Molecular Phylogenetics and Evolution 51 (2009) 19–30

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Molecular phylogenetics of Haemodoraceae in the Greater Cape and Southwest Australian Floristic Regions

Stephen D. Hopper [a,b], Rhian J. Smith [a,*], Michael F. Fay [a], John C. Manning [c], Mark W. Chase [a]

[a] Royal Botanic Gardens, Kew, Kew Green, Richmond, Surrey TW9 3AB, UK
[b] School of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia
[c] Compton Herbarium, South African National Botanical Institute, P. Bag X7, Claremont 7735, South Africa

ARTICLE INFO

Article history:
Received 4 January 2008
Revised 29 October 2008
Accepted 6 November 2008
Available online 3 December 2008

Keywords:
Haemodoraceae
Greater Cape
Southwestern Australia
SWAFR
Relictual taxa
matK
trnL intron
trnL-F spacer
Diversification
Lineage turnover

ABSTRACT

Molecular phylogenetic studies of Haemodoraceae in the Greater Cape and Southwest Australian Floristic Region (SWAFR) using *trnL*, *trnL-F* and *matK* sequence data affirm the presence of old and young rapidly radiated lineages in both regions. Commencement of tribal and generic divergence in the subfamilies occured in the Eocene in the two regions, but subsequent patterns of radiation differ slightly. The hypothesis of rapid recent speciation in these regions from the late Pliocene as the major explanation for endemic species richness is still repeated by several contemporary authors despite increasing molecular phylogenetic evidence to the contrary. Our estimates of the age of lineages in Haemodoraceae show significant lineage turnover occurring over the last 15 million years, since the mid-Miocene, with divergence of the major clades beginning in the Eocene. The search for independent evidence to date speciation episodes reliably and investigation of molecular analyses across a broad spectrum of these clades must be pursued to advance ideas rigorously concerning origins of species richness. These regions continue to confound attempts to develop theory concerning origins of global species richness, with consequent implications for conservation biology.

1. Introduction

Levyns (1938, 1964) was one of the first authors to consider the origin and speciation of the rich endemic flora of South Africa's Cape region, for which she hypothesized a predominantly tropical African origin, with minor southern (Gondwanan) and Eurasian elements. This simple view was revised and updated considerably by Goldblatt (1978). Since then, interest in the extraordinary species richness of the region has grown substantially. A feature of modern pre-molecular treatments has been persistent repetition of a rapid and recent speciation hypothesis. On the basis that the scant fossil record only shows dominance of sclerophyll fynbos shrubland elements over rainforest in the late Pliocene (Scott et al., 1997), many authors have proposed that explosive speciation has occurred in fynbos and succulent karoo genera over the past three million years (Cowling et al., 1996; Cowling and Hilton-Taylor, 1997; Cowling and Pierce, 1999; Goldblatt and Manning, 2000, 2002; Klak et al., 2004; Ellis et al., 2006; Crawford, 2008). Whereas some recent molecular phylogenetic studies refuted this rapid recent speciation hypothesis (reviewed by Linder, 2003, 2005, 2008; Galley and Linder, 2006; Edwards and Hawkins, 2007), until the Cape flora is much more comprehensively sampled it would be premature to conclude that the hypothesis has been solidly rejected. Indeed, some have speculated that such rapid rates of speciation, for example among the mesembs (Aizoaceae) of the succulent karoo, require unique genetic systems prone to diversification (Ihlenfeldt, 1994; Klak et al., 2004).

