Manuscript title: **Plastome evolution in the sole hemiparasitic genus laurel dodder (*Cassytha*) and insights into the plastid phylogenomics of Lauraceae**

Authors: Chung-Shien Wu, Ting-Jen Wang, Chia-Wen Wu, Ya-Nan Wang, Shu-Miaw Chaw

Supplementary figures and tables.

Supplementary figure 1. Dot-plot analyses of the plastomes of *Neo. delavayi* and *Cin. camphora* (A), *Cin. camphora* and *Cry. chinensis* (B), and *Cas. filiformis* and *Cin. camphora*.

Supplementary figure 2. PCR assays for detection of the two reduced fragments in *Cas. filiformis* and *Cas. pubescens*. Compared to the corresponding region in *Cin. camphora*, reduction of the two examined fragments is due to deletion of the two segments *3'ndhB–Ψycf2* and *ndhB–Ψycf1* in the plastome of *Cas. filiformis*. As a result, our PCR amplicons with the expected size indicate that both *Cas. filiformis* and *Cas. pubescens* have lost the two segments that are nevertheless, present in the IR of *Cin. camphora*. Green and blue arrows are the primer pairs used for the PCR assays.

Supplementary figure 3. Presence, pseudogenization, or absence of genes in the plastomes of *Cas. filiformis*, *Cin. camphora*, *Cry. chinensis*, and *Neo. delavayi*. Genes in bold are the 67 common protein-coding genes.

Supplementary figure 4. Box-plots comparing the length of syntenic non-genic loci among the plastomes of *Cas. filiformis*, *Cin. camphora*, *Cry. chinensis*, and *Neo. delavayi*.

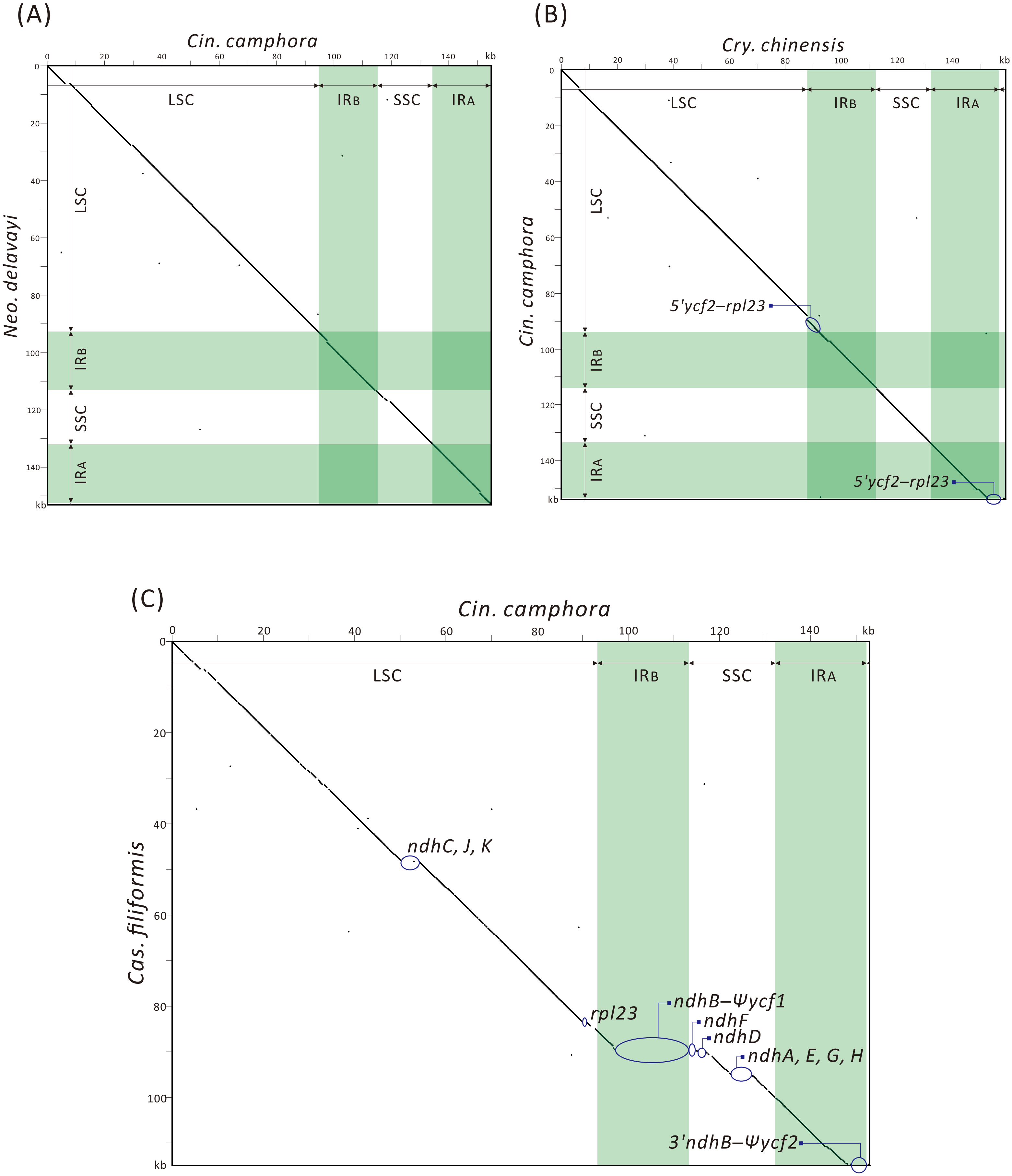
Supplementary figure 5. Two possible scenarios for IR evolution in Lauraceae.

