Plant Syst Evol (2010) 290:173–184
DOI 10.1007/s00606-010-0359-1

ORIGINAL ARTICLE

Phylogeny of the Southeast Asian endemic genus *Neocinnamomum* H. Liu (Lauraceae)

Zhi-hua Wang · Jie Li · John G. Conran · Hsi-wen Li

Received: 22 November 2009 / Accepted: 6 September 2010 / Published online: 8 October 2010

Abstract A phylogenetic analysis of *Neocinnamomum* H. Liu and related genera was conducted using *psb*A–*trn*H, *trn*K cpDNA regions, and the ITS nrDNA segment. *Neocinnamomum* was confirmed to be monophyletic, and an evolutionary series of inflorescence development within the genus was recognized. The compound thyrse seen in *N. caudatum* is reduced to the few- to many-flowered condensed inflorescences with a poorly defined branching system seen in most species and ultimately to the 1-flowered inflorescence seen in *N. atjehense*. Consensus network analysis (CNA) suggested that long-branch attraction is responsible for the observed close relationship between *Neocinnamomum* and *Cassytha* L. in a combined analysis of the complete data. In contrast, the sister relationship of *Neocinnamomum* and *Caryodaphnopsis* seen in the Bayesian analyses of the partial combined matrix was supported by CNA and is also supported by morphology and wood and bark anatomy. The close similarity of the compound thyrse of less derived *Neocinnamomum* species to the thyrsoid inflorescences of some *Caryodaphnopsis* species is also seen as strong support for their affinity.

Keywords *Neocinnamomum* · *Cassytha* · *Caryodaphnopsis* · Lauraceae · Phylogeny · Long-branch attraction · Consensus network

Z. Wang · J. Li (✉)
Laboratory of Plant Phylogenetics and Conservation Biology, Xishuangbanna Tropical Botanical Garden, The Chinese Academy of Sciences, 88 Xuefu Rd, Kunming 650223, Yunnan, People's Republic of China
e-mail: jieli@xtbg.ac.cn

Z. Wang · J. Li
Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun 666303, Yunnan, People's Republic of China

J. G. Conran
Australian Centre for Evolutionary Biology and Biodiversity, School of Earth and Environmental Sciences, The University of Adelaide, Benham Bldg DX 650 312, Adelaide, SA 5005, Australia

H. Li
Herbarium (KUN), Kunming Institute of Botany, The Chinese Academy of Sciences, Kunming 650204, Yunnan, People's Republic of China

Z. Wang
Graduate University, The Chinese Academy of Sciences, Beijing 100049, People's Republic of China

Introduction

The genus *Neocinnamomum* H. Liu is one of the most enigmatic groups in Lauraceae, with six species endemic to tropical Asia (Kostermans 1974a; Li et al. 1984, 2008b; Rohwer 1993). The genus was established originally by Liu (1934), based on the possession of four-locular anthers with collateral pollen sacs. In *Cinnamomum* Schaeff. the anther sacs are placed in pairs above each other, whereas in *Neocinnamomum* they are at the same level, with one pair being introrse or extrorse, the other lateral (Liu 1934). Although anther sac position was not accepted as a generic characteristic by Kostermans (1974a), he retained *Neocinnamomum* based on its possession of a compound thyrse (usually strongly reduced to a few- to many-flowered condensed inflorescence), a fleshy shallow fruit cup with persistent enlarged tepals, and distichous leaves (Kostermans 1974a; Li et al.

1984). Kostermans (1974a) recognized six species, four of which are easily identified. However, *N. delavayi* (Lec.) H. Liu and *N. mekongense* (Hand.-Mazz.) Kosterm. are very similar morphologically and largely distinguished from each other by the presence or absence of silvery, sericeous hairs on the branchlets (Kostermans 1974a; Li et al. 1984). Kostermans (1974a) also considered that *Neocinnamomum* was very close to *Cinnamomum*, although an earlier wood and bark anatomical study had suggested a more isolated systematic position (Richter 1981).

In contrast, recent molecular studies in Lauraceae have revealed novel but conflicting phylogenetic relationships for *Neocinnamomum* (Chanderbali et al. 2001; Rohwer and Rudolph 2005), and the previously suggested affinity with *Cinnamomum* was rejected. Instead a close relationship between *Neocinnamomum* and either *Caryodaphnopsis* Airy Shaw or *Cassytha* L. was postulated, as well as a relatively proximal position within Lauraceae, albeit with uncertainty. Chanderbali et al. (2001) found that *Neocinnamomum* was sister to *Cassytha*, with the pair placed between the Cryptocaryeae and a *Chlorocardium-Mezilaurus* clade with *Caryodaphnopsis*. This relationship received very high bootstrap support, but also displayed the longest branches in the topology. Rohwer and Rudolph (2005) explored phylogeny within Lauraceae using *trn*K intron sequences, suggesting that *Neocinnamomum* was sister to *Caryodaphnopsis,* but with low support, and agreeing with Chanderbali et al. (2001) that the *Neocinnamomum* and *Cassytha* clade was the result of long-branch attraction (LBA).

Genes for phylogenetic inference of Lauraceae are limited, with *trn*K, *psb*A–*trn*H, and ITS preferred, mainly due to their relatively fast rates of evolution (Rohwer 2000; Chanderbali et al. 2001; Li et al. 2004, 2006, 2007, 2008a; Rohwer and Rudolph 2005). The *trn*K gene has proved to be useful in inferring intergeneric relationships, especially for the basal part of Lauraceae, though few informative substitutions were detected within Laureae (Rohwer 2000; Rohwer and Rudolph 2005), limiting its usefulness there. ITS and the *psb*A–*trn*H spacer were therefore included here to help resolve the relationships of and within *Neocinnamomum* because of their faster relative evolutionary rates in the family (Chanderbali et al. 2001; Li et al. 2004, 2007, 2008a).

Accordingly, this study uses *trn*K, *psb*A–*trn*H, and ITS sequences to accomplish the following:

1. Test the hypothesis that *Neocinnamomum* is monophyletic
2. Clarify species relationships within the genus
3. Develop a more robust, better supported phylogenetic placement for *Neocinnamomum*

Materials and methods

Taxon sampling and DNA extraction

We were able to obtain material of five of the six *Neocinnamomum* species, but material of the little known *N. atjehense* Kosterm. from Sumatra was not available. Four species each from the allegedly related genera *Caryodaphnopsis* and *Cassytha* were also included as ingroup taxa, based on the studies of Chanderbali et al. (2001) and Rohwer and Rudolph (2005). Because previous studies have shown *Chlorocardium-Mezilaurus* and Perseeae-Laureae to be sister clades (Rohwer 2000; Chanderbali et al. 2001; Rohwer and Rudolph 2005), we chose only a few taxa from these for our study to represent their main lineages: *Chlorocardium rodiei* (R.H.Schomb.) Rohwer, H.G.Richt. & van der Werff and *C. venenosum* (Kosterm. & Pinkley) Rohwer, H.G.Richt. & van der Werff for the *Chlorocardium-Mezilaurus* clade and *Persea americana* Mill., *Sassafras tzumu* (Hemsl.) Hemsl., and *Umbellularia californica* (Hook. & Arn.) Nutt. to represent the Perseeae-Laureae. *Beilschmiedia robusta* C.K.Allen and *Cryptocarya metcalfiana* C.K.Allen (both Cryptocaryeae) were chosen as outgroup taxa since the Cryptocaryeae were shown to be sister to the ingroup in previous studies (Chanderbali et al. 2001; Rohwer and Rudolph 2005).

