BioOne COMPLETE

Incongruence of morphological determinations and DNA barcode sequences: a case study in Cinnamomum (Lauraceae)

Authors: Rohwer, Jens G., Trofimov, Dimitrij, Mayland-Quellhorst, Eike, and Albach, Dirk

Source: Willdenowia, 49(3) : 383-400

Published By: Botanic Garden and Botanical Museum Berlin (BGBM)

URL: https://doi.org/10.3372/wi.49.49309

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Willdenowia
Annals of the Botanic Garden and Botanical Museum Berlin

JENS G. ROHWER[1*], DIMITRIJ TROFIMOV[1], EIKE MAYLAND-QUELLHORST[2] & DIRK ALBACH[2]

Incongruence of morphological determinations and DNA barcode sequences: a case study in *Cinnamomum* (*Lauraceae*)

Version of record first published online on 4 December 2019 ahead of inclusion in December 2019 issue.

Abstract: In the course of molecular systematic studies of *Lauraceae* we received a sample of a plant cultivated under the name *Cinnamomum porrectum* in the Botanical Garden München-Nymphenburg. Preliminary determinations, both morphologically based on the *Flora of China* (Li & al. 2008) and by chloroplast sequences (*psbA-trnH* spacer, *trnK* intron including *matK* gene, *trnL* intron, *trnL-trnF* spacer and *trnQ-rps16* spacer) obtained by Sanger sequencing suggested that it was *C. camphora*, still the plant looked different from other individuals of *C. camphora* cultivated in the botanical gardens of Berlin, Hamburg, Mainz, Munich and Oldenburg. Attempts to sequence the more informative nuclear internal transcribed spacer repeatedly led to mixed signals. We therefore used Illumina sequencing on a set of pre-amplified molecular markers (ITS, *trnK* 3′ and 5′ intron, *trnL* intron, and the intergenic spacers *psbA-trnH*, *trnL-trnF*, as well as parts of the *trnQ-rps16* spacer), and downloaded available sequences of *C. camphora* and *C. parthenoxylon* from GenBank for comparison. Considerable differences were found among these sequences, but the haplotype groups do not coincide with the current species determinations. Particularly the internal transcribed spacer sequences are rather diverse, suggesting possible misidentifications, contaminations, and/or a common gene pool that is larger than anticipated. Concerning the plant in question, our results suggest that it may be a hybrid, with *C. camphora* as the maternal and another species, possibly *C. parthenoxylon*, as the paternal parent.

Key words: *Cinnamomum*, interspecific hybridization, ITS, *Lauraceae*, *matK*, NGS, *psbA-trnH* spacer, *trnK* intron, *trnL* intron, *trnL-trnF* spacer, *trnQ-rps16* spacer

Article history: Received 14 February 2019; peer-review completed 7 April 2019; received in revised form 30 August 2019; accepted for publication 18 September 2019.

Citation: Rohwer J. G., Trofimov D., Mayland-Quellhorst E. & Albach D. 2019: Incongruence of morphological determinations and DNA barcode sequences: a case study in *Cinnamomum* (*Lauraceae*). – Willdenowia 49: 383–400. doi: https://doi.org/10.3372/wi.49.49309

Introduction

In the *Lauraceae*, intraspecific hybrids are well known, especially among different varieties of the avocado, *Persea americana* Mill. (e.g. Furnier & al. 1990). Interspecific hybrids have been described between several species of *P.* subg. *Persea* (*P. americana*, *P. drymifolia* Schltdl. & Cham., *P. floccosa* Mez, *P. nubigena* L. O. Williams, *P. steyermarkii* C. K. Allen), but some of these "species" (*P. drymifolia*, *P. nubigena*) have been treated as varieties of *P. americana* by Kopp (1966), and all of them appeared to be part of *P. americana* in a wider sense in the analysis of Furnier & al. (1990). Apart from that, hybridization frequently has been invoked as an *ad hoc* explanation for occasional morphologically intermediate specimens, e.g. by Kopp (1966) and Rohwer (1993). At least to our knowledge, however, there is no unequivocal evidence of interspecific hybridization in *Lauraceae*, apart from the cases involving *P. americana* and its closest relatives.

1 Universität Hamburg, Institute of Plant Science and Microbiology, Systematics and Evolution of Plants, Ohnhorststr. 18, 22609 Hamburg, Germany; *e-mail: jens.rohwer@uni-hamburg.de (author for correspondence).
2 Carl von Ossietzky Universität Oldenburg, Fakultät V – IBU, Biodiversität und Evolution der Pflanzen, Carl-von-Ossietzky Str. 9-11, 26111 Oldenburg, Germany.

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

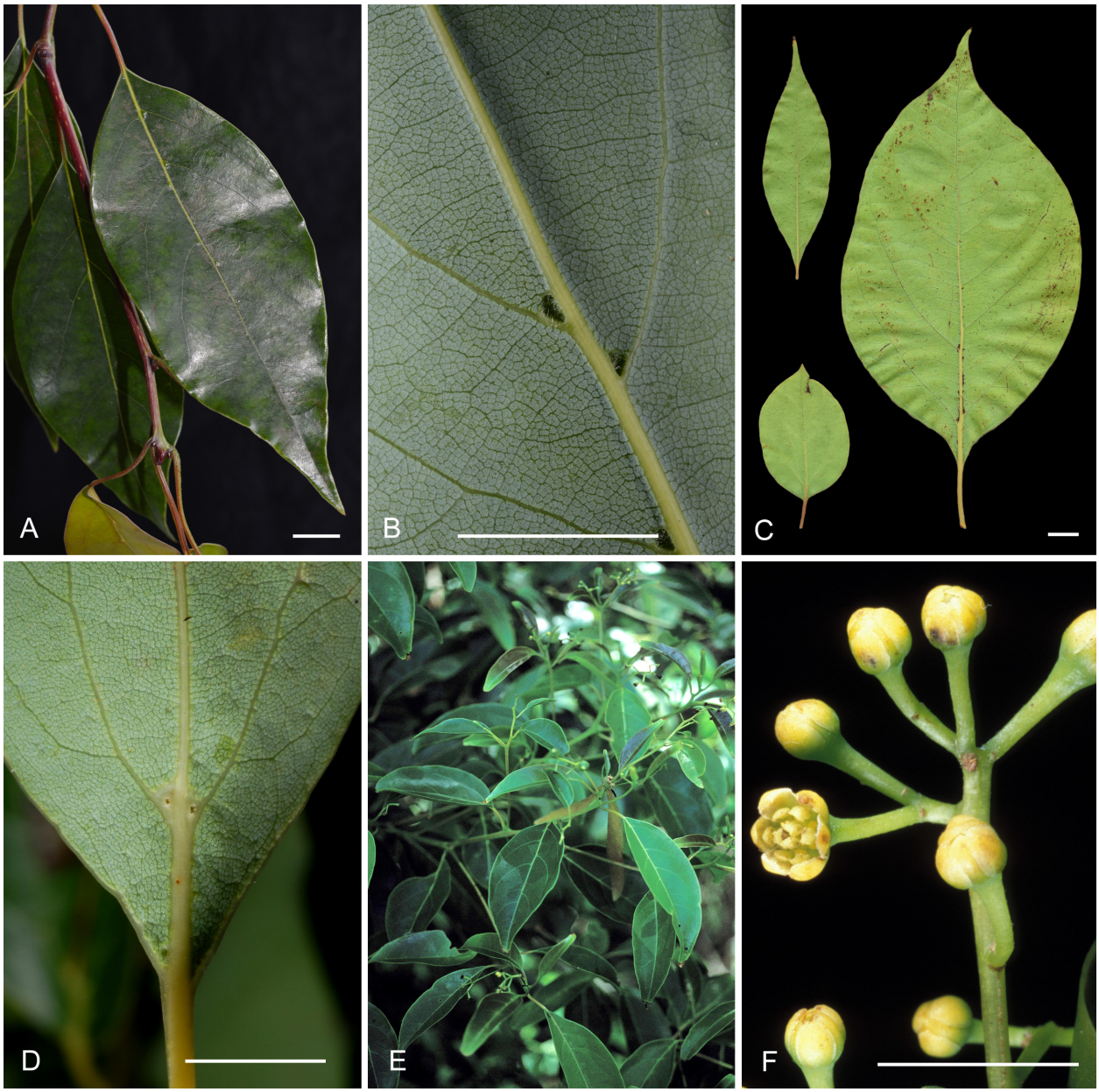


Fig. 1. A, B: *Cinnamomum* "*porrectum*", coll. *Gerlach s.n.*, Bot. Gard. München accession no. *2006/1425* (HBG); A: leafy branch; B: detail of lower leaf surface, with conspicuous domatia in axils of secondary veins. – C, D: *Cinnamomum camphora*, coll. *Rohwer s.n.* (HBG); C: leaves from sterile branches; D: detail of lower leaf surface, with domatia in axils of lowermost secondary veins. – E, F: *Cinnamomum parthenoxylon*, coll. *Rohwer 178* (MJG); E: flowering branch; F: tip of inflorescence. – Scale bars = 1 cm. – Photos by Jens G. Rohwer.

During a molecular systematic investigation on the genus *Cinnamomum* Schaeff. (Rohde & al. 2017), we received a sample of a plant cultivated in the Botanical Garden München-Nymphenburg (Munich), Germany, with the accession number 2006/1425 (Fig. 1A, B). Originally, the plant had been identified as *C. porrectum* (Roxb.) Kosterm., a synonym of *C. parthenoxylon* (Jack) Meisn. We tried to verify this determination based on the vegetative plant material received, as the individual has not yet undergone flowering. Comparison with the key and the descriptions in the *Flora of China* (Li & al. 2008) suggested *C. camphora* (L.) J. Presl or perhaps *C. micranthum* (Hayata) Hayata as the most likely determination. However, many of the leaves of the specimen in question were slightly larger than described for those two species, up to 15 cm long and 7 cm wide, with petioles up to 4 cm, compared to a maximum of 10 × 6 cm described for *C. micranthum* and 12 × 5.5 cm described for *C. camphora*, both with 3 cm petioles. In addition, the leaves of the specimen in question were elliptic to lanceolate-elliptic, usually with an acute to attenuate base, compared to mostly ovate-elliptic leaves with a broadly cuneate to rounded base in *C. micranthum* and *C. camphora*, at least on flowering branches. As a result, the general aspect of the plant in question is somewhat different from the form of *C. camphora* that is

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Table 1. Primers used in this study.