Supplementary figure 6. Trees based on non-synonymous (d_N) or synonymous (d_S) sites in the concatenated house-keeping (HK) or photosynthetic (PS) genes. All tree topologies were constrained as for the ML tree shown in Fig. 2. The species of Lauraceae are in green and the branch leading to *Cas. filiformis* is denoted by "*".

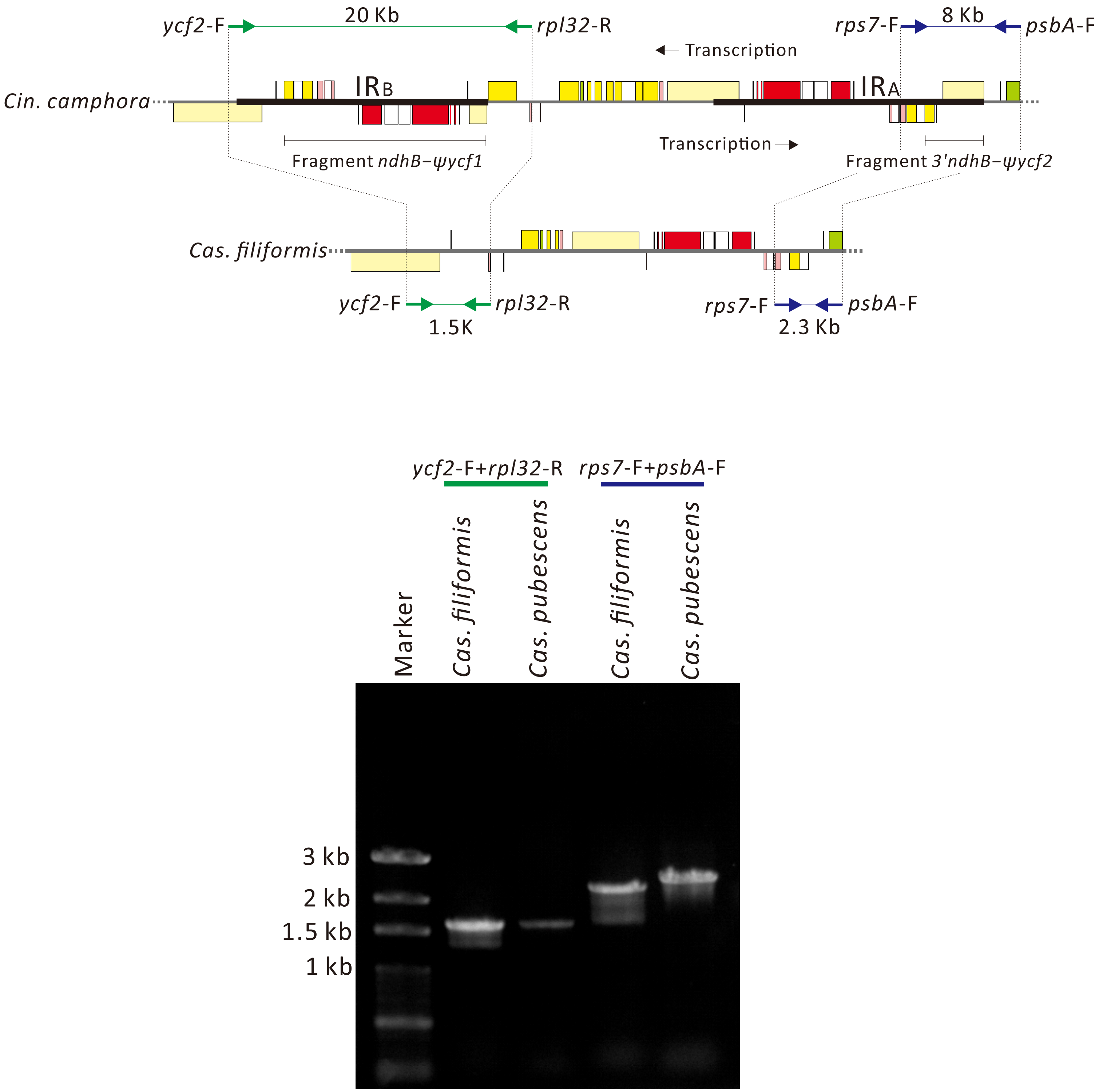
Supplementary figure 7. Estimated divergence times across 16 magnoliid species. MYA: million years ago.

Supplementary figure 8. Maximum parsimony tree inferred from the 67 shared plastid genes with *Amb. trichopoda* as the outgroup. Bootstrap support values estimated from 1,000 replicates are shown along branches. The consistency index (CI) is 0.545 and the retention index (RI) is 0.633. The scale bar for branches are indicated.

