Yu Song[a], Xin Yao[a], Bing Liu, Yunhong Tan* and Richard T. Corlett*

Complete plastid genome sequences of three tropical *Alseodaphne* trees in the family Lauraceae

https://doi.org/10.1515/hf-2017-0065
Received May 6, 2017; accepted November 8, 2017; previously published online December 7, 2017

Abstract: *Alseodaphne* is a genus of timber trees (ca. 40 spp.) belonging to the *Persea* group of the Lauraceae. It is widely distributed in tropical Asia, but is often confused with *Dehaasia* and *Nothaphoebe*, and the systematics of the genus is unclear. Here, the complete chloroplast genome sequences of *A. semecarpifolia* will be reported, the type species of *Alseodaphne*, and two China-endemic species, *A. gracilis* and *A. huanglianshanensis*. The three plastomes were 153 051 bp, 153 099 bp and 153 070 bp, respectively. Comparative genomic analyses indicate that the three *Alseodaphne* plastomes have similar genome size and those are very different with previously published plastomes of Lauraceae in length. The length difference is directly caused by inverted repeats expansion/contraction. Four highly variable loci including *psbD-trnM*, *ndhF-rpl32*, *rpl32-trnL* and *ycf1* among the three *Alseodaphne* species were identified as useful plastid candidate barcodes for *Alseodaphne* and Lauraceae species. Phylogenetic analyses based on 12 complete plastomes of Lauraceae species confirm a monophyletic *Persea* group comprising species of *Alseodaphne*, *Phoebe*, *Persea* and *Machilus*.

Keywords: *Alseodaphne*, candidate barcode, chloroplast genome, comparative genomics, Lauraceae, phylogenetic relationship

[a]**Yu Song and Xin Yao:** These authors contributed equally to this work.
***Corresponding authors: Yunhong Tan and Richard T. Corlett,** Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China; and Southeast Asia Biodiversity Research Institute, Chinese Academy of Science, Yezin, Nay Pyi Taw, Myanmar,
e-mail: tyh@xtbg.org.cn (Y. Tan); corlett@xtbg.org.cn (R.T. Corlett)
Yu Song and Xin Yao: Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China; and Southeast Asia Biodiversity Research Institute, Chinese Academy of Science, Yezin, Nay Pyi Taw, Myanmar
Bing Liu: State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

Introduction

The family Lauraceae is significant in tropical and subtropical forests worldwide and many species are harvested for their timber, particularly in East Asia. Lauraceae timbers have a wide range of properties from soft and low-density species to the exceptionally hard and dense timbers, such as the Borneo ironwood (*Eusideroxylon zwageri* Teijsm. & Binn.), which is one of the most durable woods in the world. In China, some species have specific local uses, for furniture, farm tools, or musical instruments (Wei and Werff 2008). Because of difficulties of identifying individual species, many are traded together despite their different properties. Taxonomy above the species level is also problematic in the Lauraceae and it is probably the least well understood of the major tree families. The advancement of molecular sequence identification, particularly from chloroplasts, facilitates to establish clear relationships within the family at the genus level and above (Song et al. 2017a). An improved phylogeny will allow the rapid assignment of potential timber trees to species groups with similar wood properties. This could potentially be done in the field, logging camp, or sawmill, using small, portable, sequencing equipment (Parker et al. 2017). DNA based wood identification and classification gained a lot of importance in the last two decades, especially in the field of combating illegal logging and the protection of endangered species and for differentiation of species, which are anatomically very similar (Michael et al. 2012; Jiao et al. 2014, 2015; Sandak et al. 2015; Yu et al. 2016; Chano et al. 2017; Hung et al. 2017).

The genus *Alseodaphne* Nees is confined to tropical and subtropical Asia (www.theplantlist.org), with 10 of the 40 known species in China (http://foc.eflora.cn/). *Alseodaphne semecarpifolia* Nees, a valuable timber tree from SW India, is the type species. *Alseodaphne* is part of the *Persea* group of the Lauraceae and appears to be polyphyletic as currently constituted, although the precise generic boundaries and species delineations based on morphological characters are currently unclear (Julia et al. 2009; Rohwer et al. 2009; Li et al. 2011). Several species of *Alseodaphne* are known for their high wood quality, and others, such as *A. andersonii* (King ex Hook.f.) Kosterm., *A. perakensis* (Gamble) Kosterm., *A. semecarpifolia* Nees, and *A. corneri* Kosterm, have been reported to contain potentially useful alkaloids with antiplasmodial

and antioxidant activities (Lee et al. 2001; Nafiah et al. 2011; Thakur et al. 2012; Nafiah et al. 2016; Zahari et al. 2016). In order to distinguish the *Alseodaphne* species and reconstruct their phylogenetic relationships, a molecular approach is necessary.

Using the chloroplast marker *matK* or the intron of *trnK*, previous studies suggested that *A. perakensis* and *Dehaasia cuneata* (Blume) Blume, along with *Apollonias barbujana* (Cav.) Bornm., *Persea americana* Mill. and *Phoebe formosana* (Hayata) Hayata, form a weakly supported clade (the "*Persea* group"), with *Persea indica* (L.) Spreng. and *Persea lingue* (Miers ex Bertero) Nees (Rohwer 2000; Rohwer and Rudolph 2005). Another phylogenetic study, based on different combinations of chloroplast sequences (*trnL-trnF*, *psbA-trnH*, *trnT-trnL* and *rpl16*) and nuclear barcoding markers (26S rDNA and ITS rDNA), showed a well-supported *Persea* group comprising *A. semecarpifolia*, *P. formosana*, *Dehaasia incrassata* (Jack) Kosterm., *Machilus thunbergii* Siebold & Zucc. (as *P. thunbergii*), and *A. barbujana*, along with four other *Persea* species (Chanderbali et al. 2001).