A single sample per species was employed for all taxa except *N. mekongense* (Hand.-Mazz.) Kosterm., for which three samples were sequenced due to possible problems with taxon variability and identity. Sample 01 is a typical specimen from Weixi, Yunnan, whereas sample 02, although collected from the same place, had been identified initially as *N. caudatum* (Nees) Merr. because of its larger leaves and parallel secondary veins. Sample 03 was obtained from the Kunming Botanical Garden, where it was catalogued as *N. delavayi* (Lec.) H. Liu (Kunming Botanical Garden, KIB, CAS 2006), although with uncertainty, as it lacked hairs on the branchlets and leaves. A complete list of the species sampled, along with collection and voucher information is provided in Table 1.

Total DNA was extracted from silica gel-dried leaves or herbarium specimens, using a modified CTAB procedure (Doyle and Doyle 1987), followed by purification using the QIAquick® purification Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocols.

PCR, sequencing, and sequence alignment

The main primer pairs used to amplify the *trn*K intron (including the *mat*K gene) are *trn*K–3914F and *trn*K–2R (Johnson and Soltis 1995; Steele and Vilgalys 1994). However, if these were unsuccessful, internal primers of

Table 1 Source of plant material and GenBank Accession numbers used in this study

Taxon	*trn*K	ITS	*psb*A-*trn*H	Provenance/voucher
Beilschmiedia robusta Allen	GU117736	GU082363	GU117747	Yunnan, China: Wang Z.-H. 08063 (HITBC)
Caryodaphnopsis bilocellata van der Werff & N.K.Dao	AJ627918[a]	–	AF261995[a]	GenBank
C. laotica Airy Shaw	GU117737	GU082364	GU117748	Yunnan, China: Li L. 20070059 (HITBC)
C. tonkinensis (Lec.) Airy Shaw	GU117738	GU082365	GU117749	Yunnan, China: Wang Z.-H. 07081 (HITBC)
C. tomentosa van der Werff	AJ627919[a]	–	AF268807[a]	GenBank
Cassytha ciliolata Nees	AJ247152[a]	–	–	GenBank
C. filiformis L.	GU117739	GU082366	GU117750	Guangdong, China: Zhong J.-S. 08041 (HITBC)
C. melantha R.Br.	AJ627920[a]	–	–	GenBank
C. pubescens R.Br.	AJ627921[a]	–	–	GenBank
Chlorocardium rodiei (R.H.Schomb.) Rohwer, H.G.Richt. & van der Werff	AJ247153[a]	AF272258[a]	AF268802[a]	GenBank
Ch. venenosum (Kosterm. & Pinkley) Rohwer, H.G.Richt. & van der Werff	AJ627922[a]	AF272259[a]	AF268801[a]	GenBank
Cryptocarya metcalfiana Allen	GU117740	GU082367	GU117751	Yunnan, China: Wang Z.-H. 08067 (HITBC)
Neocinnamomum caudatum (Nees) Merr.	GU117741	GU082368	GU117752	Yunnan, China: Wang Z.-H. 07085 (HITBC)
N. delavayi (Lec.) Liou	GU117742	GU082369	GU117753	Yunnan, China: Wang Z.-H. 07087 (HITBC)
N. fargesii (Lec.) Kosterm.	GU117743	GU082370	GU117754	Chongqing, China: Li L. 2007304 (HITBC)
N. lecomtei Liou	–	GU082371	–	Guangxi, China: Akiyama H. et al. 1122 (KUN 0173367)
N. mekongense (Hand.-Mazz.) Kosterm. 01	GU117744	GU082372	GU117755	Yunnan, China: Wang Z.-H. 07094 (HITBC)
N. mekongense 02	GU117745	GU082373	GU117756	Yunnan, China: Wang Z.-H. 07091 (HITBC)
N. mekongense 03	GU117746	GU082374	GU117757	Yunnan, China: Wang Z.-H. 07088 (HITBC)
Persea americana Mill.	AJ247179[a]	AF272322[a]	AF268794[a]	GenBank
Sassafras tzumu (Hemsl.) Hemsl.	AJ247188[a]	GU082375	GU117758	GenBank; Yunnan, China: Wang Z.-H. 08101 (HITBC)
Umbellularia californica (Hook. & Arn.) Nutt.	AJ247190[a]	AF268777[a]	AF272337[a]	GenBank

*trn*K–1400R (5′-TGTGTTCGCTCAAGAAAGG-3′) and *trn*K–1100F (5′-ATATACTAATACCCCACCC-3′) were designed with the help of Primer3 (Rozen and Skaletsky 2000). The *psb*A–*trn*H spacer was amplified using the primers of Sang et al. (1997). PCR amplification of the whole ITS region was successful in most cases using the primer pair ITS4–ITSF for Perseeae-Laureae or ITS4–ITS5m for other taxa (White et al. 1990; Sang et al. 1995; Chanderbali et al. 2001). For some poor-quality DNA templates, the primer combinations of ITS3–ITS4 and ITS2–ITSF or ITS2–ITS5m were used to amplify ITS1 and ITS2 (including the 5.8S nrDNA region) separately (White et al. 1990; Sang et al. 1995; Chanderbali et al. 2001). To minimize the selective amplification of pseudogenes or paralogous ITS copies, a final concentration of 10% DMSO was included in all amplifications (Buckler and Holtsford 1996; Buckler et al. 1997). All PCR amplifications included negative controls to detect contamination. Once the required fragment was amplified successfully, the PCR product was purified using the Qiagen QIAquick® PCR purification Kit. The same amplification primers were also used to sequence the purified fragments in both directions for ITS and *psb*A–*trn*H. For the *trn*K region, we obtained full-length sequences by means of the internal primer combinations of *trn*K–3914F/*trn*K–1400R and *trn*K–1100F/*trn*K–2R. Sequencing reactions were conducted on an Applied Biosystems 3100 DNA automated sequencer after purification.

Each sequence was aligned initially and edited base by base using the program SeqMan II (Lasergene software

package; DNASTAR, Madison, WI, USA), allowing uncertainties either to be resolved or recorded as ambiguities. The consensus sequences were realigned using MUSCLE 3.7 (Edgar 2004) and modified manually using BioEdit version 5.0.9 (Hall 1999).

Phylogenetic analysis

Phylogenetic analysis was performed on the combined data matrix of the *trn*K, ITS, and *psb*A–*trn*H sequences, with gaps scored as missing data. In addition, the data were analyzed using a partial combined matrix consisting of *trn*K, the 5.8S region of ITS, and both end regions of *psb*A–*trn*H. Congruence between datasets was checked using the incongruence length difference (ILD) or partition homogeneity test (Farris et al. 1995; Swofford 2002), as implemented in PAUP* version 4.01b10 (Swofford 2002). Sequences were analyzed by heuristic searching with 1,000 replicates, each with ten random sequence additions, MULTREES on, and TBR branch swapping.