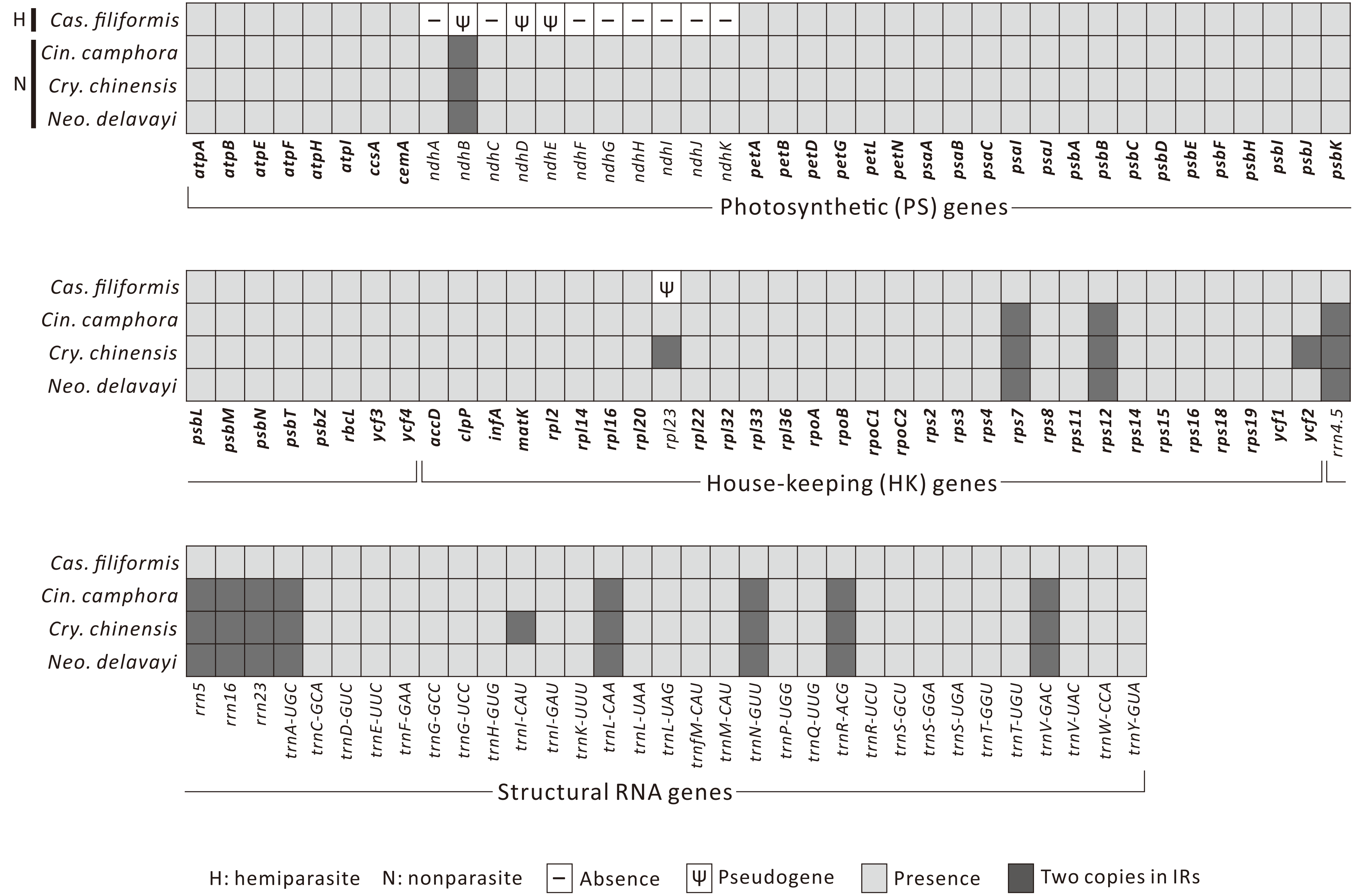
Supplementary figure 1. Dot-plot analyses of the plastomes of *Neo. delavayi* and *Cin. camphora* (A), *Cin. camphora* and *Cry. chinensis* (B), and *Cas. filiformis* and *Cin. camphora*.



Supplementary figure 2. PCR assays for detection of the two reduced fragments in *Cas. filiformis* and *Cas. pubescens*. Compared to the corresponding region in *Cin. camphora*, reduction of the two examined fragments is due to deletion of the two segments *3'ndhB–Ψycf2* and *ndhB–Ψycf1* in the plastome of *Cas. filiformis*. As a result, our PCR amplicons with the expected size indicate that both *Cas. filiformis* and *Cas. pubescens* have lost the two segments that are nevertheless, present in the IR of *Cin. camphora*. Green and blue arrows are the primer pairs used for the PCR assays.



Supplementary figure 3. Presence, pseudogenization, or absence of genes in the plastomes of *Cas. filiformis*, *Cin. camphora*, *Cry. chinensis*, and *Neo. delavayi*. Genes in bold are the 67 common protein-coding genes.



		atpA	***atpB***	***atpE***	***atpF***	***atpH***	***atpI***	***ccsA***	***cemA***	*ndhA*	*ndhB*	*ndhC*	*ndhD*	*ndhE*	*ndhF*	*ndhG*	*ndhH*	*ndhI*	*ndhJ*	*ndhK*	***petA***	***petB***	***petD***	***petG***	***petL***	***petN***	***psaA***	***psaB***	***psaC***	***psaI***	***psaJ***	***psbA***	***psbB***	***psbC***	***psbD***	***psbE***	***psbF***	***psbH***	***psbI***	***psbJ***	***psbK***
H	*Cas. filiformis*	░	░	░	░	░	░	░	░	–	Ψ	–	Ψ	Ψ	–	–	–	–	–	–	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░
N	*Cin. camphora*	░	░	░	░	░	░	░	░	░	▓	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░
	Cry. chinensis	░	░	░	░	░	░	░	░	░	▓	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░
	Neo. delavayi	░	░	░	░	░	░	░	░	░	▓	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░