Recently, the nuclear barcoding marker *ITS* was used to construct a Bayesian tree in which *Alseodaphne* species were included in four different clades (Rohwer et al. 2009). Another study, based on two nuclear markers, *ITS* and *LEAFY* intron, also confirmed that "*Alseodaphne*" is polyphyletic (Li et al. 2011). Both studies found that *Alseodaphne* species were mixed with species of *Persea*, *Nothaphoebe* Bl. and *Dehaasia* Bl. in the phylogenetic trees, and that these species were divided into two major clades, indicating that *Alseodaphne* species do not form a natural group. The *Alseodaphne-Persea* clade included *A. rugosa* Merr. & Chun, *A. hainanensis* Merr., *A. andersonii*, *A. petiolaris* Hook.f. and another unidentified *Alseodaphne* species, while the *Alseodaphne-Dehaasia* clade included *A. semecarpifolia*, *A. huanglianshanensis*, *A. gigaphylla* Kosterm and another unidentified *Alseodaphne* species. Thus both the generic delimitation and species relationships within the *Persea* group remain unclear. For a better understanding of the relationships among these species, it is necessary to reconstruct a robust phylogenetic tree of the *Persea* group based on high-throughput sequencing approaches.

The *Persea* group includes over 400 species, distributed in tropical and subtropical Asia, America, and the Macaronesian Islands (Li et al. 2011). So far, completed plastid genome sequences are available for only five species, *Machilus balansae* (Airy Shaw) F.N. Wei & S.C. Tang, *Machilus yunnanensis* Lecomte, *Phoebe sheareri* (Hemsl.) Gamble, *Phoebe omeiensis* R.H. Miao and *P. americana* (Song et al. 2015, 2016, 2017b). Here, three species of *Alseodaphne* were selected to sequence the complete chloroplast genomes. *Alseodaphne huanglianshanensis* H. W. Li & Y. M. Shui and *A. gracilis* Kosterm are endemic to tropical broad-leaved forests in Yunnan, SW China (Wei and Werff 2008), while *A. semecarpifolia* Nees occurs in tropical broad-leaved forests of peninsular India and Sri Lanka (Verma et al. 2016). Entire chloroplast genomes were then used to perform comparative analyses of *Alseodaphne* and a phylogenomic study with other Lauraceae.

Materials and methods

DNA extraction and sequencing: Silica-gel dried leaves from *A. huanglianshanensis*, *A. gracilis* and *A. semecarpifolia* were collected (Table 1). From ca. 6 cm² sections of dry leaf, genomic DNA was multiply extracted by a modified CTAB method (Yang et al. 2014). A total of 0.5 μg of each purified total DNA were directly sequenced. Fragmented to ca. 500 bp, each DNA sample was used to construct short-insert libraries according to the manufacturer's instructions (Illumina, San Diego, CA, USA). The DNA samples were indexed by tags and pooled together in one lane of a Genome Analyzer (Illumina Hiseq 2000) for sequencing at BGI-Shenzhen, and more than 2 Gb of sequence data for each *Alseodaphne* sample was obtained.

Plastome assembly and annotation: Stringent sequence filtering was applied with the NGS QC Tool Kit to select clean reads (Patel and Jain 2012) and assembled the high-quality short reads into long contigs with *de novo* assembling program CLC Genomics Workbench version 6.5 (QIAGEN). The complete chloroplast genome of *Machilus balansae* (in a related genus in the *Persea* group) served as the reference genome for contig alignment. Then the Dual Organellar Genome Annotator (DOGMA) software was applied to annotate the genes encoding proteins, transfer RNA (tRNAs) and ribosomal RNAs (rRNAs) on the *Alseodaphne* plastomes (Wyman et al. 2004). The genome map of *A. huanglianshanensis*, *A. gracilis* and *A. semecarpifolia* was drawn by the OrganellarGenomeDRAW tool (OGDRAW) (Lohse et al. 2013).

Sequence alignment and sliding window analysis: From GenBank of NCBI, five plastome sequences of the *Persea* group were collected:

Table 1: Sampled species and their voucher specimens used in this study.

Name	Herbarium	Voucher	Geographic origin	Accession number in GenBank
Alseodaphne gracilis Kosterm	HITBC-BRG	SY34114	Honghe, Yunnan, China	MG407593
Alseodaphne huanglianshanensis H. W. Li & Y. M. Shui	HITBC-BRG	SY32761	Honghe, Yunnan, China	MG407594
Alseodaphne semecarpifolia Nees	HITBC-BRG	SY33349	Sulawesi, Indonesia	MG407595

M. balansae (accession No. KT348517), *M. yunnanensis* (accession No. KT348516), *P. americana* (accession No. KX437771), *P. omeiensis* (accession No. KX437772) and *P. sheareri* (accession No. KX437773). Along with the three assembled *Alseodaphne* plastome sequences, all of these sequences were aligned by Mauve software (Darling et al. 2004). After manual adjustment with Bioedit, the matrixes including three *Alseodaphne* plastomes and eight plastomes in the *Persea* group were used to assess the variability (Pi) in DnaSP version 5 software (Librado and Rozas 2009). The window length was set to 600 base pairs and the step size to 200 base pairs.

