Molecular phylogeny of the harvestmen genus Sabacon (Arachnida: Opiliones: Dyspnoi) reveals multiple Eocene–Oligocene intercontinental dispersal events in the Holarctic

Axel L. Schönhofer, Maureen McCormack, Nobuo Tsurusaki, Jochen Martens, Marshal Hedin

1. Introduction

1.1. Biogeographic studies of the Holarctic

Over the past 20 years the biogeographic history of the Northern Hemisphere has been the subject of several detailed studies and meta-analyses (e.g., Donoghue and Smith, 2004; Enghoff, 1995; Sanmartín et al., 2001). The Holarctic is the largest planetary ecozone, spanning a large proportion of the Northern Hemisphere including parts of Europe, Asia, and North America, and many subdivisions exist according to ecology, geography, floristic and faunistic characteristics. For most biogeographic studies spanning the Holarctic, four major infraregions have been applied (e.g., Enghoff and Smith, 2004; Sanmartín et al., 2001) corresponding to zones of temperate forest endemism in the western Palearctic (Europe), eastern Palearctic (eastern Asia) and the Nearctic (eastern and western North America; see Fig. 1 in Donoghue and Smith, 2004).

The temporal connectivity between the four major Holarctic infraregions follows a complex reticulate pattern (Smith et al., 1994; summarized in Sanmartín et al., 2001). Mid- to late Cretaceous (100–80 Ma) geographic settings were characterized by the connection of Europe with eastern North America, and Asia with western North America. At the end of the Cretaceous (70–65 Ma) the American mid-Continental Seaway closed and the North American continent appeared as it does today. The opening of the Atlantic (~90 Ma) isolated North America and Europe, but major land bridges still connected these two continents, most importantly, the Thulean Bridge around 55–50 Ma. Subsequent land bridges were apparently less suitable for biotic interchange, especially for warm-adapted species (McKenna, 1983; Sanmartín et al., 2001; Tiffney, 1985). North America and Asia were repeatedly connected via the Bering Bridge, with biotic interchange thought to be
moderated mainly by climatic settings. A boreotropic forest connected eastern Asia and eastern North America during the warm and humid Eocene, ultimately disrupted by global cooling and orogeny of the Rocky Mountains at the end of this epoch (35 Ma). In the Pleistocene (1.5–1 Ma) the colder climate restricted Beringian interchange to mostly cold-adapted boreal species.

Development of novel parametric methods in biogeographic studies allows for a more comprehensive reconstruction of complex Holarctic processes than previous methods (Donoghue and Smith, 2004; Sanmartín et al., 2001 and summary herein). Most Holarctic studies implementing parametric biogeographic methods with molecular phylogenetic data focus on plant taxa, with animal studies still scarce (e.g., Burbrink and Lawson, 2007; Vieites et al., 2007). Comparisons of codistributed plant and animal taxa often reveal contrasting distributional and biogeographic patterns at comparable phylogenetic levels. For example, the degree and direction of dispersal events between Holarctic infraregions varies considerably, showing extensive interchange between floral elements of eastern North America and eastern Asia, whereas animal disjuncts across these regions are rare (Donoghue and Smith, 2004; Sanmartín et al., 2001). Holarctic plant genera are generally widespread with few plant families endemic to the Holarctic. Conversely, many Holarctic arthropod taxa show family-level endemism at continental or smaller geographic scales. The increased structuring seen in arthropods perhaps offers further insight into the complex Holarctic biogeography, particularly when conducted on groups considered as model taxa for biogeographic analyses.

1.2. Harvestmen as model taxa for historical biogeography

Recent phylogenetic studies within the arachnid order Opiliones have shown that the diversification history of many groups is tightly interconnected with Earth history processes (e.g., Boyer et al., 2007; Giribet et al., 2010; Hedlin et al., 2012; Schönhofer and Martens, 2010). The integration of phylogeny and biogeography has therefore become necessary to delineate centers of Opiliones endemism and biodiversity, as illustrated by the well-studied mite harvestmen (Boyer and Giribet, 2007; Boyer et al., 2007; Giribet et al., 2012). Despite these many biogeographic contributions, few Opiliones groups with primarily Holarctic distributions have been the focus of phylogenetic or biogeographic analyses.
attention. Relationships of Holarctic Sironidae have been difficult to resolve (Giribet and Shear, 2010), and this group seems mostly absent from continental Asia. Species of Caddo (Eupnoi) provide insight into recent biogeographic exchange between eastern North America and eastern Asia (Shultz and Regier, 2009). Regional Holarctic clades within the long-legged Phalangioidea are still being defined (Hedin et al., 2012).

Members of the harvestmen suborder Dyspnoi are a promising group for studies of Holarctic biogeography. This clade includes about 350 species in approximately 40 genera, all with north temperate distributions. Most Dyspnoi genera are restricted to single continents (Shear, 2010) while taxa above this rank are distributed across continents. An exception is the genus Sabacon (Fig. 1), which spans a larger distributional area than any other Dyspnoi genus, with centers of endemism in eastern Asia (Martens, 1972, 1989; Suzuki, 1974; Tsurusaki and Song, 1993a), eastern and western North America (Shear, 1975), and south-western Europe (Martens, 1983; Fig. 1F). In these areas, Sabacon species show strong preferences for permanently cool and moist habitats, occurring in various sheltered microhabitats in coniferous or mixed forests, in deep alpine stony gravel, and in caves. Although a phylogenetic analysis of all Sabacon taxa has never been conducted, morphological similarities between European and eastern North American taxa (Martens, 1983; Shear, 1975, 1986), as well as between North American and Asian species (Martens, 1983; Cokendolpher, 1984; Shear, 1975), suggest possible intercontinental biogeographic connections.