Primer	Direction	Sequence	Author
ITS			
ITS-18	F	5′-GTCCACTGAACCTTATCATTTAGAGG-3′	Käss & Wink (1997); Beyra Matos & Lavin (1999)
ITS-C	R	5′-GCAATTCACACCAAGTATCGC-3′	Blattner (1999)
ITS-D	F	5′-CTCTCGGCAACGGATATCTCG-3′	Blattner (1999)
ITS-H	R	5′-CGGTTCGCTCGCCGTTACTA-3′	Rohwer & al. (2014)
***trnK* intron**			
trnK-3914	F	5′-GGGGTTGCTAACTCAACGG-3′	Steele & Vilgalys (1994)
trnK-B4	R	5′-GTGAACTGGATTTATTGTC-3′	Rohwer & Rudolph (2005)
trnK-PT2	F	5′-TTTTTTCATGAATCCTAACT-3′	Rohwer & Rudolph (2005)
trnK-v510	R	5′-CRATCACTCTTTTGACTTTGG-3′	Rohwer & al. (2014)
trnL-trnF			
trnL-c	F	5′-CGAAATCGGTAGACGCTACG-3′	Taberlet & al. (1991)
trnL-d	R	5′-GGGGATAGAGGGACTTGAAC-3′	Taberlet & al. (1991)
trnL-e	F	5′-GGTTCAAGTCCCTCTATCCC-3′	Taberlet & al. (1991)
trnF-f	R	5′-ATTTGAACTGGTGACACGAG-3′	Taberlet & al. (1991)
trnQ-rps16			
*rps*16-430-F	F	5′-CATAGAATTAGTCATTTTGTCTACC-3′	Rohwer (this study)
trnQ-UUG	R	5′-TCGGAGGTTCGAATCCTTCCGTCCCAGA-3′	Hahn (2002)

commonly cultivated in European botanical gardens. Commonly applied chloroplast DNA barcode markers (*psbA-trnH* intergenic spacer, *trnK* intron including the *matK* gene, *trnL-trnF* region; Hollingsworth & al. 2011; Liu & al. 2017) and the less frequently used *trnQ-rps16* intergenic spacer obtained by Sanger sequencing likewise confirmed *C. camphora*. Attempts to sequence the more informative nuclear ribosomal internal transcribed spacer region (ITS), however, usually resulted in mixed sequences. We therefore applied high-throughput (Illumina) sequencing to a set of pre-amplified molecular markers including ITS, in order to identify the secondary signal in ITS and to investigate if there was a secondary signal in the chloroplast markers as well.

Material and methods

Material originally determined as *Cinnamomum porrectum* has been collected by Günter Gerlach in the greenhouse of the Botanical Garden München-Nymphenburg in November 2010 and on 27 October 2015. For comparison, material of *C. camphora* has been collected by JGR in the Botanical Garden of Hamburg (Fig. 1C, D). Voucher specimens for each plant are preserved in the Herbarium Hamburgense (HBG).

Total genomic DNA was extracted from silica gel dried leaves of both specimens with the "innuPREP Plant DNA Kit" (Analytik Jena, Germany) according the manufacturer's protocol, with modifications as in Rohwer & Rudolph (2005) and Trofimov & al. (2016). The molecular markers and primers used for this study are listed in Table 1. In the amplifications of the ITS region, 10% dimethyl-sulfoxide (DMSO) was added to the reaction mix, as previously described by Rohwer & al. (2009), in order to minimize problems with secondary structures caused by the high GC content of the ribosomal DNA in *Lauraceae*. The PCR products were purified by degradation of single stranded DNA and proteins with FastAP thermosensitive alkaline phosphatase and exonuclease I (Thermo Scientific), both according to manufacturers' instructions.

Reactions for Sanger sequencing and sequence analysis on a 3500 Genetic Analyzer capillary sequencer (Thermo Fisher Scientific, Waltham, USA) were performed as previously described (Rohwer & al. 2009, 2014).

For high-throughput sequencing, PCR products covering the ITS region, the *trnK* 5′ and 3′ intron regions, the *trnL* intron, the *trnL-trnF* spacer and the beginning and end of the *trnQ-rps16* spacer were sent to a sequence provider (LGC Genomics, Berlin, Germany). Multiplexed pools were built with Illumina MiSeq V3 chemistry including a normalizing step and sequenced to 300 bp paired end sequences with the aim of 10 000 sequences per specimen and locus. We received files in a first step demultiplexed to separate individuals, with the barcode indices by LGC Genomics (Berlin). These files were parsed through several steps of a semi-automated script-based pipeline to obtain multiple alignments,

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Table 2. *Cinnamomum* plant material examined and sequences downloaded from GenBank for comparison.

Species	Reference	Voucher	Provenance	ITS	*psbA-trnH*	*trnK* intron	*trnL-trnF*	*trnQ-rps16*
C. camphora	Abbott & al. (unpublished?)	*Davis 0298* (FLAS)	not indicated	—	GU135428	GU135093[1]	—	—
C. camphora	Ando & al. (2013)	isolate OGA0254	Japan: Ogasawara islands	—	—	—	AB817476	—
C. camphora	Chanderbali & al. (2001)	*Chanderbali 322* (MO)	U.S.A.: Missouri Bot. Gard.	AF272260	AF129048[2]	—	AF129020	—
C. camphora	Chao & Lee (unpublished?)	isolate CcN	China: Taiwan	JX242469	—	—	—	—
C. camphora	Chaveerach & al. (unpublished?)	*Chaveerach 738*	Thailand	—	—	JQ435499	—	—
C. camphora	Chen (unpublished?)	not indicated	not indicated	—	—	KF740401[1]	—	—
C. camphora	Chen (unpublished?)	not indicated	not indicated; China?	—	NC035882	NC035882	NC035882	NC035882
C. camphora	Chen & al. (2010)	PS1634MT02	China	—	GQ435461	—	—	—
C. camphora	Chen & al. (2017)	not indicated	China: Jiangsu, Nanjing Univ. campus	—	MF156716	MF156716	MF156716	MF156716
C. camphora	Doh & al. (unpublished?)	WKUCC65	not indicated; Korea?	KX766404	—	—	—	—
C. camphora	Elansary (2013)	*Hosam 00168*	Egypt: Cairo, Orman Garden	—	—	JX495692[1]	—	—
C. camphora	Erickson & al. (2014)	tree tag 032628 (TNM S146239)	China: Taiwan	—	KJ686728	—	—	—
C. camphora	Huang & al. (2016)	*Ci X. Q. 0265*	China: Hunan	KU139826	KU160277	—	—	—
C. camphora	Kohjyouma & al. (unpublished?)	isolate 013	not indicated	—	—	—	AB040090	—
C. camphora	Kohjyouma & al. (unpublished?)	isolate 014	not indicated	—	—	—	AB040091	—
C. camphora	Kojoma & al. (2002)	isolate 013	not indicated	—	—	—	AB040080[3]	—
C. camphora	Kojoma & al. (2002)	isolate 014	not indicated	—	—	—	AB040081[3]	—
C. camphora	Kong & al. 2014)	isolate Z	China: Guangdong	—	—	KJ510888[1]	—	—
C. camphora	Kress & Erickson (2007)	USDA TARS4956	not indicated	—	—	EF590397[1]	—	—
C. camphora	Li & al. (2016)	*Li & Huang H-DLS08* (HITBC)	China	KT248576	—	—	—	—
C. camphora	Liu & al. (2012)	PS5019MT01	China: Hubei	—	HM019386	HM019316[1]	—	—
C. camphora	Liu & al. (2012)	PS5020MT01	China: Hubei	—	HM019387	HM019317[1]	—	—
C. camphora	Liu & al. (2015)	SCBGP173_1 (IBSC)	China: Guangdong	KP092856	KP095535	KP093545[1]	—	—
C. camphora	Liu & al. (2015)	SCBGP173_2 (IBSC)	China: Guangdong	KP092857	KP095536	KP093546[1]	—	—
C. camphora	Liu & al. (2017a)	KPS0050A03	Korea: commercial material	MF096121[4]	MF096903	—	—	—

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

C. camphora	Liu & al. (2017a)	KPS0050A02	Korea: commercial material	MF096120[4]	MF096904	—	—	—
C. camphora	Liu & al. (2017a)	KPS0050A01	Korea: commercial material	MF096119[4]	MF096905	—	—	—
C. camphora	Liu & al. (2017a)	KMS0050A03	Korea: commercial material	—	MF096906	—	—	—
C. camphora	Liu & al. (2017a)	KMS0050A02	Korea: commercial material	—	MF096907	—	—	—
C. camphora	Liu & al. (2017a)	KMS0050A01	Korea: commercial material	—	MF096908	—	—	—
C. camphora	Liu & al. (2017b)	*Ci X. Q. 0224*	China: Hunan	—	KX546100	KX546013[1,3]	—	—
C. camphora	Liu & al. (2017b)	*Ci X. Q. 0384*	China: Zhejiang	KX546414[3], KX546586[4]	KX546101	KX546024[1,3]	—	—
C. camphora	Liu & al. (2017b)	*Ci X. Q. 0405*	China: Zhejiang	—	KX546102	KX546030[1,3]	—	—
C. camphora	Liu & al. (2017b)	*Li J. 2006004*	China: Guangdong	KX546537[3], KX546756[4]	KX546103	KX545833[1,3]	—	—
C. camphora	Madriñán & Chacon (unpublished?)	K6469 (ANDES)	not indicated	—	EU153948	EU153829[1]	—	—
C. camphora	Massoni & al. (2013)	*Menton s.n.* (P44475)	France: Paris, Jardin Botanique Exotique	—	—	—	KF586659, KF586691	—
C. camphora	Parveen & al. (unpublished?)	SBB-1126	not indicated; India?	JN115020	—	JN114745[1]	—	—
C. camphora	Pei & al. (2011)	tree tag 0904158	China: Guangdong	—	HQ415574	HQ415392[1]	—	—
C. camphora	Pei & al. (2011); Pei & al. (unpublished?)	isolate 176	not indicated; China?	—	HQ427102[3]	HQ427401[1,3]	—	—
C. camphora	Qiu & al. (2005)	*Qiu 102*	U.K.: Royal Bot. Gard., Kew	—	—	AJ966800[1]	—	—
C. cf. camphora[5]	Rohde & al. (2017); this study	*Gerlach s.n.* (HBG)	Germany: Bot. Gard. Munich 2006/1425	MF110039 (primary); MN480757 (secondary)	MF137959	MN482110	MN482111	MN482113
C. camphora	Rohde & al. (2017)	*Leuenberger s.n.* (HBG)	Germany: Bot. Gard. Berlin	MF110040	MF137960	—	—	—
C. camphora	Rohwer & Rudolph (2005)	*Rohwer 164* (MJG)	Germany: Bot. Gard. Mainz	—	—	AJ247154[6]	—	—
C. camphora	Sukrong (unpublished?)	CHULA-044	not indicated; Thailand?	—	LC435397	—	—	—
C. camphora	Swetha & Sasikumar (unpublished?)	isolate efgh2	India	—	MF072391[2]	MF589651[1]	—	—
C. camphora	Swetha & Sasikumar (unpublished?)	isolate ijkl3	India	—	MF072392[2]	MF589649[1]	—	—

(continued on next page)