Phylogenetic analyses: The plastome sequences of *M. balansae*, *M. yunnanensis*, *P. americana*, *P. omeiensis*, *P. sheareri*, *Cinnamomum micranthum* (Hayata) Hayata (accession No. KR014245), *Litsea glutinosa* (Lour.) C.B. Rob. (accession No. KU382356), *Endiandra discolor* Benth. (accession No. KT588615), *Endiandra globosa* Maiden & Betche (accession No. KT588614) and *Calycanthus fertilis* Walt. (accession No. NC004993) were collected from GenBank. After alignment by means of the MAFFT software with the three plastome sequences of *Alseodaphne*, the 13 sequences were manually adjusted by the BioEdit software (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Phylogenetic relationships were reconstructed based on a maximum-likelihood (ML) method in the RAxML version 7.2.6 program (Tamura et al. 2011). One thousand bootstrap replicates were performed in each analysis to obtain the confidence support. The plastome of *C. fertilis* served as an out-group.

Results

Chloroplast genome sequencing and assembly

Illumina paired-end sequencing generated 2.54 Gb, 2.59 Gb and 1.24 Gb raw reads of 150 bp in length for *A. gracilis*, *A. semecarpifolia* and *A. huanglianshanensis*, respectively. The numbers of reads after trimming for quality were 16 963 796, 17 258 104 and 8 267 212, respectively. The *de novo* assembly generated 384 563 contigs with an N50 length of 441 bp for *Alseodaphne gracilis*, 372 748 contigs with an N50 length of 357 bp for *A. semecarpifolia*, and 162 025 contigs with an N50 length of 333 bp for *A. huanglianshanensis* (Table 2). All three *Alseodaphne* chloroplast genome sequences were deposited in GenBank (deposited in GenBank: MG407593, MG407595, and MG407594).

Chloroplast genome organization of the *Alseodaphne* taxa

The chloroplast genome of *A. huanglianshanensis*, with a length of 153 070 bp, is 19 bp larger than that of *A. semecarpifolia* and 29 bp smaller than that of *A. gracilis* (Figure 1). The G + C content is 39.1% in the three species. These *Alseodaphne* chloroplast genomes include a pair of inverted repeats (IRs) of 20 280 bp in *A. huanglianshanensis* and *A. semecarpifolia*, and 20 285 bp in *A. gracilis*, separated by a small single copy (SSC) region of 18 681 bp in *A. huanglianshanensis*, 18 670 bp in *A. semecarpifolia*, and 18 682 bp in *A. gracilis*, and a large single copy (LSC) region of 93 829 bp in *A. huanglianshanensis*, 93 821 bp in *A. semecarpifolia*, and 93 847 bp in *A. gracilis* (Table 2). In the three *Alseodaphne* plastid genomes, a total of 128 genes were detected, including 84 protein-coding genes, 36 transfer RNA (tRNA) genes, and eight ribosome (rRNA) genes (Figure 1). Among

Table 2: Summary of three complete plastomes of *Alseodaphne*.

	A. gracilis	*A. huangli-anshanensis*	*A. seme-carpifolia*
Clean reads	16 963 796	8 267 212	17 258 104
Average read length	150 bp	150 bp	150 bp
Number of contigs	384 563	162 025	372 748
Total length of contigs	164 528 738 bp	55 183 597 bp	136 821 292 bp
N50 length of contigs	441 bp	333 bp	357 bp
Total cpDNA size	153 099 bp	153 070 bp	153 051 bp
Length of LSC region	93 847 bp	93 829 bp	93 821 bp
Length of IR region	20 285 bp	20 280 bp	20 280 bp
Length of SSC region	18 682 bp	18 681 bp	18 670 bp
Total GC content (%)	39.1	39.1	39.1
LSC	37.8	37.9	37.9
IR	44.3	44.3	43.9
SSC	33.8	33.9	33.8
Total number of genes	128	128	128
protein encoding	84	84	84
tRNA	36	36	36
rRNA	8	8	8