Existing biogeographic patterns imply that diversification in Sabacon must be relatively old. Fossil Sabacon dating to the Eocene are known from European amber (Dunlop, 2006), indicating both an old age and morphological stasis for this lineage. Lopez et al. (1980) mentioned the cave-dwelling Sabacon paradoxxus as an example of a Tertiary relict, and Thaler (1976) discussed S. simoni as a Holarctic relict taxon. Sabacon includes no transcontinental species, and does not feature typical boreal species, as found for example in the harvestmen suborder Eupnoi (e.g., Mitopus morio). The pattern of disjunct distributions within species groups, but narrow endemism of most Sabacon species, suggests generally poor dispersal abilities for members of the genus. Dispersal across open oceans is unlikely for Sabacon, and connection of landmasses can be assumed to be the only means for intercontinental biogeographic exchange.

1.3. Applying novel biogeographic methods to Opiliones

The study of biogeography, in combining distributional patterns of extant taxa and the reconstruction of their phylogenetic history, has undergone considerable changes over the last decade. Utilizing phylogenetic information derived from molecular data, novel modelling approaches using parsimony or likelihood reconstructions over sets of plausible trees (Lagrange: Rees and Smith, 2008; Smith, 2009; S-DIVA: Yu et al., 2010) are replacing simple area cladograms and purely parsimony-based methods. These new methodologies are able to incorporate age estimates and phylogenetic uncertainty, account for missing phylogenetic data, and add statistical confidence to reconstructions (Ree and Smith, 2008; Smith, 2009). While molecular phylogenetic analyses are becoming increasingly common within Opiliones, biogeographic results remain largely descriptive, and statistical approaches have been restricted to estimating divergence times (Boyer and Giribet, 2007; Boyer et al., 2007; Derkarabetian et al., 2010). With few exceptions (e.g., Giribet et al., 2012; Sharma and Giribet, 2011), new biogeographic methodologies have yet to be incorporated into Opiliones research.

The current study focuses on the systematics and biogeographic history of Sabacon based on sampling of major lineages including most described species diversity in North America and Europe, and representative species sampling from Japan, Nepal, and China. We use molecular phylogenetics to investigate relationships within Sabacon, and evaluate the systematic position of Sabacon within the Dyspnoi. Phylogenetic information is used to infer ancestral distribution areas of Sabacon clades, as well as the timing and directionality of Holarctic intercontinental dispersal. We also summarize differences in Sabacon morphology as these relate to recovered molecular lineages.

2. Material and methods

2.1. Taxon sampling

Sabacon species are relatively rare and sometimes difficult to collect. Many species are only known from high elevation habitats or caves, and many original descriptions are based upon single or few specimens. Nevertheless, we have obtained a comprehensive sample of North American and European taxa, and a representative sample of Asian taxa. Overall, our sample covers much of the described morphological and geographical diversity in Sabacon (see Table 1). Specimens were collected in the field via manual search and preserved in 100% EtOH, or collected into 80% and later transferred to 100%. Material was then stored at minus 80 °C until DNA extraction. Some older samples from Asia were retrieved from 70% EtOH collections. Adult specimens were used whenever possible, but due to the late-maturing seasonal phenology of Sabacon, inclusion of juveniles in this study was also necessary. The identity of juveniles was corroborated by morphological comparison with adults from the respective localities, partly loaned from the California Academy of Science (CAS), and from our own reference collections (San Diego State University Terrestrial Arthropod collection, University of Mainz Opilionid collections of A.L.S and J.M.).

2.2. Gene data collection, sequence alignment and phylogenetic analysis

Genomic DNA was extracted from leg tissue using the Qiagen DNeasy kit. The polymerase chain reaction (PCR) was used to amplify the following gene fragments: 28S rRNA (28S), using the primers ZX1, ZR2 (Mallatt and Sullivan, 1998), and the newly developed ZX1long (5′-ACCCGCTGAATTAGCATATGAC3′); mitochondrial Cytochrome Oxidase 1 (CO1), using C1-J-1718S-PIDER and C1-N-2776SPIDER (Vink et al., 2005); nuclear Elongation Factor 1-alpha (EF1α), using OP2BSAB and OPRC4 (Hedin et al., 2010). PCR protocols followed Hedin et al. (2010) for EF1α, Thomas and Hedin (2008) for CO1, and Hedin and Thomas (2010) for 28S, the latter using an annealing temperature of 56 °C. PCR products were purified on Millipore plates and amplicons were directly sequenced at Macrogen USA. Sequencer V4.5 was used to assemble and edit sequence contigs, and all ambiguous sites were scored as heterozygous using standard ambiguity codes.