Maximum parsimony (MP) and Bayesian inference (BI) methods were both used for phylogenetic reconstruction, even though MP analyses can be easily influenced by long-branch attraction (LBA; Huelsenbeck et al. 2001; Bergsten 2005; Philippe et al. 2005), a phenomenon that was observed previously for *Neocinnamomum* with *trn*K sequences by Rohwer and Rudolph (2005). BI and MP analyses were performed with MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) and PAUP* version 4.0b10 (Swofford 2002), respectively. A major advantage of using MrBayes is the possibility of partitioning the data, giving each partition its own best-fitting model of sequence evolution. However, over-partitioning may introduce unnecessary sampling variances that could influence the phylogenetic estimates (Poux et al. 2008), so only one model was employed for each gene. The models for the complete combined (*trn*K: GTG + G; ITS: GTR + I+G; *psb*A–*trn*H: K81uf + G) and partial combined datasets (*trn*K: GTG + G; 5.8S: K80 + I; adopted *psb*A–*trn*H region: TVM + G) were selected using Modeltest 3.7 (Posada and Crandall 1998) with the Akaike information criterion (AIC).

BI analysis was run with four simultaneous Markov Chain Monte Carlo (MCMC) chains for 1,000,000 generations, with trees sampled every 100 generations, giving a total of 10,001 trees in the initial sample. Before a majority-rule consensus tree was generated, trees produced prior to log likelihood stabilization and convergence were discarded as "burn in" (25%). For each dataset, MCMC chain runs were repeated twice as a safeguard against spurious results, with a Bayesian majority-rule consensus tree constructed from the remaining trees.

MP analysis was conducted using heuristic search methods with TBR branch swapping and 1,000 random addition cycles. Bootstrap tests (Felsenstein 1985) were performed using 500 replicates with heuristic settings of 100 random addition cycles per replicate, five trees saved from each addition cycle, and TBR branch swapping. The Bayesian inference and MP analysis support values are represented as posterior probability (PP) and bootstrap support (BS), respectively. The trees were edited using FigTree 1.0 (Drummond and Rambaut 2007).

In order to detect the presence of LBA, consensus networks were generated from MCMC sampled trees and those computed from MP bootstrap replicates. The networks were generated with the program Splitstree version 4.10 (Huson 1998; Huson and Bryant 2006) using a 10% threshold. This threshold ensures that only splits appearing in more than 10% of the sampled trees were included in the final consensus network.

Results

Sequence characters

Sequence variation in the ITS and *psb*A–*trn*H regions was much higher than that seen in *trn*K, with the evolutionary rate for ITS almost twice that of *trn*K, making sequence alignment very difficult, especially for *Neocinnamomum* and *Cassytha*. The *psb*A–*trn*H region seems to have evolved at a moderate rate, providing 21% of the informative sites. However, indels in this gene region were highly variable. For example, one segment of indels/deletions in *Caryodaphnopsis* was 172 bp long whereas the combined length of all indels/deletions in *Cassytha* was 200 bp long, and the total aligned *psb*A–*trn*H matrix was only 514 bp long. Nevertheless, all three datasets were found to be congruent in the ILD test ($P > 0.05$). The characteristics of each sequenced gene region and for the combined datasets are summarized in Table 2.

Within each genus, the ITS region showed variation while the cpDNA generally evolved much more slowly, with three out of the four sequenced *Neocinnamomum* species identical for their *psb*A–*trn*H sequences, as were those of *Caryodaphnopsis laotica* Airy Shaw and *C. tonkinensis* (Lec.) Airy Shaw. Similarly in the *trn*K region, no differences were detected between *N. mekongense* and *N. delavayi*.

Because only ITS amplified from the only available herbarium material of *N. lecomtei* H. Liu, it was included in the complete combined but not the partial combined analysis. Despite their morphological differences, the sequences for all three *N. mekongense* accessions were identical for all three sequenced regions.

Table 2 Character information for the individual sequence and combined datasets

Matrix	Aligned length	Uninformative sites (%)	Informative sites (%)
*psb*A–*trn*H	514	140 (27)	108 (21)
ITS	785	387 (49)	262 (33)
*trn*K	2,542	540 (21)	434 (17)
Complete combined matrix	3,814	1,067 (27)	804 (21)
Partial combined matrix	2,916	600 (20)	473 (16)

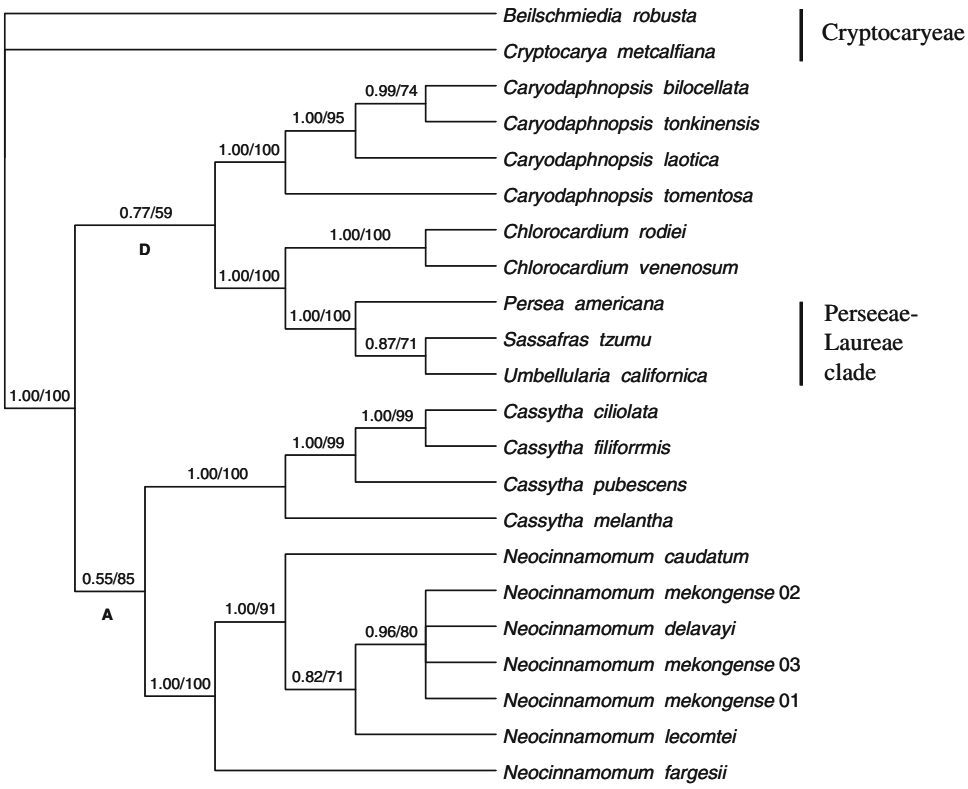

Fig. 1 Strict consensus tree for the complete combined *psb*A–*trn*H, *trn*K, and ITS dataset. The Bayesian posterior probabilities and bootstrap values greater than 50% are shown *above the branches*. *Letters below branches* refer to those in the LBA network consensus diagrams (Fig. 4)

Complete combined analyses

The Bayesian inference and MP analyses generated identical cladograms (Fig. 1) for the complete combined datasets. The analysis supports the monophyly of *Neocinnamomum* (PP = 1.00, BS = 100%), with *N. fargesii* (Lec.) Kosterm. sister to the remainder (PP = 1.00, BS = 91%) and *N. caudatum* placed below a terminal clade containing *N. lecomtei* as sister to a polytomy of the three *N. mekongense* samples plus *N. delavayi*.