Photosynthetic (PS) genes

	psbL	***psbM***	***psbN***	***psbT***	***psbZ***	***rbcL***	***ycf3***	***ycf4***	***accD***	***clpP***	***infA***	***matK***	***rpl2***	***rpl14***	***rpl16***	***rpl20***	*rpl23*	***rpl22***	***rpl32***	***rpl33***	***rpl36***	***rpoA***	***rpoB***	***rpoC1***	***rpoC2***	***rps2***	***rps3***	***rps4***	***rps7***	***rps8***	***rps11***	***rps12***	***rps14***	***rps15***	***rps16***	***rps18***	***rps19***	***ycf1***	***ycf2***	*rrn4.5*
Cas. filiformis	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	Ψ	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░
Cin. camphora	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	▓	░	░	▓	░	░	░	░	░	░	░	▓
Cry. chinensis	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	▓	░	░	░	░	░	░	░	░	░	░	░	▓	░	░	▓	░	░	░	░	░	░	▓	▓
Neo. delavayi	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	▓	░	░	▓	░	░	░	░	░	░	░	▓

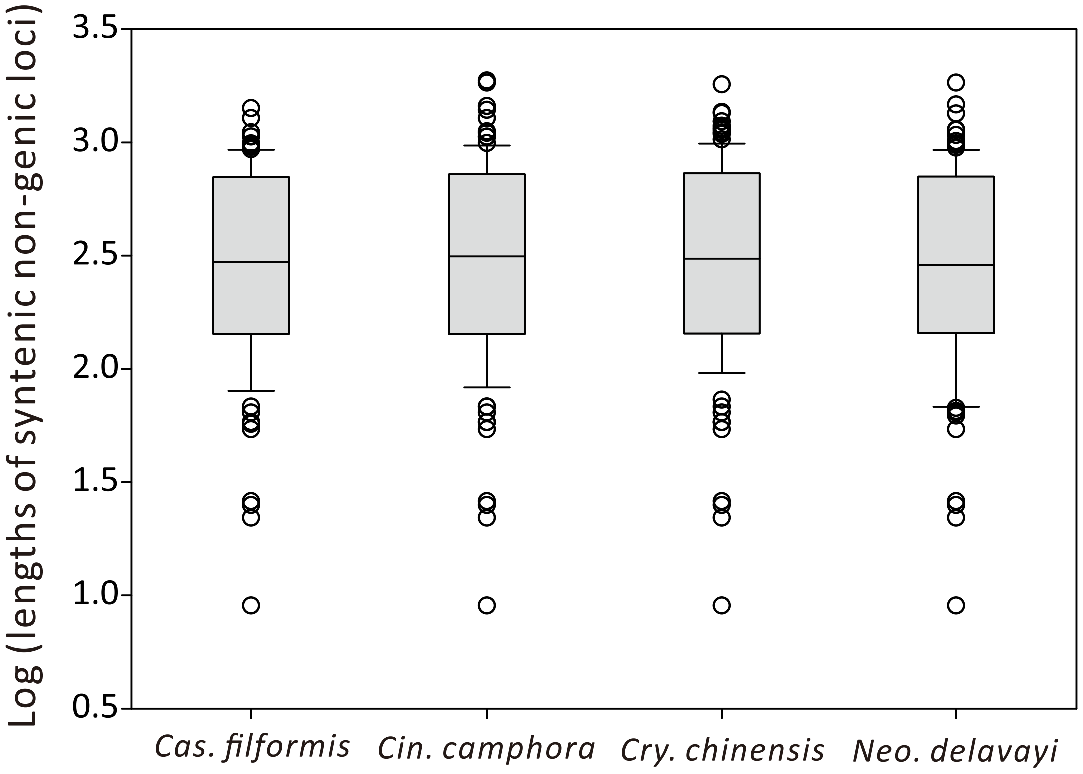
House-keeping (HK) genes

	rrn5	*rrn16*	*rrn23*	*trnA-UGC*	*trnC-GCA*	*trnD-GUC*	*trnE-UUC*	*trnF-GAA*	*trnG-GCC*	*trnG-UCC*	*trnH-GUG*	*trnI-CAU*	*trnI-GAU*	*trnK-UUU*	*trnL-CAA*	*trnL-UAA*	*trnL-UAG*	*trnfM-CAU*	*trnM-CAU*	*trnN-GUU*	*trnP-UGG*	*trnQ-UUG*	*trnR-ACG*	*trnR-UCU*	*trnS-GCU*	*trnS-GGA*	*trnS-UGA*	*trnT-GGU*	*trnT-UGU*	*trnV-GAC*	*trnV-UAC*	*trnW-CCA*	*trnY-GUA*
Cas. filiformis	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░	░
Cin. camphora	▓	▓	▓	▓	░	░	░	░	░	░	░	░	░	░	▓	░	░	░	░	▓	░	░	▓	░	░	░	░	░	░	▓	░	░	░
Cry. chinensis	▓	▓	▓	▓	░	░	░	░	░	░	░	▓	░	░	▓	░	░	░	░	▓	░	░	▓	░	░	░	░	░	░	▓	░	░	░
Neo. delavayi	▓	▓	▓	▓	░	░	░	░	░	░	░	░	░	░	▓	░	░	░	░	▓	░	░	▓	░	░	░	░	░	░	▓	░	░	░