CO1 sequences were aligned manually in MEGA 4.0 (Tamura et al., 2007) using amino acid translation, while 28S and EF1α sequences were aligned with MAFFT (vers. 6; http://mafft.cbrc.jp/alignment/software/). The G-INS-i strategy was used for the protein coding EF1α; to incorporate 28S structural information we used Q-INS-i as recommended by Katoh and Toh (2008). To account for alignment uncertainty, 28S and EF1α alignments were further trimmed using Gblocks (Castresana, 2000), applying a “less stringent” criterion with standard settings. Models of DNA sequence evolution were evaluated using jModelTest 0.1.1 (Posada, 2008) under three substitution schemes (JC, HKY, GTR) on a fixed BIONJ tree, allowing for unequal base frequencies and among-site rate variation. Final model selection was based on the Akaike
### Table 1

List of species/taxa with voucher information, locality data, GenBank accession for sequences of species used in this study. Voucher number abbreviations are: OP (Opiliones Collection Marshal Hedin, San Diego State University), AXLS (Collection Axel L. Schönhofer, University of Mainz, Germany), CJM (Collection Jochen Martens, University of Mainz, Germany).

<table>
<thead>
<tr>
<th>Species/taxon name</th>
<th>Voucher no.</th>
<th>Country: locality, GPS</th>
<th>GenBank no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ceratolasma tricantha</td>
<td>OP989</td>
<td>USA: OR, Curry Co., E. Gold Beach, 42.5465, –124.1273</td>
<td>JX573543 – JX573601</td>
</tr>
<tr>
<td>Hesperoplatystoma kephartii (Crosby and Bishop, 1924)</td>
<td>OP812</td>
<td>USA: CA, Dade Co., vic. Sutton’s Cave, 34.8597, –85.4847</td>
<td>JX573548, JX573642, JX573606</td>
</tr>
<tr>
<td>Ischiropalpus carinatus (Banks, 1905)</td>
<td>AXLS345</td>
<td>Switzerland: Kanton Luzern, Wengen, 46.83, 8.24</td>
<td>JX573544, JX573638, JX573602</td>
</tr>
<tr>
<td>Ischiropalpus hellwigii hellwigii (Panzer 1794)</td>
<td>AXLS240</td>
<td>Slovenia: OštREVa-Mountains, Sadni Travniki, 46.4570, 14.68605</td>
<td>JX573545, JX573639, JX573603</td>
</tr>
<tr>
<td>Ischiropalpus kollari</td>
<td>AXLS241</td>
<td>Austria: Steiermark, Koralpe, Wolfsberg, Großer Speikogel, 46.79874, 14.95902</td>
<td>JX573546 – JX573604</td>
</tr>
<tr>
<td>Ischiropalpus pyrenaicus Simon, 1872</td>
<td>AXLS620</td>
<td>France: Midi-Pyrénées, Dép. Ariège, Grotte de l’Estelas, 46.91413, 7.55971</td>
<td>JX573471, JX573641, JX573603</td>
</tr>
<tr>
<td>Ischiropalpus nubilus</td>
<td>OP2886</td>
<td>Japan: Shikoku, Ehime Pref., Kumakawa-mine, 36.66, 133.90</td>
<td>JX573594 – JX573603</td>
</tr>
<tr>
<td>Ischiropalpus nubilus</td>
<td>OP2888</td>
<td>Japan: Shikoku, Ehime Pref., Kumakawa-mine, 36.66, 133.90</td>
<td>JX573594 – JX573603</td>
</tr>
</tbody>
</table>

**continued...**
Information Criterion (AIC) and individual models were applied to respective partitions in all downstream analyses.

Bayesian inference using MrBayes v3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) was applied to single gene as well as concatenated datasets. Data were partitioned by gene, and for CO1 and EF1α, by codon position. Bayesian analyses were run for 5,000,000 generations, where in all cases the standard deviation of split frequencies had dropped below 0.01 (Ronquist et al., 2005). Analyses were repeated to further check for convergence. The first 40% of trees were discarded as burn-in, with remaining trees used to reconstruct a 50% majority rule consensus tree. Split frequencies were interpreted as posterior probabilities (pp) of clades.

We also investigated phylogenetic relationships for a larger set of Dyspnoi taxa to evaluate the relative phylogenetic position and distinctiveness of Sabacon within the suborder. This analysis was based on 28S rRNA, the only marker for which a comprehensive sample of Dyspnoi taxa was available. These data were generated, aligned and analyzed using the above-described methods. 28S sequences not generated for this study were downloaded from GenBank, based on studies of Giribet et al. (2010) and Schönhofer and Martens (2010; see Fig. 2).

2.3. Divergence time estimation using BEAST

We estimated taxon divergence times from the concatenated dataset using BEAST 1.6.1 (Drummond and Rambaut, 2007). Because initial analyses did not reach convergence, we conducted analyses on a reduced dataset of 37 specimens, excluding recently diverged lineages within the S. occidentalis and S. cavicolens clades. The Bayesian analysis was rerun and the reduced dataset topology compared to the full dataset topology to check for congruence. Codon partitioning was applied to CO1 (1, 2, 3) and EF1α ((1 + 2), 3) to accommodate heterogeneous rates of evolution (Brandley et al., 2011). Models for CO1 and EF1α partitions were set to the second best fitting HKY+I+G model; although other priors were high, effective sample size (ESS) values for rare substitutions within partitions for a GTR model remained low and analyses did not converge. Strongly supported nodes (pp = 0.95–1.00) from the full dataset were constrained to be monophyletic in the reduced dataset, to avoid entrapment in local optima (following Smith, 2009). No treemodel operators were removed, allowing BEAST to re-estimate the remaining topology. Priors were set to gamma if required to be modified. These settings were applied to all subsequent analyses. Analyses were run until ESS values exceeded 200 even for poorly sampled priors, which was after 100,000,000 generations. Analyses were replicated twice and checked for convergence using Tracer 1.5.