The genus formed a clade with *Cassytha*, this relationship receiving a posterior probability of 0.55 and 85% bootstrap support. The Perseeae-Laureae clade was well supported and its sister relationship with *Chlorocardium* Rohwer, H.G. Richt. & van der Werff was validated (PP = 1.00, BP = 100%). Together with *Caryodaphnopsis*, they were sister to the *Neocinnamomum-Cassytha* clade. All of the ingroup taxa were well-supported relative to the Cryptocaryeae outgroup (PP = 1.00, BP = 100%).

Partial combined analyses

The cladogram (Fig. 2) generated by MP analysis based on the partial combined matrix was somewhat similar to that of the complete combined matrix, but with less resolution. First, with *N. lecomtei* excluded, only *N. fargesii* was separated from the other species within *Neocinnamomum*, all of which formed a polytomy. The sister relationship between *Caryodaphnopsis* and the Perseeae-Laureae-*Chlorocardium* clade received less than 50% bootstrap support.

In contrast, the Bayesian inference cladogram was quite different (Fig. 3). *Neocinnamomum* and *Caryodaphnopsis*

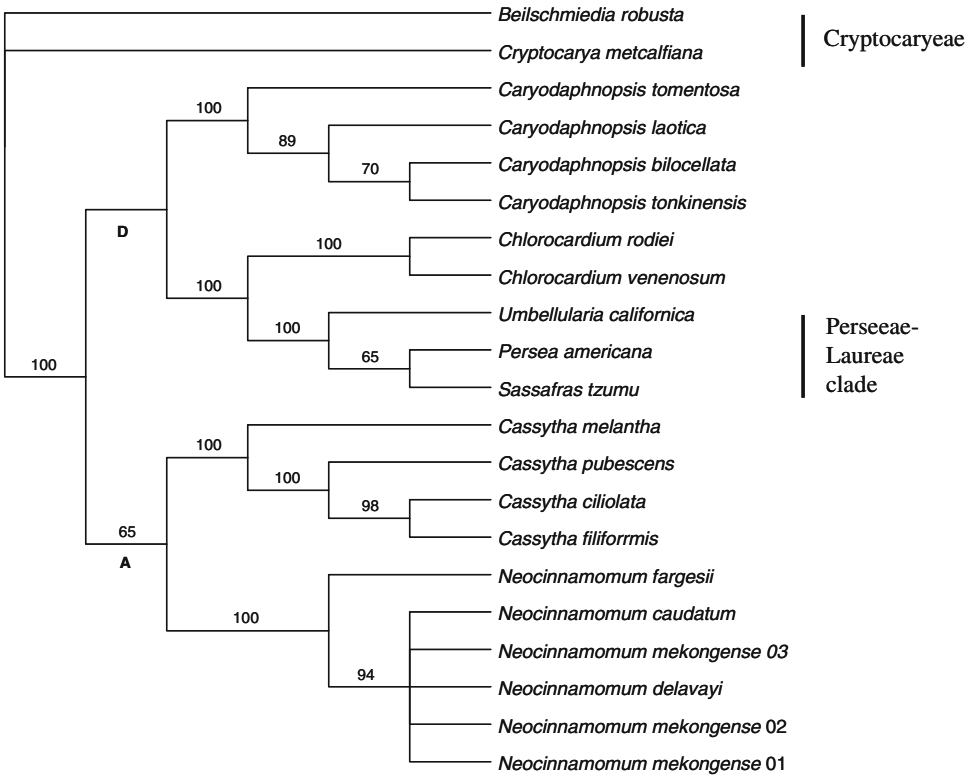

Fig. 2 Strict consensus tree for the partial combined dataset. Bootstrap values greater than 50% are shown *above the branch*. *Letters below branches* refer to those in the LBA network consensus diagrams (Fig. 4)

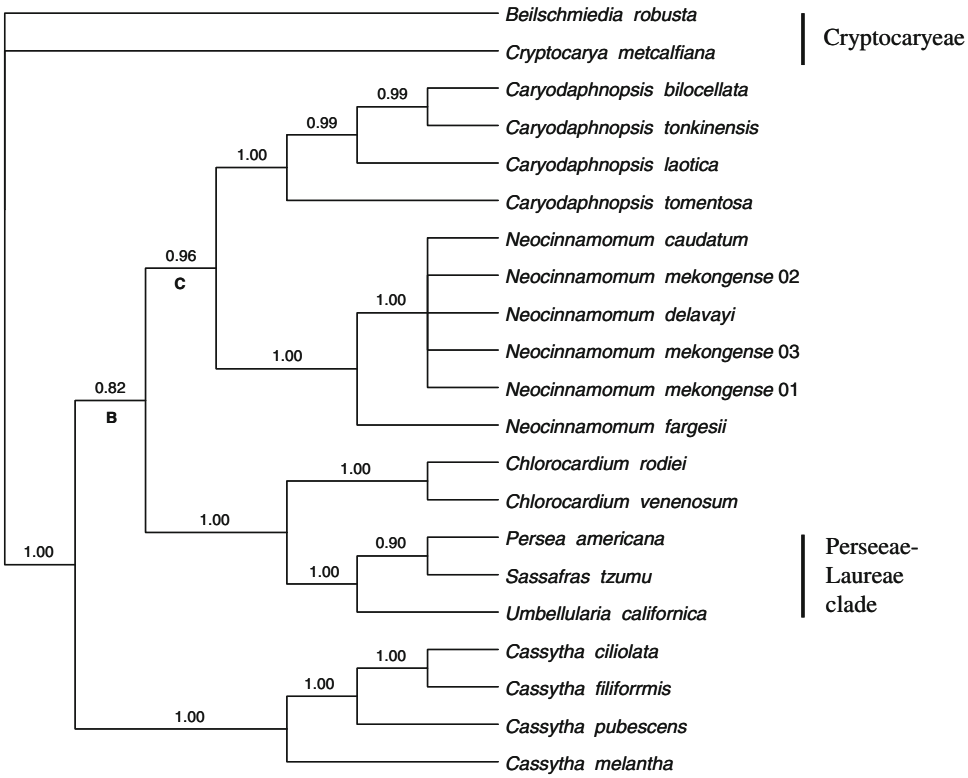

Fig. 3 Bayesian 50% majority rule consensus tree for the partial combined *psb*A–*trn*H, *trn*K, and ITS dataset. Bayesian posterior probabilities are shown *above the branches*. *Letters below branches* refer to those in the LBA network consensus diagrams (Fig. 4)

were sister clades (PP = 0.96), and together were sister to the Perseeae-Laureae-*Chlorocardium* clade (PP = 0.82). *Cassytha* was also strongly supported as sister to this group (PP = 1.00).

