ORIGINAL
ARTICLE

Dispersal patterns in space and time: a case study of Apiaceae subfamily Apioideae

Łukasz Banasiak[1], Marcin Piwczyński[2], Tomasz Uliński[3], Stephen R. Downie[4], Mark F. Watson[5], Bandana Shakya[5†] and Krzysztof Spalik[1]*

[1]*Department of Plant Systematics and Geography, Faculty of Biology, Institute of Botany, University of Warsaw, Warsaw, Poland,* [2]*Department of Animal Ecology, Institute of Ecology and Environment Protection, Nicolaus Copernicus University, Toruń, Poland,* [3]*Netezza R&D Department, SWG Lab Poland, IBM Software Group, Warsaw, Poland,* [4]*Department of Plant Biology, University of Illinois at Urbana–Champaign, Urbana, IL, USA,* [5]*Royal Botanic Garden Edinburgh, Edinburgh, Scotland, UK*

ABSTRACT

Aim To analyse spatial and temporal patterns of dispersal events in the euapioids (Apiaceae subfamily Apioideae).

Location Worldwide, with an emphasis on the Northern Hemisphere.

Methods A phylogeny of euapioids was inferred from 1194 nuclear ribosomal DNA internal transcribed spacer (nrDNA ITS) sequences using Bayesian methods. The reconstruction of ancestral areas was performed simultaneously with phylogenetic inference using a Markov discrete phylogeographical model with Bayesian stochastic search variable selection (BSSVS). Routes with significant non-zero migration rates were identified using Bayes factors. For each significant track and each tree, the distribution of dispersals in time was scored and the asymmetry of dispersals was evaluated.

Results The root of the euapioid umbellifers was reconstructed at *c.* 44.51 Ma (95% highest posterior density interval: 39.11–51.55 Ma). The Eastern Asiatic Region was reconstructed as the most probable ancestral area for the root of the tree. Seventeen migration routes have significant non-zero migration rates. Five of these tracks represent long-distance transoceanic routes: (1) western Eurasia–North America, (2) Eastern Asiatic Region–North America, (3) Australia–Neotropical Kingdom, (4) Australia–Neantarctic, and (5) Neotropical Kingdom–sub-Saharan Africa. Most dispersal events occurred among the areas that comprise the major diversification centres of apioid umbellifers: western Eurasia and the Mediterranean, Irano-Turanian and Eastern Asiatic regions. The floristic exchange among these regions was more or less symmetrical in all directions and continuous in time. The exchange between North America and the Eastern Asiatic Region was asymmetrical; the asymmetry of dispersals between western Eurasia and North America was less pronounced and not statistically significant. Floristic exchange was significantly asymmetrical for the Nearctic–Neotropical Kingdom, Nearctic–Neantarctic, western Eurasia–sub-Saharan Africa, and western Eurasia–Siberia migration tracks.

Main conclusions The observed dispersal pattern – intense and symmetrical within the same climatic zone versus scarce and often unidirectional between climatic zones or along long-distance transoceanic tracks – suggests that both the availability of suitable habitats and geographical barriers have played crucial roles in determining the present distribution of euapioid umbellifers.

*Correspondence: Krzysztof Spalik, Department of Plant Systematics and Geography, Institute of Botany, University of Warsaw, Aleje Ujazdowskie 4, PL-00-478 Warszawa, Poland. E-mail: spalik@biol.uw.edu.pl
†Present address: Bandana Shakya, International Centre for Integrated Mountain Development, Khumaltar, Kathmandu, Nepal.*

INTRODUCTION

Long-distance dispersal has recently been reinstated as a major factor shaping species distributions. Although traditionally regarded as inherently stochastic, several studies indicate that long-distance dispersal is structured in space and time. First, geographical barriers arise and disappear, opening and closing migration windows and routes (e.g.

Tiffney & Manchester, 2001). Second, migration is often physically directional, with examples including dispersals from continental land masses onto newly created islands, down river systems, and along prevailing winds or ocean currents (Cook & Crisp, 2005; and references therein). Third, successful colonization is conditional upon habitat availability, which changes over space and time. The tendency for species to retain their ancestral ecology affects their dispersal opportunities, and ecological conservatism seems to be a major factor shaping the distribution of plants and composition of biomes (Crisp et al., 2009). The importance of dispersal in explaining present distributions of plants has been well documented, but the general pattern of dispersals across space and time is still poorly understood. As pointed out by Donoghue & Smith (2004), such generalizations are difficult because they are mostly based on qualitative studies that are biased in several ways. These biases include a greater interest of the plant systematics community in specific disjunction patterns and a concentration on those groups that exhibit such patterns, as well as a preference for studying smaller groups with completely resolved phylogenies. Several biogeographical questions, however, call for a quantitative rather than a qualitative approach (Sanmartín et al., 2001, 2007; Sanmartín & Ronquist, 2004; Crisp et al., 2009). In this paper, we adopt a quantitative approach to address the question of dispersal pattern in space and time in the angiosperm family Apiaceae (Umbelliferae) subfamily Apioideae.

Apiaceae subfamily Apioideae includes c. 400 genera and 2900 species, which are widely distributed in temperate regions of the Northern Hemisphere, with fewer representatives occurring in the south (Pimenov & Leonov, 1993; Plunkett et al., in press). The Southern Hemisphere members of the subfamily either represent early-branching lineages (Downie & Katz-Downie, 1999; Calviño et al., 2006; Magee et al., 2010) or crown taxa of recent Northern Hemisphere origin (Spalik et al., 2010). The branches forming a basal paraphyly in Apioideae (i.e. protoapioids sensu Magee et al., 2010) occur almost exclusively in sub-Saharan Africa, a region that was also reconstructed as the ancestral area for the subfamily (Calviño et al., 2006). The remaining Apioideae (i.e. euapioids) probably originated in the Northern Hemisphere, but their ancestral area has yet to be formally reconstructed.

Due to extensive phylogenetic research on Apiaceae, over one-third of its species have been sampled for nuclear ribosomal DNA internal transcribed spacer (nrDNA ITS) sequence variation. These studies comprise both taxonomic analyses of major umbellifer clades, as well as studies based on geographical rather than taxonomic sampling, with all major centres of endemism covered (summarized in Downie et al., 2010). The ITS sequences of euapioids can be more or less unambiguously aligned, whereas the analysis of the entire subfamily requires more conservative markers to be used (Downie et al., 2010). The wealth of available molecular data, as well as its worldwide distribution, makes the euapioid umbellifers an excellent group in which to examine dispersal patterns among temperate taxa.

The aims of this study were to assess the dispersal model for apioid umbellifers, by identifying migration routes that explain the present distribution of apioid umbellifers, assessing the direction of dispersal along those routes, and estimating the timing of these dispersals.

