Taiwania 67(4): 560–570, **2022**
DOI: 10.6165/tai.2022.67.560

Molecular phylogenetic relationships and taxonomy position of 161 *Camellia* species in China

Zhen PANG[1,†], Yi-Ling WANG[2,†*], Nitin MANTRI[3], Yang WANG[1], Xue-Jun HUA[1], Yan-Ping QUAN[1], Xuan ZHOU[1], Zheng-Dong JIANG[1], Zhe-Chen QI[1], Hong-Fei LU[1,*]

1. *College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Zhejiang Province Key Laboratory of Plant Secondary Metabolism and Regulation, Hangzhou 310018, China.* ***2.*** *College of Life Sciences, Shanxi Normal University, Taiyuan 030024, China.* ***3.*** *The Pangenomics Lab, School of Science, RMIT University, Melbourne, Victoria 3083, Australia.* [†] *These authors contributed equally to the work.* [*]*Corresponding authors:YLW's tel: 86-18603570866; E-mail: ylwangbj@hotmail.com; HFL's tel: 86-13566997763; E-mail: luhongfei0164@163.com*

(Manuscript received 15 June 2022; Accepted 17 October 2022; Online published 24 October 2022)

ABSTRACT: *Camellia* is the largest and most important genus in the family Theaceae, with many species being of great economic, ornamental and ecological value. However, the phylogenetic resolution of these species has been difficult due to interspecific hybridization and polyploidy. Consequently, the interspecies relationships of the genus *Camellia* are still hotly debated. In this study, four chloroplast genomic regions (*matK*, *rbcL*, *ycf1*, *trnL-F*) were used as markers among 161 species representing all four sub-genera within this genus to investigate the phylogeny and interspecies relationship of the genus *Camellia*. The results showed that the 161 species of the genus *Camellia* could be grouped into 13 clades (A-M). Clades A and B mainly consisted of sect. *Camellia*. Clades C and I were made up of sect. *Theopsis* and sect. *Eriandria*. Clades D and J were composed of species from sect. *Thea*. Clade F consisted of the sect. *Paracamellia* species, whilst Clades G and M included sect. *Furfuracea* species. Clade H contained sect. *Tuberculata* and most species of sect. *Pseudocamellia*, whereas Clades K and L comprised the sect. *Chrysantha* species. These results supported that 161 *Camellia* species form paraphyletic groups, rather than a monophyletic group. And they demonstrated that the taxonomic position of related species could be resolved to some extent via sequencing markers in organelle genome, thus providing valuable cytoplasmic genetic information or maternal genetic information for accurately identifying species, clarifying taxonomy and reconstructing the phylogeny of various *Camellia* species.

KEY WORDS: *Camellia*, chloroplast gene, *matK*, molecular phylogeny, *rbcL*, section, taxonomy position, *trnL-F*, *ycf1*.

INTRODUCTION

Camellia Linnaeus is the largest genus of the family Theaceae with more than 120–300 species (Vijayan *et al*., 2009; Ly *et al*., 2022). They are mainly distributed in East and Southeast Asia (Zhao *et al*., 2022). In China, more than 239 species from this genus grow in South and Southwest regions, such as Yunnan and Sichuan provinces (Yang *et al*., 2009; Yu *et al*., 2021). The genus *Camellia* contains many species, which are economically, ornamentally and ecologically valuable (Cabrera *et al*., 2006; Khan *et al*., 2007). Some species are famous for their ornamental flowers (Chen *et al*., 2007), while others are the sources of high quality edible oils (Gramza and Korczak, 2005). Moreover, the leaves of *Camellia sinensis* (L.) O. Kuntze contain more than 700 different chemical compounds that have been proposed to be beneficial for human health (Saha *et al*., 2012), and utilized commercially as tea products in more than 25% of the countries globally (Wang *et al*., 2012). Some species, for instance, *C. japonica* Linnaeus are very popular and desirable in agriculture, horticulture and even scientific research (Ricardo *et al*., 2019; Chung *et al*., 2003).

The genus *Camellia* has been constantly supplemented and revised since its establishment (Linnaeus, 1753; Cohen-Stuart, 1916; Melchior, 1925; Nakai, 1940). Currently, three major, but to some extent, contradictory classification systems co-exist for the genus *Camellia* proposed by Sealy, Chang and Ming, respectively. Firstly, the three systems disagree with each other on the boundaries of sub-genera, sections, and species, as well as the circumscription and relationships between species, respectively. Secondly, in terms of sub-generic divisions, none was offered in Sealy's system, while four and two were offered in Chang's and Ming's system, respectively. For instances, the taxonomic system, suggested by Sealy in 1958, contained twelve sections and 82 species with no sub-genus. Subsequently, Chang proposed a new phylogenetic classification of *Camellia* including four sub-genera (Subgen. *Protocamellia* Hung T. Chang, Subgen. *Camellia*, Subgen. *Thea* (L.) Hung T. Chang and Subgen. *Metacamellia* Hung T. Chang), nineteen sections and 196 species (Chang, 1981). In 1998, Chang made a comprehensive revision of the taxonomic system of the genus *Camellia*, retaining four sub-genera, identifying 20 sections and 280 species (Chang, 1998). One year after, another newer system was proposed by Ming based on morphology, contained only two sub-genera (Subgen. *Thea* and Subgen. *Camellia*), fourteen sections and about 119 species (Ming, 1999). In addition, some sections were treated differently in the three systems. In Ming's system (Ming, 1999), sect.

Chrysantha and sect. *Pleurocarpus* Hung T. Chang were merged into sect *Archecamellia*, and sect. *Furfuracea* Hung T. Chang, sect. *Pseudocamellia* Sealy and sect. *Protocamellia* Hung T. Chang were placed into sect. *Heterogenea* Sealy. So far, these classification systems disagree in many aspects, especially in regard to the circumscription of sub-genera, sections, and species. The intragenus classification of *Camellia* is still a controversial and confusing issue. In term of the morphological characters of *Camellia*, Sealy (1958), Chang (1981, 1998) and Ming (2000) also had different views. Sealy (1958) and Ming (2000) might give a high taxonomic value to the traits of pedicel, bracteoles and sepals, whereas Chang (1981, 1998) probably valued the characters of filaments more. Thus, it is crucial to seek reliable evidence to rebuild the classification system of *Camellia* so as to establish accurate phylogenetic relationships within the genus.

During the past decades, numerous molecular phylogenetic studies have been undertaken to resolve the classification issues within *Camellia*. Using four DNA sequences, a monophyletic classification was suggested among 21 *Camellia* species, despite that the relationship among sections was not clear (Yang *et al*., 2006). Internal transcribed spacer (ITS) sequences were also used to organize 112 *Camellia* species into eight major clades, but the inter-relationships between clades remained unresolved (Vijayan *et al*., 2009). Analysis of the genomic sequences of the chloroplast from six *Camellia* species indicated that the phylogenetic relation did not agree with any of the traditional classification methods (Yang *et al*., 2014). Interestingly, chloroplast genomic sequencing of 13 *Camellia* species supported that *C. pubicosta* Merr. may be classified into sect. *Thea* (L.) Dyer, as proposed in Sealy's and Chang's systems. The analysis of five genomic regions revealed *Camellia* was paraphyletic and a widespread hybridization occurred between its species (Zhang *et al*., 2014). Orthologous nuclear *RPB2* introns 11–15 and 23, and *waxy* were sequenced for 99 taxa of *Camellia* to reconstruct its phylogenetic history. The results showed that the genus can be divided into two main clades, CI and CII. Ten supported subclades were also subsequently identified, which provided new insights into the phylogenetic relationships and systematics of *Camellia* (Zhao *et al*., 2022).