We checked for rate heterogeneity in all datasets using the random local clock model implemented in BEAST. This clock model estimates rate changes for each branch individually and identifies lineage specific rate changes (Drummond and Suchard, 2010). An uncalibrated log-normal analysis estimated different clock models for the three gene partitions, and the standard deviation of the partition frequency (stddev) was investigated for deviation from clock-like behavior.

For final analyses in BEAST a Death-Birth model of speciation and a relaxed log-normal clock were applied. Clock rates where unlinked and estimated for individual partitions. We simultaneously applied both rate and fossil calibrations for a final BEAST reconstruction. For the clocklike CO1 (see Section 3.3.) we specified a clock rate based on the standard arthropod clock from Brower (1994; 2.3% per Ma) and a newly reported arthropod clock rate (Papadopoulou et al., 2010; 2.69%/Ma). To accommodate differences of these rates a uniform prior was chosen for CO1. Meanrate to include the range of 0.0115–0.0145/Ma. The fossil taxon Sabacon
claviger from European amber was used as a calibration point for the minimum age of the cavicolens-Europe clade (see Section 3.2.). This taxon dates to the Eocene (~55–34 Ma; Dunlop, 2006) for which a log-normal distributed prior was set with a mean of 2.53, a Stdev of 0.9, and an offset of 32 so the 2.5% percentile would be 34 Ma and the median 44.18 Ma, constraining the minimum age of the node but allowing it to be significantly older.

2.4. Ancestral range reconstructions using S-DIVA and Lagrange

Ancestral distributions were reconstructed using S-DIVA (Yu et al., 2010), which evaluates the most parsimonious reconstructions of dispersal and vicariance events and integrates uncertainty over a set of analyzed trees. Here we used 5000 post burn-in trees resulting from Bayesian analysis of the full concatenated dataset. Default settings were modified for “Allowed Reconstruction” (checked) and Hold set to 32.767 (to include all alternative reconstructions per node, see Ronquist (1997), DIVA 0.1.1 manual). We specified five distributional areas for terminal taxa, including: A: continental Asia (including Nepal and China only), B: western North America, C: eastern North America, E: Europe, and J: Japan. We conducted analyses with maximum alternative scenarios at each node set to 5, 4 and 3 to examine the effect of constraints on reconstructed ancestral distributions (Bendiksby et al., 2010; Kodandaramaiah, 2010).

Likelihood reconstruction of ancestral areas using Lagrange (C++ version, Bendiksby et al., 2010; http://code.google.com/p/lagrange/downloads/detail?name=lagrange_cpp_0.1.8.tar.gz&can=28&q) was performed on the dated Bayesian consensus phylogeny from BEAST. Utilizing phylogenetic and temporal information, Lagrange reconstructs the biogeographic history of a set of discrete areas based on the distribution of extant taxa. Accounting for branch length information, Lagrange integrates the probability of extinction and dispersal as stochastic events calculated over time. Nodes that showed ambiguous area reconstruction results in the Bayesian consensus tree were further investigated by reconstructing over all 5000 retained BEAST trees, following Smith (2009) and Bendiksby et al. (2010). Lagrange output was manipulated using TextWrangler for input into “R” 2.13.1 and Excel 2007, to calculate probabilities of all ancestral area scenarios at each node, and to segregate the five best scoring scenarios per node (Smith, 2009). Lagrange C++ computation omits scenarios scoring less than 5% probability in individual reconstructions, thereby “weeding out” weakly supported scenarios. No constraints on the maximal number of areas, or on the directionality and timing of dispersal, were used in our analyses.

3. Results

3.1. Dyspnoi 28S phylogeny

Tree topologies based on 28S MAFFT versus MAFFT+Gblocks alignments were similar (results not shown), so we used the MAFFT-only alignment. The 28S phylogeny including multiple Dyspnoi and Eupnoi outgroup taxa (1268 aligned base pairs for 82 terminals) strongly supports the superfamilies Ischyropsalidoidea and Troguloidia (1.00 pp; Fig. 2). The recently emended Sabaconidae (Giribet et al., 2010), including Hesperonemastoma, Sabacon and Taracus, receives low support (0.78 pp), but Hesperonemastoma and Taracus are strongly supported as sister taxa (1.00 pp). Sabacon is strongly supported as monophyletic (1.00 pp; Fig. 2) and shows high internal 28S sequence divergence, with separation into four well supported primary clades (Fig. 2). Based on 28S phylogram branch lengths, sequence divergence within Sabacon exceeds that of any other Dyspnoi genus included in the analysis, and is indeed greater than that of most Dyspnoi families (Fig. 2).