MATERIALS AND METHODS

Taxon sampling and sequence alignment

In total, 1194 accessions of Apiaceae subfamily Apioideae were examined for ITS sequence variation (see Appendix S1 in Supporting Information). From the sequences available in GenBank we excluded duplicate and misidentified accessions (Downie et al., 2010). We also omitted accessions of the protoapioids from southern Africa, as their ITS sequences cannot be reliably aligned with those of the higher umbellifers. The included species represented all major lineages of core Apioideae. The DNA sequences were aligned using MAFFT 6.714b (Katoh & Toh, 2008), with the number of tree-rebuilding steps and maximum number of iterations set to five and options FFT and default gap opening penalty and offset parameters selected. The resulting matrix was then edited using MESQUITE 2.73 (Maddison & Maddison, 2010). Because the intervening 5.8S rDNA region was missing in its entirety for many accessions, it was excluded. The alignment of ITS1 and ITS2 regions was unambiguous and all positions were included in the analysis. Pairwise uncorrected sequence divergence was calculated using PAUP*4b10 (Swofford, 2003). The matrix is available in TreeBASE (study number S13377; http://purl.org/phylo/treebase/phylows/study/TB2:S13377).

Phylogenetic analyses and tree calibration

Because of the large size of the matrix, Bayesian analyses starting from random trees rarely converged on a stationary distribution within a reasonable period of time (several weeks). Therefore, a preliminary phylogenetic tree was inferred using maximum likelihood, as implemented in PHYML 3.0 (Guindon & Gascuel, 2003), with the nucleotide substitution model chosen using MODELTEST 3.6 (Posada & Crandall, 1998). This tree was then transformed into a chronogram with arbitrary branch lengths using FIGTREE 1.3.1 (Rambaut, 2009) and used as a starting tree for Bayesian analyses. Bayesian phylogenetic analyses and estimations of divergence times were performed simultaneously using BEAST 1.6.2 (Drummond & Rambaut, 2007). Six independent Markov chains were run for 50,000,000 generations each and sampled every 1000 generations. The initial 5000 trees were discarded as burn-in. Posterior distributions of parameter estimates and likelihood scores were examined using TRACER 1.5 (Rambaut & Drummond, 2009). All six sets of trees converged on the same stationary distribution and were used in further analyses. These sets of trees were combined and resampled with lower frequency using a custom script. The resulting subset of 10,800 trees was summarized using

TreeAnnotator onto a single target tree obtained from the analysis. The node heights were set to their median values in the summary tree.

Dating of the Apiaceae tree poses problems because there are no macrofossils for tree calibration. The family has been included in several broad-scale studies (Bremer *et al.*, 2004; Moore *et al.*, 2007; Magallón & Castillo, 2009; Bell *et al.*, 2010) and estimates from those studies may be considered as secondary calibration points. However, the estimated age of Apiaceae differs greatly among them, ranging from 29 Myr (Bell *et al.*, 2010) to 70 Myr (Bremer *et al.*, 2004). Fossil pollen provisionally attributed to Apiaceae is known from Late Cretaceous and Tertiary deposits (Gruas-Cavagnetto & Cerceau-Larrival, 1984; Farabee, 1991; Sancay *et al.*, 2006). Particularly, Gruas-Cavagnetto & Cerceau-Larrival (1984) reported pollen attributed to *Bupleurum* from the Thanetian, Ypresian, Bartonian and Priabonian ages, and pollen pertaining to Pleurospermeae from the Ypresian, Bartonian and Priabonian ages. However, in our opinion, only the fossils from the Priabonian were preserved well enough to allow identification. Additionally, fossil pollen of higher apioids was reported from the Oligocene–Miocene boundary (Sancay *et al.*, 2006). This pollen may be attributed to a monophyletic group comprising Scandiceae, Smyrnieae, Aciphylleae and the *Acronema* clade (Ł.B. & K.S., unpublished data). These three microfossils were used as calibration points (Table 1). The first calibration point was placed at the stem node of Bupleureae and constrained to a lognormal distribution with a lower bound (offset) of 33.90 Ma (the end of the Priabonian). Instead of choosing arbitrary parameters, we adjusted the mean in order to set the upper bound of distribution (95% quantile) to 58.7 Ma, corresponding to the beginning of the Thanetian and the first occurrence of presumed (but not confirmed in formal analysis) *Bupleurum*-type pollen (Gruas-Cavagnetto & Cerceau-Larrival, 1984).

Table 1 Temporal constraints used to estimate divergence times for euapioid umbellifers (in millions of years ago). Standard deviation of lognormal distribution was set to 0.5 for all calibration points.

Stem node (fossil reference)	Offset	Mean	Median	95% quantile
Bupleureae (Gruas-Cavagnetto & Cerceau-Larrival, 1984: Plate VII, Figs 13–15)	33.90	2.389	44.80	58.71
Pleurospermeae (Gruas-Cavagnetto & Cerceau-Larrival, 1984: Plate IX, Fig. 9)	33.90	2.264	43.52	55.80
Scandiceae, Smyrnieae, Aciphylleae and *Acronema* clade (Sancay *et al.*, 2006: Plate 9, Fig. 14)	23.03	1.0	25.75	29.22

Similar constraints were imposed for the stem node of Pleurospermeae, but the upper bound was constrained to the beginning of the Ypresian (55.8 Ma). For the third calibration point, the lower bound of the lognormal distribution was set at the Oligocene–Miocene boundary (23.03 Ma); arbitrary values of 1.0 and 0.5 for the mean and standard deviation were chosen (Table 1).

Biogeographical analyses

The reconstruction of ancestral areas was performed simultaneously with phylogenetic inference using a Markov discrete phylogeographical model with a Bayesian stochastic search variable selection (BSSVS) extension (Lemey *et al.*, 2009) and using the default parameters, as implemented in beast (Drummond & Rambaut, 2007). A migration graph was constructed by identifying non-zero diffusion rates with a Bayes factor (BF) test (Lemey *et al.*, 2009). We considered only rates yielding BF > 3, which is generally regarded as significant. For further description of the method and its assumptions and how this method differs from more familiar standard character optimization or explicitly biogeographical methodologies, see Lemey *et al.* (2009).

Thirteen biogeographical areas were considered (Fig. 1a) based on Takhtajan's (1986) floristic regions of the world, with some changes reflecting major centres of umbellifer endemism (Pimenov & Leonov, 1993). These areas include: (A) western Eurasia, excluding the Mediterranean coast but including the Caucasus and the Black Sea coast of Turkey; (B) northern Asia (Siberia); (C) the Mediterranean Region with Macaronesia and the Saharo-Arabian deserts; (D) the Irano-Turanian Region; (E) the Eastern Asiatic Region; (F) sub-Saharan Africa with Madagascar and the Asian part of the Sudano-Zambezian Region; (G) the Indian and Indochinese regions; (H) the Malesian Region and the Pacific archipelagos; (I) the Australian Kingdom; (J) the Neozeylandic Region of the Antarctic Kingdom; (K) the Nearctic, i.e. North America with the exclusion of Neotropical biota of Central America; (L) the Neotropical Kingdom; and (M) the Neantarctic, i.e. Chile-Patagonian and Fernandezian regions of the Antarctic Kingdom. The more detailed geographical division of the Old World adopted in this study reflects a greater diversity of apioid umbellifers in this region (Fig. 1b). This division matches geographical distributions of the species very well, as 91% of all species occurred in only one area each. Because Bayesian biogeographical analyses with a Markov model do not allow multiple areas to be assigned to terminal taxa, we assigned the area in which the sequence voucher specimen was collected for those widespread species. Therefore, we treated these accessions as representatives of a local taxon instead of a widespread one (see Appendix S1 for accession assignments).

