Iles et al. BMC Evolutionary Biology 2014, 14:102
http://www.biomedcentral.com/1471-2148/14/102

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Reconstructing the age and historical biogeography of the ancient flowering-plant family Hydatellaceae (Nymphaeales)

William J D Iles[1,2*], Christopher Lee[1], Dmitry D Sokoloff[3], Margarita V Remizowa[3], Shrirang R Yadav[4], Matthew D Barrett[5], Russell L Barrett[5], Terry D Macfarlane[6], Paula J Rudall[7] and Sean W Graham[1,2]

Abstract

Background: The aquatic flowering-plant family Hydatellaceae has a classic Gondwanan distribution, as it is found in Australia, India and New Zealand. To shed light on the biogeographic history of this apparently ancient branch of angiosperm phylogeny, we dated the family in the context of other seed-plant divergences, and evaluated its biogeography using parsimony and likelihood methods. We also explicitly tested the effect of different extinction rates on biogeographic inferences.

Results: We infer that the stem lineage of Hydatellaceae originated in the Lower Cretaceous; in contrast, its crown originated much more recently, in the early Miocene, with the bulk of its diversification after the onset of the Pliocene. Biogeographic reconstructions predict a mix of dispersal and vicariance events, but considerations of geological history preclude most vicariance events, besides a split at the root of the family between southern and northern clades. High extinction rates are plausible in the family, and when these are taken into account there is greater uncertainty in biogeographic inferences.

Conclusions: A stem origin for Hydatellaceae in the Lower Cretaceous is consistent with the initial appearance of fossils attributed to its sister clade, the water lilies. In contrast, the crown clade is young, indicating that vicariant explanations for species outside Australia are improbable. Although long-distance dispersal is likely the primary driver of biogeographic distribution in Hydatellaceae, we infer that the recent drying out of central Australia divided the family into tropical vs. subtropical/temperate clades around the beginning of the Miocene.

Keywords: Aquatic plants, Austral, Ephemeral habitats, Extinction rates, Intercontinental dispersal, ANITA-grade angiosperms, *Trithuria*, Vicariance

Background

Australia has seen widespread rainforest replaced with deserts, savannah and sclerophyll biomes since the Eocene, in response to global cooling [1]. Despite the dramatic loss of mesophytic habitat, it has a well-developed wetland flora, with many endemic species [2]. Perhaps the most unique of these habitats are ephemeral bodies of water that are home to communities characterized by extreme reduction in plant size, and annual or geophytic life histories [3,4]. Common Australian members of this ephemeral aquatic habitat include Centrolepidaceae (a family closely related to or possibly embedded within the southern rushes [5], Restionaceae), the sundew genus *Drosera* L. (Droseraceae), and *Hydrocotyle* L. (Araliaceae) [3,6], but its most noteworthy component may be the family Hydatellaceae [7]. Most members of Hydatellaceae exemplify the ephemeral aquatic syndrome, apart from a recently derived pair of perennial apomictic species that live submerged in more permanent bodies of water [8,9].

Hydatellaceae were recently recognized as the sister group of the water lilies (Cabombaceae and Nymphaeaceae), placing their divergence close to the root

* Correspondence: will.jd.iles@gmail.com
[1]Department of Botany, University of British Columbia, 3529-6270 University Blvd, Vancouver, British Columbia, V6T 1Z4, Canada
[2]UBC Botanical Garden & Centre for Plant Research, University of British Columbia, 6804 Marine Dr SW, Vancouver, British Columbia, V6T 1Z4, Canada
Full list of author information is available at the end of the article

BioMed Central

© 2014 Iles et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

of angiosperm phylogeny e.g., [10]. They have since attracted considerable attention because of the insights they may provide into the evolution of early angiosperms [11-16]. Of particular interest is the nature of the reproductive structures in the family, which may represent floral, prefloral, or pseudanthial arrangements of reproductive organs, and the incidence of unisexual and bisexual reproductive units. These may bear on our understanding of the ancestral floral *Bauplan* of angiosperms [8,9,13]. Contemporary taxonomic and phylogenetic work on the family recognizes one genus, *Trithuria* Hook. f., and 12 species in four monophyletic sections [8,9].

The distribution of the Southern Hemisphere biota has been traditionally framed in terms of vicariance events resulting from the breakup of Gondwana [17]. However, recent studies have suggested a more important role for long-distance dispersal, especially in plants [17]. Within Australia, the constriction and fragmentation of mesophytic and rainforest habitats since the Eocene [1] have led to congruent patterns of vicariant speciation across a number of plant lineages [18]. Hydatellaceae display a classic Gondwanan distribution, being present in Australia, India and New Zealand (Figure 1; the distribution is based on online herbarium records [19] and new collection records), which may imply a relictual intercontinental distribution and great antiquity for the crown clade [20]. While vicariant processes at both the hemispheric and continental scale could explain the extant distribution of Hydatellaceae, this distribution may also result from long-distance dispersal [9,17] (the wide and potentially trans-oceanic distribution of many aquatic plant species is widely suspected to be facilitated by the enhanced dispersal capabilities of their diaspores, often considered to be due to water-fowl vectors [21]). Evaluating these biogeographic hypotheses requires placing geological and geographical events in the context of dated phylogenies [22]. Most organisms lack an extensive fossil record, and so molecular dating analyses use age information from parts of the phylogeny with a good fossil record to inform the age of nodes that lack it [23].

Since the recognition that Hydatellaceae represents an ancient angiosperm lineage, a few fossils have been linked with it [7,24-26]. The most spectacular of these may be the aquatic plant *Archaefructus*, represented by whole fruiting plants from the Yixian Formation of Liaoning, China [27,28]. However, the timing and interpretation of these and other records remains contentious [7,29]. Unlike the fossil record of Hydatellaceae, water lilies have an extensive record that extends to the Lower Cretaceous