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Species	Reference	Voucher	Provenance	ITS	*psbA-trnH*	*trnK* intron	*trnL-trnF*	*trnQ-rps16*
C. camphora	Trofimov & al. (2016); Rohde & al. (2017); this study	*Rohwer s.n.* (HBG)	Germany: Bot. Gard. Hamburg A0-F-10964	KX509822	KX509882	—	—	MN482114
C. camphora	Wang & al. (unpublished?)	not indicated	not indicated; China?	—	—	—	KM056312	—
C. camphora	Wu (unpublished?)	strain JLH2	not indicated; Taiwan?	—	MF421523	MF421523	MF421523	MF421523
C. camphora	Wu (unpublished?)	not indicated	not indicated; Taiwan?	—	MG021326	MG021326	MG021326	MG021326
C. camphora	Wu & Ho (unpublished?)	clone P2	not indicated; Taiwan?	KP218517, KP218518	—	—	—	—
C. camphora	Wu & al. (2017)	*Chaw 1510*	Taiwan: Taipei Bot. Gard.	—	LC228240	LC228240	LC228240	LC228240
C. camphora	Wu & al. (2019)	CCAMP20180218 (SANU)	China: Xi'an Bot. Gard.	—	MH050970	MH050970	MH050970	MH050970
C. parthenoxylon	Chaveerach & al. (unpublished?)	*Chaveerach 737*	Thailand	—	—	JQ435498	—	—
C. parthenoxylon	Chen & al. (2010)	PS1626MT01	China	—	GQ435459	GQ434288[1]	—	—
C. parthenoxylon	Huang & al. (2016)	*Huang J. F. QSBG04* (HITBC)	Thailand: Chiang Mai	KU139870	—	—	—	—
C. parthenoxylon	Huang & al. (2016)	*Huang J. F. 201312* (HITBC)	Indonesia: Java Barat	KU139871	KU160285	—	—	—
C. parthenoxylon	Huang & al. (2016); Liu & al. (2017b)	*Ci X. Q. 09001* (HITBC)	China: Yunnan	KU139869[4], KX546594	KX546121	KX546068[1, 3]	—	—
C. parthenoxylon	Kong & al. (2014)	isolate YZ	China: Hainan	—	—	KJ510895[1]	—	—
C. parthenoxylon	Liu & al. (2015)	SCBGP017_1 (IBSC)	China: Guangdong	KP092858	KP095537	KP093276[1]	—	—
C. parthenoxylon	Liu & al. (2015)	SCBGP017_2 (IBSC)	China: Guangdong	KP096040	KP095538	KP093277[1]	—	—
C. parthenoxylon	Liu & al. (2017b)	*Ci X. Q. 0063* (HITBC)	China: Guangxi	KX546421[3], KX546593[4]	KX546119	KX545924[1]	—	—
C. parthenoxylon	Liu & al. (2017b)	*Ci X. Q. 0255* (HITBC)	China: Yunnan	—	KX546120	KX545925[1]	—	—
C. parthenoxylon	Rohde & al. (2017); this study	*Rohwer 178* (MJG)	Malaysia: Sabah	MF110054	MF137974	—	MN482112	—
C. parthenoxylon	Wu & al. (2019)	CPART20171128 (SANU)	China: Lushan Bot. Gard.	—	MH050971	MH050971	MH050971	MH050971

[1] Not overlapping with examined fragments; downloaded for comparison with AJ247154.
[2] Very short fragment; not included in haplotype analysis.
[3] Accession number not listed / different from those listed in respective publication.
[4] Short fragment, therefore not included in Table 3 (supplemental content online).
[5] Questionable plant examined here; primary sequence submitted to GenBank as *Cinnamomum* cf. *camphora*.
[6] All sequenced fragments of *Rohwer s.n.* (HBG) identical with this sequence, therefore not submitted.

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

which were then inspected manually. The first part of the pipeline deals with the raw sequence data and performs a quality control, merges paired end reads and in a second step demultiplexes the locus data for every individual. For the first step of part one, we used bbduk.sh (Bushnell & al. 2017) to quality filter and trim the reads for adapters (Trueseq-PE). Parts with a quality lower than phred 10 were removed (settings: qtrim=r trimq=10 ktrim=r k=25 mink=11 hdist=1 tbo tpe). The merging of the paired end reads was done by bbmerge.sh (Bushnell & al. 2017; settings: efilter=6 pfilter=0.00002). The third step of manipulating the raw sequence data was demultiplexing the locus data in the merged and unmerged fastq files with seal.sh (Bushnell & al. 2017), using a list of primer sequences. The second part of the pipeline sorts and streamlines the data with CAP3 (Huang & Madan 1999) to produce overlapping assemblies and multiple sequence alignments of CAP3 consensus sequences with MAFFT (Katoh & Standley 2013). As the cap3 software is not capable of multithreading we used the free GNU parallel software (Tange 2011) to compute many CAP3 assemblies side by side on a multi-core system (cap3 settings: -p 99 -g 100 -t 500 -f 2). Consensus sequences of 90% similarity were extracted with a python script using the Ace.parser module (settings: threshold=0.9, ambiguous="N", require_multiple=10) contained in Biopython (Cock & al., 2009), aligned with MAFFT (settings: -ep 0 -genafpair -maxiterate 1000 – adjustdirectionaccurately) and inspected visually. Sequence variants with a frequency of less than 5% were omitted as likely PCR or sequencing errors.

For comparison, sequences of *Cinnamomum camphora* and *C. parthenoxylon* covering at least a substantial part of the investigated genome regions were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/genbank/, accessed in March 2019), if available. Among them were seven complete chloroplast genome sequences (six of *C. camphora*, one of *C. parthenoxylon*) from which we extracted the respective chloroplast regions. These sequences were used as reference sequences for the chloroplast data. Accession numbers of all sequences are listed in Table 2. Sequences were aligned in Sequencher 4.8 (Gene Codes Corporation) or in MEGA 6 (Tamura & al. 2013). If a sequence deviated from most others in the first or the last 30 base pairs, these were deleted, in order to avoid possible primer or sequencing artefacts. Two micro-inversions in the *psbA-trnH* spacer sequences, of eight and five base pairs, respectively, were reversed and complemented in the alignments.

If more than two haplotypes were found per molecular marker, haplotype networks were constructed using the Integer Neighbor Joining algorithm (IntNJ) implemented in POPART (Leigh & Bryant 2015). Because the algorithm cannot deal with too many undefined character states, either shorter sequences or most non-overlapping alignment positions (or both) had to be excluded from the haplotype analyses. Their results therefore represent less than the entire range of variation among the sequences.

In order to detect possible misidentifications or contaminations, sequences diverging from the most frequent haplotypes were subjected to a BLAST search (Basic Local Alignment Search Tool, https://blast.ncbi.nlm.nih.gov/Blast.cgi). In addition, we inserted the ITS sequences into a subset of the ITS data matrix of Rohde & al. (2017) and performed maximum likelihood analyses of the entire ITS region and the separate ITS-1 and ITS-2 regions in MEGA 6, with taxa of the *Persea* group (*Machilus grijsii* Hance, *Phoebe sheareri* (Hemsl.) Gamble and *Persea americana*) as outgroup. Branch support was estimated by bootstrapping with 500 replicates.

Results

An examination of the leaf spectrum of the largest *Cinnamomum camphora* tree cultivated in the greenhouse of the Botanical Garden of Hamburg (Fig. 1C, D) revealed that the range of variation in leaf shape and size is much larger even in a single individual than indicated in the *Flora of China* (Li & al. 2008). On non-flowering branches we found leaves up to 19.5 cm long and 10 cm wide, with petioles up to 5 cm long. The length:width ratio of the leaf blades on this tree ranged from 1.5:1 to 3.5:1, and the leaf bases varied from rounded via obtuse, acute and cuneate to narrowly attenuate.

The complete chloroplast genome sequences can be sorted into two groups of highly similar sequences. Chloroplast (cp) group #1 consists of the sequences MF156716, MG021326 and NC035882 (all *Cinnamomum camphora*), cp group #2 of LC228240, MF421523 and MH050970 (*C. camphora*), plus MH050871 (*C. parthenoxylon*). The two groups differ considerably, by 10 substitutions in the *psbA-trnH* spacer, 13 in the *trnK* intron, one in the *trnL* intron, one in the *trnL-trnF* spacer, and nine substitutions plus two insertion/deletion events (indels) in the *trnQ-rps16* spacer. Within the groups there are only minor differences. In cp group #1, the *matK* sequences of MF156716 and NC035882 have a duplication of a single base (A) immediately before the stop codon, making the *matK* transcript 21 bases longer. In cp group #2, the *matK* sequence of LC228240 differs by a single substitution from the other members of the group.

Most of the chloroplast DNA sequences obtained from the plant in question (the suspected hybrid) by direct Sanger sequencing (*psbA-trnH* spacer, *trnK* intron, *trnL-trnF* spacer, *trnQ-rps16* spacer) turned out to be identical with the sequences of cp group #1. Only the *trnL* intron sequence agrees with cp group #2, but in this case it is the only difference between the two groups. The *trnQ-rps16* sequence likewise differs by a single substitution from cp group #1, but cp group #2 differs much more in this spacer. However, there are sometimes additional differences among the shorter sequences downloaded from GenBank.