The presence of extraordinarily species-rich genera is also seen in the Southwest Australian Floristic Region (SWAFR, sensu Hopper and Gioia, 2004) and other mediterranean climate hotspots, leading to the conclusion that explosive recent speciation has been a major evolutionary event (Hopper, 1979, 1992; Cowling et al., 1996; Hopper et al., 1996). However, particularly in relation to the SWAFR, an additional Gondwanan heritage hypothesis has been advocated for some time, proposing persistence of relictual lineages and recurrent episodes of speciation in old stable landscapes (Hopper et al., 1996). Under this scenario, species richness is due to multiple episodes of Cretaceous–Cenozoic speciation among old Gondwanan lineages in ancient stable landscapes, with significant (but not almost all) speciation associated with the late Tertiary onset of mediterranean climate. For the Greater Cape flora (*sensu* Born et al., 2007), Goldblatt and others have commented on palaeoendemics in mesic refugia of the Cape fold-mountains, but these have been regarded as of only minor relevance to the main challenge of

* Corresponding author. Fax: +44 (0)208 3325109.
E-mail address: Rhian.Smith@kew.org (R.J. Smith).

1055-7903/$ - see front matter Crown copyright © 2008 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.11.015

explaining extraordinary species richness (Cowling et al., 1996; Cowling and Hilton-Taylor, 1997; Cowling and Pierce, 1999; Goldblatt and Manning, 2000, 2002).

Molecular phylogenetics has recently generated profound new insights in this arena (Linder, 2003, 2005, 2008; Edwards and Hawkins, 2007). Furthermore, comparative molecular analyses involving taxa in the Greater Cape Floristic Region and in the SWAFR promise to add greater insights through reciprocal illumination of evolutionary processes driven by historical environmental change across the Southern Hemisphere. For example, in the Australian flora, cytogeneticists long ago hypothesized that two bursts of speciation had occurred (James, 1981). The first involved speciation in the Cretaceous to early Tertiary, resulting in the evolution of endemic Australian families, tribes and genera. This was followed by a period of relative cytoevolutionary stability as genera, subgenera and sections radiated. A more recent, late Tertiary to Quaternary second phase of extensive cytoevolution was hypothesized to have ensued, giving rise to modern species and races via frequent dysploidy and polyploidy. Do molecular phylogenetic data support such an interpretation? Could a similar pattern also apply to the Greater Cape flora?

Haemodoraceae are a small family of ca. 111 species placed in Commelinales with Commelinaceae, Hanguanaceae, Philydraceae, and Pontederiaceae (Simpson, 1990, 1998; Chase et al., 1995, 2000, 2006; Hopper et al., 1999; APG II, 2003; Graham et al., 2002). There is 100% bootstrap support for monophyly of Haemodoraceae based on plastid DNA (Hopper et al., 1999). It has been hypothesized that Haemodoraceae have many highly divergent, relictual taxa, as well as a few genera demonstrating more recent evolutionary radiations (Hopper et al., 1999).

Haemodoraceae subfamily Haemodoroideae are widespread, with nine small genera and ca. 40 species in South Africa (see Fig. 1), the Americas, and Australia-New Guinea. Conostylidoideae are endemic to Southwestern Australia (Fig. 1) and have six genera with ca. 70 species.

Conostylis is the largest genus in Haemodoraceae, with 45 species recognized by Hopper et al. (1987) and more awaiting description based on recent research (Hopper, unpublished data). The genus contains a few relictual, divergent species as well as several species complexes in which patterns of geographical variation, natural hybridization, and dysploidy and/or polyploidy are evident (Hopper et al., 2006). Other genera that have radiated in Southwestern Australia include *Anigozanthos* and *Haemodorum*.

Apart from *Conostylis* (Hopper et al., 2006), taxon sampling of other genera with multiple species in Haemodoraceae has so far been insufficient to illuminate phylogenetic relationships at the species level. Consequently, we embarked on more intensive taxon sampling to assess relationships across Haemodoraceae.

2. Materials and methods

2.1. Taxon sampling

Herbarium vouchers for all samples used in this analysis, along with the database accession numbers of their sequences are given in Table 1. With the exception of *Pyrorrhiza* all genera within Haemodoraceae were sampled (93%), and multiple species were sampled for each genus with more than one species. Over the entire family 54% of species were sampled, and every major clade within genera was sampled (based on previous molecular analyses in Hopper et al., 1999 and 2006). Six species from outside Haemodoraceae were sampled from Philydraceae, Pontederiaceae and Commelinaceae, with Commelinaceae placed as the ultimate outgroup on the basis of previous work by Dahlgren (1980) (reviewed by Simpson, 1990 and Chase et al. (2006)). Samples were collected and dried using silica gel (Chase and Hills, 1991) or were sourced from the DNA banks at the Royal Botanic Gardens, Kew (http://data.kew.org/dnabank/homepage.html) and the South African National Botanical Institute.