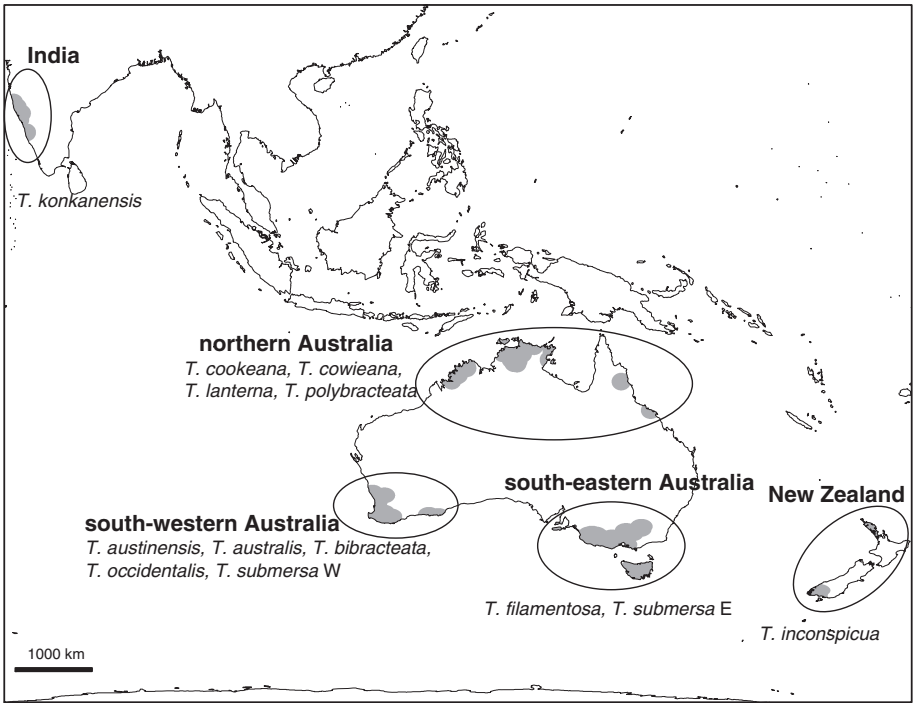

Figure 1 Global distribution of Hydatellaceae. The distribution is based on online herbarium resources [19] and new collection records; each collection locality is represented by a 120 km radial sweep. Five biogeographic areas are delineated by labelled ovals: India, New Zealand, northern Australia, south-eastern Australia, and south-western Australia. Species found in each area are listed; note that *Trithuria submersa* is treated here as south-eastern and south-western species. We created the map using ArcMAP version 9.3 (Environmental Systems Research Institute, Redlands, CA, USA); the projection is Lambert's cylindrical equal area.

[29-31]. Collectively these fossils suggest that aquatic niches were exploited early in the evolution of angiosperms, although the aquatic life-form is unlikely to be ancestral in flowering plants as a whole [32,33]. Nonetheless, an improved understanding of the diversification of Hydatellaceae may help illuminate early angiosperm ecology and how plants colonize ephemeral wetlands, which represent a unique and potentially stressful environment [4,34].

To address these questions we dated the earliest splits in Hydatellaceae using 17 plastid-genes sampled from across the seed plants, and used the resulting posterior age distributions as secondary calibrations for a species-tree analysis of the entire family, which lacks suitable fossil calibrations. Although the use of secondary calibration points has been criticized for propagating "error free" values into downstream analyses [35], here we use the entire posterior distribution from the seed-plant analysis as a prior for the subsequent analysis, accounting for the associated uncertainty [36]. We used the dated species tree to explore biogeographic hypotheses using parsimony and likelihood. In particular, likelihood-based approaches allow the estimation of parameters such as speciation rate. We also explicitly test the effect of extinction rate on biogeographic reconstruction, as this may be high in Hydatellaceae due to the patchy distribution of their habitat in space and time [4,34] and is also suggested by the "broom-and-handle" shape of the phylogeny [9,37-39].

Methods

Fossil selection and molecular dating

Fossils with unequivocal affinity to Hydatellaceae are unknown [7]. We therefore first estimated the crown-age of the family in a seed-plant analysis with 17 exemplar taxa constrained by eight fossils (Table 1; see [24,40-52]). We added *Trithuria cowieana* D.D. Sokoloff, Remizowa, T.D. Macfarl. & Rudall (*Macfarlane & al. 4217*, MW; [GenBank: JQ284074, JQ284187, JQ284224, KJ725347–KJ725349]) to an existing data set that included *T. filamentosa* Rodway and *T. submersa* Hook. f. [10,53,54] as these three species define the deepest phylogenetic splits in Hydatellaceae [9]. This seed-plant matrix includes three gymnosperms and major lineages of angiosperms [53]. Genomic sampling focused on 13 single-copy plastid genes (comprising four multi-gene clusters, *psb*D–*psb*C, *psb*E–*psb*F–*psb*L–*psb*J, *psb*B–*psb*T–*psb*N–*psb*H, and three single-gene regions, *ndh*F, *rbc*L, and *atp*B). Only protein-coding regions were considered. Details of DNA extraction, amplification, sequencing, contig assembly and alignment are described elsewhere [9,10,55,56]; for a list of accessions and the alignment matrix, see Additional files 1 and 2, respectively.

To test for and accommodate non-clocklike behaviour in the seed-plant data set we used the Bayesian random local clocks (RLC) method [57]. This accommodates molecular rate variation by allowing different sub-branches of the tree to have unique molecular clocks. Dornberg et al. [58] examined the performance of this method against the more widely used uncorrelated lognormal (UCLN) method [59] for real and simulated data sets that show high amounts of inter-clade rate variability, and found that the RLC model performed better in the presence of clade-specific rate shifts. This may be pertinent to angiosperm studies like ours, as there are known to be substantial shifts in rates among major angiosperm clades that are associated with changes in habit and life history [60]. In particular, Hydatellaceae occupy a part of the tree where there were multiple shifts in habit (for example, Hydatellaceae are mostly herbaceous annuals, water lilies are mostly perennial herbs, *Amborella* Baill. and Austrobaileyales include shrubs, small trees and lianas). The method is implemented in BEAST version 1.6.1. We used a GTR + Γ model of sequence evolution, with default priors (or those suggested by http://code.google.com/p/beast-mcmc/wiki/ParameterPriors if not automatically implemented). The BEAST analysis requires that each of the fossil calibrations have an associated prior. We used lognormal priors with 95% prior intervals of ~10–20% of the fossil age (Table 1),

Table 1 Fossil calibrations for seed-plant phylogeny

Node	Calibration	Fossil	Age (Ma)	Priors	References
A	Crown seed plants	Cordaites	315	316.0–367.5 (2, 1)	[40,41]
C	Stem angiosperms	Glossopterid, *Gangamopteris* McCoy	293.8	294.8–346.3 (2, 1)	[24,40,42,43]
G	Stem Cabombaceae	*Pluricarpellatia peltata* B. Mohr, Bernardes-de-Oliveira & D.W. Taylor	98.7*	99.1–118.0 (1, 1)	[44,45]
H	Crown Nymphaeaceae	*Monetianthus mirus* Friis, Pedersen, von Balthazar, Grimm & Crane	92.8*	93.2–112.1 (1, 1)	[46,47]
M	Stem Trimeniaceae	Unnamed seed	98.7	98.9–110.4 (0.5, 1)	[48]
N	Stem eudicots	Tricolpate pollen	124*	124.2–135.7 (0.5, 1)	[49,50]
O	Stem Araceae	*Mayoa portugallica* Friis, Pedersen & Crane	96.1*	96.5–115.4 (1, 1)	[51]
P	Stem Platanaceae	West Brothers platanoid and *Sapindopsis* Fontaine	92.8*	93.2–112.1 (1, 1)	[52]

'Node' refers to placement on phylogeny in Additional file 3. Fossil ages follow the references, except that some ages are modified following [49], indicated with an asterisk. Calibration priors are lognormal; the 95% prior interval is given, followed by the lognormal mean and standard deviation in parentheses.

consistent with some other studies, e.g., [61] (the RLC method is also more robust than the UCLN method to variation in the width of the 95% prior interval [58]). We ran seven runs of 4.0×10^7 generations, and considered four that converged on the same posterior and likelihood scores after 10% burnin. The estimated sample sizes of run statistics (posterior, prior, likelihood, parameter estimates) were all over 200 when these runs were pooled. The seed-plant chronogram and a table of divergence times is presented in Additional file 3, and the tree file is provided in Additional file 4. In all analyses we constrained Nymphaeaceae s.s. (i.e., excluding Cabombaceae) to be monophyletic, consistent with molecular and morphological analyses [62-64]. We also tested a constraint that forces *Amborella* to be the sister group of all other angiosperms; this arrangement contrasts with a clade comprising *Amborella* and Nymphaeales that we recovered in the RLC analysis, see below (these two alternative arrangements have been recovered in different studies, see [53], for example). We constrained cycads to be the sister group of angiosperms among extant seed plants, consistent with some recent studies [24,42,43,54], but also explored alternative gymnosperm sister groups to angiosperms (conifers alone, *Ginkgo* L. alone, or pairwise combinations of conifers, cycads and *Ginkgo*), or used no outgroup constraints. For these different constraint analyses we ran a single 4.0×10^7 generation replicate; they all indicated only a minimal effect on the two ages within Hydatellaceae (<1 Myr difference; data not shown).