Compared with chloroplast genome, the nuclear ones are relatively larger and more complex, and richer in genetic variation, which have been widely used in the molecular systematic study of many taxa. The rDNA ITS (ribosomal transcriptional spacer) is the best example of such a kind of marker in plant molecular systematic study (Prihatini *et al*., 2020). However, their application in *Camellia* classification is limited by many factors. *Camellia* is not only a group with frequent hybridization but also has a common phenomenon of chromosome polyploidy. When nuclear gene markers are used in the phylogenetic analysis of this genus, interspecific hybridization and chromosomal polyploidy can easily lead to misinterpretation of phylogenetic relationships due to the characteristics of parental inheritance of nuclear genes (Fang *et al*., 2010). In addition, due to relatively large molecular weight and complex structure, the amplification of ITS fragment is rather difficult (Zou *et al*., 2013), which added another factor limiting the application of ITS in *Camellia* to certain extent.

On the other hand, the chloroplast genome (cp DNA) of plants is uniparentally inherited and is not affected by genetic recombination, which is a great advantage in elucidating the complex phylogenetic relationships between the species (Zhang *et al*., 2003). Furthermore, cp DNA has the characteristics of a high degree of sequence conservation, freedom from selection pressure and independent evolutionary routes. Therefore, molecular phylogenetic trees can be constructed without having to rely on other data (Zhu *et al*., 2018). Moreover, the non-coding region generally has a higher substitution rate of nucleotide than the coding region sequence, making it more suitable for the study of genetic relationship and genetic diversity within species (Wei *et al*., 2010). In recent years, cp DNA has been successfully applied in the phylogeny and taxonomy of some plant taxa, such as *Dieffenbachia* Schott, *Sedum* Linnaeus and *Deutzia* Thunberg (Kim *et al*., 2015; Lim and Choi, 2018; Wang *et al*., 2016). The chloroplast genes *rbcL* and *matK* have also been proposed as the core barcodes of plant species by CBOL Plant Working Group (Badr *et al*., 2020).

Recently, the phylogenetic relationship of 11 yellow-flowered *Camellia* species of section *Chrysantha* was analyzed using random amplified polymorphic DNA (*RAPD*) assay (Li *et al*., 2018). The sequence analysis of four chloroplast DNA loci allowed the resolution of phylogeny of three *Camellia* sections, *Longipedicellata* Hung T. Chang, *Chrysantha* and *Longissima* Hung T. Chang, but the systematic position of *C. longissima* and the relationship between sect. *Longissima* and sect. *Longipedicellata* were still unresolved (Lu *et al*., 2008). In addition, using genome-wide SNPs from RAD Sequencing, six distinct clusters were detected by phylogeny inference, and these clusters corresponded to six *Camellia* species/varieties (Yang *et al*., 2016). Taken together, molecular phylogenetic analysis has been extensively applied to resolve the classification issues of *Camellia* but there is no apparent clear structure. This is probably because that most of the previous works focused primarily limited amount of species within specific sections or taxa. Therefore, the intrageneric and interspecies relationships within the genus *Camellia* are worth further exploring.

Here, using broad sampling, the sequence variations of four chloroplast genes among 161 species from 16 sections of four sub-genera within *Camellia* were analyzed to reassess their phylogenetic relationships. The

results would help to re-establish the taxonomic framework of the genus *Camellia*.

MATERIALS AND METHODS

Sampling

Leaf samples from *Camellia* species representing four sub-genera (Subgen. *Protocamellia*, Subgen. *Camellia*, Subgen. *Thea* and Subgen. *Metacamellia*) and 16 sections of Chang's system were collected. These included: 2 species of sect. *Archecamellia*, 1 species of sect. *Stereocarpus* (Pierre) Sealy, 13 species of sect. *Paracamellia* Sealy, 5 species of sect. *Oleifera* Hung T. Chang, 9 species of sect. *Furfuracea*, 1 species of sect. *Luteoflora* Hung T. Chang, 49 species of sect. *Camellia*, 4 species of sect. *Pseudocamellia*, 14 species of sect. *Tuberculata* Hung T. Chang, 16 species of sect. *Chrysantha*, 2 species of sect. *Longipedicellata*, 2 species of sect. *Glaberrima* Hung T. Chang, 16 species of sect. *Thea*, 1 species of sect. *Longissima*, 22 species of sect. *Theopsis* Coh. Stuart, 3 species of sect. *Eriandria* Coh. Stuart, and 1 undefined species (*Camellia* sp.). Three species: *Gordonia acuminata* Hung T. Chang (*Polyspora speciosa*), *Gordonia longicarpa* Hung T. Chang (*Polyspora longicarpa*) and *Tutcheria hexalocularia* Hu et Liang ex Chang (*Pyrenaria spectabilis*) from the closely related genera: *Gordonia* J. Ellis and *Tutcheria* Dunn (all Theaceae) were chosen as outgroups (Table S1).

DNA extraction, PCR and sequencing

Total genomic DNA was extracted from silica gel-dried or herbarium specimen leaf material by the modified CTAB protocol (Moller *et al*, 1992). Nucleotide sequences for the four chloroplast loci, *i.e.*, *matK*, *rbcL*, *ycf1*, *trnL-F* were amplified using the following primers and sent for sequencing: *matK*F and *matK*R for *matK* region, *rbcL*20F and *rbcL*1406R for *rbcL* region, *ycf1*_66f and *ycf1*_2954r for *ycf1* region, and *trn*l and *trn*f for *trnL-F* region (Taberlet *et al*., 1991) (Table S2).

The PCR reactions were performed in 50 µl volume containing 10 ng of genomic DNA, 1.0 µM of each primer, and 25 µl 2X Master Mix. The amplification cycles were: a 94 °C initial hot start for 4 min, followed by 32 cycles of 94 °C for 1 min, 52 °C for 1.5 min and 72 °C for 1 min, and a final extension of 72 °C for 10 min.

Purified PCR products were sent to Hangzhou and sequenced by TsingKe Biological Technology Limited Company. The sequences were initially assembled using Assembly program in Geneious Pro (https://www.geneious.com, Kearse *et al*., 2012) and aligned using MAFFT (http://mafft.cbrc.jp/alignment/server/, Yamada *et al*., 2016), followed by manual adjustments using Geneious software.

Phylogenetic analyses

The incongruence length difference (ILD) test (Farris *et al*., 1995) was not required for the chloroplast sequences. Based on each gene, genetic regions and the combined sequence dataset using Geneious Pro, phylogenetic analyses were performed with gaps treated as missing data and indels coded as binary characters (simple indel coding). The most suitable model of evolution was determined using jModelTest2 on XSEDE (Miller *et al*., 2010). Model parameters were estimated and optimized separately for each gene and the combined sequences, respectively. Phylogenetic relationships were conducted using maximum parsimony (MP), maximum-likelihood (ML), and Bayesian inference (BI) methods, respectively (Feng *et al*., 2020).

The MP analyses were carried out using the program PAUP* version 4.0b10 (Swofford 2003), and each consisted of a heuristic search with 1000 replicates of random sequence addition with tree bisection-reconnection (TBR) branch-swapping and the MULTREES option on, saving up to 20 most-parsimonious trees (MPTs) per replicate to reduce the time spent in swapping large islands of trees (Wilgenbusch *et al*., 2004). All characters were unordered and equally weighted. Individual gap positions were treated as missing data. Internal support for clades was evaluated by 1000 bootstrap replicates (Felsenstein, 1985), each consisting of 20 replicates of random addition, TBR branch-swapping and saving up to 20 trees per heuristic replicate.

The ML analyses were conducted using the selected model GTR+I+G and implemented in RAxML V8.2.9 of CIPRES cluster (Stamatakis, 2014), with 1,000 bootstrap replicates under the GTRCAT model on the Cipres Science Gateway (Miller *et al*., 2010). The ML searches were performed at least twice to ensure a stable topology.