3.2. Phylogenetic relationships within Sabacon

Fifty-four Sabacon specimens plus 7 outgroup taxa specimens were used to generate sequence data for 28S, CO1 and EF1α (59, 50, and 42 sequences, respectively). Only 28S could be amplified for some of the older samples and several EF1α PCR products yielded no results in direct sequencing, possibly due to the close proximity of the forward primer to additional introns reported for Sabacon (Hedin et al., 2010). The final alignments of CO1 (968 bp) and EF1α (861 bp) were unambiguous. We excluded the difficult-to-align (and sometimes partial) EF1α introns, as suggested by Gblocks. 28S was difficult to align and we compared tree topologies from the MAFFT and MAFFT+Gblocks alignments. Results based on these alternative alignments were very similar, differing only for nodes that showed low support in analyses derived from both alignments. An exception involved S. occidentalis sequence OP2591, which changed position within the occidentalis-clade with high support from alternative alignments. The placement of this particular sequence does not greatly impact the biogeographic analyses presented below. We used the MAFFT-only alignment (1236 aligned bp) in subsequent concatenated analyses. For individual gene trees see Suppl. Figs. 1–3.

In agreement with results from the larger 28S dataset (Fig. 2), analysis of the concatenated alignment (2885 total aligned bp) recovers Sabacon as monophyletic with strong support (1.00 pp, Fig. 3). Within Sabacon, four primary clades emerged that we define here for further discussion (Figs. 3–5): (1) western North American occidentalis-clade, basal placement in Sabacon, unites S. briggsi (1a), S. occidentalis and S. siskiyou (1b); (2) mitchelli-Japan-clade, unites the eastern North American S. mitchelli and the Japanese S. akiyoshiensis and S. imamurai; (3) astoriensis-Asia-clade, unites the western North American S. astoriensis and S. sheari (3a) with S. jiriensis and Sabacon sp. from Nepal and sampled Chinese Sabacon (3b) – this clade is subdivided according to geographic origin; (4) cavicolens-Europe-clade, unites the eastern North American S. cavicolens (4a) and sampled European Sabacon species. These four clades are strongly supported in both individual gene and concatenated analysis (Fig. 2). A conspicuous implication of this phylogenetic structuring is that three of four regional faunas (wDNA, eDNA, east Asia) are non-monophyletic, and that three clades include intercontinental disjuncts.

3.3. Divergence time estimation using BEAST

Rate heterogeneity as estimated by the random local clock model in BEAST (Drummond and Suchard, 2010) was generally low throughout Sabacon and Sabaconidae, although we found significantly increased rate change in the astoriensis-Asia-clade with additional increase within S. astoriensis. This result holds true for the Dyspnoi, with minor rate changes between major groups (only decreasing in Dicranolasmatidae) but marked changes in individual lower level taxa (Anelasmocephalus, Mitostoma). The uncalibrated log-normal analysis rejected clocklike evolution for EF1α (stddev: 2.1) and 28S (2.2), but not for CO1 (stddev: 0.3), justifying fetal BEAST settings as previously specified.

Unconstrained BEAST trees differed noticeably from MrBayes trees at deeper nodes, recovering a basal polytomy including Sabacon and outgroups, and implied different interrelationships of the four major Sabacon clades. As these alternative BEAST reconstructions showed low support, and BEAST is known to converge on local optima, we constrained basal nodes to force BEAST to
converge to the same topology as MrBayes, as suggested by Smith (2009). This resulted in faster convergence and higher ESS values for all parameters and allowed setting of more complex substitution models as estimated by jModelTest. The resulting BEAST chronogram is shown in Fig. 4 (see also Suppl. Fig. 4).

According to our reconstruction Sabacon originated about 85 Ma ago (53–132 highest posterior density (HPD)), with subsequent diversification into four primary clades occurring within a relative short time frame of ~20 Ma (Fig. 4). The estimated ages of the most recent common ancestor (MRCA) of the primary clades falls within a time window of less than 10 Ma (mean times of 30–40 Ma), near the Eocene–Oligocene boundary. Intraspecific divergences within nominal taxa are estimated to be relatively old, often predating the Pliocene. More recent divergences, partly dating to the Pleistocene, were detected within the more densely sampled Sabacon cavicolens (2.1–3.2 Ma) and S. occidentalis clades (1.5 Ma).

3.4. Ancestral area reconstruction using S-DIVA and Lagrange

S-DIVA analyses were conducted on a set of post-burnin trees resulting from Bayesian analysis of the full concatenated dataset. S-DIVA results vary depending upon whether the number of maximal areas per ancestral node was constrained to 3 or 4, or left unconstrained at 5. Nodes for which S-DIVA calculated several equally likely scenarios if left unconstrained often tended to favor a single area scenario in more constrained analyses (Fig. 5). Also, when optimal results could have yielded the same area scenario (e.g., a scenario including three areas), the reconstructions did not agree. Finally, nodes with alternative scenarios showed a higher number of possibilities that also differed considerably from the Lagrange C++ results.