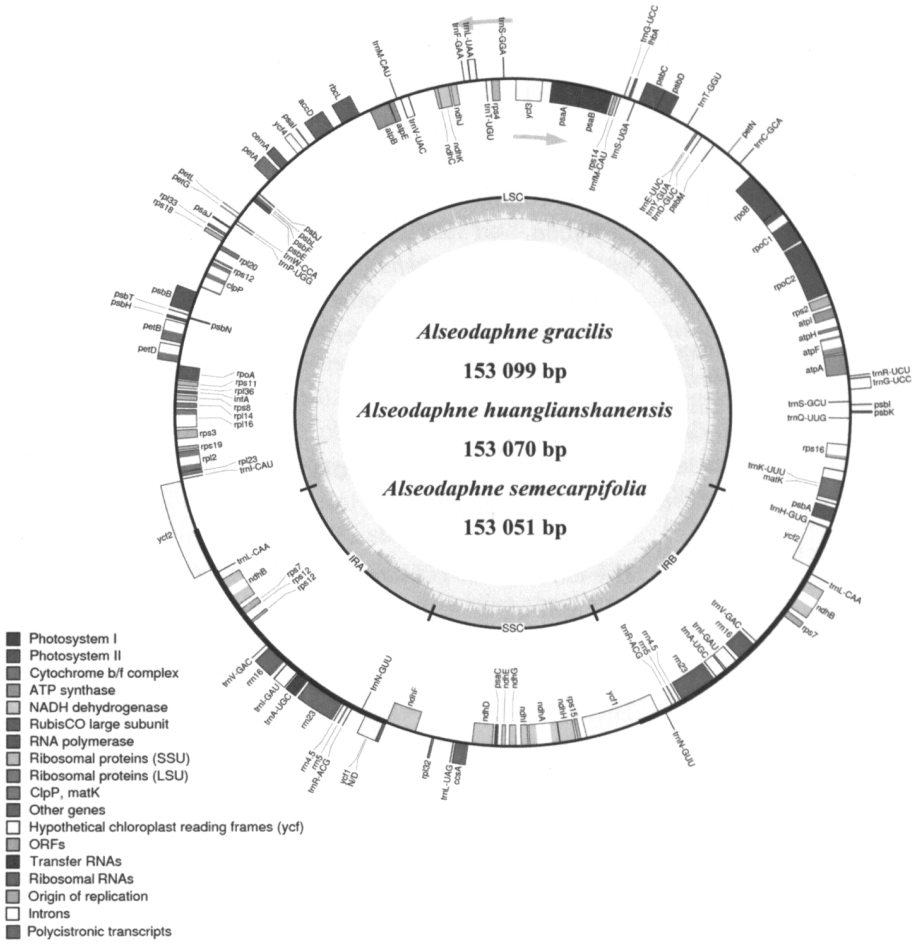

Figure 1: Gene map of *Alseodaphne* plastomes.
Genes drawn on the outside of the circle are transcribed clockwise, those on the inside counterclockwise. Small single copy (SSC), large single copy (LSC) and inverted repeats (IRa, IRb) are indicated.

these genes, five protein-coding genes, six transfer RNA (tRNA) genes, and four ribosome (rRNA) genes are duplicated in IRs (Figure 1).

Comparative analysis of chloroplast genomes

No rearrangement occurred in gene organization after verification (Figure 2). A significant degree of synteny was found within the three *Alseodaphne* chloroplast genomes, indicating that their genome structure was relatively conserved. The IR-LSC and IR-SSC boundary regions were compared across 12 sequenced Lauraceae taxa. Three thousand one hundred and sixty-two bp of 5'-*ycf2* and 1599 bp of 3'-*ycf1* are truncated at the boundaries of the IR regions in the three *Alseodaphne* plastomes (Figure 3). Double complete copies of the *ycf2* genes were detected in the two *Endiandra* plastomes, but only one complete copy and a broken fragment in the other ten sequenced Lauraceae plastomes. In contrast to *ycf2*, there is one complete *ycf1* gene and another *ycf1* fragment truncated at the boundary between the IR regions and the SSC region in all 12 sequenced chloroplast genomes (Figure 3).

Genome sequence divergence among the *Alseodaphne* taxa and related species

The level of sequence divergence among the chloroplast genomes of the three *Alseodaphne* taxa were elucidated including that and the other five species of the *Persea* group, *M. balansae*, *M. yunnanensis*, *P. americana*,

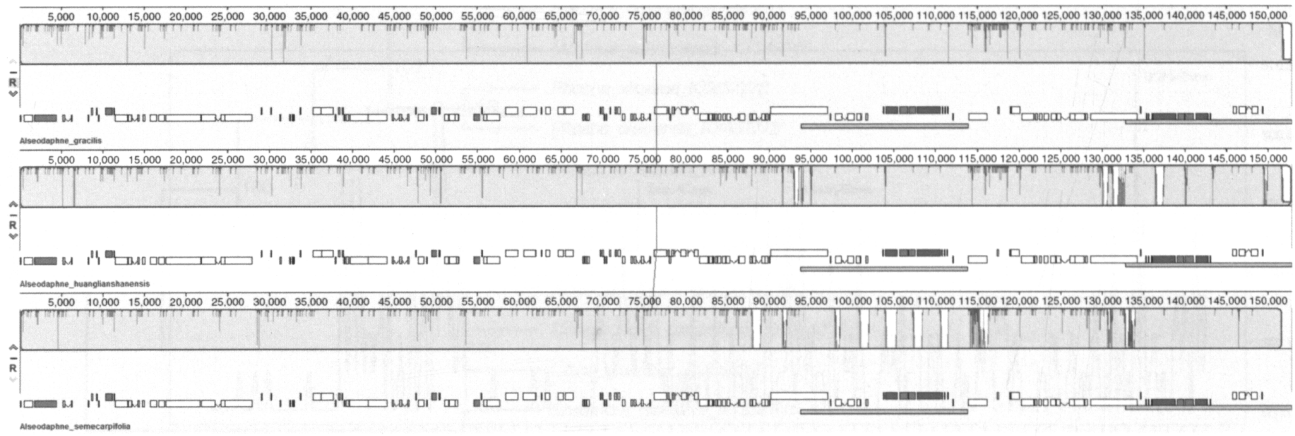

Figure 2: Synteny and rearrangements detected in *Alseodaphne* plastomes using the Mauve multiple-genome alignment program. The *Alseodaphne gracilis* genome is shown at top as the reference genome. Within each of the alignments, local collinear blocks are represented by blocks of connected lines with the same color.