Structural RNA genes

H: hemiparasite N: nonparasite – Absence Ψ Pseudogene ░ Presence ▓ Two copies in IRs

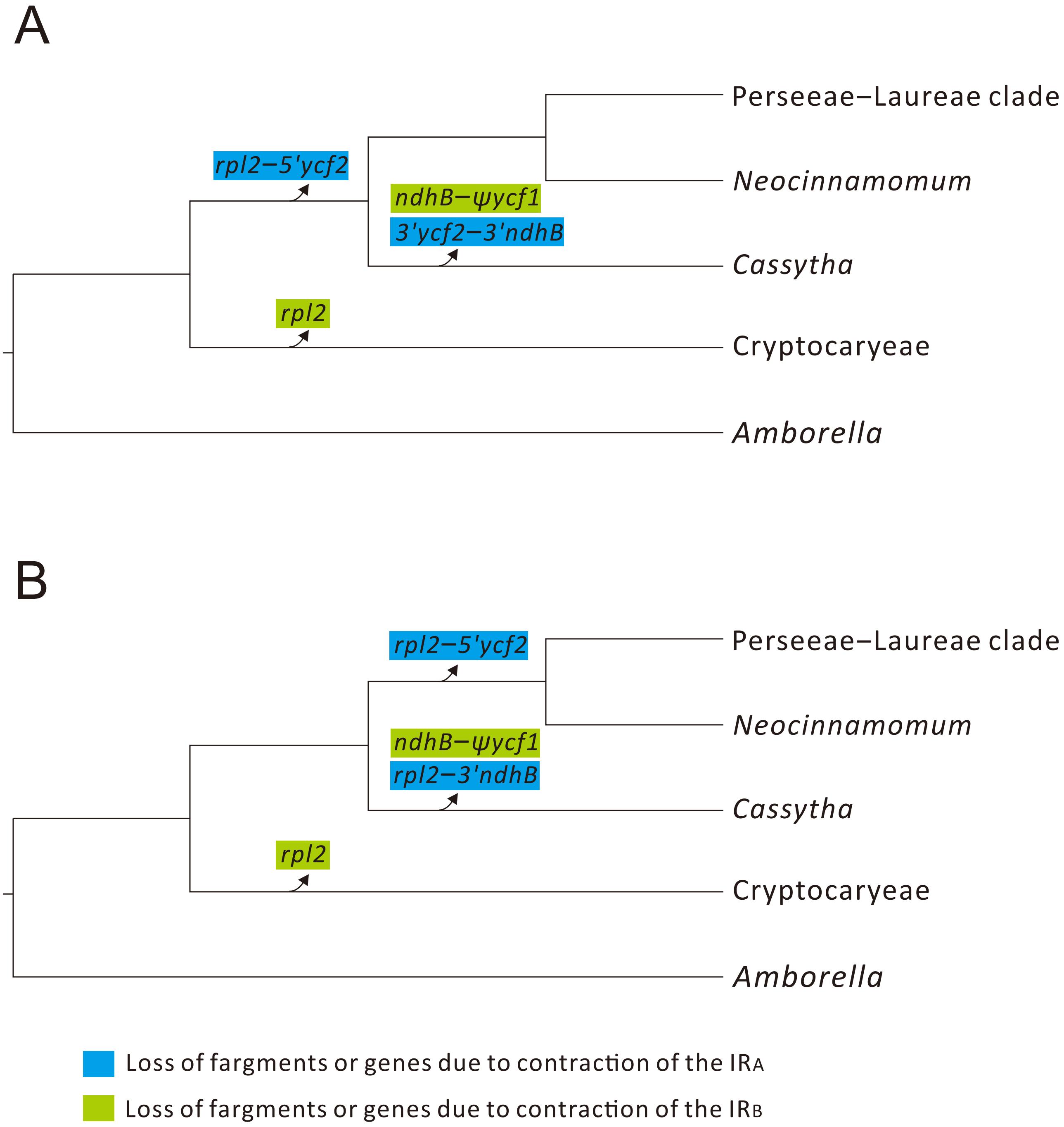
Supplementary figure 4. Box-plots comparing the length of syntenic non-genic loci among the plastomes of *Cas. filiformis*, *Cin. camphora*, *Cry. chinensis*, and *Neo. delavayi*.