To date the Hydatellaceae species tree we considered the data set of [9], which consists of two unlinked loci (four plastid regions and the nuclear ribosomal internal transcribed spacer region, ITS) for all species except *Trithuria occidentalis* Benth. which was only sampled for one plastid region. In all analyses *T. submersa* was provisionally considered to comprise separate eastern and western species, following [9]. The data were analysed with *BEAST, which estimates the species tree with a Bayesian implementation of the multi-species coalescent [65]. We used the settings outlined in [9], with the exception that we assigned the two Hydatellaceae posterior distributions determined from the RLC seed-plant analysis (see Additional file 3) as Gaussian priors for the corresponding splits in the species tree (i.e., the crown node of Hydatellaceae and the crown node of the clade consisting of sect. *Hydatella* and sect. *Trithuria*). These priors were only applied to the plastid loci (for which there was outgroup data), using the rooting of Hydatellaceae determined in the seed-plant analysis (see Additional file 3).

Biogeographic reconstructions

We reconstructed ancestral areas using three methods: maximum parsimony dispersal-vicariance (DIVA; [66]), maximum likelihood (ML) dispersal-extinction-cladogenesis (DEC; [67]), and ML ancestral-state reconstruction (ASR; cf. [68]). The DIVA and DEC methods allow the range of extant species and internal nodes (ancestral species) to encompass multiple discrete areas, and identify dispersal, vicariance or area extinction events. When vicariance events predicted in these analyses could not be explained by contemporaneous geographic division, we treated them as indicative of long-distance dispersal events. In contrast, ASR implicitly only considers dispersal/extinction, and restricts each species range and internal node to a single area. In all cases we used the species tree as the reference phylogeny (Figure 2a) and considered the five major biogeographic areas that define the range of Hydatellaceae: (1) India; (2) northern Australia; (3) New Zealand; (4) south-eastern Australia; (5) south-western Australia (Figure 1). We treated south-eastern Australia as one area, despite it being climatically variable, because species ranges there (*Trithuria filamentosa* and *T. submersa*) are essentially contiguous within Tasmania, and there is no evidence of a strong geographic barrier between them [8]. We used RASP version 2.0 Beta [66,69] to perform the DIVA analysis, and for DEC we used Lagrange version 20110117 (http://www.reelab.net/lagrange/configurator/index; [67]). We constrained some of the area connectivities for the DEC analysis, following the advice of the Lagrange website, which we based on our understanding of niche profiles of the species (unpublished data) and the underlying phylogenetic relationships. In particular, we constrained the connectivity of India to northern Australia only, and of New Zealand to south-eastern Australia only. To make comparisons more meaningful between methods we also constrained the DIVA analysis in the same way. In addition, for the DEC analysis we considered all dispersal paths to have the same rate (as the allowed distances are comparable in magnitude; ~1500–6500 km), so we set all values in the dispersal rate scaling matrix to 1.0 (i.e., all dispersal path rates are multiplied by "1"). Simulation studies show that area extinction rates in Lagrange are strongly biased towards zero [67]; as an alternative we evaluated lineage (species) extinction rates [70] as a proxy for the area extinction rate, the parameter used in our analyses (species extinction can be thought to occur when all the areas that encompass the species range go extinct, and may be an underestimate of the area extinction rate). We used the R package Diversitree version 0.7-2 [71,72] to evaluate extinction and speciation rates on the Hydatellaceae species tree alone, or on the species tree with the addition of one outgroup separated by 127 Myr (see below). Based on the results of the extinction-speciation analysis we chose six area extinction rates (0.001, 0.01, 0.05, 0.1, 0.5, and 1.0 Myr^{-1}) spanning the range of lineage extinction rates to explore the effect of different rates on biogeographic reconstructions. Dispersal rates were iteratively optimized in Lagrange for each extinction rate.