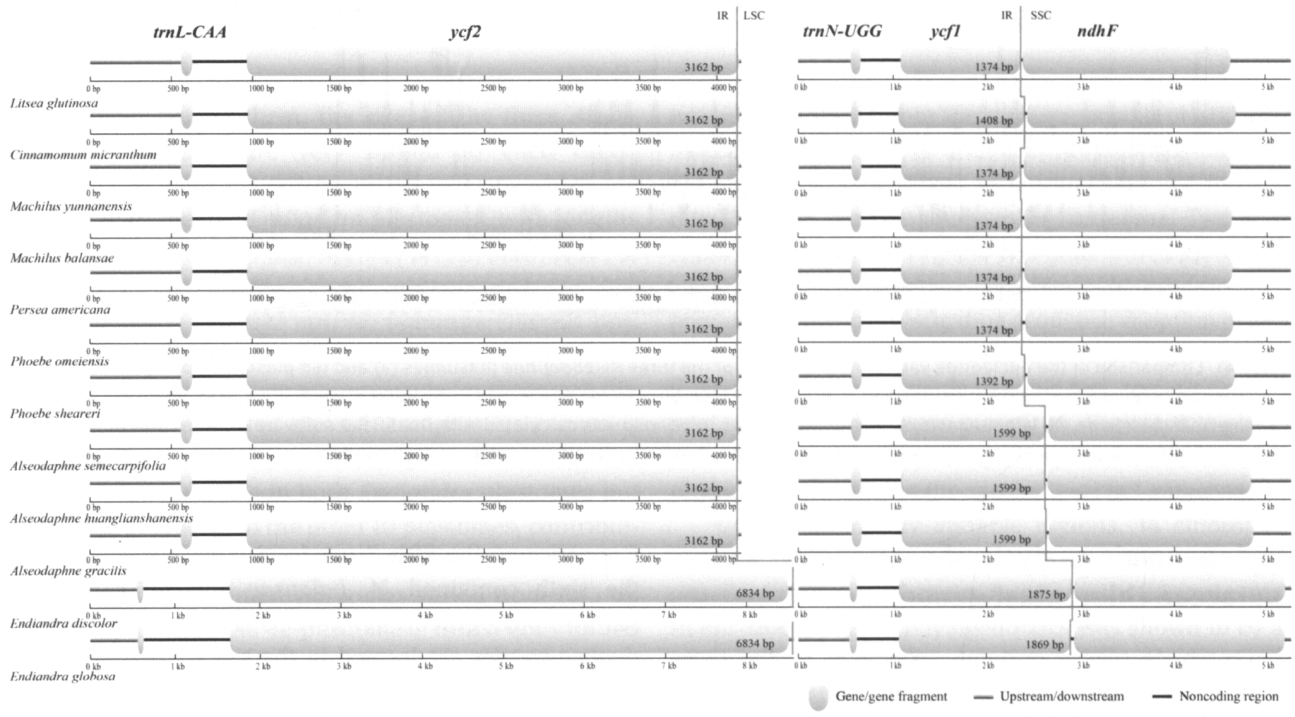

Figure 3: Comparison of LSC, IR and SSC junction positions among the three *Alseodaphne* plastomes.

P. omeiensis and *P. sheareri*. The nucleotide variability values within 600 bp in the three *Alseodaphne* plastomes varied from 0 to 0.0100, with a mean of 0.0012, and the values in the eight plastomes in the *Persea* group varied from 0 to 0.0143, with a mean of 0.0030, indicating high sequence similarity across the three *Alseodaphne* plastomes and the eight *Persea* group plastomes. However, nine highly variable loci (Pi > 0.006), including *trnG-UCC*, *accD-psaI*, *rps19-rpl3*, *ndhF-rpl32*, *rpl32-trnL*, and three regions of *ycf1*, were precisely located in the three *Alseodaphne* plastomes (Figure 4a). Among the eight Lauraceae plastomes, the five most dissimilar regions (Pi > 0.010) were *psbD-trnM*, *ndhF-rpl32*, *rpl32-trnL*, and two regions of *ycf1* (Figure 4b).