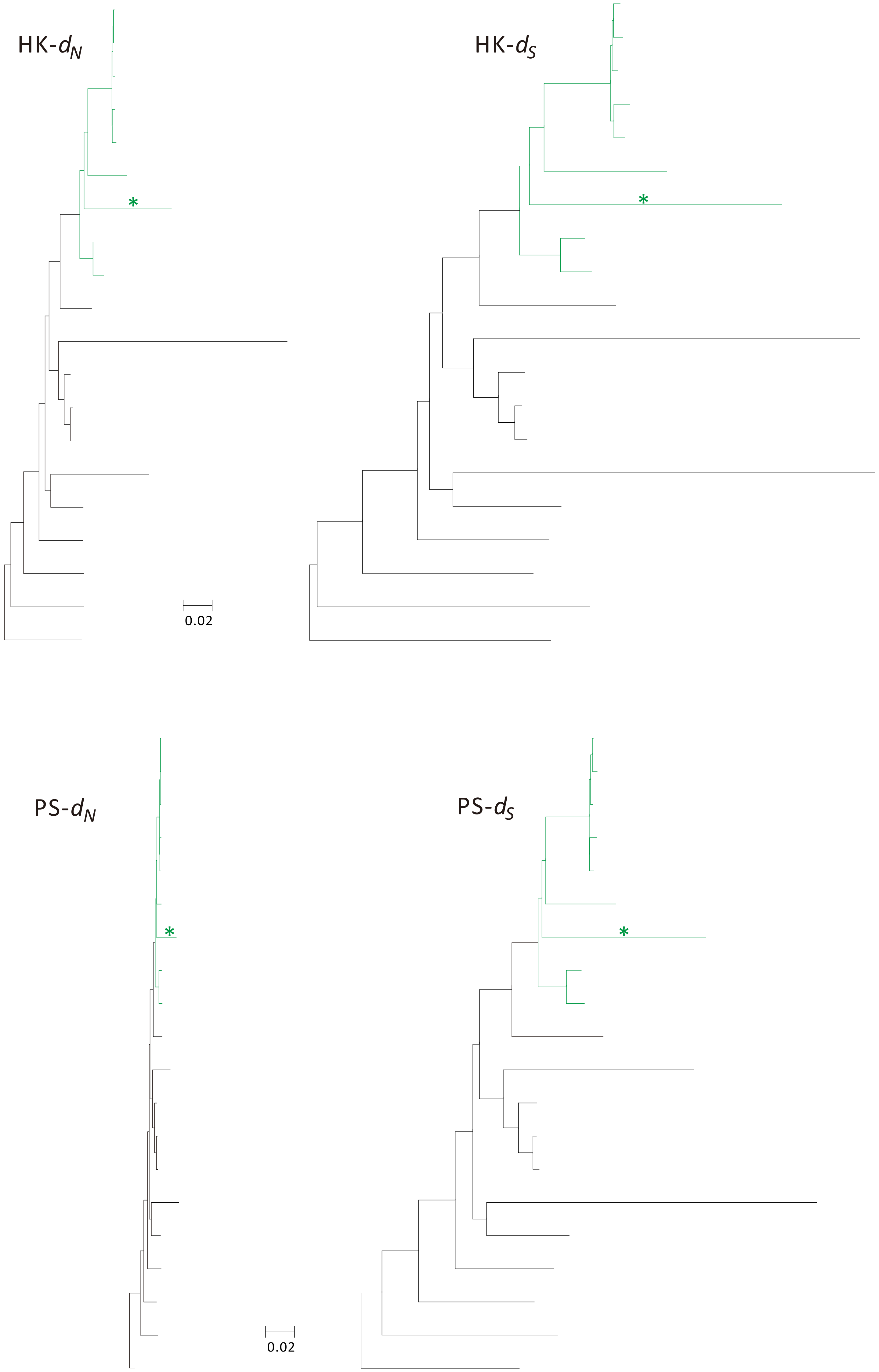
P values for 2-sided Wilcoxon rank-sum tests

Cin. camphora	0.8552		
Cry. chinensis	0.6817	0.8291	
Neo. delavayi	0.9521	0.9000	0.7110
	Cas. filiformis	*Cin. camphora*	*Cry. chinensis*

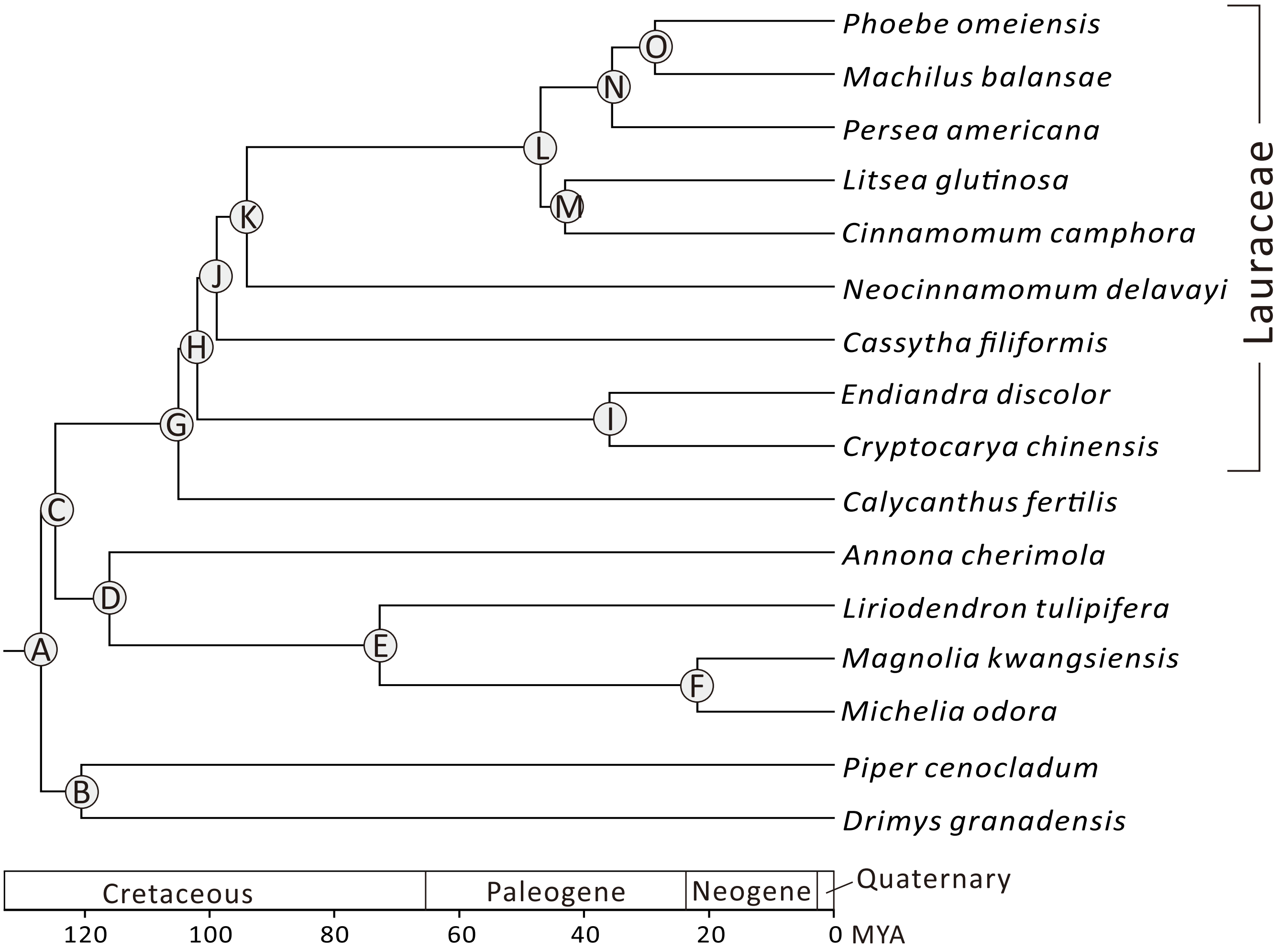
Supplementary figure 5. Two possible scenarios for the IR evolution in Lauraceae.



