.7 Ma (7.9-14 Ma). The subsequent split between *U. aedificatoria* and *Ummidia* sp. Tarifa, dated to 6.8 Ma (4.7-9.3 Ma), may have been the result of the reopening of the strait of Gibraltar (5.3 Ma) (Loget& Van Den Driessche 2006; Loget *et al.* 2005) following the closure of the late Miocene passages, which caused the Messinian Salinity Crises (MSC).

The divergence time estimate of the Iberian lineage G32, dated to ~6.6 Ma (3.6-10.1 Ma), also fits well with the same biogeographic scenario. The lineage is imbedded within the Maghrebian clade, which suggests an independent colonization of the Iberian Peninsula, most likely facilitated by the emergence of land bridges during the Messinian Salinity Crises. There is ample paleontological evidence of a wide faunal exchange between Africa and Iberia during the MSC (5.96 – 5.3Ma) (Agustí *et al.* 2006).

The Eastern Iberian clade started diversifying about 8.6 Ma (6.2-11.3 Ma). The distribution area of this clade corresponds to a region that was greatly affected by the uplifting of the Betic Cordillera between the lower Tortonian and the upper Messinian (~8-6 Ma). During this time period the region had reticulate character, where emerged areas were partly separated by shallow basins (for review see (Braga *et al.* 2003). The subsequent uplift of the different mountain ranges could have contributed to the isolation of some *Ummidia* lineages.

The younger divergences among the *Ummidia* lineage fall within the timeframe of the Quaternary glacial cycles, which started ~2.5 Ma (Gibbard *et*

al. 2010) and may be indicative that climatic oscillation may have played an additional role in shaping population structure in Mediterranean *Ummidia*.

While the time of diversification of *Ummidia* inferred from the concatenated approach closely matches the well-established geochronology of the Western Mediterranean, the coalescent approach yielded substantially younger time estimates. The confidence intervals inferred by the two approaches generally overlapped, except for the divergences within the Betic-Rift clade, which were much younger even considering the confidence intervals.

It should be noted that the time estimates inferred from concatenated genes tend to overestimate the divergence times since gene divergence may predate the actual population/species split (for review see Edwards & Beerli 2000; Arbogast *et al.* 2002). On the other hand, multi-species coalescent based methods assume that gene incongruence is solely the result of incomplete lineage sorting (ILS) (Heled & Drummond 2010), and it has been reported that in the presence of gene flow they may underestimate species divergence times (McCormack *et al.* 2011). Recent simulation studies have demonstrated a significant impact on the divergence time estimates in the presence of gene flow or horizontal transfer, specially when non sister taxa were involved in the exchange (Leache *et al.* 2014).

The position of *U. aedificatoria* (G18) in the coalescent analysis, which is supported as sister to one of the Granada lineages (G12), is incongruent with the results of the concatenated analysis, which support the sister group relationship of *U. aedificatoria* with the geographically closer *Ummida* sp. Tarifa. This observation could be explained as the results of hybridization and provide evidence that gene flow may be responsible for the much younger time estimates inferred in the Rift-Betic clade. As stated in the introduction *Ummidia* provides one of the few examples among mygalomorphs spiders of airborne dispersal by means of ballooning (Coyle *et al.* 1985; Eberhard 2006; Coyle 1983). This behaviour would explain the fact that distribution ranges of some *Ummidia* species bridge water gaps (Platnick 2014) and hence provide an explanation for gene flow among geographically isolated populations. However, the delimitation of coalescent groups within Mediterranean *Ummidia*, as revealed by the GMYC analysis, suggests low levels of dispersal. Most localities

formed distinctive coalescent groups and the only exception correspond to contiguous localities. Although some of the populations sharing GMYC lineages were more than 50 km apart (e.g. localities 31 and 32, or 43 and 44) they never involved dispersal over marine passages.

Alternatively, the underestimation of the divergence times by the coalescent approach could be due to the larger information required by coalescent methods, as has been recently suggested to explain differences in the estimate divergence times of igunanian lizards (Townsend *et al.* 2011) limitations, as point. The accuracy of time estimates in coalescent tree based methods increases with higher number of loci used in the analyses (Heled & Drummond 2010) and thus it is possible, that molecular markers utilized in present study did not contain sufficient information. On the other hand, support values in the coalescent species tree were much lower, especially within the Betic-Rift clade. Low supports in species-tree coalescent analyses may be the due to the uncertainty in the estimation of the population-level parameters (Leaché 2010). Some *Ummidia* lineages were sparsely sampled, which could have compromised the estimation of some of the relevant parameters.

An additional explanation for the contrasting divergence time estimates inferred by the two methods could be found in the use of informed priors on the substitution rates to infer absolute divergence times. It may well be that the much lower substitution rates constrained on some of the nuclear markers had a greater impact on the divergence time estimates in the species tree coalescent approach that they do in the concatenate approach.

Cryptic diversity in Mediterranean Ummidia

Although the taxonomic revision of Mediterranean *Ummidia* is beyond the scope of present study, it should be noted that the genus probably contains several undescribed species.

The actual diversity of *Ummidia* in the Mediterranean region had passed unnoticed until recently, when Decae re-erected two synonymized species and described a fourth one (Decae 2010). Furthermore, in a recent study of the molecular phylogeny of the family Ctenizidae, Opatova and collaborators suggested the existence of additional species within Mediterranean *Ummidia* based on the deep genetic divergences observed (Opatova *et al.* 2013).

The use of the GMYC method to delimitate coalescent groups, in the present study, yielded 35 lineages, which could be interpreted as representing putative species (Pons *et al.* 2006). Nevertheless, extreme caution should be applied before granting species status to lineages based on single, haploid markers, especially when involving taxa with limited dispersal capabilities (Bond *et al.* 2001). On the other hand, Hamilton and collaborators have shown the GMYC method yielded incongruent results when compared to other molecular based species delimitation methods and tend to over split lineages (Hamilton *et al.* 2014). The GMYC should be better interpreted as a starting point to define candidate lineages that are worth to be further tested for evolutionary independence (i.e. species status) by the addition of additional markers and phenotypic and ecological data (Powell 2012; Puillandre *et al.* 2012).

Regardless of the number of GMYC, our data clearly support the existence of additional lineages that may deserve species status based on the amount of genetic divergence in several markers, when compared to that observed in nominal taxa (e.g. the Huelva lineage or *Ummidia* sp. Tarifa). The examination of the female vulva, which is commonly used to identify females in the group, seems to support the distinct morphology of the divergent lineages (Opatova pers. obs.), but the lack of male specimens hampers further comparisons.

The overlooked diversity of *Ummidia* in the Mediterranean region may paralleled the situation in North America, where although about 20 species are currently recognized, many more await formal description (Bond pers. obs.). A formal taxonomic revision that integrates morphological, molecular and ecological data is needed before solid conclusion can be drawn about the actual diversity of *Ummida* in the Mediterranean Region.

Ecological preferences and niche interchangeability

The niche conservatism theory predicts that closely related taxa should shear the same niche more frequently than expected at random, although ecological preferences are seldom identical (Wiens & Graham 2005; Warren *et al.* 2008). Niches may be ecologically interchangeable, provided that they are sufficiently similar (Rader *et al.* 2005). Here, the habitat preferences of three species of Mediterranean *Ummidia* were assessed and further tested for niche

interchangeability to detect to what extent ecological preferences may have prevented the establishment of long term populations on new habitats.

The niches of *U. algarve* and *Ummidia* sp. Tarifa were found to be highly similar and hence interchangeable. Their niche was mostly defined by the Precipitation of Seasonality as climatic variable with high contribution for the model building. The closest geographic distance between these two lineages is about 160 km. However, the two lineages are separated by a stretch of unsuitable habitat (mostly wetlands), detected between the predicted distributions corresponding to the Guadalquivir Basin, which constitutes an effective barrier to dispersal. Similarly, the potential distribution of *U. picea* exceeded the known distribution range detecting suitable habitats on the coastline further to east. This observation provides further support for the low vagility of Mediterranean *Ummidia*, which have been unable to cross the unsuitable stretch of land in spite of the airborne dispersal ability reported for the genus. Interestingly in *Ummidia* sp. Tarifa, potential distribution was also predicted on the Moroccan side of the Strait of Gibraltar overlapping with the known range of its closely related species *U. aedificatoria.*

Conclusions

The diversity and distribution of the genus *Ummidia* in the Mediterranean were most likely shaped by the dynamic geological history of the region. The breakup of the Hercynian Belt and the subsequent migration of the individual blocks to their present day position was identified as the main diversification driver of the group. The closure of the last Miocene marine gateways between the Mediterranean and the Atlantic facilitated an independent colonization of the Iberian Peninsula from northern Africa, and the subsequent opening of the Strait of Gibraltar prevented further exchange between the two regions.

In spite of the reported capability f the genus for airborne dispersal, ecological niche interchangeability between geographically close species suggests that long distance dispersal events are rare and hence further support that the present day distribution is mostly the results of vicariant events.

Phylogenetic analyses further revealed several deeply divergent lineages that may correspond to unaccounted species, emphasizing the poor current

taxonomic knowledge and the need for an integrative taxonomic revision of the group.

Acknowledgements

We are thankful to E. Mora, R. Alvarez, E. Planas, P. Sousa, C. Frías, and S. Guirao for the assistance during the fieldwork, to M. Ferrández for kindly providing us with Z642 *Ummidia* sample from Southern Spain and to J. Korba and S. Macík for Z1050 *Ummidia* sample from Algeria. All *Ummidia* specimens from Tunisia were collected under a collection permit granted by the Institut National de Recherches en Génie Rural Eaux et Forêts. Cristina Frias-López and Julio Rozas provided information on the anonymous marker. This project was funded by Spanish Ministry of Science and Innovation (MICINN) grants CGL 2009-07639 and 2012-36863 (MA), additional financial support was provided by project 2009SGR1462 (MA) from the Generalitat de Catalunya. VO was supported by a graduate grant (AP 2008-01841) and graduate travelling grants FPU-2012, FPU-2013, also provided by MICINN. MA was further supported by an ICREA Academia award for excellence in research from the Generalitat de Catalunya.

References

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* **38**, W7-13.

Agustí J, Garcés M, Krijgsman W (2006) Evidence for African - Iberian exchanges during the Messinian in the Spanish mammalian record. *Palaeogeography, Palaeoclimatology, Palaeoecology* **238**, 5-14.

Altaba CR (1998) Testing vicariance: melanopsid snails and Neogene tectonics in the Western Mediterranean. *Journal of Biogeography* **25**, 541-551.

Anderson RP, Gonzalez Jr I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling* **222**, 2796-2811.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics* **33**, 707-740.

Arnedo MA, Ferrández MA (2007) Mitochondrial markers reveal deep population subdivision in the European protected spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae). *Conservation Genetics* **8**, 1147-1162.

Ayoub NA, Garb JE, Hedin M, Hayashi CY (2007) Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1g), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). *Molecular Phylogenetics and Evolution* **42**, 394-409.

Bauzà-Ribot, M.M., Juan, C., Nardi, F., Oromí, P., Pons, J., and Jaume, D. (2012) Mitogenomic Phylogenetic Analysis Supports Continental-Scale Vicariance in Subterranean Thalassoid Crustaceans. *Current Biology* **22**(21), 2069-2074.

Bidegaray-Batista L, Arnedo M (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology* **11**, 317.

Blondel J, Aronson J (1999) *Biology and wildlife of the Mediterranean region* Oxford University Press, Oxford ; New York.

Böhme M (2003) The Miocene Climatic Optimum: evidence from ectothermic vertebrates of Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology* **195**, 389-401.

Bond JE, Hedin MC, Ramirez MG, Opell BD (2001) Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology* **10**, 899-910.

Bond JE, Hendrixson BE, Hamilton CA, Hedin M (2012) A Reconsideration of the Classification of the Spider Infraorder Mygalomorphae (Arachnida: Araneae) Based on Three Nuclear Genes and Morphology. *PLoS ONE* **7**, e38753.

Bond JE, Stockman AK (2008) An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with

extreme genetic divergence and geographic structuring. *Systematic Biology* **57**, 628 - 646.

Braga JC, Martín JM, Quesada C (2003) and average rates of late Neogene–Recent uplift of the Betic Cordillera, SE Spain. *Geomorphology* **50**, 3-26.

Britton T (2007) Estimating Divergence Times in Large Phylogenetic Trees. *Systematic Biology* **56**, 741-752.

Camargo A, Werneck FP, Morando M, Sites JW, Jr., Avila LJ (2013) Quaternary range and demographic expansion of *Liolaemus darwinii* (Squamata: Liolaemidae) in the Monte Desert of Central Argentina using Bayesian phylogeography and ecological niche modelling. *Mol Ecol* **22**, 4038-4054.

Chapman J, Carlton J, Bellinger MR, Blakeslee AH (2007) Premature refutation of a human-mediated marine species introduction: the case history of the marine snail *Littorina littorea* in the Northwestern Atlantic. *Biological Invasions* **9**, 995-1008.

Cooper SJ, Harvey MS, Saint KM, Main BY (2011) Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. *Mol Ecol* **20**, 3219-3236.

Coyle EA, Greenstone MH, Hultsch AL, Morgan CE (1985) Ballooning mygalomorphs: Estimates of the masses of Sphodros and *Ummidia* ballooners (Araneae: Atypidae, Ctenizidae). *Journal of Arachnology* **13**, 291-296.

Coyle FA (1983) Aerial dispersal by mygalomorph spiderlings (Araneae, Mygalomorphae). *Journal of Arachnology*, 283-286.

Cunningham CW (2008) How to Use Genetic Data to Distinguish Between Natural and Human-Mediated Introduction of Littorina littorea to North America. *Biological Invasions* **10**, 1-6.

Cutler B, Guarisco H (1995) Dispersal aggregation of Sphodros fitchi (Araneae, Atypidae). *Journal of Arachnology*, 205-206.

Decae A (2010) The genus *Ummidia* Thorell 1875 in the western Mediterranean, a review (Araneae: Mygalomorphae: Ctenizidae). *Journal of Arachnology* **38**, 328-340.

Degnan JH, Rosenberg NA (2006) Discordance of Species Trees with Their Most Likely Gene Trees. *PLoS Genetics* **2**.

Drummond A, Ashton B, Buxton S*, et al.* (2012a) Geneious v5.6. *Available from http://www.geneious.com*.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012b) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*.

Duffey E (1956) Aerial dispersal in a known spider population. *Journal of Animal Ecology* **25**, 85-111.

Duffey E (1997) Aerial dispersal in spiders.

Eberhard WG (2006) Dispersal by Ummidia spiderlings (Araneae, Ctenizidae): Ancient roots of aerial webs and orientation? *Journal of Arachnology* **34**, 254-257.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54**, 1839-1854.

Elith J, Graham CH, Anderson RP*, et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129 -151.

Ezard T, Fujisawa T, Barraclough TG (2009) splits: SPecies' LImits by Threshold Statistics. Rpackage version 1.0-11/r29. . Available from http://R-Forge.R-project.org/projects/splits/.

Ferretti N, Pompozzi G, Copperi Sa, Schwerdt L (2013) Aerial dispersal by *Actinopus* spiderlings (Araneae: Actinopodidae). *The Journal of Arachnology* **41**, 407-408.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294-299.

Frizon de Lamotte D, Saint Bezar B, Bracéne R, Mercier E (2000) The two main steps of the Atlas building and geodynamics of the western Mediterranean. *Tectonics* **19**, 740-761.

Gibbard PL, Head MJ, Walker MJC (2010) Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science* **25**, 96-102.

Hamilton CA, Hendrixson BE, Brewer MS, Bond JE (2014) An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: a case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Mol Phylogenet Evol* **71**, 79-93.

Haussler D, O'Brien S, Ryder O*, et al.* (2009) A proposal to obtain whole-genome sequence for 10 000 vertebrate species. *Journal of Heredity* **100**, 659 - 674.

Hedin M, Bond JE (2006) Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): Conflict and agreement with the current system of classification. *Molecular Phylogenetics and Evolution* **41**, 454-471.

Hedin M, Starrett J, Hayashi C (2013) Crossing the uncrossable: novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. *Mol Ecol* **22**, 508-526.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**, 570-580.

Hendrixson BE, DeRussy BM, Hamilton CA, Bond JE (2013) An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, *Aphonopelma*). *Mol Phylogenet Evol* **66**, 327-340.

Hewitt G (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247 - 276.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907 - 913.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965-1978.

Jiménez-Moreno G, Fauquette S, Suc JP (2010) Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Palaeobotany and Palynology* **162**, 403-415.

Jolivet L, Augier R, Robin C, Suc J (2006) Lithospheric-scale geodynamic context of the Messinian salinity crisis. *Sedimentary Geology*.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature* **400**, 652 - 655.

Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst Biol* **56**, 17-24.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution* **29**, 1695-1701.

Leaché AD (2010) Species trees for spiny lizards (Genus *Sceloporus*): Identifying points of concordance and conflict between nuclear and mitochondrial data. *Molecular Phylogenetics and Evolution* **54**, 162-171.

Leache AD, Harris RB, Rannala B, Yang Z (2014) The influence of gene flow on species tree estimation: a simulation study. *Syst Biol* **63**, 17-30.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.

Loget N, Van Den Driessche J (2006) On the origin of the Strait of Gibraltar. *Sedimentary Geology* **188-189**, 341-356.

Loget N, Van Den Driessche J, Davy P (2005) How did the Messinian Salinity Crisis end? *Terra Nova* **17**, 414-419.
Lonergan L, White N (1997) Origin of the Betic-Rif mountain belt. *Tectonics* **16**, 504-522.
Maddison W (1997) Gene trees in species trees. *Systematic Biology* **46**, 523.

McCarthy D (2003) The trans-Pacific zipper effect: disjunct sister taxa and matching geological outlines that link the Pacific margins. *Journal of Biogeography* **30**, 1545-1561.

McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL (2011) Calibrating divergence times on species trees versus gene trees: implications for speciation history of Aphelocoma jays. *Evolution; international journal of organic evolution* **65**, 184-202.

McGuire G, Wright F (2000) TOPAL 2.0: improved detection of mosaic sequences within multiple alignments. *Bioinformatics* **16**, 130-134.

Michalak I, Zhang L-B, Renner SS (2010) Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. *Journal of Biogeography* **37**, 1214-1226.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*, pp. 1-8, New Orleans, LA.

Milne I, Lindner D, Bayer M*, et al.* (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* **25**, 126-127.

Myers N, Mittermeier RA, Mittermeier CC, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**, 853-858.

Noonan BP (2000) Does the phylogeny of pelomedusoid turtles reflect vicariance due to continental drift? *Journal of Biogeography* **27**, 1245-1249.

Noonan BP, Chippindale PT (2006) Vicariant origin of malagasy reptiles supports late cretaceous antarctic land bridge. *Am Nat* **168**, 730-741.

Opatova V, Bond JE, Arnedo MA (2013) Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae). *Mol Phylogenet Evol* **69**, 1135-1145.

Paulo OS, Pinheiro J, Miraldo A*, et al.* (2008) The role of vicariance vs. dispersal in shaping genetic patterns in ocellated lizard species in the western Mediterranean. *Mol Ecol* **17**, 1535-1551.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231-259.

Platnick NI (2014) The World Spider Catalog. American Museum of Natural History, online at http://research.amnh.org/entomology/spiders/catalog/index.html, [accessed 20 March 2014].

Pons J, Barraclough TG, Gomez-Zurita J*, et al.* (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology* **55**, 595-609.

Powell JR (2012) Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. *Methods in Ecology and Evolution* **3**, 1-11.

Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* **21**, 1864-1877.

Rader R, Belk M, Shiozawa D, Crandall K (2005) Empirical tests for ecological exchangeability. *Animal Conservation* **8**, 239-247.

Rambaut A (2009) FigTree v. 1.3.1

Rambaut A, Drummond AJ (2007) TRACER. Available from http://tree.bio.ed.ac.uk/software/tracer/ (accessed on November 30th, 2007).

Raven RJ (1980) The evolution and biogeography of the mygalomorph spider family Hexathelidae (Araneae, Chelicerata). *Journal of Arachnology* **8**, 251-266.

Raven RJ (1985) The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. *Bull. Amer. Mus. Nat. Hist.* **182**, 1-180.

Rehault JP, Boillott G, Mauffret A (1984) The western Mediterranean basin geological evolution. *Marine Geology* **55**, 447–477.

Roca E, Sans M, Cabrera L, Marzo M (1999) Oligocene to Middle Miocene evolution of the central Catalan margin (northwestern Mediterranean). *Tectonophysics* **315**, 209-229.
Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572 - 1574.

Rosenbaum G, Lister GS, Duboz C (2002a) Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. In: *Reconstruction of the evolution of the Alpine-Himalayan Orogen* (eds. Rosenbaum G, Lister GS), pp. 107-126. Journal of the Virtual Explorer.

Rosenbaum G, Lister GS, Duboz C (2002b) Relative motions of Africa, Iberia and Europe during Alpine orogeny. *Tectonophysics* **359**, 117–129.

Rosenbaum HC, Weinrich MT, Stoleson SA*, et al.* (2002c) The effect of differential reproductive success on population genetic structure: correlations of life history with matrilines in humpback whales of the gulf of maine. *J Hered* **93**, 389-399.

Rowe DL, Dunn KA, Adkins RM, Honeycutt RL (2010) Molecular clocks keep dispersal hypotheses afloat: evidence for trans-Atlantic rafting by rodents. *Journal of Biogeography* **37**, 305-324.

Sanmartin I (2003) Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography* **30**, 1883-1897.

Sanmartin I, Enghoff H, Ronquist F (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society* **73**, 345-390.

Satler JD, Starrett J, Hayashi CY, Hedin M (2011) Inferring Species Trees from Gene Trees in a Radiation of California Trapdoor Spiders (Araneae, Antrodiaetidae, *Aliatypus*). *PLoS ONE* **6**, e25355.

Schellart W (2002) Alpine deformation at the western termination of the Axial Zone, Southern Pyrenees. *Journal of the Virtual Explorer* **8**, 35-55.

Séranne M (1999) The Gulf of Lion continental margin (NW Mediterranean) revisited by IBS: an overview. *SPECIAL PUBLICATION-GEOLOGICAL SOCIETY OF LONDON* **156**, 15-36.

Silvestro D, Michalak I (2011) raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution*, DOI: 10.1007/s13127-13011-10056-13120.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49**, 369-381.

Simon E (1864) *Histoire naturelle des Araignées (Aranéides)* Roret edit., Paris.

Simon E (1910) Catalogue raisonné des Arachnides du Nord de l'Afrique, 1ere partie. *Annales de la Société Entomologique de France* **79**, 265-332.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.

Stephens M, Donnelly P (2003) A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data. *American Journal of Human Genetics* **73**, 1162-1169.

Stephens M, Smith NJ, Donnelly P (2001) A New Statistical Method for Haplotype Reconstruction from Population Data. *American Journal of Human Genetics* **68**, 978-989.

Stock JH (1993) Some remarkable distribution patterns in stygobiont Amphipoda. *Journal of Natural History* **27**, 807-819.

Suc JP (1984) Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* **307**, 429-438.

Townsend TM, Mulcahy DG, Noonan BP*, et al.* (2011) Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. *Molecular Phylogenetics and Evolution* **61**, 363-380.

Warren DL, Glor RE, Turelli M (2008) Enviromental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62**, 2868-2883.

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607-611.

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* **21**, 335-342.

Waters JM, Lopez JA, Wallis GP (2000) Molecular phylogenetics and biogeography of galaxiid fishes (Osteichthyes: Galaxiidae): dispersal, vicariance, and the position of Lepidogalaxias salamandroides. *Syst Biol* **49**, 777-795.

Wiens JJ, Graham CH (2005) Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology *Annual Review of Ecology and Systematics* **36**, 519–539

Wunderlich J (2000) Zwei neue Arten der Familie Falltrspinnen (Araneae: Ctenizidae) aus dem Baltischen Bernstein. *Entomologische Zeitschrift* **110**, 345-348.

Wunderlich J (2004) Fossil Spiders in Amber and Copal. *Beiträge zur Araneologie* **3A**, 1-850.

Wunderlich Jh (2011) Some fossil spiders (Araneae) in Eocene European ambers. In: *Beiträge zur Araneologie* (ed. Wunderlich J), pp. 472–538.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science* **292**, 686-693.

Supporting material

Table S1. Locality data and Genbank accession numbers for the new specimens sequenced in the present study, XXXXX = sequence present.

Species	Locality name/ Locality number	Country	Lat/Long	N	GMYC	Sample code	cox1	EF1g	*WN*	AL12541
Ummidia sp.	Alabama, Lee Co.	USA	32.5786N 854543W	1	-----	AUMS116	XXXXX	KF471439	XXXXX	XXXXX
Ummidia algarve	Sierra do Grandola, Tanganhal rd. 1	Portugal	38.0809N 8.53047W	4	G6	Z619	XXXXX			
						Z620	XXXXX			
						Z621	XXXXX			
						Z622	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Cercal - Santiago rd. 2	Portugal	37.97704N 8.69643W	5	G6	Z614	XXXXX			
						Z615	XXXXX			
						Z616	XXXXX	XXXXX	XXXXX	XXXXX
						Z617	XXXXX			
						Z618	XXXXX			
Ummidia algarve	Odemira rd. 3	Portugal	37.60564N 8.63275W	3	G5	Z611	XXXXX			
						Z612	XXXXX	XXXXX	XXXXX	XXXXX
						Z613	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Marmelete-Casais rd. 4	Portugal	37.29628N 8.62912W	4	G8	Z607	XXXXX			
						Z608	XXXXX			
						Z609	XXXXX	XXXXX	XXXXX	XXXXX
						Z610	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Cabo de Sao Vicente 5	Portugal	37.02757N 8.96902W	5	G7	Z602	XXXXX			
						Z603	XXXXX			
						Z604	XXXXX			
						Z605	XXXXX	XXXXX	XXXXX	XXXXX
						Z606	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Lagos - Odiáxere jct. 6	Portugal	37.13945N 8.67204W	5	G9	Z597	XXXXX	KF471461	XXXXX	XXXXX
						Z598	XXXXX			
						Z599	XXXXX	XXXXX	XXXXX	XXXXX
						Z601	XXXXX			

					G30	Z600	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Fuente do Benemola dirt rd. 7	Portugal	37.20745N 8.00779W	2	G4	Z593 Z594	XXXXX XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Barranco do Velho 8	Portugal	37.23917N 7.93617W	2	G4	Z595 Z596	XXXXX XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algarve	Playa Verde 9	Portugal	37.18823N 7.48643W	5	G2	Z598 Z589 Z590 Z591 Z592	XXXXX XXXXX XXXXX XXXXX XXXXX	XXXXX	XXXXX	XXXXX XXXXX XXXXX
Ummidia algarve	Villablanca 10	Spain	37.25949N 7.34308W	5	G3	Z657 Z658 Z659 Z660 Z661	XXXXX XXXXX XXXXX XXXXX XXXXX	XXXXX XXXXX	XXXXX XXXXX	XXXXX XXXXX
Ummidia sp.	Rosal de la Frontera 11	Spain	37.9743N 7.194W	3	G1	Z654 Z655 Z656	XXXXX XXXXX XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp.	St. Ana la Real 12	Spain	37.86559N 6.72768W	5	G1	Z649 Z652 Z651 Z650 Z653	XXXXX XXXXX XXXXX XXXXX XXXXX	XXXXX XXXXX XXXXX	XXXXX XXXXX XXXXX	XXXXX XXXXX XXXXX
Ummidia sp.	Valverde de Camino 13	Spain	37.55703N 6.71015W	1	G29	Z587	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp. Tarifa	Medina Sidonia – Vejer rd. 14	Spain	36.44859N 5.88949W	1	G14	Z667	XXXXX			
Ummidia sp. Tarifa	St. José del Valle	Spain	36.63092N 5.66284W	4	G14	Z662 Z664	XXXXX XXXXX	XXXXX	XXXXX	XXXXX

	15					Z665	XXXXX			
						Z666	XXXXX	XXXXX	XXXXX	XXXXX
					G33	Z663	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp. Tarifa	Puerto de Galiz 16	Spain	36.51784N 5.65624W	3	G17	Z678	XXXXX			XXXXX
						Z679	XXXXX		XXXXX	
						Z680	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp. Tarifa	Alcalá de los Gazules 17	Spain	36.38845N 5.65146W	5	G16	Z673	XXXXX			XXXXX
						Z674	XXXXX	XXXXX		
						Z675	XXXXX	XXXXX	XXXXX	XXXXX
						Z676	XXXXX	XXXXX		
						Z677	XXXXX	XXXXX	XXXXX	
Ummidia sp. Tarifa	La Ina 18	Spain	36.36373N 5.58438W	1	G35	Z642	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp. Tarifa	Sierra de Fates, Tarifa 19	Spain	36.12643N 5.64418W	4	G15	Z668	XXXXX	KF471463	XXXXX	XXXXX
						Z669	XXXXX	XXXXX		
						Z670	XXXXX			
						Z671	XXXXX		XXXXX	XXXXX
Ummidia sp.	Lentegí 20	Spain	36.84933N 3.71568W	1	G32	Z690	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp.	Jete, Motril 21	Spain	36.78444N 3.67206W	1	G12	Z689	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp.	Órgiva 22	Spain	36.90213N 3.44176W	5	G12	Z681	XXXXX	XXXXX	XXXXX	XXXXX
						Z682	XXXXX			
						Z683	XXXXX			
						Z684	XXXXX			
						Z685	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia sp.	Cherin 23	Spain	36.96039N 3.00783W	3	G13	Z686	XXXXX	XXXXX	XXXXX	XXXXX
						Z687	XXXXX	XXXXX	XXXXX	XXXXX
						Z688	XXXXX			
Ummidia sp.	San José, Cabo de Gata 24	Spain	36.80214N 2.14298W	6	G11	Z166	XXXXX			XXXXX
						Z167	XXXXX			
						Z168	XXXXX			
						Z169	XXXXX			
						Z170	XXXXX	XXXXX	XXXXX	XXXXX
						Z171	XXXXX	XXXXX	XXXXX	XXXXX

Ummidia sp.	Torre de los Lobos rd. Cabo de Gata 25	Spain	36.84535N 2.01746W	1	G31	Z164	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia picea	Vera 26	Spain	37.21327N 1.82724W	13	G24	Z121	XXXXX	KF471458	XXXXX	XXXXX
						Z122	XXXXX			
						Z123	XXXXX			
						Z124	XXXXX			
						Z125	XXXXX			
						Z126	XXXXX	XXXXX	XXXXX	XXXXX
						Z127	XXXXX			
						Z128	XXXXX			
						Z129	XXXXX			
						Z130	XXXXX			
						Z131	XXXXX			
						Z132	XXXXX			
						Z133	XXXXX			
Ummidia picea	Pulpí 27	Spain	37.38395N 1.69179W	2	G25	Z702	XXXXX	XXXXX	XXXXX	XXXXX
						Z703	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia picea	Totnana 28	Spain	37.758N 1.51071W	2	G22	Z704	XXXXX	XXXXX	XXXXX	XXXXX
						Z705	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia picea	Sierra de Espuña 29	Spain	37.86305N 1.53354W	3	G21	Z699	XXXXX	XXXXX		
						Z700	XXXXX			
						Z701	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia picea	Alhama de Murcia 30	Spain	37.85253N 1.47114W	1	G21	Z698	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia picea	El Cenajo 31	Spain	38.39757N 1.76317W	5	G19	Z706	XXXXX			
						Z707	XXXXX			
						Z708	XXXXX			
						Z709	XXXXX			
						Z710	XXXXX			
Ummidia picea	Tobarra 32	Spain	38.58025N 1.69101W	1	G19	Z711	XXXXX			
Ummidia picea	Archena 33	Spain	38.08524N 1.36514W	3	G19	Z646	XXXXX			
						Z647	XXXXX	XXXXX	XXXXX	XXXXX

						Z648	XXXXX			
Ummidia picea	Alberca 34	Spain	37.92965N 1.13036W	5	G19	Z107	XXXXX			
						Z108	XXXXX			
						Z109	XXXXX			
						Z110	XXXXX			
						Z111	XXXXX			
Ummidia picea	La Muela, Cartagena 35	Spain	37.58342N 1.09185W	3	G20	Z643	XXXXX	XXXXX	XXXXX	XXXXX
						Z644	XXXXX			
						Z645	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia picea	Cabo de Palos, Cartagena 36	Spain	37.58681N 0.83789W	5	G23	Z693	XXXXX	XXXXX	XXXXX	XXXXX
						Z694	XXXXX			
						Z695	XXXXX			
						Z696	XXXXX			
						Z697	XXXXX		XXXXX	XXXXX
Ummidia picea	Torrevieja 37	Spain	37.99311N 0.65458W	2	G19	Z691	XXXXX	XXXXX	XXXXX	XXXXX
						Z692	XXXXX			
Ummidia aedificatoria	Beni - Yder – Cherki 38	Morocco	35.38558N 5.5223W	2	G18	Z585	XXXXX	XXXXX	XXXXX	XXXXX
						Z586	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia aedificatoria	M.F.B. Bellota 39	Morocco	34.94961N 5.52892W	1	G34	Z584	XXXXX	KF471460	XXXXX	XXXXX
Ummidia sp.	Bni-Hadifa 40	Morocco	35.00545N 4.17899W	2	G28	Z582	XXXXX	XXXXX	XXXXX	XXXXX
						Z583	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algeriana	Skikda Region 41	Algeria	Not available	1	G10	Z1050	XXXXX			XXXXX
Ummidia algeriana	El Feija 42	Tunisia	36.50709N 8.3208E	5	G26	Z817	XXXXX	KF471467	XXXXX	XXXXX
						Z896	XXXXX			
						Z897	XXXXX			
						Z898	XXXXX			
						Z899	XXXXX	XXXXX	XXXXX	XXXXX
Ummidia algeriana	Tabarka - Aïn Draham rd. 43	Tunisia	36.8948N 8.72558E	1	G27	Z1043	XXXXX			
Ummidia algeriana	Aïn Sebaa 44	Tunisia	36.9604N 8.94213E	2	G27	Z900	XXXXX	XXXXX	XXXXX	XXXXX
						Z901	XXXXX	XXXXX	XXXXX	XXXXX

Table S1: Analyses of variable contributions for the 10-replicates run

Ummidia algarve

Variable	Percent contribution	Permutation contribution
Bio_4	57.9	1
Bio_15	19.7	28.4
Bio_6	16.1	39.8
Bio_3	1.6	0
Bio_16	1.5	3.2
Bio_9	1.2	5.1
Bio_8	0.9	5.8
Bio_2	0.5	0
Bio_10	0.3	16.6
alt	0.2	0

Ummidia picea

Variable	Percent contribution	Permutation contribution
Bio_16	49.9	63.4
Bio_4	31.5	21.5
Bio_15	8.1	5.6
alt	7.4	7
Bio_9	3	2.5
Bio_8	0.1	0
Bio_3	0	0.1
Bio_10	0	0
Bio_6	0	0
Bio_2	0	0

Ummidia sp. Tarifa

Variable	Percent contribution	Permutation contribution
Bio_15	55.9	86.3
Bio_2	23	12.3
Bio_16	16.5	1.4
Bio_4	1.2	0
Bio_10	0.3	0
Bio_3	0	0
Bio_9	0	0
Bio_6	0	0
Bio_8	0	0
alt	0	0

Chapter 3

Vera Opatova & Miquel A. Arnedo (*in press*) From Gondwana to Europe: inferring the origins of Mediterranean *Macrothele* spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae

De Gondwana a Europa: infiriendo los orígenes de las arañas Mediterráneas del género *Macrothele* (Araneae, Hexathelidae) y los límites de la familia Hexathelidae

Resumen

La familia Hexathelidae se encuentra entre las familias de arañas migalomorfas poco numerosas. La mayoría de las especies son endémicas de la región Austral-Asiática y la familia tradicionalmente ha sido considerada un ejemplo de un linaje de Gondwana. Sin embargo, los estudios recientes han puesto en duda la monofilia de la familia. *Macrothele* es el único género con una distribución en continentes que no habían formado la antigua Gondwana. La mayor parte de la diversidad de *Macrothele* se encuentra en el sudeste de Asia, pero algunas especies se conocen también de África central y dos especies habitan en Europa: *Macrothele calpeiana* de la Península Ibérica y *Macrothele cretica* endémica de Creta. En este estudio investigamos los orígenes de las especies europeas de *Macrothele* por medio de un enfoque filogenético multilocus e identificamos el marco temporal de la diversificación del género utilizando los métodos bayesianos de estima de edades implementando el reloj molecular relajado. Adicionalmente, también aportamos más conocimientos sobre el estado filogenético de la familia Hexathelidae. Nuestros resultados indican que la diversificación de *Macrothele* se remonta a la época de la fractura de Gondwana y su distribución actual muy probablemente refleja los movimientos subsecuentes de las placas tectónicas. Las dos especies europeas no se recuperan como especies hermanas, lo que sugiere que *Macrothele* colonizó la Cuenca Mediterránea en dos ocasiones de forma independiente. La polifilia de la familia Hexathelidae se confirmó adicionalmente y la subfamilia Atracinae se identifica como el linaje en conflicto.

From Gondwana to Europe: inferring the origins of Mediterranean *Macrothele* spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae

Vera Opatova & Miquel A. Arnedo

Institut de Recerca de la Biodiversitat & Departament de Biologia Animal, Universitat de Barcelona. Avinguda Diagonal 643, 08028 Barcelona, Spain

Abstract

The family Hexathelidae ranks among the smaller mygalomorph spider families. Most species are endemic to the Australasian region and the family was traditionally considered an example of a Gondwanan lineage. However, recent studies have cast some doubt on the monophyly of the family. *Macrothele* is the only genus with an out-of-Gondwana distribution. The bulk of the *Macrothele* diversity is found in Southeast Asia, few species are known from central Africa and two species inhabit Europe: *Macrothele calpeiana* from the Iberian Peninsula and *Macrothele cretica* endemic to Crete. Here we investigate the origins of the European *Macrothele* species by means of a multi-locus phylogenetic approach and by inferring the timeframe of the diversification of the genus using Bayesian relaxed clock methods. We also provide further insights into the phylogenetic status of the family Hexathelidae. Our results indicate that the diversification of *Macrothele* traces back to the period of the Gondwana breakup and its present day distribution most likely reflects the subsequent tectonic plate movements. The two European species were not recovered as sister taxa, suggesting that *Macrothele* colonized the Mediterranean region twice independently. The polyphyly of the family Hexathelidae is further confirmed and the subfamily Atracinae is identified as the conflicting lineage.

Introduction

Mygalomorph spiders, comprising tarantulas, funnel-web spiders and the trap-door spiders, are one of the three main evolutionary lineages within spiders (Hedin and Bond 2006; Platnick and Gertsch 1976). The group is of ancient origin and it is often described as 'primitive' because of the lack of specialization in the spinning structures and the retention of characters regarded as plesiomorphic such as the presence of four book lungs and chelicerae bearing longitudinal fangs with unsynchronized movement (Raven 1985). Although mygalomorphs inhabit all the continents except Antarctica, they are significantly less diverse both in number of genera and species than the Araneomorphae: there are less than 3,000 described mygalomorph species whereas over 41,000 araneomorphs are known (Platnick 2014).

The family Hexathelidae Simon, 1892 is one of the less diverse mygalomorph families. The bulk of the species richness is found in Australia and New Zealand. Overall, there are 75 species in nine genera, out of 113 described hexathelid species, which are endemic to this region (Platnick 2014). Along with two minor genera restricted to South America, the genus *Macrothele* is the only hexathelid distributed outside the Australasian region and the one with the most widespread distribution (Platnick 2014).

Raven (1980) erected the family Hexathelidae to group several diplurid-like genera that share the presence of numerous labial cuspules. The family includes four subfamilies, the Hexathelinae Simon 1892, Atracinae Hogg, 1901, Macrothelinae Simon, 1892 and Plesiothelinae Raven, 1980 (Gray 1988; Raven 1985). The monophyly of the group was first questioned in a morphology based analyses (Goloboff 1993). Subsequent phylogenetic studies using multi-locus molecular data (Ayoub *et al.* 2007; Hedin and Bond 2006), or a combination of molecular characters and morphological evidence (Bond *et al.* 2012) also failed to recover the monophyly of the family. Because, the former studies were based on a relatively modest sampling of hexathelid genera and species, no formal taxonomical changes were proposed.

The genus *Macrothele* with 26 described species ranks second among the most species-rich genera in the family and it is the most geographically widespread (Fig. 1). *Macrothele* spiders are robust, hairy spiders that construct extended and conspicuous funnel web sheets under large rocks, fallen trees

and logs and also on vertical slopes or embankments alongside the roads. These large spiders show an aggressive defensive behaviour when disturbed, and are of medical concern because their bite is considered harmful to humans (Hung and Wang 2004). The structure and properties of the toxins of some *Macrothele* species have been investigated (Corzo *et al.* 2003; Satake *et al.* 2004; Yamaji *et al.* 2009; Zeng *et al.* 2003) and used in cancer cell research (Gao *et al.* 2005; Liu *et al.* 2012).

The bulk of the *Macrothele* diversity is found in Southeast Asia, few species are known from central Africa and only two species with disjunct distributions inhabit Europe: *Macrothele calpeiana* (Walckenaer, 1805) (Fig. 2a) is found in southern Iberian Peninsula while *Macrothele cretica* Kulczynski, 1903 is endemic to Crete (Snazell and Allison 1989) (Fig. 2b). Interestingly, spiders identified as *Macrothele* sp. have been recently reported from the Antalya province in southern Turkey (Kunt *et al.* 2012)S. Huber pers. comment).

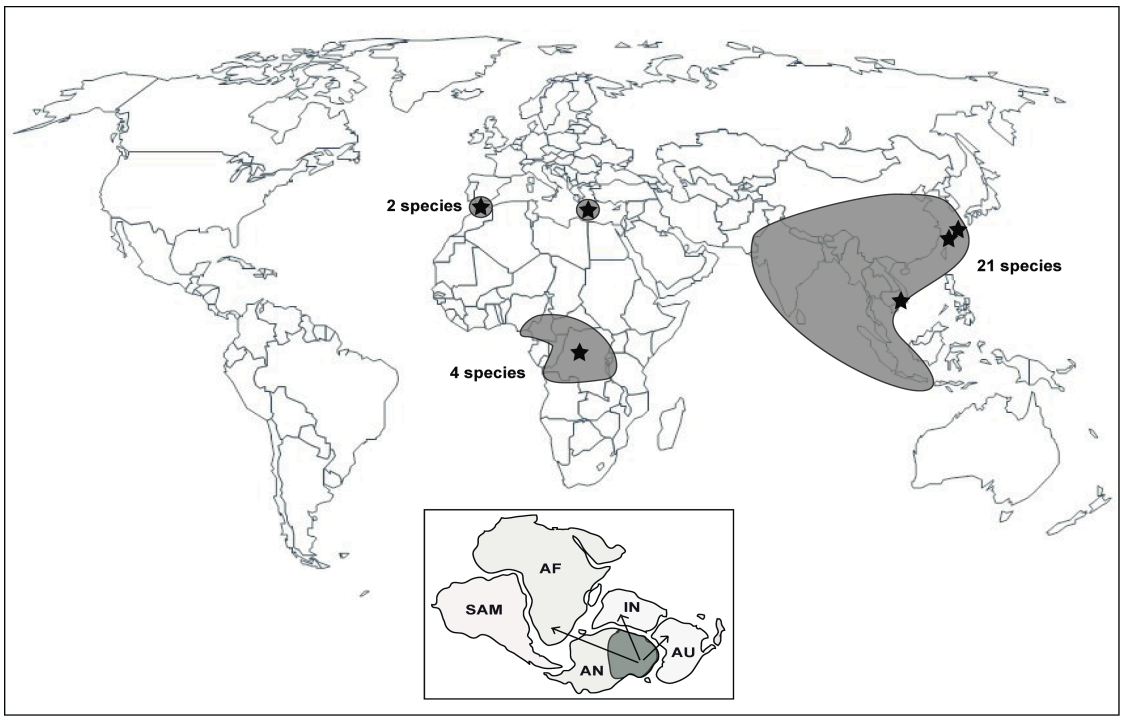


Fig. 1.- Present day distribution of the genus *Macrothele* with number of described species for each geographic area. Sampled specimen locations are marked with stars. The lower inset shows Gondwana, with the continents location, putative centre of origin and dispersal routes of *Macrothele* according to Raven (1980). AF: Africa, AN: Antarctica, AU: Australia, IN: India, SAM: South America.

Fig. 2.- *A*, Immature individual of *Macrothele cretica* from Koutsomatados, Chania, Crete. Photo credit MA. *B*, Female individual of *Macrothele calpeiana* from Sierra de Aracena, Spain. Photo credit VO.