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

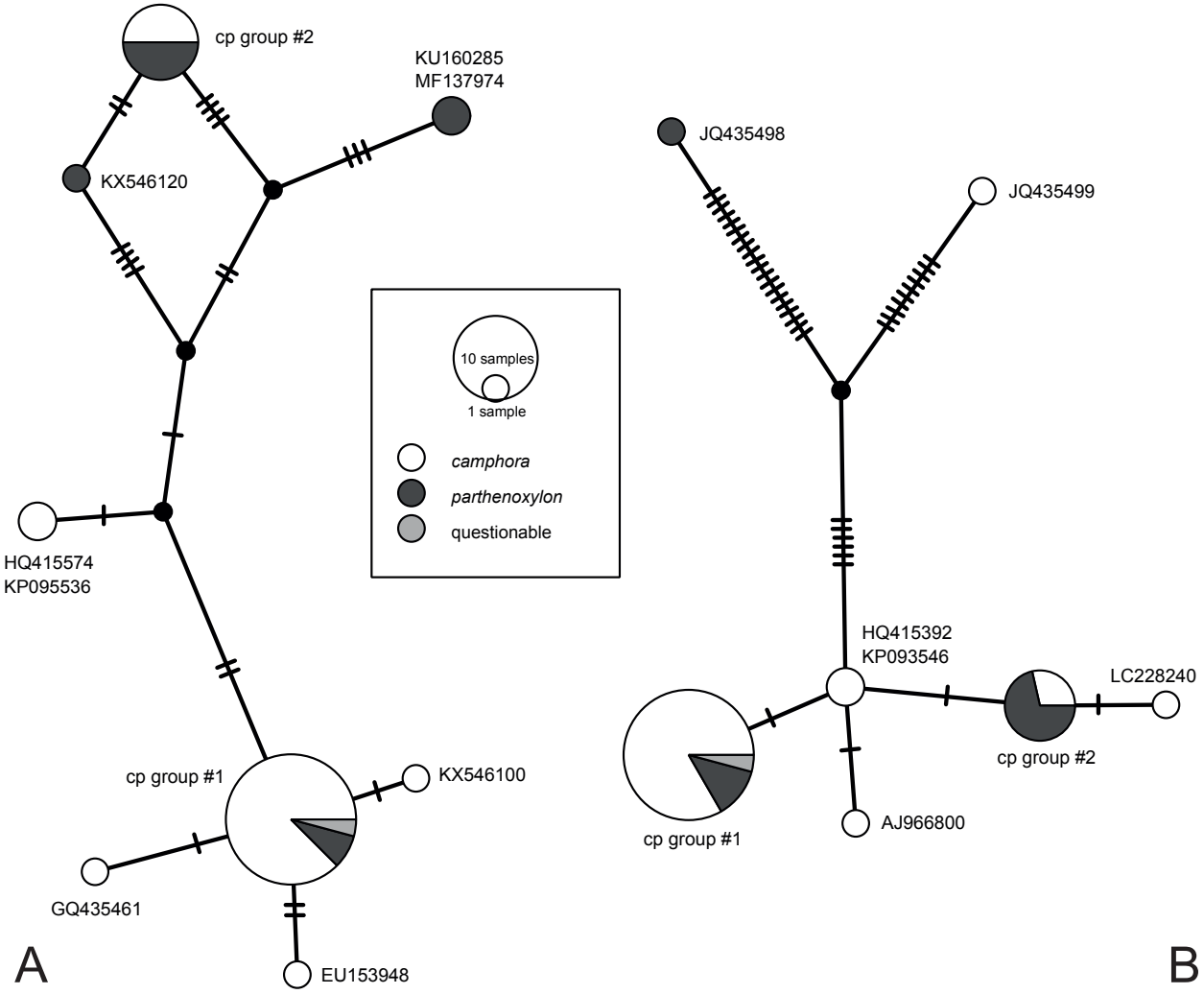


Fig. 2. Haplotype networks for the two chloroplast markers that show more than two haplotypes. – A: *psbA-trnH* intergenic spacer; B: *matK* gene, partial sequence. cp group #1 includes the total chloroplast sequences MF156716, MG021326 and NC035882, plus (A) the *psbA-trnH* sequences GQ435459, GU135428, HM019386, HM019387, HQ427102, KP095535, KU160277, KX509882, KX546101, KX546102, KX546103, KX546119, LC435397, MF096903, MF096904, MF096905, MF096906, MF096907, MF096908, MF137959 and MF137960, or (B) the *matK* sequences AJ247154, EF590397, EU153829, GQ434288, GU135093, HM019316, HM019317, HQ427401, JN114745, JX495692, KF740401, KJ510888, KJ510895, KP093545, KX545833, KX545924, KX546013, KX546024, KX546030, MF589649, MF589651, MN482110. cp group #2 includes the total chloroplast sequences MF421523, MH050970 and MH050871, plus (A) the *psbA-trnH* sequences KJ686728, KP095537, KP095538 and KX546121, as well as the *psbA-trnH* part of the total chloroplast sequence LC228240, or (B) the *matK* sequences KP093276, KP093277, KX545925 and KX546068.

Most of the downloaded *psbA-trnH* sequences of *Cinnamomum camphora* (GU135428, HM019386, HM019387, HQ427102, KP095535, KU160277, KX509882, KX546101, KX546102, KX546103, LC435397, MF072391, MF072392, MF096903, MF096904, MF096905, MF096906, MF096907, MF096908, MF137960) agree with cp group #1, disregarding repeat length differences in a poly-T single nucleotide repeat. Sequence KJ686728, however, agrees with cp group #2. Four additional haplotypes are represented by one or two sequences each (Fig. 2A). Sequences HQ415574 and KP095536 differ by three substitutions from cp group #1, of which two are shared with cp group #2. According to a BLAST search, they are 100% identical (in 100% query cover) with *C. chekiangense* Nakai (MF137961) and several sequences of *C. burmannii* (Nees & T. Nees) Blume (MF137956, MF137957, MF137958). However, many other *Cinnamomum* species are not much different, either. GQ435461 and KX546100 differ from cp group #1 by a single (different) substitution each. GQ435461 shows 99.24% similarity not only to cp group #1 sequences of *C. camphora*, but also to sequence GQ435459 of *C. parthenoxylon* and sequence HM019382 of *C. bodinieri* H. Lév. Sequence KX546100 shows 99.74% similarity to many *C. camphora* sequences, but also to sequences KX546116 and KX546117 of *C. longepaniculatum* (Gamble) N. Chao ex H. W. Li, KX546110 of *C. glanduliferum* (Wall.) Meisn., KX546119 of *C. parthenoxylon* and HM019382 of *C. bodinieri*. Sequence

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

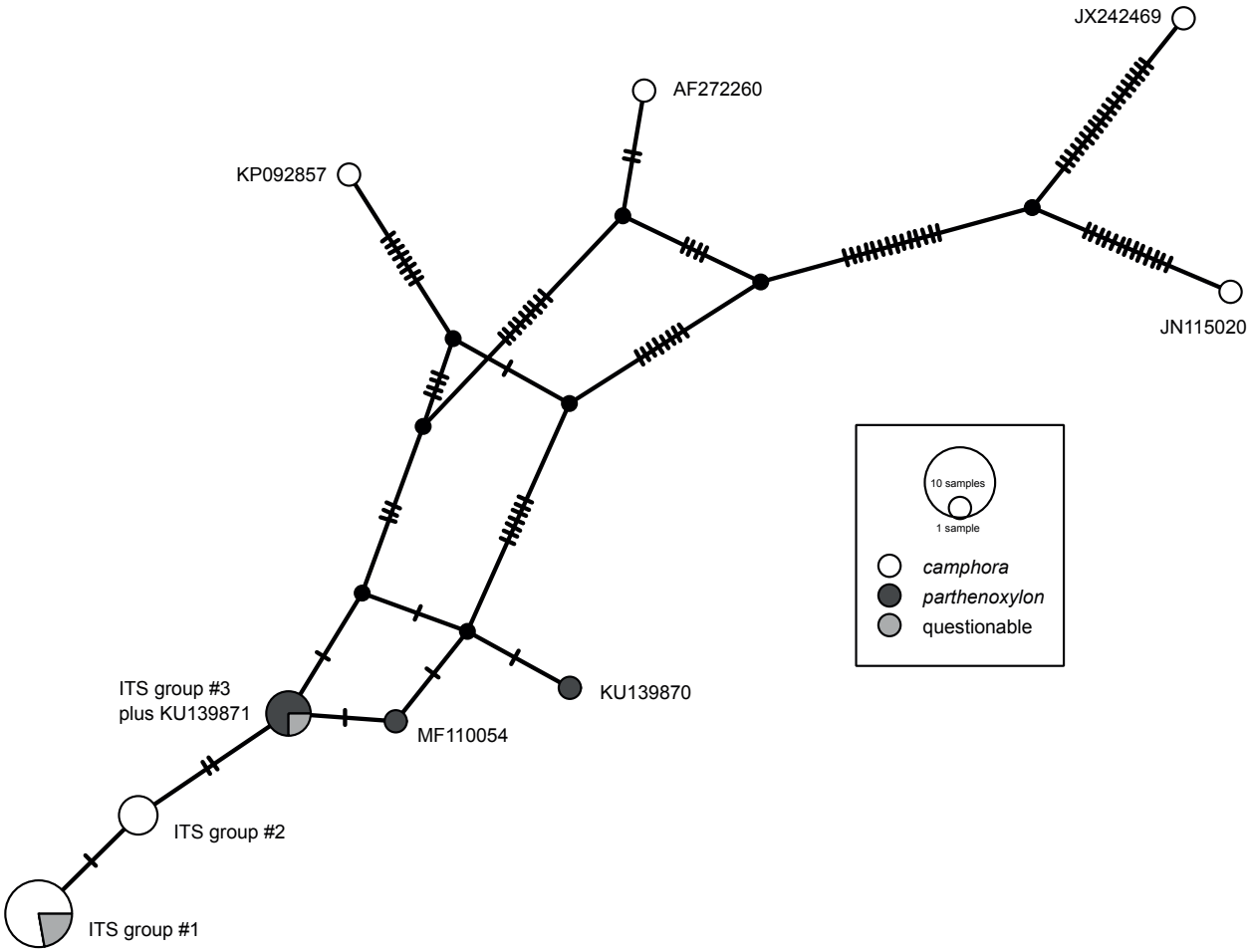


Fig. 3. Haplotype network for sequences covering the major part of the ITS region. ITS group #1 includes sequences KP092856, KT248576, KU139826, KX509822, KX546414, KX546537, MF110039 and MF110040. ITS group #2 includes sequences KP218517, KP218518 and KX776404. ITS group #3 includes sequences KP092858, KU139871 (differing by one substitution in a part of the ITS-2 region not included here), KX546421 and MN480757.

EU153948 differs by two substitutions from cp group #1, in positions 51 and 54 from its beginning, and agrees > 99 % with other *C. camphora* sequences only. Among the sequences of *C. parthenoxylon*, GQ435459 and KX546119 agree with cp group #1 and differ from cp group #2 by nine substitutions, four of them shared with KU160285 and MF137974 (see below), and two shared with KX546120. The sequences KP095537, KP095538 and KX546121 agree with cp group #2, except for differences in single nucleotide repeat lengths. In addition, there are two more haplotypes. Sequence KX546120 differs from cp group #2 by two substitutions, shared with cp group #1, in a region affected by a micro-inversion in KU160285 and MF137974. According to a BLAST search, it agrees 99.48 % with sequences of samples identified as *C. camphora* (MH050970), *C. parthenoxylon* (MH050971), *C. bodinieri* (MF137955), *C. glanduliferum* (MF137965), *C. longepaniculatum* (KX546118), *C. platyphyllum* (Diels) C. K. Allen (HM0193396), or even *Litsea ichangensis* Gamble (HM019414). The haplotype represented by KU160285 and MF137974 differs from cp group #2 by nine substitutions, four of which are shared with cp group #1. In addition, it shows two micro-inversions, of eight and five base pairs, respectively. The most similar sequence in GenBank is MF137973 (*C. paiei* Kosterm., 98.43 % identity in 100 % query cover).

Among the *trnK* intron sequences (Fig. 2B), there is only one conventional sequence (AJ247154) covering the entire *trnK* intron, identical with cp group #1. Another 29 sequences downloaded from GenBank cover only a part of the *matK* gene. Among them, 17 sequences submitted as *Cinnamomum camphora* (EF590397, EU153829, GU135093, HM019316, HM019317, HQ427401, JN114745, JX495692, KF740401, KJ510888, KP093545, KX545833, KX546013, KX546024, KX546030, MF589649, MF589651) and three sequences submitted as *C. parthenoxylon* (GQ434288, KJ510895, KX545924) are likewise identical with cp group #1. Four sequences submitted as *C. parthenoxylon* (KP093276, KP093277, KX545925, KX546068) are identical with cp group #2 and differ from cp group #1 by two substitutions. One of these substitutions is also shared with three sequences submitted as *C. camphora* (AJ966800,