Lagrange C++ analyses were conducted on a set of trees resulting from BEAST analysis of the reduced concatenated dataset. Fig. 4 pie charts show the total proportion of each scenario at nodes with possible alternatives, whereas histograms show frequencies of the five best scoring scenarios distributed over the 95% higher posterior density of node age reconstructions (see also Suppl. Fig. 4). Pie charts thereby depict the total likelihood of each scenario at the nodes, while the likelihood of timing of scenarios is shown by histograms. Large differences in the timing of alternative scenarios were generally not detectable in our dataset. While the two expanded histograms for the astoriensis-Asia and mitchelli-Japan clades, and for the most recent common ancestor of Sabacon show a shift of about 10–15 Ma at the most likely time of the individual area reconstruction, these are small differences in comparison to the overall 95% higher posterior densities and the total histogram coverage (Suppl. Fig. 4).
Lagrange (Fig. 4) and S-DIVA (Fig. 5) reconstructions support intercontinental dispersal in two of three intercontinental disjunct clades. Within the cavicolens-Europe-clade, S-DIVA unambiguously supports Europe as the center of origin. While Lagrange also supports a European center of origin with high likelihood, this reconstruction does not exclude a shared ancestral area of Europe and eastern North America (i.e., ambiguous dispersal directionality). The biogeographic connection between Europe and eastern North America supports the assignment of the fossil Sabacon claviger to a lineage of European (+eastern North American) origin, as emphasized here. Results for the mitchelli-Japan-clade differ between analyses. Lagrange shows higher support for an ancestral area combining Japan with eastern North America, while S-DIVA favors a Japanese center of origin. Neither reconstruction method provides unambiguous results for dispersal directionality in the astoriensis-Asia-clade. At deeper nodes, both methods weakly support western North America as the ancestral area for Sabaconidae (Suppl. Fig. 4).

4. Discussion

4.1. Sabacon systematics

Sabacon is the nominative genus of the originally monotypic family Sabaconidae (Dresco, 1970). Subsequent evaluations of the
family first rejected its status (Gruber, 1978; Martens, 1972; Shear, 1975) but then accepted an altered family diagnosis (Martens, 1976, 1980, 1983; Shear, 1986). Martens et al., (1980), Martens (1983) based evidence on the peculiar and derived genital morphology, unique within the Dyspnoi. Later, Shear (1986) downgraded genital morphology as genus specific for *Sabacon* and included *Taracus* in the Sabaconidae on the basis of a cladistic analysis. The latest emendation was by Giribet et al. (2010), adding *Hesperonemastoma* to Sabaconidae based on molecular phylogenetic affinities to *Taracus*. Our reassessment of Sabaconidae based upon a more comprehensive taxon sample finds low support for the relationship of *Sabacon* with *Taracus* plus *Hesperonemastoma* (Fig. 2; comparable to Giribet et al., 2010).

Compared to 28S rRNA divergence of other families within the Dyspnoi (Fig. 2), *Sabacon* emerges as an isolated and highly divergent clade, and perhaps should be re-classified into multiple genera. While we do not attempt to taxonomically subdivide *Sabacon* in this paper, we provide evidence for deep divergence encouraging such endeavors. Support for a taxonomic split also stems from morphology, where comparisons of Sabacon molecular lineages reveal clade-specific differences in male genital morphology. This situation is similar to that seen within the Dyspnoi family.
Fig. 5. S-DIVA reconstructions. Histograms on the left show differences of analyses constraint for 3, 4, and 5 areas per node. Colored circles and letters at nodes show identical results for all three constraints. Indicators (1–4) at nodes indicate clades discussed in the text.
Nemastomatidae, members of which exhibit relative stasis of body form but display greater variation in male genital and/or secondary sexual characteristics (Schönhofer and Martens, 2012).

*Sabacon* is separated into four well-defined molecular clades (Figs. 3–5); here we discuss morphological support for these clades.

1. **Occidentalis-clade:** *Sabacon briggsi* is sister to a group comprising *S. occidentalis* and *S. siskiyou*. The latter are very similar but unique within *Sabacon* in exhibiting inflated penis glands with a firmly attached, three-dimensional and partly divided stylus (Shear, 1975). From the single male *S. briggsi* available for our analysis we can confirm stylus modifications, including a regular row of backwards pointing teeth. All three species share large conical cheliceral apophyses very similar in shape and size. Within the *occidentalis* and *siskiyou* group OP2591 is genetically unique and needs to be further investigated. Specimens collected from this location are smaller than other members of this clade, but only males are available. We expect at least stylus–specific differences for males from this potentially undescribed species.

2. **Mitchelli-Japan-clade:** The eastern North American *S. mitchelli* is phylogenetically related to Japanese taxa, a result suggested by Cokendolpher (1984) based on comparative morphology. All three species in the recovered *mitchelli-Japan-clade* lack male cheliceral glands, but differ in genital morphology. The relatively simple penis of *S. mitchelli* (Shear, 1975) differs from the three-dimensional glandes penis with distinct spineation, found in the two Japanese species (Suzuki, 1974). For the nine described Japanese *Sabacon* species, Suzuki (1974) recognized the following three species groups based on external morphology (penis, ovipositor, armaments of male chelicera and palpal patella, scutum, and body size): *pygmaeus-group*, *denticapalis-group*, and *akiyoshienis-group*. It is interesting that both *S. akiyoshienis* of the *akiyoshienis-group*, which is the smallest Japanese species (ca. 1.5 mm in male body length) and *S. imamurai* of the *denticapalis-group*, which is the largest Japanese species (ca. 4 mm), are placed in the same clade despite their large difference in external morphology and body size. The phylogenetic position of the *pygmaeus-group*, not sampled in this study, remains to be determined. Also, there are *Sabacon* species from the Chinese Sichuan that appear morphologically divergent, *Sabacon cavicolens* from eastern North America probably comprises several cryptic species (estimated divergence mean between 2 and 12 Ma). Shear (1975) mentioned morphological divergence within *S. cavicolens*, a matter that requires further investigation.