The disjunct distributions of the European species promoted the discussion about their origins and colonization pathways into the Mediterranean region. The African origin of the Iberian *M. calpeiana* has been proposed by several authors, although they disagree about the timeframe of the colonization. Ferrández and collaborators (1998) proposed a colonization of the Iberian Peninsula in the early Tertiary, while Van Helsdingen and Decae (1992) argued for a very recent introduction. Divergence time estimates of the most recent common ancestor of extant *M. calpeiana* populations rejected both alternatives and suggested instead that *Macrothele* had colonized the Iberian Peninsula during the closure of the Strait of Gibraltar during the Messinian Salinity Crisis (MSC) (Arnedo and Ferrández 2007).

Species distribution modelling studies conducted on *M. calpeiana* predict the presence of the genus *Macrothele* in Northern Africa, the eastern Mediterranean and the middle east (Jiménez-Valverde *et al.* 2011; Jimenez-Valverde and Lobo 2007; Jiménez-Valverde and Lobo 2006a). While a *Macrothele* sp. has been recently found in the eastern part of the predicted distribution, *M. calpeiana* has never been found in Northern Africa, despite extensive sampling (Jiménez-Valverde 2009). The only confirmed locality in northern Africa corresponds to a garden in the Spanish city of Ceuta, and is probably due to a human mediated introduction (Ferrández and Ferrández de

Cespedes 2001). The absence of *Macrothele* from Northern Africa led Jimenénez-Valverde and Lobo (2007) to propose an Asian origin of the European species. The Asian origin of the genus was further supported by Haupt (2008), who highlighted the close resemblance in the genitalia of the European species and some Oriental species.

Both European *Macrothele* species are of conservation concern. *Macrothele calpeiana* is the only spider protected by European Union legislation. Its inclusion in both the Bern Convention listings and the European Union Habitat Directive under the Annex IV ‘‘Animal and plant species of community interest in need of strict protection’’ (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31992L0043:EN:HTML) was motivated by its putative role as quality bioindicator of the diminishing cork forest areas. The narrow affinity of *M. calpeiana* for primary cork oak forest was reported by several authors (Collins and Wells 1987; Snazell 1986; Snazell and Allison 1989), which led to the conclusion that the species, like the forest, could also be endangered (Snazell and Allison 1989). The close link between *M. calpeiana* and the primary cork oak vegetation and thus its vulnerability was challenged by subsequent field observations (Ferrández and Ferrández de Cespedes 2001; Van Helsdingen 1993; Van Helsdingen and Decae 1992). Van Helsdingen and Decae (1992) called into question both its role as bioindicator and vulnerability because the species seemed to thrive in human disturbed localities, and referred to *M. calpeiana* as an “aggressive invasive species”. The narrow ecological preference of the species was further ruled out by species niche modelling studies that showed that variables such as annual precipitation, and minimum and maximum temperatures had a more significant role in explaining the range of *M. calpeiana* than the presence of primary cork oak forest. (Jimenez-Valverde and Lobo 2007; Jiménez-Valverde and Lobo 2006b).

However, regardless of the environmental drivers of the distribution of *M. calpeiana*, the present day distribution of the species is highly fragmented. Genetic studies have shown deep population structure and restricted or no gene flow between small populations qualifying them as vulnerable (Arnedo and Ferrández 2007).

There is virtually no information about the biology or life cycle of *M.*

cretica, but given the fact that this large spider is only known from few localities in Western Crete, where the natural habitat might be negatively influenced by human activity such as construction or tourism development, the species has been included in the IUCN red list under the data deficient category.

Aims

This study aims to uncover the origins and colonization pathways of the genus *Macrothele* in Europe. To answer this question we use a multi-locus approach on a representative sample of *Macrothele* diversity, including specimens from all major biogeographic regions inhabited by the genus, namely Europe, tropical Africa and Southeast Asia, along with exemplar species from all Hexathelidae subfamilies, except the monotypic Plesiothelinae, and representatives of all major mygalomorph lineages. We inferred the time frame of *Macrothele* diversification using Bayesian relaxed clock models to test alternative evolutionary scenarios. Our analyses, conducted in the context of a wider mygalomorph phylogeny framework provided by previous studies (Ayoub *et al.* 2007; Bond *et al.* 2012), also enabled us to investigate the status and limits of the whole family Hexathelidae.

Materials and methods

Taxonomic sampling

The concatenated molecular data matrix of Bond *et al*. (2012), available at the Dryad Data repository http://dx.doi.org/10.5061/dryad.7sq2j, provided the starting point for the present analyses. The matrix was further modified by removing the *Megahexura* MY152 terminal, which shown a highly divergent 18S sequence, and adding the *Porrhothele* MY858 and *Macrothele* MY1024 “Myanmar” sequences available in GenBank (Hedin and Bond 2006), because of their relevance for the present analyses.

Six *Macrothele* specimens were newly sequenced in this study for the three genes sampled (see below). The specimens representing *M. calpeiana* and *M. cretica* were collected by the authors (permit SGYB-AFR-CMM-2835), *Macrothele* sp. from Congo was kindly loaned by the Royal Museum of Central Africa (RMCA) Belgium and the South East Asian *Macrothele* specimens were donated by colleagues. Locality data and Genbank accession numbers of the

Macrothele specimens sequenced in the study are summarized in Table 1. For Information on the additional specimens used in the analyses see Hedin and Bond (2006) and Bond *et al.* (2012).

Species (fam. Hexathelidae)	**Country**	**Locality**	**Lat/Long**	**18S**	**28S**	**EF-1γ**
Macrothele Z220 "Congo"	Congo D.R.	Masaco	n/a	KJ628316	KJ628316	---
Macrothele calpeiana Z47	Spain	Granada, Iznalloz	N37.36583 W3.46917	KJ628312	KJ628318	KJ628324
Macrothele cretica LB289	Greece	Crete, nr. Topolia	N35.3941 E 23.6716	KJ628311	KJ628317	KJ628323
Macrothele gigas LB169	Japan	Iriomote Is., Funaura	n/a	KJ628314	KJ628320	KJ628326
Macrothele yaginumai Z913	Japan	Iriomote Is.	n/a	KJ628313	KJ628319	KJ628325
Macrothele Z914 "Vietnam"	Vietnam	Cat Ba	n/a	KJ628315	KJ628321	KJ628327

Table 1. Locality data and Genbank accession numbers for the new specimens sequenced in the present study.

DNA extraction, PCR amplification, cloning and sequencing

Whole genomic DNA was extracted from the specimens using a single leg per individual. Extractions were performed using the SpeedTools Tissue Extraction Kit (Biotools) following the manufacturer's guidelines. Partial fragments of nuclear genes 28S rDNA (*28S*), 18S rDNA (*18S*) and Elongation factor-1 gamma (*EF1*γ) were amplified using the following primer combination; *28S* was amplified in three fragments with ZX1/rd5b, rd4.8a/rd7b1 and 2012/2762 (primer sequences available at Giribet *et al.* 1999; Hedin and Maddison 2001; Mallatt and Sullivan 1998; Schwendinger and Giribet 2005; Winchell *et al.* 2002). 18S was amplified in two overlapping fragments 1F/5R and 4F/9R (Giribet *et al.* 1996). The primers from Ayoub *et al* (2007) were used to amplify the *EF1*γ, first the pre-amplification was carried out using the "touchdown" PCR with ER1gF78/EF1gR1258 and a second "nested" PCR was performed with EF1gF218/EF1gR1090 combination.

The PCR conditions for *28S* and *18S* amplification were as follows: 5min at 94°C followed by 35 cycles of denaturation at 94°C for 30s, annealing at 64-62°C for 60s and extension for 60s at 72°C and final extension at 72°C for 5min. *EF1*γ was amplified in "nested" pcr at 52°C. For details about the "touchdown" steps see Ayoub *et al* (2007). All the reactions were carried out in a total reaction volume of 25μl of 1.25 U *Taq* polymerase (Promega), 2.5 mM $MgCl_2$ (Promega), 0.2 mM of each dNTP, 0.2 μM of each primer and 1.5μl of DNA and the amount of *Taq* buffer recommended by the manufacturer.

Macrothele sp. from Congo presented constant sequencing problems in the *28S* fragment. To obtain standard quality sequences, cloning was performed on purified PCR products with pGem®-T (Promega) set following the manufacturer's instructions. The cells with incorporated vector were plated on LB/ampicillin/IPTG/X-Gal plates and incubated at 37°C overnight. Cloned fragments were amplified using commercial vector primers 5´-ATTTAGGTGACACTATAG -3´/ 5´- TAATACGACTCACTATAGGG -3´ and following PCR conditions: 5min at 94°C followed by 30 cycles of denaturation at 94°C for 30s, annealing at 50°C for 60s, extension for 60s at 72°C and final extension at 72°C for 5min.

PCR products were purified using ExoSAP-IT (USB Corporation) and sequenced in both directions using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The sequencing was performed on ABI 3700 automated sequencer at *Centre de Recursos de Biodiversitat Animal*, University of Barcelona (CRBA, www.ub.edu/crba) Spain.

The chromatograms were assembled and edited in Geneious v. 5.3.6. (Drummond *et al.* 2010).

Phylogenetic analysis

All gene fragments used in this study presented length polymorphism due to the presence of insertions and deletions. Final sequence alignments of *28S* and *18S* were obtained using the online version of MAFFT v. 6 (available at http://mafft.cbrc.jp/alignment/server/ Katoh and Toh 2008) using the Q-INS-i approach considering the secondary structure of RNA with default settings, (gap opening penalty GOP = 1.53 and offset value set to 0.0). An online version

of multiple alignment program TranslatorX (available at http://www.translatorx.co.uk/ Abascal *et al.* 2010) was used to build the alignment of *EF1γ*, taking into account amino acid information. All three genes were concatenated in a single matrix using Geneious v. 5.3.6. Poorly aligned positions were detected and automatically removed with the help of the program trimAl (Capella-Gutiérrez *et al.* 2009). Gap information was included in the phylogenetic analyses by scoring gaps as presence/absence data following the simple coding method of Simmons and Ochoterena (2001), as implemented in the program SeqState (Muller 2006).

The concatenated data set was subject to Parsimony, Maximum Likelihood and Bayesian phylogenetic inference. A search for most parsimonious trees was conducted with the program TNT v.1.1 (Goloboff *et al.* 2003; Goloboff *et al.* 2008). Heuristic searches included 1,000 iterations of Wagner trees followed by a final round of branch swapping. Clade support was assessed by inferring trees from 1,000 Jack-knife replicates (36% probability removal) of the original dataset.

The best partition scheme and corresponding evolutionary models were determined with the help of the program PARTITIONFINDER (Lanfear *et al.* 2012) using the Bayesian Information Criterion. Maximum Likelihood (ML) analyses were conducted with RAxML v.7.2.8 (Stamatakis 2006) run under the graphical interface raxmlGUI (Silvestro and Michalak 2011). The combined matrix was partitioned by gene and *EF1γ* was further partitioned by codon position (see below). An independent GRT+G substitution model was assigned to each partition. Gaps were considered as binary characters under a general BINGAMMA model. The best maximum likelihood tree was selected from 100 iterations starting with distinct randomized parsimony trees. The support of the branches was inferred from 1000 bootstrapped matrices.

Bayesian inference analyses were conducted with MrBayes v. 3.2 (Ronquist *et al.* 2012) remotely run at the CIPRES server (http://www.phylo.org/, Miller *et al.* 2010). Unlinked models were specified for each partition as selected by PARTITIONFINDER and a standard discrete model was implemented for the gaps scored as absence/presence data, selecting the gamma rates and “all” coding options. Two independent runs of 10 million generations with 8 MCMC (Markov Chain Monte Carlo) chains each,

starting from random trees and resampling every 1000 generations were run simultaneously. The convergence of the MCMC chains was monitored examining the standard deviation of split frequencies and the correct mixing of each chain with the effective sample size (ESS) as visualized in TRACER v. 1.5 (Rambaut and Drummond 2009). The first $4x10^6$ generations were discarded as a *burn-in* for further analyses.

Fast maximum likelihood searches as implemented in RAxML (command line f a -x -m GTRCAT -N 100) were conducted to investigate the effect of missing data due to the presence of incomplete taxa and/or gene fragments of unequal length. New matrices were constructed by eliminating incomplete taxa (2 or 1 genes missing), or by removing a predefined percentage of missing data (ie. 25%, 50%, 75%) with the help of the Phyutility program (Smith and Dunn 2008).

Additionally, the trees recovered (see results) displayed an unequal distribution of branch lengths. Such asymmetry may be indicative of the existence of strong rate-heterogeneity across terminals, which may have an insidious effect on phylogenetic inference (ie. Long Branch Attraction, see Bergsten 2005 for a review). A Bayesian relative rates test (Wilcox *et al.* 2004) was conducted to detect significant long-branch terminals. The posterior probability distribution of branch lengths for all branches was obtained by resampling 5,000 trees from the MrBayes runs, after burn-in and combining runs. For each tree the distance from the most recent common ancestor (TMRCA) of the ingroup to each of the terminal taxa was calculated with the program Cadence v.1.0.1 (Wilcox 2004, available at http://www.biosci.utexas.edu/antisense/). A new maximum likelihood search was conducted on the data matrix after removal of the terminals with longer branches than the average, as revealed by the Bayesian relative rates test, using the same settings described above.

All trees were visualized and manipulated in the program FigTree v. 1.3.1 (Rambaut 2009).

The Approximately Unbiased (AU) topology test (Shimodaira 2002) as implemented in the program CONSEL (Shimodaira and Hasegawa 2001) was used to test statistical significance of alternative topologies (see below), including monophyly of the family Hexathelidae, of its nominal subfamilies, as

well as of the two Mediterranean species. Constrained searches were conducted in RAxML using the preferred partition scheme with an unlinked GTR+G model for each partition, and selecting the best-constrained tree out of 50 replicates.

Divergence time estimation

A timeframe for the diversification of *Macrothele* was inferred by means of relaxed clock models as implemented in the program BEAST v.1.7.5 (Drummond *et al.* 2012). Lineage-specific substitution rates and paleontological data were included as informed priors.

The dataset was split into the preferred partitions and their corresponding evolutionary models. Three relaxed lognormal clocks were defined, one for each gene. The *EF1γ* substitution rates estimated in Opatova *et al*. (2013) were incorporated into the divergence time analysis as normal priors for the ucld. mean (mean= 0.00117, stdev= 0.00014) of the lognormal clock of the combined *EF1γ* partitions. The rates of the other two clocks were included as parameters to estimate. Priors on the ucld.mean were defined as uniform with starting value 0.00115 and maximum and minimum bounds 0.0115 and 0.0001. The maximum bound corresponds to a universal substitution rate estimated for arthropod mtDNA (Brower 1994) and the starting value to one order of magnitude slower, based on the average slower pace of nuclear genes compared to mitochondrial ones.

Fossil information was incorporated as follows. A uniform prior was defined for the Mygalomorphae crown group, with a minimum bound 242 million years ago (Ma), based on the age of the oldest mygalomorph fossil, *Rosamygale grauvogely* Selden and Gall, 1992 from the Middle Triassic (Anisian) from the Vosges region in France, classified as a hexathelid (Selden and Gall 1992). Similarly, a uniform prior was defined for the root node, with a minimum bound at 299 Ma, which corresponds to the oldest *Mesothele* fossil, *Palaeothele montceauensis* Selden 2000, from the Upper Carboniferous (Stephanian) of Montceau-les-Mines, France (Selden 1996). Additionally, a maximum bound of 530 Ma was defined for the root, which corresponds to the average age of the earliest stem chelicerates, known from the Lower Cambrian Maotianshan Shale (Chen 2009)

Three independent chains of $5x10^7$ generations were run remotely on the CIPRES facility. Convergence and correct mixing was assessed as in the former analyses. The BEAST accompanying programs LOGCOMBINER and TREEANNOTATOR were used to combine independent runs following burn-in and to estimate a consensus chronogram.

Results

Sequences and alignments

The concatenated matrix of the three nuclear genes resulted in 6,538 aligned characters that were reduced to 5,773 after removal of uncertain alignment positions. The length by gene fragment was: 1296 characters of the *EF1γ*, 1633 of the *18S*, and 2844 of the *28S*. Scoring gaps as presence/absence added 162 characters to the matrix. The matrix included 71 terminals. Removal of taxa missing 2 genes and missing 1 gene resulted in matrices of 70 and 67 terminals, respectively. The complete data set had a 27.6% missing data, and 8.2% (4474 characters), 7% (4350) and 4.2% (4012) following removal of fragments with 25%, 50% and 75% of missing data, respectively.

Phylogenetic analysis

The parsimony analyses of the complete data matrix with gaps scored as absence/presence yielded 4 trees of 8414 steps.

The BIC criterion as implemented in PARTITIONFINDER selected the partition by gene and by codon position of the *EF1γ* as the best partition scheme. The preferred models for each partition were as follows: K80+I+G for the *EF1γ* 1st positions, GTR+I+G for both the *EF1γ* 2nd and 3rd positions, the SYM+I+G for the *18S* and GTR+I+G for the *28S*.

Both Maximum Likelihood (best ML tree score= -lnL 33480.472331) and Bayesian analyses of the concatenated matrix inferred very similar tree topologies. The parsimony topology was less resolved and showed some conflict with the model-based topology, especially among the deepest lineages. The majority rule consensus tree recovered with Bayesian analysis, including Bayesian posterior probabilities, ML bootstrap, and Parsimony Jack-knife

supports, is shown in Fig. 3. Overall, Bayesian support was slightly higher than ML support, which was in turn higher than parsimony Jack-knife.

The genus *Macrothele* was recovered as monophyletic and sister group to the remaining Avicularoidea, excluding Dipluridae, the Hexathelinae and the *Porrhothele* and *Macrothele* MY1024 individuals, in the model-based analyses, with high support in the Bayesian analysis. The position of *M.* "Congo" remained unresolved in the parsimony analyses. *Macrothele* "Congo" was recovered as sister group to the other species, with high support in most analysis, which in turn supported *M. yaginumai* as sister group to the remaining species. None of the model based analyses recovered *M. calpeiana* and *M. cretica* as sister groups, and their relationships remained unresolved in the parsimony analysis. None of the analyses supported the monophyly of Hexathelidae. The Atracinae formed a well-supported group with the Actinopodidae, and the Hexathelinae were closely related to the Dipluridae in the model based analyses, although with low support. The Macrothelinae were not monophyletic as *Porrhothele* was included in the model-based analyses within a clade with the Hexathelinae and *Macrothele* MY1024, although with low support.

Traditional mygalomorph higher-level taxonomic groups, namely the Bipectina, Crassitarsae and Domiothelina clades, already supported in Bond *et al*. (2012) were recovered without change in the present analysis.

Results of the fast ML analyses conducted on matrices with reduced percentages of missing data are summarized in Table 2. Overall, higher completeness of the data matrix did not result in better-resolved or supported trees. Similarly, removal of the terminals with significant longer branches as indicated by the Bayesian relative rate test (atypoids plus *Euagrus* and *Australothele*) resulted in a topology fully congruent with the tree inferred from the complete data matrix (data not shown).

The AU topology test (Table 3) rejected monophyly of Hexathelidae and the sister group relationships of the Atracinae with either the Macrothelinae or the Hexathelinae, while the sister group relationship of Macrothelinae and Hexathelinae could not be fully rejected based on this data set. The monophyly

of the European *Macrothele*, the inclusion of *Macrothele* MY1024 in *Macrothele* or the Macrothelinae could no be significantly rejected.

Timeframe of diversification

Results of the divergence time analysis are shown in Fig. 4. The *Macrothele* crown group started diversifying in the mid Cretaceous to early Paleogene (118.8 Ma, 197.4-53.8 Ma). The most recent common ancestor of Non-African *Macrothele* originated in the late Cretaceous to mid Eocene (60.7 Ma, 84.4-41.7 Ma) and *M. calpeiana* stem group in the Eocene to Oligocene (40.77 Ma, 58-26.6 Ma).

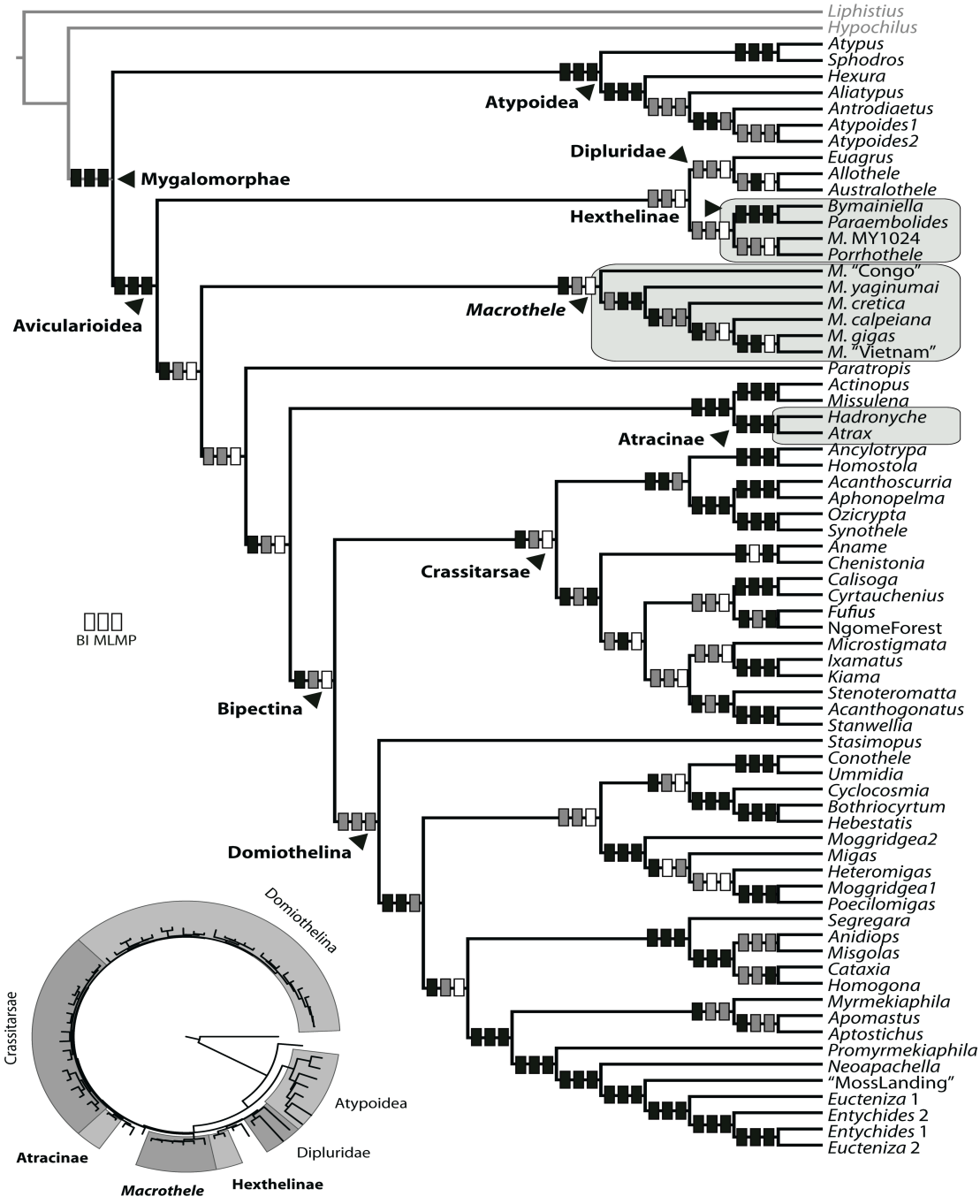


Fig. 3.- Topology obtained in the Bayesian analyses. Outgroup taxa are in grey. All hexathelid specimens are highlighted in boxes. Dots on nodes denote supports as follows: left rectangle are Bayesian posterior probabilities (PP), middle ones are maximum likelihood bootstraps, and right ones are Parsimony Jack-knifes. Black rectangles= PP> 0.95, ML bootstrap, MP Jack-knife support > 80%, grey= clade recovered but with support values less than thresholds above, white= clade not recovered.

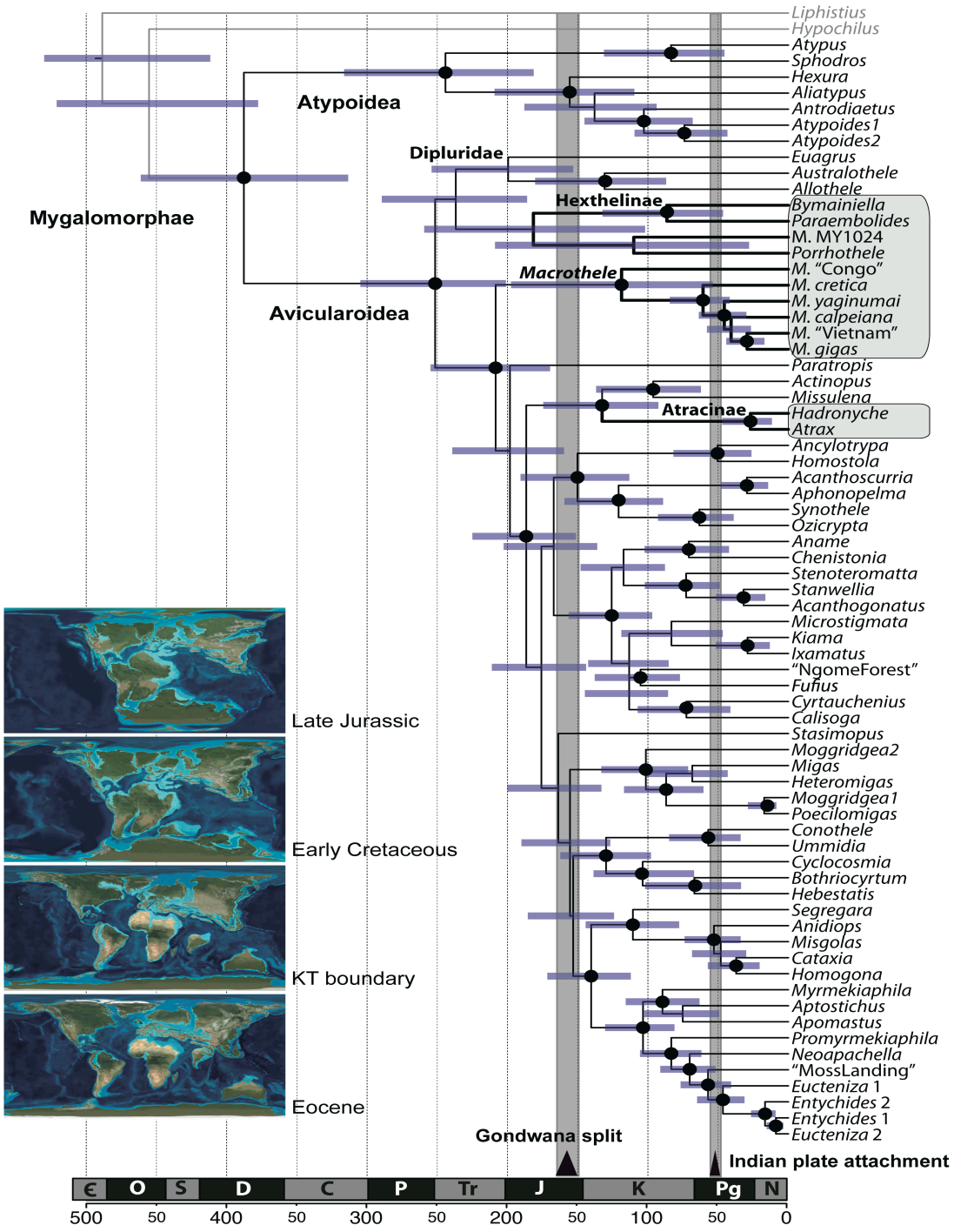


Fig. 4.- Chronogram inferred with a lognormal relaxed clock in BEAST. Dots on nodes denote Bayesian posterior probabilities >0.95. Node bars indicate the 95% HPD confidence intervals of the divergence time. The x-axis is time in million years before present (Ma). Grey vertical bars mark relevant geological events that may have affected diversification of *Macrothele* lineages. C: Cambrian, O: Ordovician, S: Silurian, D: Devonian, C: carboniferous, P: Permian, T: Triasic, J: Jurassic, K: Cretaceous, Pg: Paleogene (right to left). Insets correspond to paleogeographic reconstructions of tectonic plates at relevant time windows in the past (Blakey 2008).

Table 2.- Summary of the results of fast RAxML analyses on matrices with reduced amounts of missing data. All analyses conducted with gaps as missing data. Numbers in cells correspond to bootstrap supports. Clade: relevant clades, Full-71: complete matrix, all characters included, Trimmed-71: complete matrix, uncertain alignment position removed, *EF1γ*, *18S* and *28S*: partial gene trees, Full-70: Taxa with 2 genes missing removed, all characters, Full-67: Taxa with 1 or 2 gene missing removed, all characters included, 25%: chars with more than 25% missing removed, 50%: chars with more than 50% missing removed, 75%: chars with more than 75% missing removed NA: non applicable, NO: not recovered, *: *Macrothele* MY1024 "Myanmar" sister to *Paraembolides*, **: *M. cretica* sister to *M. calpeiana* (BS: 20).

Clade	**Full- 71**	**Trimmed 71**	***EF1γ***	***18S***	***28S***	**Full-70**	**Full-67**	**25%**	**50%**	**75%**
Hexathelidae	no	no	no	no	no	no	no	no	no	no
Atracinae	100	100	100	73	97	100	100	100	100	100
Hexathelinae	100	100	100	no*	83	100	100	100	99	99
Hexathelinae+*Porrhothele*	no	no	na	na	no	na	na	no	no	43
Hexathelinae+*Porrhothele*+*Macrothele* MY1024	22	27	na	42	no	no	na	34	15	no
Atracina +Actinopodidae	98	99	86	no	no	98	99	99	99	98
Macrothele, excluding MY1024	4	6	na	no	4	4	na	no	no	no
M. Z220 "Congo" sister to remaining *Macrothele*	no	95	100	61	no	no	100	91	94	86
M. cretica sister to remaining Eurasian *Macrothele*	no	78	76	no*	no	no	75	60	70	74
M. yaginumai sister to remaining *Macrothele,* excluding *M. cretica*)	77	77	24	no	no	82	80	73	83	73
M. Z914 "Vietnam" sister to *M. gigas*	99	97	41	96	no	100	98	95	97	94

Discussion

On the origins of Mediterranean Macrothele *species*

Geological events causing vicariance contributed greatly to the present day diversity and distribution of organisms (eg. McCarthy 2003; Sanmartin *et al.* 2001; Sanmartin and Ronquist 2004). The ancient events such as break ups of former continents or tectonic plate rearrangements are still traceable in old groups with low levels of dispersal capacity or restrictive habitat preferences (Bauzà-Ribot *et al.* 2012; Stock 1993).

Because of their low vagility, mygalomorph spiders are particularly good candidates to investigate large-scale biogeographic patterns. The low dispersal capability of mygalomorphs has been confirmed in phylogeographic studies conducted on several families, including trapdoor (Bond *et al.* 2001; Bond and Stockman 2008) and funnel web spiders (Arnedo and Ferrández 2007), and even in the family Atypidae (Pedersen and Loeschcke 2001), which uses airborne dispersal (ie. ballooning) and hence higher dispersal rates should be expected. The present day distribution of ground dwelling spiders, including mygalomorphs, is generally regarded as reflecting continental drift better than rare long distance dispersal events (Bidegaray-Batista and Arnedo 2011; Platnick and Gertsch 1976; Platnick and Nelson 1978).

According to Raven (1980), the family Hexathelidae presumably originated in East Antarctica, when it was still part of the continent forming Gondwana and spread across the continent before it split up, around 165-150 Ma (see Sanmartin and Ronquist 2004)(Fig. 1). This suggestion was based on the observation that most Hexathelidae genera are found in the Australasian region. However, hexathelids are most likely polyphyletic (Bond *et al.* 2012; Hedin and Bond 2006 and the present study). Our results suggests that *Macrothele* is the sister group of the remaining Avicularoidea with the exception of diplurids and Hexathelinae but with low support. Because the external relationships of *Macrothele* remain contentious, the reconstruction of its ancestral area is not feasible. However, the internal relationships of the genus and the inferred time frame provide some hints on the origin of the Mediterranean taxa.

The sister group relationship of the African representative to the European and Asian *Macrothele* specimens is recovered in most analyses and receives high support in the Bayesian analyses. Our time estimates indicate a mid Mesozoic age for this split (118.8 Ma, 197.4-53.8 Ma). Interestingly, this timeframe includes some major tectonic events (Fig. 4), namely the break up of Pangea about 190–180 Ma (Veevers 2004) and the subsequent disintegration of Gondwana. It is generally agreed that Africa split from Eastern Gondwana, including the Indian subcontinent sometime between 165 and 155 Ma (see Yoder and Nowak 2006 and references herein).

The geological scenario that better fits the inferred time window of the basal split between the African and Eurasian *Macrothele* corresponds to the African plate break off of Eastern Gondwana. Moreover, the estimated diversification of the Eurasian *Macrothele* (60.6 Ma, 87.4-41.7 Ma) coincides well with the collision of the Indian subcontinent to the Asian plate dated about 55 Ma (Ali and Aitchison 2008; Briggs 2003). Bringing these two pieces of information together, we suggest that *Macrothele* may have colonized Asia by rafting on the Indian subcontinent. Unfortunately, little is known about the only Indian species *Macrothele vidua* Simon 1906, other than it belongs among the large species (30 mm long), has a long slender bulb similar to *M. calpeiana* (Simon 1906) and occurs across the eastern side of the lower Himalayan plateau (West Bengal) (Gravely 1915).

Alternatively, the ancestor of *Macrothele* could have been distributed across Pangea, and the separation of the African and Eurasian species would have resulted from the split between Laurasia and Gondwana. The geological time frame of the Pangean break off overlaps marginally with our time estimates for the basal split in *Macrothele*. Interestingly, the oldest fossil evidence of a mygalomorph, *Rosamygale grauvogeli*, from the Middle Triassic of France has been interpreted as an hexathelid, which indicates that hexathelids were probably present throughout Pangea prior to its break-up (Selden and Gall 1992). However, provided Hexathelidae is a polyphyletic group, it is difficult to interpret the relevance of these results for the origin of *Macrothele*.

The next connection between Africa and Eurasia allowing the dispersal of *Macrothele* was not established until the closing of the Tethys sea in middle Miocene (ca. 16-12 Ma) (Bosworth *et al.* 2005). However this timeframe

postdates by far the diversification of most Eurasian lineages and hence could be discarded with the data at hand.

Surprisingly, the two Mediterranean species were not recovered as sister-groups. This finding suggests that the Mediterranean region was colonized by two different *Macrothele* lineages, presumably originating in Asia. Although the topological test failed to reject monophyly of the two Mediterranean species, there are morphological features that support their independent origin.

Snazell and Allison (1989) described in both sexes of *M. calpeiana* in all but the first instars, the presence of well-developed lyrae on the anterolateral face of the first leg coxa. A similar structure, located on female palpal trochanter was also found in the Chinese species *M. palpator* Pocock, 1901, and in both sexes of the species *M. gigas* Shimojana and Haupt, 1998, *M. guizhouensis* Hu and Li, 1986 and *M. raveni* Zhu and Song, 2000 from China and the Ryukyu Islands. Unlike other East Asian species of smaller size, all the former species, including *M. calepiana*, do not bear specialized setae on the male pedipalpal tibia (Hu and Li 1986; Shimojana and Haupt 1998; Zhu and Song 2000) and have female receptacula which are thin and lengthy (Haupt 2008).

The similarities between *M. calpeiana* and the Chinese species led Haupt to suggest the possibility that *M. calpeiana* was in fact the result of a recent introduction of a Chinese species in the Iberian Peninsula (Haupt 2008). This suggestion, however, was rejected by the deep intraspecific genetic divergence found between *M. calpeiana* populations (Arnedo and Ferrández 2007).

Our results provide support for the close affinities between *M. calpeiana* and *M. gigas*, which in most analyses form one clade with the *Macrothele* specimen from Vietnam. We did not observe any trace of lyrae in the Vietnam specimen, which was an immature. Unlike *M. calpeiana*, there is no information in the literature about the presence of lyrae on immature stages of the large Asian species. Confirmation of the presence of lyrae in the *M. calpeiana* + *M. gigas* + *M.* "Vietnam" clade will have to await further collections of adult specimens of the last species. Interestingly, *M. yaginumai*, which belongs to the

group of small East Asian species that lack lyrae, fell outside the clade including *M. calpeiana* and *M. gigas* in all analyses.

There is no reference to lyrae in the original description of *M. cretica,* based on a male, and were neither found in the female specimen used in the species redescription (Snazell and Allison 1989) or recently collected immature individuals. *Macrothele cretica* differs from *M. calpeiana* in additional characters such as the distance between the anterior eyes and bulb and palpal tibia morphology (Blasco and Ferrández 1986). *Macrothele cretica* also lacks the specialized setae on male pedipalpal tibia exhibited by the small East Asian species, which is congruent with the observation that *M. cretica* and *M. yaginumai* never clustered together in any analyses.

Although most authors traditionally agree on the African ancestry of *Macrothele calpeiana* (Arnedo and Ferrández 2007; Ferrández *et al.* 1998; Van Helsdingen and Decae 1992), the origin of *M. cretica* was never discussed at length. The absence of a natural African population of *M. calpeiana* and the presence of *Macrothele* spiders on the eastern part of the potential distribution inferred by species distribution modelling (Jimenez-Valverde and Lobo 2007) has been interpreted as indicative of an Asian origin of Mediterranean *Macrothele* suggesting that *Macrothele* would have colonised Europe during the late Oligocene-Early Miocene, and *M. cretica* would represent a leftover stepping stone vestige of the ancient continuous distribution (Jimenez-Valverde and Lobo 2007).

Although our results are compatible with an Asian origin of Mediterranean *Macrothele*, they also challenge the sister-group relationship of the species and hence provide a better fit for a two independent colonization events scenario. Our inferred timeframe is also older than that suggested by the single Asian origin hypothesis, Late Cretaceous-Middle Tertiary for *M. cretica* and Eocene-Oligocene for *M. calpeiana*. Climatic reconstructions of the Eocene suggest that the climate was temperate to tropical and that the forests were certainly wet (Selden and Nudds 2012). The climatic deterioration endured by the Mediterranean region since the Miocene climatic optimum may have displaced *Macrothele* species southwards to their current isolated ranges, as has been suggested for the ground-dwelling spider *Harpactocrates* (Bidegaray-Batista *et al.* 2013).

Similar to *Macrothele*, the azure-winged magpie *Cyanopica cyanus* shows an extremely disjunct distribution, which was traditionally considered as the result of an introduction from the Far East to the Iberian Peninsula by Portuguese sailors. However, a recent mitochondrial study revealed that the Iberian populations are native to the region and the extreme disjunction is better explained as the result of major extinction events throughout the previously continuous distribution, probably as a result of the Pleistocene climatic oscillations (Fok *et al.* 2002).

The taxonomic status and the limits of the family Hexathelidae

Our results further confirm the polyphyly of the family Hexathelidae found in previous molecular studies (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin and Bond 2006) and already suspected based on morphology (Goloboff 1993). Hexathelidae monophyly was based on a single character, namely the presence of numerous labial cuspules (Raven 1978; Raven 1980; Raven 1985).

Raven (1980) proposed the Hexathelidae as the Dipluridae sister family, although he subsequently modified this view to hypothesize hexathelids as sister group to the Quadrithelinae, which would include the diplurids and the Crassitarsae. Alternatively, the analyses of morphological data (Goloboff 1993) suggested Dipluridae as a grade at the base of the remaining Avicularoidea excluding hexathelids. Molecular analyses of mygalomorph interfamily relationships (Ayoub *et al.* 2007 parsimony tree, mygalomorphs only; Bond *et al.* 2012; Hedin and Bond 2006) recovered the hexathelids and the diplurids as successive sister groups to the remaining Avicularoidea, but with the hexathelids being non monophyletic. The monophyly of the family Hexathelidae was statistically rejected by Bayes factors using a *18S+28S* combined data set (Hedin and Bond 2006).

Here, we have conducted likelihood based topological tests on a data set containing *EF1*γ along the nuclear ribosomal genes. We hypothesize that the Atracinae taxa are likely contributors to hexathelid polyphyly. Topological tests could not reject any other alternative taxa combination, although the result of the Hexathelinae+Macrothelinae monophyly was close to significance. All analyses recovered the Atracinae as the sister group to the Actinopodidae with high support, which is concordant with previous analyses (Bond *et al.* 2012).

Traditionally, the Atracinae have been placed as the sister group to the Macrothelinae (Raven 1980). However, they differ in relevant characters such as the presence of a wider embolus, strong retromarginal cheliceral teeth, maxillary lobe and procurved fovea (Gray 2010). In addition, the venom protein similarities between Actinopodidae and Atracinae provides further support for this relationship, as already pointed out by Bond and collaborators (2012).

None of the analyses supported the monophyly of Hexathelinae+Macrothelinae. Wunderlich (1986; 2004) considered the Macrothelinae as a distinct lineage from the remaining hexathelids more closely related to diplurids. Our results support in part this suggestion by including the Macrothelinae *Porrhothele* along with the Hexathelinae as the sister group to the sampled diplurids. Unfortunately, this part of the tree is poorly supported and thus inconclusive in most analyses.

The monophyly of Macrothelinae was already challenged by former molecular studies that failed to retrieve a close relationship of *Porrhothele* and *Macrothele* (Hedin and Bond 2006). Our results further confirm the polyphyly of the Macrothelinae. The group was not recovered in any analyses and the separation of the bulk of *Macrothele* received high support in the Bayesian analyses. Morphological characters further corroborate the non-sister relation of *Macrothele* and *Porrhothele*. Both genera should present an incrassate first tibia in the male (Raven 1980), but such a character does not in fact occur in known males of *Macrothele* (Snazell and Allison 1989).

Interestingly, in our analyses the sample MY1024 representing *Macrothele* from Myanmar clustered with low support with *Porrhothele*, but did not cluster with the six *Macrothele* specimens newly added to the original data matrix. This might indicate the potential polyphyly of the genus *Macrothele* as defined based on morphological characters. The specimen was placed in the Hexathelinae+Dipluridae clade in most of our analyses, but the topology test could not reject that it is congeneric with the remaining *Macrothele* specimens.

A timeframe of the diversification of the mygalomorph spiders

Although the reconstruction of the family level relationships of the infraorder Mygalomorphae was beyond the scope of the present study, our time divergence results cast new light onto the timing of diversification of this group.

Ayoub and collaborators (2007) inferred a timeframe of mygalomorph diversification by using multiple fossil calibrations and implementing the NPRS method (Sanderson 1997). Although Bond and collaborators re-examined the family level phylogenetic relationships of mygalomorphs, they did not conduct a time divergence analyses. Here we reanalyse Bond *et al*'s data (2012), including new *Macrothele* sequences, and infer a time frame of diversification using Bayesian relaxed clock methods (Drummond *et al.* 2006).

Overall our time estimates are much older than those in Ayoub *et al*. This observation may be partially explained by the reinterpretation of some key taxa. In Ayoub *et al*. (2007) a maximum age on the split of araneomorphs from mygalomorphs was placed at 392 Ma, based on the age of the, at that time, oldest known spider fossil, *Attercopus* (Selden *et al.* 1991). However, more recently, *Attercopus* and other fossils of similar age have been reinterpreted as members of the new plesion Uraraneida, which may be the sister group of Araneae (Penney and Selden 2011; Selden *et al.* 2008). Therefore the age of *Attercopus* should be better interpreted as the minimum age for the origin of spiders, which is compatible with our time estimates.

Our results set the origin of mygalomorphs more than 60 Ma earlier than former studies, around 455.3 Ma (confidence interval 521.1 – 377.6 Ma), while the extant lineages started diversifying about 387.8 Ma (461.2 – 313.5 Ma).

Overall, the divergence of the major Avicularoidea lineages trace back to the mid to late Mesozoic. We further confirm the ancient origin of the families Atypidae, Dipluridae and the Hexathelinae and Macrothelinae hexathelids, which would have originated in the Triassic period predating the divergence time of most other families occurring in the late Jurassic to early Cretaceous.

Conclusions

The reanalysis of available multilocus sequences of representative mygalomorph family diversity with the addition of new sequences encompassing the diversity of *Macrothele* spiders offers new insights into the origins of a component of the European fauna of conservation concern and the taxonomic status of the family Hexathelidae. The Bayesian relaxed clock analysis reveals that the divergence of the genus *Macrothele* dates back to the period of Gondwana breakup and its present day distribution most likely reflects

the complex geological history and tectonic plate movements. The two European species may have originated in Asia but they probably colonized the Mediterranean region independently. The polyphyly of the family Hexathelidae is further confirmed, and the subfamily Atracinae is identified as the main cause for this result. Our results also set the timeframe for the mygalomorph spiders diversification further back into the past, suggesting an even more ancient origin of the group and its main evolutionary lineages.

