Received: 1 March 2019 | Revised: 6 June 2019 | Accepted: 12 June 2019

DOI: 10.1111/1755-0998.13050

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

Testing and using complete plastomes and ribosomal DNA sequences as the next generation DNA barcodes in *Panax* (Araliaceae)

Yunheng Ji[1,2] | Changkun Liu[1] | Zhenyan Yang[1] | Lifang Yang[1,3] | Zhengshan He[4] | Hengchang Wang[5] | Junbo Yang[4] | Tingshuang Yi[4]

[1]CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

[2]Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Population, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

[3]School of Life Science, Yunnan University, Kunming, China

[4]Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

[5]CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China

Correspondence
Yunheng Ji, CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China.
Email: jiyh@mail.kib.ac.cn
Junbo Yang and Tingshuang Yi, Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China.
Emails:jbyang@mail.kib.ac.cn(JY) and tingshuangyi@mail.kib.ac.cn(TY)

Funding information
the Major Program of National Natural Science Foundation of China, Grant/Award Number: 31590820 and 31590823; National Natural Science Foundation of China, Grant/Award Number: 31872673; Chinese Academy of Sciences the Large-scale Scientific Facilities, Grant/Award Number: 2017-LSF-GBOWS-02; the NSFC-Joint Foundation of Yunnan Province, Grant/Award Number: U1802287

Abstract

Complete plastid genome (plastome) sequences and nuclear ribosomal DNA (nrDNA) regions have been proposed as candidates for the next generation of DNA barcodes for plant species discrimination. However, the efficacy of this approach still lacks comprehensive evaluation. We carried out a case study in the economically important but phylogenetically and taxonomically difficult genus *Panax* (Araliaceae). We generated a large data set of plastomes and nrDNA sequences from multiple accessions per species. Our data improved the phylogenetic resolution and levels of species discrimination in *Panax*, compared to any previous studies using standard DNA barcodes. This provides new insights into the speciation, lineage diversification and biogeography of the genus. However, both plastome and nrDNA failed to completely resolve the phylogenetic relationships in the *Panax bipinnatifidus* species complex, and only half of the species within it were recovered as monophyletic units. The results suggest that complete plastome and ribosomal DNA sequences can substantially increase species discriminatory power in plants, but they are not powerful enough to fully resolve phylogenetic relationships and discriminate all species, particularly in evolutionarily young and complex plant groups. To gain further resolving power for closely related species, the addition of substantial numbers of nuclear markers is likely to be required.

KEYWORDS
next-generation DNA barcodes, *Panax*, plastome, ribosomal DNA, species discrimination, standard DNA barcodes

1 | INTRODUCTION

The mitochondrial gene cytochrome oxidase 1 (*COI*) has been shown to be effective and reliable as the standard animal DNA barcode for species identification and species discovery (Costa et al., 2007; Hajibabaei, Janzen, Burns, Hallwachs, & Hebert, 2006; Kim et al., 2012; Pons et al., 2006; Smith, Poyarkov, & Hebert, 2008; Ward, Zemlak, Innes, Last, & Hebert, 2005). However, in plants, reliable species discrimination based on plastid sequences (i.e. *rbcL*, *matK*, *trnH-psbA*), and the internal transcribed spacer (ITS) of the nuclear ribosomal DNA (nrDNA) (hereafter "standard DNA barcodes"), remains problematic, especially in recently diverged and rapidly radiated taxa (Coissac, Hollingsworth, Lavergne, & Taberlet, 2016; Hollingsworth, 2011; Hollingsworth et al., 2009; Hollingsworth, Graham, & Little, 2011; Hollingsworth, Li, Michelle, & Twyford, 2016; Li et al., 2011). With the advent of next-generation DNA sequencing technologies, the costs of genome sequencing have decreased, making it possible to generate large amounts of genomic data to extend the concept of DNA barcoding for plant species identification (Coissac et al., 2016; Hollingsworth et al., 2016; Li et al., 2015). The development of extended DNA barcoding approaches for species identification has been referred to as "ultra-barcoding" (Kane et al., 2012), or "plant barcoding 2.0" (Hollingsworth et al., 2016).