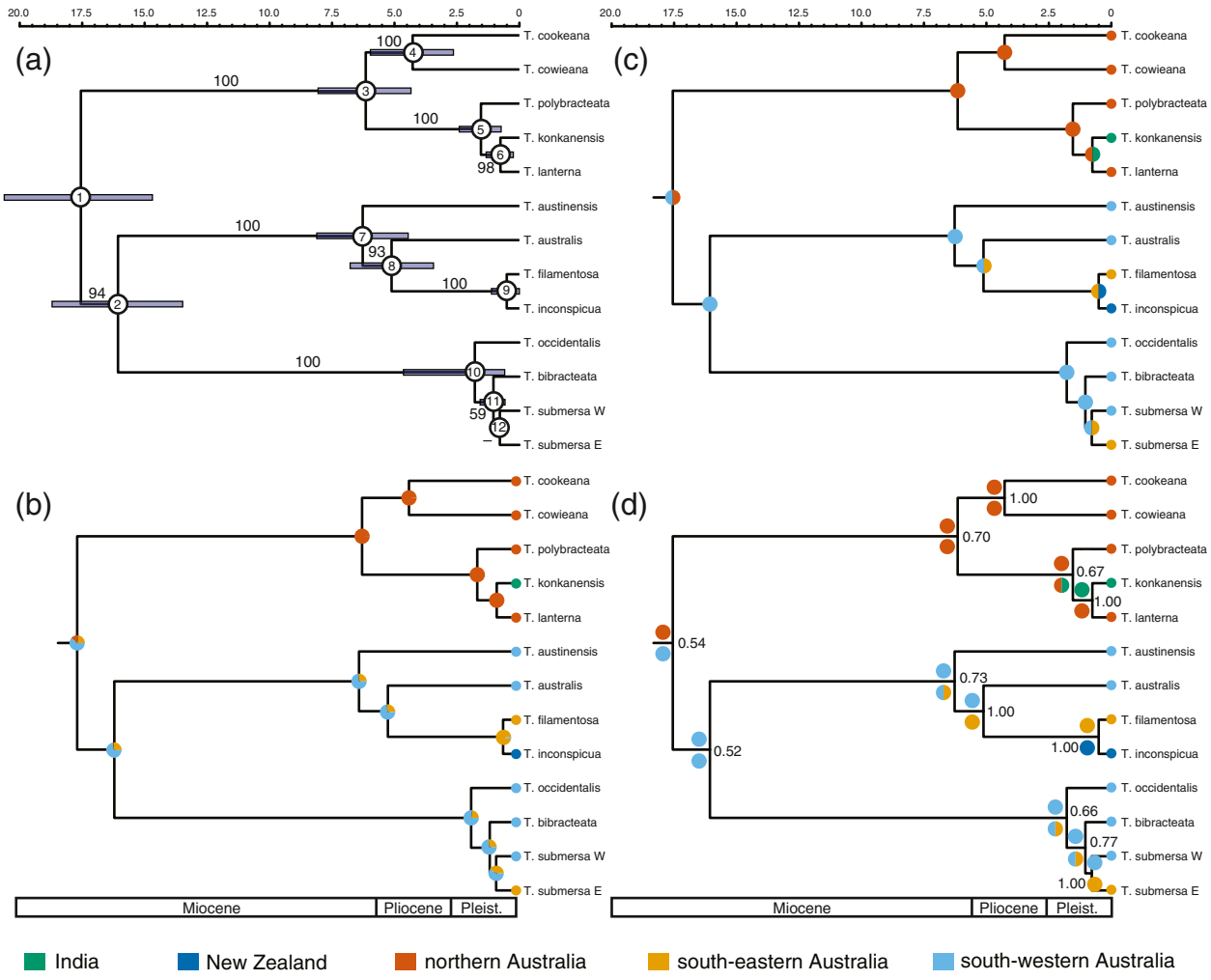

Figure 2 Dated species tree and biogeographic inferences for Hydatellaceae. (a) Timing of divergences in Hydatellaceae inferred using a multi-species coalescent analysis, based on four plastid genes and nuclear ITS regions, with prior dating estimates for the first two nodes derived from a separate seed-plant analysis (see Additional file 3; labelled nodes are referred to in Additional file 5). Numbers beside branches are support values (posterior probabilities expressed as percentages); dashes indicate <50% support. Divergence time uncertainty is noted by blue bars, representing 95% HPD. The time scale is in Ma. Letters adjacent to tips represent: E = east, T. =*Trithuria*, W = west. Historical biogeography inferred in **(b)** with the full model of ML ancestral-state reconstruction, in **(c)** with MP based dispersal-vicariance analysis, and in **(d)** with ML based dispersal-extinction cladogenesis analysis. Pie fractions in **(b)** represent relative likelihoods; in **(c)** and **(d)** they represent joint areas where the species is inferred to have existed in multiple areas. The relative likelihood of the best geographic range pair is shown in **(d)** adjacent to individual nodes.

The ASR analyses were performed with BayesTraits version 1.0 (www.evolution.rdg.ac.uk; [73]). We considered three nested models which were evaluated using the corrected Akaike information criterion (AICc; [74]). The most complex model (hereafter the 'full model') assumed three separate symmetrical transition rates: between Australia and India or New Zealand (assuming trans-oceanic dispersals to be equivalent), between south-western and south-eastern Australia (dispersals across the Nullarbor Plain), and between northern Australia and south-western or south-eastern Australia (dispersals across the arid zone). The simplest model ('simple model') consists of a single rate between all the allowed transition rates in the full model. The two-rate transition model has symmetric rates between Australia and India or New Zealand, contrasting with a separate rate for all transitions within Australia ('continental model'). Root state frequencies were set to empirical values.

Results

Molecular dating and diversification

Hydatellaceae are estimated to have diverged from the water lilies 126.7 Ma (120.6–133.2 Ma, 95% HPD), in the Lower Cretaceous, with a crown clade age of 19.1 Ma (15.7–23.4 Ma, 95% HPD), in the early Miocene (see Additional file 3). The estimated multi-species coalescent

age for the crown of Hydatellaceae is 17.6 Ma (14.7–20.6 Ma, 95% HPD), in the early Miocene, with most diversification occurring after ~6 Ma, in the late Miocene (Figure 2a; see Additional file 5 for a table of divergence times, and Additional file 6 for the tree file). For the speciation-extinction analysis we estimated a speciation rate of 0.430 Myr^{-1} (0.107–0.881 Myr^{-1}; 95% HPD) and a lineage extinction rate of 0.446 Myr^{-1} (0.003–0.955 Myr^{-1}; 95% HPD). Including a distantly related outgroup did not substantially change speciation or extinction parameter estimates (data not shown).

Biogeographic reconstructions

The full ASR model had the best AICc score (Figure 2b; differences between best and alternative models: simple $\Delta = 1.06$; continental $\Delta = 2.135$). It shows a split between the tropical (northern Australia and India) and subtropical/temperate (south western Australia, south eastern Australia, and New Zealand) clades (Figure 2b). Within the tropical clade we infer that the Indian species, *Trithuria konkanensis* S.R. Yadav & Janarth, represents a relatively recent long-distance dispersal event from northern Australia (Figure 2b). Within the subtropical/temperate clade, the New Zealand species *T. inconspicua* represents a long-distance dispersal event from south-eastern Australia. There is no significant support for a particular direction of dispersal between south-western and south-eastern Australia (Figure 2b).

The DIVA analysis reconstructed several vicariance events across the phylogeny (Figure 2c). There is an inferred vicariance event at the root of the family, between the tropical and subtropical/temperate clades (Figure 2c). An additional one was predicted between northern Australia and India in the tropical clade, and in the subtropical/temperate clade two more were predicted between south-western Australia and south-eastern Australia, and one between south-eastern Australia and New Zealand (Figure 2c). The DEC analysis recovered similar patterns of vicariance across Hydatellaceae, although it also inferred dispersals, for example from south-eastern Australia to New Zealand (Figure 2d). In both analyses, vicariance is implausible in the context of dating analyses, except for the one involving the root split between the tropical and subtropical/temperate clade. Increasing the extinction rate in the DEC analysis served to depress confidence in the estimated ancestral ranges for all nodes, considering both the relative likelihood of the best reconstructed range pair (Figure 3) and the relative likelihood of the individual ranges at nodes (Figure 4).

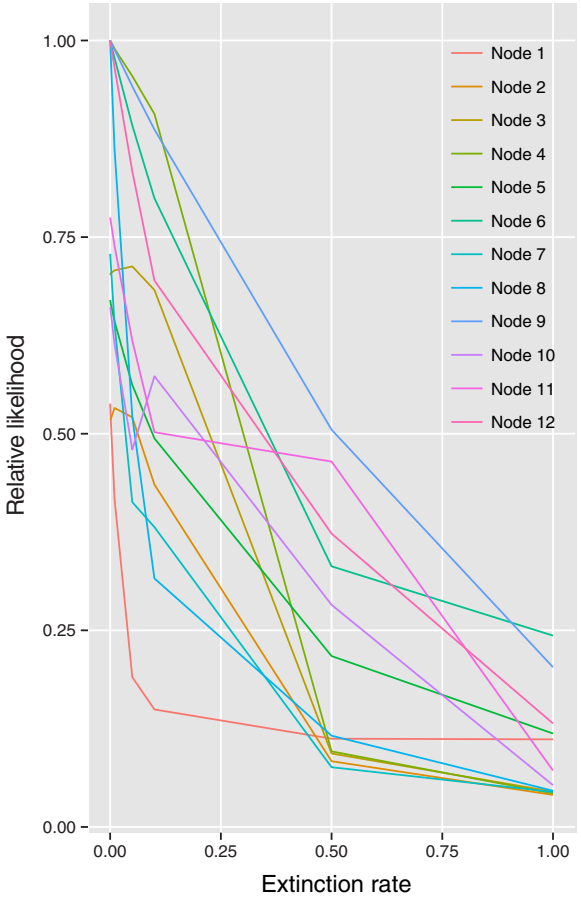

Figure 3 Relationship between area extinction rate and relative likelihood of range pairs in the dispersal-extinction-cladogenesis (DEC) analysis. The plot depicts the value of the range pair with the largest relative likelihood at each node and extinction rate; note that the top pair of ranges is not necessarily consistent across extinction rates. The 'zero' extinction value is the auto-optimized estimate (4.285×10^{-9} Myr^{-1}). Node numbers correspond to those in Figure 2a.