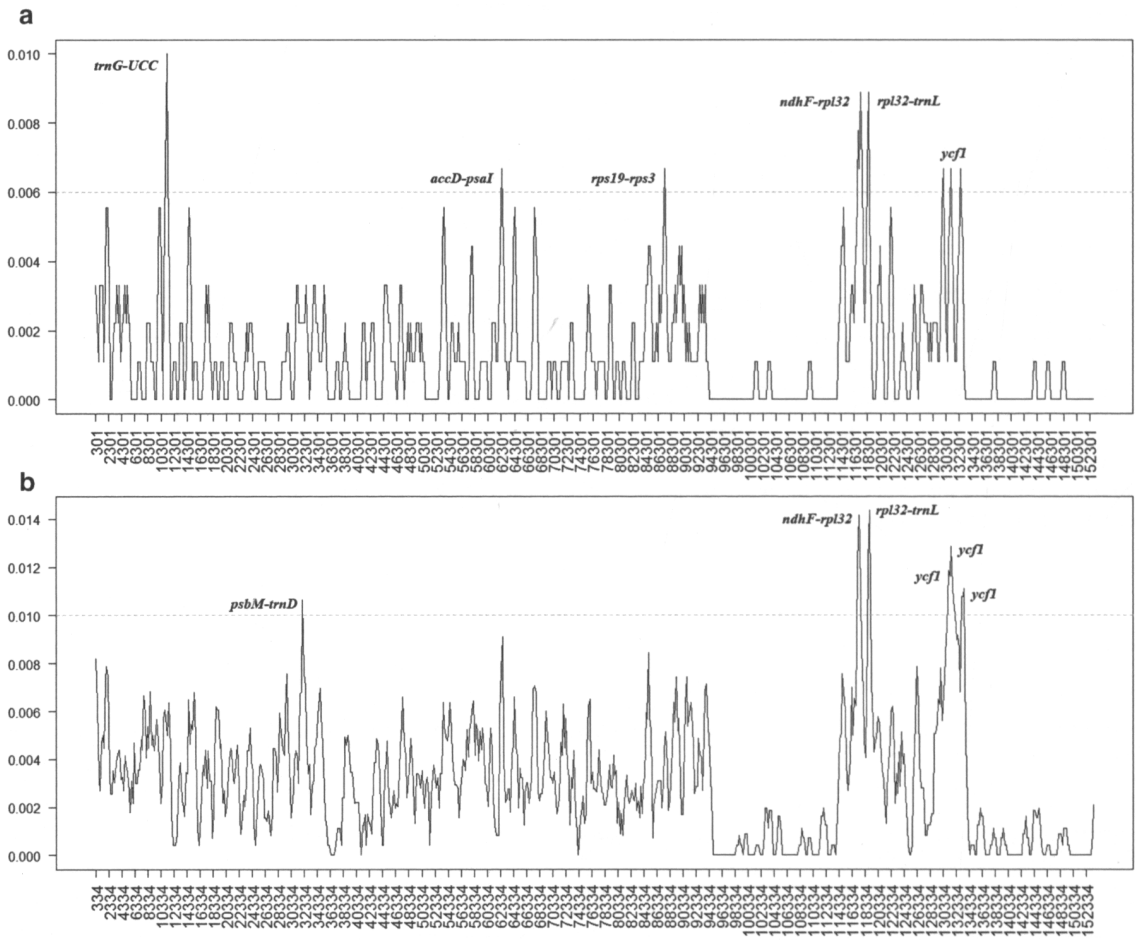

Figure 4: Comparison of the nucleotide variability (Pi) values of the three *Alseodaphne* plastomes (a) and ten plastomes of Lauraceae (b). The genetic divergence among the plastomes of *A. gracilis*, *A. huanglianshanensis*, *A. semecarpifolia*, *P. americana*, *P. omeiensis*, *P. sheareri*, *M. balansae*, *M. yunnanensis*, *C. micranthum* and *L. glutinosa* was calculated with DnaSP 5.0 software (window length: 600 bp, step size: 200 bp). X-axis: position of the midpoint of a window, Y-axis: nucleotide diversity of each window.

Phylogenomic analysis of sequenced Lauraceae chloroplast genomes

The complete plastid genome sequences of the three *Alseodaphne* species and the other nine sequenced Lauraceae taxa including *C. micranthum*, *M. balansae*, *M. yunnanensis*, *E. discolor*, *E. globosa*, *P. americana*, *P. omeiensis*, *P. sheareri* and *L. glutinosa*, formed the base to perform a phylogenetic analysis, with *C. fertilis* (Calycanthaceae, Laurales; GenBank accession No. NC004993) as outgroup. A maximum likelihood analysis yielded a tree topology with 100% bootstrap (BS) values at each node. The topology divided the 12 Lauraceae taxa into six clades (Figure 5). The first divergent clade contained *E. discolor* and *E. globosa*, the second *C. micranthum* and *L. glutinosa*, the third *A. gracilis*, *A. huanglianshanensis* and *A. semecarpifolia*, the fourth *P. americana*, the fifth *P. omeiensis* and *P. sheareri*, and the sixth contained *M. balansae* and *M. yunnanensis*. The phylogenetic tree reveals that the three *Alseodaphne* species are located in the same clade, which is closely related to *P. americana*, *P. omeiensis*, *P. sheareri*, *M. balansae* and *M. yunnanensis*.

Discussion and conclusions

This study presents three complete chloroplast genomes for woody plants in the genus *Alseodaphne* obtained by Illumina high-throughput sequencing technology (Figure 1). To date, there are now seven genera and 12 species, whose chloroplast genomes have been sequenced within the Lauraceae (Rossetto et al. 2015; Song et al. 2015; Song et al. 2016; Wu

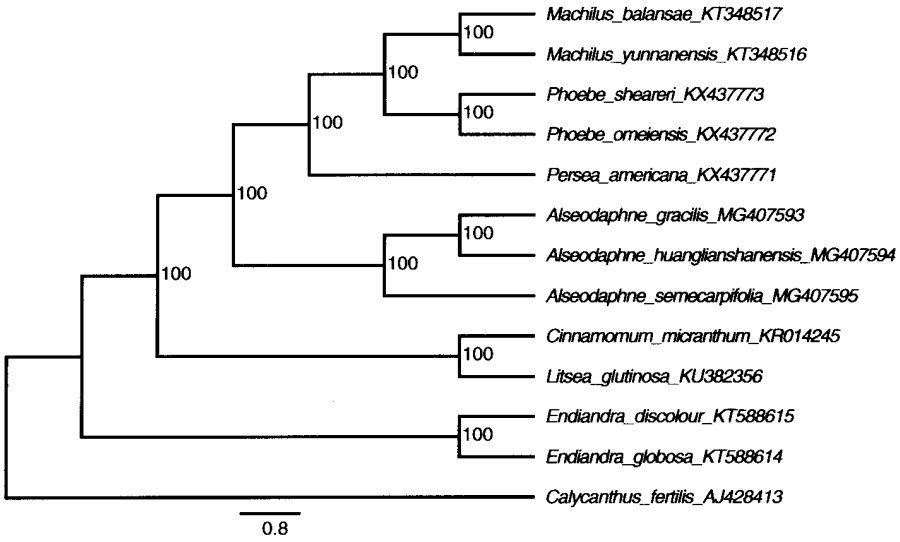

Figure 5: Molecular phylogenetic tree of 12 species of Lauraceae based on complete plastome sequences.
The phylogenetic tree was constructed from available chloroplast genome sequences of 12 species in the family Lauraceae by the maximum likelihood (ML) method with bootstrapping analysis (1000 reiterations) in the RAxML version 7.2.6 program. The tree is rooted with the plastome sequence of *Calycanthus fertilis*.

et al. 2016). Among them, the sizes of the three *Alseodaphne* plastomes are very similar to each other, with lengths from 153 051 bp to 153 099 bp, which are smaller than the two *Endiandra* plastomes, with lengths of 158 525 bp and 158 507 bp, but larger than those of *L. glutinosa* (152 618 bp), *C. micranthum* (152 700 bp), *M. yunnanensis* (152 622 bp), *M. balansae* (152 721 bp), *P. americana* (152 723 bp), *P. omeiensis* (152 855 bp) and *P. sheareri* (152 876 bp). Two main reasons for these size differences were identified through comparative genomics analysis. First, one copy of the *ycf2* gene with a length of 6834 bp was complete in both *Endiandra* plastomes but truncated to 3162 bp in the other 10 plastomes, contributing almost 3700 bp to the length difference. Second, the lengths of the truncated *ycf1* genes varied from 1875 bp to 1374 bp among the 12 plastomes, which contributed around 250 bp to the length difference. The two genes *ycf1* and *ycf2* are located at the boundaries between the IR regions and the LSC or SSC regions (Figure 3), and the length mutations are responsible for the contraction of the IR regions in the plastomes.