Compared to standard DNA barcodes, whole plastid genomes (plastomes) and entire nrDNA arrays possess more variable characters and therefore have the potential to improve discriminatory power (Kane et al., 2012; Nock et al., 2011; Ruhsam et al., 2015; Tontifilippini, Nevill, Dixon, & Small, 2017). In addition, they are present in multiple copies in each plant cell, making assembly feasible from low-coverage genome sequencing (Coissac et al., 2016; Hollingsworth et al., 2016; Straub et al., 2012; Zeng et al., 2018). Because of these advantages, plastomes and nrDNA sequences have been recommended for consideration as candidates for the next generation of DNA barcodes for plants (Hollingsworth et al., 2016; Kan & Cronk, 2008; Kane et al., 2012; Nock et al., 2011; Ruhsam et al., 2015; Tontifilippini et al., 2017; Yang, Tang, Li, Zhang, & Li, 2013).

The discriminatory power of complete plastomes and nrDNA sequences in plants has been evaluated in some recent studies (e.g., Fu et al., 2019; Kane et al., 2012; Ruhsam et al., 2015; Turner, Paun, Munzinger, Chase, & Samuel, 2016; Yang et al., 2013). However, to our knowledge, very few studies have been undertaken which sample multiple individuals from multiple congeneric species. Where only single individuals are sampled per taxon, it is likely that unique characters will be found (including sequencing errors), given the large size of the data set. However, this variation will not necessarily translate to robust data for species discrimination. In contrast, sampling multiple individuals per species allows us to test for species-level monophyly. This is particularly important given the known extent of plastid and rDNA introgression in plants (Rieseberg & Soltis, 1991). For instance, fewer than half of the New Caledonian *Araucaria* species with multiple accessions sampled were resolved as monophyletic by complete plastome sequencing (Ruhsam et al., 2015). Therefore, the efficacy of plastome and complete nrDNA barcoding approaches needs to be further assessed.

Panax L. (the ginseng genus; Araliaceae) is one of the most medically important plant groups in the world, and is disjunctly distributed in East Asia and eastern North America (Lee & Wen, 2004; Wen & Zimmer, 1996; Zuo et al., 2011). The genus includes seven well-defined species (*Panax ginseng*, *P. japonicus*, *P. notoginseng*, *P. pseudoginseng*, *P. quinquefolius*, *P. stipuleanatus* and *P. trifolius*) and one species complex (*P. bipinnatifidus* species complex). Almost all species within the genus has been used as medicinal herbs in East Asia, especially in China (Choi & Wen, 2000). The *P. bipinnatifidus* species complex (Zuo et al., 2011) is morphologically diverse and taxonomically difficult, including six species (*P. bipinnatifidus*, *P. major*, *P. sinensis*, *P. vietnamensis*, *P. wangianus* and *P. zingiberensis*) distributed in the Sino-Himalayan region according to the taxonomic treatment of Wen and Zimmer (1996). The *P. bipinnatifidus* species complex comprises several recently diverged lineages (Zuo et al., 2011; Zuo, Wen, Ma, & Zhou, 2015; Zuo, Wen, & Zhou, 2017), which may have radiated since the Oligocene/Miocene transition (Zuo et al., 2017). Neither standard DNA barcodes nor AFLP have provided adequate data to fully reconstruct robust phylogenetic relationships or successfully distinguish all of the species within the species complex (Lee & Wen, 2004; Wen & Zimmer, 1996; Zhu, Fushimi, Cai, & Komatsu, 2003; Zuo et al., 2011, 2015, 2017).

To date, no DNA markers have been developed which can fully discriminate among all species and resolve all evolutionary relationships within *Panax*. The genus thus provides an ideal model to evaluate the usefulness of whole plastomes and nrDNA sequences for phylogenetic resolution and extended barcodes for species discrimination. By sampling multiple individuals per species, we use low-coverage genome sequencing technology to generate plastome and nrDNA sequences to investigate whether they can improve phylogenetic resolution and species discrimination in this phylogenetically and taxonomically difficult genus. Specifically, we addressed the following questions. (a) Compared to standard DNA barcodes, do plastomes and nrDNA sequences improve phylogenetic resolution and species identification in the genus *Panax*? (b) If so, to what extent do the plastomes and ribosomal DNA sequences match the species boundaries?