Acknowledgements

In the first place, we would like to acknowledge the groundbreaking studies on mygalomorph molecular systematics carried out by Jason Bond, Marshall Hedin and Nadia Ayoub, which provided the foundations for the present study. We are thankful to J. Král and J. Haupt for kindly providing us with the specimens of *M. yaginumai, M. gigas* and *Macrothele* sp. from Vietnam, Rudy Joqué (RMCA) for *Macrothele* sp. from Congo and A. Schonhofer for specimen of *M. cretica*. Miguel Angel Ferrández for calling up our attention on the distribution of the lyrae in certain Asian *Macrothele* species and in *M calpeiana*. We would like to thank J. Bond, M. Rix and one anonymous reviewer for their insightful comments that greatly improved the manuscript. This project was funded by Spanish Ministry of Science and Innovation (MICINN) grant CGL 2009-07639 (MA) and additional financial support was provided by project 2009SGR1462 from the Generalitat de Catalunya. VO was supported by a graduate grant (AP 2008-01841). MA was further supported by an ICREA Academia award for excellence in research from the Generalitat de Catalunya.

References

Abascal, F., Zardoya, R., and Telford, M.J. (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research*.

Ali, J.R., and Aitchison, J.C. (2008) Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166 - 35 Ma). *Earth-Science Reviews* **88**, 145-166.

Arnedo, M.A., and Ferrández, M.A. (2007) Mitochondrial markers reveal deep population subdivision in the European protected spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae). *Conservation Genetics* **8**(5), 1147-1162.

Ayoub, N.A., Garb, J.E., Hedin, M., and Hayashi, C.Y. (2007) Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1g), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). *Molecular Phylogenetics and Evolution* **42**(2), 394-409.

Bauzà-Ribot, M.M., Juan, C., Nardi, F., Oromí, P., Pons, J., and Jaume, D. (2012) Mitogenomic Phylogenetic Analysis Supports Continental-Scale Vicariance in Subterranean Thalassoid Crustaceans. *Current Biology* **22**(21), 2069-2074.

Bergsten, J. (2005) A review of long-branch attraction. *Cladistics* **21**(2), 163-193.

Bidegaray-Batista, L., and Arnedo, M.A. (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *Bmc Evolutionary Biology* **11**.

Bidegaray-Batista, L., Ferrández, M.Á., and Arnedo, M.A. (2013) Winter is coming: Miocene and Quaternary climatic shifts shaped the diversification of Western-Mediterranean Harpactocrates (Araneae, Dysderidae) spiders. *Cladistics*, DOI: 10.1111/cla.12054

Blakey, R.C. (2008) Gondwana paleogeography from assembly to breakup—A 500 m.y. odyssey. *The Geological Society of America Special Paper* **441**, 1-28.

Blasco, A., and Ferrández, M.A. El género *Macrothele* Ausserer, 1871 (Araneae: Dipluridae) en la Península Ibérica. In 'X International Congress of Arachnology', 1986, Jaca, Spain, pp. 311-320

Bond, J.E., Hedin, M.C., Ramirez, M.G., and Opell, B.D. (2001) Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology* **10**(4), 899-910.

Bond, J.E., Hendrixson, B.E., Hamilton, C.A., and Hedin, M. (2012) A Reconsideration of the Classification of the Spider Infraorder Mygalomorphae (Arachnida: Araneae) Based on Three Nuclear Genes and Morphology. *PLoS ONE* **7**(6), e38753.

Bond, J.E., and Stockman, A.K. (2008) An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology* **57**(4), 628 - 646.

Bosworth, W., Huchon, P., and McClay, K. (2005) The Red Sea and Gulf of Aden Basins. *Journal of African Earth Sciences* **43**(1-3), 334-378.

Briggs, J.C. (2003) The biogeographic and tectonic history of India. *Journal of Biogeography* **30**(3), 381-388.

Brower, A.V.Z. (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA* **91**, 6491-6495.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*.

Chen, J.-Y. (2009) The sudden appearance of diverse animal body plansduring the Cambrian explosion. *The International journal of developmental biology* **53**(5-6), 733-751. [In English]

Collins, N.M., and Wells, S. (1987) Invertebrates in need of special protection in Europe. In '. Vol. 35.' (Ed. A H.) pp. 162. (Council of Europe: Strasbourg)

Corzo, G., Gilles, N., Satake, H., Villegas, E., Dai, L., Nakajima, T., and Haupt, J. (2003) Distinct primary structures of the major peptide toxins from the venom of the spider Macrothele gigas that bind to sites 3 and 4 in the sodium channel. *FEBS Letters* **547**(1-3), 43-50.

Drummond, A., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., and Wilson, A. (2010) Geneious v.5.3. *Available from* http://www.geneious.com

Drummond, A.J., Ho, S.Y., Phillips, M.J., and Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**(5), 699-710.

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*.

Ferrández, M.A., and Ferrández de Cespedes, H. (2001) Arachnida. In 'Los Invertebrados no insectos de la "Directiva Hábitat" en España.'. (Eds. MA Ramos, D Bragado and J Fernández) pp. 133-144. (Ministerio de Medio Ambiente. Dirección Genaral de la Conservación de la naturaleza)

Ferrández, M.A., Ferrández de Cespedes, H., and Perucho, A. (1998) *Macrothele calpeiana*, la araña negra de los alcornocales. *Quercus* **146**, 14-18.

Fok, K.W., Wade, C.M., and Parkin, D.T. (2002) Inferring the phylogeny of disjunct populations of the azure–winged magpie Cyanopica cyanus from mitochondrial control region sequences. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**(1501), 1671-1679.

Gao, L., Shan, B.-e., Chen, J., Liu, J.-h., Song, D.-x., and Zhu, B.-c. (2005) Effects of spider Macrothele raven venom on cell proliferation and cytotoxicity in HeLa cells. *Acta Pharmacol Sin* **26**(3), 369-376.

Giribet, G., Carranza, S., Baguñà, J., Riutort, M., and Ribera, C. (1996) First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Molecular Biology And Evolution* **13**(1), 76-84.

Giribet, G., Rambla, M., Carranza, S., Baguñà, J., Riutort, M., and Ribera, C. (1999) Phylogeny of the arachnid order Opiliones (Arthropoda) inferred from a combined approach of complete 18S and partial 28S ribosomal DNA sequences and morphology. *Molecular Phylogenetics and Evolution* **11**(2), 296-307.

Goloboff, P.A. (1993) A reanalysis of mygalomorph spider families (Araneae). *American Museum Novitates* **3056**, 1–32.

Goloboff, P.A., Farris, J.S., and Nixon, K.C. (2003) TNT: Tree Analysis Using New Technologies. Program and documentation available at http://www. zmuc. dk/public/phylogeny/TNT/. In '.' 1.0 edn.)

Goloboff, P.A., Farris, J.S., and Nixon, K.C. (2008) TNT, a free program for phylogenetic analysis. *Cladistics* **24**(5), 774-786.

Gravely, F.H. (1915) Records of the Indian Museum. In '. Vol. 11.' pp. 257-287. (Pub. by order of the trustees of the Indian Museum, Calcutta

Gray, M. (2010) A Revision of the Australian Funnel-web Spiders (Hexathelidae: Atracinae). *Records of the Australian Museum* **62**(2-3), 285-392. [In English]

Gray, M.R. (1988) Aspects of the systematics of the Australian funnel web spiders (Araneae: Hexathelidae: Atracinae) based upon morphological and electrophoretic data. *Aust. ent. Soc. misc. Publ.* **5**, 113-125.

Haupt, J. (2008) An organ of stridulation in East Asian Hexathelid spiders (Araneae, Mygalomorphae). *Revista Ibérica de Aracnología* **15**, 19-23.

Hedin, M., and Bond, J.E. (2006) Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): Conflict and agreement with the current system of classification. *Molecular Phylogenetics and Evolution* **41**(2), 454-471.

Hedin, M.C., and Maddison, W.P. (2001) A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). *Molecular Phylogenetics and Evolution* **18**(3), 386-403.

Hu, J.L., and Li, F.J. (1986) On two species of Macrothele from China (Araneae: Dipluridae). *Acta Zootaxonomica Sinica* **11**, 35-39.

Hung, S.-W., and Wang, T.-L. (2004) Arachnid Envenomation in Taiwan. *Annals of Disaster Medicine* **3**(Suppl 1), S12-S17.

Jiménez-Valverde, A. (2009) Absence points of Macrothele calpeiana (Walckenaer, 1805)(Araneae: Hexathelidae) in Morocco (North Africa). *Boletın de la Sociedad Entomológica Aragonesa* **44**, 559-561.

Jiménez-Valverde, A., Decae, A.E., and Arnedo, M.A. (2011) Environmental suitability of new reported localities of the funnelweb spider Macrothele calpeiana: an assessment using potential distribution modelling with presence-only techniques. *Journal of Biogeography* **38**(6), 1213-1223.

Jimenez-Valverde, A., and Lobo, J.M. (2007) Potential distribution of the endangered spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae) and the impact of climate warming. *Acta Zoologica Sinica* **53**, 865-876.

Jiménez-Valverde, A., and Lobo, J.M. (2006a) Distribution determinants of endangered Iberian spider *Macrothele calpeiana* (Araneae, Hexathelidae). *Environmental Entomology* **35**(6), 1491-1499.

Jiménez-Valverde, A., and Lobo, J.M. (2006b) Establishing reliable spider (Araneae, Araneidae and Thomisidae) assemblage sampling protocols: estimation of species richness, seasonal coverage and contribution of juvenile data to species richness and composition. *Acta Oecologica* **30**(1), 21-32.

Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**(4), 286-298.

Kunt, K.B., Kaya, R.S., Özkütük, R.S., Danisman, T., Yagmur, E.A., and Elverici, M. (2012) Additional notes on the spider fauna of Turkey (Araneae). *Turkish Journal of Zoology* **36**(5), 637-651.

Lanfear, R., Calcott, B., Ho, S.Y.W., and Guindon, S. (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution* **29**(6), 1695-1701.

Liu, Z., Zhao, Y., Li, J., Xu, S., Liu, C., Zhu, Y., and Liang, S. (2012) The venom of the spider Macrothele raveni induces apoptosis in the myelogenous leukemia K562 cell line. *Leukemia Research* **36**(8), 1063-1066.

Mallatt, J., and Sullivan, J. (1998) 28S and 18S rDNA Sequences Support the Monophyly of Lampreys and Hagfishes. *Molecular Biology and Evolution* **15**(12), 1706-1718.

McCarthy, D. (2003) The trans-Pacific zipper effect: disjunct sister taxa and matching geological outlines that link the Pacific margins. *Journal of Biogeography* **30**, 1545–1561.

Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 'Proceedings of the Gateway Computing Environments Workshop (GCE).' pp. 1-8: New Orleans, LA)

Muller, K. (2006) Incorporating information from length-mutational events into phylogenetic analysis. *Molecular Phylogenetics and Evolution* **38**, 667-676.

Opatova, V., Bond, J.E., and Arnedo, M.A. (2013) Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae). *Molecular Phylogenetics and Evolution* **69**(3), 1135-1145.

Pedersen, A.A., and Loeschcke, V. (2001) Conservation genetics of peripheral populations of the mygalomorph spider *Atypus affinis* (Atypidae) in northern Europe. *Molecular Ecology* **10**(5), 1133-1142.

Penney, D., and Selden, P.A. (2011) 'Fossil Spiders: The Evolutionary History of a Mega-diverse Order.' (Siri Scientific Press: Manchester, UK)

Petrunkevitch, A. (1946) Fossil spiders in the collection of the American Museum of Natural History. *American Museum Novitates* **1328**, 1-36.

Platnick, N.I. (2014) The World Spider Catalog. In '.' v. 14.5 edn. (American Museum of Natural History, New York, online at http://research.amnh.org/iz/spiders/catalog, [accessed 26 March 2014])

Platnick, N.I., and Gertsch, W.J. (1976) The suborder of spiders: A cladistic analysis. *American Museum Novitates* **2607**, 1-15.

Platnick, N.I., and Nelson, G. (1978) A Method of Analysis for Historical Biogeography. *Systematic Biology* **27**(1), 1-16.

Rambaut, A. (2009) FigTree v.1.3.1. In 1.3.1 edn. (Available from http://tree.bio.ed.ac.uk/software/figtree/ (accessed on March 26th, 2010))

Rambaut, A., and Drummond, A.J. (2009) TRACER v.1.5. In 1.4 edn. (Available from http://tree.bio.ed.ac.uk/software/tracer/ (accessed on December 1st, 2009))

Raven, R. (1978) Systematics of spider subfamily Hexathelinae (Dipluridae-Mygalomorphae-Arachnida). *Australian Journal of Zoology*, 1-75.

Raven, R.J. (1980) The evolution and biogeography of the mygalomorph spider family Hexathelidae (Araneae, Chelicerata). *Journal of Arachnology* **8**(3), 251-266.

Raven, R.J. (1985) The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. *Bulletin of American Museum of Natural History* **182**(1), 1-180.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology* **61**(3), 539-542.

Sanderson, M.J. (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* **14**(12), 1218-1231.

Sanmartin, I., Enghoff, H., and Ronquist, F. (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society* **73**(4), 345-390.

Sanmartin, I., and Ronquist, F. (2004) Southern Hemisphere Biogeography Inferred by Event-Based Models: Plant versus Animal Patterns. *Systematic Biology* **53**(2), 216-243.

Satake, H., Villegas, E., Oshiro, N., Terada, K., Shinada, T., and Corzo, G. (2004) Rapid and efficient identification of cysteine-rich peptides by random screening of a venom gland cDNA library from the hexathelid spider Macrothele gigas. *Toxicon* **44**(2), 149-156.

Schwendinger, P.J., and Giribet, G. (2005) The systematics of the south-east Asian genus *Fangensis* Rambla (Opiliones: Cyphophthalmi: Stylocellidae). *Invertebrate Systematics* **19**(4), 297-323.

Selden, P.A. (1996) First fossil mesothele spider, from the Carboniferous of France. *Revue suisse de Zoologie hors série*, 585-596.

Selden, P.A., and Gall, J.C. (1992) A Triassic mygalomorph spider from the northern Vosges, France. *Palaeontology* **35**, 211–235.

Selden, P.A., and Nudds, J.R. (2012) ' Evolution of Fossil Ecosystems.' second edn. (Manson Publishing Ltd.: London) 288

Selden, P.A., Shear, W.A., and Bonamo, P.M. (1991) A spider and other arachnids from the Devonian of New York, and a reinterpretation of Devonian Araneae. *Palaeontology* **34**, 241-281.

Selden, P.A., Shear, W.A., and Sutton, M.D. (2008) Fossil evidence for the origin of spider spinnerets, and a proposed arachnid order. *Proceedings of the National Academy of Sciences* **105**(52), 20781-20785.

Shimodaira, H. (2002) An Approximately Unbiased Test of Phylogenetic Tree Selection. *Systematic Biology* **51**, 492 - 508.

Shimodaira, H., and Hasegawa, M. (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246 - 1247.

Shimojana, M., and Haupt, J. (1998) Taxonomy and natural history of the funnel-web spider genus Macrothele (Araneae: Hexathelidae: Macrothelinae) in the Ryukyu Islands (Japan) and Taiwan. *Species Diversity* **3**(1), 1-15.

Silvestro, D., and Michalak, I. (2011) raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution*, DOI: 10.1007/s13127-011-0056-0.

Simmons, M.P., Ochoterena, H., and Carr, T.G. (2001) Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Systematic Biology* **50**(3), 454-462.

Simon, E. (1906) Arachnides (2e partie). In, Voyage de M. Maurice Maindron dans l'Inde méridionale. *Annales de la Société entomologique de France* **75**, 279-314.

Smith, S.A., and Dunn, C.W. (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* **24**(5), 715-716.

Snazell, R.G. (1986) The spider genus *Macrothele* Ausserer in Spain (Araneae, Dipluridae). *Bulletin of the British Arachnological Society* **17**(1), 80-83.

Snazell, R.G., and Allison, R. (1989) The genus Macrothele Ausserer (Araneae; Hexathelidae) in Europe. *Bulletin of the British Arachnological Society* **8**, 65-72.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**(21), 2688-2690.

Stock, J.H. (1993) Some remarkable distribution patterns in stygobiont Amphipoda. *Journal of Natural History* **27**(4), 807-819.

Van Helsdingen, P.J. (1993) Can *Macrothele calpeiana* (Walckenaer) (Araneae, Hexathelidae) be used as a bio - indicator? *Bulletin de la Société neuchâteloise des sciences naturelles* **116**, 253-258.

Van Helsdingen, P.J., and Decae, A. (1992) Ecology, distribution and vulnerability of *Macrothele calpeiana* (Walckenaer) (Araneae, Hexathelidae). *Tljdschrift voor Entomologie* **135**, 169-178.

Veevers, J.J. (2004) Gondwanaland from 650 - 500 Ma assembly through 320 Ma merger in Pangea to 185, 100 Ma breakup: supercontinental tectonics via stratigraphy and radiometric dating. *Earth-Science Reviews* **68**, 1-132.

Wilcox, T.P. (2004) Cadence v. 1.0, available at http://www.biosci.utexas.edu/antisense/. In '.' 1.0 edn.)

Wilcox, T.P., Garcia de Leon, F.J., Hendrickson, D.A., and Hillis, D.M. (2004) Convergence among cave catfishes: long-branch attraction and a Bayesian relative rates test. *Molecular Phylogenetics and Evolution* **31**(3), 1101-1113.

Winchell, C.J., Sullivan, J., Cameron, C.B., Swalla, B.J., and Mallatt, J. (2002) Evaluating Hypotheses of Deuterostome Phylogeny and Chordate Evolution with New LSU and SSU Ribosomal DNA Data. *Molecular Biology and Evolution* **19**, 762-776.

Wunderlich, J. (1986) 'Spinnenfauna gestern und heute.' (Publishing House Joerg Wunderlich: 69493 Hirschberg, Germany) 283

Wunderlich, J. (2004) Fossil Spiders in Amber and Copal. *Beiträge zur Araneologie* **3A**, 1-850.

Yamaji, N., Little, M.J., Nishio, H., Billen, B., Villegas, E., Nishiuchi, Y., Tytgat, J., Nicholson, G.M., and Corzo, G. (2009) Synthesis, Solution Structure, and Phylum Selectivity of a Spider δ-Toxin That Slows Inactivation of Specific Voltage-gated Sodium Channel Subtypes. *Journal of Biological Chemistry* **284**(36), 24568-24582.

Yoder, A.D., and Nowak, M.D. (2006) Has Vicariance or Dispersal Been the Predominant Biogeographic Force in Madagascar? Only Time Will Tell. *Annual Review of Ecology, Evolution, and Systematics* **37**(1), 405-431.

Zeng, X., Xiao, Q., and Liang, S. (2003) Purification and characterization of raventoxin-I and raventoxin-III, two neurotoxic peptides from the venom of the spider *Macrothele raveni*. *Toxicon* **41**(6), 651-656.

Zhu, M.S., and Song, D.X. (2000) Review of the Chinese funnel-web spiders of the genus Macrothele, with descriptions of two new species (Araneae: Hexathelidae). *Raffles Bulletin of Zoology* **48**(1), 59-64.

Chapter 4

Vera Opatova & Miquel A. Arnedo. Spiders on a hot volcanic roof: Colonisation pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae)

Las arañas sobre el tejado volcánico caliente: Las vías de colonización y la filogeografía de la araña migalomorfa endémica a las Islas Canarias *Titanidiops canariensis* (Araneae, Idiopidae)

Resumen

Los estudios realizados en las islas volcánicas han contribuido en gran medida a la comprensión actual de la diversificación de los organismos. El archipiélago de las Islas Canarias, situado en el Océano Atlántico, al noroeste de la costa del norte de África, alberga un gran número de especies endémicas. Debido a la baja capacidad de dispersión, las arañas migalomorfas son generalmente ausentes de las islas oceánicas. *Titanidiops canariensis*, que habita las islas orientales del archipiélago, constituye una excepción a esta regla. En este estudio, se utiliza un enfoque multilocus que combina tres genes mitocondriales y cuatro genes nucleares para investigar los orígenes y la filogeografía de esta extraordinaria especie. Proporcionamos un marco temporal para la colonización de las Islas Canarias con dos enfoques alternativos: concatenación y la reconstrucción del árbol de la especie en un marco de inferencia bayesiana con reloj molecular relajado. Adicionalmente, se investiga la existencia de especies crípticas en las islas por medio de un método bayesiano de delimitación de especies a partir de múltiples loci. Nuestros resultados indican que *T. canariensis* colonizó las Islas Canarias una vez, muy probablemente durante el Mioceno, aunque las discrepancias entre las estimas de edad obtenidas en los diferentes enfoques no permiten aproximar el marco temporal exacto. Un patrón filogeográfico complejo reveló que la colonización de Lanzarote, podía haber transcurrido en dos eventos independientes o una colonización de ida y vuelta entre Fuerteventura y Lanzarote. Adicionalmente, los datos corroboran un refugio volcánico reconocido previamente, destacando el impacto de la dinámica historia volcánica de las islas en los patrones filogeográficos de los taxones endémicos. *T. canariensis* abarca por lo menos dos especies diferentes, una que habita la península de Jandía y el centro de Fuerteventura y la segunda ocupa la zona entre el centro de Fuerteventura a Lanzarote norte. Los datos también sugieren que los linajes de *Titanidiops* del norte de África se podían haber extendido a la región después de que las islas habían sido colonizadas. Además, la

diversidad de *Titanidiops* en el norte de África está probablemente infraestimada.

Spiders on a hot volcanic roof: Colonisation pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae)

Vera Opatova and Miquel A. Arnedo

[1]*Institut de Recerca de la Biodiversitat & Departament de Biologia Animal, Universitat de Barcelona. Avinguda Diagonal 643, 08028 Barcelona, Spain.*

Abstract

Studies conducted on volcanic islands have greatly contributed to the current understanding of the diversification of organisms. The Canary Islands archipelago, located in the Atlantic Ocean, northwest of the coast of northern Africa, harbours a large number of endemic taxa. Because of their low vagility, mygalomorph spiders are usually absent from oceanic islands. The spider *Titanidiops canariensis*, which inhabits the easternmost islands of the archipelago, constitutes an exception to this rule. Here, we use a multi-locus approach that combines three mitochondrial and four nuclear genes to investigate the origins and phylogeography of this remarkable trap-door spider. We provide a timeframe for the colonisation of the Canary Islands using two alternative approaches: concatenation and species tree inference in a Bayesian relaxed clock framework. Additionally, we investigate the existence of cryptic species on the islands by means of a Bayesian multi-locus species delimitation method. Our results indicate that *T. canariensis* colonised the Canary Islands once, most likely in the Miocene, although discrepancies between the timeframes from different approaches make the exact timing uncertain. A complex evolutionary history for the species in the archipelago is revealed, which involves either two independent colonisations of Lanzarote or a back-and-forth dispersal between Fuerteventura and Lanzarote. The data further corroborate a previously proposed volcanic refugium, highlighting the impact of the dynamic volcanic history of the island on the phylogeographic patterns of the endemic taxa. *T. canariensis* includes at least two different species, one inhabiting the Jandia peninsula and central Fuerteventura and one spanning

from central Fuerteventura to Lanzarote. Our data also suggest that the extant northern African *Titanidiops* species may have expanded to the region after the islands were colonised and, hence, are not the source of colonisation. In addition, the present day diversity of *Titanidiops* in northern Africa is greatly underestimated.

Key words: Oceanic archipelago; species delimitation; concatenation; species tree; volcanic refugia

Introduction

Oceanic islands of volcanic origin are ideal systems for evolutionary studies [1]. Episodes of volcanic activity have left their fingerprints on the genetic diversity and distribution of endemic terrestrial organisms. Recurrent range shifts, geographic isolation and population bottlenecks driven by lava flows have shaped the complex phylogeographic patterns and have led speciation in local organisms [2-9].

The Canary Islands archipelago lies in the Atlantic Ocean, approximately 110 km from the north-western coast of Africa, comprising seven main islands and several smaller islets (Fig. 1). The region harbours a significant number of endemic organisms; 50 % of the known invertebrates and 27 % of the vascular plants inhabiting the archipelago are local endemics. This extraordinary biological richness has been traditionally interpreted as a relict of the Tertiary Mediterranean diversity, but the advent of molecular phylogenetics revealed a large amount of in situ diversification [10]. Some groups, however, have colonised the archipelago repeatedly [11-14]. The archipelago was built by several cycles of volcanic activity tracing back to the Miocene and persisting until today [15-17].

The islands on the eastern side were the first to emerge, and the remaining islands appeared subsequently, following an east to west pattern [15]. Fuerteventura, Lanzarote and the surrounding islets, hereafter referred to as the Eastern Canary Islands, are the emergent parts of a volcanic ridge that runs parallel to the African coast. The islands are separated by shallow waters and have been repeatedly connected during marine transgressions [18]. The

sub-aerial stage of Fuerteventura began approximately 20 million years ago (Ma), and the volcanism progressed in a SSW-NNE direction. After several rounds of volcanic activity followed by periods of erosional quiescence, post-erosional volcanic activity began in Lanzarote approximately 1.6 Ma, and eruptions have been documented even in historic times [18,19].

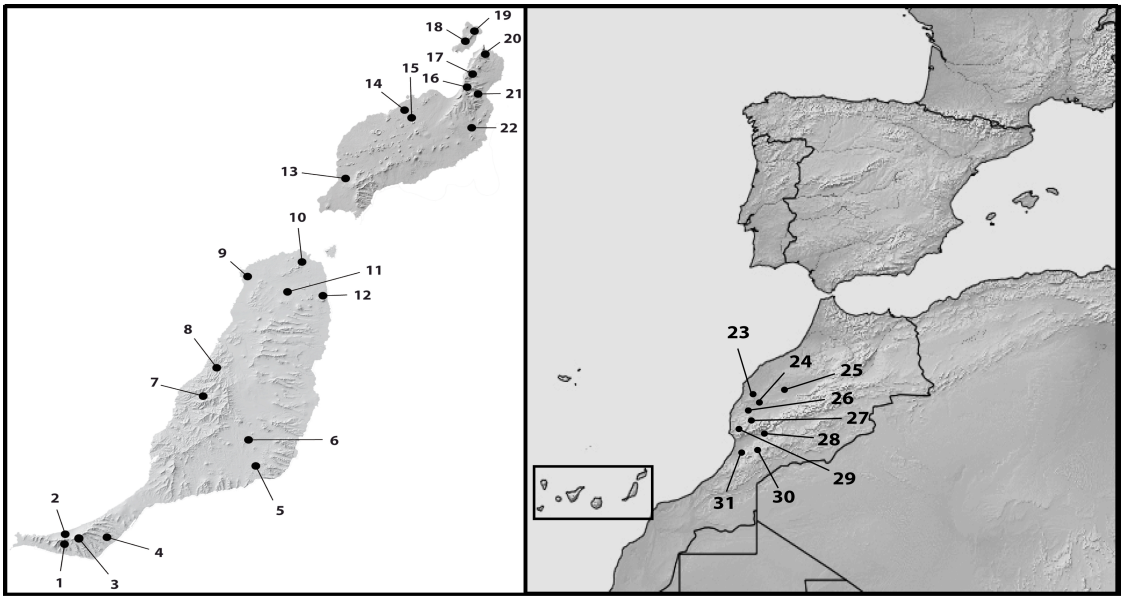


Figure 1.- *Titanidiops* sampling locations, 1: Barranco del Ciervo, 2: Cofete, 3: Pico de Fraile, 4: Barranco Mal Nombre, 5: Tequital, 6: Caldera la Laguna, 7: Betancuria, 8: Valle de Aguas Verdes, 9: Faro Tscón, 10: Corralejo, 11: Villaverde, 12: Caldería de la Roja, 13: Salinas de Janubio, 14: Tinajo, 15: Tinache, 16: Valle de Malpaso, 17: Valle de Guinate, 18: Montaña de Mojón, 19: El Vallichuelo, 20: Mirador del Río, 21: Barranco Hondo del Valle, 22: Tejía, 23: Smi-Mou, 24: Kemis-oulat-el-Hadj, 25: Ouzoud Falls rd., 26: Jbele Amsittene, 27: Tamanat - Aid-Beoude rd., 28: Iguer rd., 29: Imoza rd. nr. Tourarin, 30: Ait-Aisa, 31: Aid-Baha. For detailed information, see Table S1. The map was created using SimpleMappr http://www.simplemappr.net/

The infraorder Mygalomorphae is one of the three main lineages recognised within spiders [20], representing approximately 6.3 % of extant spider diversity [21]. Mygalomorphs are generally mid-size, robust spiders that lack the ability to spin complicated web structures but present other characters,

such as four book lungs and chelicerae with unsynchronised movement of the longitudinal fangs [22].

Mygalomorphs show high levels of local endemism, which is generally attributed to their low dispersal ability [23,24]. They show phenological and behavioural sex dimorphism; females tend to be long-lived and sedentary, while short-lived adult males actively search for mates after moulting and, thus, mediate gene flow among the populations. The taxonomy of the group is challenging, as most of the closely related taxa are morphologically homogenous and the majority of diagnostic characters are found in the reproductive organs of adult males, which are usually present in the field for a short period of time. Because of their uniform morphology and poor dispersal ability, which cause deep genetic structuring even among geographically close populations, mygalomorphs have become a model system to test species boundaries [24-27].

Mygalomorphs are notoriously absent from oceanic islands, and the few exceptions involve species either in the Caribbean [21], belonging to groups for which airborne dispersal has been reported [28-30], or in the Australian region, where the existence of land bridges during sea level changes cannot be ruled out [31]. In this regard, the presence of the trap-door spider *Titanidiops canariensis* Wunderlich, 1992 (Fig. 2) on the Canary Islands has great biogeographic relevance, as the archipelago was never connected to the continent, and the family Idiopidae, to which it belongs, has had no reported cases of aerial dispersal.

Idiopids are widespread across Australia, New Zealand, South-East Asia, Sub-Saharan Africa, Madagascar and South America but also include a few disjoint species in North Africa and the Middle East [21]. This wide distribution is often attributed to the Gondwanan origin of the family [20,32]. The family Idiopidae comprises 22 genera and 318 species, half of which are in Australia, and it ranks among the richest mygalomorph groups [21]. The family is well defined on a morphological basis [22], and its monophyly has been subsequently supported by molecular studies [20,32,33]. The genus *Titanidiops* currently comprises only three species: *T. canariensis*, endemic to the Canary Islands; *T. maroccanus* Simon, 1909 from Morocco; and the type species *T. compactus* Gerstäcker, 1873, which occurs in East Africa. *T. canariensis*

inhabits the Eastern Canary Islands, where it can be found in most habitats with the exception of sand dunes and barren lava badlands (i.e., lava flows of recent origin, devoid of soil). The spiders are mid-sized and ground dwelling, and they construct underground, silk-lined burrows that open to the surface with a trapdoor [34,35]. There is almost no information known regarding its ecology, life cycle or phylogenetic affinities.

Figure 2.- Picture of *Titanidiops canariensis* from loc. 7: Betancuria, photo credit VO

This study aims to uncover the phylogenetic origins of *T. canariensis*, one of the few examples of trap-door spiders endemic to an oceanic archipelago, and to infer the temporal framework for the colonisation of the islands using a multi-locus approach. Because of the low vagility of trap-door spiders and the dynamic volcanic history of the Eastern Canaries, we hypothesise deep and complex phylogeographic patterns, which may have led to the formation of cryptic species.

Materials and Methods

Taxonomic sampling

Most of the *Titanidiops* samples used in the present study were collected by the first author between 2009 and 2010. Additional specimens of *Titanidiops canariensis* and one specimen of the genus *Segregara* were kindly donated by colleagues. Three other representatives of the family Idiopidae were included in the analyses: *Idiops syriacus* from Israel, *Idiops* sp. from South Africa [20] and

Segregara sp. from South Africa. The last species was used as an outgroup to root the trees. Detailed locality data are included in Table S1. The specimens of Idiops syriacus were collected under the collection permit 2011/38207 granted by the Israeli government to Y. Lubin and extended to the 26th European Congress of Arachnology participants.

DNA extraction, PCR amplification, cloning and sequencing

Whole genomic DNA was extracted from the samples using the SpeedTools Tissue Extraction Kit (Biotools) following the manufacturer's guidelines. Partial fragments of four mitochondrial and four nuclear genes were sequenced in present study: 5' half of the Cytochrome oxidase I (*cox1*) (the animal barcode), the 3' half of the 16s rDNA (*16S*), the *tRNA-Leu* (*L1*) and the 5' half of the NADH dehydrogenase subunit I (*nad1*), a fragment of the 28S rDNA (*28S*), Elongation factor-1 gamma (*EF1*γ), Histone H3 (*H3*) and Heat shock protein Hsp 70 (*Hsp70*), respectively.

The PCR amplifications were carried out with the following primer combinations. *Cox1* with the primer pair C1-J-1490/C1-N-2198 [36], the fragment comprising the 16S, L1 and *nad1* with the primer pair LR-N-13398 [37] / and N1-J-12261 [38] or, alternatively, only the 3' half of the 16S with LR-N-13398 combined with LR-J-12864 [4]. All mitochondrial fragments were successfully amplified at annealing temperature range of 43-46°C. For the nuclear genes, *28S* was amplified with 28S-O/28S-B or 28S-O/28S-C [39,40]at 62-64°C annealing temperature. The *EF1*γ fragment was amplified with primers ER1gF78/EF1γR1258 [32] using a pre-amplification "touchdown" PCR step (see [32] for details), followed by a "nested" PCR using the primer pair designed in the present study EF-gIDIf 5´- GGCAACAACCAGCTCGTGGA -3´/ EF-gIDIr 5´-GTGCTGTTATTATCTTCGCC -3´, at 52°C annealing temperature. Histone *H3* was amplified with the primer combination H3a F/H3a R [41]. Heat shock protein *Hsp70* was amplified with primers MT70-MF3/MT70-R4 [42] and complemented with a newly designed primer pair for *Titanidiops canariensis* Hsp70-TitMF3 5´- AGCGACATGATGCCGAGAGT -3´/ Hsp70-TitR4 5´-GGAGGATGCAGTGGACATGG -3'. Both primer combinations yielded amplifications at an annealing temperature between 52-54°C. Some individuals

showed *Mt70* amplicons of different length and were further purified and cloned using the pGem®-T kit (Promega), following the manufacture's guidelines.

All the reactions were carried out in total reaction volume of 25μl mixing 1.25 U *Taq* polymerase (Promega), 2.5 mM $MgCl_2$ (Promega), 0.2 mM of each dNTP, 0.2 μM of each primer and 1.5μl of DNA and the 5μl of *Taq* buffer. PCR products were purified using ExoSAP-IT (USB Corporation) and sequenced in both directions using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on ABI 3700 automated sequencer at *Centres Científics i Tecnològics* of the University of Barcelona (CCiTUB, www.ccit.ub.edu) Spain.

The chromatograms were assembled and edited in Geneious v. 5.3.6. [43].

Sequence alignment, allele phasing and recombination testing

All gene fragments with exception of *nad1* and *H3* presented length polymorphism due to indel mutations. Sequence alignments of *28S*, *16S, tRNA-Leu* and *Hsp70* were obtained with the online version of MAFFT v. 6 (available at http://mafft.cbrc.jp/alignment/server/], [44]) using the Q-INS-i approach with default settings (gap opening penalty GOP = 1.53 and offset value set to 0.0). Online version of the alignment program TranslatorX (available at http://www.translatorx.co.uk/, [45]), which uses amino acid back translation to guide the nucleotide alignment, was used to build the alignments of protein coding genes *cox1* and *EF1*γ. The alignments of *nad1* and *H3* were trivial since no length variation was observed in this gene fragments.

The allelic phases of the nuclear gene fragments were determined using the PHASE algorithm [46,47] as implemented in DnaSP 5.10.1 [48].

Recombination was tested by means of the difference of sum of squares method (DSS) as implemented in TOPALi v 2.5 [49,50]. The size of sliding window was set to 500bp in all fragments except for *H3*, where a 300bp frame was used instead.

Delimitation of putative independent evolutionary lineages

The General Mixed Yule Coalescence model (GMYC) [51] was used to delimitate coalescent groups (i.e. putative independent evolutionary lineages)

within the complete *Titanidiops cox1* sequence data set. The computer program BEAST was used to infer an ultrameric tree for the whole *cox1* data set running two chains and defining a lognormal relaxed clock, with a single partition, the ucld.mean parameter set to 1 and selecting a constant population size coalescent tree prior [52]. The GMYC analysis was carried out in the R (http://www.r-project.org) environment using the SPLITS package [53].

Phylogenetic analysis

Phylogenetic inference was conducted under two different approaches: (1) by assuming a common underlaying tree for the different genes (i.e. concatenation approach) and (2) by assuming independent gene trees and species tree (i.e. species tree/gene tree approach). The *cox1*, *16S*, *tRNA-Leu*, *nad1*, *28S*, *EF1γ, H3* gene fragments of a single individual from each *Titanidiops canariensis* GMYC cluster identified above were concatenated in a single matrix using Geneious v. 5.3.6. [43]. The gaps were coded as presence/absence characters following the simple coding approach of [54] as implemented in FastGap 1.2 (available at http://www.aubot.dk/FastGap_home.htm, [55]. The *Mt70* gene fragment was not included in the concatenated matrix because of the deep differences observed within individual alleles the absence of data for Moroccan samples.

The best partitioning scheme and evolutionary model were selected using the greedy algorithm in the program PARTITIONFINDER [56]. The *L1* and *16S* genes were combined in a single partition.

The Bayesian inference analyses were conducted in MrBayes v. 3.1.2 [57] and run remotely at the CIPRES portal [58]. A ten-partition scheme with the corresponding evolutionary models as selected by PARTITIONFINDER and an additional restriction model for binary scored gaps were defined (see Table 1). Two independent runs of $5x10^7$ generations with 8 MCMC (Markov Chain Monte Carlo) chains each, starting from random trees and resampling each 1000 generations were run simultaneously. The first 20% of the generations were discarded as a *burn-in* for the analyses. Convergence of the runs was assessed by monitoring the standard deviation of split frequencies (<0.01) with the help of the program TRACER v.1.5 [59], which was further used to assess correct mixing within each chain.

Maximum Likelihood (ML) analyses were conducted in RaxML v.7.2.8. [60]. Independent GRT+G+I substitution models were assigned to each partition of the ten-partition scheme (see above), and a binary model was applied to the binary scored gaps. The best maximum likelihood tree was selected from 100 iterations and support assessed with 1000 replicates of bootstrap resampling.

All trees were visualized and manipulated with the program FigTree v. 1.3.1 [61].

Estimation of divergence times

Divergence time was estimated using two complementary approaches: concatenation and a coalescent-based approach. Time estimates were conducted in a Bayesian framework with the help of the program BEAST 1.7.4. [62]. The concatenation analysis was aimed to maximize the taxon sampling and only included mitochondrial data (*Cox*, *16S, tRNA-Leu* and *nad1*). Additionally, distant outgroups were removed from the analyses to reduce branch length disparity (i.e. only *Titanidiops* and *Idiops syriacus* samples were included). In an attempt to facilitate convergence, a simplified partition scheme by gene was implemented with corresponding substitution model provided from PARTIONFINDER. All genes were set to share the same tree and a Yule speciation model was set as tree prior. Because of the lack of fossil data and relevant biogeographic events, we relied on a spider specific mitochondrial substitution rate available in the literature [63]. A normal prior was assigned to the ucld.mean parameter of the lognormal relaxed clock, with initial and mean value 0.012722067 and standard deviation 0.00454. Three independent runs of $5x10^7$ generations were conducted remotely at the BIOPORTAL computer resource of the University of Oslo (http://www.bioportal.uio.no/). Convergence between runs and correct mixing within each run were visualized with TRACER [59]. Individual runs were combined in BEAST accompanying program LOGCOMBINER. The first 10% of the generations of each run was discarded as a burn-in. A consensus chronogram was inferred with TREEANNOTATOR.

To ascertain the timeframe of the evolution of *Titanidiops canariensis*, a multi-gene coalescent approach as implemented in the program *BEAST [64] was performed. All GMYC lineages of *T. canariensis* with at least two

sequences per gene were included. The GMYC 1 clade from *Titanidiops* sp. from Morocco was used as outgroup. Similarly to the concatenated analysis, the partition scheme was simplified to speed up computation and facilitate convergence. Mitochondrial genes were combined in a single partition and a single evolutionary model, lognormal relaxed clock and tree were defined for this partition. Conversely, independent models, lognormal relaxed clocks and trees were allowed for each nuclear gene. Along with mitochondrial substitution rate implemented in the concatenated analysis, a substitution rate for the *EF1*γ was set to 0.00117/per site/million years (normal distribution prior ucld.mean=0.00117, sdev=0.00014) as recently estimated for mygalomorph spiders [65]. Uniform priors to the ucld.mean were assigned for the *28S*, *H3* and *Hsp70,* with lower and upper bounds 0.0001 and 0.0115, respectively, and starting value 0.001, under the assumption that the nuclear genes are about one order of magnitude slower than mitochondrial and generally no nuclear protein coding gene will show higher rates than the mitochondrial genes. Finally, we used the oldest subaerial age of Fuerteventura (22 My) [66] as a maximum bound for the tree root.

Four independent runs of 100 millions of generations were run remotely on the CIPRES portal [58]. Results were monitored and analysed as already specified in the concatenated analysis.

BPP

The status of mitochondrial GMYC clusters as candidate species was further tested in a Bayesian multi-species coalescent framework using the species delimitation method in the program BPP 2.2 [36,37]. This method calculates the posterior probabilities of alternative species delimitation models using a multispecies coalescent approach and interprets gene tree incongruence as the result of ancestral polymorphism.

Following Leaché & Fujita [38], we used the rjMCMC algorithm 0 with the fine-tuning parameter ε set to 15.0. Each species delimitation model was assigned an equal prior probability. Because of the volcanic nature of the Canary Islands, and in the absence of information about the ancestral population size of the target organism, we assumed an evolutionary scenario where the colonising ancestral population was of a small size due to the

frequent bottlenecks associated with volcanism and island colonisation (θs ~*G*(2, 2000)). The effects of alternative species divergence time scenarios on species delimitation were tested by implementing either shallow (τ0 ~*G*(2, 2000)) or deep divergence priors (τ0 ~*G*(1, 10)). The species tree obtained in the *BEAST analysis was used as the fixed topology. The species delimitation analyses were restricted to those clades determined with high support (i.e., p>0.95) in the *BEAST analyses (see results). The convergence and sensitivity of the results to the initial condition were assessed by running three independent chains for each parameter combination, each time starting with a different species delimitation model: one that lumped all candidate species into a single species, one that considered an intermediate number of species and one that considered the full range of species.

Population structure

Standard genetic diversity indices, including the nucleotide (π) and haplotype diversity (H) indices, were calculated in DnaSP 5.10.1 [39] for *cox1*, *EF1γ*, *H3* and *Hsp70* for all GMYC clades consisting of at least two individuals. Additionally, values were also calculated for the *28S 16S + tRNA-Leu* and *nad1* fragments in the dataset, the Moroccan *Titanidiops* and *T. canariensis* samples and the two *T. canariensis* main clades determined in the *BEAST analyses.

Allele networks were constructed using the minimum spanning tree method in the program HapStar [40] based on the output provided by Arlequin [41].

We looked for patterns of isolation-by-distance in the Canarian samples using the online version of the program IBDWS 3.23 [42]. The *cox1* F_{ST} pairwise estimates calculated in IBDSW from raw sequences and the geographic distances between the locations obtained in Geographic Matrix Distance Generator v. 1.2.3 (Ersts, American Museum of Natural History, Center for Biodiversity and Conservation, http://biodiversityinformatics.amnh.org/open_source/gdmg) were correlated in three sets of analyses: one including all *T. canariensis* samples and the other two using the two main clades identified in *Beast.

Results

Sampling, sequencing and recombination testing

Information about specimens, localities and sequence GenBank accession numbers are listed in Table S1, and the localities are shown in the map in Fig. 1. A total of 100 specimens were sequenced in the present study. The following gene fragments were obtained for *Titanidiops*: the mitochondrial *cox1* (673bp, 231 variables), *16S-L1* (600bp, 177 variables) and *nad1* (382bp, 149 variables) and the nuclear *28S (7*62, 22 variables), *EF1*γ (844bp, 52 variables), *H3* (347bp, 46 variables) and *Hsp70* (606bp, 86 variables). No recombination was detected within any of the gene fragments used in this study.

GMYC-based lineage delimitation

The complete *cox1* data matrix, including 98 *Titanidiops* specimens from the Canaries and Morocco, was analysed using the single-threshold option of the GMYC algorithm, which was shown not to be significantly worse than the multiple-threshold option (p=0.23). The GMYC algorithm identified 32 entities/clusters (CI: 27-34) ($p=4.6*10^{-9}$), of which 9 were Moroccan and 23 were from the Canary Islands (Table S1, Fig. 3). In most cases, the GMYC clusters corresponded to single localities, and each locality included a single GMYC cluster (62%). Exceptions to this pattern included 4 instances of GMYC clusters found in more than one locality, usually including one or more nearby localities, and 5 instances of localities with more than one GMYC, usually involving closely related clusters. Interestingly, two different GMYC clusters belonging to distant clades (G20 and G10) were sampled from locality 6, in central Fuerteventura.