Compared with *ycf2*, *ycf1* has higher genetic divergence among the eight sequenced plastomes of the *Persea* group (Figure 4). Three highly variable loci were also identified in the complete *ycf1* gene sequences. With *ndhF-rpl32* and *rpl32-trnL*, the five highly variable loci showed promising levels of variation for further development in applications such as DNA-barcoding and phylogenetic studies at the species level in Lauraceae. These results are partially congruent with those of Dong et al. (2015), who reported that two regions of *ycf1* were the most promising plastid DNA barcodes for land plants, and those of Hinsinger and Strijk (2016), who suggested that *rpl32-trnL* should be incorporated in DNA barcode analyses in Lauraceae. Two other regions, *ndhF-rpl32* and *psbD-trnM*, identified among the eight sequenced Lauraceae plastomes, were the focus of previous analyses investigating sequence variation in seed plants (Shaw et al. 2007; Korotkova et al. 2014). In addition, three rarely reported highly variable loci, *trnG*-UCC, *accD-psaI* and *rps19-rps3*, were present in *Alseodaphne* plastomes. All of these highly variable regions will be useful for phylogenetic studies at the species level in Lauraceae.

Unlike our highly variable regions, 14 chloroplast genomic markers previously failed to resolve the phylogenetic problems within the *Persea* group (Li et al. 2011). In our study, plastid phylogenomics was shown to be an efficient way to resolve relationships in both the *Persea* group (Figure 5) and the Lauraceae family (Song et al. 2017a). With species from six genera of Lauraceae, our phylogenomic analysis based on 12 published chloroplast genomes supported a monophyletic *Persea* group comprising species of *Alseodaphne*, *Phoebe*, *Persea* and *Machilus*, as in previously published topologies (Rohwer et al. 2009; Li et al. 2011). Species of *Endiandra* formed the basal group in the phylogeny, and *Litsea* Lam. and *Cinnamomum* Trew are separated from both the *Endiandra* group and the *Persea* group, as in previously published phylogenetic trees (Rohwer 2000; Chanderbali et al. 2001; Rohwer and Rudolph 2005). These results suggest that plastid phylogenomics could be used to reconstruct a

robust phylogeny for the *Persea* group or even the whole family Lauraceae with systematic sampling.

Acknowledgments: The authors would like to thank Prof. Jie Li and Dr. Lang Li from the Plant Phylogenetics & Conservation Group of Xishuangbanna Tropical Botanical Garden, Chinese Academy Sciences, for plant tissue exchange of *Alseodaphne gracilis*, and acknowledge Jun-bo Yang, Jing Yang, Juan-hong Zhang, Chun-yan Lin and Ji-xiong Yang at the Germplasm Bank of Wild Species in Southwest China, Kunming Institute of Botany, Chinese Academy Sciences, for sequencing technology. This work was supported by the project of the Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences (Y4ZK111B01), the National Natural Science Foundation of China (No. 31600531), the CAS "Light of West China" Program, and the 1000 Talents Program (WQ20110491035). We sincerely thank the anonymous referees and Dr. Oskar Faix for their critical and invaluable comments that greatly improved our manuscript.

Data Archiving Statement: The complete chloroplast genome sequence data of the three *Alseodaphne* species will be submitted to Genebank of NCBI through the revision process. All of the accession numbers from NCBI must be supplied prior to final acceptance of the manuscript.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.

References

Chanderbali, A.S., van der Werff, H., Renner, S.S. (2001) Phylogeny and historical biogeography of Lauraceae: Evidence from the chloroplast and nuclear genomes. Ann. Mo. Bot. Gard. 88:104–134.

Chano, V., López de Heredia, U., Collada, C., Soto, Á. (2017) Transcriptomic analysis of juvenile wood formation during the growing season in *Pinus canariensis*. Holzforschung 71: 919–937.

Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T. (2004) Mauve: Multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14:1394–1403.

Dong, W.P., Xu, C., Li, C.H., Sun, J.H., Zuo, Y.J., Shi, S., Cheng, T., Guo, J.J., Zhou, S.L. (2015) *ycf1*, the most promising plastid DNA barcode of land plants. Scientific Reports 5:8348.

Hinsinger, D.D., Strijk, J.S. (2017) Toward phylogenomics of Lauraceae: the complete chloroplast genome sequence of *Litsea glutinosa* (Lauraceae), an invasive tree species on Indian and Pacific Ocean islands. Plant Gene 9:71–79.

Hung, K. H., Lin, C. H. Ju, L. P. (2017) Tracking the geographical origin of timber by DNA fingerprinting: a study of the endangered species Cinnamomum kanehirae in Taiwan. Holzforschung 71:853–862.

Jiao, L., Yin, Y., Cheng, Y., Jiang, X. (2014) DNA barcoding for identification of the endangered species *Aquilaria sinensis*: comparison of data from heated or aged wood samples. Holzforschung 68:487–494.

Jiao, L., Liu, X., Jiang, X., Yin, Y. (2015) Extraction and amplification of DNA from aged and archaeological *Populus euphratica* wood for species identification. Holzforschung 69:925–931.

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

Korotkova, N., Nauheimer, L., Ter-Voskanyan, H., Allgaier, M., Borsch, T. (2014) Variability among the most rapidly evolving plastid genomic regions is lineage-specific: implications of pairwise genome comparisons in *Pyrus* (Rosaceae) and Other Angiosperms for Marker Choice. PLoS One 9:e112998.