Consensus network

The consensus networks derived from the analyses (Fig. 4) were simplified to reflect the main crown lineages. The

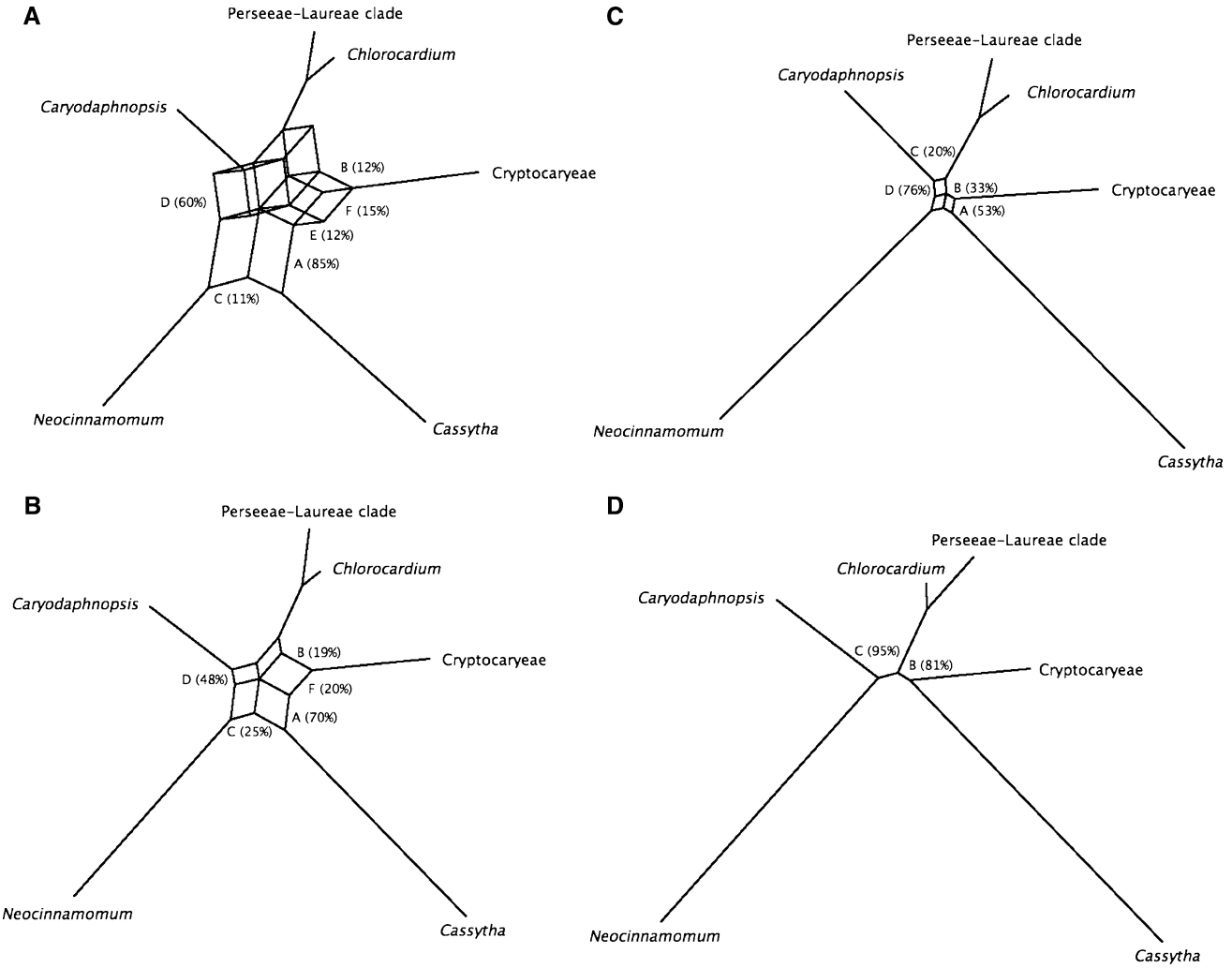

Fig. 4 Consensus networks derived from trees produced by maximum parsimony (MP) bootstrap and Bayesian MCMC analyses using a threshold of 10%. Consensus networks **a** and **b** were computed from the trees produced during 500 MP bootstrap replicates, based on the complete and partial combined *psb*A–*trn*H, *trn*K, and ITS datasets, respectively. Consensus network **c** and **d** were computed from the MCMC sampled trees, based on the complete and partial combined datasets, respectively. The lengths of the splits and edges are proportional to the tree branch lengths. *Numbers* are the percentage support for those scenarios

edges and splits (parallel edges) in consensus networks are equal to branches or clades in phylogenetic trees, and the splits represent the conflicting signals. As an example, for the conflict displayed in Fig. 4a, the 3D box reflects the uncertainty over the placement of *Caryodaphnopsis*, the three options being as follows: grouping with Cryptocaryeae (split E), or with *Neocinnamomum* (split C), or with *Chlorocardium* and the Perseeae-Laureae (split D), each with different confidence values (12, 11, and 60%, respectively).

For the consensus networks with conflicting signals, three boxes (four splits) invariably existed: the conflict of whether *Neocinnamomum* and *Cassytha* (split A) formed a sister group, or if *Neocinnamomum* and *Caryodaphnopsis* (split C) formed a sister group; the box showed two possible sister groups for *Caryodaphnopsis*, either *Neocinnamomum* (split C) or *Chlorocardium*-Perseeae-Laureae (split D). The box on the unsettled position of *Cassytha* linked together with either *Neocinnamomum* (split A) or Cryptocaryeae (split B). In contrast, the consensus network for Bayesian analysis of the partial combined matrix showed no conflict signals. Branch lengths are not shown in the MP and Bayesian trees, as the length of the splits and edges in the consensus networks are proportional to the branch lengths of their corresponding parent trees.

Discussion

Sequence characteristics

Analysis of the individual datasets showed that they gave substantially less resolved results (not shown), both for infrageneric relationships within *Neocinamomum* (due to fewer variable sites) or for intergeneric relationships. Accordingly, only the results for the combined analyses are presented here. The complete combined analysis was performed mainly to clarify infrageneric relationships within *Neocinamomum*, whereas the partial combined analysis helped to resolve the generic relationships of *Neocinamomum* and examine the influence of evolutionary rates in these taxa on LBA.