2 | MATERIALS AND METHODS

2.1 | Samples

In total, 52 individual plants from nine *Panax* species were sampled (Table 1). Plant material of *Panax ginseng*, *P. notoginseng* and *P. quinquefolius* was harvested from medicinal plantations. Other species were collected from the wild using records from herbarium specimens to locate populations. Leaf tissues were dried with silica gel, and vouchers were deposited at the Herbarium of Kunming

TABLE 1 Taxa included in this study with locality, voucher and GenBank accession numbers

Sample	Locality	Voucher	Accession number of plastome	Accession number of ribosomal DNA
Panax bipinnatifidus	Mianning, Sichuan, China	JYH-2016486	MK408959	MK408795
P. bipinnatifidus	Mianning, Sichuan, China	JYH-2016488	MK408944	MK408763
P. bipinnatifidus	Zhatong, Yunnan, China	JYH-2016481	MK408926	MK408776
P. bipinnatifidus	Wuxi, Chongqing, China	JYH-2016547	MK408957	MK408786
P. bipinnatifidus	Shiyan, Hubei, China	JYH-2016548	MK408918	MK408792
P. ginseng	Cultivated in Wuding, Yunnan, China	JYH-2016473	MK408938	MK408780
P. major	Jianchuan, Yunnan, China	JYH-2016483	MK408962	MK408761
P. major	Lijiang, Yunnan, China	JYH-2016484	MK408948	MK408782
P. major	Heqing, Yunnan, China	JYH-2016485	MK408940	MK408800
P. notoginseng	Maguan, Yunnan, China	LCK-1	MK408927	MK408771
P. notoginseng	Maguan, Yunnan, China	LCK-2	MK408945	MK408777
P. notoginseng	Maguan, Yunnan, China	LCK-3	MK408937	MK408762
P. notoginseng	Xichou, Yunnan, China	LCK-4	MK408946	MK408810
P. notoginseng	Xichou, Yunnan, China	LCK-5	MK408925	MK408769
P. notoginseng	Xichou, Yunnan, China	LCK-7	MK408931	MK408766
P. notoginseng	Xichou, Yunnan, China	LCK-8	MK408928	MK408804
P. notoginseng	Qiubei, Yunnan, China	LCK-9	MK408955	MK408784
P. notoginseng	Qiubei, Yunnan, China	LCK-10	MK408954	MK408775
P. quinquefolius	Cultivated in Wuding, Yunnan, China	JYH-2016493	MK408923	MK408799
P. quinquefolius	Cultivated in Wuding, Yunnan, China	JYH-2016494	MK408953	MK408785
P. sinensis	Enshi, Hubei, China	JYH-2016608	MK408932	MK408812
P. sinensis	Hefeng, Hubei, China	JYH-2016609	MK408961	MK408789
P. sinensis	Shiyan, Hubei, China	JYH-2016610	MK408958	MK408774
P. sinensis	Zhuxi, Hubei, China	JYH-2016611	MK408919	MK408767
P. sinensis	Zhenping, Hubei, China	JYH-2016612	MK408924	MK408798
P. sinensis	Tianquan, Sichuan, China	JYH-2016474	MK408967	MK408778
P. sinensis	Baoxing, Sichuan, China	JYH-2016475	MK408956	MK408796
P. sinensis	Shimian, Sichuan, China	JYH-2016476	MK408939	MK408770
P. stipuleanatus	Maguan, Yunnan, China	JYH-2016466	MK408920	MK408807
P. stipuleanatus	Pingbian, Yunnan, China	JYH-2016435	MK408936	MK408811
P. stipuleanatus	Pingbian, Yunnan, China	JYH-2016437	MK408965	MK408781
P. wangianus	Tengchong, Yunnan, China	JYH-2016540	MK408964	MK408801
P. wangianus	Tengchong, Yunnan, China	JYH-2016539	MK408930	MK408806
P. wangianus	Tengchong, Yunnan, China	JYH-2016541	MK408941	MK408764
P. wangianus	Longlin, Yunnan, China	JYH-2016543	MK408943	MK408779
P. wangianus	Mt. Emei, Sichuan, China	JYH-2016544	MK408934	MK408773
P. wangianus	Mt. Emei, Sichuan, China	JYH-2016545	MK408921	MK408797
P. wangianus	Zhenfeng, Guizhou, China	JYH-2016550	MK408935	MK408765
P. wangianus	Zhenfeng, Guizhou, China	JYH-2016551	MK408963	MK408805
P. wangianus	Zhenfeng, Guizhou, China	JYH-2016552	MK408951	MK408809
P. zingiberensis	Maguan, Yunnan, China	JYH-2016465	MK408933	MK408794
P. zingiberensis	Maguan, Yunnan, China	JYH-2016472	MK408922	MK408783
P. zingiberensis	Maguan, Yunnan, China	JYH-2016468	MK408960	MK408808
P. zingiberensis	Maguan, Yunnan, China	JYH-2016467	MK408949	MK408791