Discussion

The phylogenetic origin of Hydatellaceae near the root of angiosperm phylogeny [10] and lack of reliable fossils [7] make consideration of the family age infeasible outside the context of angiosperm divergence times. Unfortunately the crown age and subsequent timing of diversification of angiosperms remains one of the most vexing questions in evolutionary biology, with some molecular estimates [42,49,61] substantially older (~100 Myr) than the oldest reported crown angiosperm fossils [29]. Our estimated age of 158.7 Ma (151.0–167.7 Ma, 95% HPD; see Additional file 3) is more in-line with less extreme results reported elsewhere [75,76]. A stem age for Hydatellaceae of ~127 Ma (see Results and Additional file 3) suggests that stem lineage Hydatellaceae were colonizing aquatic environments in the Lower Cretaceous, although when Hydatellaceae acquired the unique suite of traits suited for ephemeral aquatic habitats is unclear.

A crown age for Hydatellaceae in the early Miocene (~18 Ma, see Figure 2a and Additional file 5) indicates

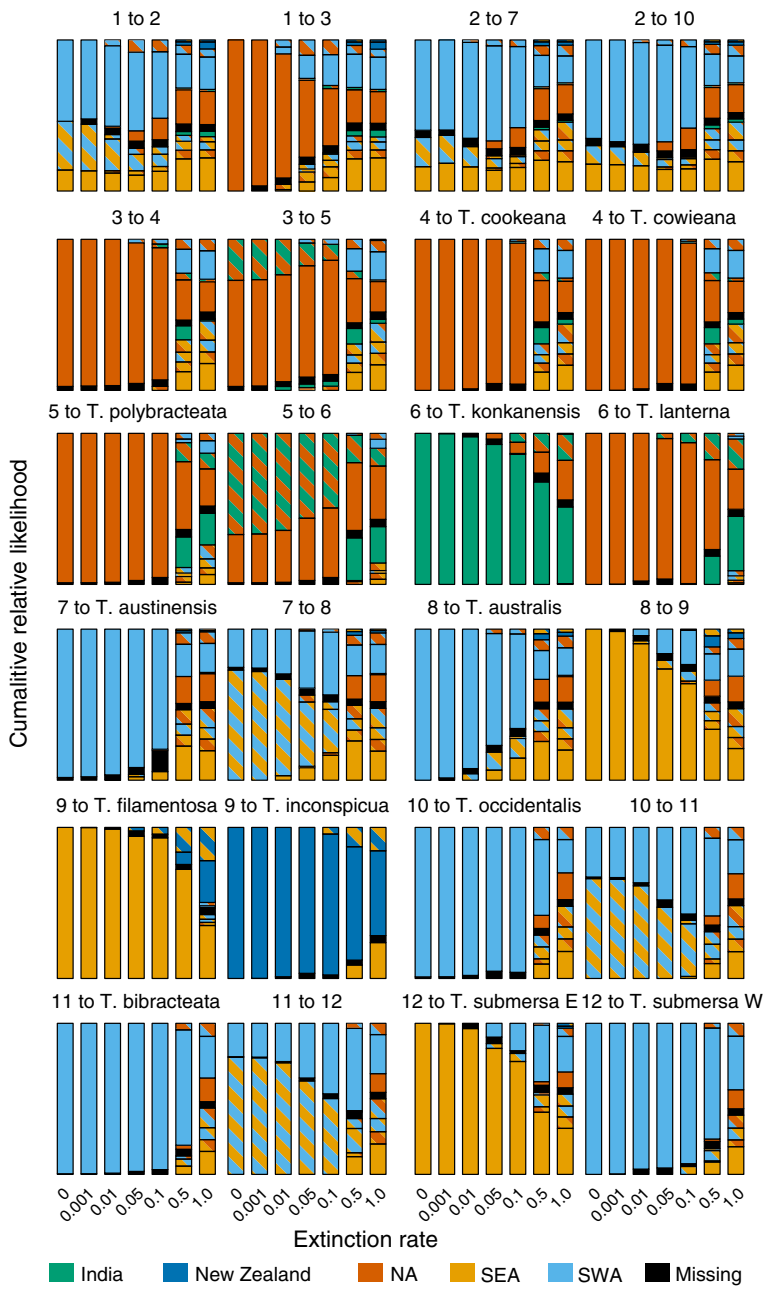

Figure 4 Relationship between extinction rate and relative likelihood of geographic ranges in each descendant lineage of a range-pair in the dispersal-extinction-cladogenesis (DEC) analysis. The likelihoods sum to unity in each bar; subdivisions represent the relative likelihood of each range. Hashed ranges comprise more than one area. NA, northern Australia; SEA, south-eastern Australia; SWA, south-western Australia. The 'zero' extinction rate is the auto-optimized estimate (4.285×10^{-9} Myr^{-1}). Node numbers correspond to those in Figure 2a; the range pair depicted in Figure 2d reports the best pair descending from each node (e.g., node 1 to 2 and 1 to 3) at the auto-optimized rate.

that a proposed Gondwanan explanation for the current intercontinental distribution [20] is incorrect, as it would require that the Indian and north Australian species pair *Trithuria konkanensis* and *T. lanterna* diverged ~125 Ma, according to the timing of the breakup of East Gondwana [77], instead of the estimated divergence time of 0.76 Ma (0.24–1.33 Ma, 95% HPD; see Figure 2a and Additional file 5). This highlights the importance of assessing proposed vicariant patterns with a careful consideration of phylogeny, geology, and estimated divergence times [22].

Within Australia, climate driven vicariance events are more plausible, although here as well, the last submersion of the Nullarbor Plain (~15 Ma), which separates

the south-eastern and south-western regions, substantially predated the relevant phylogenetic splits (Figure 2a; [18]). However, the DIVA and DEC analyses indicate a continent-scale vicariance event at the root of extant Hydatellaceae (Figure 2c,d). The interior of Australia was still relatively wet in the early Miocene (up to the mid-Miocene), and although there were permanent lakes, there was also a marked dry season, indicating the potential for ephemeral aquatic habitats [1]. The continued aridification of central Australia presumably led to this vicariance event. Our analyses therefore support a minimum of four long-distance dispersal events in Hydatellaceae (Australia to India, Australia to New Zealand, and two instances from south-western to south-eastern Australia; Figure 2). The inferred long-distance dispersal events likely involved selfers or apomicts, consistent with Baker's Law [78]. The New Zealand species *Trithuria inconspicua* and its Tasmanian sister species *T. filamentosa* are both thought to be perennial apomicts [79,80]; selfing is thought to characterize the Indian *T. konkanensis* and its sister species, *T. lanterna*, in northern Australia [8,81]. Baker's Law has been extended to dispersal in general, not just islands, and as a result we expect selfing taxa to have wider distributions than outcrossing ones [82]. This seems to be the case in Hydatellaceae, where dioecious species are generally much more limited in distribution than related cosexual species [8].