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

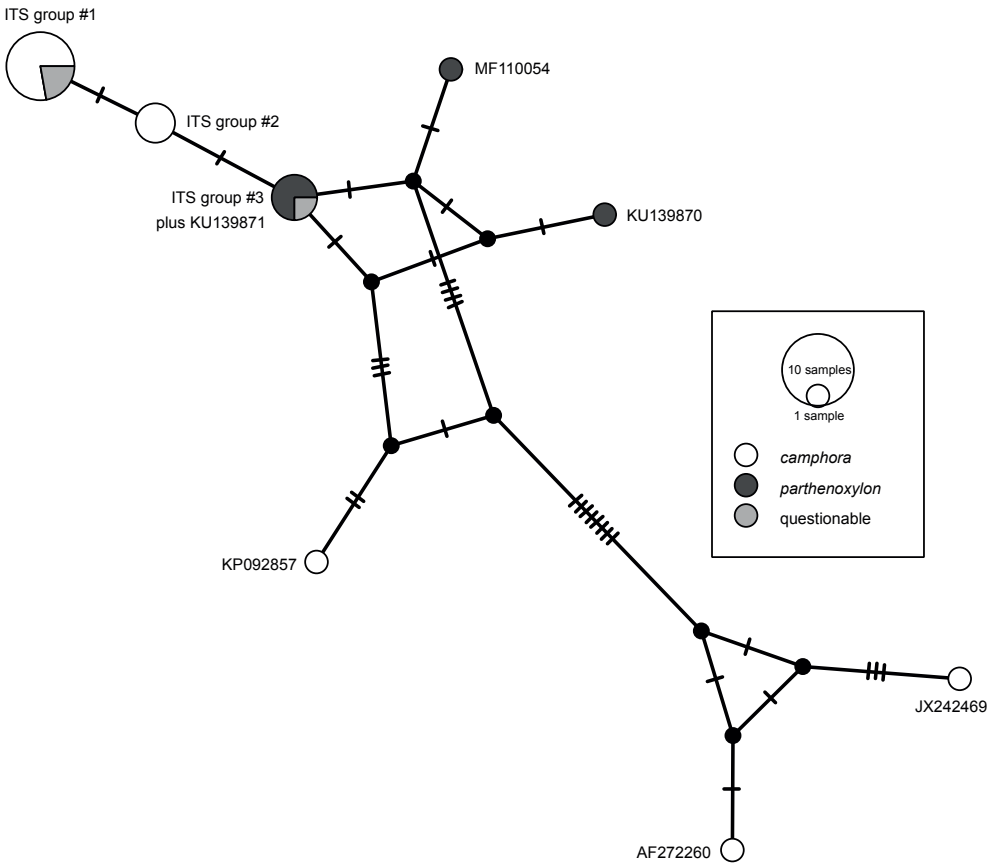


Fig. 4. Haplotype network for sequences covering the major part of the ITS-1 region. ITS group #1 includes sequences KP092856, KT248576, KU139826, KX509822, KX546414, KX546537, MF110039 and MF110040. ITS group #2 includes sequences KP218517, KP218518 and KX776404. ITS group #3 includes sequences KP092858, KU139871 (differing by one substitution in ITS-2), KX546421 and MN480757.

HQ415392, KP093546), which otherwise agree with cp group #1, except for one additional substitution in AJ966800. The sequences JQ435499 ("*C. camphora*") and JQ435498 ("*C. parthenoxylon*") differ considerably from all other sequences. JQ435499 differs by 13 substitutions from cp group #1 and by 15 substitutions from cp group #2. JQ435498 differs by 24 substitutions from cp group #1, by 25 substitutions from cp group #2, and by an insertion and a deletion of a single nucleotide from all other sequences. Nevertheless, according to a BLAST search these deviating sequences agree 97.53% and 95.89% with numerous other *Lauraceae*, mainly *Cinnamomeae*, including taxa of the Neotropical *Ocotea* complex. The indels in JQ435498 are 15 positions apart in the coding region of the *matK* gene, so that the reading frame is relatively quickly restored.

In the *trnL-trnF* region, our new sequence of *Cinnamomum parthenoxylon* (MN482112) is identical with that of the plant in question (MN482111), i.e. it agrees with cp group #2 in the *trnL* intron and with cp group #1 in the *trnL-trnF* spacer. The four short *trnL* intron sequences (AB040090, AB040091, AB817476, KF586691) and the five *trnL-trnF* spacer sequences (AB040080, AB040081, AF129020, KF586659, KM056312) of *C. camphora* downloaded from GenBank are identical with cp group #1. Apart from the complete chloroplast genomes, there are no sequences of the *trnL-trnF* region of *C. parthenoxylon* in GenBank, and no previously submitted sequences of the *trnQ-rps16* intergenic spacer of any *Cinnamomum* species.

Sanger sequencing of the internal transcribed spacer (ITS) region was less straightforward in the plant in question. In most attempts based on the first collection (from 2010) we obtained a mixed signal, in which only a part of the sequence was clearly readable, while most of it contained overlapping, non-identical chromatogram signals. We therefore requested new material from Munich in 2015, but the results were mostly the same. In two of six attempts, however, the signal was relatively clear, with a much lower secondary signal and only a few ambiguous peaks in the chromatogram, of which one remained in the consensus sequence (MF110039).

The ITS sequences submitted to GenBank as *Cinnamomum camphora* or as *C. parthenoxylon* are much more diverse than the chloroplast sequences (Fig. 3–5). The pairwise distances (p-distances) among the reasonably complete sequences (> 500 base pairs; 14 sequences of *C. camphora*, five of *C. parthenoxylon*) are summarized in Table 3 (supplemental content online). The sequences KP218517 and KP218518 are from the

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

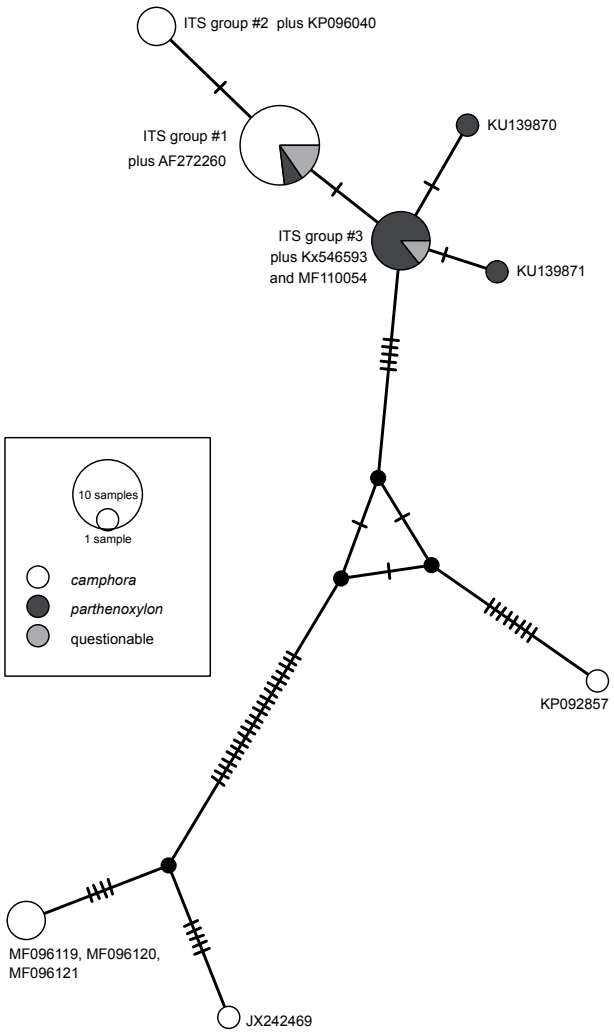


Fig. 5. Haplotype network for sequences covering the major part of the ITS-2 region. ITS group #1 includes sequences AF272260 (strongly different in ITS-1), KP092856, KT248576, KP096040 (ITS-2 only), KU139826, KX509822, KX546414, KX546537, KX546586 (ITS-2 only), MF110039 and MF110040. ITS group #2 includes sequences KP218517, KP218518 and KX776404. ITS group #3 includes sequences KP092858, KU139869 (ITS-2 only), KX546421, KX546593 (ITS-2 only), KX546594 (ITS-2 only), MF110054 (differing by one substitution in ITS-1) and MN480757.

same voucher, therefore only one of them is included in Table 3. Completely identical (in their overlapping parts) are three pairs or groups of sequences: KX509822 and MF110040 (in the following called ITS group #1), KP218517, KP218518 and KX766404 (ITS group #2), as well as KP092858 and KX546421 (ITS group #3). ITS group #3 differs from both ITS group #1 and ITS group #2 by 11 substitutions, whereas ITS groups #1 and #2 differ by six substitutions. Figures 3–5 show far fewer substitutions between these groups because many of substitutions are in positions either not covered, or affected by an indel, or with an ambiguity in at least one of the additional sequences, so that they are not recognized by POPART. Sequences KX546537 and MF110039 are similar to ITS group #1 but include one ambiguous position (the same in both of them), KU139826 two ambiguous positions (one of them shared with KX546537 and MF110039), which causes p-distances of zero to sequences that differ only in these positions. In addition to the identical sequences KX509822 and MF110040 (ITS group #1), and the three sequences with ambiguous positions, the group of sequences differing by not more than one substitution in the ITS-1 and/or the ITS-2 region includes also the sequences KP092856, KT248576 and KX546414. All these sequences are shown as members of ITS group #1 in Fig. 3–5.

While most ITS sequences show not more than 13 pairwise differences, and no differences at all in the 5.8S region, four of them are highly divergent. The most divergent sequence, JX242469, differs from the other sequences by an average of 24.4 substitutions in the ITS-1 region, 14.7 substitutions in the 5.8S region, and 39.9 substitutions in the ITS-2 region. Sequence JN115020 differs from the others by an average of 15.2 substitutions in the ITS-1 region, 10.9 substitutions in the 5.8S region, and 27.7 substitutions in the ITS-2 region, KP092857 differs by an average of 12.0 substitutions in the ITS-1 region, 2.4 in the 5.8S region, and 17.6 in the ITS-2 region, and AF272260 by 20.4 in the ITS-1 region, 2.3 in the 5.8S region, and 6.3 in the ITS-2 region. In the ITS-2 region, however, it is identical with the ITS group #1 sequences. In the 5.8S region, both AF272260 and KP092857 differ from most sequences by only a single mutation; the much higher averages are due to the rather aberrant sequences JN115020 and JX242469. According to a BLAST search, the ITS-1 part of JX242469 agrees 99.54 % with JX242468, submitted as *Cinnamomum burmannii*, whereas the ITS-2 part agrees 99.38 % with *C. micranthum* f. *kanehirae* (Hayata) S. S. Ying (JX242470). All these sequences have been submitted by the same authors. The most similar *C. camphora* sequences are AF272260 (ITS-1, 94.42 %) and the rather short (240 bases) sequences MF096119–MF096124 (ITS-2, 96.25 %). Sequence KP092857 agrees best (99.83 %) with several sequences of *C. burmannii* (FM957802, KP092854, KP092855, KX766400, MF110036, MF110037, MF110038). It is identical with most of them in the ITS-2 region. There is no other *C. camphora* sequence among the top 100 BLAST hits, neither for ITS-1, nor for ITS-2, nor for the entire ITS-region. The ITS-1 part of AF272260 agrees 94.42 % with JX242469 (submitted as *C. camphora*, see above) and 93.95 % with JX242468 (submitted as *C. burmannii*, see above). The only other hit for *C. camphora* among the top 100 BLAST results is JN115020 (96.03 %), but with only 56 % query cover.

The result of the Illumina sequencing in the plant in question was the same for the chloroplast markers investigated, i.e. the by far most frequent reads for all of them were identical with the result of the Sanger sequencing (*Cinnamomum camphora* cp group #1, but *trnL* intron sequence like cp group #2).