(Continues)

TABLE 1 (Continued)

Sample	Locality	Voucher	Accession number of plastome	Accession number of ribosomal DNA
P. zingiberensis	Simao, Yunnan, China	JYH-2016436	MK408942	MK408772
P. zingiberensis	Simao, Yunnan, China	JYH-2016438	MK408952	MK408790
P. zingiberensis	Simao, Yunnan, China	JYH-2016439	MK408966	MK408787
P. zingiberensis	Simao, Yunnan, China	JYH-2016440	MK408947	MK408768
P. zingiberensis	Mengzi, Yunnan, China	JYH-2016441	MK408929	MK408802
P. zingiberensis	Mengzi, Yunnan, China	JYH-2016463	MK408969	MK408788
P. zingiberensis	Mengzi, Yunnan, China	JYH-2016464	MK408968	MK408803
P. zingiberensis	Maguan, Yunnan, China	JYH-2016471	MK408950	MK408793

Institute of Botany, Chinese Academy of Sciences (KUN). Except for *P. ginseng*, multiple individuals within a species representing different localities were included in the study.

Beside the newly collected material for DNA sequencing, publicly available plastomes (25 accessions, Table S1) and nrDNA sequences (13 accessions, Table S2) of *Panax* were also included. To avoid sequence misidentification, only sequences from published papers or with available vouchers were selected. The total data set of de novo sequenced individuals and samples from GenBank consisted of 77 individuals of 12 species. For *Panax vietnamensis*, multiple individuals were available for plastid genome sequences from the GenBank database, but not for nrDNA, meaning 10 species in total were available for tests of species-level monophyly for plastome sequences and nine species for nrDNA sequences.

2.2 | Illumina sequencing, assembly and annotation

Genomic DNA was extracted from ~20 mg leaf samples by a modified CTAB method (Doyle & Doyle, 1987). Approximately 5 μg of purified genomic DNA was sheared by sonication to generate fragments ~500 bp in length for constructing a paired-end library. Illumina libraries were prepared according to the manufacturer's protocol. Paired-end sequencing was performed on the Illumina HiSeq 2000 system.

Adaptors and low-quality reads were removed from the raw data via the NGS QC Toolkit (Patel & Jain, 2012), by setting the cut-off value for the percentage of read length to 80 and PHRED quality scores to 30. The clean reads were assembled into contigs with SPAdes version 3.10.1 (Bankevich et al., 2012). The complete plastome sequence of *P. ginseng* (GenBank accession: KM067394) was downloaded as a reference for the plastid genome assembly. Contigs matching the reference were extracted and assembled de novo in Bowtie version 2.2.6 (Langmead & Salzberg, 2012) with default parameters. Plastomes were annotated with the Dual Organellar Genome Annotator database (Wyman, Jansen, & Boore, 2004) integrated in Geneious version 10.2 (Kearse et al., 2012). Start and stop codons and intron/exon boundaries for protein-coding genes were checked manually. Annotated tRNA genes were further verified by tRNAscan-SE 1.21 (Schattner, Brooks, & Lowe, 2005) with default parameters. The boundaries of the large-single copy (LSC), small-single copy (SSC) and inverted-repeat (IR) regions for each plastome were visually examined and manually adjusted according to those of the reference plastome.

The ribosomal DNA sequence (including 26S, 18S and 5.8S ribosomal RNA genes, and ITS) of *P. ginseng* (GenBank accession: KM036395) was used as the reference to assemble nrDNA sequences. The internal gene spacer (IGS) region in *Panax* species possesses too many repeat sequences and inversions that may make the assembly inaccurate, and therefore we did not assemble the region. Contigs mapping to reference ribosomal DNA sequence were assembled using the processes described above. The ribosomal RNA genes and ITS regions of each newly generated sequence were annotated by comparison with the reference.