Phylogenetic analyses

All genes except *Hsp70*, which could only be reliably amplified and sequenced in the Canarian specimens, were concatenated with the outgroup sequences in a single matrix for subsequent Bayesian and ML tree inference. The concatenated data matrix consisted of 3,583 bp (some alignment positions with a high proportion of missing data were removed) and 32 binary coded gaps scored for 51 terminals. The *Titanidiops canariensis* lineages were represented

by single representatives of each GMYC cluster to increase the speed of the analyses. The partition scheme and corresponding evolutionary models are summarised in Table 1.

The Bayesian and maximum likelihood (-lnL 15343.170107) analyses of the concatenated data matrix resulted in similar tree topologies, although the Bayesian inference yielded higher clade supports (Fig. 4). Both analyses supported the reciprocal monophyly of *T. canariensis* and a clade comprising *Idiops syriacus* and the Moroccan *Titanidiops* lineages. The internal topology of *T. canariensis* was well supported and highly congruent between the methods. Conversely, most relationships within the deeply divergent Moroccan *Titanidiops* lineages were poorly supported.

Both analyses indicated a well-supported clade (hereafter referred as the JSF clade) that includes the specimens from the Jandia Peninsula in southern Fuerteventura except for the individual Z37, which constituted an independent lineage, along with some specimens from central Fuerteventura. These Fuerteventura lineages were shown to be sister lineages to the remaining Canarian representatives, albeit with low support. Similarly, most Lanzarote specimens formed one well-supported clade (hereafter referred as the L clade), including two reciprocal geographic clades (northeast and southwest), which were sisters, albeit with low support, to a clade (hereafter referred as the FL clade) that includes representatives from northern and central Fuerteventura and two additional Lanzarote lineages, one from the southwest and one from the northeast. Although the internal relationships of the last clade were poorly supported, the central Fuerteventura lineage G13 (Z232) was determined to be a sister group to all remaining lineages within the clade.

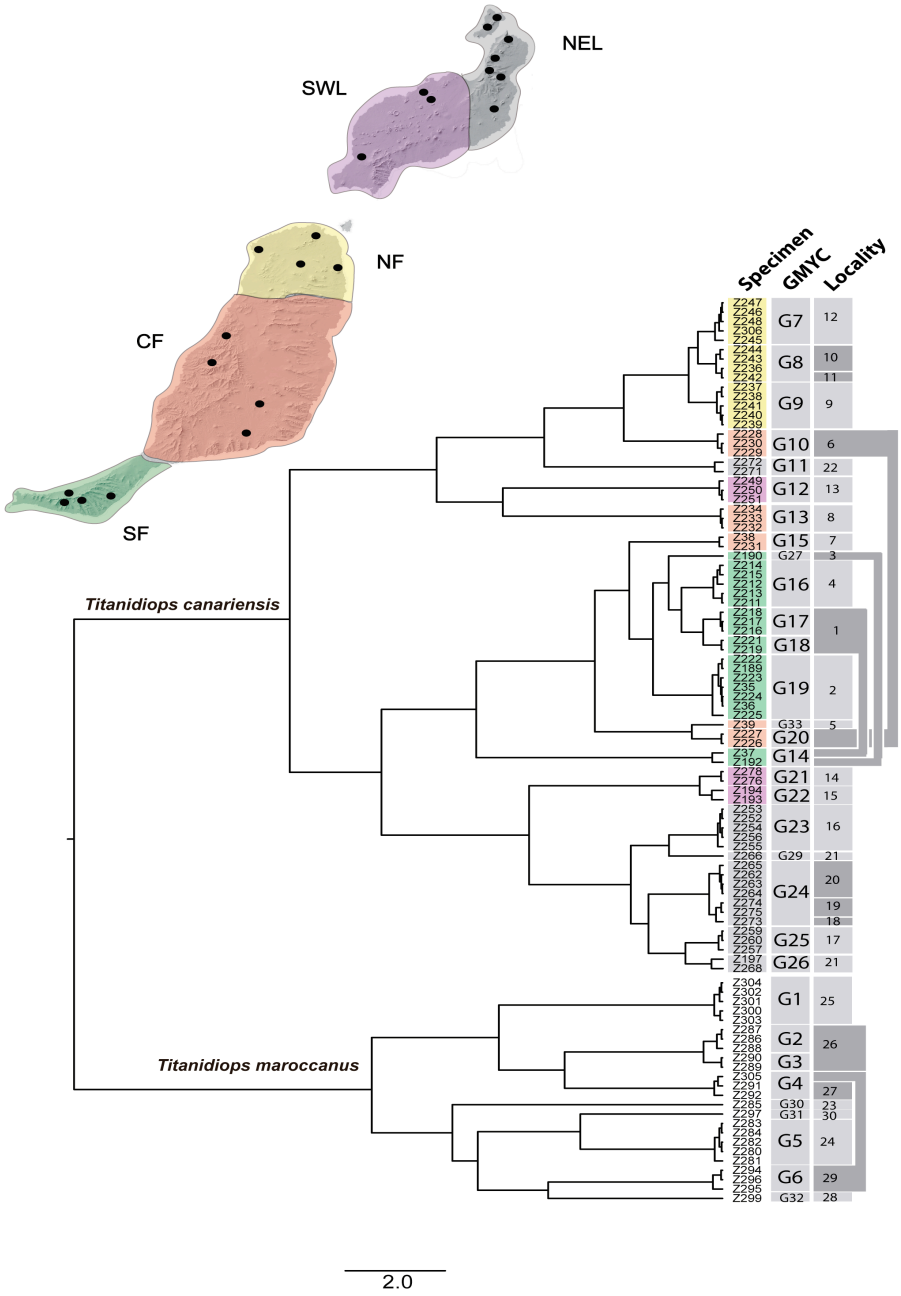


Figure 3.- Ultrametric *cox1* BEAST tree. GMYC column: clades identified as independent GMYC clusters in the SPLITS analyses, labelled as in Table S1. Locality column: localities where the respective GMYC cluster was collected. Localities in dark grey correspond to those either including more than one GMYC cluster or those where the GMYC cluster was found in an additional locality (connected by bars). Terminal colour codes represent geographic location as shown in the map. SF: southern Fuerteventura (Jandia Peninsula), CF: central Fuerteventura, NF: northern Fuerteventura, SWL: south-western Lanzarote, NEL: north-eastern Lanzarote (includes La Graciosa islet).

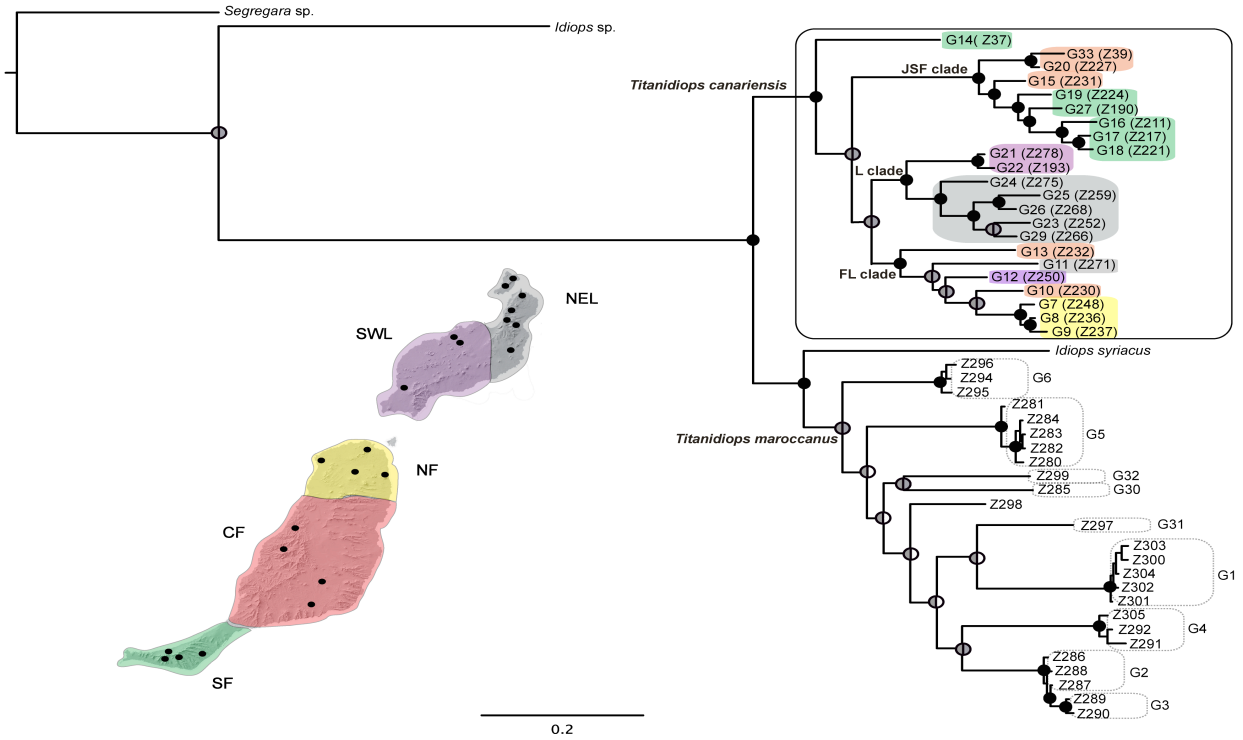


Figure 4.- Topology obtained in the concatenated Bayesian analyses. *Titanidiops canariensis* is represented by single representatives from each GMYC cluster identified (Fig. 3, Table S1). *T. maroccanus* GMYC clusters are highlighted in boxes. Dots on nodes denote support as follows: left semi-circles are Bayesian posterior probabilities (PP) and right ones are maximum likelihood bootstraps, black = PP > 0.95, ML bootstrap support >80%, grey = clade determined but with support values less than the thresholds above, white = topology not determined. Terminal colour codes as in Fig. 3.

Estimation of divergence times

Overall, the tree topology and the clade supports were similar to those found in the Bayesian and ML analyses. The root was assigned to the split between *T. canariensis* and the clade formed by the Moroccan *Titanidiops* lineages and *I. syriacus* (Fig. 5), and it was estimated to have occurred approximately 12 million years ago (Ma) (24.65–6.41 Ma). The TMRCA of *T. canariensis* was estimated at 8.08 Ma (16.01– 4.16 Ma). The diversification of the JSF clade began 2.86 Ma (5.74-1.39), while the L clade and the FL began diversifying earlier, 6.98 Ma (13.82-3.57).

The split between *Titanidiops maroccanus* and *I. syriacus* was traced back to approximately 11 Ma (21.79–5.41), although this relationship was not well supported. Diversification of the deeper *T. maroccanus* lineages occurred from the late Miocene to the early Pliocene. As in the previous analyses, most relationships within the Moroccan *Titanidiops* remained unresolved.

Coalescent approach

The coalescent approach resulted in a similar topology to the concatenated analyses but with lower support. The resulting *T. canariensis* species tree (Fig. 5) was divided into two well-supported clades, one corresponding to the JSF clade determined in the previous analyses, and the other one (hereafter referred to as the A clade) included the remaining individuals, which were also determined in the previous analysis albeit with lower support. The internal relationships of the JSF clade corresponded approximately to those found in the previous analyses. Conversely, clade A showed major lineage rearrangements compared to the previous analyses, mostly involving lineages G11, G12 and G13. Lineage G11, which was shown to be part of the clade FL in the previous analyses, was determined to be a sister lineage to the remaining lineages in clade A. The lineage G12, also shown to be part of the FL clade in the previous analyses, was nested within the former L clade. Finally, lineage G13, which was formerly supported as the first split within the FL clade, was nested with high support within a clade formed by the lineages from northern Fuerteventura. Overall, these differences resulted in more geographically congruent clades.

Lineage divergence times were much younger than those determined in the mtDNA concatenated analysis. The origin of the *T. canariensis* stem was placed at 4.25 Ma (13.6-0.8 Ma), while its TMRCA dated back to 2.29 Ma (4.28-0.05 Ma). The diversification of clade A began approximately 1.36 Ma (1.71-0.27 Ma). The origins and diversification of the main geographic lineages, namely the JSF clade, the central-northern Fuerteventura clade and the Lanzarote clade, occurred approximately 1 Ma (2.34-0.44 Ma).

Concatenated mtDNA + nucDNA		**Concatenated mtDNA time**		**Coalescent mtDNA + nucDNA**	
Partition	**Model**	**Partition**	**Model**	**Partition**	**Model**
cox1 1st	HKY+I+G	*cox1*	GTR+I+G	mtDNA	GTR+G
cox1, *nad1* 2nd	K81+I+G	*16S-L1*	GTR+I+G	*28S*	HKY+I
cox1 3rd	K81+G	*nad1*	GTR+I+G	*EF1γ*	HKY+I+G
nad1 3rd	HKY+G			*H3*	K80+I
nad 1st, *16SL1*, *H3* 3rd	GTR+I+G			*Hsp70*	K80+G
EF1γ 1st	JC+I				
EF1γ 2nd	F81+I				
EF1γ 3rd	K80+G				
28S	HKY+I				
H3 2nd, 3rd	K80+I				

Table 1.- Evolutionary models and partition schemes selected by PARTITIONFINDER for the different analyses conducted. Concatenated mtDNA time refers to concatenated analyses conducted with BEAST including informed substitution rate priors.

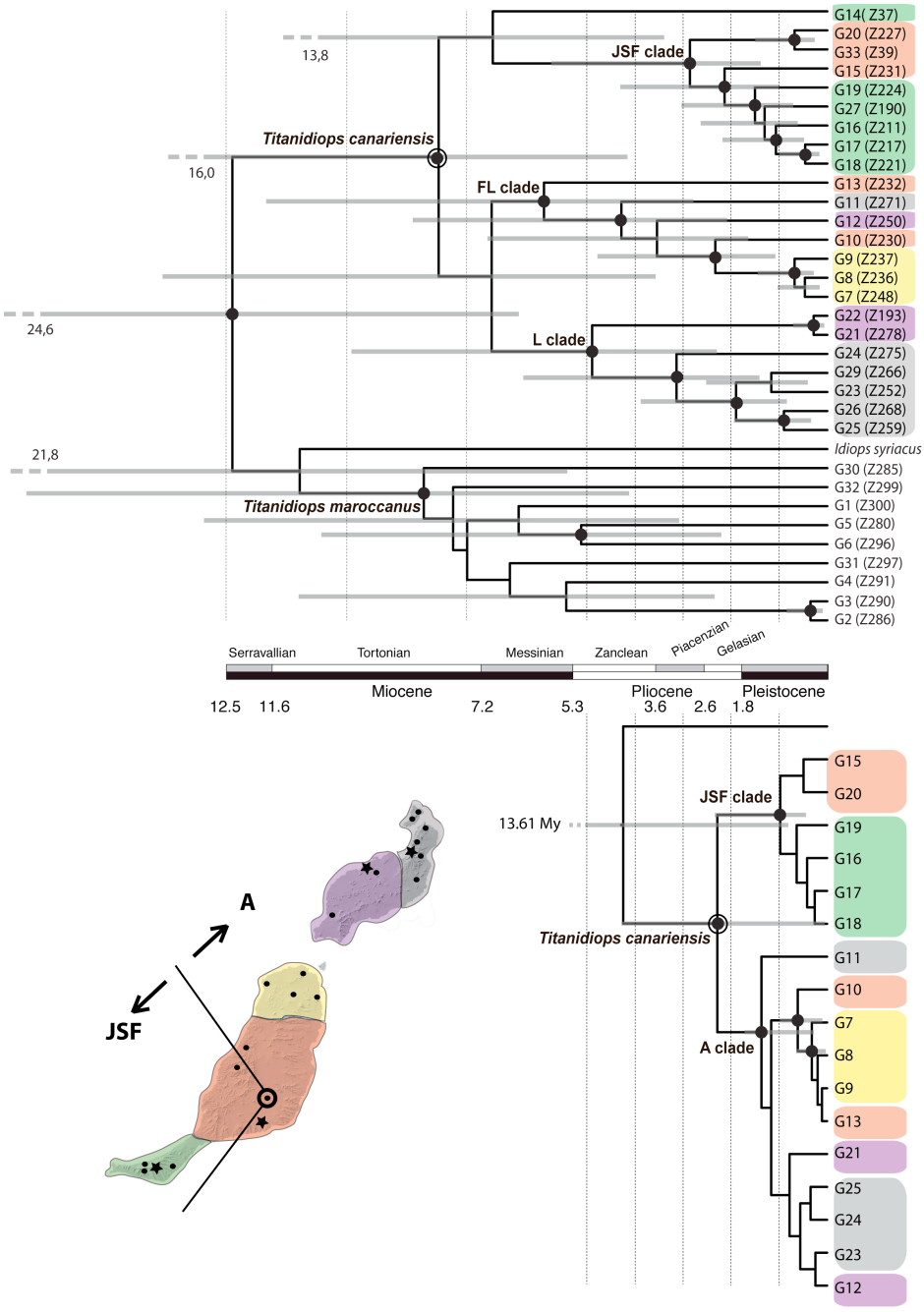


Figure 5.- Chronograms obtained with (a) the concatenated approach using BEAST and with (b) the multispecies coalescent (species tree) approach using *BEAST. Dots on nodes denote Bayesian posterior probabilities above 0.95. Node bars indicate the 95% HPD confidence intervals of the divergence time. The common x-axis is time in million years (My). Terminal colour codes are as in the figure inset. Samples from the localities marked as stars on the map were only included in the concatenated approach.

BPP

The following lineages, delineated with high support in the *BEAST analysis, were included in the species delimitation analyses: G15+G20 (candidate species 1, sp. 1), G16-G19 (sp. 2), G11 (sp. 3), G10, (sp. 4) and G7-G9+G13 (=sp. 5). In addition, although not supported, the clade G21+G23-G25 (=sp 6) was also tested as it was the sister group of a well-supported clade and similar in age to the remaining clades analysed.

Under the first scenario tested, i.e., shallow divergences, all runs starting from K=1 and K=2 supported (i.e., a posterior probability of the split >0.95) a two species hypothesis, in agreement with the basal split between the JSF and A clades (sp1+sp2 and sp3+sp4+sp5+sp6). Runs starting from K=3 resulted in three delimited species because clade A was further divided into sp3 and sp4+sp5+sp6. Finally, runs under K=4, K=5 and K=6 supported a five species hypothesis, sp1+sp2, sp3, sp4, sp5 and sp6 (Fig. 6). In all cases, sp1 and sp2 were lumped into a single species. The results of the second scenario, deep divergences, were directly dependent on the selected K values. These results were identical to the defined starting models.

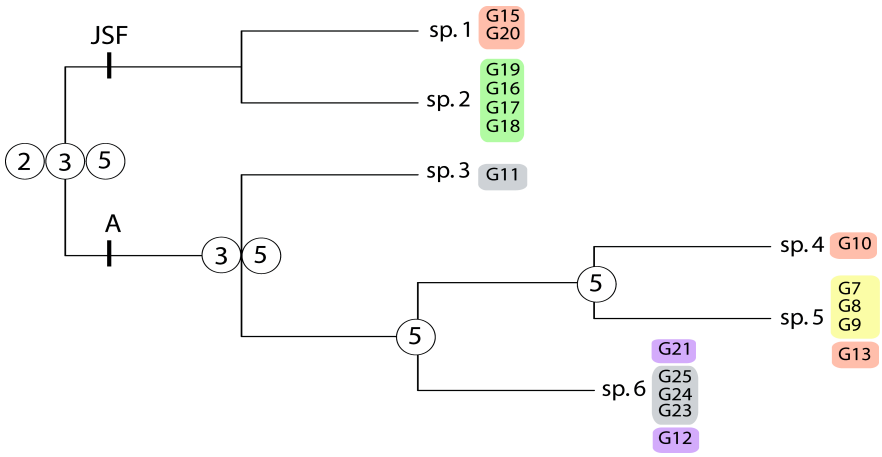


Figure 6.- Multi-locus Bayesian species delimitation with BPP. Only results obtained under the small population sizes (θs ~*G*(2, 2000)) and shallow divergence (τ0 ~*G*(2, 2000)) scenarios are shown. Circles on nodes indicate lineages (descendants) supported as independent evolutionary lineages (i.e., species) and the total number of independent lineages supported under alternative models of number of starting species (K). Terminal colour codes as in Fig. 3.

Population analyses

Standard genetic diversity indices, nucleotide diversity (π) and haplotype diversity (H), were calculated for the *cox1*, *EF1γ*, *H3* and *Hsp70* genes for each GMYC lineage comprising two or more individuals from the entire *Titanidiops* dataset, for *T. canariensis*, *T. moroccanus* and for the JSF and A clades. The results are summarised in Table S2.

The *T. canariensis EF1γ* allele network (Fig. S1) included 27 unique alleles. The two most common alleles, which differed by a single substitution and were shared by 18 and 12 individuals, were collected from all geographical zones. Overall, the network was congruent with the main mtDNA lineages, assuming the most frequent alleles were shared across lineages due to ancestral polymorphism.

The resulting *Hsp70* network (Fig. S2) contained 45 distinct alleles. The highest allele diversity was found among the northern Lanzarote samples (20 alleles), and the northern Fuerteventura samples showed the lowest (3 alleles). The remaining three geographic zones presented a moderate number of alleles (central and southern Fuerteventura had 10 alleles, respectively, and south-western Lanzarote had 7). Unlike the *EF1γ* network, the *Hsp70* one had little geographic structure, and most mtDNA lineages were not determined. The JSF clade was an exception, and the alleles belonging to individuals carrying this mtDNA lineage were all closely related. A group of alleles found mostly in northern Lanzarote (one from southern Lanzarote) was separated from all remaining alleles by at least 14 missing mutations.

Isolation by distance

There was no significant correlation between genetic and geographic distances, either for all *T. canariensis* samples together or for any of the two main lineages (JSF and A) (r (P) = -0.0036 (0. 5746), 0.2879 (0.3391), -0.0710 (0.7247), respectively).

Discussion

Mygalomorph spiders on the oceanic islands

Mygalomorph spiders are notoriously absent from oceanic archipelagos, most likely due to their low dispersal abilities [23,24,26,43]. The very few reported exceptions include members of the ctenizid genus *Ummidia* in the Caribbean, one of the few mygalomorphs that use airborne dispersion by ballooning, and several species of the family Barychelidae found on oceanic archipelagos in the Pacific [44-46]. Especially relevant in the present context is the presence of the genera *Idioctis* L. Koch, 1874 and *Nihoa* Raven & Churchill, 1992 on remote Pacific archipelagos, including the Hawaiian Islands, one of the most isolated archipelagos in the world. Both genera have been considered a challenge to the vicariant wisdom [45]. However, although the intertidal and littoral habitats of *Idioctis* point towards dispersal by floating on foam or rafting after being washed out by waves, thus far, no explicit biogeographic hypotheses have been put forward to explain their origin. Here, we provide, for the first time, dated phylogenetic information to decipher the origins of an oceanic mygalomorph.

In spite of the relatively short distance between the Canary Islands and northern Africa, approximately 110 km at the narrowest point, the well-supported monophyly of the Canarian lineages suggests a single event of colonisation of the Canary Island by *Titanidiops*.

The discovery of 12 My-old fossil ostrich-like eggs on Lanzarote cast some doubts on the oceanic origin of the Eastern Canaries, suggesting that the islands were continental in origin and that a land connection might have still existed in the lower Pliocene [47]. However, geophysical evidence accumulated in subsequent years indicate that these islands were built on oceanic crust [48]. Our results may support their oceanic origin, as the presence of land bridges would have facilitated recurrent colonisation of the islands by more than one *Titanidiops* lineage.

Although the mechanism by which *Titanidiops* colonised the Canaries remains speculative, it has been suggested that the ground-dweller spider *Dysdera* colonised the Canary Islands from the neighbouring north-eastern African coast by transporting itself on floating islands [2]. It is interesting to note that some of the *Titanidiops* localities sampled in Morocco are distributed

around the valley of the Sous River, one of the most important rivers in Morocco. It is possible that in the more humid climate of the past, the river had high flows and torrential currents, facilitating the formation of large mats of floating vegetation that may have acted as rafts for ground-dwelling arthropods, including *Titanidiops*. In addition, recurrent sea level regressions recorded during the Triassic [49] would have further facilitated marine dispersal by exposing submarine banks, which reduce the distance from the continent, as suggested by the arrival of the extinct endemic rodent *Malpaisomys* [50].

A large gap was observed between the timeframes of *Titanidiops* diversification determined from only the concatenated mitochondrial data and from the multi-locus species tree approaches. In general, longer divergence time estimates are expected from gene trees because a fraction of gene divergence may pre-date population/species divergences [51,52]. There is ample empirical evidence that the use of species-tree approaches usually results in younger estimates of species divergence times [53-55], sometimes as much as three times younger than those inferred using concatenated gene approaches [56]. In our study, the species tree estimates of the timing of most recent common ancestor of the Canarian *Titanidiops* (2.3 Ma, 4.3-0.05 Ma) were about half as young as the mtDNA gene estimates (8.8 Ma, 16-4.16 Ma). Other studies, however, have found that the divergence times obtained under both approaches did not differ substantially [57-59]. Differences in time estimates under both approaches may be exacerbated in the case of large ancestral populations [54] or in the presence of gene flow among lineages [53]. Because of the inherently small populations involved in colonisation events, in the case of *Titanidiops,* the existence of gene flow may provide a better explanation. The effect of such events may also explain some of the incongruences observed in the *Hsp70* allele network (Fig. S2).

The mismatch between the estimated times using the two approaches, either during the Miocene or, more likely, in the Pliocene, hinders the proposal of a geological scenario for the colonisation of the islands. Although the random nature of the colonisation process could explain the absence of *Titanidiops* from the western Canaries, the strong environmental differences between the arid

and low-lying Eastern Canaries and the lush and elevated Western Canaries may suggest specific ecological or habitat preferences.

If *Titanidiops* are better suited to the present day environmental conditions in the Eastern Canaries, its arrival to the island probably occurred after the onset of such conditions. It is worth noting that there are other cases of eastern Canarian endemic lineages with northern African close relatives, especially among plants (see [60]). Examples of faunal connections include the spider *Dysdera lancerotensis* [3], the *Calathus* beetles [12], the *Chalcides* skinks [61], the *Tarentola* geckos [11] and the fossil rodent *Malpaisomys* [50]. Available time estimates for these lineages suggest a mid-Pliocene to early Pleistocene colonisation of the Eastern Canaries. Interestingly, this timeframe coincides with the onset of a major dryness event in the region following the establishment of the cold Canarian sea current ~4 Ma [62]. Our species-tree estimates of the Canarian *Titanidiops* stem fits well with the former scenario, which would suggest a preadaptation to xeric conditions. However, the time estimates in the cited studies were obtained either from single mitochondrial genes or from gene concatenation.

Our concatenated *Titanidiops* divergence time estimates are much older, tracing back to the Miocene epoch, and are closer to the values reported for the Canarian lineages that have undergone species radiations, such as the *Pholcus* [63] spiders and most of the *Dysdera* [4,64] and the *Gallotia* lizards [65]. Similarly, the coalescent time of the *Titanidiops* mtDNA haplotypes (4.25 Ma, 13.6-0.8 Ma) pre-dates most values reported in the literature for Eastern Canaries endemics, including the lizard *Gallotia atlantica* (1.9 Ma, confidence interval 3.3-1.3 Ma) [7], the darkling beetle *Hegeter politus* and the spiders *Dysdera lancerotensis* (1.5 Ma, 2.3-0.7 Ma) [3], *D. nesiotes* (0.8 Ma, 1.1-0.53 Ma) and *D. alegranzaensis* (0.83-0.4 Ma) [5].

Further support for the old (Miocene) colonisation of the Canaries by *Titanidiops* may come from the observation that the sampled population from the continental source, Morocco, formed a monophyletic group, indicating that the continental lineages that gave rise to the island populations may have gone extinct and, hence, that the colonisation preceded the current diversification of *T. maroccanus*.

How many species of Titanidiops inhabit the Canary Islands?

Mygalomorph spiders are phenotypically conservative and pose a challenge to classical taxonomy. The application of molecular tools to investigate phylogeographic patterns in this group has uncovered a great amount of hidden diversity [23,43,66-70]. However, because of their low dispersal capability and burrow fidelity, mygalomorph spiders tend to form clustered aggregations [71,72], which usually lead to extensive genetic structures [24,27,66,73], making it difficult to distinguish between fragmented populations and independent evolutionary lineages [26]. As stated above, the estimated time of the common ancestor of the *T. canariensis* individuals sampled in the present study is similar to estimates obtained for lineages that have diversified into several endemic species. Consequently, it should not come as a surprise that the Canaries might harbour more than one *Titanidiops* species.

Although it is generally agreed that species delimitation should be based on the integration of multiple lines of evidence (i.e., genotypes, phenotypes, ecology), [74], there have been little advances to date in integrating such diverse data sources in a quantitative framework [75]. Novel statistical approaches for the quantitative assessment of species boundaries mostly focus on genetic data [76,77] because of the inherent advantages of these data in terms of abundance, standardisation and codification.

Among the multiple methods developed to determine species boundaries based on multi-locus data, the Bayesian multi-species coalescent model in the program BPP has become increasingly popular [26,38,78-82]. Although the method assumes no gene flow, efficient performance has been proven with the low gene flow level of 0.1 migrants per generation [83].

In this study, we have applied the BPP method to validate whether the large number of genetically divergent lineages revealed by single markers corresponds to distinct evolutionary lineages. We tested two alternative scenarios, corresponding to either shallow or deep divergences between candidate species. Because of the dynamic nature of volcanic landscapes, involving recurrent extirpations and recolonisations of populations, and the low vagility of mygalomorph spiders, small ancestral population sizes were assumed for both scenarios. However, given the inconsistent results from the

deep divergences scenario and the relatively recent estimates from the species-tree approach for the diversification of *T. canariensis*, which were mostly confined to the Pleistocene, we argue that the scenario assuming shallow divergences between the candidate species is probably the most appropriate.

Our results are sensitive to the starting species number model (K), which, to the best of our knowledge, has not been reported in any previous study. The basal split of *T. canariensis* into two species was the only consistently supported alternative across all starting species models. Higher numbers of species were only supported by particular K values. The BPP method has been found to over split species diversity, most likely because it does not integrate over the species tree parameter space [26]. Here, we take a conservative approach and propose that *T. canariensis* includes two distinct species.

Interestingly, although the two putative species are, for the most part, geographically isolated, they do overlap in at least one locality, suggesting similar habitat preferences. In addition, the overlapping of the two putative species offers the possibility to test for reproductive isolation. Unfortunately, only 5 individuals (2 belonging to JSF and 3 to clade A) were sampled from this locality. Although this reduced sample does not show clear evidence of gene flow, most shared alleles may be explained by ancestral polymorphism; a larger sample and more variable nuclear markers would be required to assess the actual level of genetic isolation.

The “JSF” species was found in all localities sampled on the Jandia peninsula. This massif is effectively isolated from the rest of Fuerteventura by a low-elevation isthmus, covered by aeolian sands, which explains the presence of several examples of vicariant species on the peninsula [84]. However, it was also found in localities across the isthmus that do not show any obvious geographic discontinuities with the populations of the second species, including the co-occurring locality. The branching pattern determined in the better-sampled concatenated analyses is compatible with at least two rounds of back-and-forth colonisation between the main island and Jandia. Therefore, the north-wise displacement of the contact zone between the two putative species could be explained by a secondary colonisation of the Fuerteventura main island from a lineage originally isolated in the peninsula.

It is not strange that all the specimens collected in the present study were either females or juveniles, because they were mostly extracted from burrows, which adult males are known to abandon in search of females. In fact, the only collected *T. canariensis* males (3 specimens) were found in a pitfall trap in Pico Zarza, Jandia, near our localities 1 and 2 (JSF species). Preliminary observations revealed subtle differences in morphology between the putative species; however, in the absence of male material from localities throughout the Eastern Canaries, we have refrained from formally describing a new species.

Phylogeographic patterns in T. canariensis

As was expected based on former phylogeographic studies on trap-door spiders [26,27,73], a high number (23) of narrowly distributed mtDNA GMYC clusters, found frequently in single localities, were determined in *T. canariensis*. These results provide further support for the limited dispersal ability of trap-door spiders.

Although low vagility usually promotes isolation by distance, which has been suggested to play an important role in speciation in mygalomorph spiders [23,43,66-70], there was no evidence of such a pattern in *T. canariensis*. The explanation for this lack of correlation may be found in the dramatic relief of the islands. Close localities are frequently separated by steep ridges or deep valleys, which most likely represent an effective barrier to dispersal.

In spite of its low dispersal ability, *T. canariensis* successfully spread through the Eastern Canary Islands. The position of the Jandia lineage G14 in the concatenated analysis, shown to be a sister group to the remaining *T. canariensis* lineages, is compatible with a south to north colonisation pathway of the islands, as reported for *Hegeter* darkling beetles [85] and *Gallotia* lizards [7]. Unfortunately, this topology was poorly supported and could not be further corroborated in the species tree analyses that did not include G14 due the small sample size. Similarly, the low support of most basal nodes prevented the determination of the exact number and direction of the inter-island colonisation events. However, our results provide support for the existence of two independent lineages in Lanzarote, suggesting either two colonisation events for the island or back-and-forth island hopping. Similar complex phylogeographic patterns within Lanzarote and between Lanzarote and

Fuerteventura have been reported in other Eastern Canaries endemic taxa [3,5,7] and have been explained as the consequence of frequent population extinctions due to lava flows and the recurrent connections of the two islands during Neogene marine transgressions.

It is noteworthy that one of the Lanzarote lineages includes a locality in the Zonzamas area (G11, locality 22), which has been identified as a refugium during episodes of volcanic activity [5,7,86].

Our results also suggest a recent colonisation of La Graciosa from Lanzarote, as samples from La Graciosa belong to the same GMYC cluster (G24) as those from the nearby locality of Mirador del Rio in Lanzarote. La Graciosa is separated from Lanzarote by a narrow seaway approximately 2 km wide and of shallow water, and most of the shelf connecting the islands was exposed during the Pleistocene sea level oscillations [18], facilitating terrestrial dispersal between the two islands.

This pattern differs from that observed in the ground spider *Dysdera lancerotensis*, for which La Graciosa populations seem to have originated from the northern islet of Alegranza [3]. The La Graciosa populations of *D. alegranzaensis,* conversely, show genetic affinities with both their Alegranza and Lanzarote populations [5]. *Titanidiops* has not been reported in the islets north of La Graciosa, suggesting that it never colonised them or that it went extinct, which could explain the more recent origin of its present day populations.

Insights into continental Titanidiops diversity

Although determining the diversity and phylogenetic relationships of the continental *Titanidiops* were beyond the scope of the present study, our results provide some insights into the evolutionary patterns of *T. maroccanus* and improve the currently poor knowledge of this species.

Several highly divergent lineages were identified within *T. maroccanus*, showing a similar diversification timeframe to that estimated for *T. canariensis*, which may suggest the existence of cryptic species in Morocco. The divergence times estimated within *T. maroccanus* are far older than those reported in any other organism inhabiting the region. The basal split into the two main lineages of the wolf spider species complex *Lycosa oculata*, which comprises five

putative species, occurred approximately 3.3 Ma, while present day species diversified between 2.19 to 0.89 Ma. Similarly, most genetic diversity within Moroccan reptile species has been dated to the period of Pleistocene climatic oscillations (starting at 2.3 Ma) [87,88]. A more thorough sampling, additional variable nuclear genes and, in particular, adult males will be needed to determine the actual *Titanidiops* species diversity in Morocco.

The concatenated analyses supported the sister group relationship of the Moroccan *Titanidiops* lineages with *Idiops syriacus* (O.P. Cambridge 1870). Interestingly, *I. syriacus,* although originally described as *Idiops* (Perty 1833)*,* was transferred to the genus *Titanidiops* by E. Simon [89], where it remained until R. Raven [22], in the revision of the suborder Mygalomorphae, synonymised both genera. When the genus *Titanidiops* was re-erected by J. Wunderlich [34,35], the placement of *I. syriacus* and other former members of *Titanidiops* was not indicated. Interestingly, in the original description of *T. maroccanus*, Simon had already indicated that *I. syriacus* was the most similar species to the Moroccan *Titanidiops* [90]. Our results confirm this observation. However, additional representatives of the genus *Idiops,* as well as and other genera within the family Idiopidae, will be required to establish the phylogenetic limits of the genus *Titanidiops*.

The sister group relationship of *T. maroccanus* and *I. syriacus* has important implications for inferring the origins of the Canarian *Titanidiops*. The origin of *T. maroccanus* postdates the split of *T. canariensis* from their common ancestor and, most likely, the colonisation of the Canary Islands. This observation suggests that present day Moroccan diversity may be the result of a secondary colonisation of the region by the *Titanidiops* lineage.

Conclusion

T. canariensis colonised the Canary Islands once, presumably from northern Africa, although no closely related Moroccan *Titanidiops* lineage has been detected. The time of colonisation remains undetermined, mostly due to the discrepancy between the time estimates obtained with the concatenated and the species tree approaches. Several lines of evidence, however, point towards a Miocene origin. A complex phylogeographic pattern was revealed,

involving either two independent colonisations of Lanzarote or a back-and-forth colonisation between Fuerteventura and Lanzarote. Volcanic activity may have contributed to the determined geographic patterns of genetic diversity, which further supports the existence of a volcanic refugium in the Zonzamas area in west-central Lanzarote. Our results are compatible with the existence of at least two species in the Canaries, one inhabiting the Jandia Peninsula and southern Fuerteventura and the second ranging from central Fuerteventura to northern Lanzarote. The two species co-occur in at least one locality.

Several highly divergent lineages were also detected within *T. maroccanus*, most likely representing cryptic species. In this study, a close phylogenetic affinity between the Moroccan *Titanidiops* lineages and *Idiops syriacus* was also uncovered, in agreement with the previous placement of *I. syriacus* in the genus *Titanidiops*, which suggests a different origin for the Canarian and the present day Moroccan lineages of *Titanidiops*.

Acknowledgements

We are thankful to E. Planas and C. Ribera for the assistance during the fieldwork, to N. Macías-Hernández and X. Giet for kindly providing us with additional *T. canariensis* samples and to M. Forman for providing us with the specimen of *Segregara* sp.

References

1. Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. Molecular Ecology 11: 951-966.

2. Arnedo MA, Oromí P, Ribera C (2001) Radiation of the spider genus *Dysdera* (Araneae, Dysderidae) in the Canary Islands: Cladistic assessment based on multiple data sets. Cladistics 17: 313-353.

3. Bidegaray-Batista L, Macías-Hernández N, Oromí P, Arnedo MA (2007) 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. Molecular Ecology 16: 3198-3214.

4. Macías-Hernández N, Oromí P, Arnedo MA (2008) Patterns of diversification on old volcanic islands as revealed by the woodlouse-hunter spider genus *Dysdera* (Araneae, Dysderidae) in the eastern Canary Islands. Biological Journal of the Linnean Society 94: 589–615.

5. Macías-Hernández N, Bidegaray-Batista L, Oromí P, Arnedo MA (2013) The odd couple: contrasting phylogeographic patterns in two sympatric sibling species of woodlouse-hunter spiders in the Canary Islands. Journal of Zoological Systematics and Evolutionary Research 51: 29-37.

6. Macías-Hernández N, Bidegaray-Batista L, Emerson BC, Oromí P, Arnedo M (2013) The Imprint of Geologic History on Within-Island Diversification of Woodlouse-Hunter Spiders (Araneae, Dysderidae) in the Canary Islands. Journal of Heredity 104: 341-356.

7. Bloor P, Kemp SJ, Brown RP (2008) Recent volcanism and mitochondrial DNA structuring in the lizard Gallotia atlantica from the island of Lanzarote. Molecular Ecology 17: 854–866.

8. Vandergast AG, Gillespie RG, Roderick GK (2004) Influence of volcanic activity on the population genetic structure of Hawaiian Tetragnatha spiders: fragmentation, rapid population growth and the potential for accelerated evolution. Mol Ecol 13: 1729-1743.

9. Vandergast AG, Bohonak AJ, Weissman DB, Fisher RN (2007) Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: Stenopelmatus). Mol Ecol 16: 977-992.

10. Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: Towards a phylogeographic synthesis for the Canary Islands. Trends in Ecology & Evolution 15: 104-109.

11. Carranza S, Arnold EN, Mateo JA, Geniez P (2002) Relationships and evolution of the North African geckos, *Geckonia* and *Tarentola* (Reptilia: Gekkonidae), based on mitochondrial and nuclear DNA sequences. Mol Phylogenet Evol 23: 244-256.

12. Emerson BC, Oromí P, Hewitt GM (2000) Interpreting colonization of the *Calathus* (Coleoptera: Carabidae) on the Canary Islands and Madeira through the application of the parametric bootstrap. Evolution 54: 2081-2090.

13. Nogales M, López M, Jiménez-Asensio J, Larruga JM, Hernández M, et al. (1998) Evolution and biogeography of the genus *Tarentola* (Sauria: Gekkonidae) in the Canary Islands, inferred from mitochondrial DNA sequences. Journal Of Evolutionary Biology 11: 481-494.

14. Stüben PE, Astrin JJ (2010) Molecular phylogeny in endemic weevils: revision of the genera of Macaronesian Cryptorhynchinae (Coleoptera: Curculionidae). Zoological Journal of the Linnean Society 160: 40-87.

15. Carracedo JC, Day S, Guillou H, Badiola ER, Canas JA, et al. (1998) Hotspot volcanism close to a passive continental margin: The Canary Islands. Geological Magazine 135: 591-604.

16. Ibáñez JM, De Angelis S, Díaz-Moreno A, Hernández P, Alguacil G, et al. (2012) Insights into the 2011–2012 submarine eruption off the coast of El Hierro (Canary Islands, Spain) from statistical analyses of earthquake activity. Geophysical Journal International 191: 659-670.

17. Carracedo JC, Torrado FP, González AR, Soler V, Turiel JLF, et al. (2012) The 2011 submarine volcanic eruption in El Hierro (Canary Islands). Geology Today 28: 53-58.

18. Carracedo JC, Singer B, Jicha B, Guillou H, RodríguezE., et al. (2003) La erupción y el tubo volcánico del Volcán Corona (Lanzarote, Islas Canarias). Estudios Geológicos 59: 277–302.

19. Carracedo J, Rodríguez-Badiola E (1993) Evolución geológica y magmática de la isla de Lanzarote (Islas Canarias). Revista de la Academia Canaria de Ciencias v: 25-58.

20. Hedin M, Bond JE (2006) Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): Conflict and agreement with the current system of classification. Molecular Phylogenetics and Evolution 41: 454-471.

21. Platnick NI (2014) The World Spider Catalog. v. 14.5 ed: American Museum of Natural History, New York, online at http://research.amnh.org/iz/spiders/catalog, [accessed 26 March 2014].

22. Raven RJ (1985) The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. Bull Amer Mus Nat Hist 182: 1-180.

23. Bond JE, Hedin MC, Ramirez MG, Opell BD (2001) Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Molecular Ecology 10: 899-910.

24. Bond JE, Stockman AK (2008) An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. Systematic Biology 57: 628 - 646.

25. Stockman A, Bond JE (2007) Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. Molecular Ecology 16: 3374-3392.

26. Satler JD, Carstens BC, Hedin M (2013) Multilocus Species Delimitation in a Complex of Morphologically Conserved Trapdoor Spiders (Mygalomorphae, Antrodiaetidae, Aliatypus). Systematic Biology 62: 805-823.

27. Hendrixson BE, DeRussy BM, Hamilton CA, Bond JE (2013) An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, Aphonopelma). Molecular Phylogenetics and Evolution 66: 327-340.

28. Coyle EA, Greenstone MH, Hultsch AL, Morgan CE (1985) Ballooning mygalomorphs: Estimates of the masses of Sphodros and Ummidia ballooners (Araneae: Atypidae, Ctenizidae). Journal of Arachnology 13: 291-296.

29. Eberhard WG (2006) Dispersal by Ummidia spiderlings (Araneae, Ctenizidae): Ancient roots of aerial webs and orientation? Journal of Arachnology 34: 254-257.

30. Coyle FA (1985) Balloning behavior of *Ummidia* spiderlings (Araneae, Ctenizidae). Journal of Arachnology 13: 137-138.

31. Raven RJ (1980) The evolution and biogeography of the mygalomorph spider family Hexathelidae (Araneae, Chelicerata). Journal of Arachnology 8: 251-266.

32. Ayoub NA, Garb JE, Hedin M, Hayashi CY (2007) Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1g), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). Molecular Phylogenetics and Evolution 42: 394-409.

33. Bond JE, Hendrixson BE, Hamilton CA, Hedin M (2012) A Reconsideration of the Classification of the Spider Infraorder Mygalomorphae (Arachnida: Araneae) Based on Three Nuclear Genes and Morphology. PLoS ONE 7: e38753.

34. Wunderlich J (1992) Die Spinnen-fauna der Makaronesischen Inseln. Beiträge zur Araneologie 1: 1-619.

35. Wunderlich J (1995) Zu Ökologie, Biogeographie, Evolution und Taxonomie einiger Spinnen der Makaronesischen Inseln. Beiträge zur Araneologie 4: 385-439.

36. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294-299.

37. Simon C, Frati F, Beckenbach A, Crespi B, Liu B, et al. (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87: 651 - 701.

38. Hedin MC (1997) Speciational history in a Diverse clade of habitat-specialized spiders (Araneae: Nesticidae: *Nesticus*): Inferences from geographic-based sampling. Evolution 51: 1929-1945.

39. Giribet G, Edgecombe G, Wheeler WC (1999) Sistematica y filogenia de Artropodos: estado de la cuestion con enfasis en datos moleculares. Boletín de la Sociedad Entomológica Aragonesa 26: 197-212.

40. Hedin MC, Maddison WP (2001) A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). Molecular Phylogenetics and Evolution 18: 386-403.

41. Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, et al. (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Australian Journal of Zoology 46: 419-437.

42. Satler JD, Starrett J, Hayashi CY, Hedin M (2011) Inferring Species Trees from Gene Trees in a Radiation of California Trapdoor Spiders (Araneae, Antrodiaetidae, *Aliatypus*). PLoS ONE 6: e25355.

43. Drummond A, Ashton B, Cheung M, Heled J, Kearse M, et al. (2010) Geneious v.5.3. Available from http://wwwgeneiouscom.

44. Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. Briefings in Bioinformatics 9: 286-298.

45. Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res 38: W7-13.

46. Stephens M, Donnelly P (2003) A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data. American Journal of Human Genetics 73: 1162-1169.

47. Stephens M, Smith NJ, Donnelly P (2001) A New Statistical Method for Haplotype Reconstruction from Population Data. American Journal of Human Genetics 68: 978-989.

48. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

49. McGuire G, Wright F, Prentice MJ (2000) A Bayesian model for detecting past recombination events in DNA multiple alignments. J Comput Biol 7: 159-170.

50. Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, et al. (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics 25: 126-127.

51. Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, et al. (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. Systematic Biology 55: 595-609.

52. Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, et al. (2009) Accelerated Species Inventory on Madagascar Using Coalescent-Based Models of Species Delineation. Systematic Biology 58: 298-311.

53. Ezard T, Fujisawa T, Barraclough TG (2009) splits: SPecies' LImits by Threshold Statistics. Rpackage version 1.0-11/r29. . Available from http://R-Forge.R-project.org/projects/splits/.

54. Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. Systematic Biology 49: 369-381.

55. Borchsenius F (2009) FastGap 1.2.

56. Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Molecular Biology and Evolution 29: 1695-1701.

57. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

58. Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE). New Orleans, LA. pp. 1-8.

59. Rambaut A, Drummond AJ (2009) TRACER v.1.5. 1.4 ed: Available from http://tree.bio.ed.ac.uk/software/tracer/ (accessed on December 1st, 2009).

60. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688 - 2690.

61. Rambaut A (2009) FigTree v. 1.3.1

62. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution.

63. Bidegaray-Batista L, Arnedo M (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. BMC Evolutionary Biology 11: 317.

64. Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27: 570-580.

65. Opatova V, Bond JE, Arnedo MA (2013) Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae). Mol Phylogenet Evol 69: 1135-1145.

66. Carracedo JC, Day S (2002) Canary Islands. Harpenden, Hertfordshire, UK: Terra Publishing. 294 p.

67. Rannala B, Yang Z (2013) Improved Reversible Jump Algorithms for Bayesian Species Delimitation. Genetics 194: 245-253.

68. Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences 107: 9264-9269.

69. Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). Proceedings of the Royal Society B: Biological Sciences 277: 3071-3077.

70. Teacher AGF, Griffiths DJ (2011) HapStar: automated haplotype network layout and visualization. Molecular Ecology Resources 11: 151-153.

71. Excoffier L, Laval G, Schneider S (2005) ARLEQUIN, Version 3.0: an Integrated Software Package for Population Genetics Data Analysis. Evolutionary Bioinformatics Online 1: 47–50.

72. Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. BMC Genetics 6: 13.

73. Hedin M, Starrett J, Hayashi C (2013) Crossing the uncrossable: novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, Antrodiaetus) from California. Mol Ecol 22: 508-526.

74. Raven RJ (1988) A revision of the mygalomorph spider genus Idioctis (Araneae, Barychelidae). American Museum novitates. American Museum Novitates 2929: 1-47.

75. Churchill TB, Raven RJ (1992) Systematics of the intertidal trapdoor spider genus Idioctis (Mygalomorphae: Barychelidae) in the western Pacific with a new genus from the northeast. Memoirs of The Queensland Museum 32: 9-30.

76. Raven RJ (1994) Mygalomorph spiders of the Barychelidae in Australia and the western Pacific. Memoirs of the Queensland Museum 35: 291-706.

77. Sauer EGF, Rothe P (1972) Ratite eggshells from Lanzarote. Canary Islands. Science 176: 43-45.

78. Haq BU, Hardenbol J, Vail PR (1987) Chronology of fluctuating sea level since the Triassic. Science 235: 1156–1167.

79. Pagès M, Chevret P, Gros-Balthazard M, Hughes S, Alcover JA, et al. (2012) Paleogenetic Analyses Reveal Unsuspected Phylogenetic Affinities between Mice and the Extinct <italic>Malpaisomys insularis</italic>, an Endemic Rodent of the Canaries. PLoS ONE 7: e31123.

80. Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution 54: 1839-1854.

81. Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annual Review of Ecology and Systematics 33: 707-740.

82. McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL (2011) Calibrating divergence times on species trees versus gene trees: implications for speciation history of Aphelocoma jays. Evolution; international journal of organic evolution 65: 184-202.

83. Kubatko LS, Gibbs HL, Bloomquist EW (2011) Inferring Species-Level Phylogenies and Taxonomic Distinctiveness Using Multilocus Data in Sistrurus Rattlesnakes. Systematic Biology 60: 393-409.

84. Sánchez-Gracia A, Castresana J (2012) Impact of Deep Coalescence on the Reliability of Species Tree Inference from Different Types of DNA Markers in Mammals. PLoS ONE 7: e30239.

85. Jennings WB, Edwards SV (2005) Speciational history of Australian grass finches (Poephila) inferred from thirty gene trees. Evolution 59: 2033-2047.

86. Bidegaray-Batista L, Ferrández MÁ, Arnedo MA (2013) Winter is coming: Miocene and Quaternary climatic shifts shaped the diversification of Western-Mediterranean Harpactocrates (Araneae, Dysderidae) spiders. Cladistics: n/a-n/a.

87. Oaks JR (2011) A Time-calibrated species tree of Crocodylia reveals a recent radiation of the true crocodiles. Evolution 65: 3285-3297.

88. Townsend TM, Mulcahy DG, Noonan BP, Sites Jr JW, Kuczynski CA, et al. (2011) Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. Molecular Phylogenetics and Evolution 61: 363-380.

89. Santos Guerra A (2002) Flora Vascular Nativa. In: Fernández-Palacios JM, Martín-Esquivel JL, editors. Naturaleza de las Islas Canarias Ecología y Conservación. San Cristobal de La Laguna: Turquesa. pp. 185-192.

90. Brown RP, Yang Z (2010) Bayesian Dating of Shallow Phylogenies with a Relaxed Clock. Syst Biol 59: 119-131.

91. Meco J (2008) Historia geológica del clima en Canarias. Las Palmas de Gran Canaria: Laboratorio de Paleontología, Departamento de Biología, Universidad de Las Palmas de Gran Canaria (ULPGC).

92. Dimitrov D, Arnedo MA, Ribera C (2008) Colonization and diversification of the spider genus Pholcus Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: Evidence for long-term occupancy yet rapid recent speciation. Molecular Phylogenetics and Evolution 48: 596-614.

93. Macias-Hernandez N, Oromi P, Arnedo MA (2010) Integrative taxonomy uncovers hidden species diversity in woodlouse hunter spiders (Araneae, Dysderidae) endemic to the Macaronesian archipelagos. Systematics and Biodiversity 8: 531-553.

94. Cox SC, Carranza S, Brown RP (2010) Divergence times and colonization of the Canary Islands by Gallotia lizards. Molecular Phylogenetics and Evolution 56: 747-757.

95. Arnedo MA, Ferrández MA (2007) Mitochondrial markers reveal deep population subdivision in the European protected spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae). Conservation Genetics 8: 1147-1162.

96. Hedin MC (1997) Molecular phylogenetics at the population/species interface in cave spiders of the Southern Appalachians (Araneae: Nesticidae: *Nesticus*). Molecular Biology and Evolution 14: 309-324.

97. Starrett J, Hedin M (2007) Multilocus genealogies reveal multiple cryptic species and biogeographical complexity in the California turret spider Antrodiaetus riversi (Mygalomorphae, Antrodiaetidae). Molecular Ecology 16: 583-604.

98. Hendrixson BE, Bond JE (2007) Molecular phylogeny and biogeography of an ancient Holarctic lineage of mygalomorph spiders (Araneae: Antrodiaetidae: Antrodiaetus). Molecular Phylogenetics and Evolution 42: 738-755.

99. Decae AE (1996) Systematics of the trapdoor spider genus *Cyrtocarenum* Ausserer, 1871 (Araneae, Ctenizidae). Bulletin of British arachnological Society 10: 161-170.

100. Bond JE, Beamer DA, Lamb T, Hedin M (2006) Combining genetic and geospatial analyses to infer population extinction in mygalomorph spiders endemic to the Los Angeles region. Animal Conservation 9: 145-157.

101. Hamilton CA, Formanowicz DR, Bond JE (2011) Species Delimitation and Phylogeography of *Aphonopelma hentzi* (Araneae, Mygalomorphae, Theraphosidae): Cryptic Diversity in North American Tarantulas. PLoS ONE 6: e26207.

102. Padial J, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. Frontiers in Zoology 7: 16.

103. Yeates D, Seago A, Nelson L (2011) Integrative taxonomy, or iterative taxonomy? Systematic Entomology 36: 209-217.

104. Fujita MK, Leachè AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. Trends in ecology & evolution (Personal edition).

105. Carstens BC, Pelletier TA, Reid NM, Satler JD (2013) How to fail at species delimitation. Molecular Ecology: n/a-n/a.

106. Burbrink FT, Pyron RA (2011) The impact of gene-tree/species-tree discordance on diversification-rate estimation. Diversification rates from gene trees 65: 1851-1861.

107. Brown RM, Siler CD, Lee Grismer L, Das I, McGuire JA (2012) Phylogeny and cryptic diversification in Southeast Asian flying geckos. Molecular Phylogenetics and Evolution 65: 351-361.

108. Pepper M, Doughty P, Fujita MK, Moritz C, Keogh JS (2013) Speciation on the Rocks: Integrated Systematics of the <italic>Heteronotia spelea</italic> Species Complex (Gekkota; Reptilia) from Western and Central Australia. PLoS ONE 8: e78110.

109. Sistrom M, Donnellan SC, Hutchinson MN (2013) Delimiting species in recent radiations with low levels of morphological divergence: A case study in Australian Gehyra geckos. Molecular Phylogenetics and Evolution 68: 135-143.

110. Páez-Moscoso DJ, Guayasamin JM (2012) Species limits in the Andean toad genus Osornophryne (Bufonidae). Molecular Phylogenetics and Evolution 65: 805-822.

111. Zhang C, Zhang D-X, Zhu T, Yang Z (2011) Evaluation of a Bayesian Coalescent Method of Species Delimitation. Systematic Biology 60: 747-761.

112. Arnedo MA, Oromí P, Ribera C (2000) Systematics of the genus *Dysdera* (Araneae, Dysderidae) in the Eastern Canaries. Journal of Arachnology 28: 261-292.

113. Juan C, Ibrahim KM, Oromi P, Hewitt GM (1998) The phylogeography of the darkling beetle, Hegeter politus, in the eastern Canary Islands. Proceedings Of The Royal Society Of London Series B Biological Sciences 265: 135-140.

114. Bidegaray-Batista L, Macias-Hernandez N, Oromi P, Arnedo MA (2007) 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. Molecular Ecology 16: 3198-3214.

115. Pinho C, Harris DJ, Ferrand N (2007) Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. Molecular Ecology 16: 1191-1205.

116. Santos X, Rato C, Carranza S, Carretero MA, Pleguezuelos JM (2012) Complex phylogeography in the Southern Smooth Snake (Coronella girondica) supported by mtDNA sequences

Compleja filogeografía en la culebra lisa meridional (Coronella girondica) confirmada con secuencias de ADN mitocondrial. Journal of Zoological Systematics and Evolutionary Research 50: 210-219.

117. Simon E (1903) Histoire naturelle des araignées. Paris: Roret. 669-1080.

118. Simon E (1909) Etude sur les Arachnides recuellis au Maroc par M. Martinez de la Escalera en 1907. Memorias de la Real Sociedad Española de Historia Natural 6: 1-43.

Supporting material

Table S1.- Specimen information, locality data and GenBank accession numbers. Sequence present if not marked otherwise.

Locality name/ Locality number		Geographic code	Lat/Long	N	GMYC	Sample code	cox1	16s	nad1	EF1g	28S	H3	mt70
Midreshet Ben-Gurion		Israel	30.85179N 34.77654W		*Idiops siriacus*	Z579							No data
		South Africa			*Segregara*	Z912							No data
		South Africa			*Idiops*	MY189					DQ639920		No data
Kemis-oulat-el-Hadj	24	Morocco	31.3964N 9.31604W	5	G5	Z280 Z281 Z282 Z283 Z284							No data
Smi-Mou	23	Morocco	31.50096N 9.6158W	1	G30	Z285							
Jbele Amsittene	26	Morocco	31.15739N 9.69767W	5	G2	Z286 Z287 Z288							No data
					G3	Z289 Z290							No data
Tamanat - Aid-Beoude rd.	27	Morocco	31.00613N 9.59705W	2	G4	Z291 Z292							No data
Imoza rd. nr. Tourarin	29	Morocco	30.64256N 9.70562W	1	G4	Z305							No data
				3	G6	Z294 Z295							No data

Ait-Aisa	30	Morocco	30.17806N 8.29407W	1	G31	Z296 Z297		No data
Aid-Baha	31	Morocco	29.84522N 8.93745	1	Not assigned	Z298	No data	No data
Iguer rd.	28	Morocco	30.90627N 8.31932W	1	G32	Z299		No data
Ouzoud Falls rd.	25	Morocco	31.95973N 6.76811W	5	G1	Z300 Z301 Z302 Z303 Z304		No data
Barranco de Mal Nombre	4	SF	28.09139N 14.28589W	5	G16	Z211 Z212 Z213 Z214 Z215		
Barranco del Ciervo	1	SF	28.0854N 14.37196W	6	G17	Z216 Z217 Z218		
					G18	Z219 Z221		
					G14	Z37		
Pico de Fraile	3	SF	28.10178N 14.35557W	2	G14 G27	Z192 Z190		
Cofete	2	SF	28.10265N 14.38232W	7	G19	Z35 Z36 Z189 Z222 Z223 Z224		

						Z225
Tequital	5	CF	28.27382N 13.98401W	1	G33	Z39
Caldera de la Laguna	6	CF	28.33501N 13.99525W	3	G10	Z228 Z229 Z230
				2	G20	Z226 Z227
Betancuria	7	CF	28.41644N 14.05922W	2	G15	Z38 Z231
Valle de Aguas Verdes	8	CF	28.47845N 14.05128W	4	G13	Z232 Z233 Z234
Caldería de la Roja	12	NF	28.62996N 13.83421W	5	G7	Z245 Z246 Z247 Z248 Z306
Corralejo	10	NF	28.72216N 13.8795W	4	G8	Z242 Z243 Z244
Villaverde	11	NF	28.65061N 13.91518W	1	G8	Z236
Faro Toscón	9	NF	28.70336N 14.00779W	5	G9	Z237 Z238 Z239 Z240 Z241
Salinas de Janubio	13	SWL	2.94288N 13.81886W	3	G12	Z249 Z250 Z251

Tinajo	14	SWL	29.0614N 13.6777W	4	G21	Z276 Z278
Montaña de Tinache, Tinajo	15	SWL			G22	Z193 Z194
Tejía	22	NEL	29.02955N 13.51675W	2	G11	Z271 Z272
Barranco Hondo del Valle	21	NEL	29.14139N 13.48203W	3	G26 G29	Z197 Z268 Z266
Valle de Malpaso	16	NEL	29.1128N 13.51668W	5	G23	Z252 Z253 Z254 Z255 Z256
Valle de Guinate	17	NEL	29.17536N 13.50576W	3	G25	Z257 Z259 Z260
Mirador del Río	20	NEL	29.21336N 1.48111W	7	G24	Z262 Z263 Z264 Z265
Montaña de Mojón, La Graciosa	18	NEL	29.2472N 13.518W	7	G24	Z273
El Vallichuelo, La Graciosa	19	NEL	29.26577N 13.48683W	7	G24	Z274 Z275

Table S2.- Standard genetic diversity indices, nucleotide diversity (π) and haplotype diversity (H) of the *cox1*, *EF1γ*, *H3* and *Hsp70* genes for each GMYC lineage formed by two or more individuals for the entire *Titanidiops* dataset, *T. canariensis*, *T. maroccanus* and the JSF and A clades.

Group	**Locus**	**N_{ind}**	**N_{seq}**	**Pi ± SD**	**H**	**Hd ± SD**
Overall	*cox1*	96	96	0.11848 ± 0.00258	47	0.980 ± 0.004
Titanidiops samples	*16S+tRNALeu*	60	60	0.07353 ± 0.00306	30	0.979 ± 0.007
	nad1	20	20	0.12501 ± 0.01064	18	0.984 ± 0.024
	EF1γ	100	113	0.00668 ± 0.00044	44	0.941 ± 0.012
	28S	53	56	0.00565 ± 0.00054	15	0.837 ± 0.032
	H3	56	89	0.00818 ± 0.00082	26	0.839 ± 0.030
Overall	*cox1*	72	72	0.10317 ± 0.00187	33	0.971 ± 0.007
Titanidiops canariensis	*16S+tRNALeu*	43	43	0.07353 ± 0.00312	30	0.979 ± 0.011
	nad1	17	17	0.10809 ± 0.00982	15	0.978 ± 0.031
	EF1γ	65	82	0.00298 ± 0.00023	23	0.896 ± 0.021
	28S	38	39	0.00181 ± 0.00029	6	0.684 ± 0.047
	H3	41	64	0.00465 ± 0.00063	11	0.721 ± 0.041
	Hsp70	56	70	0.01285 ± 0.00095	38	0.957 ± 0.013
Overall	*cox1*	24	24	0.11129 ± 0.00473	16	0.957 ± 0.025
Titanidiops maroccanus	*16S+tRNALeu*	17	17	0.06566 ± 0.00487	11	0.912 ± 0.056
	nad1	3	3	0.15754 ± 0.04306	3	1.000 ± 0.272
	EF1γ	25	31	0.00527 ± 0.00048	24	0.976 ± 0.016
	28S	15	17	0.00313 ± 0.00049	10	0.897 ± 0.056
	H3	15	25	0.02246 ± 0.00196	22	0.990 ± 0.014
G1	*cox1*	5	5	0.00149 ± 0.00052	3	0.700 ± 0.218
	EF1γ	5	8	0.00146 ± 0.00042	5	0.786 ± 0.151
	H3	2	3	0.01458 ± 0.00524	3	1.000 ± 0.272
G2	*cox1*	3	3	0.00102 ± 0.00048	2	0.667 ± 0.314
	EF1γ	3	4	0.00306 ± 0.00094	4	1.000 ± 0.177

	H3	2	4	0.02875 ± 0.00798	4	1.000 ± 0.177
G3	*cox1*	2	2	0	1	0
	EF1γ	2	2	0.00376 ± 0.00188	2	1.000 ± 0.500
	H3	2	3	0.03128 ± 0.01083	3	1.000 ± 0.272
G4	*cox1*	3	3	0.00346 ± 0.00122	3	1.000 ± 0.272
	EF1γ	3	5	0.00376 ± 0.00114	5	1.000 ± 0.126
	H3	2	4	0.02875 ± 0.00761	4	1.000 ± 0.177
G5	*cox1*	5	5	0.00060 ± 0.00036	2	0.400 ± 0.237
	EF1γ	5	5	0.00075 ± 0.00022	2	0.600 ± 0.175
	H3	2	3	0.03899 ± 0.01191	3	1.000 ± 0.272
G6	*cox1*	3	3	0.00303 ± 0.00143	2	0.667 ± 0.314
	EF1γ	3	3	0.00326 ± 0.00094	3	1.000 ± 0.272
	H3	2	2	0.01163 ± 0.00581	2	1.000 ± 0.500
G7	*cox1*	5	5	0.00119 ± 0.00045	3	0.700 ± 0.218
	EF1γ	5	9	0.00137 ± 0.00016	4	0.806 ± 0.089
	H3	2	3	0.00580 ± 0.00158	3	1.000 ± 0.272
	Hsp70	4	6	0.00515 ± 0.00332	2	0.333 ± 0.215
G8	*cox1*	4	4	0.00077 ± 0.00041	2	0.500 ± 0.265
	EF1γ	4	5	0.00293 ± 0.00071	3	0.800 ± 0.164
	H3	2	3	0.00585 ± 0.00205	3	1.000 ± 0.272
	Hsp70	4	4	0	1	0
G9	*cox1*	5	5	0	1	0
	EF1γ	5	9	0.00208 ± 0.00047	6	0.889 ± 0.091
	H3	2	3	0.00388 ± 0.00183	2	0.667 ± 0.314
	Hsp70	5	6	0.00640 ± 0.00207	2	0.533 ± 0.172
G10	*cox1*	3	3	0.00099 ± 0.00047	2	0.667 ± 0.314
	EF1γ	3	4	0.00210 ± 0.00056	4	1.000 ± 0.177
	H3	2	2	0	1	0
	Hsp70	3	4	0.00348 ± 0.00094	4	1.000 ± 0.177
G11	*cox1*	2	2	0.00446 ± 0.00223	2	1.000 ± 0.500
	EF1γ	2	2	0	1	0
	H3	2	2	0	1	0
	Hsp70	2	2	0.00168 ± 0.00084	2	1.000 ± 0.500
G12	*cox1*	3	3	0	1	0

	EF1γ	3	3	0	1	0
	H3	3	3	0.00220 ± 0.00104	2	0.667 ± 0.314
	Hsp70	3	6	0.00981 ± 0.00255	5	0.933 ± 0.122
G13	*cox1*	4	4	0	1	0
	EF1γ	4	6	0.00171 ± 0.00054	3	0.800 ± 0.172
	H3	2	3	0.00192 ± 0.00091	2	0.667± 0.314
	Hsp70	4	5	0.00842 ± 0.00197	4	0.900 ± 0.161
G14	*cox1*	2	2	0.00594 ± 0.00297	2	1.000 ± 0.500
	EF1γ	2	3	0.00084 ± 0.00040	2	0.667 ± 0.314
G15	*cox1*	2	2	0.00149 ± 0.00074	2	1.000 ± 0.500
	EF1γ	2	2	0	1	0
	H3	2	4	0.02216 ± 0.00824	4	1.000 ± 0.177
	Hsp70	2	3	0.01433 ± 0.00581	3	1.000 ± 0.272
G16	*cox1*	5	5	0.00297 ± 0.00062	4	0.900 ± 0.161
	EF1γ	5	5	0.00146 ± 0.00059	3	0.700 ± 0.218
	H3	2	3	0.03488 ± 0.01473	3	1.000 ± 0.272
	Hsp70	5	6	0.01404 ± 0.00208	5	0.933 ± 0.122
G17	*cox1*	3	3	0	1	0
	EF1γ	3	3	0.00082 ± 0.00039	2	0.667 ± 0.314
	H3	2	3	0.03661 ± 0.01392	3	1.000 ± 0.272
	Hsp70	3	3	0.00771 ± 0.00364	2	0.667 ± 0.314
G18	*cox1*	2	2	0	2	0
	EF1γ	2	2	0	1	0
	H3	2	3	0.00601 ± 0.00211	3	1.000 ± 0.272
	Hsp70	2	2	0.00360 ± 0.00180	2	1.000 ± 0.500
G19	*cox1*	6	6	0	1	0
	EF1γ	3	5	0.00174 ± 0.00048	4	0.900 ± 0.161
	H3	3	6	0.03168 ± 0.00625	6	1.000 ± 0.096
	Hsp70	2	2	0	1	0
G20	*cox1*	2	2	0	2	0
	EF1γ	2	2	0	1	0
	H3	2	3	0.01198 ± 0.00565	2	0.667 ± 0.314
	Hsp70	2	2	0.00176 ± 0.00088	2	1.000 ± 0.500
G21	*cox1*	2	2	0	1	0
	EF1γ	2	2	0.00123 ± 0.00062	2	1.000 ± 0.500

	H3	2	4	0.00192 ± 0.00059	2	0.667 ± 0.204
	Hsp70	2	2	0	1	0
G22	*cox1*	2	2	0.00902 ± 0.00451	2	1.000 ± 0.500
G23	*cox1*	5	5	0	1	0
	EF1γ	5	5	0.00075 ± 0.00022	2	0.600 ± 0.175
	H3	2	3	0.00440 ± 0.00147	3	1.000 ± 0.272
	Hsp70	4	5	0.01136 ± 0.00330	5	1.000 ± 0.126
G24	*cox1*	7	7	0.00340 ± 0.00089	3	0.667 ± 0.160
	EF1γ	7	8	0.00138 ± 0.00038	5	0.786 ± 0.151
	H3	2	4	0.00434 ± 0.00158	3	0.833 ± 0.222
	Hsp70	7	9	0.00761 ± 0.00125	7	0.889 ± 0.091
G25	*cox1*	3	3	0.00101 ± 0.00047	2	0.667 ± 0.314
	EF1γ	3	3	0.00246 ± 0.00086	3	1.000 ± 0.272
	H3	2	4	0.00528 ± 0.00127	4	1.000 ± 0.177
	Hsp70	2	3	0.00790 ± 0.00219	3	1.000± 0.272
G26	*cox1*	2	2	0.00625 ± 0.00313	2	1.000 ± 0.500
	H3	2	2	0.00288 ± 0.00144	2	1.000 ± 0.500
	cox1	20	20	0.03854 ± 0.00530	10	0.889 ± 0.051
Titanidiops canariensis clade JSF	*16S+tRNALeu*	15	15	0.01995 ± 0.00213	12	0.962 ± 0.040
	nad1	8	8	0.04987 ± 0.00732	6	0.893 ± 0.111
	EF1γ	17	19	0.00201 ± 0.00029	6	0.795 ± 0.060
	28S	12	12	0.00079 ± 0.00017	2	0.485 ± 0.106
	H3	13	22	0.00655 ± 0.00042	11	0.853 ± 0.063
	Hsp70	14	16	0.01522 ± 0.00242	8	0.858 ± 0.063
	cox1	43	43	0.07935 ± 0.00286	19	0.950 ± 0.013
Titanidiops canariensis clade A	*16S+tRNALeu*	23	23	0.07338 ± 0.00467	18	0.980 ± 0.018
	nad1	6	6	0.11581 ± 0.02510	6	1.000 ± 0.096
	EF1γ	43	56	0.00259 ± 0.00019	19	0.916 ± 0.017
	28S	20	21	0.00193 ± 0.00057	7	0.595 ± 0.108
	H3	18	29	0.00468 ± 0.00042	7	0.828 ± 0.037
	Hsp70	40	52	0.01094 ± 0.00071	29	0.939 ± 0.022

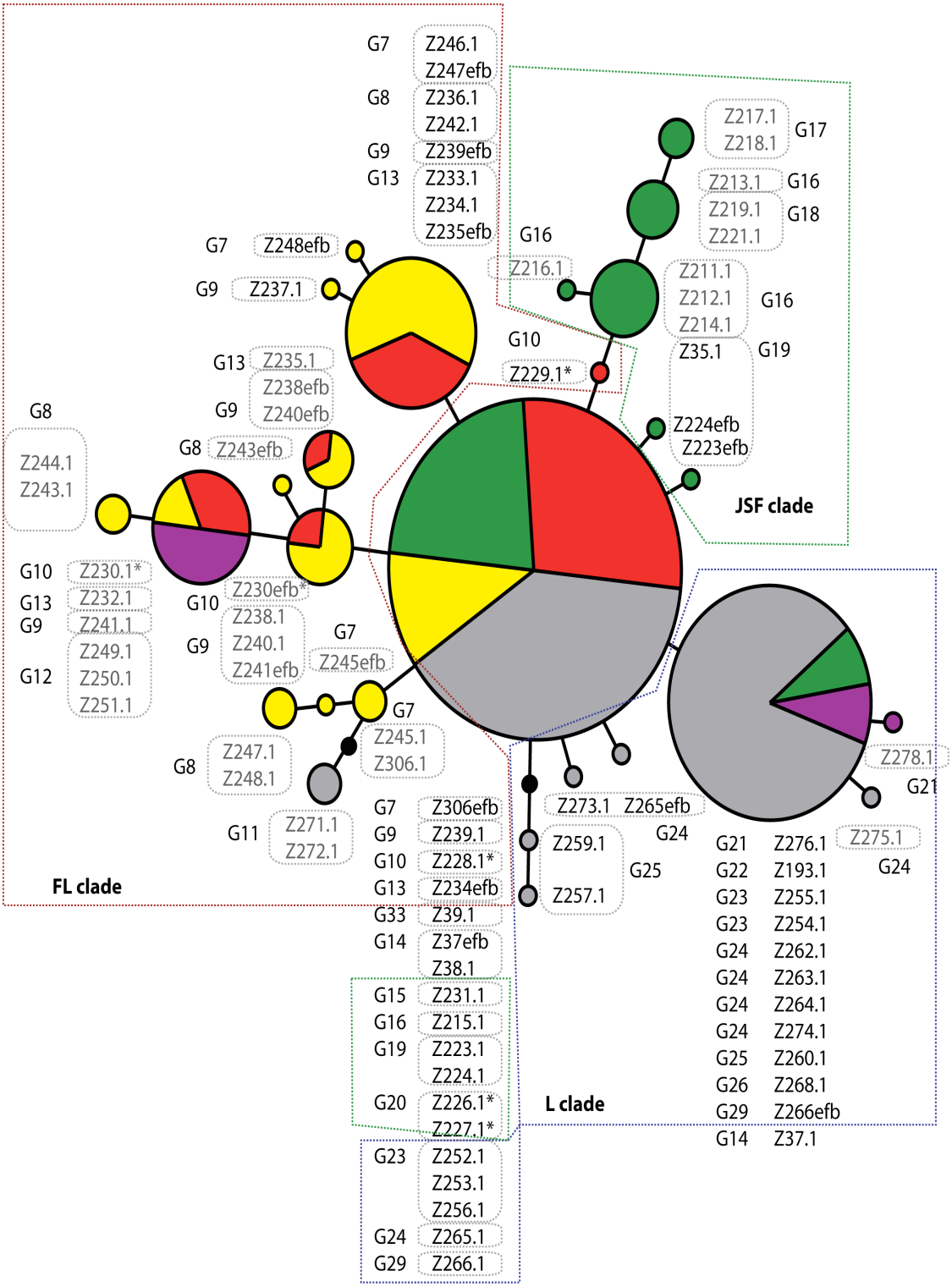


Figure S1.- EF1 g allele network. Circle size is proportional to the allele frequency. Small filled circles represent missing alleles. Each allele is labelled with the individuals and GMYC cluster in which it was found. Asterisks indicate samples from the single locality (6) where the two putative species co-occurred (G10, G20).

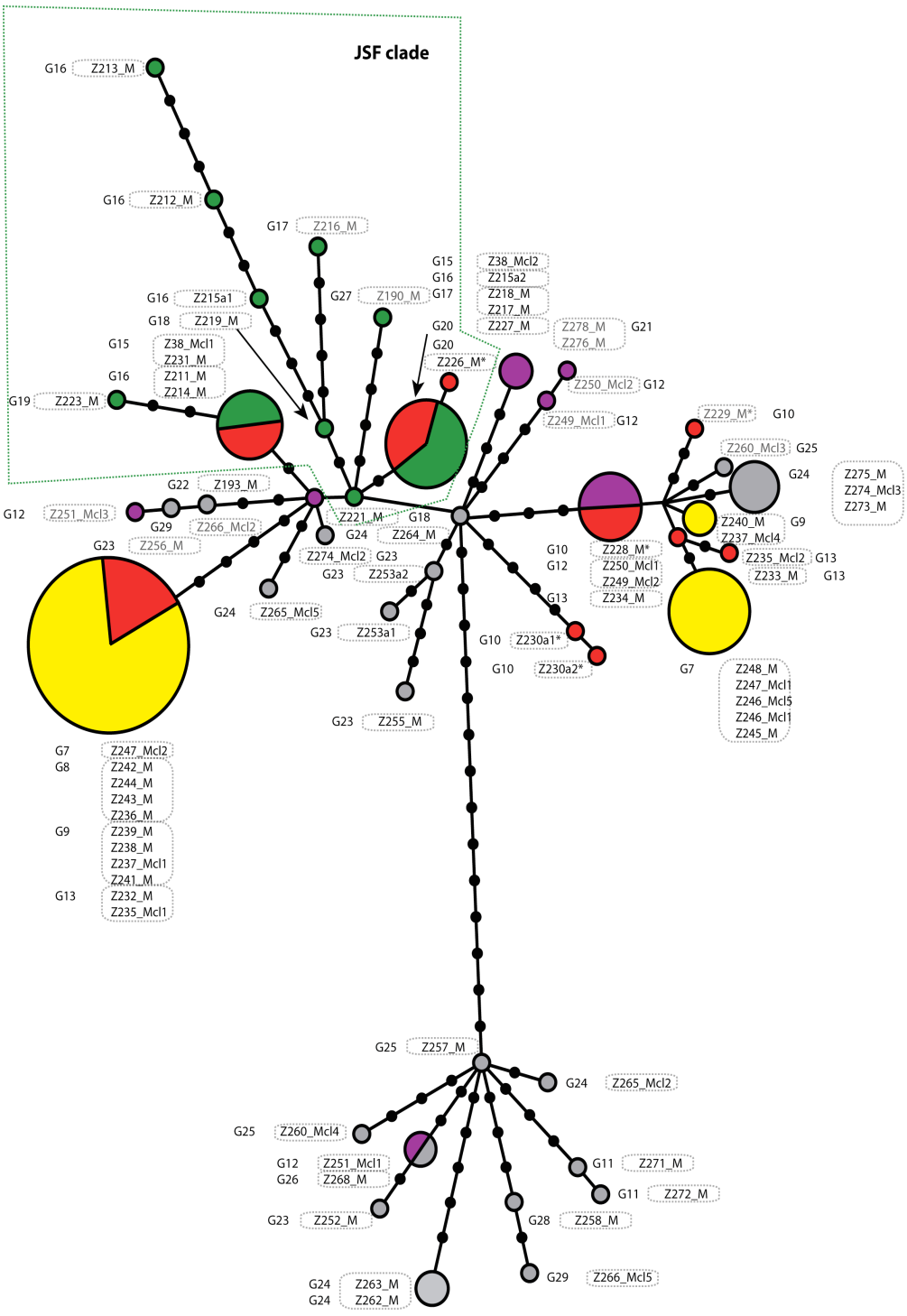


Figure S2.- mt70 allele network. Circle size is proportional to the allele frequency. Small filled circles represent missing alleles. Each allele is labelled with the individuals and GMYC cluster in which it was found. Asterisks indicate samples from the single locality (6) where the two putative species co-occurred (G10, G20).

Annex I

Vera Opatova, Kanchon K Dasmahapatra, Miguel-Ángel Ferrández, Miquel A. Arnedo, Threatened or Threatening: Inferring the population structure of Iberian endangered funnel-web spider *Macrothele calpeiana* (Araneae,Hexathelidae)

Threatened or Threatening: Inferring the population structure of Iberian endangered funnel-web spider *Macrothele calpeiana* (Araneae,Hexathelidae)

Vera Opatova[1], Kanchon K Dasmahapatra[2], Miguel-Ángel Ferrández[3], Miquel A. Arnedo[1]

[1]Institut de Recerca de la Biodiversitat & Departament de Biologia Animal, Universitat de Barcelona. Avinguda Diagonal 643, 08028 Barcelona, Spain. [2]Department of Biology, University of York, Wentworth Way, York YO10 5DD, United Kingdom. [3]Sociedad para el Estudio y la Conservación de las Arañas, c/Villafranca 24, 1_C, 28028 Madrid, Spain.

General framework

Both species of the genus *Macrothele* that inhabit Europe are of conservation concern. *Macrothele calpeiana* (Walckenaer, 1805) is found in the southern Iberian Peninsula and *Macrothele cretica* Kulczynski, 1903 is endemic to Crete (Snazell & Allison 1989). *Macrothele cretica* has been included in the IUCN red list under the data deficient category and *M. calpeiana* is the only spider protected by European Union legislation. The inclusion of *M. calpeiana* in both the Bern Convention listings and the European Union Habitat Directive under the Annex IV ‘‘Animal and plant species of community interest in need of strict protection’’ (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31992L0043:EN:HTML) was motivated by its putative role as quality bioindicator of the diminishing cork forest areas, which led to the conclusion, that the species could also be endangered (Snazell & Allison 1989).

However, the close affinity of *M. calpeiana* towards the primary cork oak vegetation and thus its vulnerability has been challenged by field observations (Ferrández & Ferrández de Cespedes 2001; Van Helsdingen 1993; Van Helsdingen & Decae 1992). *M. calpeiana* has been found in human disturbed localities, which called into question both its bioindicator role and vulnerability

status. Moreover the fact that the species thrives on anthropogenic habitats made the authors referred to *M. calpeiana* as an "aggressive invasive species" (Van Helsdingen & Decae 1992).

The narrow ecological preferences of the species were further questioned by species niche modelling (Jimenez-Valverde & Lobo 2007; Jiménez-Valverde & Lobo 2006). On the other hand, genetic studies have shown deep population structure across its highly fragmented present day distribution and restricted or no gene flow between small populations, which may qualify as vulnerable (Arnedo & Ferrández 2007).

Aims

In the present study we conducted an extensive sampling throughout the know distribution of the species to reveal the population structure of *M. calpeiana* and determine the level of fragmentation of the species using a multi locus approach. We combine sequence data with genome-wide screening method (i.e. AFLP) to test if deeply divergent lineages detected within the species might actually correspond to fully independent species. This information is fundamental to assess the actual conservation status of the species. We further test the suggestion that *M. calpeiana* localities recently reported far from its known distribution range might be the result of human mediated introductions.

Preliminary Conclusions

The results corroborate the deep divergences found within *M. calpeiana* already reported in previous studies (Arnedo & Ferrández 2007), but also reveal shared haplotypes or GMYC lineages between the localities within the main clades, suggesting gene flow across localities sampled in continuously inhabited geographical ranges.

Five deeply divergent clades were detected in the DNA sequence concatenated analyses (Portugal, Huelva, S1, SE, S2), while only two clusters were inferred from the AFLP data (Portugal+Hueva and S1+SE+S2). Interestingly, significantly less variability (i.e. mixed ancestry in the individuals and its percentage) was detected in the S1+SE+S2 clade, which supports a relatively recent area expansion in this lineage as proposed by Arnedo &

Ferrández (2007). The Portugal+Huelva cluster, on the other hand, show higher AFLP diversity which may be interpreted as ancestral gene polymorphisms in otherwise completely isolated populations.

The deepest divergences trace back to Pleistocene, indicating that the Quaternary climatic oscillations (starting ~2.5 Ma) may have driven pulses of population extinctions and expansions and area range shifts, which contributed to generate the patchy phylogeorgaphic pattern. We propose to further tests this hypothesis by means of an Approximate Bayesian Computation approach (ABC) in collaboration with Dr. Alejandro Sánchez-Garcia, from the Genetics Department of the Universitat de BArcelona. Finally, we will assess the effect of climatic changes on the distribution range of *M. calpeiana* by using Species Distribution Modelling (SDM) tools to predict its distribution under both a Last Glacial Maximum and a Global Warming scenario. This is an on-going study in collaboration with Dr. Alberto Jiménez Valverde of the Museo Naciaonal de Ciencias Naturales in Madrid.

Materials and Methods

Taxonomic sampling

The samples used in present study were obtained by the authors from extensive sampling campaigns covering the known *M. calpeiana* distribution in the Iberian Peninsula (Fig. 1) under the collection permit SGYB-AFR-CMM-2835. Additional specimens from other parts of the Mediterranean region representing putative human mediated introductions of *M. calpeiana* were kindly donated by colleagues. Detailed locality data is included in Table S1. Given the protection status of the target species, most samples were collected by a non-lethal technique inducing a leg autotomy (III pair), previously applied successfully in protected mygalomorph spiders (Longhorn 2002; Longhorn *et al.* 2007; Machkour M'Rabet *et al.* 2009). EF1g sequences of *M. cretica* and *M. gigas* (Opatova & Arnedo 2014) representing outgroups in present study were downloaded from GenBank.

DNA extraction, PCR amplification and sequencing

Whole genomic DNA was extracted with the help of the SpeedTools Tissue Extraction Kit (Biotools). Partial fragments of four mitochondrial and

three nuclear genes were sequenced in present study: 5' half of the Cytochrome oxidase I (*cox1*) (the animal barcode), the 3' half of the 16s rDNA (*16S*), the *tRNA-Leu* (*L1*) and the 5' half of the NADH dehydrogenase subunit I (*nad1*), a fragment of the Elongation factor-1 gamma (*EF1*γ) and a fragment comprising Internal transcribed spacer I and II (*ITS I-II*), respectively.

The PCR amplifications were carried out with the following primer combinations. *Cox1* with the primer pair (Folmer *et al.* 1994), the fragment comprising the *16S*, *L1* and *nad1* with the primer pair LR-N-13398 (Simon *et al.* 1994) / and N1-J-12261 (Hedin 1997) or, alternatively, only the 3' half of the 16S with LR-N-13398 combined with LR-J-12864 (Macías-Hernández *et al.* 2008). All mitochondrial fragments were successfully amplified at annealing temperature range of 43-46°C. The *EF1*γ fragment was amplified with primers ER1gF78/EF1γR1258 and EF1gF218/EF1gR1090 (Ayoub *et al.* 2007). The *ITS* fragment was amplified with primer pair CAS18sF1/CAS28sB1d (Ji *et al.* 2003) at 48°C.

All the PCR reactions were carried out in reaction volume of 25μl, purified using ExoSAP-IT (USB Corporation) and sequenced in both directions using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on ABI 3700 automated sequencer at *Centres Científics i Tecnològics* of the University of Barcelona (CCiTUB, www.ccit.ub.edu) Spain.

The chromatograms were assembled and edited in Geneious v. 5.3.6. (Drummond *et al.* 2012a)

Sequence alignment and recombination testing

The alignments of *cox1*, *nad1* and *EF1G* were trivial since no length variation was observed in the sequences. The *16S*, *tRNA-Leu* and *ITS* gene fragments presented length polymorphism due to the indel mutations and their alignments were obtained with the online version of MAFFT v. 6 (available at http://mafft.cbrc.jp/alignment/server/, (Katoh & Toh 2008)) using the Q-INS-i approach with default settings (gap opening penalty GOP = 1.53 and offset value set to 0.0).

Recombination was tested in TOPALi v 2.5 (McGuire *et al.* 2000; Milne *et al.* 2009) by means of the difference of sum of squares method (DSS). The size

of sliding window was set to 500bp in all fragments except for *16s+L1* and *nad1* where a 300bp frame was used instead.

Delimitation of putative independent evolutionary lineages

The General Mixed Yule Coalescence model (GMYC) (Pons *et al.* 2006) was used to delimitate the putative independent evolutionary lineages within the complete *M. calpeiana cox1* haplotype data set. RaxML v.7.2.8. (Stamatakis 2006) under the graphical interface raxmlGUI (Silvestro & Michalak 2011) was used to infer the topology from Maximum Likelihood search based on 100 iterations and independent GTRCAT model applied to each codon position. The ultrametric tree was obtained in PATHd8 (Britton 2007) The GMYC analysis was carried out in the R (http://www.r-project.org) environment using the SPLITS package (Ezard *et al.* 2009).

Phylogenetic analyses

In present study, we conducted the phylogenetic analyses using two complementary approaches: (1) the concatenation approach assuming a common tree for all the sequenced genes and (2) coalescent-based species tree/gene tree approach allowing each gene an independent gene tree (see *Estimation of divergence times* section). The concatenated approach was conducted on two different data sets: first, on the mitochondrial genes matrix representing the individuals unique *cox1* haplotypes, and second, on matrix with all mitochondrial and nuclear genes used in this study sampled for one individual per each GMYC clade (see *Estimation of divergence times* section). The concatenated matrices were created using Geneious v. 5.3.6. (Drummond *et al.* 2012a).

Gap in the *16s-L1* fragment were coded using a presence/absence coding approach of Simons and Ochoterena (Simmons & Ochoterena 2000) as implemented in FastGap 1.2 (available at http://www.aubot.dk/FastGap_home.htm, (Borchsenius, 2009)). The greedy algorithm implemented in the program PARTITIONFINDER (Lanfear *et al.* 2012) was used to select the best evolutionary models and partitioning schemes.