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

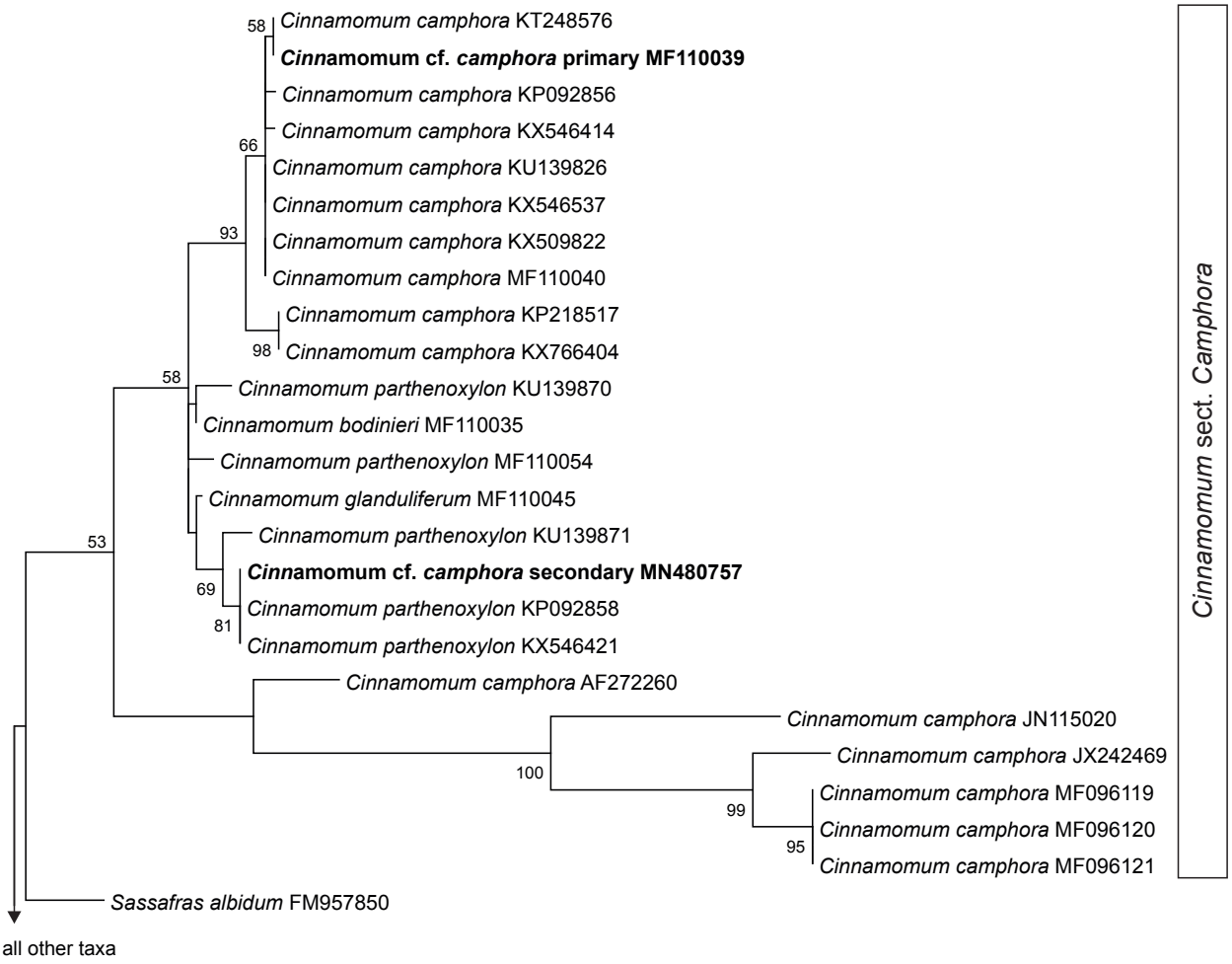


Fig. 6. Result of a maximum likelihood analysis of the ITS data, part 1. Numbers next to branches are bootstrap values ≥ 50%. Sequences of the plant in question in **boldface**.

In the ITS sequences, we found different copies for both the ITS-1 and the ITS-2 region. The most frequent copies, with 443 reads (73% of the total reads) for the ITS-1 and 3390 reads (52%) for the ITS-2, were likewise identical with the sequences obtained by Sanger sequencing (MF137959, *Cinnamomum camphora*, ITS group #1), except for a few ambiguities, all but one near the ends (cut off for the analysis in Table 3, supplemental content online). For the ITS-1 region, we found a single alternative sequence, with much lower coverage (162 reads, 27%). It shows five non-ambiguous base pairs difference compared to the most frequent copy, but agrees with the ITS-1 part of two of the five *C. parthenoxylon* sequences in GenBank (KP092858 and KX546421 = ITS group #3). The second most frequent copy for the ITS-2 region, with 2186 reads (34%), differs from the most frequent copy by four substitutions, one ambiguity, and an indel of ten base pairs, and agrees with the ITS-2 part of the same two *C. parthenoxylon* sequences (KP092858 and KX546421). A third and a fourth variant retrieved for the ITS-2 region, with 617 (9%) and 315 (5%) reads, respectively, agree with one of the two most frequent copies in their first half, and with the other in the second. For the phylogenetic analysis, we combined the second most frequent copies of ITS-1 and ITS-2 into a single sequence (MN480757).

The modeltest implemented in MEGA suggested a Tamura 3-parameter model with discrete Gamma distribution (T92+G) for the ITS data set. The result for the entire ITS region (Fig. 6) shows the *Laureae* (*Laurus nobilis* L., *Lindera benzoin* (L.) Blume and *Neolitsea sericea* (Blume) Koidz.) as sister to the *Cinnamomeae*, i.e. to all other taxa except the outgroup. Among the *Cinnamomeae*, there is a trichotomy consisting of (1) *Cinnamomum* sect. *Cinnamomum*, (2) a clade consisting of *Aiouea* Aubl. as sister group of the *Ocotea* complex, and (3) *Sassafras* J. Presl as sister taxon to *C.* sect. *Camphora*. However, the sequence KP092857, submitted to GenBank as *C. camphora*, is nested among the species of *C.* sect. *Cinnamomum*. The relationships among the species of *Cinnamomum* and within the *Ocotea* complex have been described in detail in previous papers (Huang & al. 2016, Rohde & al. 2017, Trofimov & al. 2019). All sequences added to the dataset here, except KP092857 (see above), form a common clade with the other sequences of species of *C.* sect. *Camphora*. The sequences AF272260,

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

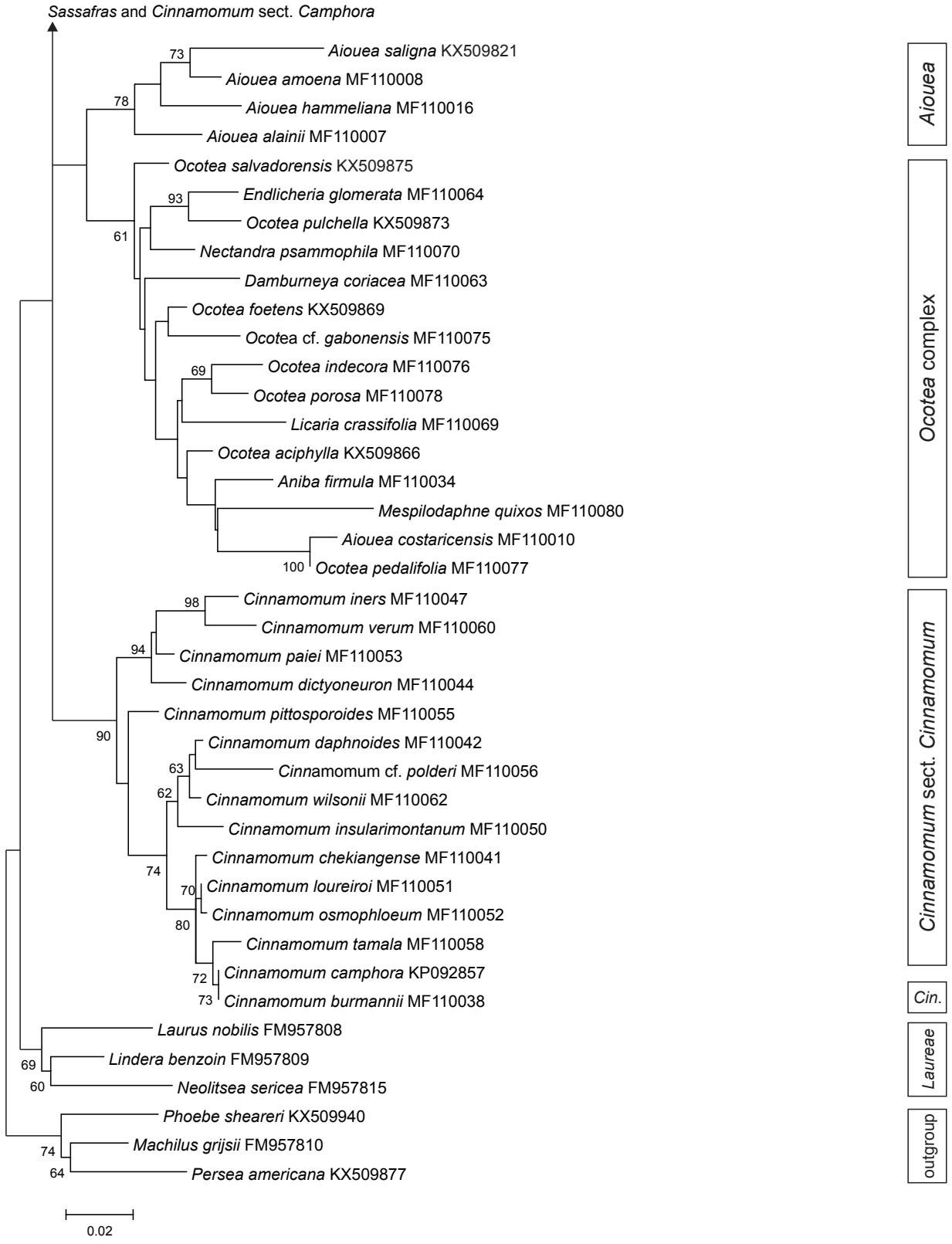


Fig. 7. Result of a maximum likelihood analysis of the ITS data, part 2. Numbers next to branches are bootstrap values ≥ 50%. *Cin.* = *Cinnamomum* sect. *Cinnamomum*.

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

JN115020, JX242469 and MF096119–MF096121 form the sister group to all remaining taxa, separated from them (and often also among each other) by rather long branches. The most frequent ITS copy of the plant in question (MF110039) is nested in a clade consisting of *C. camphora* sequences, whereas the second most frequent copy (MN480757) is nested among *C. parthenoxylon* sequences. However, only a few of the nodes in this analysis reached significant bootstrap support. Exclusion of the aberrant sequences AF272260, JN115020, JX242469 and the sequences covering the ITS-2 region only does not change most bootstrap values significantly, except that the one for *C.* sect. *Camphora* increases to 99%.

If the ITS-1 part is analysed separately (not shown), the sequences AF272260 and JX242469 form the sister group to the remaining *Cinnamomeae*, after separation of *Cinnamomum* sect. *Cinnamomum*, and *Sassafras* is unresolved with respect to *C.* sect. *Camphora* and the clade consisting of *Aiouea* and the *Ocotea* complex. Sequence JN115020 was not included because it covers only about half of the ITS-1 region.