The BI was performed with Mrbayes on XSEDE (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) to select the best model of sequence evolution. According to the results of jModelTest, the models were chosen by the Akaike information criterion (AIC) and determined by AIC scores. Two runs were conducted in parallel with four Markov chains (one cold and three heated), with each running for 20,000,000 generations from a random. The first 5,000 trees (25%) from each run were discarded as burn-in. Resulting trees from the two independent runs were then pooled to produce one 50% majority-rule consensus tree. Bayesian posterior probabilities (PP) from the sampled trees after the burn-in period were used for generating the final tree.

RESULTS

The combined sequence was 5,661 bps long with 111 parsimony informative (2.0%) and 192 variation sites (3.4%) (Table 1). The data from individual gene analysis did not produce strongly supported phylogenetic trees. Therefore, the data based on the combined sequences

Table 1. Statistics of four chloroplast gene (*matK*, *rbcL*, *trnL-F*, *ycf1*) and the combined sequence.

DNA fragment	Matrix length	Number of simple sites(%)	Number of variant sites(%)
matK	2241	50(2.2)	81(3.6)
rbcL	1157	18(1.5)	25(2.1)
trnL-F	793	18(2.3)	34(4.2)
ycf1	1394	25(1.8)	52(3.7)
Combined	5661	111(2.0)	192(3.4)

Table 2. Parameters of best model from JmodelTest.

Gene	Combined	*matK*	*rbcL*	*trnL-F*	*ycf1*
aligned length	5661	2322	1157	793	1394
Length variation	5613	2317-2322	1156-1157	758-779	1370-1391
model	GTR+G+F	GTR+G+F	GTR+I+G+F	GTR+F	GTR+I+F
-lnL	8926.3397	2715.731	1855.2079	1342.0122	2361.0613
K	331	331	331	330	331
freqA	0.3213	0.3253	0.2645	0.3207	0.3684
freqC	0.1678	0.1592	0.2006	0.1445	0.1657
freqG	0.1849	0.1631	0.2526	0.1792	0.1566
freqT	0.3259	0.3524	0.2823	0.3556	0.3092
R(a) [AC]	2.3660	1.0000	1.0000	2.7288	2.6033
R(b) [AG]	2.3660	2.4412	2.6035	2.7288	2.6033
R(c) [AT]	0.2607	0.0001	1.0000	0.2732	0.1000
R(d) [CG]	1.0000	1.0000	2.6035	1.0000	0.1000
R(e) [CT]	2.3660	2.4412	2.6035	2.7288	2.6033
R(f) [GT]	1.0000	1.0000	1.0000	1.0000	1.0000
gamma shape	0.0210	0.1690	0.8160		
p-inv			0.9110		0.8520

were chosen to construct the phylogenetic trees (Table 2).

The analyses combined data confirmed that the genus *Camellia* was monophyletic (PP = 100, MLBS = 74%, MPBS = 44%). The ML and BI analyses of the combined dataset yielded topologies similar to the MP phylogeny (Fig. 1, Fig. S1, S2). It was worth noting that the MP tree had almost no structure, which may be caused by the attraction of long branches of sequences with large variation. The Bayesian tree that based on the combined chloroplast genes was selected to represent our results. Within the resulting phylogenetic tree, all the species analyzed clustered into 13 clades (Clade A, B, C, D, E, F, G, H, I, J, K, L and M) and five highly supported dichotomous branches. Among the 13 clades, A-E clades clustered into a large branch with a high bootstrap in ML and BI trees (PP = 100, MLBS = 69%).

Clade A mainly contained the species from sect. *Camellia*. Clustering *C. hunanica* and *C. tunganica* together was strongly supported by a high bootstrap values (PP = 99, MLBS = 67%, MPBS = 63%), whereas *C. phelloderma* and *C. villosa* were clustered into another group (PP = 100, MLBS = 87%, MPBS = 64%). The remaining species of sect. *Camellia* formed individual branch and were sororal to each other. Interestingly, this clade encompassed two species from different sections, namely, *C. macrosepala* of sect. *Theopsis* and *C. odorata* of sect. *Paracamellia* (Fig. 1, Fig. S1, S2).

The Clade B included all remaining species of sect. *Camellia*. Among them, seven species, namely *C. bailinshanica*, *C. pitardii*, *C. glabriperulata*, *C. stictoclada*, *C. paucipetala*, *C. reticulate* and *C. omeiensis*, were clustered into a group with high bootstrap values (PP = 99, MLBS = 57%). Similarly, *C. borealiyunnanica*, *C. brachygyna*, *C. jinshajiangica*, *C. tenuivalvis*, *C. weiningensis* and *C. oligophlebia* formed a strongly supported group with high bootstrap values (PP = 99, MLBS = 67%, MPBS = 60%). The third group included *C. albosericea* and *C. magniflora* with high fidelity (PP = 100, MLBS = 99%, MPBS = 96%). These three groups displayed paraphyletic relationship. The species *C. pitardii* var. *alba* formed a sole branch (Fig. 1, Fig. S1, S2).

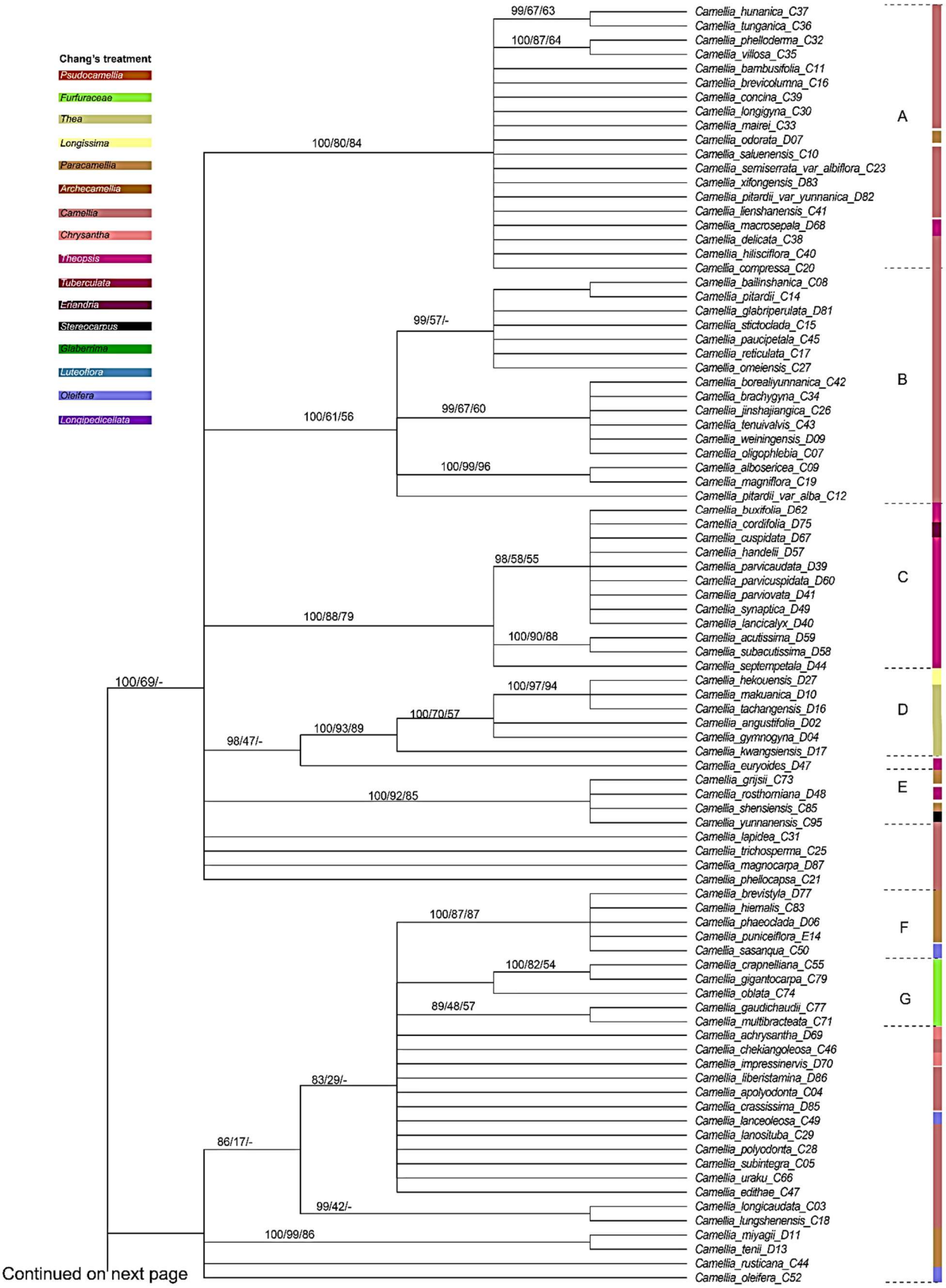
The Clade C consisted of species from sect. *Theopsis* and *C. cordifolia* of sect. *Eriandria*. There were two strongly supported subclades. One included eight species of sect. *Theopsis* (*C. buxifolia*, *C. cuspidate*, *C. handelii*, *C. parvicaudata*, *C. parvicuspidata*, *C. parviovata*, *C. synaptica* and *C. lancicalyx*) and *C. cordifolia* of sect. *Eriandria* (PP = 98, MLBS = 58%, MPBS = 55%), while the other included two species of sect. *Theopsis*, *C. acutissima* and *C. subacutissima* with high bootstrap values (PP = 100, MLBS = 90%, MPBS = 88%). The species *C. septempetala* formed a separate branch and demonstrated a paraphyletic relationship with other two branches (Fig. 1, Fig. S1, S2).