The Bayesian inference (BI) analyses were conducted in MrBayes v. 3.1.2 (Ronquist & Huelsenbeck 2003). A partition scheme with the corresponding evolutionary models were defined for each gene fragment and the restriction model was assigned for the binary coded gaps. Two independent runs of $5x10^7$ generations with 8 MCMC (Markov Chain Monte Carlo) chains each, starting from random trees and resampling each 1000 generations were run remotely at the CIPRES server (http://www.phylo.org/ Miller et al. 2010). The first 20% of the generations were discarded as a *burn-in* for the analyses. Convergence of the runs was assessed by monitoring the standard deviation of split frequencies (<0.01) in TRACER v.1.5 (Rambaut & Drummond 2009).

Maximum Likelihood (ML) analyses were conducted in RaxML v.7.2.8. (Stamatakis 2006) under the graphical interface raxmlGUI (Silvestro & Michalak 2011). Independent GRT+G+I substitution models were assigned to each partition, and a binary model was applied to the gaps. The best maximum likelihood tree was selected from 100 independent iterations and node support assessed with 1000 replicates of bootstrap resampling.

All trees were visualized and manipulated with the program FigTree v. 1.3.1 (Rambaut, 2009 available at http://tree.bio.ed.ac.uk/software/figtree/).

Estimation of divergence times

Divergence time analyses were conducted in a Bayesian framework using both concatenation and a coalescent-based approach with the help of the program BEAST 1.8.0. (Drummond *et al.* 2012b). *Cox1*, *16s, L1, nad1, EFG1* and *ITS* gene fragments from a single individual representing each *M. calpeiana* GMYC cluster and *M. cretica* representing an outgroup were concatenated in the single matrix. In an attempt to facilitate convergence and reduce computation effort, partitioning by codon position was omitted in the analysis and all the mitochondrial genes were treated as a single partition. A substitution model selected in PARTIONFINDER was assigned to each gene fragment.

Because of the lack of fossil record and relevant biogeographic events that could be used as calibration points, we used mitochondrial substitution rate for ground-dwelling spider genus *Parachtes* (Bidegaray-Batista & Arnedo 2011) and mygalomorph specific rate for *EF1*γ gene (Opatova *et al.* 2013). A normal

prior was assigned to the ucld.mean parameter of the lognormal relaxed clock for both *mitochondrial* and *EF1γ* partitions, initial and mean value 0.012722067 and standard deviation 0.00454 for *mitochondrial* and 0.00117 with standard deviation of 0.00014 for *EF1γ* respectively. Because of the very little variation in the *ITS* fragment, a strict clock with a uniform prior was assigned to its ucld.mean with lower and upper bounds set between 0.0001 and 0.0115, with starting value at 0.001. The upper constrain corresponds to the universal mitochondrial substitution rate selected under the assumption that the nuclear genes are about one order of magnitude slower than mitochondrial genes. Three independent runs of $8x10^7$ generations were run remotely at the CIPRES server (http://www.phylo.org/, Miller et al. 2010) sampling every 1000 generations. The results were visualized with TRACER (Rambaut & Drummond 2007) in order to check the convergence between the runs and correct mixing of the chains. Subsequently, the results of the individual runs were combined and resampled at lower frequency in BEAST accompanying program LOGCOMBINER. The first 20% of the generations of each run was discarded as a burn-in. A consensus chronogram was inferred with TREEANNOTATOR.

The species tree and *M. calpeiana* divergence timeframe was assessed by means of multi-gene coalescent approach as implemented in *BEAST (Heled & Drummond 2010). Five groups (Portugal, Huelva, South1, South2 and South-east) corresponding to the deeply divergent clades with continuous geographic distribution recovered in the analyses performed on the concatenated mitochondrial matrix based on unique haplotypes. *M. cretica* was used as outgroup in this analyses. The simplified partition scheme and lognormal relaxed and strict clock were applied respectively as described in BEAST analysis. Three independent runs of 150 millions of generations, with sampling every 1000 generations, were run remotely on the CIPRES portal (http://www.phylo.org/, Miller et al. 2010). The runs were combined and resampled at lower frequency in LOGCOMBINER. A consensus chronogram was inferred with TREEANNOTATOR.

Haplotype networks and population structure inferred from DNA sequences

Standard genetic diversity indices, including the nucleotide (π) and haplotype diversity (H) indices, were calculated in DnaSP 5.10.1 (Librado &

Rozas 2009) for *cox1* for the entire dataset, for all sampled locations with at least two records and for the four main geographic clades defined as: Portugal, Huelva, South and East. Haplotype networks were constructed using the minimum spanning tree method as implemented in the program HapStar (Teacher & Griffiths 2011) based on the output provided by Arlequin (Excoffier *et al.* 2005).

Genotyping and population clustering analyses

The amplified fragment length polymorphism (AFLP) (Vos *et al.* 1995) became useful technique to address population genetics, phylogeographic and species delimitation issues especially in non-model organisms, where sufficiently variable nuclear loci are not usually readily available. The method does not require any previous knowledge about the genome of the studied organism and with relatively simple protocol, hundreds of loci can be generated in short time in relatively economic way (Bensch & Akesson 2005; Dasmahapatra *et al.* 2009, Meudt & Clarke 2007; Mueller & Wolfenbarger 1999).

Spectrophotometry performed on NanoDrop 3300 (Thermo Scientific) was performed on the samples sequenced for *Cox1* in order to assess the DNA quality. Only samples showing normal absorbance curves and concentrations higher than 10ng/μl were retained for further AFLP analyses. A modified Vos *et al.*'s protocol (Vos *et al.* 1995) described in Madden *et al*. (2004) was used in AFLP genotyping with four primer combinations TaqI-CAC + EcoRI-ACA; TaqI-CAG + EcoRI-ACA; TaqI-CTG + EcoRI-ATG; TaqI-CCA + EcoRI-ACA (Madden *et al.* 2004). The sequencing was performed at University College of London sequencing facility. The profiles were visualized and analyzed in Peak Scanner v 1.0 (Applied Biosystems). Samples presenting aberrant profiles were discarded and only peaks of minimum height of 100 rfu well distinguishable from the background noise were retained for the analysis (Bonin *et al.* 2004).

Final 187 AFLP profiles, where all four geographic areas were represented + 14 replicates representing approximately 8% of the dataset generated form the extraction stage were scored by automatic scoring algorithm implemented in RawGeno 2.0.1 (Arrigo *et al.* 2009). The peak between 100 – 300bp were scored using default binning options and reviewed manually, the bin reproducibility was set to 95 - 100% yielding a combined matrix with 1.1%

mean error rate per locus.

The number of genetically different groups within the dataset was inferred in STRUCTURE v 2.3.4 (Falush *et al.* 2003; Falush *et al.* 2007; Pritchard *et al.* 2000) under admixture model with correlated allele frequencies. The value of the parameter λ was estimated from the whole dataset assuming one population (K=1) as recomended in the user's manual (Falush *et al.* 2003), the obtained value (λ= 0.6865) was then implemented in 10 replicates runs for K values between 1 and 6. Each run consisted of 10^5 burn-in and 10^5 data collection. We used the ΔK method (Evanno *et al.* 2005) to select the best-supported number of inferred clusters as implemented in STRUCTURE HARVESTER (Earl 2012).

Results

Sampling, sequencing and recombination testing

Detailed information about specimens, sampling locations, geographic clade assignment and sequence GenBank accession numbers is listed in Table S1. Sampling localities are shown in the map in Fig. 1. A total of 284 *M. calpeiana* specimens and two outgroups: *M. cretica* and *M. gigas* were sequenced in the present study. The length and variability of the gene fragments for *M. calpeiana* ingroup was as follows: *cox1* (675bp, 443 variables, 330 parsimony informative (pi)), *16S-L1* (533bp, 65 variables, 52 pi) and *nad1* (370bp, 142 variables, 131 pi) and the nuclear *EF1γ* (881bp, 48 variables, 47 pi), *ITS* (989bp, 9 variables, 7pi). No recombination was detected within any of the gene fragments used in this study.

GMYC-based lineage delimitation

The *cox1* haplotype data matrix, including 95 *M. calpeiana* terminals, was analysed using the single-threshold option of the GMYC algorithm, which was shown not to be significantly worse than the multiple-threshold option. The GMYC algorithm identified 24 independent entities/clusters, confidence interval (CI)= 24-37 (p= 0.0001956283) (Table S1, Fig. 2). In nearly 60% of the cases, the GMYC clusters represented single localities, in the remaining portion GMYC membership was widely shared across localities, up to 3 in Portugal clade, up to 6 in both Huelva and eastern clades and up to 10 localities belonged to the

same cluster in the southern clade. On the other hand, 6 geographically proximate localities from Portugal and Southeastern Iberian peninsula presented more than one GMYC cluster, usually closely related.

Phylogenetic analyses

All mitochondrial genes of the individuals presenting a unique *cox1* haplotype were concatenated with the outgroup sequences in a single matrix for ML and Bayesian analyses. The concatenated matrix consisted of 97 terminals of 1,615 bp (572 variable, 322pi) and 12 binary coded gaps. The partition scheme and corresponding evolutionary models selected by PARTITIONFINDER are summarised in Table 1.

The Bayesian and maximum likelihood (-lnL 6877.700764) analyses of the concatenated data matrix resulted in similar tree topologies, although the Bayesian inference yielded higher supports (Fig. 3). In both analyses the monophyly of *M. calpeiana* was supported and a monophyletic clade formed by the Portuguese samples was recovered as a sister clade to all remaining *M. calpeiana* samples. Both analyses recovered four well-supported clades inside Spanish *M. calpeiana* roughly corresponding to the geographic regions, although the relationship among them remains mostly unresolved.

The Huelva clade was placed, although with low support, as a sister clade to the clade comprising the samples from southern and eastern Iberian Peninsula which was further divided into two clades. A clade (hereafter referred to as S1) comprising few samples from the easternmost verge of the Southern area was recovered as a sister group to a poorly supported clade that included two monophyletic groups comprising southern and eastern samples (SE clade) and clade composed exclusively of southern samples (S2).

Despite the relatively close relationships among most the samples within the clades, their exact relationships were mostly incongruent between the ML and BI topologies. Deeply divergent haplotypes were detected both in Huelva and S2 clade.

The *M. calpeiana* samples collected across the Mediterranean, well outside its native area of occurrence, were placed in SE clade in case of RA258 and in the S2 clade in case of PK117, G63 and G64 sharing the Z312 haplotype.

Estimation of divergence times

The divergence time estimates were performed on a concatenated matrix of mitochondrial and nuclear genes sampled for one individual representing each GMYC cluster, *M. cretica* was used to root the analyses.

Overall, the tree topology and the clade supports inferred by Beast (Fig. 4) were similar to those found in the ML and Bayesian analyses with exception of the position of Huelva clade, which was placed as a sister group to the Portuguese clade albeit with low support. The root age assigned to the split between *M. calpeiana* and *M. calpeiana* clade was estimated to 52 million years ago (Ma) (100.1–23.7 Ma). The most common recent ancestor (TMRCA) of *M. calpeiana* was estimated at 4.44 Ma (8.2– 2.09 Ma) corresponding to the split between the Portugal + Huelva clades from the remaining *M. calpeiana* diversity (S1, SE, S2 clades). The split between the Portuguese and Huelva clade was dated to 3.71 Ma (7.04-1.68), while the diversification between the S1 and SE, S2 clades began 2.78 Ma (5.22-1.33). The SE diverged from S2 clade about, 2.41 Ma (4.44-1.13) although this particular node did not received sufficient support.

Coalescent approach

The coalescent approach performed on five groups representing deeply divergent clades (see Materials and Methods section) resulted in a same topology and supports as the concatenated analyses. The resulting *M. calpeiana* species tree (Fig. 4) was divided into two clades, one comprising the Portugal and Huelva populations, albeit the sister species relationship between them was not supported, and the other one included the remaining individuals from S1, SE and S2 populations. In the second clade, similarly to the concatenated analyses, the sister species relationship of S1 to SE, S2 populations was recovered with high support, while the SE – S2 sister group relationship was not supported.

Lineage divergence times were slightly younger than those determined in the concatenated analysis. The exception was the root age, which was estimated more than twice as old in the coalescent analyses: 52.24 Ma (100-23.7 Ma) than in the concatenated approach results 121.76 Ma (242.53-50.37

Ma) respectively. The TMRCA of *M. calpeiana* dated back to 3.1 Ma (5.94-1.37 Ma). The divergence between Portugal and Huelva clades occurred approximately 2.26 Ma (4.52-0.84 Ma). The S1 clade diverged from SE and S2 clades approximately 1.85 Ma (3.29-0.58 Ma), while the divergence between SE and S2 clades dates back to 1.12 Ma (2.36-0.35 Ma).

Population analyses

Standard genetic diversity indices, nucleotide diversity (π) and haplotype diversity (H), were calculated for the *cox1* for each locality represented by two or more individuals, for the entire *M. calpeiana* dataset, and for the four geographic regions (Portugal, Huelva, South, East). The results are summarised in Table S2.

Similarly, *M. calpeiana cox1* haplotype networks were constructed independently or each geographic area (Fig. 5, Table S3). The Portuguese group included 10 unique haplotypes separated at most by 12 steps between the most divergent ones. The two most common haplotypes were shared by 11 and 6 individuals respectively and divided by 9 steps. The Huelva group yielded 20 unique haplotypes, most of them separated by less than 12 mutation steps. The two most common haplotypes shared across the localities included 19 and 16 individuals respectively. A highly divergent haplotype corresponding to the samples from Valverde de Camino sampling location was separated from its closest relative haplotype by more than 40 missing mutation steps. In the southern clade most of the haplotypes were very closely related, separated from the central haplotype by four mutation steps as maximum. The most common haplotype was shared across localities. Three different groups of highly divergent haplotypes were detected in the data set, the fist was separated by 24 mutation steps from the main haplotype group and was distantly related to the two remaining divergent groups separated by 30 and 37 steps to its closest relative, which were separated from the main haplotype groups by approximately 50 and 60 mutation steps. These groups corresponded to the samples recovered in S1 and ES clades instead of S2. The Eastern group included 15 unique haplotypes. Most of the samples presented the most commonly shared haplotype 15 (19 individuals) or were closely related

to it i.e. separated by one or two mutation steps. The two most divergent haplotypes were separated from the central one by 14 and 21 steps.

In all groups haplotypes shared across the localities were detected, especially both Southern and Eastern groups presenting star-like patterns, where most of the haplotypes are separated from the most common central haplotype by only few mutation steps.

Genotyping and population clustering analyses

The AFLP method yielded 189 loci (1.1% mean error rate per locus) scored for 187 representing all four geographic areas. Two clusters were identified from the samples in the STRUCTURE analyses by ΔK method from 10 independent runs (maximum at K2 =2867.575703). Most of the samples from Portugal and Huelva clades were placed in the cluster 1, while most of the remaining samples were placed in cluster 2 (Fig. 6). Some degree of mixed ancestry was detected in more than 60% of the samples. Interestingly, cluster 2 ancestry over 80% were inferred in about 23% Portugal + Huelva samples (cluster 1), while similar degree of ancestry of cluster 1 was detected in only about 7% of the S1, S2 and SE samples.

Concatenated mtDNA		**Concatenated mtDNA + nucDNA time**		**Coalescent mtDNA + nucDNA**	
Partition	**Model**	**Partition**	**Model**	**Partition**	**Model**
cox1 1st	HKY+I	mtDNA	HKY+I +G	mtDNA	GTR+I+G
cox1, *nad1* 2nd	HKY+I	*EF1*γ	HKY+I	*EF1*γ	HKY+I
cox1 3rd	HKY+G	*ITS*	JC	*ITS*	JC
nad 1st, *16SL1*	GTR+I+G				
nad 3rd	HKY+G				

Table 1.- Evolutionary models and partition schemes selected by PARTITIONFINDER for the different analyses. Concatenated mtDNA + nucDNA time refers to concatenated analyses conducted with BEAST including informed substitution rate priors.

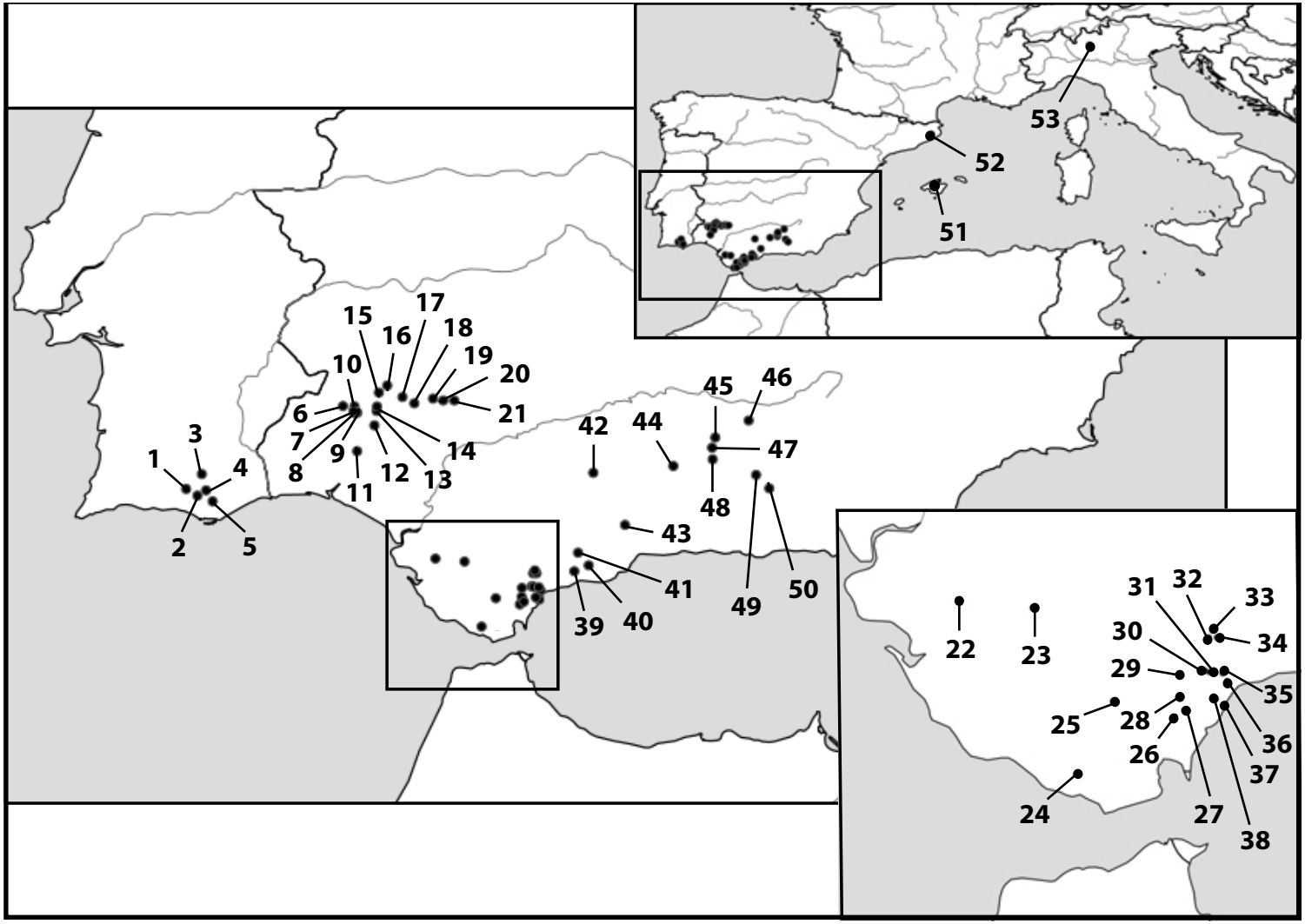


Fig. 1. Sampling locations of *Macrothele calpeiana*, detailed locality information can be found in Table S1

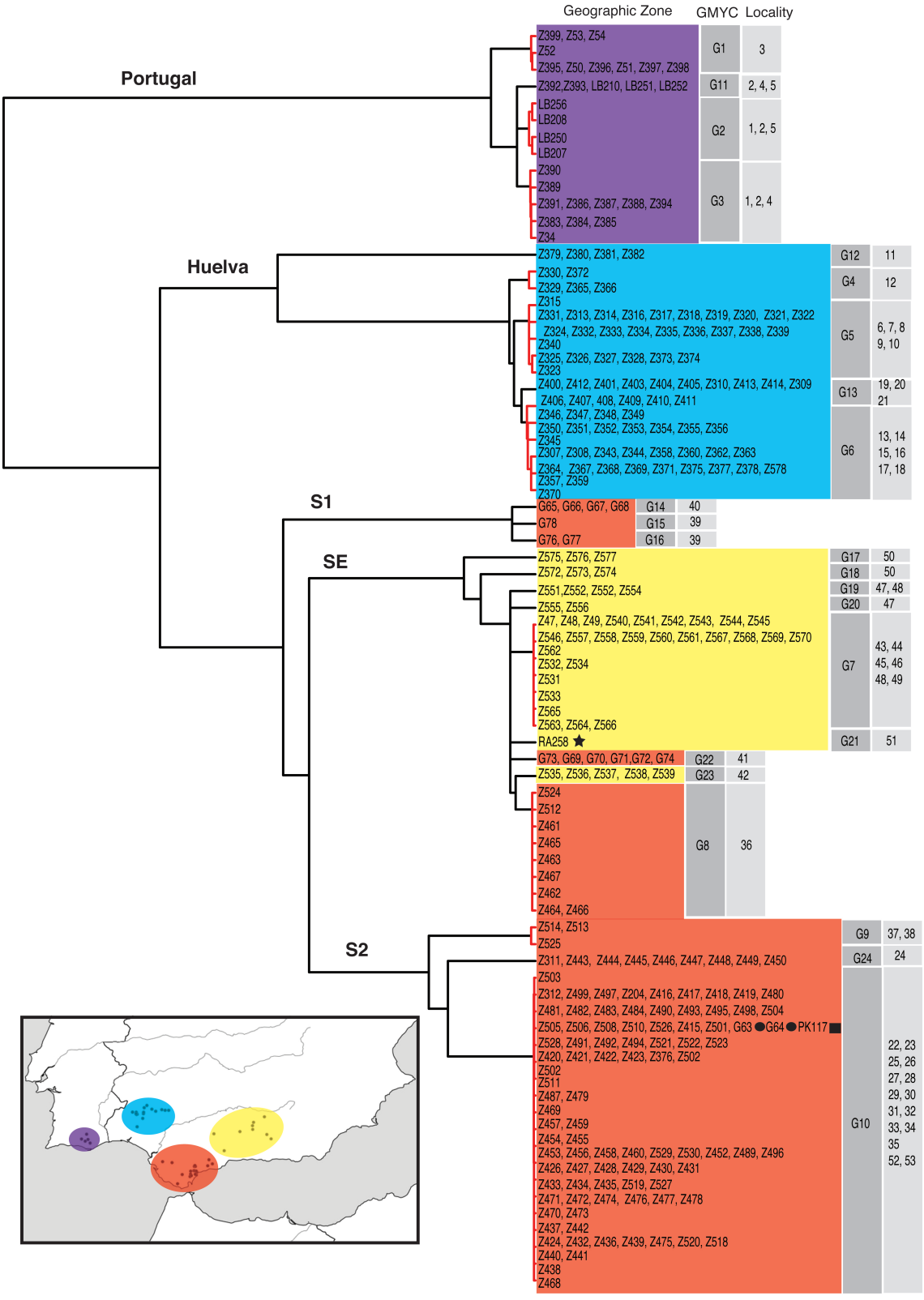


Fig. 2. Ultrametric *cox1* tree: clades identified as independent GMYC clusters in the SPLITS analyses, labelled as in Table S1. Locality column: localities where the respective GMYC cluster was collected. Colour codes represent geographic areas as shown in the map, Purple: Portugal, Blue: Huelva, Red: South, Yellow: East. Clade names: S1: South 1, SE: South East, S2: South 2

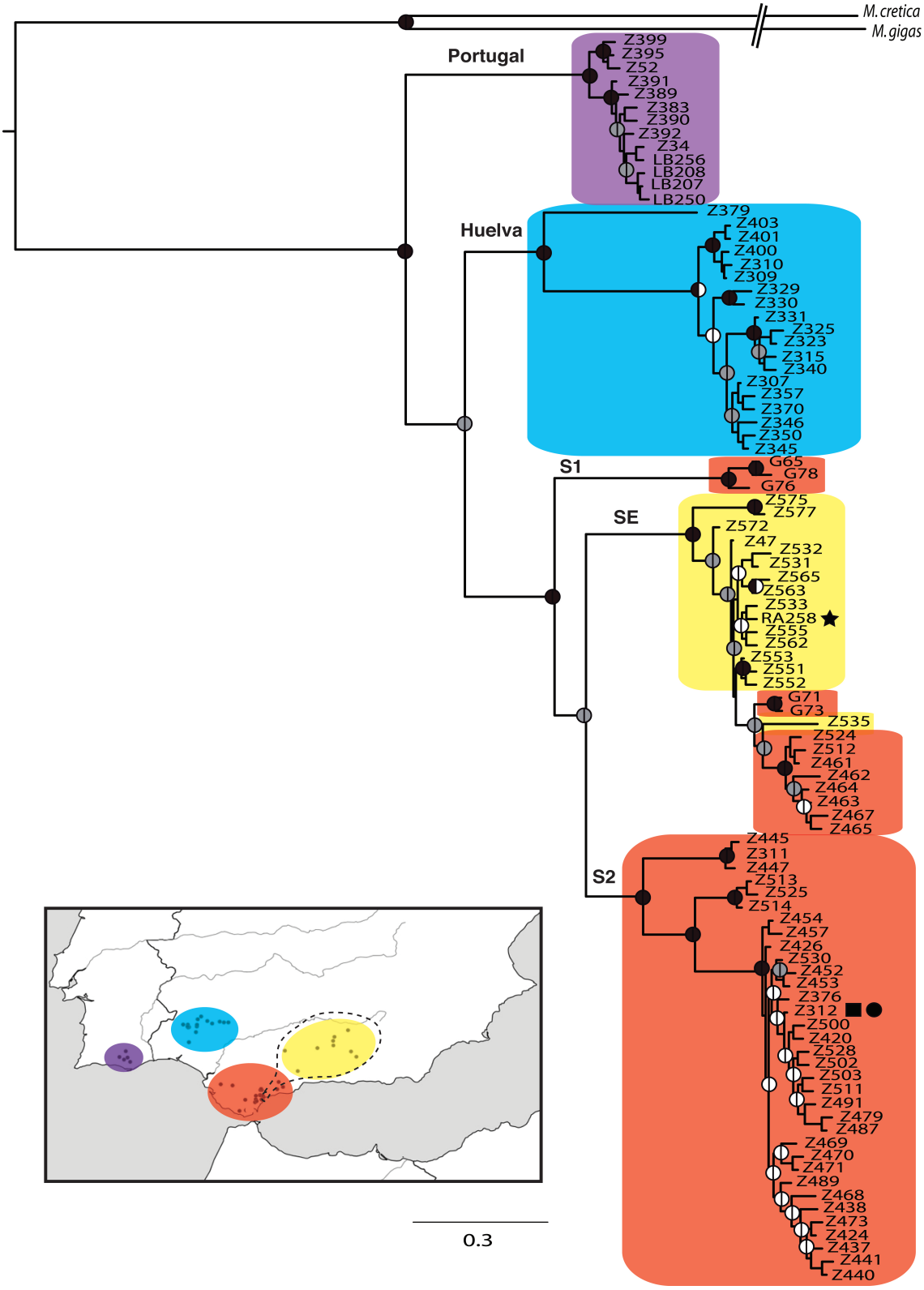


Figure 3.- Topology obtained in the concatenated Bayesian analyses. Terminals represent unique cox1 haplotypes. Dots on nodes denote support as follows: left semi-circles are Bayesian posterior probabilities (PP) and right ones are maximum likelihood bootstraps, black = PP > 0.95, ML bootstrap support >80%, grey = clade determined but with support values less than the thresholds above, white = topology not determined. Star: Sample from Balearic Islands, Square and Dot next to Z312: Same haplotype in samples from Italy and Girona, Catalonia

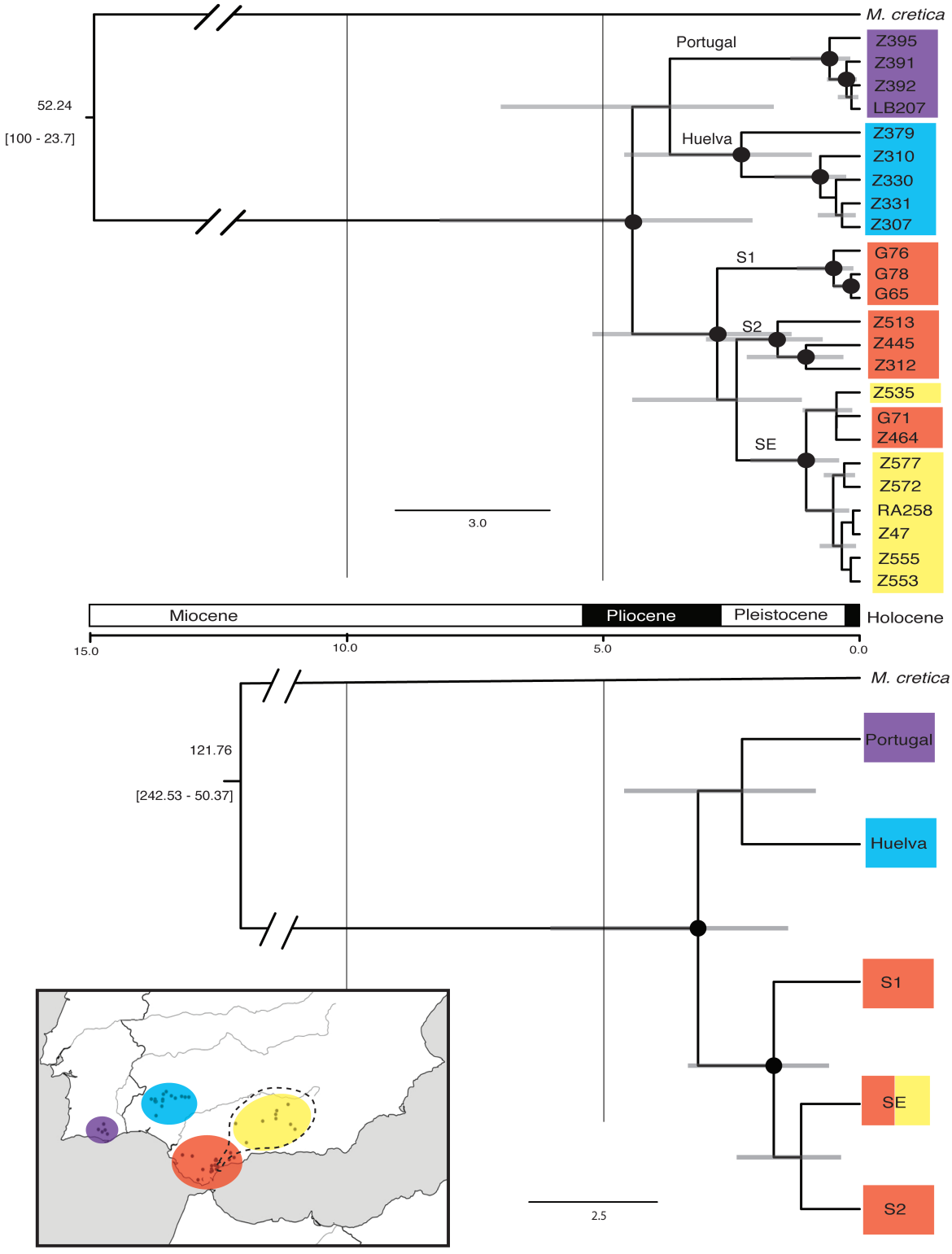


Figure 4.- Divergence time estimates inferred in program BEAST. Dots on nodes denote the PP support > 0.95. Upper half: concatenated approach, each GMYC clade is represented by one individual. Lower half: coalescent approach, topology inferred by *BEAST algorithm. Numbers at the bases of the trees represent the root age and confidence intervals. Colour codes correspond to the geographic ranges marked in the map. Clade abbreviations: S1: South 1, SE: South-east, S2: South 2. Geologic scale bar is in million years (MY).

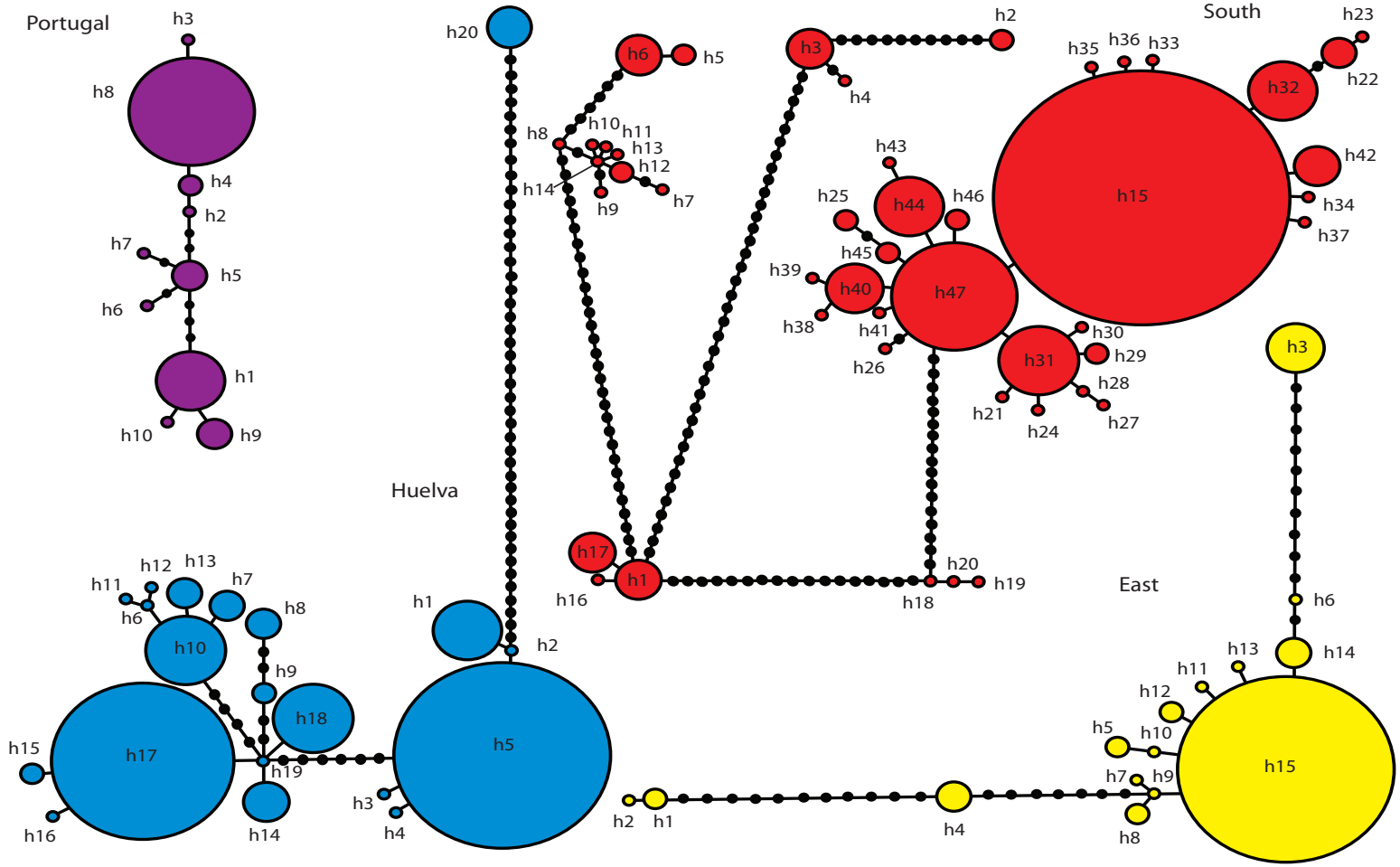


Figure 5.- *Cox1* haplotype network. Circle size is proportional to the haplotype frequency. Small filled circles represent missing haplotypes. See Table S3 for individual haplotype assignment.

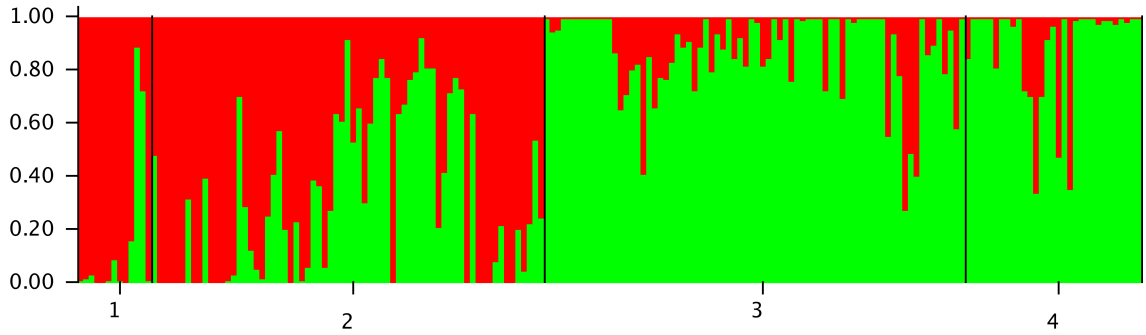


Figure 6.- STRUCTURE clustering analyses results for K=2 performed in on AFLP data. Red: Cluster 1, Green: Cluster 2. Each bar represents a single individual ordered according to its geographic origin: 1: Portugal, 2: Huelva, 3: South, 4: East. Scale on the right refers to an ancestry percentage.

References

Arnedo MA, Ferrández MA (2007) Mitochondrial markers reveal deep population subdivision in the European protected spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae). *Conservation Genetics* **8**, 1147-1162.

Arrigo N, Tuszynski JW, Ehrich D, Gerdes T, Alvarez N (2009) Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics* **10**, 33.

Ayoub NA, Garb JE, Hedin M, Hayashi CY (2007) Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1g), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). *Molecular Phylogenetics and Evolution* **42**, 394-409.

Bensch S, Akesson M (2005) Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology* **14**, 2899-2914.

Bidegaray-Batista L, Arnedo M (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology* **11**, 317.

Bonin A, Bellemain E, Eidesen Bronken P*, et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Mol Ecol* **13**, 3261 - 3273.

Borchsenius F (2009) FastGap 1.2. http://www.aubot.dk/FastGap_home.htm.

Britton T (2007) Estimating Divergence Times in Large Phylogenetic Trees. *Systematic Biology* **56**, 741-752.

Dasmahapatra KK, Elas M, Hill RI, Hoffman JI, Mallet J (2009) Mitochondrial DNA barcoding detects some species that are real, and some that are not. *Molecular Ecology Resources* **9999**.

Drummond A, Ashton B, Buxton S*, et al.* (2012a) Geneious v5.6. *Available from http://www.geneious.com*.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012b) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*.

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN, Version 3.0: an Integrated Software Package for Population Genetics Data Analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.

Ezard T, Fujisawa T, Barraclough TG (2009) splits: SPecies' LImits by Threshold Statistics. Rpackage version 1.0-11/r29. . Available from http://R-Forge.R-project.org/projects/splits/.

Falush D, Stephens M, Pritchard J (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567 - 1587.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**, 574-578.

Ferrández MA, Ferrández de Cespedes H (2001) Arachnida. In: *Los Invertebrados no insectos de la "Directiva Hábitat" en España.* (eds. Ramos MA, Bragado D, Fernández J), pp. 133-144. Ministerio de Medio Ambiente. Dirección Genaral de la Conservación de la naturaleza.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294-299.

Hedin MC (1997) Speciational history in a Diverse clade of habitat-specialized spiders (Araneae: Nesticidae: *Nesticus*): Inferences from geographic-based sampling. *Evolution* **51**, 1929-1945.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**, 570-580.

Ji YÄ, Zhang DÄ, He LÄ (2003) Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. *Molecular Ecology Notes* **3**, 581-585.

Jimenez-Valverde A, Lobo JM (2007) Potential distribution of the endangered spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae) and the impact of climate warming. *Acta Zoologica Sinica* **53**, 865-876.

Jiménez-Valverde A, Lobo JM (2006) Establishing reliable spider (Araneae, Araneidae and¬†Thomisidae) assemblage sampling protocols: estimation of¬†species richness, seasonal coverage and¬†contribution of¬†juvenile data to¬†species richness and¬†composition. *Acta Oecologica* **30**, 21-32.

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**, 286-298.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution* **29**, 1695-1701.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.

Longhorn S (2002) Non-lethal DNA sampling from CITES II protected "tarantula" spiders of Belize. *Las Cuevas Newsletter* **9**, 8-9.

Longhorn SJ, Nicholas M, Chuter J, Vogler AP (2007) The utility of molecular markers from non-lethal DNA samples of the CITES II protected "tarantula" Brachypelma vagans (Araneae, Theraphosidae). *Journal of Arachnology* **35**, 278-292.

Machkour M'Rabet S, Hènaut Y, Dor A*, et al.* (2009) ISSR (Inter Simple Sequence Repeats) as molecular markers to study genetic diversity in tarantulas (Araneae, Mygalomorphae). *Journal of Arachnology* **37**, 10-14.

Macías-Hernández N, Oromí P, Arnedo MA (2008) Patterns of diversification on old volcanic islands as revealed by the woodlouse-hunter spider genus *Dysdera* (Araneae, Dysderidae) in the eastern Canary Islands. *Biological Journal of the Linnean Society* **94**, 589–615.

Madden JR, Lowe TJ, Fuller HV*, et al.* (2004) Neighbouring male spotted bowerbirds are not related, but do maraud each other. *Animal Behaviour* **68**, 751-758.

McGuire G, Wright F, Prentice MJ (2000) A Bayesian model for detecting past recombination events in DNA multiple alignments. *J Comput Biol* **7**, 159-170.

Meudt HM, Clarke AC (2007) Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science* **12**, 106-117.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*, pp. 1-8, New Orleans, LA.

Milne I, Lindner D, Bayer M*, et al.* (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* **25**, 126-127.

Mueller U, Wolfenbarger L (1999) AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* **14**, 389 - 394.

Opatova V, Arnedo MA (2014) From Gondwana to Europe: inferring the origins of Mediterranean Macrothele spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae. *Invertebr Syst*. In press.

Opatova V, Bond JE, Arnedo MA (2013) Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae). *Mol Phylogenet Evol* **69**, 1135-1145.

Pons J, Barraclough TG, Gomez-Zurita J*, et al.* (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology* **55**, 595-609.

Pritchard JK, Stephens M, Donnelly PJ (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Rambaut A, Drummond AJ (2009) TRACER v.1.5. Available from http://tree.bio.ed.ac.uk/software/tracer/ (accessed on December 1st, 2009).

Rambaut A (2009) FigTree v. 1.3.1 http://tree.bio.ed.ac.uk/software/figtree/

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572 - 1574.

Silvestro D, Michalak I (2011) raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution*, DOI: 10.1007/s13127-13011-10056-13120.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49**, 369-381.

Simon C, Frati F, Beckenbach A*, et al.* (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am* **87**, 651 - 701.

Snazell RG, Allison R (1989) The genus Macrothele Ausserer (Araneae; Hexathelidae) in Europe. *Bulletin of the British Arachnological Society* **8**, 65-72.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.

Teacher AGF, Griffiths DJ (2011) HapStar: automated haplotype network layout and visualization. *Molecular Ecology Resources* **11**, 151-153.

Van Helsdingen PJ (1993) Can *Macrothele calpeiana* (Walckenaer) (Araneae, Hexathelidae) be used as a bio - indicator? *Bulletin de la Société neuchâteloise des sciences naturelles* **116**, 253-258.

Van Helsdingen PJ, Decae A (1992) Ecology, distribution and vulnerability of *Macrothele calpeiana* (Walckenaer) (Araneae, Hexathelidae). *Tljdschrift voor Entomologie* **135**, 169-178.

Vos P, Hogers R, Bleeker M*, et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**, 4407 - 4414.

Supporting information

Table S1.- Specimen information, locality data and GenBank accession numbers, XXXX indicates sequenced fragment.