Statistical biogeographic methods such as DEC allow not only an examination of the biogeographic history of a clade and an estimate of the processes involved in producing that history (dispersal, vicariance and extinction), but also quantification of how confident we are in these reconstructions, via consideration of (relative) likelihoods. A strong bias towards estimating zero area extinction rates may occur in the DEC framework, both for real and simulated data sets [67]. We examined the effect that this may have on our reconstructions by manually varying the extinction rate based on the range of values seen in our speciation-extinction analysis (see Results). Our confidence in reconstructing both (a) range pairs (thereby indicating possible processes such as vicariance or dispersal; Figure 3), and (b) each individual descendent lineage's range (as indicated by the relative likelihoods for each range across all possible range pairs; Figure 4), is compromised at higher rates. For the estimated extinction rate based on tree shape (~0.5 Myr^{-1}), there is very little confidence in any particular range pair (relative likelihoods are <0.6, Figure 3), and in the ranges of individual descendent lineages, besides a few of the very shallowest and youngest nodes (Figure 4). Estimating extinction rates from phylogenies is contentious and often leads to large confidence intervals [37,83,84], which is what we infer with our data. Nevertheless, our results are potentially in line with estimates for other herbaceous groups [85]. Even relatively moderate extinction rates may limit our ability to confidently reconstruct biogeographic history, and so inferences based on the very low optimal extinction rate predicted in the DEC analysis should be treated cautiously. However, despite the greater uncertainty in biogeographic reconstructions at higher extinction rates, the New Zealand and Indian species must represent recent long-distance dispersal events, given their very recent separation from closely related Australian species. The Indian species was discovered only recently (1994; [20]) and yet has a relatively extensive range [20], which may add further weight to the possibility that the global distribution of the family may be more extensive than is currently reported [7,10,86]. Further phylogeographic work in individual species may also reveal additional instances of intra-specific migration and extinction (e.g., with regards to the substantially disjunct distribution of *Trithuria inconspicua* in New Zealand).

Conclusions

Our analyses suggest the Hydatellaceae lineage arose in the Lower Cretaceous, but that extant species diversity dates from the Miocene. The former age highlights the early exploitation of aquatic environments by angiosperms. Our results also emphasize the potentially high extinction rate associated with ephemeral aquatic habitats. Despite having a classical Gondwanan intercontinental pattern, the young age of the crown clade of Hydatellaceae contradicts the role of vicariance events in shaping the family's distribution. This suggests instead that long-distance dispersal is predominately responsible for its disjunct distribution both within and outside Australia.

Availability of supporting data

The new data sets supporting the results of this article are included within the article (and its additional files).

Additional files

Additional file 1: Voucher and accession information for seed-plant molecular dating.

Additional file 2: Seed plant alignment matrix.

Additional file 3: Chronogram and table of inferred ages from the Bayesian random local clock molecular dating analysis of the seed plants.

Additional file 4: Tree file for the seed-plant Bayesian random local clock tree.

Additional file 5: Table of inferred ages in the Bayesian multi-species coalescent molecular dating of Hydatellaceae.

Additional file 6: Tree file for the Hydatellaceae Bayesian multi-species coalescent tree.

Abbreviations

AICc: Corrected Akaike information criterion; ASR: Ancestral state reconstruction; BI: Bayesian inference; DEC: Dispersal extinction cladogenesis; DIVA: Dispersal vicariance; HPD: Highest posterior density; Ma: Millions of

years ago; ML: Maximum likelihood; Myr: Millions of years; RLC: Random local clock; UCLN: Uncorrelated lognormal.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
WJDI designed the study and carried out the analyses. CL constructed the distribution map. All authors contributed to drafting the manuscript; writing was led by WJDI and SWG. All authors read and approved the final manuscript.

Acknowledgements
We thank Marc Jopson and Rick Ree for technical assistance with programming and data analysis, John Conran for advice on ephemeral aquatic plant ecology, and Darren Irwin for advice on biogeographic analyses. This work was supported by a University of British Columbia Graduate Scholarship to WJDI and an NSERC (Natural Sciences and Engineering Research Council of Canada) Discovery Grant to SWG.

Author details
[1]Department of Botany, University of British Columbia, 3529-6270 University Blvd, Vancouver, British Columbia, V6T 1Z4, Canada. [2]UBC Botanical Garden & Centre for Plant Research, University of British Columbia, 6804 Marine Dr SW, Vancouver, British Columbia, V6T 1Z4, Canada. [3]Department of Higher Plants, Biological Faculty, M.V. Lomonosov Moscow State University, 119234 Moscow, Russia. [4]Department of Botany, Shivaji University, Kolhapur 416 004, Maharashtra, India. [5]Botanic Gardens and Parks Authority, Kings Park and Botanic Garden, West Perth, WA 6005, Australia and School of Plant Biology, University of Western Australia, Nedlands, WA 6009, Australia. [6]Department of Parks & Wildlife, Western Australian Herbarium, Science Division, Brain Street, Manjimup 6258, WA, Australia. [7]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, UK.