Supplementary figure 6. Trees based on non-synonymous (d_N) or synonymous (d_S) sites in the concatenated house-keeping (HK) or photosynthetic (PS) genes. All tree topologies were constrained as for the ML tree shown in Fig. 2. The species of Lauraceae are in green and the branch leading to *Cas. filiformis* is denoted by “*”.



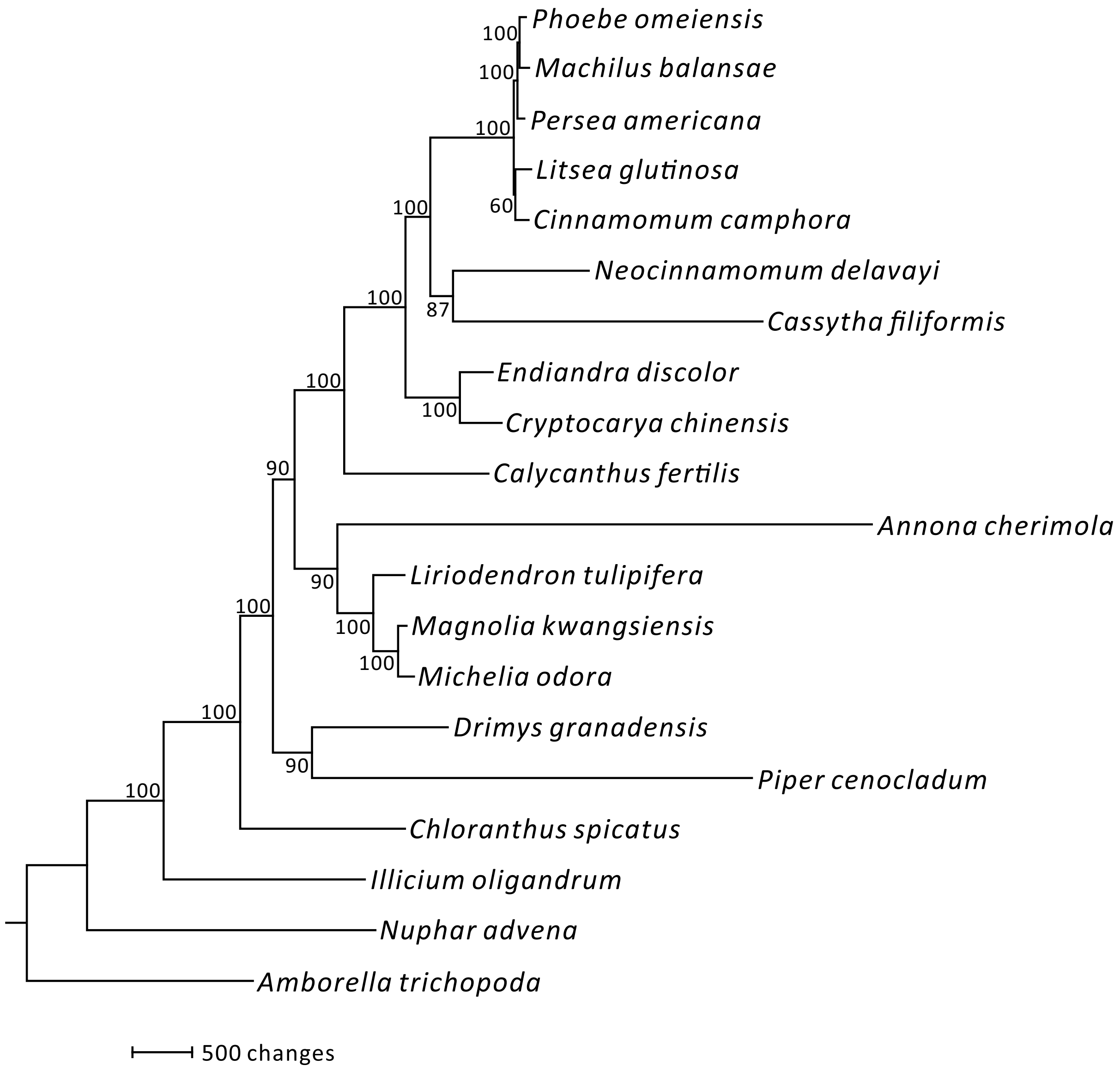
Supplementary figure 7. Estimated divergence times across 16 magnoliid species. MYA: million years ago.