Group/Locality	Locality number	Geographic code	Lat/Long	N	GMYC	Sample code	cox1	16s	nad1	EF1g	ITS
Rocha da Pena	1	Portugal	37.25063N 8.09819W	6	2	LB207	XXXX	XXXX			
						LB208	XXXX	XXXX			
					3	Z386	XXXX				
						Z387	XXXX				
						Z388	XXXX				
						Z389	XXXX				
Fonte Benemola	2	Portugal	37.19662N 8.00646W	6	2	LB256	XXXX	XXXX	XXXX		
					3	Z34	XXXX	XXXX	XXXX	XXXX	XXXX
						Z383	XXXX				
						Z384	XXXX				
						Z385	XXXX				
					11	LB210	XXXX	XXXX			
Ameixal	3	Portugal	37.37123N 7.97087W	10	1	Z50	XXXX	XXXX	XXXX	XXXX	XXXX
						Z51	XXXX	XXXX	XXXX	XXXX	XXXX
						Z52	XXXX				
						Z53	XXXX				
						Z54	XXXX				
						Z395	XXXX	XXXX	XXXX	XXXX	XXXX
						Z396	XXXX				
						Z397	XXXX				
						Z398	XXXX				
						Z399	XXXX				
Barranco do Velho	4	Portugal	37.23917N 7.93617W	5	3	Z390	XXXX				
						Z391	XXXX	XXXX	XXXX	XXXX	XXXX
						Z394	XXXX				

					11	Z392	XXXX			
						Z393	XXXX			
San Brás de Alpontel	5	Portugal	37.15N 7.8833W	3	2	LB250	XXXX			
					11	LB251	XXXX	XXXX		
						LB252	XXXX	XXXX		
Molino de la Romera	6	Huelva	37.92345N 6.825W	6	5	Z325	XXXX			
						Z326	XXXX			
						Z327	XXXX			
						Z328	XXXX			
						Z373	XXXX			
						Z374	XXXX			
Aguafría	7	Huelva	37.88387N 6.73797W	3	5	Z338	XXXX			
						Z339	XXXX			
						Z340	XXXX			
Santa Ana la Real, finca La Hoyuela	8	Huelva	37.86995N 6.70882W	6	5	Z313	XXXX			
						Z314	XXXX			
						Z316	XXXX			
						Z315	XXXX			
						Z317	XXXX			
						Z318	XXXX			
Santa Ana la Real	9	Huelva	37.8659N 6.72768W	6	5	Z319	XXXX			
						Z320	XXXX			
						Z321	XXXX			
						Z323	XXXX			
						Z322	XXXX			
						Z324	XXXX			
Jabugo	10	Huelva	37.92124N 6.73201W	7	5	Z331	XXXX	XXXX	XXXX	XXXX
						Z332	XXXX			
						Z333	XXXX			
						Z334	XXXX			
						Z335	XXXX			
						Z336	XXXX			
						Z337	XXXX			
Valverde de Camino	11	Huelva	37.55703N 6.71015W	4	12	Z379	XXXX	XXXX	XXXX	XXXX
						Z380	XXXX			

						Z381	XXXX				
						Z382	XXXX				
Campofrío	12	HU	37.7665N	5	4	Z329	XXXX				
			6.57006W			Z330	XXXX				
						Z372	XXXX				
						Z365	XXXX				
						Z366	XXXX				
Aracena	13	Huelva	37.88378N	4	6	Z307	XXXX	XXXX	XXXX	XXXX	
			6.55223W			Z308	XXXX	XXXX	XXXX	XXXX	XXXX
						Z343	XXXX				
						Z344	XXXX				
Carboneras	14	Huelva	37.92007N	5	6	Z367	XXXX				
			6.54878W			Z368	XXXX				
						Z369	XXXX				
						Z370	XXXX				
						Z371	XXXX				
Casa Vallecutillo	15	Huelva	38.03289N	5	6	Z362	XXXX				
			6.53377W			Z363	XXXX				
						Z364	XXXX				
						Z377	XXXX				
						Z378	XXXX				
Finca los Helechales	16	Huelva	38.09032N	6	6	Z357	XXXX				
			6.46621W			Z358	XXXX				
						Z359	XXXX				
						Z360	XXXX				
						Z375	XXXX				
						Z578	XXXX				
Cala	17	Huelva	37.99675N	7	6	Z350	XXXX				
			6.3422W			Z351	XXXX				
						Z352	XXXX				
						Z353	XXXX				
						Z354	XXXX				
						Z355	XXXX				
						Z356	XXXX				
Crío las Lanchas	18	Huelva	37.94531N	5	6	Z345	XXXX				
			6.24619W			Z346	XXXX				

						Z347	XXXX				
						Z348	XXXX				
						Z349	XXXX				
Hoya de St. María - el Real de la Jara rd.	19	Huelva	37.98419N 6.09432W	6	13	Z400	XXXX				
						Z401	XXXX				
						Z403	XXXX				
						Z404	XXXX				
						Z405	XXXX				
						Z411	XXXX				
Mirador Sierra de Padrona	20	Huelva	37.96881N 6.01126W	5	13	Z406	XXXX				
						Z407	XXXX				
						Z408	XXXX				
						Z409	XXXX				
						Z410	XXXX				
Embalse el pintado	21	Huelva	37.96851N 5.92228W	5	13	Z309	XXXX	XXXX	XXXX	XXXX	XXXX
						Z310	XXXX	XXXX	XXXX	XXXX	XXXX
						Z412	XXXX	XXXX	XXXX		XXXX
						Z413	XXXX				
						Z414	XXXX				
Estella del Marques	22	South	36.685N 6.07472W	11	10	Z468	XXXX				
						Z469	XXXX				
						Z470	XXXX				
						Z471	XXXX				
						Z473	XXXX				
						Z472	XXXX				
						Z474	XXXX				
						Z476	XXXX				
						Z477	XXXX				
						Z478	XXXX				
						Z475	XXXX				
San Jose del Valle	23	South	36.661N 5.83889W	2	10	Z529	XXXX				
						Z530	XXXX	XXXX	XXXX	XXXX	XXXX
Facinas	24	South	36.1333N 5.7W	9	24	Z311	XXXX	XXXX	XXXX	XXXX	XXXX
						Z443	XXXX				
						Z444	XXXX				
						Z445	XXXX	XXXX	XXXX	XXXX	XXXX

La Ina	25	South	36.36373N 5.58438W	9	10	Z446 Z447 Z448 Z449 Z450 Z452 Z453 Z454 Z455 Z456 Z457 Z458 Z459 Z460	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX	XXXX	XXXX	XXXX
El Picacho, Alcala de los Gazules	26	South	36.31302N 5.38991W	7	10	Z436 Z437 Z438 Z439 Z440 Z441 Z442	XXXX XXXX XXXX XXXX XXXX XXXX XXXX				
Pto Galiz – Ubrique	27	South	36.33553N 5.35676W	3	10	Z518 Z519 Z520	XXXX XXXX XXXX				
Embalse de Guadalgacin-Ubrique	28	South	36.37462N 5.37467W	13	10	Z376 Z424 Z426 Z427 Z428 Z429 Z430 Z431 Z432 Z433 Z434 Z435	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX				

						Z527	XXXX				
El Bosque	29	South	36.4474N 5.37214W	21	10	Z312	XXXX	XXXX	XXXX	XXXX	XXXX
						Z479	XXXX				
						Z480	XXXX				
						Z481	XXXX				
						Z482	XXXX				
						Z483	XXXX				
						Z484	XXXX				
						Z485	XXXX				
						Z487	XXXX				
						Z489	XXXX				
						Z490	XXXX				
						Z491	XXXX				
						Z492	XXXX				
						Z493	XXXX				
						Z494	XXXX				
						Z495	XXXX				
						Z496	XXXX				
						Z497	XXXX				
						Z498	XXXX				
Benamahoma	30	South	36.45647N 5.27727W	2	10	Z204	XXXX	XXXX	XXXX	XXXX	XXXX
						Z526	XXXX				
El Bosque – Grazalema rd.	31	South	36.46061N 5.29297W	7	10	Z504	XXXX				
						Z505	XXXX				
						Z506	XXXX				
						Z508	XXXX				
						Z510	XXXX				
						Z511	XXXX				
Río Gudalete	32	South	36.56585N 5.2813W	9	10	Z415	XXXX				
						Z416	XXXX				
						Z417	XXXX				
						Z418	XXXX				
						Z419	XXXX				
						Z420	XXXX				
						Z421	XXXX				
						Z422	XXXX				

						Z423	XXXX				
Coripe	33	South	36.58895N	3	10	Z521	XXXX				
			5.26799W			Z522	XXXX				
						Z523	XXXX				
El castaño	34	South	36.56634N	5	10	Z499	XXXX				
			5.25406W			Z501	XXXX				
						Z500	XXXX				
						Z502	XXXX				
						Z503	XXXX				
Pto. Boyar	35	South	36.45298N	1	10	Z528	XXXX				
			5.23604W								
Villaluenga del	36	South	36.42642N	8	8	Z512	XXXX				
Rosario			5.22146W			Z524	XXXX				
						Z461	XXXX				
						Z462	XXXX				
						Z463	XXXX				
						Z464	XXXX	XXXX	XXXX	XXXX	XXXX
						Z465	XXXX				
						Z466	XXXX				
						Z467	XXXX				
El Colmenar	37	South	36.3524N	2	9	Z513	XXXX	XXXX	XXXX	XXXX	XXXX
			5.23311W			Z514	XXXX				
Cortes de la Frontera	38	South	36.37045N	1	9	Z525	XXXX				
			5.25973W								
Istan to Monda	39	South	36.5823N	3	15	G78	XXXX				
			4.9476W				XXXX				
					16	G76	XXXX	XXXX	XXXX		XXXX
						G77	XXXX	XXXX	XXXX		XXXX
Monda-Marbella rd	40	South	36.63018N	4	14	G65	XXXX				
			4.8311W			G66	XXXX				
						G67	XXXX				
						G68	XXXX				
						G65	XXXX	XXXX	XXXX	XXXX	XXXX
Pto. de las Abejas,	41	South	36.733N	6	22	G70	XXXX				
Yunquera			4.917W			G71	XXXX	XXXX	XXXX	XXXX	XXXX
						G72	XXXX				

						G73	XXXX	XXXX	XXXX	XXXX	XXXX
						G74	XXXX				
Puente Gentil	42	East	37.38147N	5	23	Z535	XXXX				
			4.79338W			Z536	XXXX				
						Z537	XXXX				
						Z538	XXXX				
						Z539	XXXX				
El Torcal de	43	East	36.95858N	4	7	Z531	XXXX				
Antequera			4.53598W			Z532	XXXX				
						Z533	XXXX				
						Z534	XXXX				
Priego de Córdoba	44	East	37.43641N	6	7	Z540	XXXX				
			4.14345W			Z541	XXXX				
						Z542	XXXX				
						Z543	XXXX				
						Z544	XXXX				
						Z545	XXXX				
Los Villares	45	East	37.66796N	5	7	Z557	XXXX				
			3.80346W			Z558	XXXX				
						Z559	XXXX				
						Z560	XXXX				
						Z561	XXXX				
Mancha Real	46	East	37.80628N	5	7	Z562	XXXX				
			3.53201W			Z563	XXXX				
						Z564	XXXX				
						Z565	XXXX				
						Z566	XXXX				
Valdepeñas de Jaén	47	East	37.58545N	5	19	Z552	XXXX				
			3.82922W			Z553	XXXX	XXXX	XXXX	XXXX	XXXX
						Z554	XXXX				
					20	Z555	XXXX				
						Z556	XXXX				
Frailes	48	East	37.49108N	2	7	Z546	XXXX				
			3.82442W								
					19	Z551	XXXX				

Iznalloz	49	East	37.36468N 3.47183W	7	7	Z47	XXXX	XXXX	XXXX	XXXX	XXXX
						Z48	XXXX	XXXX	XXXX	XXXX	
						Z49	XXXX				
						Z567	XXXX				
						Z568	XXXX				
						Z569	XXXX				
						Z570	XXXX				
Sierra de Huétor	50	East	37.25549N 3.36518W	6	17	Z575	XXXX				
						Z576	XXXX				
						Z577	XXXX				
					18	Z572	XXXX	XXXX	XXXX	XXXX	XXXX
						Z573	XXXX				
						Z574	XXXX				
Inca	51	Mallorca	39.71875N 2.91554E	1	21	RA258	XXXX				
Girona region	52	Catalonia	---	1	10	PK117	XXXX				
Villa de Alme	53	Italy	45.75N 9.6167E	2	10	G63	XXXX				
						G64	XXXX				
M. cretica		Crete, nr. Topolia, Greece	35.3941N 23.6716E			LB289	XXXX			KJ628323	
M. gigas		Iriomote Is., Funaura, Japan	---			LB169	XXXX			KJ628326	

Table S2.- Standard genetic diversity indices, nucleotide diversity (π) and haplotype diversity (H) of the *cox1* for each locality represented by two or more individuals for the entire *M. calpeiana* dataset and the main geographic clades.

Group/Locality	**Locality number**	**GMYC**	**N_{ind}**	**Pi ± SD**	**H**	**Hd ± SD**
Overall	---	---	290	0.06184±0.0000014	57	0.920±0.010
Portugal	---	---	30	0.00424±0.00048	10	0.823±0.052
Huelva	---	---	84	0.01424±0.00262	17	0.888±0.018
South	---	---	126	0.02780±0.00301	41	0.921±0.016
East	---	---	45	0.00945±0.00180	15	0.807±0.055
Rocha da Pena	1	2, 3	6	0.00561±0.00161	4	0.8±0.172
Fonte Benemola	2	2, 3, 11	6	0.02920±0.01649	5	0.933±0.122
Ameixal	3	1	10	0.00101±0.00028	3	0.6±0.131
Barranco do Velho	4	3, 11	5	0.00148±0.00041	3	0.8±0.164
San Brás de Alpontel	5	2, 11	3	0.00141±0.00067	2	0.667±0.314
Molino de la Romera	6	5	6	0	1	0
Aguafría	7	5	3	0.00099±0.00047	2	0.667±0.314
Santa Ana la Real, finca La Hoyuela	8	5	6	0.00049±0.00032	2	0.333±0.215
Santa Ana la Real	9	5	6	0.00049±0.00032	2	0.333±0.215
Jabugo	10	5	7	0	1	0
Valverde de Camino	11	12	4	0	1	0
Campofrío	12	4	5	0.00267±0.00078	2	0.6±0.175
Aracena	13	6	4	0	1	0
Carboneras	14	6	5	0.00059±0.00035	2	0.4±0.237
Casa Vallecutillo	15	6	5	0	1	0
Finca los Helechales	16	6	6	0.00079±0.00026	2	0.533±0.172
Cala	17	6	7	0	1	0
Crío las Lanchas	18	6	5	0.00119±0.0007	2	0.4±0.237
Hoya de St. María - el Real de la Jara rd.	19	13	6	0.00158±0.00039	4	0.8±0.172
Mirador Sierra de Padrona	20	13	5	0	1	0
Embalse el pintado	21	13	5	0.0009±0.00026	2	0.6±0.175
Estella del Marques	22	10	11	0.00312±0.00084	6	0.727±0.144
San Jose del Valle	23	10	2	0.00148±0.00074	2	1.0±0.5

Facinas	24	24	9	0.00142±0.00033	3	0.667±0.105
La Ina	25	10	9	0.00313±0.00063	4	0.778±0.110
El Picacho, Alcala de los Gazules	26	10	7	0.00268±0.00064	5	0.905±0.103
Pto Galiz – Ubrique	27	10	3	0.00099±0.00047	2	0.667±0.314
Embalse de Guadalgacin-Ubrique	28	10	13	0.00065±0.00026	3	0.41±0.154
El Bosque	29	10	21	0.00222±0.00048	5	0.652±0.101
Benamahoma	30	10	2	0	1	0
El Bosque – Grazalema rd.	31	10	7	0.02074±0.01392	3	0.524±0.209
Río Gudalete	32	10	9	0.00082±0.00013	2	0.556±0.090
Coripe	33	10	3	0	1	0
El castaño	34	10	5	0.00187±0.0005	4	0.9±0.161
Villaluenga del Rosario	36	8	8	0.00392±0.00082	7	0.964±0.077
El Colmenar	37	9	2	0.00148±0.00074	2	1.0±0.5
Istan to Monda	39	15, 16	3	0.01216±0.00573	2	0.667±0.314
Monda-Marbella rd	40	14	4	0	1	0
Pto. de las Abejas, Yunquera	41	22	6	0.00081±0.00026	2	0.533±0.172
Puente Gentil	42	23	5	0	1	0
El Torcal de Antequera	43	7	4	0.00346±0.00108	3	0.833±0.222
Priego de Córdoba	44	7	6	0	1	0
Los Villares	45	7	5	0	1	0
Mancha Real	46	7	5	0.00237±0.00089	3	0.7±0.218
Valdepeñas de Jaén	47	19, 20	5	0.00267±0.00062	3	0.8±0.164
Frailes	48	7, 19	2	0.00296±0.00148	2	1.0±0.5
Iznalloz	49	7	7	0	1	0
Sierra de Huétor	50	17, 18	6	0.01205±0.00253	3	0.733±0.155
Villa de Alme	53	10	2	0	1	0

Table S3: Haplotype network sample assignment

Geographic Zone	Haplotype	Individuals
Portugal	1	Z50, Z398, Z395, Z396, Z397, Z51
	2	LB256
	3	LB250
	4	LB208, Z34
	5	Z383, Z384, Z385
	6	Z389
	7	Z390
	8	LB207, Z386, Z387, Z388, Z391, Z394, Z392, Z393, LB210, LB251, LB252
	9	Z399, Z53, Z54
	10	Z52
Huelva	1	Z325, Z326, Z327, Z328, Z373, Z374
	2	Z323
	3	Z315
	4	Z340
	5	Z313, Z314, Z316, Z317, Z318, Z319, Z320, Z321, Z322, Z324, Z331, Z332, Z333, Z334, Z335, Z336, Z337, Z338, Z339
	6	Z412
	7	Z413, Z414, Z310
	8	Z329, Z365, Z366
	9	Z330, Z372
	10	Z309, Z406, Z407, Z408, Z409, Z410, Z411
	11	Z400
	12	Z401
	13	Z403, Z404, Z405
	14	Z346, Z347, Z348, Z349
	15	Z357, Z359
	16	Z370
	17	Z307, Z308, Z343, Z344, Z358, Z360, Z362, Z363, Z64, Z367, Z368, Z369, Z371, Z375, Z377, Z378
	18	Z350, Z351, Z352, Z353, Z354, Z355, Z356
	19	Z345
	20	Z379, Z380, Z381, Z382

South	1	Z311, Z443, Z444, Z446
	2	G76, G77
	3	G65, G66, G67, G68
	4	G78
	5	G69 G71
	6	G70 G72 G73 G74
	7	Z462
	8	Z467
	9	Z512
	10	Z524
	11	Z461
	12	Z464, Z466
	13	Z465
	14	Z463
	15	Z497, Z499, Z204, Z312, Z416, Z417, Z418, Z419, Z480, Z481, Z482, Z483, Z484, Z485, Z490, Z493, Z495, Z498, Z501, Z504, Z505, Z506, Z508, Z510, Z526, Z415
	16	Z445
	17	Z447, Z448, Z449, Z450
	18	Z514
	19	Z525
	20	Z513
	21	Z468
	22	Z479, Z486, Z488
	23	Z487
	24	Z438
	25	Z457, Z459
	26	Z469
	27	Z441
	28	Z440
	29	Z437, Z442
	30	Z473
	31	Z424, Z432, Z436, Z439, Z475, Z518, Z520
	32	Z491, Z492, Z494, Z521, Z522, Z523
	33	Z500
	34	Z502
	35	Z503

	36	Z511
	37	Z528
	38	Z452
	39	Z530
	40	Z453, Z456, Z458, Z460, Z529
	41	Z376
	42	Z420, Z421, Z422, Z423
	43	Z470
	44	Z471, Z472, Z474, Z476, Z477, Z478
	45	Z454, Z455
	46	Z489, Z496
	47	Z426, Z427, Z428, Z429, Z430, Z431, Z433, Z434, Z435, Z519, Z527
East	1	Z575, Z576
	2	Z577
	3	535, Z536, Z537, Z538, Z539
	4	Z572, Z573, Z574
	5	Z532, Z534
	6	Z565
	7	Z551
	8	Z552, Z554
	9	Z553
	10	Z531
	11	Z533
	12	Z555, Z556
	13	Z562
	14	Z563, Z564, Z566
	15	Z47, Z48, Z49, Z540, Z541, Z542, Z543, Z544, Z545, Z546, Z557, Z558, Z559, Z560, Z561, Z567, Z568, Z569, Z570

General Discussion

The results of the present study have greatly contributed to improve our current understanding of the diversity, biogeography and phylogeny of the mygalomorph spiders inhabiting the Mediterranean region and the Canary Islands. Because of its secluded habits and conservative morphology, the group has so far been neglected by taxonomists. Before this study, the original species description – oftentimes short and poorly detailed, and lists of localities were the only scientific information available for some of the target groups. In few cases, extensive information about habits and ecology was available, but the taxonomic status and phylogenetic relationships of the groups were a matter of debate.

An integrative approach, combining sequence data of multiple mitochondrial and nuclear genes, DNA genotyping and species distribution modelling has been used here to address a whole set of different questions concerning the diversity of this taxonomically challenging group and to set the grounds for further research on its ecology and evolution. Each chapter focuses on a specific group and evolutionary question and includes an extensive discussion on its own. In this section, I summarize the general trends driving mygalomorph diversification and discuss future prospects on Mediterranean mygalomorph spider research.

What do we know about the diversity of mygalomorph spiders in the Mediterranean region and the Canary Island?

As already stated, closely related mygalomorph spiders tend to show homogeneous external morphology (Bond *et al.* 2001). For example, until recently only one species of the genus *Ummidia* was recognized in the Mediterranean region (Decae 2010) and the independent status of the ctenizid genera *Cyrtocarenum* and *Cteniza* were a matter of debate (Decae 1996; Raven 1985). However, our reconstruction of the phylogeny of the family Ctenizidae confirmed the independent status of the Mediterranean *Ummidia* species re-erected by Decae (2010), and further revealed two additional lineages pending formal description, which points towards a higher diversity in Mediterranean *Ummidia* that previously suspected (Opatova *et al.* 2013).

Similarly, our findings confirm the sister group relationship between *Cteniza* and *Cyrtocarenum* (discussed in Chapter 1), already recognized by Ausserer (1871) in the original description of *Cyrtocarenum*, but indicate that in spite of their morphological similarity the two genera diverged about 75 Ma (Opatova *et al.* 2013).

The former results prompted us to investigate further the phylogeographic patterns of *Ummidia* in Western Mediterranean (Chapter 2). The extensive sampling conducted in our study, which covered most of its reported distribution range (Decae 2011; Platnick 2014), revealed several deeply divergent lineages, which origins traced back to the Miocene. Although we drove no taxonomic conclusions, since the main focus of our study was the evolutionary history and biogeography of the genus, it is clear that *Ummidia* is in need of a thorough taxonomic revision based on an integrative approach to fully apprehend its real diversity in the Mediterranean region. In the meantime, our results regarding the discovery of a large number of deeply divergent lineages that might correspond to new species are in agreement with on-going research on *Ummidia* in the New World that suggest that the diversity of the genus has also been severely underestimated there (Bond & Coyle 1995).

The preliminary results of an AFLP-based clustering and the divergence time estimates inferred from the nuclear and mitochondrial genes in the population study of *Macrothele calpeiana* revealed a presence of two deeply divergent lineages, which origins trace back to the lower Pliocene. One lineage corresponds to the populations from Portugal and Huelva, while the second one includes the rest of *M. calpeiana* populations. Both lineages may potentially correspond to cryptic species, but it would be necessary to conduct DNA based coalescent delimitation methods to provide further support for this hypothesis (Annex I).

Bayesian multi-species coalescent delimitation methods were used to investigate the existence of cryptic species within *Titanidiops canariensis* (Chapter 4). Our results revealed at least two species with partially sympatric distributions. Several highly divergent lineages were further detected within *T. maroccanus*, which may also represent cryptic species. In all cases, however, we did refrained from formally describing the new species due to the absence of male material.

Overall, our results regarding unaccounted diversity are in agreement with the findings of other authors. A significant number of Mediterranean mygalomorphs species has recently been described (Decae 2010; Decae *et al.* 2006; Decae 2005, Isaia, 2012; Isaia & Decae 2012), including a new genus (Decade & Cardoso 2005), which confirms that the group as a whole had received little taxonomic attention to date. Formal taxonomic treatment of some of our findings has been hampered by the lack of adult males, which are essential for the species descriptions. However, our results set the ground for future taxonomic revisions of Mediterranean mygalomorphs, providing information on the distribution and localities of the putative species based on an extensive sampling. Specific sampling strategies to collect males (e.g. pitfalls) on the know localities may be used in future projects to complete species description.

What factors shaped the distribution patterns and drove the diversification of the mygalomorph spiders in the Mediterranean region and the Canary Islands?

The Mediterranean Basin, including the Macaronesian archipelagos, ranks among the top twenty-five global biodiversity hotspots (Myers *et al.* 2000). The region harbours a high number of local endemics among vascular plants, invertebrates and terrestrial vertebrates (Blondel 2010; Medail & Quezel 1997). The extraordinary species richness and high level of endemism of this area is likely a consequence of its unique location at the crossroad of the Euroasian and African plates, the frequent climatic changes and a dynamic geological history (Myers *et al.* 2000).

Geological processes such as the continental break ups or the rearrangement of tectonic microplates contributed greatly to the present day distribution of the organisms (Altaba 1998; McCarthy 2003; Noonan & Chippindale 2006; Waters *et al.* 2000) and the imprint of these events can be recognized especially in the phylogeographic patterns of lineages with low dispersal ability (Bauzà-Ribot *et al.* 2012; Raven 1978; Stock 1993).

Our results reveal that the past geological events had a great impact on the present day distribution of Mediterranean mygalomorphs. The amphi-

Atlantic distribution of *Ummidia*, once suggested to be the result of a human mediated introduction (Simon 1864, 1910), has been found to trace back to a former Laurasian distribution, as supported by our molecular divergence time estimates. The vicariant origin of the Mediterranean lineages was already anticipated by Decae (2010), but our study (Opatova *et al.* 2013) provided independent support for this hypothesis.

The rearrangement of microplates and, more specifically the Hercynian Belt break up, drove the diversification of two Ctenizidae genera and shaped their present day distribution. Our findings support that the split between the species *Cteniza sauvagesi* and *C. moggridgei* fits the time frame of the initial disintegration of the Hercynian Belt, i.e. the detachment of the Corsica-Sardinia terrain from its previous location and the subsequent opening of the Gulf of Lyon, dated approximately ~30 Ma. Therefore, we hypothesized that the ancestor of the two extant *Cteniza* species already inhabited the region at the time of the break up of the former continuous landmass and that the present day distribution is the result of these vicariant event (discussed in detail in Chapter 1). A similar explanation was been put forward to explain the distribution and diversification of the ground-dwelling spiders of the genus *Parachtes* (Bidegaray-Batista& Arnedo 2011).

Similarly, the divergences among *Ummidia* lineages closely mirror the well-dated geological events involving the disintegration of the other components of the Hercynian Belt. The split of *U. algarve* + Huelva lineages from the rest of the *Ummidia* diversity estimated about 24 Ma, matches the time of the opening of the Valencia Trough, dated to the late Oligocene (25 Ma) (Roca *et al.* 1999). The subsequent divergence of the clade comprising the Northern Africa lineage dated to 16.5 Ma (11.9 – 22.2 Ma) roughly corresponds to the southward rifting of the Kabylies from the Balearic Islands and Betic-Rif blocks, which started in the late Miocene (~21 Ma) (Rehault *et al.* 1984) (for details see Chapter 2).

In general, the split between the Iberian and Northern African lineages of other organisms has been explained by range shifts induced by the Pleistocene climatic oscillations (Paulo *et al.* 2008; Pinho *et al.* 2007; Santos *et al.* 2012), the Messinian Salinity Crisis (5.96 – 5.3Ma) (Agustí *et al.* 2006) or, exceptionally, have been traced back to the lower Messinian (Paulo *et al.* 2008).

In this respect, the *Ummidia* model represents a unique distribution pattern where the deepest divergences trace back to the late Oligocene – early Miocene. However, an event of dispersal during the MSC, presumably from North Africa to Iberia, was also documented by our data, providing an exception to a general vicariant pattern in trap-door spiders.

The extremely disjunct distribution in *Macrothele* was of great biogeographic interest and inspired debates on both the origin of the European species and the timeframe of the colonisation of the Mediterranean region (Arnedo & Ferrández 2007; Ferrández *et al.* 1998; Haupt 2008; Jimenénez-Valverde & Lobo 2007; Van Helsdingen & Decae 1992).

Our findings support that the genus *Macrothele* colonised the Mediterranean region during the Eocene in two independent waves, presumably from Asia, which in turn may have been colonised from Africa by rifting on the Indian subcontinent.

Similar extremely disjunct distribution have only also been reported in the azure-winged magpie *Cyanopica cyanus*, which, similarly to *Macrothele* (Haupt 2008), has been attributed to human mediated introduction. However, a recent study revealed that its present day disjunct distribution is better explained as the result of major extinction throughout the previously continuous range during the Pleistocene climatic oscillations (Fok *et al.* 2002).

We have demonstrated that the Canary Islands were colonised once by the ancestor of *Titanidiops canariensis*. Although the exact time of colonisation remains elusive due to the incongruence among dating methods (i.e. concatenated vs. coalescent approaches), the arrival of *Titanidiops* to the islands most likely preceded the colonisation by other endemic lineages. We further hypothesized that colonisation occurred by rafting from the nearby mainland, as already suggested in he case of the ground dwelling spiders of the genus *Dysdera* (Arnedo *et al.* 2001).

These finding are of great biogeographic interest given the notorious lack of mygalomorph spiders from oceanic islands. The only few exceptions to this observation correspond to groups with documented aerial dispersal by means of ballooning (discussed in detail in Chapter 4).

Insights into the high level phylogeny of mygalomorph spiders

Our study sheds additional light on the high level phylogeny of mygalomorph spiders. The limits of the families Ctenizidae and Hexathelidae were tested in the phylogenetic framework provided by previous studies (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin & Bond 2006).

The monophyly of the family Ctenizidae was not supported in recent molecular phylogenetic studies, although no alternative placement of the component genera received any significant support in the analyses (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin & Bond 2006).

Our study provides the first ever molecular phylogeny of the family that includes representatives of all ctenizid genera. Once again, Ctenizidae monophyly has not been recovered. According to our results, the lack of monophyly is mostly due to, first, the genus *Stasimopus*, already reported in former analyses although we found it to cluster with the remaining ctenizids in some analyses (i.e. ML), and, second, the clade *Cteniza*+*Cyrtocarenum*, which did not clustered with the remaining ctenizids in any analyses. Unfortunately, topology tests could not reject the single origin of the family. On the other hand, our findings provide significant evidence against the traditional subfamily division of the family (discussed in detail in Chapter 1).

Similarly, our results further confirmed the polyphyly of the family Hexathelidae (see Chapter 3) revealed in previous molecular studies (Ayoub *et al.* 2007; Bond *et al.* 2012; Hedin & Bond 2006) and already suspected based on morphology (Goloboff 1993). We demonstrate that hexathelid polyphyly is mostly due to the inclusion of the Atracinae. Although none of the analyses supported the monophyly of Hexathelinae+Macrothelinae, the topology test could not reject a single origin of the two lineages. Interestingly, our results also suggest polyphyly of the genus *Macrothele* as currently defined based on morphological characters.

Our findings emphasize the need for redefining the limits of the families Ctenizidae and Hexathelidae and re-evaluating the morphological characters used in the current taxonomy of both groups. Additionally, it brings to light the inability of household molecular markers used in arthropod higher phylogenies

to recover well-supported relationships within the Mygalomorphae. The advent of the so-called Next Generation Sequencing technologies promises to revolutionize phylogenetic inference in non-model organisms, and may provide a definitive answer to the phylogenetic relationships of the mygalomorph spiders.

Conclusions

Chapter 1: **Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae)**

1. The North American and Mediterranean species of the genus *Ummidia* form reciprocal monophyletic groups that probably originated from a former Laurasian ancestor as a result of the opening of the Atlantic Ocean.

2. The opening of the western Mediterranean basin drove the diversification of the Mediterranean *Ummidia* and the genus *Cteniza*.

3. The split of the extant *Cyrtocarenum* species predates the breakup of the former continuous landmass between Greece and Turkey.

4. Despite the morphological similarity between *Cteniza* and *Cyrtocarenum,* they form reciprocal sister clades with deep divergence times, which supports their status as independent evolutionary lineages.

5. Conversely, the taxonomic status of *Conothele* remains unclear; more material collected throughout its distributional range would be required to evaluate its putative synonymy with the genus *Ummidia*.

6. The current taxonomy of the family Ctenizidae is in need of a major revision. Molecular-based phylogenetic reconstructions disagree with the suprageneric divisions of the family.

Chapter 2: **Loosening the belt: The Hercynian Belt break up shaped the distribution of the trap-door spider genus *Ummidia* (Araneae, Ctenizidae) in the Western Mediterranean**

1. The present day distribution of the genus *Ummidia* in the Western Mediterranean was shaped by the dynamic geological history of the region. The break up of the Hercynian Belt and the subsequent migration

of the individual blocks to their present day position triggered the basal diversification of the group.

2. Two cases of backward colonisation of the Iberian Peninsula from Morocco have been detected, most likely as a result of dispersal events across the Strait of Gibraltar during the Messinian Salinity Crisis.

3. Species Distribution Modelling analyses suggest that *Ummidia algarve* and *Ummidia* sp. Tarifa are ecologically exchangeable.

4. Phylogenetic analyses reveal several independent lineages within the Mediterranean Ummidia that correspond to nominal species and putative new species.

Chapter 3: **From Gondwana to Europe: inferring the origins of Mediterranean *Macrothele* spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae**

1. The Bayesian relaxed clock analysis reveals that the divergence of the genus *Macrothele* dates back to the period of the Gondwana breakup and its present day distribution most likely reflects the subsequent continental drift.

2. The two European species of the genus *Macrothele* originated presumably in Asia and colonised the Mediterranean region independently.

3. The polyphyly of the family Hexathelidae is further corroborated, and the subfamily Atracinae is identified as the main responsible for the non-monophyly.

4. The timeframe for the mygalomorph spiders diversification is set further back into the past, suggesting a more ancient origin of the group and its main evolutionary lineages.

Chapter 4: **Spiders on a hot volcanic roof: Colonisation pathways and phylogeography of the Canary Islands endemic trap-door spider *Titanidiops canariensis* (Araneae, Idiopidae)**

1. *Titanidiops canariensis* colonised the Canary Islands once, presumably by rafting from nearby Northern Africa.

2. The time of colonisation of the archipelago remains unsettled due to the discrepancy observed between the time estimates obtained with the concatenated and the coalescent approaches, but several lines of evidence point towards a Miocene origin.

3. *Titanidiops canariensis* shows a complex phylogeographic pattern, involving either two independent colonisations of Lanzarote or a back and forth colonisation between Fuerteventura and Lanzarote.

4. Volcanic activity may have played a key role in shaping the geographic patterns of genetic diversity in this species. The existence of a volcanic refugium in the Zonzamas area, in western-central Lanzarote, is further supported by our data.

5. The Eastern Canaries are most likely inhabited by two different *Titanidiops* species; one is found in the Jandia Peninsula and extends its distribution to southern Fuerteventura and the second ranges from central Fuerteventura to north Lanzarote. The two species overlap at least in one locality.

6. Several highly divergent lineages were also detected within *T. maroccanus*, which may probably represent additional cryptic species.

7. *Titanidiops maroccanus* may be more closely related to middle-east species, suggesting a secondary colonisation of Northern Africa by the genus.

Annex I: **Threatened or Threatening: Inferring the population structure of Iberian endangered funnel-web spider *Macrothele calpeiana* (Araneae, Hexathelidae)**

1. The results corroborate the existence of deeply divergent lineages within *M. calpeiana*, but also show low levels of genetic differentiation across localities within continuous distribution ranges.

2. Five deeply divergent clades were detected in the concatenated analyses performed on mitochondrial data set, while only two clusters were inferred from the AFLP data. The cluster comprising samples from southern and eastern parts of the Iberian Peninsula shows lower variability, potentially a consequence of recent demographic expansion.

3. Main lineages trace back to Pleistocene, indicating that the population fluctuations and area range shifts during the glaciations may have been responsible for the present-day geographic patterns of genetic diversity.

References

REFERENCES

Agustí J, Garcés M, Krijgsman W (2006) Evidence for African - Iberian exchanges during the Messinian in the Spanish mammalian record. *Palaeogeography, Palaeoclimatology, Palaeoecology* 238, 5-14.

Altaba CR (1998) Testing vicariance: melanopsid snails and Neogene tectonics in the Western Mediterranean. *Journal of Biogeography* 25, 541-551.

Alvarado-Serrano DF, Knowles LL (2014) Ecological niche models in phylogeographic studies: applications, advances and precautions. *Mol Ecol Resour* 14, 233-248.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics* 33, 707-740.

Arnedo MA, Ferrández MA (2007) Mitochondrial markers reveal deep population subdivision in the European protected spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae). *Conservation Genetics* 8, 1147-1162.

Arrigo N, Tuszynski JW, Ehrich D, Gerdes T, Alvarez N (2009) Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics* 10, 33.

Ausserer A (1871) Beiträge zur Kenntniss der Arachniden-Familie der Territelariae Thorell (Mygalidae Autor). *Verhandlungen der Kaiserlich-Koniglichen Zoologisch-Botanischen Gesellschaft in Wien* 21, 117-224.

Ayoub NA, Garb JE, Hedin M, Hayashi CY (2007) Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1g), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). *Molecular Phylogenetics and Evolution* 42, 394-409.

Bauzà-Ribot MM, Juan C, Nardi F*, et al.* (2012) Mitogenomic Phylogenetic Analysis Supports Continental-Scale Vicariance in Subterranean Thalassoid Crustaceans. *Current Biology* 22, 2069-2074.

Bensch S, Akesson M (2005) Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology* 14, 2899-2914.

Bidegaray-Batista L, Gillespie R, Arnedo MA (2011) Bringing spiders to the multilocus era: novel anonymous nuclear markers for Harpactocrates ground-dwelling spiders (Araneae: Dysderidae) with application to to related genera. *Journal of Arachnology* 39, 506-510.

Blondel J (2010) *The Mediterranean region : biological diversity in space and time*, 2nd edn. Oxford University Press, Oxford ; New York.

Blondel J, Aronson J (1999) *Biology and wildlife of the Mediterranean region* Oxford University Press, Oxford; New York.

Bond J (2012) Phylogenetic treatment and taxonomic revision of the trapdoor spider genus Aptostichus Simon (Araneae, Mygalomorphae, Euctenizidae). *ZooKeys* 252, 1-209.

Bond JE, Coyle FA (1995) Observations on the natural history of an Ummidia trapdoor spider from Costa Rica (Araneae, Ctenizidae). *Journal of Arachnology* 23, 157-164.

Bond JE, Hedin M (2006) A total evidence assessment of the phylogeny of North American euctenizine trapdoor spiders (Araneae, Mygalomorphae, Cyrtaucheniidae) using Bayesian inference. *Molecular Phylogenetics and Evolution* 41, 70-85.

Bond JE, Hedin MC, Ramirez MG, Opell BD (2001) Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology* 10, 899-910.

Bond JE, Hendrixson BE, Hamilton CA, Hedin M (2012) A Reconsideration of the Classification of the Spider Infraorder Mygalomorphae (Arachnida: Araneae) Based on Three Nuclear Genes and Morphology. *PLoS ONE* 7, e38753.

Bond JE, Stockman AK (2008) An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology* 57, 628 - 646.

Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* 16, 3737-3758.

Bowen D, Richmond G, Fullerton D*, et al.* (1986) Correlation of Quaternary glaciations in the Northern Hemisphere. *Quaternary Science Reviews* 5, 509-510.

Braga JC, Martín JM, Quesada C (2003) Patterns and average rates of late Neogene–Recent uplift of the Betic Cordillera, SE Spain. *Geomorphology* 50, 3-26.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA* 91, 6491-6495.

Brown RM, Siler CD, Lee Grismer L, Das I, McGuire JA (2012) Phylogeny and cryptic diversification in Southeast Asian flying geckos. *Molecular Phylogenetics and Evolution* 65, 351-361.

Burbrink FT, Yao H, Ingrasci M*, et al.* (2011) Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (*Lampropeltis pyromelana*). *Mol Phylogenet Evol* 60, 445-454.

Campbell D, Duchesne P, Bernatchez L (2003) AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. *Molecular Ecology* 12, 1979-1991.

Carracedo J, Rodríguez-Badiola E (1993) Evolución geológica y magmática de la isla de Lanzarote (Islas Canarias). *Revista de la Academia Canaria de Ciencias* v, 25-58.

Carracedo JC, Day S, Guillou H*, et al.* (1998a) Hotspot volcanism close to a passive continental margin: The Canary Islands. *Geological Magazine* 135, 591-604.

Carracedo JC, Day SJ, Guillou H*, et al.* (1998b) Origen y evolucion del volcanismo de las Islas Canarias. In: *Ciencia y Cultura en Canarias*, p. 165. OACIMC, Santa Cruz de Tenerife.

Carracedo JC, Singer B, Jicha B*, et al.* (2003) La erupción y el tubo volcánico del volcán Corona (Lanzarote, Islas Canarias). *Estudios Geológicos* 59, 277-302.

Chapman J, Carlton J, Bellinger MR, Blakeslee AH (2007) Premature refutation of a human-mediated marine species introduction: the case history of the marine snail Littorina littorea in the Northwestern Atlantic. *Biological Invasions* 9, 995-1008.

Claridge MF, Dawah HA, Wilson MR, Systematics Association. (1997) *Species : the units of biodiversity*, 1st edn. Chapman & Hall, London ; New York.

Coello J, Cantagrel JM, Hernán F*, et al.* (1992) Evolution of the eastern volcanic ridge of the Canary Islands based on new K-Ar data. *Journal of Volcanology and Geothermal Research* 53, 251-274.

Collins NM, Wells S (1987) Invertebrates in need of special protection in Europe (ed. H. A), p. 162. Council of Europe, Strasbourg.

Cooper SJ, Harvey MS, Saint KM, Main BY (2011) Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. *Mol Ecol* 20, 3219-3236.

Coyle EA, Greenstone MH, Hultsch AL, Morgan CE (1985) Ballooning mygalomorphs: Estimates of the masses of *Sphodros* and *Ummidia* ballooners (Araneae: Atypidae, Ctenizidae). *Journal of Arachnology* 13, 291-296.

Coyle FA (1983) Aerial dispersal by mygalomorph spiderlings (Araneae, Mygalomorphae). *Journal of Arachnology*, 283-286.

Coyne JA, Orr HA (2004) *Speciation* Sinauer Associates, Sunderland, Mass.

Cracraft J (1989) Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: *Speciation and its consequences* (eds. Otte D, Endler JA), pp. 28-57. Sinauer Assoc., Sunderland, Mass.

Crawford LA, Koscinski D, Keyghobadi N (2012) A call for more transparent reporting of error rates: the quality of AFLP data in ecological and evolutionary research. *Molecular Ecology* 21, 5911-5917.

Creutzburg N (1963) Paleogeographic evolution of Crete from Miocene till our days. *Cretan Annals* 15, 336-342.

Cunningham CW (2008) How to Use Genetic Data to Distinguish Between Natural and Human-Mediated Introduction of Littorina littorea to North America. *Biological Invasions* 10, 1-6.

Cutler B, Guarisco H (1995) Dispersal aggregation of Sphodros fitchi (Araneae, Atypidae). *Journal of Arachnology*, 205-206.

Dasmahapatra KK, Elias M, Hill RI, Hoffman JI, Mallet J (2009) Mitochondrial DNA barcoding detects some species that are real, and some that are not. *Molecular Ecology Resources* 9999.

de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation. A conceptual unification and terminological recommendations. In: *Endless Forms. Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 57-75. Oxford University Press, New York.

de Queiroz K (2005) Different species problems and their resolution. *BioEssays* 27, 1263-1269.

de Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology* 56, 879 - 886.

Decae A (2010) The genus Ummidia Thorell 1875 in the western Mediterranean, a review (Araneae: Mygalomorphae: Ctenizidae). *Journal of Arachnology* 38, 328-340.

Decae A, Cardoso P (2005) *Iberesia*, a new genus of trapdoor spiders (Araneae, Nemesiidae) from Portugal & Spain. *Revista ibérica de Aracnología*, 3.

Decae A, Cardoso P, Selden P (2006) Taxonomic review of the portuguese Nemesiidae (Araneae, Mygalomorphae). *Revista ibérica de Aracnología*, 1-18.

Decae AE (1996) Systematics of the trapdoor spider genus *Cyrtocarenum* Ausserer, 1871 (Araneae, Ctenizidae). *Bulletin of British arachnological Society* 10, 161-170.

Decae AE (2005) Trapdoor spiders of the genus *Nemesia* Audouin, 1826 on Majorca and Ibiza: taxonomy, distribution and behaviour (Araneae, Mygalomorphae, Nemesiidae). *Bulletin of the British Arachnological Society* 13, 145.

Decae AE (2011) *Diversity and distribution of mygalomorph spiders in the Mediterranean region*, Ghent University.

Degnan JH, Rosenberg NA (2006) Discordance of Species Trees with Their Most Likely Gene Trees. *PLoS Genetics* 2.

Dermitzakis D (1990) Paleogeography, geodynamic processes and event stratigraphy during the Late Cenozoic of the Aegean area. *Atti Convegni Lincei* 85, 263-288.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*.

Duffey E (1956) Aerial dispersal in a known spider population. *Journal of Animal Ecology* 25, 85-111.