Received: 21 February 2014 Accepted: 1 May 2014
Published: 13 May 2014

References
1. Martin HA: **Cenozoic climatic change and the development of the arid vegetation in Australia.** *J Arid Environ* 2006, **66:**533–563.
2. Aston HI: *Aquatic Plants of Australia.* Melbourne: Melbourne University Press; 1973:368.
3. Diels L: **Die Pflanzenwelt von West-Australien südlich des Wendekreises mit einer Einleitung über die Pflanzenwelt Gesamt-Australiens in Grundzügen**. In *Die Veg der Erde.* 7th edition. Edited by Engler A, Drude O. Leipzig: Verlag von Wilhelm Engelmann; 1906:1–413.
4. Deil U: **A review on habitats, plant traits and vegetation of ephemeral wetlands—a global perspective.** *Phytocoenologia* 2005, **35:**533–705.
5. Briggs BG, Marchant AD, Perkins AJ: **Phylogeny of the restiid clade (Poales) and implications for the classification of Anarthriaceae, Centrolepidaceae and Australian Restionaceae.** *Taxon* 2014, **63:**24–46.
6. Pignatti E, Pignatti S: **Centrolepidi-Hydrocotyletea alatae, a new class of ephemeral communities in Western Australia.** *J Veg Sci* 1994, **5:**55–62.
7. Sokoloff DD, Remizowa MV, Macfarlane TD, Yadav SR, Rudall PJ: **Hydatellaceae: a historical review of systematics and ecology.** *Rheedea* 2011, **21:**115–138.
8. Sokoloff DD, Remizowa MV, Macfarlane TD, Rudall PJ: **Classification of the early-divergent angiosperm family Hydatellaceae: one genus instead of two, four new species and sexual dimorphism in dioecious taxa.** *Taxon* 2008, **57:**179–200.
9. Iles WJD, Rudall PJ, Sokoloff DD, Remizowa MV, Macfarlane TD, Logacheva MD, Graham SW: **Molecular phylogenetics of Hydatellaceae (Nymphaeales): sexual-system homoplasy and a new sectional classification.** *Am J Bot* 2012, **99:**663–676.
10. Saarela JM, Rai HS, Doyle JA, Endress PK, Mathews S, Marchant AD, Briggs BG, Graham SW: **Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree.** *Nature* 2007, **446:**312–315.
11. Friedman WE, Bachelier JB, Hormaza JI: **Embryology in *Trithuria submersa* (Hydatellaceae) and relationships between embryo, endosperm, and perisperm in early-diverging flowering plants.** *Am J Bot* 2012, **99:**1083–1095.
12. Rudall PJ, Sokoloff DD, Remizowa MV, Conran JG, Davis JI, Macfarlane TD, Stevenson DW: **Morphology of Hydatellaceae, an anomalous aquatic family recently recognized as an early-divergent angiosperm lineage.** *Am J Bot* 2007, **94:**1073–1092.
13. Rudall PJ, Remizowa MV, Prenner G, Prychid CJ, Tuckett RE, Sokoloff DD: **Nonflowers near the base of extant angiosperms? Spatiotemporal arrangement of organs in reproductive units of Hydatellaceae and its bearing on the origin of the flower.** *Am J Bot* 2009, **96:**67–82.
14. Rudall PJ, Eldridge T, Tratt J, Ramsay MM, Tuckett RE, Smith SY, Collinson ME, Remizowa MV, Sokoloff DD: **Seed fertilization, development, and germination in Hydatellaceae (Nymphaeales): implications for endosperm evolution in early angiosperms.** *Am J Bot* 2009, **96:**1581–1593.
15. Friedman WE: **Hydatellaceae are water lilies with gymnospermous tendencies.** *Nature* 2008, **453:**94–97.
16. Sokoloff DD, Remizowa MV, Macfarlane TD, Tuckett RE, Ramsay MM, Beer AS, Yadav SR, Rudall PJ: **Seedling diversity in Hydatellaceae: implications for the evolution of angiosperm cotyledons.** *Ann Bot* 2008, **101:**153–164.
17. Sanmartín I, Ronquist F: **Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns.** *Syst Biol* 2004, **53:**216–243.
18. Crisp MD, Cook LG: **A congruent molecular signature of vicariance across multiple plant lineages.** *Mol Phylogenet Evol* 2007, **43:**1106–1117.
19. Council of Heads of Australasian Herbaria: **Australia's virtual herbarium.** 2012, [http://avh.chah.org.au]
20. Yadav SR, Janarthanam MK: **Hydatellaceae: a new family to Indian flora with a new species.** *Rheedea* 1994, **4:**17–20.
21. Les DH, Crawford DJ, Kimball RT, Moody ML, Landolt E: **Biogeography of discontinuously distributed hydrophytes: a molecular appraisal of intercontinental disjunctions.** *Int J Plant Sci* 2003, **164:**917–932.
22. Donoghue MJ, Moore BR: **Toward an integrative historical biogeography.** *Integr Comp Biol* 2003, **43:**261–270.
23. Sauquet H: **A practical guide to molecular dating.** *Comptes Rendus Palevol* 2013, **12:**355–367.
24. Doyle JA: **Integrating molecular phylogenetic and paleobotanical evidence on origin of the flower.** *Int J Plant Sci* 2008, **169:**816–843.
25. Endress PK, Doyle JA: **Reconstructing the ancestral angiosperm flower and its initial specializations.** *Am J Bot* 2009, **96:**22–66.
26. Hofmann C-C, Zetter R: **Upper Cretaceous sulcate pollen from the Timerdyakh Formation, Vilui Basin (Siberia).** *Grana* 2010, **49:**170–193.
27. Sun G, Dilcher DL, Zheng S, Zhou Z: **In search of the first flower: a Jurassic angiosperm, *Archaefructus*, from northeast China.** *Science* 1998, **282:**1692–1695.
28. Sun G, Ji Q, Dilcher DL, Zheng S, Nixon KC, Wang X: **Archaefructaceae, a new basal angiosperm family.** *Science* 2002, **296:**899–904.
29. Friis EM, Crane PR, Pedersen KR: *Early Flowers and Angiosperm Evolution.* Cambridge: Cambridge University Press; 2011:385.
30. Cevallos-Ferriz SRS, Stockey RA: **Permineralized fruits and seeds from the Princeton Chert (Middle Eocene) of British Columbia: Nymphaeaceae.** *Bot Gaz* 1989, **150:**207–217.
31. Collinson ME: **Recent and Tertiary seeds of the Nymphaeaceae sensu lato with a revision of *Brasenia ovula* (Brong.) Reid and Chandler.** *Ann Bot* 1980, **46:**603–632.
32. Drew BT, Ruhfel BR, Smith SA, Moore MJ, Michael J, Briggs BG, Gitzendanner MA, Soltis PS, Soltis DE: **Another look at the root of the angiosperms reveals a familiar tale.** *Syst Biol.* In press.
33. Feild TS, Arens NC, Doyle JA, Dawson TE, Donoghue MJ: **Dark and disturbed: a new image of early angiosperm ecology.** *Paleobiology* 2004, **30:**82–107.
34. Husband BC, Barrett SCH: **Spatial and temporal variation in population size of *Eichhornia paniculata* in ephemeral habitats: implications for metapopulation dynamics.** *J Ecol* 1998, **86:**1021–1031.
35. Graur D, Martin W: **Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision.** *Trends Genet* 2004, **20:**80–86.
36. Dos Reis M, Inoue J, Hasegawa M, Asher RJ, Donoghue PCJ, Yang Z: **Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny.** *Proc R Soc B Biol Sci* 2012, **279:**3491–3500.
37. Nee S, Holmes EC, May RM, Harvey PH: **Estimating extinction from molecular phylogenies**. In *Extinction Rates.* Edited by Lawton JH, May RM. Oxford: Oxford University Press; 1995:164–182.
38. Crisp M, Cook L, Steane D: **Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities?** *Philos Trans R Soc London B Biol Sci* 2004, **359:**1551–1571.