Except *C. hekouensis* of sect. *Longissima*, Clade D included species of sect. *Thea*. *C. kwangsiensis* formed a separate branch, which was the earliest differentiation and sister to other groups (PP = 100, MLBS = 70%, MPBS = 57%). A strongly supported group (PP = 100, MLBS = 97%, MPBS = 94%), which consisting of *C. hekouensis*, *C. makuanica* and *C. tachangensis*, was paraphyletic with the group containing *C. angustifolia* and *C. gymnogyna*. In the ML (Fig. S2) and BI (Fig. 1) phylogenetic tree, Clade D had a sister relationship with the branch of *C. euryoides* from sect. *Theopsis* (PP = 98, MLBS = 47%).

The Clade E contained four species. Two species (*C. grijsii* and *C. shensiensis*) were from the sect. *Paracamellia*, one (*C. rosthorniana*) from sect. *Theopsis* and one (*C. yunnanensis*) from sect. *Stereocarpus*. These four species formed a single branch (Fig. 1, Fig. S1, S2).

The Clade F was strongly supported and consisted of four species (*C. brevistyla*, *C. hiemalis*, *C. phaeoclada* and *C. puniceiflora*) from sect. *Paracamellia* and one (*C. sasanqua*) from sect. *Oleifera*. These five species formed a single subclade (Fig. 1, Fig. S1, S2).

The Clade G mainly comprised five species from sect. *Furfuracea*, namely *C. crapnelliana*, *C. gigantocarpa*, *C. oblata*, *C. gaudichaudii* and *C. multibracteata*. However,



Continued on next page

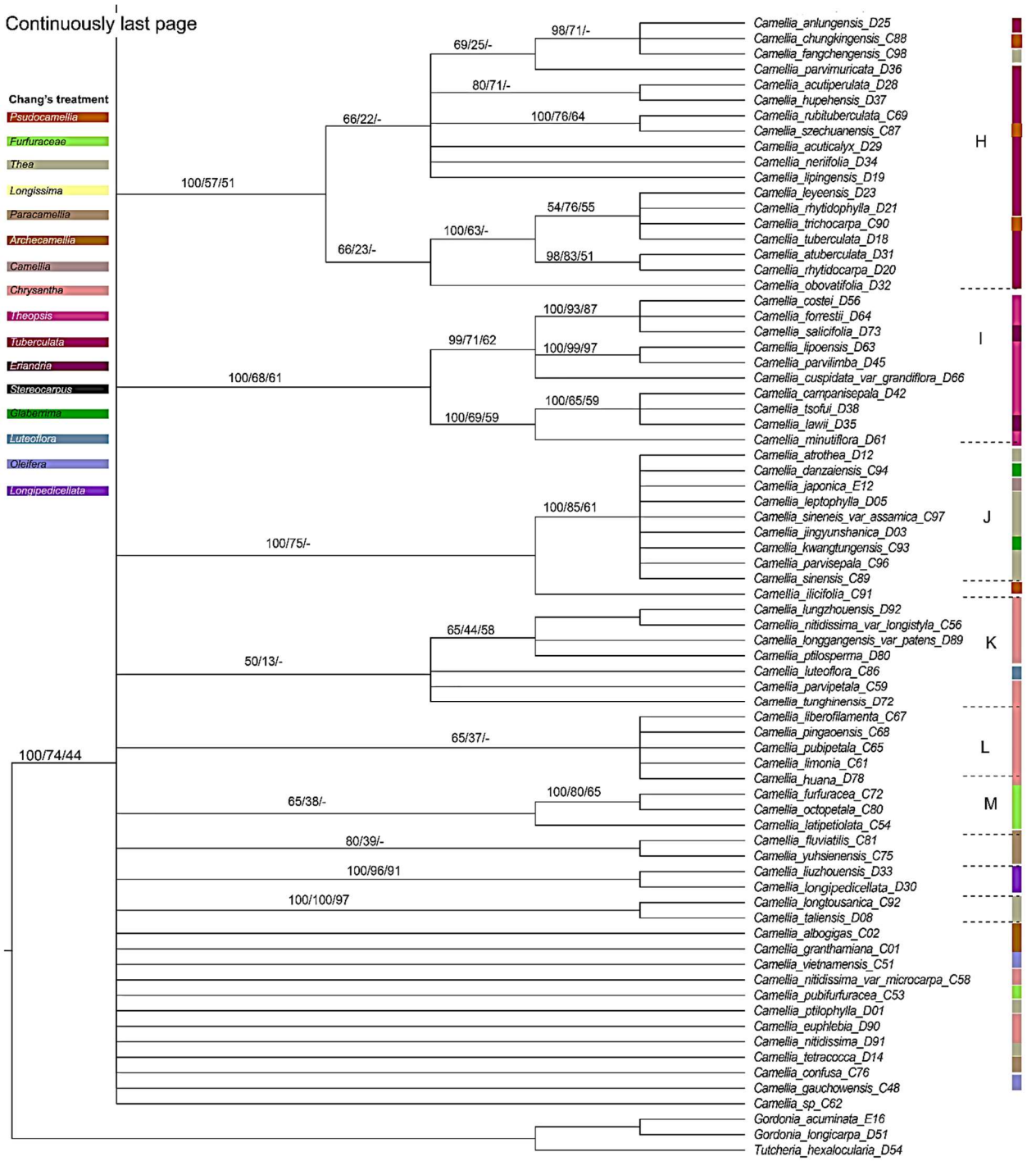


Fig. 1. Bayesian inference phylogenetic tree of *Camellia* inferred from all combined four sequenced chloroplast DNA regions (matK, rbcL, ycf1, trnL-F). Bayesian posterior probability, ML bootstrap and MP bootstrap (PP/ML-BS/MP-BS) were indicated on the major branch of the tree. The colored bars on the right of the tree indicate the sectional classification of Chang's treatment.

this clade in MP (Fig. S1) was not consistent with ML (Fig. S2) and BI (Fig. 1). Within BI, it didn't recognize this branch but was divided into two paraphyletic groups, one consisted of *C. crapnelliana*, *C. gigantocarpa* and *C. oblata*, and the other consisted of *C. gaudichaudii* and *C. multibracteata*. Within MP, the Clade G was not strongly supported as a monophyletic clade. Therefore, it's not suitable to treat it as a clade alongside the other 12 clades. Interestingly, the grouping together of *C. crapnelliana* and *C. gigantocarpa* was strongly supported in MP, ML and BI (PP = 100, MLBS = 82%, MPBS = 54%) (Fig. 1, Fig. S1, S2).