Node	Divergence time (MYA)	95% confidence interval
A*	127	–
B	120.495	120.2482–120.7418
C	124.7039	124.6142–124.7936
D*	116	–
E	72.74	71.27456–74.20544
F	21.875	21.21487–22.53513
G*	105	–
H*	102	–
I	35.89455	35.44649–36.3426
J	98.84955	98.73001–98.96908
K	93.9975	93.77836–94.21664
L*	47	–
M	43.05909	42.7554–43.36279
N	35.49955	34.64733–36.35176
O	28.63386	27.70872–29.55901

*constrained age (Time Tree, Hedges et al. 2015)

Supplementary figure 8. Maximum parsimony tree inferred from the 67 shared plastid genes with *Amb. trichopoda* as the outgroup. Bootstrap support values estimated from 1,000 replicates are shown along branches. The consistency index (CI) is 0.545 and the retention index (RI) is 0.633. The scale bar for branches are indicated.



Supplementary table 1. Primers used for gap closing

Species	Primer
Cinnamomum camphora	psbA: CAGCTTGGCCTGTAGTAGGTATCTG
	ycf2: GAATAGAGGAACTCCCAAGCATTTG
	ndhF: GCAAGTTTCTTTGTAGGAGAAGG
	trnN: ACCGCTCTACCACTGAGCTACTGAG
Neocinnamomum delavayi	psbJ: TGTAACTGGTATTCCTGTGATCGGTTC
	petA: CTGCGAGATGTTGGGAATTGC
	psbA: AATCTGTAGTTGACAGTCAAGGTCGTG
	ycf2: TGATCCAATAGCGTTCCGTTAGATAG
	ycf1: GTAGTTTTGTTGGTTGGTTAATTGGTC
	ndhF: CCCATGAATCGGACAATACTATGC
Cassytha filiformis	psbA: AATCTGTAGTTGACAGTCAAGGTCGTG
	rps7: AGAGACTCATAGAATGGCAGAGGC
	ycf2: ACGGAACGCTATTGGATCAAATG
	rpl32: AATGCCCGGTAGAAATGGATTTAG
	rps18: GATTAATTACTATTGCTATAAAACAAGCTCG
	rpl20: GGAGACATAGAACAAAAATGCGTTC
	accD: GCTCGATCTATGGTCCGAAACAAC
	rbcL: CTAGCTGCTGCTTGTGAGGTATGG
	psbD: CGTTGCATTGGAGGAGCGA
	trnT: TGGACCTGACCCTTTGAATCATG
	rps16: CCATTTCAAAGAAGGCGGAGG
	trnK: TTTCTTTGTTGGTGGAGTCCTAACG
Cryptocarya chinensis	trnN: TCGTAGTGGCTCCGTTCCATG
	ndhF: CAAAATTGATGGGAATGTAAGAAATCTG
	psbA: AAGTTATGCATGAACGTAATGCTCAC
	ycf2-1: TTGTTGAACCCTGTGAATCACGTG
	rps19: TGGTAGATGCCCGGGACCAAG
	ycf2-2: CACCGAATCGAATTTGGTCCATG

Supplementary Table 2. The 20 basal angiosperm species and their plastome accession numbers in this study. *Newly sequenced species.

Classification	Order	Family	Species	Accession
Magnoliids	Magnoliales	Annonaceae	*Annona cherimola*	NC_030166
		Magnoliaceae	*Liriodendron tulipifera*	NC_008326
			Magnolia kwangsiensis	NC_015892
			Michelia odora	JX280398
	Laurales	Calycanthaceae	*Calycanthus fertilis*	NC_004993
		Lauraceae	*Cassytha filiformis**	LC210517
			*Cinnamomum camphora**	LC228240
			*Neocinnamomum delavayi**	LC213014
			*Cryptocarya chinensis**	LC212965
			Endiandra discolor	KT588615
			Litsea glutinosa	KU382356
			Machilus balansae	NC_028074
			Persea american	KX437771
			Phoebe omeiensis	KX437772
	Canellales	Winteraceae	*Drimys granadensis*	NC_008456
	Piperales	Piperaceae	*Piper cenocladum*	NC_008457
Other basal angiosperm	Amborellales	Amborellaceae	*Amborella trichopoda*	NC_005086
	Nymphaeales	Nymphaeaceae	*Nuphar advena*	NC_008788
	Austrobaileyales	Schisandraceae	*Illicium oligandrum*	NC_009600
	Chloranthales	Chloranthaceae	*Chloranthus spicatus*	NC_009598