39. Crisp MD, Cook LG: **Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies.** *Evolution* 2009, **63:**2257–2265.
40. Taylor TN, Taylor EL: *The Biology and Evolution of Fossil Plants.* Prentice Hall: Englewood Cliffs, New Jersey; 1993.
41. Hughes NF: *The Enigma of Angiosperm Origins.* Cambridge: Cambridge University Press; 1994.
42. Magallón SA: **Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms.** *Syst Biol* 2010, **59:**384–399.
43. Mathews S, Clements MD, Beilstein MA: **A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants.** *Philos Trans R Soc London B Biol Sci* 2010, **365:**383–395.
44. Mohr BAR, Bernardes-de-Oliveira MEC, Taylor DW: ***Pluricarpellatia*, a nymphaealean angiosperm from the Lower Cretaceous of northern Gondwana (Crato Formation, Brazil).** *Taxon* 2008, **57:**1147–1158.
45. Taylor DW, Brenner GJ, Basha SH: ***Scutifolium jordanicum* gen. et sp. nov. (Cabombaceae), an aquatic fossil plant from the Lower Cretaceous of Jordan, and the relationships of related leaf fossils to living genera.** *Am J Bot* 2008, **95:**340–352.
46. Friis EM, Pedersen KR, Crane PR: **Fossil evidence of water lilies (Nymphaeales) in the early Cretaceous.** *Nature* 2001, **410:**357–360.
47. Friis EM, Pedersen KR, von Balthazar M, Grimm GW, Crane PR: ***Monetianthus mirus* gen. et sp. nov., a nymphaealean flower from the early Cretaceous of Portugal.** *Int J Plant Sci* 2009, **170:**1086–1101.
48. Yamada T, Nishida H, Umebayashi M, Uemura K, Kato M: **Oldest record of Trimeniaceae from the Early Cretaceous of northern Japan.** *BMC Evol Biol* 2008, **8:**135.
49. Clarke JT, Warnock RCM, Donoghue PCJ: **Establishing a time-scale for plant evolution.** *New Phytol* 2011, **192:**266–301.
50. Hughes NF, McDougall AB: **Barremian-Aptian angiospermid pollen records from southern England.** *Rev Palaeobot Palynol* 1990, **65:**145–151.
51. Friis EM, Pedersen KR, Crane PR: **Araceae from the Early Cretaceous of Portugal: evidence on the emergence of monocotyledons.** *Proc Natl Acad Sci U S A* 2004, **101:**16565–16570.
52. Doyle JA, Endress PK: **Integrating early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots.** *J Syst Evol* 2010, **48:**1–35.
53. Graham SW, Iles WJD: **Different gymnosperm outgroups have (mostly) congruent signal regarding the root of flowering plant phylogeny.** *Am J Bot* 2009, **96:**216–227.
54. Rai HS, Reeves PA, Peakall R, Olmstead RG, Graham SW: **Inference of higher-order conifer relationships from a multi-locus plastid data set.** *Botany* 2008, **86:**658–669.
55. Graham SW, Olmstead RG: **Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms.** *Am J Bot* 2000, **87:**1712–1730.
56. Graham SW, Reeves PA, Burns ACE, Olmstead RG: **Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference.** *Int J Plant Sci* 2000, **161:**S83–S96.
57. Drummond AJ, Suchard MA: **Bayesian random local clocks, or one rate to rule them all.** *BMC Biol* 2010, **8:**114.
58. Dornburg A, Brandley MC, McGowen MR, Near TJ: **Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea).** *Mol Biol Evol* 2012, **29:**721–736.
59. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A: **Relaxed phylogenetics and dating with confidence.** *PLoS Biol* 2006, **4:**e88.
60. Smith SA, Donoghue MJ: **Rates of molecular evolution are linked to life history in flowering plants.** *Science* 2008, **322:**86–89.
61. Smith SA, Beaulieu JM, Donoghue MJ: **An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants.** *Proc Natl Acad Sci U S A* 2010, **107:**5897–5902.
62. Löhne C, Borsch T, Wiersema JH: **Phylogenetic analysis of Nymphaeales using fast-evolving and noncoding chloroplast markers.** *Bot J Linn Soc* 2007, **154:**141–163.
63. Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi Z, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu Y-L, Hilu KW, Davis CC, Sanderson MJ, Beaman RS, Olmstead RG, Judd WS, Donoghue MJ, Soltis PS: **Angiosperm phylogeny: 17 genes, 640 taxa.** *Am J Bot* 2011, **98:**704–730.
64. Taylor DW: **Phylogenetic analysis of Cabombaceae and Nymphaeaceae based on vegetative and leaf architectural characters.** *Taxon* 2008, **57:**1082–1095.
65. Heled J, Drummond AJ: **Bayesian inference of species trees from multilocus data.** *Mol Biol Evol* 2010, **27:**570–580.
66. Ronquist F: **Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography.** *Syst Biol* 1997, **46:**195–203.
67. Ree RH, Smith SA: **Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis.** *Syst Biol* 2008, **57:**4–14.
68. Sanmartín I, van der Mark P, Ronquist F: **Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands.** *J Biogeogr* 2008, **35:**428–449.
69. Yu Y, Harris AJ, He X: **S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories.** *Mol Phylogenet Evol* 2010, **56:**848–850.
70. Nee S, May RM, Harvey PH: **The reconstructed evolutionary process.** *Philos Trans R Soc London B Biol Sci* 1994, **344:**305–311.
71. FitzJohn RG: **Diversitree: comparative phylogenetic analyses of diversification in R.** *Methods Ecol Evol* 2012, **3:**1084–1092.
72. Core Team R: *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing; 2011 [http://www.R-project.org]
73. Pagel M, Meade A, Barker D: **Bayesian estimation of ancestral character states on phylogenies.** *Syst Biol* 2004, **53:**673–684.
74. Sugiura N: **Further analysis of the data by Akaike's information criterion and the finite corrections.** *Commun Stat Theory Methods A7* 1978, **7:**13–26.
75. Bell CD, Soltis DE, Soltis PS: **The age and diversification of the angiosperms re-revisited.** *Am J Bot* 2010, **97:**1296–1303.
76. Magallón S, Hilu KW, Quandt D: **Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates.** *Am J Bot* 2013, **100:**556–573.
77. Ali JR, Aitchison JC: **Gondwana to Asia: plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma).** *Earth-Science Rev* 2008, **88:**145–166.
78. Baker HG: **Self-compatibility and establishment after "long-distance" dispersal.** *Evolution* 1955, **9:**347–349.
79. Remizowa MV, Sokoloff DD, Macfarlane TD, Yadav SR, Prychid CJ, Rudall PJ: **Comparative pollen morphology in the early-divergent angiosperm family Hydatellaceae reveals variation at the infraspecific level.** *Grana* 2008, **47:**81–100.
80. Rudall PJ, Remizowa MV, Beer AS, Bradshaw E, Stevenson DW, Macfarlane TD, Tuckett RE, Yadav SR, Sokoloff DD: **Comparative ovule and megagametophyte development in Hydatellaceae and water lilies reveal a mosaic of features among the earliest angiosperms.** *Ann Bot* 2008, **101:**941–956.
81. Sokoloff DD, Remizowa MV, Yadav SR, Rudall PJ: **Development of reproductive structures in the sole Indian species of Hydatellaceae, *Trithuria konkanensis*, and its morphological differences from Australian taxa.** *Aust Syst Bot* 2010, **23:**217–228.
82. Pannell JR, Barrett SCH: **Baker's Law revisited: reproductive assurance in a metapopulation.** *Evolution* 1998, **52:**657–668.
83. Rabosky DL: **Extinction rates should not be estimated from molecular phylogenies.** *Evolution* 2010, **64:**1816–1824.
84. Nee S: **Birth-death models in macroevolution.** *Annu Rev Ecol Evol Syst* 2006, **37:**1–17.
85. Lancaster LT: **Molecular evolutionary rates predict both extinction and speciation in temperate angiosperm lineages.** *BMC Evol Biol* 2010, **10:**162.
86. Hamann U: **Hydatellaceae**. In *Fam Genera Vasc Plants IV Flower Plants - Monocotyledons - Alismatanae Commelinanae.* Edited by Kubitzki K. Berlin: Springer; 1998:231–234.

doi:10.1186/1471-2148-14-102
Cite this article as: Iles *et al.*: Reconstructing the age and historical biogeography of the ancient flowering-plant family Hydatellaceae (Nymphaeales). *BMC Evolutionary Biology* 2014 **14**:102.