If the ITS-2 part is analysed separately (not shown), not even the outgroup is recovered as monophyletic. *Machilus grijsii* and *Phoebe sheareri* appear as sister group to a clade consisting of the *Laureae*, *Sassafras* and the aberrant *Cinnamomum camphora* sequences JX242469, MF096119, MF096120 and MF096121, whereas *Persea americana* appears as sister taxon to the remainder of *C.* sect. *Camphora*. *Cinnamomum* sect. *Cinnamomum* appears as sister group to the clade consisting of *Persea americana* and *C.* sect. *Camphora*, and all these taxa appear as sister group of the clade consisting of *Aiouea* and the *Ocotea* complex. The changed relationships persist even if the aberrant sequences are excluded, but none of them reaches 50% bootstrap support.

Discussion

A comparison of our results with previously published sequences revealed on the one hand considerable differences among sequences (supposedly) belonging to the same species, and on the other hand unexpected agreement among sequences of (supposedly) different taxa. Differences among DNA sequences (supposedly) belonging to the same taxon can be due to a variety of causes. Ideally, they would be evidence of intraspecific variability, but other possible causes such as incorrect determinations, contaminations, inadvertent swapping of samples before or during lab work, paralogous sequences or pseudogenes should be taken into account as well. It therefore makes sense to compare the results of different studies, in order to detect possible errors. Agreement of sequences among different taxa can be (and frequently is) due to low divergence in the molecular markers examined. Nevertheless, it also may help to detect errors.

In the chloroplast sequences, at least those of the single copy regions used here, there should be no problems with paralogous sequences or pseudogenes. It is therefore intriguing that such different chloroplast sequences have been found in such a widespread and well-known species as *Cinnamomum camphora*. The differences between the cp groups #1 and #2 observed here in some cases (*psbA-trnH*, *matK*) are larger than those found in the same molecular markers among closely related but morphologically clearly different species of *Lauraceae* in recent studies (e.g. Rohwer & Rudolph 2005; Huang & al. 2016; Rohde & al. 2017; Trofimov & al. 2016, 2019). On the other hand, sequences submitted as *C. camphora* and *C. parthenoxylon* were frequently found to share the same haplotype. This supports the statement of Wu & al. (2019) that "the species delimitation and interrelationship of *C. camphora* and *C. parthenoxylon* may need further investigation."

The two species are usually treated as distinct, but scarcely any of the key characters separating them (e.g. in Li & al. 2008) appears to be entirely constant. The leaves are mostly triplinerved at the base in *Cinnamomum camphora* and penninerved in *C. parthenoxylon*, but not all leaves of *C. camphora* are triplinerved, especially on sterile branches (Fig. 1C). The lower leaf surface is usually glaucous in *C. camphora*, whereas it may be glaucous or not in *C. parthenoxylon*. Domatia in the axils of the (lower) secondary veins are common in *C. camphora*, but sometimes they are inconspicuous or even missing in a few leaves, whereas they are generally inconspicuous or sometimes absent in *C. parthenoxylon*. It is therefore quite possible that the two species may be mixed up, especially if determinations are based on sterile branches only. In addition, specimens identified as *C. parthenoxylon*, even by researchers familiar with the group, e.g. by A. J. G. H. Kostermans in the Naturalis Leiden herbarium (L, available on https://bioportal.naturalis.nl/), show a rather wide range of morphological variation, so that it is possible that more than one species might be involved in this complex.

Misidentifications are certainly responsible for some of the sequence diversity observed here. Numerous *Lauraceae* species, even in different genera, are so similar to one another that even experienced experts have a substantial rate of errors in their identifications (Liu & al. 2017b), especially of sterile material. Misidentifications can be detected by BLAST searches, if different sequences obtained from the same voucher consistently point to another species, of which several sequences prepared by different authors are available in GenBank. Among the sequences examined here, those based on "isolate SCBGP173_2" (Liu & al. 2015) may be such a case. The ITS sequence KP092857 differs by 27–29 substitutions from the majority of the other *Cinnamomum camphora* sequences and agrees best with *C. burmannii*. In the ITS-2 region, KP092857 is completely identical with *C. burmannii*, and in the result of the Maximum Likelihood analysis it nested among the species of *C.* sect. *Cinnamomum*, as sister tax-

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

on to *C. burmannii*. The *psbA-trnH* sequence KP095536 differs by only three substitutions from most specimens of *C. camphora*, but is likewise identical with several accessions of *C. burmannii*. The *matK* sequence KP093546 appears compatible with *C. camphora*, but several sequences of *C. burmannii* differ by just a single mutation within the small, conserved region covered by this fragment. We therefore assume that "isolate SCBGP173_2" (IBSC) may indeed represent *C. burmannii*. *Cinnamomum burmannii* is a species of *C.* sect. *Cinnamomum*, which is characterized by (sub)opposite, strongly triplinerved or trinerved leaves without domatia, vegetative buds with inconspicuous or no bud scales, and fruits with persistent tepal bases on the margin of the cupule, whereas *C. camphora* and *C. parthenoxylon* are species of *C.* sect. *Camphora*, characterized by alternate, penninerved to weakly triplinerved leaves with domatia, vegetative buds with distinct bud scales (perulate), and cupules without remnants of tepals. Both *C. burmannii* and *C. camphora* are frequently cultivated, e.g. as street trees in S China. They are usually quite different, at least when flowering or fruiting, but sometimes the leaves are not exactly opposite in *C. burmannii*, and the lowermost secondary veins may be less prominent, so that misidentification of sterile material becomes conceivable.

Large distances separating a particular sequence from the majority of its supposedly conspecific sequences may also point to either a misidentification or a contamination. The *matK* sequences JQ435499 ("*Cinnamomum camphora*") and JQ435498 ("*C. parthenoxylon*") differ not only from these two species but from all core *Lauraceae*, more than these differ among each other. In this case, however, we could not find any highly similar sequence, so that we cannot tell which taxon they might represent.

The situation is more complicated in the much more variable ITS sequences. If both ITS and chloroplast sequences are available of a particular voucher specimen, then ITS group #1 usually corresponds to cp group #1 and ITS group #3 corresponds to cp group #2. However, there are exceptions. Among the sequences covering only the ITS-2 part, the *Cinnamomum parthenoxylon* sequences KX546421 and KX546593, both based on the collection Ci X. Q. 0036 (HITBC), are in ITS group #3, but the *psbA-trnH* and *matK* sequences from the same voucher (KX546119 and KX546924) are in cp group #1. Sequence JX242469 differs considerably from all other sequences, even in the 5.8S rDNA region, which otherwise is highly conserved among the *Cinnamomeae*. If it had been the only aberrant sequence, we would have dismissed it as an occasional sequencing error. However, it shares (at least) 34 substitutions and 5 indels compared to ITS group #1 with JN115020, 18 substitutions and 2 indels with the ITS-1 part of AF272260 (of which the ITS-2 part is identical with ITS group #1), and 31 substitutions and 6 indels with the ITS-2 sequences MF096119–MF096121. Most of the substitutions and all of the indels are shared with at least one of the other aberrant sequences, often with all of them (if they cover the respective region). JN115020 even shares eight of its 15 substitutions in the 5.8S rDNA with JX242469. This suggests that ITS pseudogenes may have been sequenced in these cases. In any case, one should be wary if an ITS sequence shows an unexpectedly large divergence from those of supposedly closely related taxa.

If we exclude suspected misidentifications and pseudogene sequences, there still is considerable diversity among the sequences submitted as *Cinnamomum camphora* or *C. parthenoxylon*. Based on morphology alone, the questionable plant with the accession number 2006-1425 in the botanical garden of Munich could easily be accommodated in *C. camphora*, as the variability of the leaves on sterile branches is much larger in this species than on the fertile branches that are commonly photographed or preserved in herbaria. Nevertheless, our results are consistent with the assumption that the plant may be of hybrid origin. In the chloroplast genome, all or by far most reads of a certain amplicon were identical. The chloroplast genome is assumed to be maternally inherited, although to our knowledge this has not been checked in the *Lauraceae* so far. Only in one part of the *trnK* intron we found indications of different copies, but these were at least twelve times less frequent than the dominant copy. We cannot really explain these additional copies. Occasional transmission of a proplastid from the male gametophyte might be an option, but this is purely speculative. A PCR error or PCR recombination in an early cycle appears even more likely (Simon & al. 2012).

In the biparentally inherited nuclear ITS region, on the other hand, there is considerable evidence for different copies in this individual. High-throughput DNA sequencing has replaced cloning as the means to investigate the presence of multiple, divergent copies of nuclear ribosomal cistrons including ITS. For example, Zozomová-Lihová & al. (2014) supported previous hypotheses on the parents of *Cardamine* ×*schulzii* Urbanska-Worytkiewicz using 454 sequencing of ribosomal DNA. Furthermore, they found a bias toward the maternal copy. Similarly, in our case the second most frequent copy was found to be only about $^{1}/_{3}$ to $^{2}/_{3}$ as frequent as the dominant copy, and this is obviously sufficient to lead to heavily disturbed sequences in direct Sanger sequencing. In both the haplotype network analysis and the maximum likelihood analysis the dominant copy is found among sequences of *Cinnamomum camphora*, as also the key characters suggested, but the secondary copy is found in a different group, consisting mainly of sequences submitted as *C. parthenoxylon*. The amount of intraspecific sequence variation is not yet well known in the *Lauraceae*, although a first attempt at its quantification has been made by Liu & al. (2012). However, given the amount of sequence difference between the two copies, intraspecific variation is unlikely.

It still is uncertain if the second parent really is *Cinnamomum parthenoxylon*, because of the morphological

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

and molecular heterogeneity of the material attributed to this taxon (see above). For comparison, photos of a plant that we identified as *C. parthenoxylon* are shown in Fig. 1E, F. The ITS sequence (MF110054) obtained from this plant (*Rohwer 178*, MJG), however, differs by eight substitutions from the sequence of the suspected hybrid, and in the maximum likelihood analysis it is unresolved with respect to most other clades in *C.* sect. *Camphora*.

In our experience, intraspecific differences in ITS sequences are not really rare, at least if a species is widespread (compare, e.g., *Ocotea aciphylla* (Nees & Mart.) Mez, GenBank accession numbers DQ787422, GQ480374 and KX509866). In contrast, intraindividual variation is seldom recognized by direct amplification and sequencing, likely due to rarity of hybridization of individuals with divergent ITS-copies. In cases of offspring of parents with divergent but closely related ITS-copies just one or two double peaks are invariably retrieved in repeated sequencing attempts, suggesting polymorphisms in the respective sequence positions. Such sequences have been described, e.g. by Morden & al. (2015) for a few collections of *Cryptocarya mannii* Hillebr. They remain readable and can be included in phylogenetic analyses. Length mutations (indels) are more problematic. A single length mutation between two primer sites, most frequently a simple repeat of an adjacent motif, often can be resolved by sequencing toward this point from both sides. If there are several length mutations, however, cloning and (Sanger) sequencing of several clones used to be necessary to sort the different copies apart. In practice, limited resources often restricted these efforts and made it more cost-effective to omit a particular specimen and try another one. Since specific and generic delimitations among the *Lauraceae* are still a difficult issue, it is important to assess the possible role of hybridization in the evolution of the family. High-throughput sequencing now offers the opportunity to include collections yielding mixed signals with conventional methods, giving new insights into a still poorly explored field.