Beast analyses resulted in 10,800 chronograms (after discarding the initial burn-in trees and resampling with lower frequency) with internal nodes annotated with ancestral area reconstructions. These trees were subsequently used in the
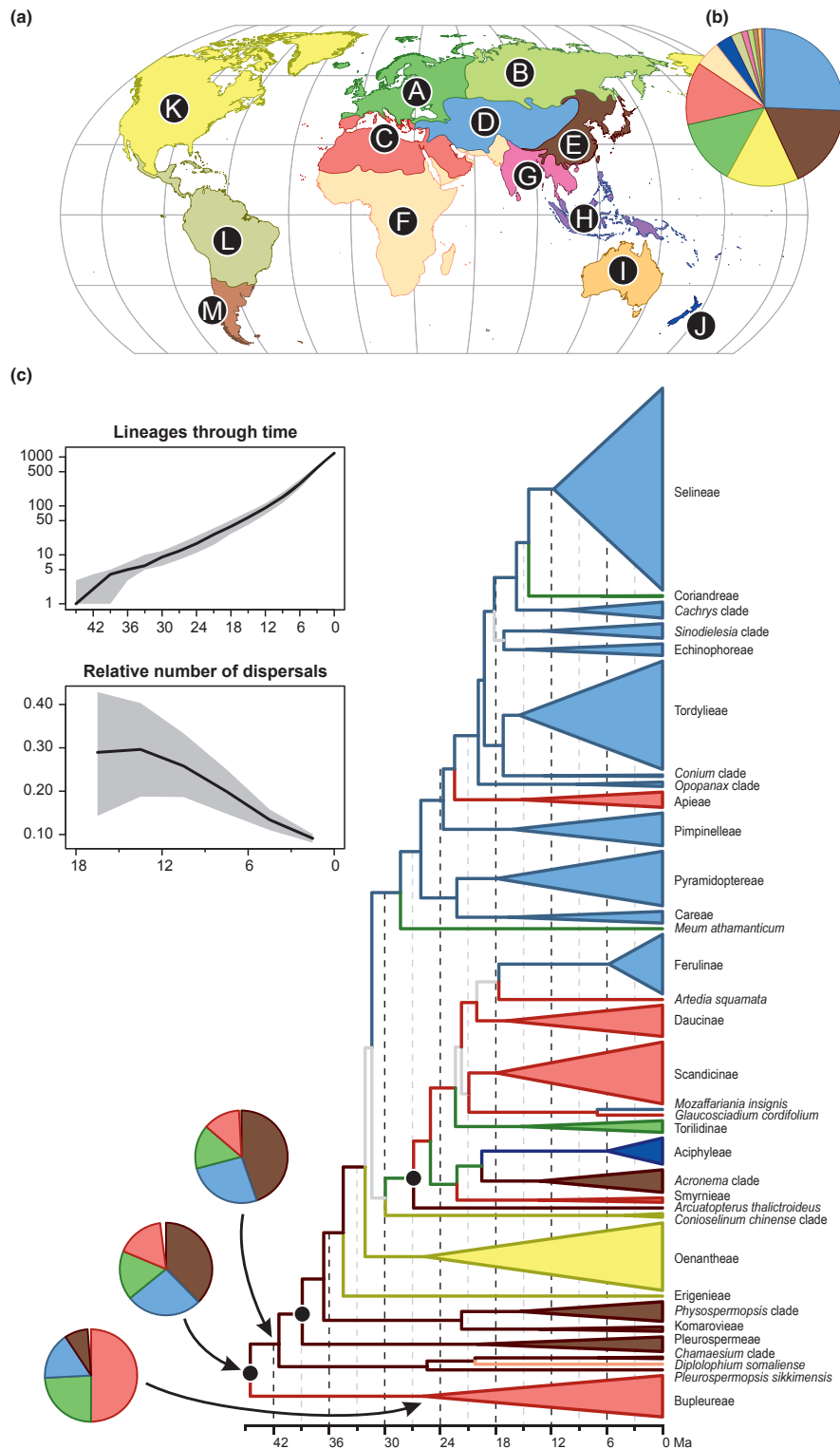
Figure 1 (a) Floristic divisions adopted in this study. The areas are based on Takhtajan (1986) with minor changes reflecting the centres of endemism of umbellifers (Pimenov & Leonov, 1993). (b) Proportion of species sampled from particular geographical areas. (c) Simplified chronogram of 1194 nrDNA ITS euapioid sequences calibrated with Tertiary fossil pollen. The calibration points are marked with black circles. The scale axis is in millions of years ago (Ma). Triangles represent tribes and major clades of euapioids; the width of each triangle is proportional to the number of taxa sampled. Colours denote the most probable ancestral area reconstructed using the Bayesian method. Branches marked in grey have posterior probabilities below 50%. Pie charts indicate areas included in a 95% Bayesian confidence interval of ancestral state reconstructions for indicated nodes. Insets show cumulative log-number of lineages through time and the number of dispersals relative to the number of branches estimated for 3-Myr intervals.

study of dispersal patterns in time and space. For the analysis of dispersal windows, each dated tree was divided into 3-Myr intervals starting from the tips and the following data were recorded for each interval: the number of branches, the number of overall reconstructed dispersals, and the number of dispersals in both directions along tracks with significant non-zero rates. Both median and 95% highest posterior density (HPD) for the number of dispersals relative to the number of branches were estimated for the time span of 0–18 Ma, but not for other time spans, because for earlier intervals the median of the number of branches was below 30. When estimating the timing of dispersals we assumed that they took place in the stem nodes of the respective branches (i.e. they were associated with allopatric speciation).

For each of the tracks with a significant non-zero rate, a dominant direction of dispersals over all trees was established. Subsequently, for each track and each tree, a dispersal asymmetry index (DAI) was calculated. We define this index as:

$$DAI = [2d_1/(d_1 + d_2)] - 1$$

where d_1 is the number of dispersals in the dominant direction and d_2 is the number of dispersals in the opposite direction along the analysed track. The DAI may vary between 1 (all dispersals inferred for a given tree are in the dominant direction) and −1 (all dispersals for a given tree are in the opposite direction), whereas its mean for the whole set of trees varies between 0 (symmetrical bidirectional dispersal) and 1 (unidirectional dispersal). When calculating the mean DAI, we treated all trees as equally probable. We considered an exchange along a particular track as significantly asymmetrical if 95% of DAIs were larger than zero (i.e. if the one-tailed 95% posterior density interval did not include zero). The calculations were performed using the package cot written in R (R Development Core Team, 2010). Additionally, in order to verify cot results and to demonstrate capabilities of distributed computing, a Hadoop MapReduce implementation, CotMR, was developed. Both implementations will be presented in a separate publication (Ł.B., T.U., M.P. & K.S., in prep.).