Lee, S.S., Chang, S.M., Chen, C.H. (2001) Chemical constituents from Alseodaphne andersonii. J. Nat. Prod. 64:1548–1551.

Li, L., Li, J., Rohwer, J.G., van der Werff, H., Wang, Z.H., Li, H.W. (2011) Molecular phylogenetic analysis of the Persea group (Lauraceae) and its biogeographic implications on the evolution of tropical and subtropical Amphi-Pacific disjunctions. Am. J. Bot. 98:1520–1536.

Librado, P., Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lohse, M., Drechsel, O., Kahlau, S., Bock, R. (2013) OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 41:W575–W581.

Michael, H.A., Schröder, H., Wischnewski, N., Degen, B., Magel, E., Fladung, M. (2012) Development of DNA-based methods to identify CITES-protected timber species: a case study in the Meliaceae family. Holzforschung 66:97–104.

Nafiah, M.A., Mukhtar, M.R., Omar, H., Ahmad, K., Morita, H., Litaudon, M., Awang, K., Hadi, A.H.A. (2011) N-Cyanomethylnorboldine: A new aporphine isolated from *Alseodaphne perakensis* (Lauraceae). Molecules 16:3402–3409.

Nafiah, M.A., Tan, S.P., Khoo, J.H.C., Hasnan, M.H.H., Awang, K., Hadi, A.H.A., Ahmad, K. (2016) A new aporphine alkaloid from the leaves of *Alseodaphne corneri* Kosterm (Lauraceae). Tetrahedron. Lett. 57:1537–1539.

Patel, R.K., Jain, M. (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7:e30619.

Parker, J., Helmstetter, A.J., Devey, D., Wilkinson, T., Papadopulos, A.S.T. (2017) Field-based species identification of closely-related plants using real-time nanopore sequencing. Scientific Reports 7:8345.

Rohwer, J.G. (2000) Toward a phylogenetic classification of the Lauraceae: evidence from matK sequences. Syst. Bot. 25:60–71.

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. Mo. Bot. Gard. 92:153–178.

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.

Rossetto, M., Kooyman, R., Yap, J.Y.S., Laffan, S.W. (2015) From ratites to rats: the size of fleshy fruits shapes species' distributions and continental rainforest assembly. Proc. Biol. Sci. 282:1820.

Sandak, J., Sandak, A., Cantini, C., Autino, A. (2015) Differences in wood properties of *Picea abies* L. Karst. in relation to site of provenance and population genetics. Holzforschung 69:385–397.

Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L. (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 94:275–288.

Song, Y., Dong, W., Liu, B., Xu, C., Yao, X., Gao, J., Corlett, R.T. (2015) Comparative analysis of complete chloroplast genome sequences of two tropical trees *Machilus yunnanensis* and *Machilus balansae* in the family Lauraceae. Front. Plant Sci. 6:662.

Song, Y., Yao, X., Tan, Y.H., Gan, Y., Corlett, R.T. (2016) Complete chloroplast genome sequence of the avocado: gene organization, comparative analysis, and phylogenetic relationships with other Lauraceae. Can. J. For. Res. 46:1293–1301.

Song, Y., Yu, W.B., Tan, Y.H., Liu, B., Yao, X., Jin, J.J., Padmanaba, M., Yang, J.B., Corlett, R.T. (2017a) Evolutionary comparisons of the chloroplast genome in Lauraceae and insights into loss events in the Magnoliids. Genome Biol. Evol. 9: 2354–2364.

Song, Y., Yao, X., Tan, Y.H., Gan, Y., Yang, J.B., Corlett, R.T. (2017b) Comparative analysis of complete chloroplast genome sequences of two subtropical trees, *Phoebe sheareri* and *Phoebe omeiensis* (Lauraceae) Tree Genet. Genomes 13:120.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739.

Thakur, B.K., Anthwal, A., Rawat, D.S., Rawat, B., Rashmi, Rawat, M.S.M. (2012) A review on genus *Alseodaphne*: phytochemistry and pharmacology. Mini. Rev. Org. Chem. 9:433–445.

Verma, R.S., Kumar, A., Mishra, P., Kuppusamy, B., Padalia, R.C., Sundaresan, V. (2016) Essential oil constituents of *Alseodaphne semecarpifolia* from central western Ghats, India. Chem. Nat. Compd. 52:516–517.

Wei, F.N., Werff, H.V.D. (2008) "Flora of China". In Dipterocarpaceae, Vol. 13. Eds. Wu, Z.Y., Raven, P.H., Hong, D.Y. Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis, MO, pp. 201–224.

Wu, C.C., Ho, C.K., Chang, S.H. (2016) The complete chloroplast genome of *Cinnamomum kanehirae* Hayata (Lauraceae). Mitochondrial DNA Part A 27:2681–2682.

Wyman, S.K., Jansen, R.K., Boore, J.L. (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252–3255.

Yang, J.B., Li, D.Z., Li, H.T. (2014) Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. Mol. Ecol. Resour. 14:1024–1031.

Yu, M., Liu, K., Zhou, L., Zhao, L., Liu, S. (2016) Testing three proposed DNA barcodes for the wood identification of *Dalbergia odorifera*. Holzforschung 70:127–136.

Zahari, A., Ablat, A., Sivasothy, Y., Mohamad, J., Choudhary, M.I., Awang, K. (2016) In vitro antiplasmodial and antioxidant activities of bisbenzylisoquinoline alkaloids from *Alseodaphne corneri* Kosterm. Asian Pac. J. Trop. Med. 9:328–332.