The Clade H mainly included 14 species of sect. *Tuberculata* (*C. anlungensis*, *C. parvimuricata*, *C. acutiperulata*, *C. hupehensis*, *C. rubituberculata*, *C. acuticalyx*, *C. neriifolia*, *C. lipingensis*, *C. leyeensis*, *C. rhytidophylla*, *C. tuberculate*, *C. atuberculata*, *C. rhytidocarpa* and *C. obovatifolia*). Additionally, this clade also contained three species of sect. *Pseudocamellia* (*C. chungkingensis*, *C. szechuanensis* and *C. trichocarpa*) and one species of sect. *Thea* (*C. fangchengensis*). Within ML (Fig. S2) and BI (Fig. 1), Clade H was divided into two sister groups. However, these two sister groups were weakly supported.

Apart from one species (*C. lawii*) of sect. *Eriandria*, Clade I mainly contained species of sect. *Theopsis* within two sister groups. *C. costei*, *C. forrestii*, *C. salicifolia*, *C. lipoensis*, *C. parvilimba* and *C. cuspidata* var. *grandiflora* formed one group, while *C. campanisepala*, *C. tsofui* and *C. minutiflora* clustered into another group (Fig. 1, Fig. S1, S2).

The Clade J contained some species of sect. *Thea*, two (*C. kwangtungensis* and *C. danzaiensis*) of sect. *Glaberrima* and one (*C. japonica*) of sect. *Camellia*. All species in this clade formed a sole branch. In ML (Fig. S2) and BI (Fig. 1), Clade J was sister to the branch of *C. ilicifolia* from sect. *Pseudocamellia* with a high bootstrap value (PP = 100, MLBS = 75%).

The Clade K contained species from sect. *Chrysantha* and only one (*C. luteoflora*) from sect. *Luteoflor*. However, this clade was not supported in the MP tree (Fig. S1). The group was made up of *C. lungzhouensis*, *C. nitidissima* var. *longistyla*, *C. longgangensis* var. *patens* and *C. ptilosperma* that were well-supported in all three topology trees (Fig. 1).

The Clade L comprised five species of sect. *Chrysantha*, i.e. *C. liberofilamenta*, *C. pingaoensis*, *C. pubipetala*, *C. limonia* and *C. huana*. This clade formed a similar topology structure in ML (Fig. S2) and BI (Fig. 1) trees, but was not supported within the MP tree (Fig. S1).

The Clade M was not well supported in the MP tree (Fig. S1). In the ML tree (Fig. S2) and BI (Fig. 1) tree, *C. latipetiolata* as the earliest differentiated taxon, formed a sister group with *C. furfuracea* and *C. octopetala* (PP = 100, MLBS = 80%, MPBS = 65%).

On comparison of the three phylogenies, there were three groups well supported in all topology trees (Fig. 1, S1, S2). One consisted of *C. liuzhouensis* and *C. longipedicellata* of sect. *Longipedicellata* (PP = 100, MLBS = 96%, MPBS = 91%), one made up of *C. longtousanica* and *C. taliensis* of sect. *Thea* (PP = 100, MLBS = 100%, MPBS = 97%), and one of *C. miyagii* and *C. tenii* of sect. *Paracamellia* (PP = 100, MLBS = 99%, MPBS = 86%) (Fig. 1, Fig. S1, S2). Within the ML and BI trees (Fig.1, S2), one group of *C. longicaudata* and *C. lungshenensis* from sect. *Camellia* was well supported (PP = 99, MLBS = 42%). *C.fluviatilis* and *C. yuhsienensis* of sect. *Paracamellia* clustered into a group that was weakly supported in ML and BI (PP = 80, MLBS = 39%) (Fig.1, S2). Meanwhile, the remaining species, including *C. albogigas*, *C. granthamiana*, *C. vietnamensis*, *C. nitidissima* var. *microcarpa*, *C. pubifurfuracea*, *C. ptilophylla*, *C. euphlebia*, *C. nitidissima*, *C. tetracocca, C. confuse*, *C. gauchowensis* and *Camellia sp*., formed a sole branch individually, with poor phylogenetic relationship and position in the topology trees (Fig. 1, S1, S2).

Comparison of our classification with Chang's classification system

In Fig. 1, some differences in intragenus phylogenetic relationship and position of *Camellia* were observed, in comparison with Chang's classification system (Chang, 1981, 1998), which was presented using different color stripe. Most species of sect. *Camellia* in Chang's system were mainly distributed in Clade A and B, whereas the remaining species scattered into some groups without support. Sect. *Theopsis* was mainly distributed in two strongly supported Clades C and I. Clade K and L mainly encompassed the species of sect. *Chrysantha*.

Many species of sect. *Thea* mainly distributed in Clades D and J, whilst the remaining species gathered into other groups, such as *C. fangchengensis* in Clade H, and *C. longtousanica* and *C. taliensis* clustered into a separate group.

Species of sect. *Paracamellia* gathered into different groups. Among these groups, *C. odorata* is in Clade A, *C. grijsii*, *C. shensiensis* and *C. rosthorniana* of sect. *Theopsis* and *C. yunnanensis* of sect. *Stereocarpus* formed Clade E, whereas other four species *C. brevistyla*, *C. hiemalis*, *C. phaeoclada*, *C. puniceiflora* and *C. sasanqua* of sect. *Oleifera* were well supported into Clade F, and the remaining two species *C. fluviatilis* and *C. yuhsienensis* clustered into another group.

Unexpectedly, two species *C. albogigas* and *C. granthamiana* of sect. *Archecamellia* did not cluster with any species from the thirteen Clades. Among the four species of sect. *Pseudocamellia*, three species: *C. chungkingensis*, *C. szechuanensis*, and *C. trichocarpa*, were in Clade H, and clustered together with most species of sect. *Tuberculata*. The remaining one species, *C. ilicifolia*, was sister to Clade J, but not supported in the MP analyses (Fig. S1).

Sect. *Furfuracea* formed two Clades G and M. However, the species *C. pubifurfuracea* did not cluster in anyone of the thirteen Clades. Species from Sect. *Tuberculata* mainly clustered into Clade H. Whereas two species: *C. albogigas* and *C. granthamiana*, did not cluster within the 13 clades. The species from sect. *Eriandria* were mostly presented in Clades C and I, which surprisingly encompassed the species of sect. *Theopsis.* Two species *C. danzaiensis* and *C. kwangtungensis* of sect. *Glaberrima* were presented in Clade J. Whilst *C. sasanqua* of sect. *Oleifera* clustered with sect. *Paracamellia* into Clade F, and other four species of sect. *Oleifera* did not cluster into a well-supported group (Fig.1). The two species from sect. *Longipedicellata*, *C. liuzhouensis* and *C. longipedicellata*, clustered into a group.