Duffey E (1997) Aerial dispersal in spiders.

Eberhard WG (2006) Dispersal by Ummidia spiderlings (Araneae, Ctenizidae): Ancient roots of aerial webs and orientation? *Journal of Arachnology* 34, 254-257.

Edwards DL, Knowles LL (2014) Species detection and individual assignment in species delimitation: can integrative data increase efficacy? *Proc Biol Sci* 281, 20132765.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54, 1839-1854.

Eldredge N (2001) The Sixth Extinction. *American Institute of Biological Sciences, Washington, DC Available from http://www. actionbioscience. org/newfrontiers/eldredge2. html*.

Elith J, Graham CH, Anderson RP*, et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129 -151.

Ereshefsky M (1992) *The units of evolution. Essays on the nature of species* MIT press, Massachusetts.

Esselstyn JA, Evans BJ, Sedlock JL, Anwarali Khan FA, Heaney LR (2012) Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proc Biol Sci* 279, 3678-3686.

Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positvely misleading. *Syst Zool* 27, 401 - 410.

Ferrández MA, Ferrández de Cespedes H (2001) Arachnida. In: *Los Invertebrados no insectos de la "Directiva Hábitat" en España.* (eds. Ramos MA, Bragado D, Fernández J), pp. 133-144. Ministerio de Medio Ambiente. Dirección Genaral de la Conservación de la naturaleza.

Ferrández MA, Ferrández de Cespedes H, Perucho A (1998) *Macrothele calpeiana*, la araña negra de los alcornocales. *Quercus* 146, 14-18.

Ferretti N, Pompozzi G, Copperi Sa, Schwerdt L (2013) Aerial dispersal by *Actinopus* spiderlings (Araneae: Actinopodidae). *The Journal of Arachnology* 41, 407-408.

Fok, K.W., Wade, C.M., and Parkin, D.T. (2002) Inferring the phylogeny of disjunct populations of the azure–winged magpie Cyanopica cyanus from mitochondrial control region sequences. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269 (1501), 1671-1679.

Frizon de Lamotte D, Saint Bezar B, Bracéne R, Mercier E (2000) The two main steps of the Atlas building and geodynamics of the western Mediterranean. *Tectonics* 19, 740-761.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA* 91, 6491-6495.

Gattacceca J, Deino A, Rizzo R*, et al.* (2007) Miocene rotation of Sardinia: New paleomagnetic and geochronological constraints and geodynamic implications. *Earth and Planetary Science Letters* 258, 359-377.

Gibbard PL, Head MJ, Walker MJC (2010) Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science* 25, 96-102.

Goloboff PA (1993) A reanalysis of mygalomorph spider families (Araneae). *American Museum Novitates* 3056, 1–32.

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135, 147-186.

Hamilton CA, Formanowicz DR, Bond JE (2011) Species Delimitation and Phylogeography of *Aphonopelma hentzi* (Araneae, Mygalomorphae, Theraphosidae): Cryptic Diversity in North American Tarantulas. *PLoS ONE* 6, e26207.

Hamilton CA, Hendrixson BE, Brewer MS, Bond JE (2014) An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: a case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Mol Phylogenet Evol* 71, 79-93.

Haupt J (2008) An Organ of Stridulation in East Asian Hexathelid Spiders (Araneae, Mygalomorphae). *Revista ibérica de Aracnología*, 19-23.

Heads M (2011) Old Taxa on Young Islands: A Critique of the Use of Island Age to Date Island-Endemic Clades and Calibrate Phylogenies. *Systematic Biology* 60, 204-218.

Hedin M, Bond JE (2006) Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): Conflict and agreement with the current system of classification. *Molecular Phylogenetics and Evolution* 41, 454-471.

Hedin M, Starrett J, Hayashi C (2013) Crossing the uncrossable: novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. *Mol Ecol* 22, 508-526.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27, 570-580.

Hendrixson BE, DeRussy BM, Hamilton CA, Bond JE (2013) An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, Aphonopelma). *Mol Phylogenet Evol* 66, 327-340.

Herrmann M, Holderegger R, Van Strien MJ (2013) Influence of parameter settings in automated scoring of AFLPs on population genetic analysis. *Mol Ecol Resour* 13, 128-134.

Hewitt G (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58, 247 - 276.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405, 907 - 913.

Hey J (2001) The mind of the species problem. *Trends Ecol Evol* 16, 326-329.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965-1978.

Ibáñez JM, De Angelis S, Díaz-Moreno A*, et al.* (2012) Insights into the 2011–2012 submarine eruption off the coast of El Hierro (Canary Islands, Spain) from statistical analyses of earthquake activity. *Geophysical Journal International* 191, 659-670.

Isaia M, Decae A (2012) Revalidation of *Nemesia meridionalis* Costa, 1835 (Araneae, Mygalomorphae, Nemesiidae), and first description of the male. *Arachnology* 15, 280-284.

Jiménez-Moreno G, Fauquette S, Suc JP (2010) Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Palaeobotany and Palynology* 162, 403-415.

Jimenez-Valverde A, Lobo JM (2007) Potential distribution of the endangered spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae) and the impact of climate warming. *Acta Zoologica Sinica* 53, 865-876.

Jiménez-Valverde A, Lobo JM (2006) Distribution determinants of endangered Iberian spider *Macrothele calpeiana* (Araneae, Hexathelidae). *Environmental Entomology* 35, 1491-1499.

Jolivet L, Augier R, Robin C, Suc J (2006) Lithospheric-scale geodynamic context of the Messinian salinity crisis. *Sedimentary Geology*.

Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: Towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution.* 15, 104-109.

Kekkonen M, Hebert PD (2014) DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Mol Ecol Resour*.

Kodandaramaiah U (2011) Tectonic calibrations in molecular dating. *Current Zoology* 57.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature* 400, 652 - 655.

Kubatko L, Carstens B, Knowles L (2009) STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 25, 971.

Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst Biol* 56, 17-24.

Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). *Proceedings of the Royal Society B: Biological Sciences* 277, 3071-3077.

Leache AD, Harris RB, Rannala B, Yang Z (2014) The influence of gene flow on species tree estimation: a simulation study. *Syst Biol* 63, 17-30.

Leavitt SD, Johnson LA, Goward T, St Clair LL (2011) Species delimitation in taxonomically difficult lichen-forming fungi: an example from morphologically and chemically diverse Xanthoparmelia (Parmeliaceae) in North America. *Mol Phylogenet Evol* 60, 317-332.

Liu L, Pearl DK (2007) Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst Biol* 56, 504-514.

Lonergan L, White N (1997) Origin of the Betic-Rif mountain belt. *Tectonics* 16, 504-522.

Macías-Hernández N, Oromí P, Arnedo MA (2010) Integrative taxonomy uncovers hidden species diversity in woodlouse hunter spiders (Araneae, Dysderidae) endemic to the Macaronesian archipelagos. *Systematics and Biodiversity* 8, 531 - 553.

Maddison W (1997) Gene trees in species trees. *Systematic Biology* 46, 523.

Maddison WP, Knowles LL (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55, 21–30.

Mallet J (2001) The speciation revolution. *Journal of Evolutionary Biology* 14, 887-888.

Mariette Sp, Chagné D, Lézier Cl*, et al.* (2001) Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity* 86, 469-479.

Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In: *Species: The units of biodiversity.* (eds. Claridge MF, Dawah HA, Wilson MR), pp. 381-424. Chapman and Hall, London.

Mayden RL, Wood RM (1994) Systematics, Species Concepts, and the Evolutionarily Significant Unit in Biodiversity and Conservation Biology, 58-113.

Mayr E (1982) *The growth of biological thought: Diversity, evolution, and inheritance* Harvard University Press, Cambridge.

McCarthy D (2003) The trans-Pacific zipper effect: disjunct sister taxa and matching geological outlines that link the Pacific margins. *Journal of Biogeography* 30, 1545-1561.

Medail F, Quezel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Annals of the Missouri Botanical Garden* 84, 112-127.

Meudt HM, Clarke AC (2007) Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science* 12, 106-117.

Mueller U, Wolfenbarger L (1999) AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* 14, 389 - 394.

Myers N, Mittermeier RA, Mittermeier CC, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403, 853-858.

Noonan BP (2000) Does the phylogeny of pelomedusoid turtles reflect vicariance due to continental drift? *Journal of Biogeography* 27, 1245-1249.

Noonan BP, Chippindale PT (2006) Vicariant origin of malagasy reptiles supports late cretaceous antarctic land bridge. *Am Nat* 168, 730-741.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13, 1143-1155.

Opatova V, Bond JE, Arnedo MA (2013) Ancient origins of the Mediterranean trap-door spiders of the family Ctenizidae (Araneae, Mygalomorphae). *Mol Phylogenet Evol* 69, 1135-1145.

Padial J, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. *Frontiers in Zoology* 7, 16.

Paez-Moscoso DJ, Guayasamin JM, Yanez-Munoz M (2011) A new species of Andean toad (Bufonidae, Osornophryne) discovered using molecular and morphological data, with a taxonomic key for the genus. *ZooKeys*, 73-97.

Paulo OS, Pinheiro J, Miraldo A*, et al.* (2008) The role of vicariance vs. dispersal in shaping genetic patterns in ocellated lizard species in the western Mediterranean. *Mol Ecol* 17, 1535-1551.

Pedersen AA, Loeschcke V (2001) Conservation genetics of peripheral populations of the mygalomorph spider *Atypus affinis* (Atypidae) in northern Europe. *Molecular Ecology* 10, 1133-1142.

Penney D, Selden PA (2011) *Fossil Spiders: The Evolutionary History of a Mega-diverse Order* Siri Scientific Press, Manchester, UK.

Peterson AT (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics* 3, 59-72.

Pinho C, Harris DJ, Ferrand N (2007) Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. *Molecular Ecology* 16, 1191-1205.

Platnick NI (2014) The World Spider Catalog. American Museum of Natural History, online at http://research.amnh.org/entomology/spiders/catalog/index.html, [accessed 20 March 2014].

Platnick NI, Gertsch WJ (1976) The suborder of spiders: A cladistic analysis. *American Museum Novitates* 2607, 1-15.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* 6, 847-859.

Pons J, Barraclough TG, Gomez-Zurita J*, et al.* (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology* 55, 595-609.

Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol Ecol* 21, 1864-1877.

Ratnasingham S, Hebert PD (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE* 8, e66213.

Raven R (1978) Systematics of spider subfamily Hexathelinae (Dipluridae-Mygalomorphae-Arachnida). *Australian Journal of Zoology*, 1-75.

Raven RJ (1980) The evolution and biogeography of the mygalomorph spider family Hexathelidae (Araneae, Chelicerata). *Journal of Arachnology* 8, 251-266.

Raven RJ (1985) The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. *Bull. Amer. Mus. Nat. Hist.* 182, 1-180.

Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG (2007) Applications of Ecological Niche Modeling for Species Delimitation: A Review and Empirical Evaluation Using Day Geckos (Phelsuma) from Madagascar. *Syst Biol* 56, 907-923.

Rehault JP, Boillott G, Mauffret A (1984) The western Mediterranean basin geological evolution. *Marine Geology* 55, 447–477.

Ridley M (2004) *Evolution*, 3rd edn. Blackwell Pub., Malden, MA.

Rissler L, Apodaca J (2007) Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the Black Salamander (*Aneides flavipunctatus*). *Syst Biol* 56, 924 - 942.

Roca E, Sans M, Cabrera L, Marzo M (1999) Oligocene to Middle Miocene evolution of the central Catalan margin (northwestern Mediterranean). *Tectonophysics* 315, 209-229.

Rosenbaum G, Lister GS, Duboz C (2002) Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. In: *Reconstruction of the evolution of the Alpine-Himalayan Orogen* (eds. Rosenbaum G, Lister GS), pp. 107-126. Journal of the Virtual Explorer.

Rowe DL, Dunn KA, Adkins RM, Honeycutt RL (2010) Molecular clocks keep dispersal hypotheses afloat: evidence for trans-Atlantic rafting by rodents. *Journal of Biogeography* 37, 305-324.

Sanderson MJ (2003) r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19, 301-302.

Santos X, Rato C, Carranza S, Carretero MA, Pleguezuelos JM (2012) Complex phylogeography in the Southern Smooth Snake (*Coronella girondica*) supported by mtDNA sequences

Satler JD, Starrett J, Hayashi CY, Hedin M (2011) Inferring Species Trees from Gene Trees in a Radiation of California Trapdoor Spiders (Araneae, Antrodiaetidae, *Aliatypus*). *PLoS ONE* 6, e25355.

Schellart W (2002) Alpine deformation at the western termination of the Axial Zone, Southern Pyrenees. *Journal of the Virtual Explorer* 8, 35-55.

Séranne M (1999) The Gulf of Lion continental margin (NW Mediterranean) revisited by IBS: an overview. *Special Publication-Geological Society of London* 156, 15-36.

Shaffer H, Thomson R (2007) Delimiting species in recent radiations. *Syst Biol* 56, 896 - 906.

Simon E (1864) *Histoire naturelle des Araignées (Aranéides)* Roret edit., Paris.

Simon E (1910) Catalogue raisonné des Arachnides du Nord de l'Afrique, 1ere partie. *Annales de la Société Entomologique de France* 79, 265-332.

Sistrom M, Donnellan SC, Hutchinson MN (2013) Delimiting species in recent radiations with low levels of morphological divergence: a case study in Australian Gehyra geckos. *Mol Phylogenet Evol* 68, 135-143.

Snazell RG (1986) The spider genus *Macrothele* Ausserer in Spain (Araneae, Dipluridae). *Bulletin of the British Arachnological Society* 17, 80-83.

Snazell RG, Allison R (1989) The genus Macrothele Ausserer (Araneae; Hexathelidae) in Europe. *Bulletin of the British Arachnological Society* 8, 65-72.

Soberon J, Peterson AT (2005) *Interpretation of Models of Fundamental Ecological Niches and Species, Distributional Areas*.

Speranza F, Villa IM, Sagnotti L*, et al.* (2002) Age of the Corsica-Sardinia rotation and Liguro-Provençal Basin spreading: new paleomagnetic and Ar/Ar evidence. *Tectonophysics* 347, 231-251.

Stock JH (1993) Some remarkable distribution patterns in stygobiont Amphipoda. *Journal of Natural History* 27, 807-819.

Stockman A, Bond JE (2007) Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. *Molecular Ecology* 16, 3374-3392.

Suc JP (1984) Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* 307, 429-438.

Talavera G, Dinca V, Vila R (2013) Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution* 4, 1101-1110.

Van Helsdingen PJ (1993) Can *Macrothele calpeiana* (Walckenaer) (Araneae, Hexathelidae) be used as a bio - indicator? *Bulletin de la Société neuchâteloise des sciences naturelles* 116, 253-258.

Van Helsdingen PJ, Decae A (1992) Ecology, distribution and vulnerability of *Macrothele calpeiana* (Walckenaer) (Araneae, Hexathelidae). *Tljdschrift voor Entomologie* 135, 169-178.

Vos P, Hogers R, Bleeker M*, et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23, 4407 - 4414.

Waters JM, Lopez JA, Wallis GP (2000) Molecular phylogenetics and biogeography of galaxiid fishes (Osteichthyes: Galaxiidae): dispersal, vicariance, and the position of *Lepidogalaxias salamandroides*. *Syst Biol* 49, 777-795.

Whitlock R, Hipperson H, Mannarelli M, Butlin R, Burke T (2008) An objective, rapid and reproducible method for scoring AFLP peak height data that minimizes genotyping error. *Molecular Ecology Resources* 8, 725-735.

Wiens JJ, Penkrot TA (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Systematic Biology* 51, 69-91.

Wiley EO (1981) *Phylogenetics. The theory and practice of phylogenetic systematic* John Wiley & sons, New York.

Wilson EO (1992) *The diversity of Life* The Belknap Press of Harvard university press, Cambridge, Mass., USA.

Wunderlich J (1992) Die Spinnen-Fauna der Makaronesischen Inseln: Taxonomie, Ökologie, Biogeographie und Evolution. *Beitr. Araneol*, 1-619.

Wunderlich J (1995) Zu Ökologie, Biogeographie, Evolution und Taxonomie einiger Spinnen der Makaronesischen Inseln (Arachnida: Araneae). *Beitr. Araneol*, 385-439.

Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences* 107, 9264-9269.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science* 292, 686-693.

Zhang C, Zhang D-X, Zhu T, Yang Z (2011) Evaluation of a Bayesian Coalescent Method of Species Delimitation. *Systematic Biology* 60, 747-761.

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: *Evolving Genes and Proteins* (eds. Bryson V, Vogel HJ), pp. 97–166. Academic Press.

Resumen en castellano

INTRODUCCIÓN GENERAL

Descubriendo la diversidad de grupos morfológicamente uniformes y con patrones biogeográficos complejos

La biodiversidad de la Tierra está amenazada por la acción del hombre (Eldredge 2001). Se estima que alrededor de 30.000 especies se extinguen cada año, una proporción significativa de las cuales ni siquiera ha sido catalogada por la ciencia (Wilson 1992). Por lo tanto, el diseño de planes eficientes de conservación de la biodiversidad basados en protocolos de evaluación rápida es un requerimiento para la detección y protección de aquellas áreas ricas en diversidad de especies y con características ecológicas únicas (Myers *et al*. 2000).

La cuenca mediterránea es uno de los veinticinco puntos calientes de la biodiversidad mundial que merecen una atención especial de conservación (Myers *et al*. 2000). Tanto la cuenca mediterránea como las Islas Canarias presentan un nivel elevado de especies endémicas (Blondel 2010; Juan *et al*. 2000; Medail & Quezel 1997).

La capacidad de reconocer especies, un paso fundamental para la comprensión de la biodiversidad, puede ser particularmente difícil en grupos morfológicamente uniformes. El infraorden Mygalomorphae es uno de los tres linajes evolutivos principales reconocidos dentro de las arañas (Hedin & Bond 2006; Platnick & Gertsch 1976). Las relaciones filogenéticas a nivel de familia han sido objeto de estudios recientes (Ayoub *et al* 2007; Bond *et al* 2012; Hedin & Bond 2006), pero su diversidad permanece en general poco estudiada.

Hasta hace pocos años, se reconocía una sola especie del género *Ummidia* (fam. Ctenizidae) en el Mediterráneo, localizada en el sur de la Península Ibérica y el norte de África (Decae 2010). Del mismo modo, la familia Cyrtaucheniidae, está representada por una solo especie, *Cyrtauchenius walckenaeri*, a lo largo de todo el Mediterráneo occidental y *Atypus affinis*, de la familia Atypidae, ocurre desde el sur de la Península Ibérica hasta zonas del sur del Reino Unido y Suecia (Platnick 2014). Teniendo en cuenta la limitada capacidad de dispersión de las arañas migalomorfas, parece probable que esas especies pueda en realidad esconder una mayor diversidad que habría

pasado desapercibida debido a su gran uniformidad morfológica. En ese sentido, existen estudios que ponen de manifiesto la existencia de una elevada estructuración genética y endemismos locales en arañas migalomorfas (Bond *et al*. 2001).

Los factores que contribuyen al escaso conocimiento taxonómico del grupo en el Mediterráneo se relacionan con las dificultades en la recolección de material y la homogeneidad en la morfología externa de taxones relacionados. La mayoría de los grupos mediterráneos construyen nido subterráneos que son difíciles de detectar. Mediante captura directa se recolectan sobre todo hembras y juveniles, porque los machos adultos abandonan el nido en búsqueda de hembras. Por lo general, los machos se colectan mediante trapas de caída durante la fase de dispersión, que se limita a un período corto del año. Las distintas estrategias de recolección resultan a menudo en la descripción de especies sobre la base de uno solo de los sexos. Por ejemplo, en el género *Nemesia* (fam. Nemesiidae), el más diverso de migalomorfos en el Mediterráneo, aproximadamente la mitad de las especies están descritas por uno solo sexo (Platnick 2014).

Por otro lado, los hábitos sedentarios de los migalomorfos les convierten en un perfecto sistema modelo para estudios de biogeografía, debido a que los procesos geológicos, tales como la deriva continental, suele reflejarse en los organismos con baja dispersión o preferencias por hábitats específicos (Bauza-Ribot *et al* 2012; Hedin *et al* 2013; Stock 1993).

Los patrones de distribución de algunos de los géneros Mediterráneos han sido objeto de discusión acerca del origen de los mismos. Por ejemplo, el género *Ummidia* se encuentra distribuido a ambos márgenes del atlántico, aunque es más diverso en el Nuevo Mundo, lo que podría explicarse tanto por una distribución ancestral en Laurasia y su posterior fragmentación o una introducción humana reciente en el Mediterráneo (Decae 2010). Por otro lado, el posible origen africano o asiático de la especies europeas de *Macrothele* (Fam. Hexathelidae) también sido objeto de debate (Arnedo & Ferrández 2007; Ferrández *et al* 1998; Haupt 2008; Jiménez-Valverde & Lobo 2007; Van Helsdingen & Decae 1992), y la presencia de *Titanidiops canariensis* (fam. Idiopidae) en las Islas Canarias orientales presenta un reto a la visión tradicional vicariante de la de la distribución de los migalomorfos.

En esta tesis aporta nueva información sobre la diversidad críptica, los patrones de distribución y las relaciones filogenéticas de varias familias de migalomorfas que habitan el Mediterráneo Occidental y las Islas Canarias. Estos aspectos se investigan mediante el uso de marcadores moleculares múltiples y nuevos métodos de análisis filogenéticos. Así mismo, se infiere el marco temporal de la diversificación de los grupos de interés para contrastar los diferentes escenarios biogeográficos sobre su origen e identificar los principales factores que promovieron la diversificación de los migalomorfos en la cuenca mediterránea.

Por primera vez en arañas migalomorfas se aplican métodos bayesianos coalescentes para la delimitación de especies a partir de datos moleculares. En concreto se investigan los límites de la especie *Titanidiops canariensis*, que puede constituir un ejemplo para la delimitación de especies en grupos morfológicamente uniformes.

OBJETIVOS:

El principal objetivo de esta tesis es estudiar los patrones filogenéticos y filogeográficos de las arañas migalomorfas del Mediterráneo Occidental y las Islas Canarias, mediante métodos basados en secuencias de ADN y genotipados nucleares (i.e. AFLP). Se incorpora la estimación del tiempo de divergencia y las preferencias del nicho ecológico para identificar los factores que promovieron la diversificación de los grupos seleccionados y poder contrastar los diferentes escenarios biogeográficos responsables de su distribución actual.

Los objetivos específicos del trabajo son:

1. Proporcionar nuevos conocimientos sobre la filogenia de la familia Ctenizidae e inferir el marco temporal para su diversificación con el objetivo de interpretar los orígenes de los taxones Mediterráneos y aportar información nueva que permita la evaluación de su estatus taxonómico.

2. Interpretar los patrones de biogeográficos del género *Ummidia* en el contexto de la historia geológica dinámica del Mediterráneo Occidental; proporcionar un marco temporal para su diversificación y comparar las preferencias ecológicas de las tres especies nominales de *Ummidia* ibéricas con el fin de evaluar su capacidad para el intercambio ecológico.

3. Desentrañar los orígenes del género *Macrothele* y sus vías de colonización en el Mediterráneo e investigar los límites de la familia Hexathelidae en el contexto de un filogenia con una amplia representación de arañas migalomorfas proporcionada por estudios anteriores.

4. Estudiar los orígenes filogenéticos de *Titanidiops canariensis*, uno de los pocos ejemplos de arañas migalomorfas endémicas de un archipiélago oceánico, e inferir el marco temporal de la colonización de las islas. Evaluar la posible existencia de especies crípticas dentro de *T. canariensis* mediante métodos de delimitación de especies bayesiana basados en datos moleculares

DISCUSIÓN

Los resultados de esta tesis contribuyen a mejorar notablemente el conocimiento actual sobre la diversidad críptica, la biogeografía y la filogenia de arañas migalomorfas de la cuenca mediterránea e Islas Canarias. El grupo ha sido poco estudiado por los taxónomos debido a los hábitos crípticos de la mayoría de las especies y su notable uniformidad morfológica. Los conocimientos científicos previos disponibles para la mayoría de los grupos estudiados, se limitaban a la descripción original de la especie, a menudo escueta y poco informativa y a localidades de colecta. En algunos casos existían conocimientos adicionales sobre hábitos y ecología, pero el estatus taxonómico y la relaciones filogenéticas de estos taxones era objeto de debate.

En esta tesis se ha implementado una metodología integrativa que combina secuencias de DNA de múltiples genes, mitocondriales y nucleares, el genotipado (AFLPs) y la modelación de distribución de especies, para abordar una serie de cuestiones relacionadas con la diversidad de este grupo y para sentar las bases del conocimiento necesarias para futuros estudios en estas interesantes arañas. Cada capítulo representa un estudio individual que incluye su propia discusión detallada. Sin embargo se ha considerado pertinente discutir de forma global las tendencias evolutivas en la diversificación de los migalomorfos, así como las perspectivas sobre investigaciones futuras sobre arañas migalomorfas mediterráneas.

¿Qué sabemos sobre de la diversidad de arañas migalomorfas en la cuenca mediterránea y las Islas Canarias?

Como se ha mencionado anteriormente, las arañas migalomorfas suelen tener una morfología externa homogénea en taxones relacionados (Bond *et al.* 2001). Hasta hace poco, se reconocía una sola especie del género de ctenizidos *Ummidia* en la región Mediterránea (Decae 2010) y se ponía en duda el estatus independiente de los géneros *Cyrtocarenum* y *Cteniza,* en la misma familia (Decae, 1996; Raven, 1985).

Nuestros resultados confirman la existencia de varias especies de *Ummidia* mediterráneas, incluyendo las recientemente restablecidas por Decae (Decae 2010) y dos especies adicionales desconocidas (Opatova *et al*. 2013). Por otra parte, aunque se confirma la estrecha relación entre los grupos hermanos *Cteniza* y *Cyrtocarenum* (discutido en el Capítulo 1), reconocida ya por Ausserer (1871) y a pesar de su similitud morfológica, nuestros resultados revelan que los dos géneros tienen un origen antiguo, que se remonta probablemente a 75 Ma (Opatova *et al*. 2013), lo que apoya su estatus como géneros distintos.

Estos resultados abrieron el camino para el proyecto posterior dedicado a desentrañar los patrones filogeográficos del género *Ummidia* en el Mediterráneo Occidental (Capítulo 2). El muestreo realizado para este estudio abarcó la mayor parte de la distribución conocida de *Ummidia* (Decae 2011; Platnick 2014) y reveló la existencia de varios linajes divergentes, cuyos

orígenes se remontan al Mioceno. Aunque el objetivo del estudio fue la biogeografía de *Ummidia* y, por tanto, no se han realizado descripción formal de los linajes detectados, resulta obvio que es necesario realizar una revisión taxonómica completa del genero utilizando un enfoque integrativo. Además, el descubrimiento del elevado número de linajes divergentes que pueden corresponder a nuevas especies, concuerda con las observaciones preliminares disponibles de *Ummidia* en el Nuevo Mundo, donde la diversidad de género parece también haber sido subestimada (Bond & Coyle, 1995).

Por otra parte, los resultados preliminares de agrupamiento mediante genotipado por AFLPs, así como los tiempos de divergencia estimados a partir de la matriz concatenada de los genes nucleares y mitocondriales en el estudio poblacional de *Macrothele calpeiana* han revelado la existencia de dos linajes divergentes con orígenes en el Plioceno inferior, candidatos a ser especies distintas. Uno de los linajes se distribuye en Portugal y Huelva, mientras que el segundo habita el resto de la distribución de *M. calpeiana*. La confirmación de que ambos linajes constituyen de hecho especies crípticas distintas, sin embargo, deberá esperar a la aplicación de métodos coalescentes de delimitación de especies basados en ADN (Anexo I).

Se ha investigado la existencia de especies crípticas dentro de *Titanidiops canariensis* por medio de métodos bayesianos de delimitación de especies (Capítulo 4). Los resultados indican la existencia de al menos dos especies con una distribución parcialmente simpátrica. También se han detectado varios linajes divergentes dentro de *T. maroccanus*, que pueden representar especies crípticas. Sin embargo la ausencia de machos adultos para comparar ha hecho desestimar en ambos casos la descripción formal de las nuevas especies.

En general, los resultados de este trabajo corroboran los resultados de otros autores, ya que recientemente se han descrito un número importante de migalomorfos Mediterráneos (Decae 2010; Decae *et al* 2006; Decae 2005; Isaia & Decae 2012), incluso un nuevo género (Decae y Cardoso 2005). Como se ha mencionado anteriormente, una limitación importante en los estudios realizados en esta tesis ha sido la falta de machos adultos, imprescindibles para la descripción de nuevas especies. Sin embargo, gracias a la información

detallada sobre la distribución de los diferentes linajes proporcionada por este estudio, será posible plantear en el futuro cercano colectas específicas destinadas a la obtención de machos de los linajes de interés mediante el uso de las trampas de caída.

Los factores responsables por los patrones de distribución y la diversificación de las arañas migalomorfas en la región del Mediterráneo y las Islas Canarias

La cuenca mediterránea es uno de los veinticinco puntos calientes de biodiversidad mundial (Myers *et al* 2000) y alberga, junto a las Islas Canarias, un elevado número de especies endémicas, incluyendo plantas vasculares, invertebrados y vertebrados terrestres (Blondel 2010; Medail & Quézel 1997). La extraordinaria riqueza de especies y el alto nivel de endemismo de la región, es probablemente el resultado de su localización en el punto de conexión ente las placas Euroasiática y Africana y a su compleja historia climática y geológica (Myers *et al*. 2000).

Los procesos geológicos como la deriva continental y la formación de microplacas ha condicionado la distribución actual de muchos grupos (Altaba 1998; McCarthy 2003; Noonan y Chippindale 2006). La influencia de estos eventos es especialmente evidente en linajes con baja capacidad de dispersión (Waters *et al* 2000; Bauzà-Ribot *et al* 2012; Cuervo 1978; Stock 1993).

Nuestros resultados ponen en evidencia que los eventos geológicos pasados tuvieron un gran impacto en la distribución actual de los migalomorfos Mediterráneos. La distribución a ambos lados del océano Atlántico de *Ummidia*, que había sido explicada como una introducción humana (Simon 1864, 1910), es en realidad consecuencia de un distribución ancestral en Laurasia. Esta hipótesis está apoyada por nuestras estimas del tiempo de divergencia molecular, que datan de la divergencia entre el clado Mediterráneo y el del Nuevo Mundo en el marco temporal correspondiente a la desintegración de Laurasia. Esta hipótesis había sido ya adelantada por Decae (Decae 2010), pero hasta el uso de métodos de datación molecular no ha sido posible contrastarla (Opatova *et al.* 2013).

Los movimientos de las placas tectónicas como consecuencia de la desintegración del denominado Cinturón Herciniano, promovieron la

diversificación de dos géneros de la familia Ctenizidae y contribuyeron en gran medida a perfilar su distribución actual. Según nuestros resultados, la divergencia entre *Cteniza sauvagesi* y *C. moggridgei* corresponde a la primera fase de la desintegración del Cinturón Herciniano correspondiente a la formación del Golfo de Lyon (~30 Ma), cuando un bloque que contenía la futuras Córcega y Cerdeña se separó y empezó a alejarse de su posición original. Por lo tanto, asumiendo que el ancestro de las dos especies de *Cteniza* existentes habitaba la región en el momento de la fractura de la masa de tierra continua, su distribución actual sería el resultado de un evento de vicariancia (discutido en detalle en el Capítulo 1), tal y como se ha descrito también para la diversificación de arañas de género *Parachtes* (Bidegaray-Batista & Arnedo 2011).

Del mismo modo, las divergencias entre los linajes principales de *Ummidia* se corresponden bien con los eventos geológicos bien datados que dieron lugar a la desintegración del Cinturón Herciniano, incluyendo la formación de Golfo de Valencia, datado a finales del Oligoceno alrededor de 25 Ma (Roca *et al*. 1999), que encajaría con la separación del linaje *U. algarve* + Huelva del resto de la *Ummidia* del Mediterráneo, estimado en 24 Ma. La siguiente divergencia del clado circunscrito mayormente al norte de África, se ha datado en 16.5 Ma (11,9-22,2 Ma) y corresponde aproximadamente a la separación de la placa de la Kabilia del bloque Bético-Rifeño, que se inició a finales del Mioceno, have unos 21 Ma (Rehault *et al*. 1984) (para más detalles véase el Capítulo 2).

En grupos con distribución similar a *Ummidia*, la presencia de especies o poblaciones a ambos lados del estrecho suele explicarse por el desplazamiento de las especies durante las oscilaciones climáticas del Pleistoceno (Paulo *et al* 2008; Pinho *et al* 2007; Santos *et al* 2012), la Crisis de Salinidad del Messiniense (CSM, 5.96 - 5.3 Ma) (Agustí *et al* 2006), o excepcionalmente, se obtienen dataciones anteriores al Messiniense (Paulo *et al* 2008). En este aspecto *Ummidia* presenta un patrón de distribución único, donde las divergencias más profundas se remontan a finales del Oligoceno – principios del Mioceno. Sin embargo, también se ha observado posibles eventos de dispersión a través de conexiones terrestres establecidas durante la CSM

presumiblemente desde el norte de África hacia Iberia, lo que supone una excepción a un patrón de diversificación general en *Ummidia* por vicariancia.

La distribución disjunta del género *Macrothele* ha sido objeto de un gran interés biogeográfico, inspirando debates tanto sobre el origen de las especies europeas, como del marco temporal de la colonización de la región mediterránea (Arnedo & Ferrández 2007; Ferrández *et al.* 1998; Haupt 2008; Jimenénez-Valverde & Lobo 2007; Van Helsdingen & Decae 1992).

Nuetrso resultados indican que el género *Macrothele* llegó a la cuenca mediterránea en el Eoceno, mediante dos eventos de colonización independientes, presumiblemente desde Asia, que a su vez fue probablemente colonizada desde África, a través de la India. El rabilargo *Cyanopica cyanus* muestra un patrón de distribución disjunta Europa-Asia similar a *Macrothele*. Tradicionalmente se consideraba que dicha distribución era el resultado una introducción humana reciente, de forma similar a lo que había sido propuesto también por algunos autores para *Macrothele* (Haupt 2008). Sin embargo, un estudio reciente reveló que la distribución discontinua presente probablemente es el resultado de extinciones a gran escala durante las oscilaciones climáticas del Pleistoceno de un área anteriormente habitada por la especie (Fok et al. 2002).

Titanidiops canariensis colonizó las Islas Canarias una sola vez desde el continente, probablemente utilizando dispersión pasiva mediante islas flotantes (“rafting”), tal y como se ha señalado en el caso las arañas del género *Dysdera* (Arnedo *et al*. 2001). Estos resultados son de un gran interés biogeográfico ya que la fauna de las islas oceánicas suelen caracterizarse por la ausencia de migalomorfos, salvo pocas excepciones de especies con dispersión aérea mediante “ballooning” (discutido en detalle en Capítulo 4).

Aportación al conocimiento de la filogenia de alto nivel de las arañas migalomorfas

Esta tesis también ha contribuido a un mejor conocimientos de las relaciones filogenéticas y los limites de las familias de arañas migalomorfas. Los límites de las familias Ctenizidae y Hexathelidae han sido investigados en

un marco filogenético proporcionados por estudios anteriores (Ayoub *et al* 2007; Bond *et al* 2012; Hedin & Bond 2006).

La monofilia de la familia Ctenizidae no había sido recuperada en los últimos filogenéticos moleculares del infraorden Mygalomorphae, pero las posiciones alternativas en la filogenia de algunos géneros tampoco recibió un apoyo sustancial (Ayoub *et al* 2007; Bond *et al* 2012; Hedin & Bond 2006). La monofilia de la familia Ctenizidae solo se recuperó en análisis combinados de caracteres morfológicos y datos moleculares (Bond *et al*. 2012).

En esta tesis se ha llevado a cabo por primera vez una filogenia molecular que incluía representantes de todos los géneros de ctenizidos. Nuestros resultados confirman la polifilia de Ctenizidae debido, tal y como se había observado en los estudios anteriores, a la posición del género *Stasimopus*, que se agrupaba con el resto de los ctenizidos sólo en el análisis de ML, pero también a la posición del linaje *Cteniza* + *Cyrtocarenum*, que nunca formaba un clado monofilético con resto de los ctenizidos, lo cual constituye una novedad. Sin embargo, el test de topología no pudo rechazar el origen único de la familia. En cualquier caso, los resultados de los análisis filogenéticos cuestionan la división taxonómica tradicional de la familia Ctenizidae (discutido en detalle en el Capítulo 1).

Los resultados de esta tesis confirman la polifilia de la familia Hexathelidae (véase el Capítulo 3), también detectada en estudios moleculares anteriores (Ayoub *et al* 2007; Bond *et al* 2012; Hedin & Bond 2006) y ya sospechada en base a caracteres morfológicos (Goloboff 1993). Según los resultados, los taxones de la subfamilia Atracinae son los principales responsables de la polifilia de los hexathelidos. Ninguno de los análisis apoya la monofilia de Hexathelinae + Macrothelinae, pero el test de topología no pudo rechazar un origen único de ambos linajes. Curiosamente, nuestros resultados también sugieren la polifilia del género *Macrothele*, tal y como se define actualmente en base a caracteres morfológicos.

Estos resultados ponen de manifiesto la necesidad de reevaluar los caracteres morfológicos utilizados en la taxonomía actual de las familias Ctenizidae y Hexathelidae. Así mismo, el bajo poder de resolución de los marcadores moleculares clásicamente utilizados para la inferencia de filogenias

de alto nivel de artrópodos sugiere que el establecimiento de hipótesis filogenéticas robustas entre familias de migalomorfos tendrá que esperar al desarrollo de nuevos loci nucleares. En este sentido, el desarrollo de las nuevas metodologías de secuenciación masiva prometen revolucionar el estudio de la filogenia de organismos no modelos, y constituye la principal fuente de esperanza para la resolución de las relaciones filognéticas de alto nivel en migalomorfos.

CONCLUSIONES

Capítulo 1: **Orígenes antiguos de las arañas migalomorfas Mediterráneas de la familia Ctenizidae (Araneae, Mygalomorphae)**

1. Las especies Mediterráneas y Norteamericanas del género *Ummidia* forman grupos monofiléticos recíprocos que probablemente se originaron a partir de un ancestro común que habitó Laurasia antes de la formación del Océano Atlántico.

2. La formación del la cuenca mediterránea occidental probablemente catalizó la diversificación de los representantes Mediterráneos de los géneros *Ummidia* y *Cteniza.*

3. La especiación del género *Cyrtocarenum* es anterior a la desintegración de la antigua masa de tierra continua entre Grecia y Turquía.

4. A pesar de la similitud morfológica de *Cteniza* y *Cyrtocarenum* ambos forman clados hermanos exclusivos con tiempo de divergencia profunda, observación que apoya su condición de linajes evolutivos independientes.

5. La posición de *Conothele* aún no está resuelta; la obtención de material adicional a lo largo de su rango de distribución será necesario para la correcta evaluación de su posible sinonimia con *Ummidia*.

6. Es necesaria una re-evaluación de los caracteres morfológicos utilizados en la taxonomía actual de la familia Ctenizidae. Las relaciones filogenéticas están en contradicción con algunas agrupaciones taxonómicas de alto nivel.

Capítulo 2: **Aflojarse el cinturón: La desintegración del Cinturón Herciniano modeló la diversidad y distribución de las arañas migalomorfas del género *Ummidia* (Araneae, Ctenizidae) en el Mediterráneo Occidental**

1. La distribución actual del género *Ummidia* en el Mediterráneo Occidental es resultado de la dinámica historia geológica de la región. La desintegración del Cinturón Herciniano y la posterior migración de las microplacas resultantes se identifica como el motor principal de vicarianza responsable de la diversificación de los principales linajes de *Ummidia*.

2. Se han detectado dos casos de colonización de la Península Ibérica desde el Norte de África, en ambos casos los eventos de dispersión a través del Estrecho de Gibraltar ocurrieron presumiblemente a través de conexiones terrestres surgidas en la Crisis de Salina del Messiniense.

3. *Ummidia algarve* y *Ummidia* sp. Tarifa son ecológicamente intercambiables.

4. Los análisis filogenéticos revelan numerosos linajes evolutivos independientes que podrían corresponder a nuevas especies.

Capítulo 3: **De Gondwana a Europa: infiriendo los orígenes de las arañas Mediterráneas del género *Macrothele* (Araneae, Hexathelidae) y los límites de la familia Hexathelidae**

1. Los métodos bayesianos de estima de edades mediante reloj molecular relajado revelan que la divergencia del género *Macrothele* se remonta a

la época de la desintegración de Gondwana y su distribución actual probablemente refleja la deriva continental.

2. Las dos especies europeas del género *Macrothele* se originaron probablemente en Asia y colonizaron el Mediterráneo de forma independiente.

3. Se confirma la polifília de la familia Hexathelidae y se identifica la subfamilia Atracinae como la causa principal de dicha observación.

4. El marco temporal de la diversificación de las arañas migalomorfas se desplaza aún más en el pasado, lo que sugiere un origen más antiguo del grupo y de sus linajes evolutivos principales.

Capítulo 4: **Las arañas sobre el tejado volcánico caliente: Las vías de colonización y la filogeografía de la araña migalomorfa endémica a las Islas Canarias *Titanidiops canariensis* (Araneae, Idiopidae)**

1. *Titanidiops canariensis* colonizó las Islas Canarias una sola vez, presumiblemente desde el norte de África

2. Las discrepancias entre las estimas de edad obtenidas a partir de los análisis concatenados y los de la reconstrucción del árbol de especie no permiten determinar el marco temporal de la colonización, pero varias líneas de evidencia apuntan hacia el Mioceno.

3. Se pone de manifiesto un patrón filogeográfico complejo, que incluye dos eventos independientes de colonización de Lanzarote o una colonización de ida y vuelta entre Fuerteventura y Lanzarote.

4. La actividad volcánica probablemente contribuyó a los patrones geográficos de diversidad genética y se confirma la existencia de un

refugio volcánico en Zonzamas, en la zona centro-occidental de Lanzarote, ya observado en otras especies endémicas.

5. Los métodos moleculares de delimitación de especies sugieren al menos dos especies en las Islas Canarias; una en la península de Jandía y el sur de Fuerteventura y una segunda en el centro de Fuerteventura y Lanzarote norte. Las dos especies coocurren al menos en una localidad.

6. Se detectan varios linajes evolutivos divergentes también dentro de *T. maroccanus*, que pueden corresponderse a nuevas especies crípticas.

7. *Titanidiops maroccanus* parece estar más relacionado con especies del oriente próximo, lo que puede sugerir dos colonizaciones independientes del norte de África.

Anexo 1: **Amenaza o amenazada: Infiriendo la estructura poblacional de la araña negra de los alcornoques *Macrothele capeiana* (Araneae, Hexathelidae), una especie protegida**

1. Los resultados corroboran la existencia de linajes divergentes dentro de *M. calpeiana*, pero también niveles bajos de diferenciación genética entre localidades con rangos de distribución continuos.

2. Se han detectado cinco linajes divergentes en los análisis concatenados basados en genes nucleares, mientras que el genotipado mediante AFLPs revela solo dos grupos. El grupo que incluye las muestras del sur y este de la Península Ibérica muestran una variabilidad significativamente menor, potencialmente debido a expansiones poblacionales recientes.

3. La mayor parte de linajes se remontan al Pleistoceno, lo que sugiere que las fluctuaciones en los tamaños poblacionales y el desplazamientos del área de distribución durante los los ciclos glaciales pueden ser responsables de los patrones actuales de diversidad genética.

Annex II

Invertebrate Systematics

Systematics, Phylogeny and Biogeography

Just Accepted

From Gondwana to Europe: inferring the origins of Mediterranean Macrothele spiders (Araneae, Hexathelidae) and the limits of the family Hexathelidae

Vera Opatova, Miquel Arnedo

Abstract

The family Hexathelidae ranks among the smaller mygalomorph spider families. Most species are endemic to the Australasian region and the family was traditionally considered an example of a Gondwanan lineage. However, recent studies have cast some doubt on the monophyly of the family. Macrothele is the only genus with an out-of-Gondwana distribution. The bulk of the Macrothele diversity is found in Southeast Asia, few species are known from central Africa and two species inhabit Europe: Macrothele calpeiana from the Iberian Peninsula and Macrothele cretica endemic to Crete. Here we investigate the origins of the European Macrothele species by means of a multi-locus phylogenetic approach and by inferring the timeframe of the diversification of the genus using Bayesian relaxed clock methods. We also provide further insights into the phylogenetic status of the family Hexathelidae. Our results indicate that the diversification of Macrothele traces back to the period of the Gondwana breakup and its present day distribution most likely reflects the subsequent tectonic plate movements. The two European species were not recovered as sister taxa, suggesting that Macrothele colonized the Mediterranean region twice independently. The polyphyly of the family Hexathelidae is further confirmed and the subfamily Atracinae is identified as the conflicting lineage.

IS14004 Accepted 05 April 2014