Although LBA artifacts may be overcome by combining multiple data sets, this might also reinforce the wrong tree, especially when the data are all the same kind but with high evolutionary rates (Bergsten 2005; Kennedy et al. 2005). Because the ILD test of our data showed that the three datasets were the "same kind" ($P > 0.05$), the highly variable regions of the ITS and *psb*A–*trn*H sequences were discarded for the partial analyses. The *psb*A–*trn*H noncoding region is one of the most widely used segments in phylogenetic studies, but numerous, long indels in this region have handicapped its application (Aldrich et al. 1988; Štorchová and Olson 2007), and in our case, the total indel length in *Cassytha* was close to 200 bp, possibly reflecting the accelerated rates of plastid evolution seen in many parasitic plants (Nickrent et al. 1998; Bungard 2004), although this phenomenon does not seem to have been studied in detail for *Cassytha*. In addition, the regions involving the long segments of indels are also the ones evolving fastest in the *psb*A–*trn*H spacer. Because the positions of the indels are difficult to determine and we don't know how they may influence the phylogenetic analysis, as a conservative choice, the regions involving long indel segments were discarded, leaving only the end regions near to *psb*A and *trn*H in the partial combined matrix. Similarly, only the 5.8S region of ITS was included in the partial combined matrix as the available flanking 18S nrDNA and 26S nrDNA region sequences were too short.

Monophyly and infrageneric phylogeny of *Neocinnamomum*

The monophyly of *Neocinnamomum* was strongly supported. This is consistent with its distinctive wood and bark anatomy (Richter 1981) that could not be accommodated satisfactorily into any of the three Lauraceae wood types proposed by van der Werff and Richter (1996). In terms of morphology, *Neocinnamomum* can be also easily separated from other Lauraceae by the characteristic combination of the inflorescence usually strongly reduced to condensed few-flowered thyrses; a fleshy shallow fruit cup with persistent enlarged tepals; and triplinerved, distichously alternate (not spiral) leaves (Kostermans 1974a; Li et al. 1984, 2008b).

The infrageneric phylogeny of *Neocinnamomum* is based mainly on trees from the complete combined matrix because the ITS sequence for *N. lecomtei* and the variable regions of the ITS and *psb*A–*trn*H datasets were excluded in the partial combined matrix. *Neocinnamomum fargesii* was the most proximal species. It can be separated from the remainder by its only sparsely pilose flowers and bracts (also seen in the branchlets and leaves). The analyses then placed *N. caudatum* as sister to the remainder, and it is the only species in the genus with a compound thyrse consisting of a main peduncle with much reduced side branchlets of few- to many-flowered, condensed cymes. The distal taxon placement showed the densely pilose *N. lecomtei* to be sister to *N. delavayi* and *N. mekongense*.

Neocinnamomum delavayi and *N. mekongense* are difficult to distinguish on both morphology and molecular characteristics (Fig. 1), and *N. mekongense* has been treated previously as a variety of *N. delavayi* (Handel-Mazzetti 1925), although Kostermans (1974a) raised it to specific rank based on the possession of larger leaves and completely glabrous branchlets. However, these two species are so similar morphologically that it is difficult to separate them, especially in the herbarium, as leaf size ranges overlap and the silvery sericeous hairs of *N. delavayi* are much less obvious on older leaves and branchlets. Given that the distribution of *N. mekongense* in western and northwestern Yunnan and southeastern Xizang completely overlaps with that of *N. delavayi* (Yunnan, southeastern Xizang and southern Sichuan) (Li et al. 1984, 2008b), we support the reduction of *N. mekongense* to a localized variant of *N. delavayi*, as proposed by Handel-Mazzetti (1925). In addition, the variation seen among the three samples of *N. mekongense* used in this study shows that parallel secondary veins are not restricted to *N. caudatum,* so the character should be used with caution for identification.

Although reliable morphological characters for classification in Lauraceae are relatively sparse, inflorescence structure is generally accepted as useful for inter- and suprageneric classification (Rohwer et al. 1991; van der Werff and Richter 1996; van der Werff 2001; Li and Li 2004; Li et al. 2004, 2006, 2007; Julia et al. 2009). Nevertheless, a detailed understanding of inflorescence patterns and development is essential in Lauraceae (Li et al. 2004).

In *Neocinnamomum*, there are two possible primitive inflorescence types: the compound thyrse with condensed cymes (Fig. 5b) seen in *N. caudatum*, or a few-flowered condensed inflorescence with a strongly reduced peduncle

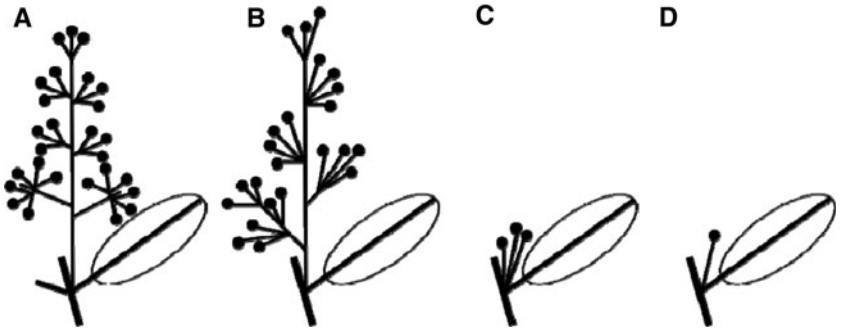

Fig. 5 Inflorescence types in *Neocinnamomum* and *Caryodaphnopsis*. A thyrse in *Caryodaphnopsis* (**a**), a compound thyrse with condensed lateral cymes as in *N. caudatum* (**b**), few-flowered, condensed inflorescence with poorly defined branching, as seen in most species of *Neocinnamomum* (**c**), a 1-flowered inflorescence as occurs in *N. atjehense* (**d**)

and pedicellate flowers (the remainder; Fig. 5c). If the latter condition is ancestral, it only requires one change from the few-flowered condensed inflorescence to the compound thyrse of *N. caudatum* when inflorescence patterns are mapped onto the tree topology (Fig. 1).

In contrast, if thyrses with axillary condensed cymes are ancestral, two changes are required to create the condensed, few-flowered inflorescence seen in most species of the genus: first from a compound thyrse to the few-flowered condensed inflorescence of *N. fargesii*; second from the compound thyrse exhibited in *N. caudatum* to the few-flowered condensed inflorescences of the remainder. However, although the second hypothesis appears to be more parsimonious when using *Neocinnamomum* alone, an ancestral condition of a thyrse with condensed cymes is more plausible when the inflorescence of the sister genus *Caryodaphnopsis* is included for comparison.

The inflorescences of *Caryodaphnopsis* are mainly thyrses, corresponding to the type II "cymose-panicles" of van der Werff and Richter (1996) and are similar to most members of the Perseeae (Kostermans 1957, 1974b; Rohwer 1993). However, in some species, such as *C. tonkinensis* (Lec.) Airy Shaw, *C. inaequalis* (A.C.Sm.) van der Werff & H.G.Richt., *C. metallica* Kosterm., and *C. tomentosa* van der Werff, the cymes of the thyrses are subopposite (Fig. 5a) (Kostermans 1974b; van der Werff and Richter 1985; van der Werff 1991) corresponding more closely to the type III inflorescences of van der Werff and Richter (1996). Furthermore, in *C. tonkinensis* the cymes of the thyrses possess short to very short branchlets and flowers sometimes attached directly to the main peduncle (Kostermans 1974b). This reduced inflorescence form could be changed easily to the compound thyrse with condensed branches seen in *N. caudatum* by reduction of the pedicels and branchlets.