DISCUSSION

Application of nuclear genes in taxonomy of *Camellia*

Although the ITS sequence analysis is not ideal for the phylogenetic inference of *Camellia*, genetic markers in other regions of the genome might still be suitable for research on the evolution of *Camellia*. Recently, Zhao *et al*. made a taxonomically comprehensive phylogenetic analysis of 99 representative samples of *Camellia* using three orthologous nuclear DNA regions. It was suggested that orthologous nuclear *RPB2* (introns11–15 and 23) and *waxy* regions can be used for phylogenetic analysis of the genus *Camellia*, while ITS is more appropriate for the analysis of gene evolution rather than phylogenetic inference for the genus. The genomic region surrounding *LEAFY* gene that bears non-homologous copies should be interpreted with great caution for taxonomic significance in the family Theaceae (Zhao *et al*., 2022). A single-copy nuclear gene called phenylalanine ammonia-lyase (*PAL*) is also used to study the population genetic structure and phylogeography of *Camellia flavida* (Wei *et al*., 2017). Chen *et al* also used two nuclear genes (*PAL* and *waxy*) to infer species relationships among *Camellia chrysanthoides* and its closely related species (Chen *et al*., 2021). Therefore, some regions of nuclear DNA, such as *PAL*, *waxy*, and *RPB2*, play an important role in the study of species phylogeny and should be paid attention to.

Sect. *Pseudocamellia* and sect. *Tuberculata*

Sealy (1958) set up sect. *Pseudocamellia*, which including only two species: *C. szechuanensis* and *C. tuberculata*. Then Chang (1981) removed *C. tuberculata* out of sect. *Pseudocamellia* and established sect. *Tuberculata.* Our result showed that *C. szechuanensis* and *C. tuberculata* clustered into a group. These two species were further differentiated into sister groups in BI and ML trees.

Some species of sect. *Pseudocamellia* had been proposed to be merged into sect. *Tuberculata* (Ming and Zhong, 1993). For example, Ming and Zhong (1993) suggested that *C. chungkingensis* and *C. ilicifolia* of sect. *Pseudocamellia* is placed into the sect. *Tuberculata*. They suggested that sect. *Tuberculata* may be regarded as the geographical vicarious taxon of sect. *Pseudocamellia*. Vijayan *et al*. (2009) also found that *C. chungkingensis* and *C. szechuanensis* of sect. *Pseudocamellia* and five species of sect. *Tuberculata* clustered into a group, despite a polyphyletic relationship between these two sections. According to criteria in classic taxonomy, the species from these two sections shared high similarity in their flower structure and capsule characteristic. Their major differentiation is only based on the pericarp feature. Our previous results also showed that the species from these two sections display similar leaf architecture and were hence clustered into one group (Lu *et al*., 2012). Based on the combined chloroplast sequences, *C. chungkingensis*, *C. szechuanensis* and *C. trichocarpa* gathered into a group with the sect. *Tuberculata*. Whereas, *C. ilicifolia* shared a closer relationship with the sect. *Thea*, which was well supported in the ML and BI trees.

Sect. *Glaberrima* and sect. *Thea*

In Ming's classification of the genus *Camellia*, sect. *Glaberrima* was merged into the sect. *Thea* (Ming, 1999). Our previous quantitative and qualitative leaf anatomy characterization also indicated a close phylogenetic relationship between the two sections (Lu *et al*., 2008). However, Chang (1996) suggest that they should still remain in two sections since there were significant morphological differentiation in flowers, fruits, leaves and branches between the two sections. The results of this study strongly supported that the species of sect. *Glaberrima* clustered with the species of sect. *Thea*, in agreement with Ming's classification and our earlier opinion.

Sect. *Furfuracea* and sect. *Camellia*

Based on their morphological characteristics, such as large flowers, terminal, white, sessile, undifferentiated bracts, caduceus and furfuraceous capsule, Chang (1981) constructed the sect. *Furfuracea*. Nevertheless, Ming (2000) thought that furfuraceous capsule is a characteristic also shared by other sections. Therefore, this section was subsequently removed into sect. *Heterogenea* in Ming's system (Ming, 2000). Whereas sect. *Camellia* was treated as a separate taxon in both classification systems.

Our results showed that sect. *Furfuracea* was not monophyletic and nested with taxa of many sections, such as sect. *Camellia*, sect. *Paracamellia* and sect. *Oleifera*, etc. However, some species of sect. *Furfuracea* could be well grouped into one branch and clustered on a large branch with some species of sect. *Camellia*, which indicated these two sections had a close phylogenetic relationship. Moreover, a strongly supported group was formed by *C. weiningensis* and some species of sect. *Camellia*, which supported the view of Chang (1998) who placed this species into sect. *Camellia*.

Sect. *Oleifera* and Sect. *Paracamellia*

The sect. *Paracamellia* was firstly proposed by Sealy in 1958, and the sect. *Oleifera* was firstly established by Chang in 1981. Ming (1999) pointed that sect. *Oleifera* and sect. *Paracamellia* have many similar morphological features, for example, axillary or subterminal flowers, short stamens, 1/2 as long as petals, short style, and no clear differentiation between these two sections. Meanwhile, Chang (1981, 1998) reported that sect. *Oleifera* had style longer than 1cm, larger flowers and larger capsule, while sect. *Paracamellia* species had style shorter than 1cm, smaller flowers and smaller capsule. The sect. *Oleifera* and sect. *Paracamellia* should be treated as two different groups owing to significantly different characteristics. According to their leaf

morphological and anatomy characteristics, the significant differentiation of leaf structure was observed between these two sections (Lin *et al*., 2008). Using FTIR (Fourier Transform Infrared spectroscopy), the spectra feature of sect. *Oleifera* was different with sect. *Paracamellia* (Shen *et al*., 2008). The results based on three orthologous nuclear DNA regions showed that *C. grijsii* and *C. shensiensis* were nested in the clade *Camellia* II and not closely related with other oil camellias that form the clade *Paracamellia* (Zhao *et al*., 2022).

In Figure 1, *C. sasanqua* of sect. *Oleifera* and four species of sect. *Paracamellia* formed a strongly supported group. Other two species, *C. lanceoleosa* and *C. oleifera*, were presented in the group with other species of sect. *Paracamellia*, though it was weakly supported. The position of the remaining two species, *C. vietnamensis* and *C. gauchowensis*, could not be identified in this study. The phylogenetic relationship of sect. *Oleifera* and sect. *Paracamellia* therefore needs to be further explored.

Sect. *Chrysantha* and sect. *Longipedicellata*

Currently, the phylogenetic relationship among species of sect. *Chrysantha* is controversial. Some researchers showed that the species among sect. *Chrysantha* had a close phylogenetic relationship (Shi *et al*., 1998; Xiao *et al*., 2002; Tang *et al*., 2004). However, based on four chloroplast gene sequences, the species formed three paraphyletic groups that indicated sect. *Chrysantha* might be a paraphyletic or polyphyletic group (Fang *et al*., 2010). The results of our study support the view of Fang *et al* (2010).

C. longipedicellata from sect. *Longipedicellata* clustered together with some species of sect. *Chrysantha* distributed in Vietnam, which were proposed to be closely related to sect. *Longipedicellata* (Fang *et al*., 2010). Contrary to the view of Fang *et al*. (2010), *C. longipedicellata* and *C. liuzhouensis* from sect. *Longipedicellata* clustered into a strongly supported group. *C. longipedicellata* not gathered with sect. *Chrysantha* and was paraphyletic to the sect. *Chrysantha*. Therefore, our results indicate sect. *Longipedicellata* is a separate taxon.

Sect. *Longissima*

Sect. *Longissima*, which including three species, was firstly set up by Chang (1981). In Ming's classification system, Sect. *Longissim* was merged into sect. *Longipedicellata* (Ming, 1999, 2000). Fang *et al*. (2010) found that sect. *Longissima* was not a monophyletic group. On the one hand, *C. hekouensis* sister to the other species of *Camellia* genus indicated that sect. *Longissima* may be a primitive group of *Camellia*. On the other hand, *C. longissima* had no relationship with *C. hekouensis*, but was related to sect. *Furfuracea* and sect. *Oleifera* (Fang *et al*., 2010). Within our phylogenetic analyses, *C. hekouensis* formed a strongly supported group with sect. *Thea*. This result was contrary to the view of Fang *et al*. (2010). Thus, it was suggested that *C. hekouensis* should be merged into sect. *Thea*.