3. **Astoriensis-Asia-clade:** This group is subdivided according to geographic origin. The western North American clade includes *S. sheari* and *S. astoriensis*. Cokendolpher (1984) considered *S. sheari* to be closely related to *S. mitchelli* based upon external morphology, but this relationship is not supported by our data. There is considerable sequence divergence between the Oregon/Washington versus Idaho populations of *S. astoriensis*. Investigation of specimens from these disjunct areas reveals differences in male chelicerae and penis morphology (pers. obs.), perhaps consistent with species-level differentiation. There remains uncertainty in the correct phylogenetic placement of taxa in this clade. Different analyses swapped the position of *S. sheari* (Figs. 2 and 3), and basal relationships in the *astoriensis-sheari-clade* of the full dataset (Fig. 3) are partly unresolved. We also found considerable rate change within this clade suggesting alignment or systematic errors, which might be overcome by incorporating additional Asian taxa. Our current sparse sampling prevents strong conclusions regarding the phylogenetic structuring of mainland Asian taxa. According to our results and morphological evidence we argue that most of the Chinese and Nepalese divergence is relatively recent, but we acknowledge that this fauna is morphologically disparate, and that our current sample does not capture most of this morphological divergence.

4. **Cavicolens-Europe-clade:** Based on comparative morphology, Shear (1975) and Martens (1983) suggested a close relationship between *S. cavicolens* and European *Sabacon*, corroborated in this study. Also, speciation in this clade is clearly older than the Pleistocene, consistent with the hypothesis of Martens (1983). Basal nodes within the group receive only low support, placing *S. simoni* closest to *S. cavicolens*. Strongly supported European relationships include *S. paraadoxus* with *S. viscayanus* (as expected by Martens, 1983), and *S. pasomianus* with the *Sabacon* juvenile from the Cevennes in southern France (MCZDNA100711). These latter taxa are not in geographic proximity and are genetically divergent. *Sabacon cavicolens* from eastern North America probably comprises several cryptic species (estimated divergence mean between 2 and 12 Ma). Shear (1975) mentioned morphological divergence within *S. cavicolens*, a matter that requires further investigation.

4.2. **Historical biogeography**

While the northern latitudes climate of the Eocene was more tropical, a sharp decrease in average global mean temperatures at the Eocene–Oligocene boundary (~ 33 Ma; Zachos et al., 2001) marks the almost simultaneous intercontinental biogeographic exchange observed in all *Sabacon* clades (Fig. 4). Most extant *Sabacon* taxa show preferences to permanently cool and moist habitats, even to include subtropical cloud forest taxa occurring in Nepal (some included in our sample; Martens, 1972, 1983). We hypothesize that the onset of global cooling enabled cryophillic *Sabacon* to disperse into more suitable habitats, and to diversify into the many lineages we see today, a causal connection also found for cold adapted snakes (Lynch, 2009) and salamanders (Vieites et al., 2007; Zhang et al., 2008). Biotic conditions during connection of the trans-Atlantic de Geer land bridge provide an illustrative example. At the Eocene–Oligocene boundary the climate is reconstructed as becoming more seasonal with cooler winters but moderate annual temperatures (Eldrett et al., 2009; Ivany et al., 2000). In Greenland, as part of the de Geer Bridge, these conditions included high annual precipitation promoting a humid coniferous forest (Eldrett et al., 2009). Modern humid coniferous forests house rich *Sabacon* faunas, as for example, in western North America. Fennoscandian *Sabacon* fossils from Baltic amber (Dunlop, 2006), also from the de Geer Bridge, most likely originated from Sciadopityaceae (Wolfe et al., 2009), which together with other dominant conifers, went almost globally extinct in the late Oligocene. This climatic and fossil evidence suggests high latitudinal dispersal conditions likely suitable for *Sabacon*. We hypothesize that global cooling created large and maybe continuous suitable habitats, enabling a general expansion of cryophillic *Sabacon* and connecting habitat patches from the warmer Eocene.

Phylogenetic relationships within the *cavicolens*-Europe-clade clearly favor Europe as the center of origin for this clade. A possible scenario is that the complex and repeated fragmentation of Europe (Smith et al., 1994) caused early and deep divergence of European *Sabacon*, while the trans-Atlantic de Geer land bridge facilitated dispersal to eastern North America. Further speciation within *S. cavicolens* is considerably younger and must have happened long after the land bridge had been severed.

For the *mitchelli*-Japan-clade, ancestral areas are either Japan or a combination of Japan and eastern North America; the former fa-
vored by S-DIVA and the latter by Lagrange. Major events disrupting a connection between Asia and eastern North America include the uplift of the Rocky Mountains during the Eocene (45–36 Ma) followed by a sharp decrease in average temperatures during the Oligocene (~33 Ma) to the present (Zachos et al., 2001). While decreasing temperatures might have facilitated dispersal of cold adapted Sabacon, the increasing aridification of mid-continenental North America could have prohibited subsequent faunistic interchange.