RESULTS

Phylogeny and reconstruction of ancestral areas

Concatenated ITS1 and ITS2 sequences ranged from 404 to 448 bp. The total number of aligned positions was 629, with 88 of these positions constant and 541 variable. The uncorrected pairwise distance between aligned sequences (P-distance) ranged from 0.000 to 0.476, with a mean value of 0.219. Modeltest with the Akaike information criterion (AIC) selected the GTR+G+I model of nucleotide substitution as fitting these data best. Phylogenetic analyses of this matrix resulted in trees comprising hitherto recognized monophyletic tribes and informal clades of subfamily

Apioideae (Fig. 1c, Appendix S2; an annotated tree with clade support, ancestral area reconstructions, node ages and 95% HPD intervals was deposited in TreeBASE as study number S13377), although the relationships among several clades and within those clades differed from those of previous analyses. Nevertheless, these differences concerned nodes that were poorly supported in previous analyses and the trees herein are consistent with the consensus tree presented in Downie *et al.* (2010). The accumulation of lineages through time is practically log-linear and the relative number of dispersals decreases with time (insets in Fig. 1c).

The divergence between tribe Bupleureae and the remaining euapioids (i.e. core euapioids) was estimated at 44.51 Ma, with a 95% HPD interval of 39.11–51.55 Ma. All major clades of euapioids originated and diversified between 25 and 15 Ma (Fig. 1c). The Eastern Asiatic Region was reconstructed as the most probable ancestral area for the root of the tree; however, the Mediterranean Region, western Eurasia and the Irano-Turanian Region were also included in the 95% Bayesian confidence interval (pie charts in Fig. 1c). For the two primary branches, the Mediterranean Region was ancestral for Bupleureae, whereas the Eastern Asiatic Region was the ancestral area for core euapioids. An eastern Asian origin was also inferred for several early-branching lineages of core euapioids (viz. the *Chamaesium* clade, *Pleurospermopsis sikkimensis*, Pleurospermeae and the *Physospermopsis* clade). However, the Irano-Turanian and Mediterranean regions and western Eurasia were also included in this 95% confidence interval. From among the 34 major lineages of euapioids considered in this study, 12 originated in the Irano-Turanian Region, seven in the Mediterranean Region, seven in the Eastern Asiatic Region, three in western Eurasia, three in North America, one in New Zealand, and one in sub-Saharan Africa.

Dispersal routes and asymmetry of floristic exchange

Out of 78 possible migration tracks, only 17 have non-zero rates supported by BF > 3 and were included in the migration graph (Fig. 2a, Appendix S3). Twelve of these tracks represent inland dispersals, mostly within the Old World, or relatively short-distance dispersals across seas, such as between Australia and New Zealand or between Australia and the Malesian Region. The long-distance transoceanic routes include: (1) western Eurasia–North America; (2) Eastern Asiatic Region–North America; (3) Australia–Neotropical Kingdom; (4) Australia–Neantarctic; and (5) Neotropical Kingdom–sub-Saharan Africa. From among the inland routes, three long-distance dispersal tracks are noteworthy: (1) western Eurasia–Eastern Asiatic Region omitting Siberia; (2) western Eurasia–sub-Saharan Africa, jumping over the Mediterranean; and (3) Nearctic–Neantarctic, passing by tropical South America. In contrast, some short-distance migration routes, i.e. those between adjacent areas, appear to be of lesser importance and therefore were not included in the dispersal graph. The Eastern Asiatic and
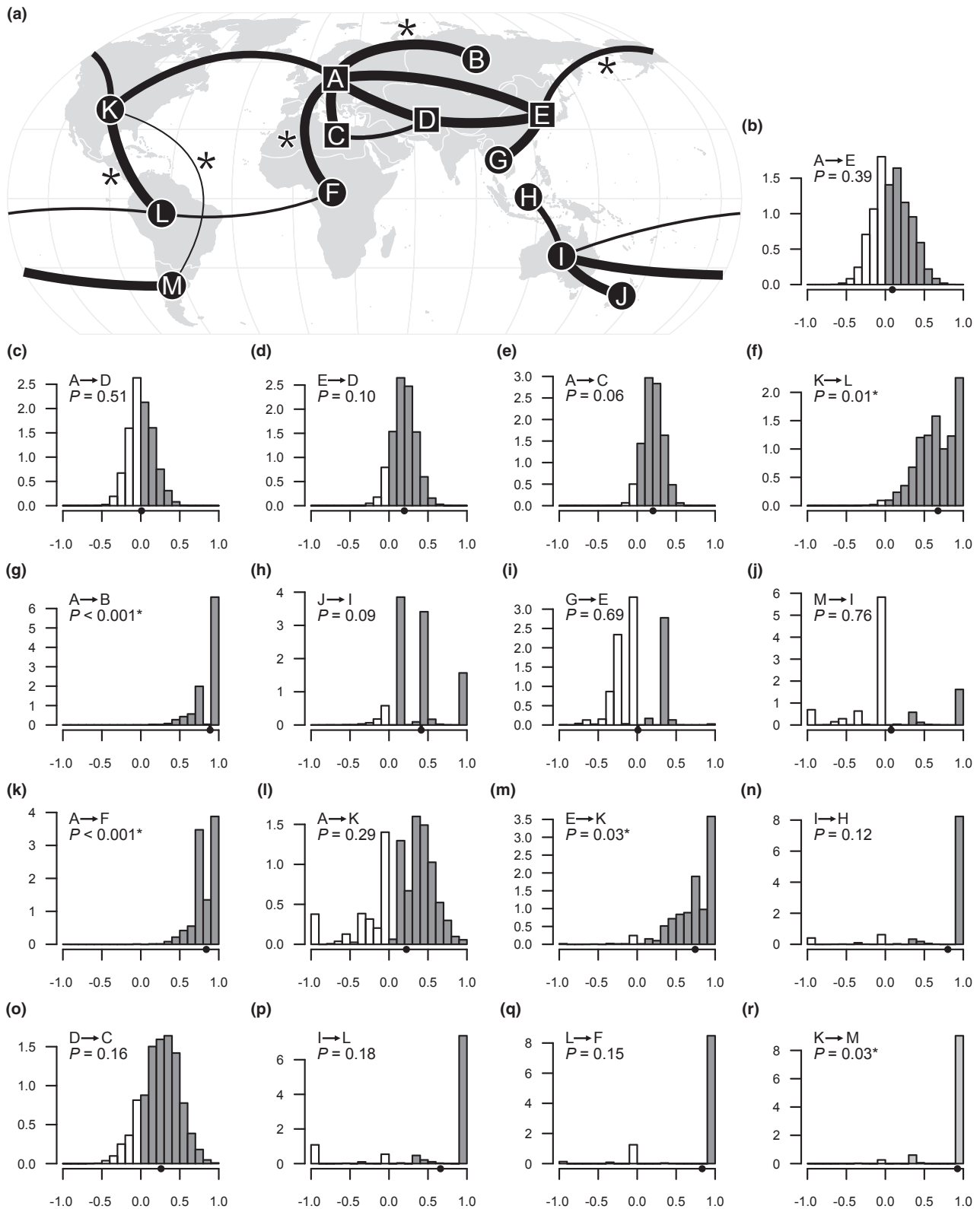
Figure 2 (a) Migration graph of euapioid umbellifers inferred in this study. Only edges with rates supported by Bayes factors greater than 3 are indicated. The thickness of the line represents the relative strength by which the rates are supported. Squares mark the areas included within a 95% Bayesian confidence interval for the ancestral area at the root of the euapioids. Asterisks indicate tracks with significantly asymmetrical floristic exchange. (b–r) Posterior density histograms of dispersal asymmetry indices (DAI). Charts are in decreasing order of Bayes factor. For each histogram, a prevailing direction of dispersal is indicated and the P-value of the asymmetry test is given. Mean DAI values are indicated by dots on the x-axis.