Identification of other species of *Camellia*

It has always been controversial if *C. hiemalis* should be treated as a varietas of *C. sasanqua* from sect. *Oleifera*. In Chang's classification system (Chang, 1981), *C. hiemalis* was treated as an independent species. However, in our study, *C. hiemalis* clustered with *C. sasanqua* into a group which was highly supported, and it may be reasonable that *C. hiemalis* is treated as a varietas of *C. sasanqua*.

On the other hand, *C. yuhsienensis* of sect. *Oleifera* and *C. fluviatilis* of sect. *Paracamellia* clustered into a well-supported group, which shared a distant relationship with *C. lanceoleosa* from the same section.

Three species of sect. *Furfuracea*, *C. furfuracea*, *C. oblata* and *C. parafurfuracea*, were merged into *C. furfuracea* by Ming (1999, 2000). Meanwhile, in our analyses, *C. furfuracea* and *C. oblata* were in two different highly supported groups and did not support the merger of them.

It was worth noting that *C. japonica* and some species of the sect. *Glaberrima* and sect. *Thea* formed a well-supported group, which indicated a closely polyphyletic relationship among them.

In the ML and BI phylogenetic trees, both showed that *C. obovatifolia* and *C. anlungensis* were in different sister groups under the same subclade, which did not support the merger of them by Ming's system (Ming, 1999; 2000).

In Ming's classification system (Ming, 1999), *C. gigantocarpa*, *C. multibracteata* and *C. octopetala* were merged into *C. crapnelliana*. However, in our analysis, *C. gigantocarpa* clustered with *C. crapnelliana* into a strongly supported group, whereas *C. multibracteata* and *C. gaudichaudii* gathered into a group, and *C. octopetala* and *C. furfuracea* formed a branch. These results suggest that *C. gigantocarpa* could merge into *C. crapnelliana*, but the other species should not be merged into *C. crapnelliana*. Similarly, *C. angustifolia* and *C. parvisepala* formed a separate group in ML and BI analyses, which did not support the merger of them by Ming's system. They should be treated as different species.

ACKNOWLEDGMENTS

The work was supported by Zhejiang Provincial Top Key Discipline of Biology, Zhejiang Provincial Universities Key Discipline of Botany and National Natural Science Foundation of China (31970358). Authors on the manuscript declare no conflict of interest. We are very grateful to two anonymous reviewers for their review of this article, and they have put forward many valuable suggestions, which are of great help to the improvement of the quality of this paper. Thanks very much for their careful work.

LITERATURE CITED

Badr, A., N. El-Sherif, S. Aly, S.D. Ibrahim and M. Ibrahim 2020. Genetic diversity among selected *Medicago sativa* cultivars using inter-retrotransposon-amplified polymorphism, chloroplast DNA barcodes and morpho-agronomic trait analyses. Plants-Basel **9(8)**: 995.

Cabrera, C., R. Artacho and R. Giménez 2006. Beneficial effects of green tea - a review. J. Am. Coll. Nutr. **25(2)**: 79–99.

Chang, H.T. 1981. A taxonomy of the genus *Camellia*. J. Sun. Yatsen. Univ. **1**: 1–180.

Chang, H.T. 1996. Diagnosis on the systematic development of Theaceae - the taxonomic problem of section *Glaberrima* Chang of *Camellia*. Acta Sci. Nat. Univ. Sunyatseni **35**: 87–90.

Chang, H.T. 1998. Theaceae. in: Delectis Florae Republicae Popularis Sinicae, Agendae Academiae Sinicae Edita, Tomus (ed.), Flora of Reipublicae Popularis Sinicae, vol. **49(3)**: 1–251. Science Press, Beijing.

Chen, H.L., S.L. Liu, Q.Q. Ye, Y.B. Lu and S.Q. Tang 2021. Inferring species relationships among *Camellia chrysanthoides* and its closely related species for developing conservation strategies. Plant Spec. Biol. **36(3)**: 390–398.

Chen, L., Z.X. Zhou and Y.J. Yang 2007. Genetic improvement and breeding of tea plant (*Camellia sinensis*) in China: from individual selection to hybridization and molecular breeding. Euphytica **154(1-2)**: 239–248.

Chung, M.Y., B.K. Epperson and M.G. Chung 2003. Genetic structure of age classes in *Camellia japonica* (Theaceae). Evolution **57(1)**: 62–73.

Cohen-Stuart, C.P. 1916. Voorbereidende onderzoekingen ten dienste van de selecte der theeplant. Meded. Proefst. Thee. **40**: 1–328.

Fang, W., J.B. Yang, S.X. Yang and D.Z. Li 2010. Phylogeny of *Camellia* sect. *Longipedicellata*, *Chrysantha*, and *Longissima* (Theaceae) based on sequence data of four choloroplast DNA loci. Acta. Bot. Yunnan. **32(1)**: 1–13.

Farris, J.S., M. Källersjö, A.G. Kluge and C. Bult 1995. Testing significance of incongruence. Cladistics **10(3)**: 315–319.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39(4)**: 783–791.

Feng, J., L.L. Jiang, J.L. Zhang, H. Zheng, Y.F. Sun, S.N. Chen, M.L. Yu, W. Hu, D.F. Shi, X.H. Sun and H.F. Lu 2020. Nondestructive determination of soluble solids content and pH in red bayberry (*Myrica rubra*) based on color space. J. Food Sci. Tech. Mys. **57(12)**: 4541–4550.

Gramza, A. and J. Korczak. 2005. Tea constituents (*Camellia sinensis* L.) as antioxidants in lipid systems. Trends Food Sci. Tech. **16(8)**: 351–358.

Huelsenbeck, J.P. and F. Ronquist 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17(8)**: 754–755.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Mentjies and A. Drummond 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28(12)**: 1647–1649.

Khan, N. and H. Mukhtar. 2007. Tea polyphenols for health promotion. Life Sci. **81(7)**: 519–533.

Kim, C., T. Deng, J. Wen, Z.L. Nie and H. Sun 2015. Systematics, biogeography, and character evolution of *Deutzia* (Hydrangeaceae) inferred from nuclear and chloroplast DNA sequences. Mol. Phylogenet. Evol. **87**: 91–104.

Li, X.L., Z.Q. Fan, H.B. Guo, N. Ye, T. Lyu, W. Yang, J. Wang, J.-T. Wang, B. Wu, J. Li and H. Yin 2018. Comparative genomics analysis reveals gene family expansion and changes of expression patterns associated with natural adaptations of flowering time and secondary metabolism in yellow *Camellia*. Funct. Integr. Genomic. **18(6)**: 659–671.

Lim, M.S. and S.H. Choi 2018. Estimation of phylogeny of nineteen sedoideae species cultivated in Korea inferred from chloroplast DNA analysis. Horticult. J. **87(1)**: 132–139.

Lin, X.Y., Q.F. Peng, H.F. Lu, Y.Q. Du and B.Y. Tang 2008. Leaf anatomy of *Camellia* sect. *Oleifera* and sect. *Paracamellia* (Theaceae) with reference to their taxonomic significance. J. Syst. Evol. **46(2)**: 183–193.

Linnaeus, C. 1753. *Species plantarum*. Vol I. Stockholm: Impensis Laurentii Salvii.

Lu, H.F., W. Jiang, M. Ghiassi, S. Lee and M. Nitin. 2012. Classification of *Camellia* (Theaceae) species using leaf architecture variations and pattern recognition techniques. PLoS One **7(1)**: e29704.

Lu, H.F., J.B. Shen, X.Y. Lin and J.L. Fu. 2008. Relevance of fourier transform infrared spectroscopy and leaf anatomy for species classification in *Camellia* (Theaceae). Taxon **57(4)**: 1274–1288.