Acknowledgements

We thank Günter Gerlach for sending (twice) material of the plant cultivated as *Cinnamomum porrectum* in the Botanical Garden of Munich. Anna Maria Vogt is gratefully acknowledged for extracting the DNA from this specimen (and many others) and the pre-amplifications of the genome regions examined. An anonymous reviewer is also thanked for comments on an earlier version of this article.

References

Ando H., Setsuko S., Horikoshi K., Suzuki H., Umehara S., Inoue-Murayama M. & Isagi Y. 2013: Diet analysis by next-generation sequencing indicates the frequent consumption of introduced plants by the critically endangered red-headed wood pigeon (*Columba janthina nitens*) in oceanic island habitats. – Ecol. Evol. **3:** 4057–4069.

Beyra Matos A. & Lavin M. 1999: Monograph of *Pictetia* (*Leguminosae–Papilionoideae*) and review of the *Aeschynomeneae*. – Syst. Bot. Monographs **56:** 1–93.

Blattner F. R. 1999: Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. – BioTechniques **27:** 1180–1186.

Bushnell B., Rood J. & Singer E. 2017: BBMerge – accurate paired shotgun read merging via overlap. – PLoS ONE **12(10):** e0185056.

Chen S., Yao H., Han J., Liu C., Song J., Shi L., Zhu Y., Ma X., Gao T., Pang X., Luo K., Li Y., Li X., Jia X., Lin Y. & Leon C. 2010: Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. – PLoS ONE **5(1):** e8613.

Chen C., Zheng Y., Liu S., Zhong Y., Wu Y., Li J., Xu L. A. & Xu M. 2017: The complete chloroplast genome of *Cinnamomum camphora* and its comparison with related *Lauraceae* species. – PeerJ **5:** e3820.

Cock P. J. A., Antao T., Chang J. T., Chapman B. A., Cox C. J., Dalke A., Friedberg I., Hamelryck T., Kauff F., Wilczynski B. & de Hoon M. J. L. 2009: Biopython: freely available Python tools for computational molecular biology and bioinformatics. – Bioinformatics **25:** 1422–1423.

Elansary H. O. 2013: Towards a DNA barcode library for Egyptian flora, with a preliminary focus on ornamental trees and shrubs of two major gardens. – DNA Barcodes (Berlin) **1:** 46–55.

Furnier G. R., Cummings M. P. & Clegg M. T. 1990: Evolution of the avocados as revealed by DNA restriction fragment variation. – J. Heredity **81:** 183–188.

Hahn W. J. 2002: A phylogenetic analysis of the arecoid line of palms based on plastid DNA sequence data. – Molec. Phylogen. Evol. **23:** 189–204.

Hollingsworth P. M., Graham S. W. & Little D. P. 2011: Choosing and using a plant DNA barcode. – PLoS ONE **6(5):** e19254.

Huang J. F., Li L., van der Werff H., Li H. W., Rohwer J. G. & Crayn D. M. 2016: Origins and evolution of cinnamon and camphor: a phylogenetic and historical biogeographical analysis of the *Cinnamomum* group (*Lauraceae*). – Molec. Phylogen. Evol. **96:** 33–44.

Huang X. & Madan A. 1999: CAP3: a DNA sequence assembly program. – Genome Res. **9:** 868–877.

Käss E. & Wink M. 1997: Molecular phylogeny and phylogeography of the genus *Lupinus* (family *Leguminosae*) inferred from nucleotide sequences of the *rbcL* gene and ITS 1+2 sequences of rDNA. – Pl. Syst. Evol. **208:** 139–167.

Katoh K. & Standley D. M. 2013: MAFFT multiple sequence alignment software version 7: improvements

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

in performance and usability. – Molec. Biol. Evol. **30:** 772–780.

Kojoma M., Kurihara K., Yamada K., Sekita S., Satake M. & Iida O. 2002: Genetic identification of cinnamon (*Cinnamomum* spp.) based on the *trnL-trnF* chloroplast DNA. – Pl. Med. (Stuttgart) **68:** 94–96.

Kong D., Ma C., Zhang Q., Li L., Chen X., Zeng H. & Guo D. 2014: Leading dimensions in absorptive root trait variation across 96 subtropical forest species. – New Phytol. **203:** 863–872.

Kopp L. E. 1966: A taxonomic revision of the genus *Persea* in the western hemisphere (*Perseae–Lauraceae*). – Mem. New York Bot. Gard. **14:** 1–120.

Kress W. J. & Erickson D. L. 2007: A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding *trnH-psbA* spacer region. – PLoS ONE **2(6):** e508.

Leigh J. W. & Bryant D. 2015: POPART: full-feature software for haplotype network construction. – Methods Ecol. Evol. **6:** 1110–1116.

Li X. W., Li J. & van der Werff H. 2008: *Cinnamomum* Schaeffer. – Pp. 166–187 in: Wu Z. Y., Raven P. H. & Hong D. Y. (ed.), Flora of China **7.** *Menispermaceae* through *Capparaceae*. – Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. – Online at http://www.efloras.org/florataxon.aspx?flora_id=2&taxon_id=107109.

Liu J., Shi L., Song J., Sun W., Han J., Liu X., Hou D., Yao H., Li M. & Chen S. 2017a: BOKP: a DNA barcode reference library for monitoring herbal drugs in the Korean pharmacopeia. – Frontiers Pharmacol. **8:** 931.

Liu J., Yan H. F., Newmaster S. G., Pei N., Ragupathy S. & Ge X. J. 2015: The use of DNA barcoding as a tool for the conservation biogeography of subtropical forests in China. – Diversity & Distrib. **21:** 188–199.

Liu Z., Chen S. L., Song J. Y., Zhang S. J. & Chen K. L. 2012: Application of deoxyribonucleic acid barcoding in *Lauraceae* plants. – Pharmacogn. Mag. **8:** 4–11.

Liu Z. F., Ci X. Q., Li L., Li H. W., Conran J. G. & Li J. 2017b: DNA barcoding evaluation and implications for phylogenetic relationships in *Lauraceae* from China. – PLoS ONE **12(4):** e0175788.

Massoni J., Forest F. & Sauquet H. 2013: Increased sampling of both genes and taxa improves resolution of phylogenetic relationships within *Magnoliidae*, a large and early-diverging clade of angiosperms. – Molec. Phylogen. Evol. **70:** 84–93.

Morden C. W., Harbin S. C., Rohwer J. G., Portner T. & Yorkston M. 2015: Characterization of Hawaiian *Cryptocarya* (*Lauraceae*): recognition of a critically endangered species and relation to non-Hawaiian congeners. – Pacific Sci. **69:** 103–115.

Pei N., Lian J. Y., Erickson D. L., Swenson N. G., Kress W. J., Ye W. H. & Ge X. J. 2011: Exploring tree-habitat associations in a Chinese subtropical forest plot using a molecular phylogeny generated from DNA barcode loci. – PLoS ONE **6(6):** e21273.

Qiu Y. L., Dombrovska O., Lee J., Li L., Whitlock B. A., Bernasconi-Quadroni F., Rest J. S., Davis C. C., Borsch T., Hilu K. W., Renner S. S., Soltis D. E., Soltis P. S., Zanis M. J., Cannone J. J., Gutell R. R., Powell M., Savolainen M., Chatrou L. W. & Chase M. W. 2005: Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. – Int. J. Pl. Sci. **166:** 815–842.

Rohde R., Rudolph B., Ruthe K., Lorea-Hernández F. G., Moraes P. L. R. de, Li J. & Rohwer J. G. 2017: Neither *Phoebe* nor *Cinnamomum* – the tetrasporangiate species of *Aiouea* (*Lauraceae*). – Taxon **66:** 1085–1111.

Rohwer J. G. 1993: Flora neotropica monographs **60.** *Lauraceae*: *Nectandra*. – New York: New York Botanical Garden.

Rohwer J. G., Moraes P. L. R. de, Rudolph B. & van der Werff H. 2014: A phylogenetic analysis of the *Cryptocarya* group (*Lauraceae*), and relationships of *Dahlgrenodendron*, *Sinopora*, *Triadodaphne*, and *Yasunia*. – Phytotaxa **158:** 111–132.

Rohwer J. G., Li J., Rudolph B., Schmidt S. A., van der Werff H. & Li H. W. 2009: Is *Persea* (*Lauraceae*) monophyletic? Evidence from nuclear ribosomal ITS sequences. – Taxon **58:** 1153–1167.

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

Simon U. K., Trajanoski S., Kroneis T., Sedlmayr P., Guelly C. & Guttenberger H. 2012: Accession-specific haplotypes of the internal transcribed spacer region in *Arabidopsis thaliana* – a means for barcoding populations. – Molec. Biol. Evol. **29:** 2231–2239.

Steele K. P. & Vilgalys R. 1994: Phylogenetic analyses of *Polemoniaceae* using nucleotide sequences of the plastid gene *matK*. – Syst. Bot. **19:** 126–142.

Taberlet P., Gielly L., Pautou G. & Bouvet J. 1991: Universal primers for amplification of three non-coding regions of chloroplast DNA. – Pl. Molec. Biol. **17:** 1105–1109.

Tamura K., Stecher G., Peterson D., Filipski A. & Kumar S. 2013: MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. – Molec. Biol. Evol. **30:** 2725–2729.

Tange O. 2011: GNU Parallel – the command-line power tool. – login: USENIX Mag. **36:** 42–47.

Trofimov D., Moraes P. L. R. de & Rohwer J. G. 2019: Towards a phylogenetic classification of the *Ocotea* complex (*Lauraceae*): classification principles and reinstatement of *Mespilodaphne*. – Bot. J. Linn. Soc. **190:** 25–50.

Trofimov D., Rudolph B. & Rohwer J. G. 2016: Phylogenetic study of the genus *Nectandra* (*Lauraceae*), and reinstatement of *Damburneya*. – Taxon **65:** 980–996.

Wu C. S., Wang T. J., Wu C. W., Wang Y. N. & Chaw S. M. 2017: Plastome evolution in the sole hemiparasitic

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

genus laurel dodder (*Cassytha*) and insights into the plastid phylogenomics of *Lauraceae*. – Genome Biol. Evol. **9:** 2604–2614.

Wu Y., Wei W., Liu L., Liu G., Guo Q. Q. & Qian Z. Q. 2019: The complete chloroplast genomes of the evergreen tree species *Cinnamomum camphora* and *Cinnamomum parthenoxylon* (*Laurales*: *Lauraceae*). – Mitochondrial DNA Part B, **4:** 813–814.

Zozomová-Lihová J., Mandáková T., Kovaříková A., Mühlhausen A., Mummenhoff K., Lysak M. A., Kovařík A. 2014: When fathers are instant losers: homogenization of rDNA loci in recently formed *Cardamine* × *schulzii* trigenomic allopolyploid. – New Phytol. **203:** 1096–1108.

Willdenowia
Open-access online edition bioone.org/journals/willdenowia
Online ISSN 1868-6397 · Print ISSN 0511-9618 · Impact factor 1.156
Published by the Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin
© 2019 The Authors · This open-access article is distributed under the CC BY 4.0 licence

Downloaded From: https://complete.bioone.org/journals/Willdenowia on 29 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use