According to the MrBayes reconstruction of the full dataset (Fig. 3), the astorienisis-Asia-clade comprises two well supported geographic lineages, with western North American S. astorienisis and S. shearii as sister to the Asian Sabacon. The separating split, dating to 32 Ma (17–53 HPD) falls in the time frame for increased faunistic interchange between western Nearctic and eastern Palearctic (Sammartin et al., 2001), while no recent Pleistocene exchange is apparent. While our data does not allow an inference of dispersal directionality in this clade, high morphological diversity of recent Asian Sabacon indicates continental Asia as the potential source area.

Although some biogeographic patterns seen in Sabacon are also found in other taxa, Sabacon is peculiar in several respects. For example, one connection for which Donoghue and Smith (2004) recovered the highest proportion of disjunct distributions, Asia and Europe, is completely absent from our biogeographic scenarios. Also, there is no apparent phylogenetic affinity between species in eastern versus western North America. Donoghue and Smith (2004) re-analyzed the animal data of Sammartin et al. (2001) and found few animal disjuncts between eastern North America and Europe (7%), with a low proportion (4%) of these reconstructed dispersal events occurring from Europe to eastern North America. The eastern North America to Japan connection is rare for animal disjuncts and dispersal corridors (Brewer et al., 2012; Donoghue and Smith, 2004; Sammartin et al., 2001). Another peculiarity concerning the mitichelli-Japan-clade is the estimated divergence time for the geographical groups (mean = 32 Ma, 16–53 HPD). Xiang et al. (2000) dated splits between eleven species pairs of plants, and found most estimates in the last 10 Ma (max. 12.6 Ma) in the late Miocene and Pliocene. These estimates probably correspond to other Opiliones taxa, e.g., Crosbycus dasyneurus (Shear, 1986), Acropopilio boopis (Suzuki, 1976), and two Caddo species (Suzuki, 1976; Shultz and Regier, 2009) with disjunct distributions in Japan (and China for Crosbycus; Tsurusuki and Song, 1993b) and eastern North America. Divergence in these harvestmen taxa is likely relatively young, as evidenced by morphological similarity and low genetic divergence (when estimated).

Our molecular clock-derived divergence date estimates differ considerably from similar molecular clock estimates reported by Gribet et al. (2010). They estimated the split between S. cavicolens and European Sabacon at ~75 Ma, which greatly exceeds our estimates (mean of 33 Ma, 24–46 HPD). Similarly, the Gribet et al. (2010) estimate for the MRCA of Ceratolasma and Ischyropsalis (~160 Ma) again exceeds our estimate (65 Ma, 31–109 HPD), as does the overall age of the Ischyropsalidoidea (243 versus 137 Ma). These differences may stem from the predominantly ancient calibration points used by Gribet et al. (2010), which may not account for fluctuations along the calibrated regression and influence younger nodes (Conroy and van Tuinen, 2003; Rutschmann et al., 2007). Applying their root age of 243 Ma to our analysis caused considerable increase in overall HPD ranges (results not shown), and we therefore refrained from using this possible calibration. Ultimately, incorporation of additional reliable calibration points and more DNA sequence data will be needed to resolve this discrepancy.

To summarize, the onset of diversification within the four major Sabacon clades, as well as intercontinental exchange in three clades, is nearly simultaneous and seems connected to global cooling, allowing dispersal across northern land bridges. Our data supports the perspective that the study of arthropods provides more fine-scaled views of Holarctic biogeography. For example, while cryophilic Sabacon were able to survive in suitable microhabitats, the aforementioned coniferous megaflora went mostly extinct and is no longer available for molecular based reconstructions.

Acknowledgements

We thank Jared Grummer, Dean Leavitt, Casey Richard and Jordan Satler for help with the phylogenetic methods and for general discussion. Richard Ree and Stephen Smith provided help with the Lagrange analysis and methods used in Smith (2009). Fred Coyle, Shahan Derkarabetian, Lucia Labrada, Dean Leavitt, Steven Thomas, Robin Keith, Wolfgang Schwaller, Joachim Schmidt, Jim Starrett, Jordan Satler, Adrienne Richard, Dale Richard, Dan Richard, Casey Richart and Peter Scott helped to collect specimens. Research efforts of M. McCormack were funded by the American Arachnological Society, and a U.S. Fish and Wildlife Service contract to M.H. A.L.S. was funded by a postdoc grant of the Deutsche Forschungsgemeinschaft (DFG) to work on systematics and biogeography of Holarctic Opiliones, and collected Sabacon in the southwestern Alps with the help of an EIDT/STIB-M travel Grant (2008). Field research in Asia was funded by Feldhausch Stiftung and Wagner Stiftung, both at Fachbereich Biologie, Mainz University, as annual grants to J.M.

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.2012.10.001.

References


Lynch, V.J., 2009. Live-birth in vipers (Viperidae) is a key innovation and adaptation to global cooling during the Cenozoic. Evolution 63, 2457–2465.