Irano-Turanian regions did not contribute much to the colonization of adjacent Siberia. Floristic links between sub-Saharan Africa and the Mediterranean Region or Central Asia are not significant. Within South America, the exchange between the tropical and temperate parts of the continent, i.e. Neotropical Kingdom and Neantarctic, was unimportant. The Indian and Indochinese regions are connected with the Eastern Asiatic Region but not with the neighbouring Irano-Turanian Region or with sub-Saharan Africa, comprising the Asian portion of the Sudano-Zambezian Region.

The median number of reconstructed dispersals was 334 (95% HPD: 309–358); 90.9% of these dispersals took place along the 17 paths included in the diffusion model. Most dispersal events occurred among the four Eurasian areas that constitute the major diversification centres of apioid umbellifers (the Mediterranean, Irano-Turanian and Eastern Asiatic regions and western Eurasia; Fig. 3a). Floristic exchange between these regions accounted for 72.7% of the total number of dispersal events. The most frequently used migration routes include those between the Mediterranean Region and adjacent western Eurasia (23.8% of reconstructed dispersals), between the Irano-Turanian Region and western Eurasia (19.5%), and between the Irano-Turanian and Eastern Asiatic regions (12.9%). The floristic exchange among these four areas occurred in both directions, with mean dispersal asymmetry indices for each pair within the range 0.01–0.26. Dispersal asymmetry was not significant (Fig. 2b–e,o), although only barely so for the track between western Eurasia and the Mediterranean (Fig. 2e, $P = 0.06$). The exchange between western Eurasia and Siberia accounted for 2.1% of the total number of dispersals and was strongly and significantly asymmetrical (Fig. 2g; mean DAI = 0.89; $P < 0.001$), with prevailing dispersals from Europe to Siberia. Similarly, the dispersals from Europe to sub-Saharan Africa were significantly more numerous than those northwards (Fig. 2k; mean DAI = 0.84; $P < 0.001$).

The dispersal events between western Eurasia and the Nearctic accounted for only 2.5% of the total floristic exchange, with a prevalence of dispersals from Europe to North America; this asymmetry was, however, not significant (Fig. 2l; mean DAI = 0.22; $P = 0.29$). An alternative migration route across the Pacific accounted for 2.5% of inferred dispersals. The exchange across the Pacific was significantly asymmetrical, with the dominance of dispersals from eastern Asia to North America (Fig. 2m; mean DAI = 0.75; $P = 0.03$). The exchange between the Nearctic and the Neotropical Kingdom was significantly asymmetrical, with the dominant direction from north to south (Fig. 2f; mean DAI = 0.68; $P = 0.01$). A similar asymmetry was inferred for the track between the Nearctic and Neantarctic (Fig. 2r; mean DAI = 0.93; $P = 0.03$). The exchange between the Neotropical Kingdom and sub-Saharan Africa was not significantly asymmetrical (Fig. 2q; mean DAI = 0.83; $P = 0.15$), with the smallest number of dispersals (representing only 0.3% of the total number of events).

The dispersals to and from Australia made up 3.0% of all inferred dispersal events. The floristic exchange between Australia and New Zealand and between Australia and Patagonia occurred in both directions, with mean dispersal asymmetry indices of 0.41 and 0.07, respectively (Figs 2h,j). Mean DAIs for the migration routes connecting Australia with the Malesian Region and Neotropics have higher values (0.80 and 0.66, respectively), but dispersal asymmetry was not significant for any track to or from Australia. The median number and 95% HPD of inferred dispersals and mean DAI values are given in Appendix S3.

Temporal pattern of dispersals

The temporal distributions of dispersals along significant migration routes are presented in Fig. 3. Continuous floristic exchange without any distinct dispersal windows occurred between western Eurasia and the Irano-Turanian Region (Fig. 3c), between the Mediterranean Region and western Eurasia (Fig. 3e), and between the Mediterranean and Irano-Turanian regions (Fig. 3o). These areas comprise three of the four major areas of endemism of umbellifers in the Old World where nearly half of the species included in this study occur. The floristic exchange between eastern Asia and, respectively, western Eurasia (Fig. 3b) and Central Asia (Fig. 3d) occurred throughout the analysed period of time but was most intense during the mid- to late Miocene. More or less distinct dispersal peaks in the mid- to late Miocene are also visible for the exchange between the Eastern Asiatic and the Indian and Indochinese regions (Fig. 3i), between the Old World and North America (Figs 3l,m) and between western Eurasia and sub-Saharan Africa (Fig. 3k). The main peak of dispersals from temperate North America to the Neotropical Kingdom occurred 18–15 Ma (Fig. 3f), whereas the dispersal of umbellifers from North America to southern South America started at the Miocene–Pliocene boundary (Fig. 3r). The colonization of Siberia from Europe intensified *c.* 9 Ma (Fig. 3g). A distinct increase in the relative number of dispersals beginning from the Miocene–Pliocene boundary occurred for the tracks connecting Australia with New Zealand (Fig. 3h), the Neantarctic (Fig. 3j), the Neotropical Kingdom (Fig. 3p) and the Malesian Region (Fig. 3n). A direct connection between the Neotropical Kingdom and sub-Saharan Africa is a very recent route, active during the Quaternary, that was used only in the biogeographical history of *Lilaeopsis* (Fig. 3q).