2.2. DNA extraction and sequencing

Total genomic DNA was extracted using a 2× CTAB protocol (Doyle and Doyle, 1987) and subsequently cleaned on a cesium chloride/ethidium bromide gradient (1.55 g/ml density) for purification (Creeth and Denborough, 1970). All samples are lodged in the DNA bank of the Royal Botanic Gardens, Kew or the South African National Botanical Institute.

The plastid regions chosen for the analysis were the *trnL-F* region (encompassing the *trnL* intron and *trnL-F* intergenic spacer), and partial exon of the *matK* gene. Amplification of these regions was carried out separately in 25 μl reactions, containing 22 μl of 2.5 mM Mg PCR master mix (Abgene Ltd., Epsom, UK), 1 μl bovine serum albumin (0.04%), 0.5 μl each of forward and reverse primers (at a concentration of 100 ng/μl) and approximately 50 ng DNA template. The primers used to amplify the *trnL-F* region were c and f (Taberlet et al., 1991) and *matK* were 390F and 1326R (Sun et al., 2001).

The amplification profile for *trnL-F* consisted of initial denaturation of samples at 94 °C for 3 min, followed by 28 cycles of denaturation at 94 °C for 1 min, annealing at 48 °C for 1 min and extension at 72 °C for 1 min and a final phase extension phase of 7 min at 72 °C. Amplification of *matK* consisted of initial denaturation of samples at 94 °C for 1 min, followed by 38 cycles of denaturation at 94 °C for 30 s, annealing at 52 °C for 40 s and extension at 72 °C for 40 s, and a final phase extension phase of 5 min at 72 °C.

Products were purified with Machery-Nagel Nucleospin® Extract II DNA purification columns using the manufacturer's protocol (GmbH and Co., Düren, Germany). Dideoxy cycle sequencing was then performed with the chain termination method using ABI (Applied Biosystems Inc. Warrington, UK) Prism Big Dye version 3.1 reaction kit, following the manufacturer's protocols. The products were run on an ABI 3700 Genetic Analyzer. Sequence editing and assembly were performed using Sequencher v.4.5 (Genecodes Corporation, Ann Arbor, Michigan, USA). Sequences were aligned by eye in PAUP v4.0b2A (Swofford, 2001) in accordance with the guidelines of Kelchner (2000).

2.3. Data analysis

Aligned sequence data from the plastid regions were analyzed using the Fitch parsimony algorithm (Fitch, 1971) in the phylogenetic software package PAUP v4.0b2A. Heuristic searches were carried out with 1000 replicates of random taxon entry using tree bisection-reconnection (TBR) and limiting the number of trees held at each step to 20 to reduce the amount of time spent swapping on large numbers of sub-optimal trees. Support for clades was obtained by performing bootstrap analysis with 1000 replicates of simple taxon addition with TBR swapping and retaining 20 trees per replicate.

A Bayesian analysis was carried out on the same data using MrBayes v3.1.2 (Huelsenbeck and Ronquist, 2001 J.P. Huelsenbeck and F. Ronquist, MRBAYES: Bayesian inference of phylogeny, *Bioinformatics* 17 (2001), pp. 754–755. Full Text via CrossRef | View Record in Scopus | Cited By in Scopus (3666) Ronquist and Huelsenbeck, 2003). The model of substitution for each plastid region was determined using MrModelTest 2.2 (Nylander, 2004), which specified a general time-reversible model (Yang, 1994) corrected for invariable sites and among site rate variation using a discrete gamma distribution