Ly, N.S., V.D. Luong, T.H. Le, N.D. Do, N. Tran, A.D. Nguyen, T.L. Nguyen, C. Uematsu and H. Katayama 2022. *Camellia annamensis* (Theaceae), a new species from central Vietnam. Taiwania **67(2)**: 243–249.

Melchior, H. 1925. Theaceae. pp. 109–154 in: Engler, A. and Prantl, K. (eds.), Die Natürlichen Pflanzenfamilien. II, vol. 21. W. Engelmann, Leipzip, Germany.

Miller, M.A., W. Pfeiffer and T. Schwartz 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. pp. 1-8 in: Proceedings of the Gateway Computing Environments Workshop (GCE), vol. **14**. New Orleans.

Ming, T.L. 1999. A systematic synopsis of the genus *Camellia*. Acta Bot. Yunnan **21**: 149–159.

Ming, T.L. 2000. *Monograph of the genus Camellia.* Yunnan Science and Technology Press, Kunming.

Ming, T.L. and Y.C. Zhong 1993. A revision of genus *Camellia* sect. *Tuberculata*. Acta Bot.Yunnan **15**: 123–130.

Moller, E.M., G. Bahnweg and H. Sandermann. 1992. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res. **20(22)**: 6115–6116.

Nakai, T. 1940. A new classification of the Sino-Japanese genera and species which belong to the tribe Camellieae (II). J. Jap. Bot. **16**: 27–35.

Prihatini, I., A.Y.P.B.C. Widyatmoko, I.L.G. Nurtjahjaningsih, V. Yuskianti and S.A. Danarto 2020. DNA barcoding of *Pericopsis mooniana* from two different populations in Indonesia based on rDNA ITS (Internal Transcribed Spacer). IOP Conference Series: MSE **935(1)**: 012024.

Ricardo, N.M.J. Páscoa, A.M. Teixeira and C. Sousa. 2019. Antioxidant capacity of *Camellia japonic*a, cultivars

assessed by near- and mid-infrared spectroscopy. Planta **249(4)**: 1053–1062.

Ronquist, F. and J.P. Huelsenbeck 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19(12)**: 1572–1574.

Saha, D., S. Mandal and A. Saha 2012. Copper induced oxidative stress in tea (*Camellia sinensis*) leaves. J. Environ. Biol. **33**: 861–866.

Sealy, J.R. 1958. A revision of the genus *Camellia*. Royal Horticultural Society.

Shen, J.B., H.F. Lu, Q.F. Peng, J.F. Zheng and Y.M. Tian 2008. FTIR spectra of *Camellia* sect. *Oleifera*, sect. *ParaCamellia*, and sect. *Camellia* (Theaceae) with reference to their taxonomic significance. J. Syst. Evol. **46**: 194–204.

Shi, S.H., S.Q. Tang, Y.Q. Chen, L.H. Qu and H.T. Chang. 1998. Phylogenetic relationships among eleven yellow-flowered *Camellia* species based on random amplified polymorphic DNA. Acta. Phytotax. Sin. **36**: 314–322.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30(9)**: 1312–1313.

Taberlet, P., L. Gielly, H. Pautou and J. Bouvet 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. **17(5)**: 1105–1109.

Tang, S.Q., S.H. Shi, Y. Zhong and Y. Wang 2004. Phylogenetic relationships of golden camellias (sect. *Chrysantha*, *Camellia*) in China: evidence from ITS sequences of nuclear ribosomal DNA. Guihaia **24**: 488–492.

Vijayan, K., W.J. Zhang, and C.H. Tsou 2009. Molecular taxonomy of *Camellia* (Theaceae) inferred from nrITS sequences. Am. J. Bot. **96(7)**: 1348–1360.

Wang, B., L.M. Han, C. Chen and Z.Z. Wang 2016. The complete chloroplast genome sequence of *Dieffenbachia seguine* (Araceae). Mitochondrial DNA **27(4)**: 2913–2914.

Wang, Y.S., L.P. Gao, Z.R. Wang, Y.J. Liu, M.L. Sun, D.Q. Yang, C.L. Wei, Y. Shan and T. Xia 2012. Light-induced expression of genes involved in phenylpropanoid biosynthetic pathways in callus of tea (*Camellia sinensis* (L.) O. Kuntze). Sci. Hortic. **133**: 72–83.

Wei, S.J., Y.B. Lu, Q.Q. Ye and S.Q. Tang. 2017. Population genetic structure and phylogeography of *Camellia flavida* (Theaceae) based on chloroplast and nuclear DNA sequences. Front. Plant Sci. **8**:718.

Wei, Z.Z., L.Q. Guo, J.F. Zhang, B.L. Li, D.Q. Zhang and H. Guo 2010. Phylogenetic relationship of *Populus* by *trnL-F* sequence analysis. Journal of Beijing Forestry University **32**: 27–33.

Wilgenbusch, J.C., D.L. Warren and D.L. Swofford. 2004. AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Xiao, T.J. and C.R. Parks 2002. Molecular analysis of genus *Camellia*. American *Camellia* Society: American *Camellia* Year-book, 52–58.

Yamada, K.D., K. Tomii and K. Katoh 2016. Application of the MAFFT sequence alignment program to large data - reexamination of the usefulness of chained guide trees. Bioinformatics **32(21)**: 3246–3251.

Yang, J.B., H.T. Li, S.X. Yang, D.Z. Li and Y.Y. Yang 2006. The application of four DNA sequences to studying molecular phylogeny of *Camellia* (Theaceae). Acta Bot. Yunnan **28**: 108–114.

Yang, J.B., J. Yang, H.T. Li, Y. Zhao and S.X. Yang 2009. Isolation and characterization of 15 microsatellite markers from wild tea plant (*Camellia taliensis*) using FIASCO method. Conserv. Genet. **10(5)**: 1621–1623.

Yang, H., C.L. Wei, H.W. Liu, J.L. Wu, Z.G. Li, L. Zhang, J.B. Jian, Y.Y. Li, Y.L. Tai, J. Zhang, Z.Z. Zhang, C.J. Jiang, T. Xia and X.C. Wan 2016. Genetic divergence between *Camellia sinensis* and its wild relatives revealed via genome-wide SNPs from RAD sequencing. PLoS One **11(3)**: e0151424.

Yang, J.B., S.X. Yang, H.T. Li, J. Yang and D.Z. Li 2014. Comparative chloroplast genomes of *Camellia* species. PLoS One **8(8)**: e73053.

Yu, X.Q., E.D. Liu, Z.W. Liu, B.O. Xiao and S.X. Yang 2021. *Camellia luteocalpandria* (Theaceae), a new species and the first discovery of sect. *Calpandria* in China. Phytotaxa **489(2)**: 223–228.

Zhang, Q., Y. Liu and Sodmergen 2003. Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. Plant Cell Physiol. **44(9)**: 941–951.

Zhang, W., S.L. Kan, H. Zhao, Z.Y. Li and X.Q. Wang 2014. Molecular phylogeny of tribe Theeae (Theaceae *s.s.*) and its implications for generic delimitation. PLoS One **9(5)**: e98133.

Zhao, D.W., T.R. Hodkinson and J.A.N. Parnell 2022. Phylogenetics of global *Camellia* (Theaceae) based on three nuclear regions and its implications for systematics and evolutionary history. J. Syst. Evol. : jse 12837.

Zhu, H., X.G. Yi, S.X. Zhu, H.C. Wang, Y.F. Duan and X.R. Wang. 2018. Analysis on relationship and taxonomic status of some species in subg.*Cerasus* koehne with chloroplast DNA *atp*B-*rbc*L fragment. Bull. Bot. Res. **38(1)**: 820–827.

Zou, J.B., X. Wang, L.L. Li and J.Q. Liu 2013. The combination of chloroplast and nuclear DNA fragments could improve the discrimination power for barcoding spruce (*Picea*) species. Plant Divers. Resour. **35**: 375–385.

Supplementary materials are available from Journal Website.