DISCUSSION

Examining the pattern of dispersals in space and time: possible limitations of the analyses

In this study, we show that dispersals within extensive land masses, such as within Eurasia and particularly those extending through one main climatic zone, are more numerous, more symmetrical, and occur more or less continuously in
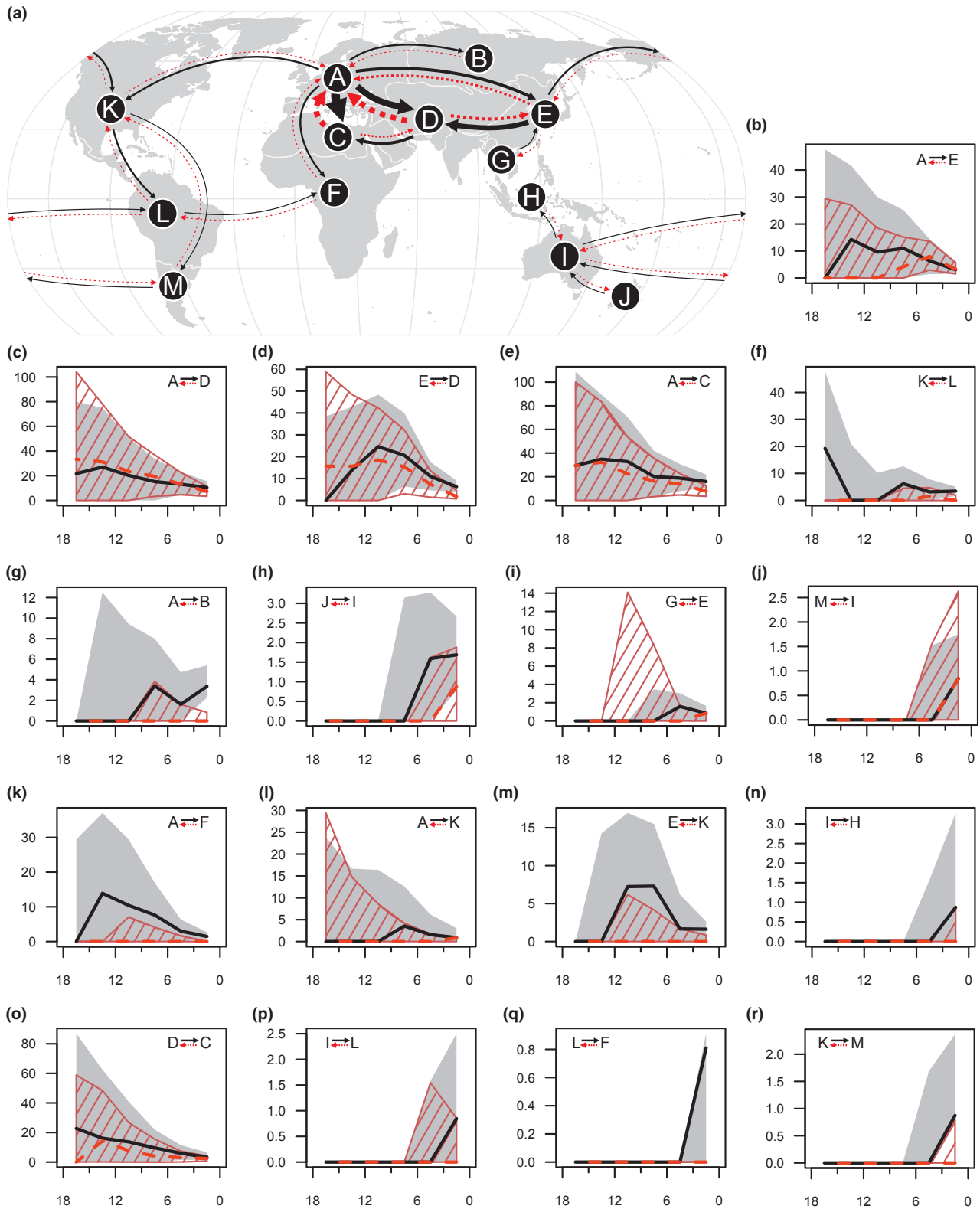
Figure 3 (a) Migration graph of euapioid umbellifers inferred in this study. Line thickness reflects the number of reconstructed dispersals; arrows indicate the direction. (b–r) Temporal distribution of inferred dispersal events relative to the number of branches in 3-Myr intervals. Charts are in decreasing order of Bayes factor. Percentile median (lines) and 95% highest posterior density intervals (polygons) are given. The prevailing direction of dispersal is marked with a solid black line and grey shading while the opposite is indicated with a dashed red line and red hatching.

time, whereas transoceanic dispersals or those between different climatic zones, particularly between temperate and tropical climates, are less numerous, often more asymmetrical, and tend to exhibit dispersal windows.

The Markov discrete phylogeographical model used in this study is particularly useful for the analysis of large data sets, but it has limitations. Contrary to the dispersal–extinction–cladogenesis (DEC) model (Ree *et al.*, 2005; Ree & Smith, 2008) or dispersal–vicariance analysis (DIVA; Ronquist, 1997), it explicitly excludes widespread ancestors, requiring that broad distributions of extant taxa be trimmed to single areas. However, widespread species are uncommon among umbellifers, with only 9% of species included in this study occupying more than one area. Moreover, unless these species are recent invaders, they often show distinct phylogeographical structure and sometimes constitute vicariant species or subspecies, as exemplified by our studies on the umbellifer genera *Berula* and *Sium* (Spalik & Downie, 2006; Spalik *et al.*, 2009). With broadly defined individual areas, colonization of a new area results in immediate genetic isolation by distance, i.e. dispersal is immediately followed by divergence. Such an assumption was also adopted in a Bayesian approach to island biogeography (Sanmartín *et al.*, 2008). The advantage of the BSSVS extension is that it decreases the number of estimated parameters and the variance of estimates. However, by putting a strong prior in favour of any given transition rate being zero, it may underestimate the number of inferred tracks. In addition, the cut-off value of Bayes factor test is arbitrary: any value higher than 2 may be chosen. We chose the value of 3 because it was the highest value for which the migration graph remained coherent. Despite these limitations, Bayesian analysis incorporating BSSVS is a valuable tool for biogeographical inference as it allows the analysis of extensive data sets and provides statistical rather than solely descriptive answers.

The origin and early diversification of Apioideae

Apiaceae subfamily Apioideae originated in southern Africa, as suggested by the present distribution of early-branching apioids or protoapioids (excluded from our analyses because of ambiguous ITS sequence alignment) and biogeographical reconstructions (Calviño *et al.*, 2006). Protoapioids mostly occur in the subtropical climatic zone and include several woody species. In contrast, euapioids comprise mostly herbaceous plants of the temperate climatic zone, and are especially abundant in arid regions (Plunkett *et al.*, in press). The adaptation of a euapioid ancestor to the temperate climatic zone that took place in the early to mid Tertiary somewhere in the Northern Hemisphere (see Fig. 1c) was a key change determining the subsequent evolutionary success of umbellifers. The occurrence of presumed protoapioid pollen in early Tertiary deposits of Europe has been explained by reference to tropical climatic conditions; the subsequent demise of these plants and the appearance of euapioid pollen

are suggestive of a successive climatic cooling (Gruas-Cavagnetto & Cerceau-Larrival, 1984).

As shown by our study, the radiation of euapioids occurred in Eurasia, with Bupleureae originating in Europe (an archipelago during the early Tertiary) and core euapioids diversifying in eastern Asia. However, when elucidating dispersal scenarios it is important to consider not only the posterior probabilities for state reconstructions at a given node, but also the dominant reconstructions throughout the phylogeny (Lemey *et al.*, 2009). Therefore, the location of the root of euapioid umbellifers in the Eastern Asiatic Region should be treated with caution, as all four major centres of umbellifer diversity are equally good candidates (Fig. 1c). Early-branching euapioids are absent from the Southern Hemisphere and the migration of euapioids to the south occurred during the mid- to late Tertiary (Spalik *et al.*, 2010; and this study). This coincides with the cooling of the global climate (particularly the onset of permanent glaciation in the Antarctic; Zachos *et al.*, 2001) and an increase in temperate and dry biomes in the Southern Hemisphere (Palazzesi & Barreda, 2007; Salzmann *et al.*, 2008).