Compound thyrses with condensed cymes and/or 1-flowered inflorescences also occur occasionally in *N. delavayi*, *N. mekongense*, and possibly *N. lecomtei* and *N. fargesii*. An evolutionary series of inflorescences might therefore be postulated for the genus as follows: an ancestral compound thyrse becomes reduced to a few- to many-flowered condensed inflorescence and ultimately to a 1-flowered inflorescence (Fig. 5). According to the tree topology, the change from compound thyrse to a few-flowered condensed inflorescence appears to have happened twice.

Long-branch attraction

LBA, the erroneous grouping of two or more long branches as sister groups due to methodological artifacts, is a well known source of misleading phylogeny (Holland et al. 2004, 2005; Bergsten 2005; Kennedy et al. 2005). Using long branch extraction and Bayesian inference (in small datasets also maximum likelihood), Rohwer and Rudolph (2005) suggested that a sister group relationship between *Neocinnamomum* and *Cassytha* may be the result of LBA and that *Neocinnamomum* was closer to *Caryodaphnopsis*, even though the latter relationship was not strongly supported. In order to obtain a more reliable phylogeny for *Neocinnamomum* and its sister group, we applied consensus network to examine the conflicting signals of trees produced by MP and Bayesian inference.

As in Rohwer and Rudolph (2005), in all analyses that we performed *Neocinnamomum* and *Cassytha* had the longest branches (Fig. 4). In the consensus networks based on complete combined analyses (Fig. 4a, b), either from MP bootstrap replicates or MCMC sampled trees, conflict signals over the placement of *Neocinnamomum* relative to *Cassytha* are present in the corresponding MP and Bayesian trees. Performing the analysis on a complete combined matrix should be expected to improve its reliability. However, no method is completely immune to LBA (Kennedy et al. 2005; Bergsten 2005), including Bayesian

inference and maximum likelihood, so that the topology concordance from different methods is not a guarantee against LBA.

Because LBA artifacts may be reinforced by combining multiple data sets of the same kind but with high evolutionary rates (Bergsten 2005; Kennedy et al. 2005), a partial combined analysis with the more variable regions excluded may help to minimize such artifacts. In the consensus networks computed from the trees collected during MP bootstrap replicates for the partial combined matrix (Fig. 4c), conflict signals can still be detected for the placement of *Neocinnamomum*; however, a relationship between *Neocinnamomum* and *Cassytha* (albeit not very close) received the strongest support. In contrast, in the first three consensus networks (Fig. 4a–c), the signals of the sister relationship between *Neocinnamomum* and *Caryodaphnopsis* can always be detected and the consensus network for the Bayesian analyses of the partial combined matrix (Fig. 4d) also affirmed a sister relationship between *Neocinnamomum* and *Caryodaphnopsis*. This indicates that the relationship of *Neocinnamomum* and *Cassytha* in previous analyses was influenced by LBA, with a sister relationship of *Neocinnamomum* and *Caryodaphnopsis* more likely.

Intergeneric relationships

Though a sister relationship between *Caryodaphnopsis* and *Neocinnamomum* is credible, the branch lengths connecting them are short. Chanderbali et al. (2001) suggested that *Caryodaphnopsis* and *Neocinnamomum* may be the descendants of Laurasian Cretaceous Lauraceae, with the fossil wood taxon *Caryodaphnopsoxylon richteri* Gottwald (1992) and fossil flower *Neusenia tetrasporangiata* Eklund (2000) comparing favorably with both of them. However, this suggestion of an ancient origin implies that there will probably be few extant characteristics common between *Caryodaphnopsis* and *Neocinnamomum*. Nevertheless, some connections can be seen in their morphology, wood, and bark anatomy.

Septate wood fibers and sparse to incompletely vasicentric wood parenchyma link *Neocinnamomum* with *Caryodaphnopsis* and the Perseeae (Richter 1981), whereas simple vessel perforations are shared among *Neocinnamomum*, *Caryodaphnopsis*, and the Cryptocarya group (Richter 1981). *Neocinnamomum* is nevertheless somewhat isolated, possessing the smallest intervascular pits in Lauraceae besides *Chlorocardium*. In contrast, *Caryodaphnopsis* shares more wood and bark anatomical similarities with Richter's (1981) *Cryptocarya* group.

Morphologically, *Neocinnamomum* and *Caryodaphnopsis* are similar, sharing triplinerved venation, a very short, small perianth tube, 4-locular anthers, and fruits almost free on a rather swollen pedicel (Li et al. 1984; Rohwer and Rudolph 2005). Added to this is their similarity in inflorescence structure, with some species of *Caryodaphnopsis* possessing more or less similar inflorescences to *N. caudatum* as discussed above. Because *Neocinnamomum*, *Caryodaphnopsis*, and *Cassytha* are distal to the Cryptocaryeae, the ancestral inflorescence form for *Neocinnamomum* and *Caryodaphnopsis* may have been derived from the inflorescences now seen in the Cryptocaryeae and related groups. However, there is also the possibility that these inflorescences arose independently from a common ancestral inflorescence form (as yet uncertain), with subsequent divergence. In either case, interpretation of inflorescence evolution for *Cassytha* is difficult, as it has become very specialized in many aspects, including its inflorescence. Nevertheless our study using three gene regions and wider sampling agrees with the *mat*K study of Rohwer and Rudolph (2005) that *Cassytha* apparently lacks close living relatives.

In conclusion, the multi-gene sister relationship seen between *Neocinnamomum* and *Cassytha* is apparently the result of LBA, supporting the *mat*K-based findings of Rohwer and Ruldoph (2005). However, it may be more accurate to say that on the basis of our study, *Neocinnamomum* is more closely related to *Caryodaphnopsis* than it is to *Cassytha*. *Cassytha* is clearly related to the core Cryptocaryeae and *Caryodaphnopsis*, albeit still very different from these taxa and with its position still uncertain, probably due in part to the extreme morphological and genetic changes resulting from its hemiparasitic lifestyle.

Acknowledgments We would like to thank Prof. Jens G. Rohwer for providing assistance with translating the wood and bark anatomical papers published in German. This research was supported by the Chinese Academy of Sciences (KSCX2-YW-Z-001) and National Natural Science Foundation of China (30870170).