Given the relatively stable ecological characteristics of euapioids across the phylogeny, the observed pattern of dispersals is expected. Limited dispersal between different climatic regimes relative to those within climatic regimes results from niche conservatism (Wiens & Graham, 2005). A recent study of ecologically diverse samples of more than 11,000 plant species from the Southern Hemisphere demonstrated that the tendency for species to retain their ancestral ecology over tens of millions of years and across continents is a common phenomenon, both for speciation events within land masses and those following transoceanic colonizations (Crisp *et al.*, 2009). The absence of early-branching euapioids from most of the Southern Hemisphere may, therefore, be attributed to a lack of climatic suitability throughout most of the Tertiary rather than to limited dispersal opportunities.

Floristic exchange between Eurasia and North America

Plant dispersals between Eurasia and North America occur both over the Atlantic and the Pacific. It is generally assumed that the most intense migrations took place when these continents were connected by land bridges: the North Atlantic land bridge (NALB) and Beringia (Tiffney & Manchester, 2001). The NALB functioned throughout the early Eocene with gradually decreasing dispersal opportunities later in the Tertiary as the land connections turned into an archipelago and the climate became cooler and more seasonal. Beringia functioned throughout the Tertiary and during the glacial period it was a major refuge of temperate biota. It was traditionally regarded as a major dispersal track, and this belief was supported by numerous cases of eastern Asian–North American plant disjunctions. This disjunction pattern, initially recognized in the 19th century, has been since intensely studied (reviewed by Wen, 1999). For Apioideae, the trans-Atlantic connection was

equally important as the trans-Pacific one. For both tracks, the prevailing direction of dispersal was from the Old World to the New World, although this asymmetry was significant only for the eastern Asia–North America exchange.

Nearctic, Neotropical and Holantarctic kingdoms

South America was isolated from North America throughout most of the Tertiary. The Panama Isthmus formed around 3 Ma, starting the Great American Biotic Exchange, but migrations through island-hopping may have started much earlier. Cody et al. (2010) reviewed migration times estimated on the basis of molecular phylogenies for several American taxa and concluded that plants dispersed earlier than animals. They reported that one-third of examined plant migrations have stem-node divergence dates that are over 20 Ma, whereas all animals have stem-node divergence dates that are below 20 Ma. Our data also support an early dispersal of plants to the south, preceding the establishment of the land connection by a long time. However, in the context of the Great American Biotic Exchange, the Nearctic and the Neotropical floristic kingdoms are not equivalent to North America and South America, respectively. The majority of Neotropical umbellifers occur in Central America rather than in South America; therefore, their occurrence in the Neotropical Kingdom reflects an adaptation to warmer climatic conditions or the colonization of mountain enclaves of cooler habitats instead of crossing a biogeographical barrier. The exception is Lilaeopsis, which includes some true South American species. Detailed studies of this genus revealed multiple dispersals between South America and North America during its recent history (Spalik et al., 2010; Bone et al., 2011). Lilaeopsis comprises true aquatics and dispersal by migratory waterfowl might have been responsible for its broad distribution.

Dispersal by birds has also been proposed to explain the amphitropical disjunctions of several species between temperate North America and South America (Raven, 1963). Our results corroborate Raven's (1963) hypothesis, suggesting a northern temperate and relatively recent origin of these taxa. Traditional phytogeographical studies placed the southern part of South America, the Juan Fernández Archipelago and New Zealand in the Holantarctic Kingdom. Such links were also evident in this study, as well as in our earlier studies of disjunct Southern Hemisphere euapioids, including all euapioid lineages occurring at present in Australasia (Spalik et al., 2010). The almost concurrent migration of euapioid umbellifers to Australasia and the Neantarctic suggests that the major factor determining their success in the Southern Hemisphere was the increasing availability of temperate habitats resulting from the climatic cooling in the late Tertiary.

Connections between western Eurasia and northern and eastern Asia

One interesting pattern to emerge in this study is a long-distance track between western Eurasia and the Eastern Asiatic

Region, excluding the connecting areas of Siberia and the Irano-Turanian Region. Siberia (or northern Asia) was only recently colonized from the west rather than from the south. Such a pattern is in agreement with commonly adopted biogeographical divisions; for instance, Takhtajan (1986) included temperate Europe and Siberia in the Circumboreal Region, as opposed to the Irano-Turanian and Eastern Asiatic regions. The connection between the Eastern Asiatic Region and western Eurasia may be relictual as these areas constitute two of four major centres of temperate deciduous forest diversity, traditionally regarded as remnants of a mesophytic boreotropical flora that dominated the Northern Hemisphere in the early Tertiary (Donoghue & Smith, 2004). This direct link may therefore have resulted from the extinction of some taxa in connecting areas (Siberia, Irano-Turanian Region) instead of from long-distance dispersals. Because of the greater impact of glaciations in Europe than in eastern Asia, the European survivors were probably better adapted to colder climatic conditions than the Asian ones and rapidly colonized the eastern Palaearctic when the glaciers retreated. Long distance was therefore a less important barrier for these species than ecological constraints.

CONCLUSIONS

Although Apiaceae subfamily Apioideae is probably of southern African origin, the euapioids originated and diversified during the early Tertiary in Eurasia, particularly in western Eurasia and the Mediterranean, Irano-Turanian and Eastern Asiatic regions. The floristic exchange among these regions was symmetrical and more or less continuous over time. The migration of umbellifers to North America occurred throughout the late Tertiary (particularly the Miocene), partly coincided with the existence of the North Atlantic Land Bridge and Beringia, and was significantly asymmetrical along at least one route (Beringia). This pattern suggests that geographical barriers were key factors limiting dispersal in the Northern Hemisphere.

With the exception of the colonization of sub-Saharan Africa, which started in the mid-Miocene, the migration to the Southern Hemisphere (Australasia, Neantarctic) intensified from the Miocene–Pliocene boundary. Euapioids are mostly plants of open habitats of the temperate climatic zone, and the observed dispersal pattern – intense and symmetrical within the same climatic zone, but scarce and often unidirectional between climatic zones – suggests that the increasing availability of suitable temperate habitats at the end of Tertiary has been a major factor determining their concurrent colonization of the Southern Hemisphere.

ACKNOWLEDGEMENTS

REFERENCES

Bell, C.D., Soltis, D.E. & Soltis, P.S. (2010) The age and diversification of the angiosperms re-revisited. *American Journal of Botany*, **97**, 1296–1303.

Bone, T.S., Downie, S.R., Affolter, J.M. & Spalik, K. (2011) A phylogenetic and biogeographic study of the genus *Lilaeopsis* (Apiaceae tribe Oenantheae). *Systematic Botany*, **36**, 789–805.

Bremer, K., Friis, E.M. & Bremer, B. (2004) Molecular phylogenetic dating of asterid flowering plants shows Early Cretaceous diversification. *Systematic Biology*, **53**, 496–505.

Calviño, C.I., Tilney, P.M., van Wyk, B.-E. & Downie, S.R. (2006) A molecular phylogenetic study of southern African Apiaceae. *American Journal of Botany*, **93**, 1828–1847.

Cody, S., Richardson, J.E., Rull, V., Ellis, C. & Pennington, R.T. (2010) The Great American Biotic Interchange revisited. *Ecography*, **33**, 326–332.

Cook, L.G. & Crisp, M.D. (2005) Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. *Journal of Biogeography*, **32**, 741–754.

Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, **458**, 754–756.

Donoghue, M.J. & Smith, S.A. (2004) Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1633–1644.

Downie, S.R. & Katz-Downie, D.S. (1999) Phylogenetic analysis of chloroplast *rps*16 intron sequences reveals relationships within the woody southern African Apiaceae subfamily Apioideae. *Canadian Journal of Botany*, **77**, 1120–1135.

Downie, S.R., Spalik, K., Katz-Downie, D.S. & Reduron, J.-P. (2010) Major clades within Apiaceae subfamily Apioideae as inferred by phylogenetic analysis of nrDNA ITS sequences. *Plant Diversity and Evolution*, **128**, 111–136.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Farabee, M.J. (1991) Botanical affinities of some Triprojectacites fossil pollen. *American Journal of Botany*, **78**, 1172–1181.

Gruas-Cavagnetto, C. & Cerceau-Larrival, M.-T. (1984) Apports des pollens fossiles d'Ombellifères à la connaissance paléoécologique et paléoclimatique de l'Éocène français. *Review of Palaeobotany and Palynology*, **40**, 317–345.

Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

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

Lemey, P., Rambaut, A., Drummond, A.J. & Suchard, M.A. (2009) Bayesian phylogeography finds its roots. *PLoS Computational Biology*, **5**, e1000520.

Maddison, W.P. & Maddison, D.R. (2010) *Mesquite: a modular system for evolutionary analysis, version 2.73*. Available at: http://mesquiteproject.org.

Magallón, S. & Castillo, A. (2009) Angiosperm diversification through time. *American Journal of Botany*, **96**, 349–365.

Magee, A.R., Calviño, C.I., Liu, M., Downie, S.R., Tilney, P.M. & van Wyk, B.-E. (2010) New tribal delimitations for the early diverging lineages of Apiaceae subfamily Apioideae. *Taxon*, **59**, 567–580.

Moore, M.J., Bell, C.D., Soltis, P.S. & Soltis, D.E. (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences USA*, **104**, 19363–19368.

Palazzesi, L. & Barreda, V. (2007) Major vegetation trends in the Tertiary of Patagonia (Argentina): a qualitative paleoclimatic approach based on palynological evidence. *Flora – Morphology, Distribution, Functional Ecology of Plants*, **202**, 328–337.

Pimenov, M.G. & Leonov, M.V. (1993) *The genera of the Umbelliferae*. Royal Botanic Gardens, Kew, UK.

Plunkett, G.M., Pimenov, M.G., Reduron, J.P., Kljuykov, E.V., Lee, B.Y., van Wyk, B.E., Tilney, P.M., Watson, M.F., Ostroumova, T.A., Spalik, K., Hart, J.M., Henwood, M.J., Webb, C.J., Pu, F.T., Mitchell, A.D. & Muckensturm, B. (in press) Apiaceae. *The families and genera of vascular plants* (ed. by K. Kubitzky). Springer, Berlin.

Posada, D. & Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

R Development Core Team (2010) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A. (2009) *FigTree v1.3.1*. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. Available at: http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A. & Drummond, A.J. (2009) *Tracer v1.5.0*. Available at: http://beast.bio.ed.ac.uk/Tracer.

Raven, P.H. (1963) Amphitropical relationships in the floras of North and South America. *Quarterly Review of Biology*, **38**, 151–177.

Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.

Ree, R.H., Moore, B.R., Webb, C.O. & Donoghue, M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.

Ronquist, F. (1997) Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46**, 195–203.

Salzmann, U., Haywood, A.M., Lunt, D.J., Valdes, P.J. & Hill, D.J. (2008) A new global biome reconstruction and data-model comparison for the Middle Pliocene. *Global Ecology and Biogeography*, **17**, 432–447.

Sancay, R.H.N., Batı, Z., Işık, U., Kırıcı, S. & Akça, N. (2006) Palynomorph, Foraminifera, and calcareous nannoplankton biostratigraphy of Oligo–Miocene sediments in the Muş basin, eastern Anatolia, Turkey. *Turkish Journal of Earth Sciences*, **15**, 259–319.

Sanmartín, I. & Ronquist, F. (2004) Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology*, **53**, 216–243.

Sanmartín, 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.

Sanmartín, I., Wanntorp, L. & Winkworth, R.C. (2007) West Wind Drift revisited: testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. *Journal of Biogeography*, **34**, 398–416.

Sanmartín, I., van der Mark, P. & Ronquist, F. (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, **35**, 428–449.

Spalik, K. & Downie, S.R. (2006) The evolutionary history of *Sium* sensu lato (Apiaceae): dispersal, vicariance, and domestication as inferred from ITS rDNA phylogeny. *American Journal of Botany*, **93**, 747–761.

Spalik, K., Downie, S.R. & Watson, M.F. (2009) Generic delimitations within the *Sium* alliance (Apiaceae tribe Oenantheae) inferred from cpDNA *rps*16–5′*trn*K$^{(UUU)}$ and nrDNA ITS sequences. *Taxon*, **58**, 735–748.

Spalik, K., Piwczyński, M., Danderson, C.A., Kurzyna-Młynik, R., Bone, T.S. & Downie, S.R. (2010) Amphitropic amphiantarctic disjunctions in Apiaceae subfamily Apioideae. *Journal of Biogeography*, **37**, 1977–1994.

Swofford, D.L. (2003) *PAUP*: phylogenetic analysis using parsimony (*and other methods)*. Version 4.0b10. Sinauer Associates, Sunderland, MA.

Takhtajan, A. (1986) *The floristic regions of the world*. University of California Press, Berkeley, CA.

Tiffney, B.H. & Manchester, S.R. (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.

Wen, J. (1999) Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology and Systematics*, **30**, 421–455.

Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 519–539.

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 INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Accessions of Apiaceae subfamily Apioideae examined for variation in nuclear ribosomal DNA ITS sequences.

Appendix S2 Detailed chronogram of 1194 nrDNA ITS euapioid sequences.

Appendix S3 Floristic exchange along tracks with significant non-zero migration rates (Bayes factor test value > 3.0).

BIOSKETCHES

Łukasz Banasiak and **Krzysztof Spalik** work on the molecular phylogeny, biogeography and taxonomy of the family Apiaceae and the evolution of morphological traits in plants.

Author contributions: Ł.B. and K.S. conceived the ideas of the manuscript; Ł.B. and T.U. wrote the software; B.S. and M.F.W. helped with the biogeographical data; Ł.B. prepared the matrices and performed the analyses; M.P. and T.U. helped with the analyses; K.S. and Ł.B. wrote the manuscript; S.R.D. edited the manuscript.

Editor: Mark Carine