References

Aldrich J, Cherney BW, Merlin E (1988) The role of insertions/deletions in the evolution of the intergenic region between *psb*A and *trn*H in the chloroplast genome. Curr Genet 14:137–146

Bergsten J (2005) A review of long-branch attraction. Cladistics 21:163–193

Buckler ES, Holtsford TP (1996) *Zea* systematics: ribosomal ITS evidence. Molec Biol Evol 13:612–622

Buckler ES, Ippolito A, Holtsford TP (1997) The evolution for ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145:821–832

Bungard RA (2004) Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. Bioessays 26:235–247

Chanderbali AS, van der Werff H, Renner SS (2001) Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. Ann Missouri Bot Gard 88:104–134

Doyle JJ, Doyle JS (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Drummond AJ, Rambaut A (2007) Beast: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Eklund H (2000) Lauraceous flowers from the Late Cretaceous of North Carolina, USA. Bot J Linn Soc 132:397–428

Farris JS, Källersjö M, Kluge AG, Bult CJ (1995) Constructing a significance test for incongruence. Syst Biol 44:570–572

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Gottwald H (1992) Hölzer aus marinen Sanden des oberen Eozän von Helmstedt (Niedersachsen). Palaeontographica B 225:27–103

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98

Handel-Mazzetti H (1925) *Cinnamomum delavayi* Lecomte var. *mekongense*. Anz Akad Wiss Wien Math-Naturwiss Kl 62:218

Holland BR, Huber KT, Moulton V, Lockhart PJ (2004) Using consensus networks to visualize contradictory evidence for species phylogeny. Molec Biol Evol 21:1459–1461

Holland BR, Delsuc F, Moulton V, Baker AJ (2005) Visualizing conflicting evolutionary hypotheses in large collections of trees: using consensus networks to study the origins of placentals and hexapods. Syst Biol 54:66–76

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310–2314

Huson D (1998) SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14:68–73

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Molec Biol Evol 23:254–267

Johnson LA, Soltis DE (1995) Phylogenetic inference in Saxifragaceae *sensu stricto* and *Gilia* (Polemoniaceae) using *mat*K sequences. Ann Missouri Bot Gard 82:149–175

Julia S, Soepadmo E, Yahud W (2009) Problem in the generic delimitation between *Alseodaphne*, *Dehaasia* and *Nothaphoebe* (Lauraceae) in Borneo. Blumea 54:192–197

Kennedy M, Holland BR, Gray RD, Spencer HG (2005) Untangling long branches: identifying conflicting phylogenetic signals using spectral analysis, neighbor-net, and consensus networks. Syst Biol 54:620–633

Kostermans AJGH (1957) Lauraceae. Pengumuman Balai Besar Penjelidikan Kehutanan Indonesia 57:1–64

Kostermans AJGH (1974a) A monograph of the genus *Neocinnamomum* Liou Ho. Reinwardtia 9:85–96

Kostermans AJGH (1974b) A monograph of *Caryodaphnopsis* A. Shaw. Reinwardtia 9:123–137

Kunming Botanical Garden, Kunming Institute of Botany, Chinese Academy of Sciences (2006) An enumeration of plants growing in Kunming Botanical Garden. Yunnan Science and Technology Press, Kunming, China

Li J, Li H-W (2004) Advances in Lauraceae systematic research on the world scale. Acta Bot Yunnan 26:1–11

Li H-W, Pai PY, Lee SK, Wei FN, Wei YT, Yang YC, Huang PH, Tsui HP, Shia ZD, Li JL (1984) Lauraceae. In: Li H-W (ed) Flora Republicae Popularis Sinicae, vol. 31. Science Press, Beijing

Li J, Christophel DC, Conran JG, Li H-W (2004) Phylogenetic relationships within the *Litsea* complex (Lauraceae) inferred from sequences of the chloroplast gene *mat*K and nuclear ribosomal DNA ITS regions. Pl Syst Evol 246:19–34

Li Z-M, Li J, Li H-W (2006) Polyphyly of the genus *Actinodaphne* (Lauraceae) inferred from the analyses of nrDNA ITS and ETS sequences. Acta Phytotax Sinica 44:272–285

Li L, Li J, Conran JG, Li H-W (2007) Phylogeny of *Neolitsea* (Lauraceae) inferred from Bayesian analysis of nrDNA ITS and ETS sequences. Pl Syst Evol 269:203–221

Li J, Conran JG, Christophel DC, Li Z-M, Li L, Li H-W (2008a) Phylogenetic relationships of the *Litsea* complex and core Laureae (Lauraceae) using ITS and ETS sequences and morphology. Ann Missouri Bot Gard 95:580–599

Li X-W (Li H-W), Li J, van der Werff H (2008b) *Neocinnamomum* H. Liou. In: Wu Z-Y, Raven, PH, Hong, D-Y (eds) Flora of China, vol. 7. Science Press and Missouri Botanical Garden Press, Beijing and St. Louis, pp. 187–189

Liu H (Liou H) (1934) Lauracées de Chine et d'Indochine: contribution à l'étude systématique et phytogéographique. Hermann et Cie, Paris

Nickrent DL, Duff RJ, Colwell AE, Wolfe AD, Young ND, Steiner KE, de Pamphilis CW (1998) Molecular phylogenetic and evolutionary studies of parasitic plants. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular systematics of plants II. DNA sequencing. Kluwer Academic, Boston, pp 211–241

Philippe H, Zhou Y, Brinkmann H, Rodrigue N, Delsuc F (2005) Heterotachy and long-branch attraction in phylogenetics. BMC Evol Biol 5:50

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818

Poux C, Madsen O, Glos JW, de Jong W, Vences M (2008) Molecular phylogeny and divergence times of Malagasy tenrecs: influence of data partitioning and taxon sampling on dating analyses. BMC Evol Biol 8:102

Richter HG (1981) Anatomie des sekundären Xylems und der Rinde der Lauraceae. Sonderb Naturwiss Vereins Hamburg 5:1–148

Rohwer JG (1993) Lauraceae. In: Kubitzki K, Rohwer JG, Bittrich V (eds) The families and genera of vascular plants, vol 2. Springer, Berlin, pp 366–391

Rohwer JG (2000) Toward a phylogenetic classification of the Lauraceae: evidence from *mat*K sequences. Syst Bot 25:60–71

Rohwer JG, Rudolph B (2005) Jumping genera: the phylogenetic positions of *Cassytha*, *Hypodaphnis*, and *Neocinnamomum* (Lauraceae) based on different analyses of *trn*K intron sequences. Ann Missouri Bot Gard 92:153–178

Rohwer JG, Richter HG, van der Werff H (1991) Two new genera of Neotropical Lauraceae and critical remarks on their generic delimitation. Ann Missouri Bot Gard 78:388–400

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

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics: methods and protocols (methods in molecular biology). Humana, Totowa, pp 365–386

Sang T, Crawford DJ, Stuessy TF (1995) Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proc Natl Acad Sci USA 92:6813–6817

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am J Bot 84:1120–1136

Steele KP, Vilgalys R (1994) Phylogenetic analysis of Polemoniaceae using nucleotide sequences of plastid gene *mat*K. Syst Bot 19:126–142

Štorchová H, Olson MS (2007) The architecture of the chloroplast *psb*A-*trn*H non-coding region in angiosperms. Pl Syst Evol 268:235–256

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer and Associates, Sunderland

van der Werff H (1991) New species of Lauraceae from Ecuador and Peru. Ann Missouri Bot Gard 78:409–423

van der Werff H (2001) An annotated key to the genera of Lauraceae in the *Flora Malesiana* region. Blumea 46:125–140

van der Werff H, Richter HG (1985) *Caryodaphnopsis* Airy-Shaw (Lauraceae), a genus new to the neotropics. Syst Bot 10: 166–173

van der Werff H, Richter HG (1996) Toward an improved classification of Lauraceae. Ann Missouri Bot Gard 83:409–418

White TJ, Bruns TD, Lee SB, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand GH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, San Diego, pp 315–322