2.3 | Data analysis

We investigated the utility of the complete plastomes and nrDNA sequences for species identification using tree-based and genetic distance methods. The number of species with multiple accessions resolving as monophyletic was recorded, as was the branch support for each node. To compare the discriminatory power of standard and next-generation DNA barcodes, plastid *rbcL*, *matK*, *trnH-psbA* and nuclear ITS sequences were subsampled from each accession and compared with plastome and complete nrDNA sequences.

The DNA sequences from the four data sets: (a) whole plastomes, (b) nrDNA sequences, (c) combination of *rbcL*, *matK* and *trnH-psbA*, and (d) ITS, were respectively aligned using the program mafft (Katoh & Standley, 2013) with manual adjustment where necessary. Phylogenetic analysis for each data set was performed using maximum likelihood (ML) and Bayesian inference (BI) methods. For all analyses, *Aralia elata* was used to root the tree. The best-fit substitution model GTR + G was determined by modeltest 3.7 (Posada & Crandall, 1998) using the Akaike information criterion (Posada & Buckley, 2004). ML analyses were performed in the software RAxML-HPC BlackBox version 8.1.24 (Stamatakis, 2006). The best scoring ML tree for each data set was generated with 1,000 rapid bootstrap (BS) replicates to provide support values. BI analyses were performed in MrBayes version 3.2 (Ronquist & Huelsenbeck, 2003).

TABLE 2 Comparison of characteristics of standard barcodes and complete plastome and nrDNA sequences in *Panax*

Sequence	Aligned length (bp)	No. of variable sites (% divergence)	No. of parsimony-informative sites (%divergence)
*mat*K + *rbc*L+trnH-*psb*A	3,365	141 (4.19)	85 (2.53)
ITS	612	75 (12.25)	54 (8.82)
rDNA	5,852	138 (2.36)	103 (1.76)
Entire plastid genome	158,602	3,948 (2.49)	2,195 (1.38)

Two independent Markov chain Monte Carlo (MCMC) runs were performed with 1,000,000 generations, sampling every 100 generations. Trees from the first 250,000 generations were discarded as burn-in, with posterior probability (PP) values computed from the remaining trees.

Pairwise distance was calculated using the Kimura 2-parameter (K2P) distance in MEGA 6.0 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013) to detect barcoding gaps for each data set (whole plastomes, nrDNA sequences, combination of *rbcL*, *matK* and *trnH-psbA*, and ITS). Barcoding gaps were identified by the scatter plot of the minimum interspecific distance versus maximum intraspecific distance for each species. A species with its minimum interspecific distance larger than its maximum intraspecific distance is considered to be successfully identified (Collins & Cruickshank, 2013; Liu et al., 2014).

3 | RESULTS

3.1 | Low-coverage genome sequencing

Illumina sequencing generated between 15 and 53.7 million paired-end clean reads per sample; of those, 1.8E5 to 4.2E6 and 4.2E4 to 4.0E5 reads were mapped to the reference chloroplast genome and ribosomal DNA sequences, respectively. Based on these data, the de novo assembly covered the entire plastome and nrDNA for all samples, with average coverage ranging from 176× to 4,021×, and from 1,072× to 10,122×. In most analysed accessions, the sequencing depth of plastomes was lower than that of the nrDNA sequences (Table S3).

3.2 | Analyses of plastomes

The de novo assembly produced 52 *Panax* plasomes, whose size ranged from 156,006 to 156,403 bp (Table S3). The plastomes exhibited a typical quadripartite structure, with two copies of a large IR separated by an LSC and an SSC. The overall feature of the plastomes is conserved in *Panax*. All plastomes encoded 114 unigenes, including 80 protein-coding genes, 30 tRNAs and four rRNA genes (Table S4). These genes were arranged in an identical order (Figure S1). Alignment of the plastomes yielded a matrix of 158,602 positions. In this alignment, we identified 3,948 variable sites, 2,195 of which were parsimony-informative (Table 2).

Tree topologies resulting from ML and BI analysis were identical (Figure 1). Four major clades were recovered within *Panax*. *Panax trifolius* and *P. stipuleanatus* were resolved as the two most basally diverged clades, and the remaining species grouped into two fully supported monophyletic clades (BS = 100%, PP = 1.00). Within the first clade, *Panax ginseng* was sister to *P. quinquefolius* (BS = 100%, PP = 1.00), and the clade including these two species was sister to *P. japonicus* (BS = 100%, PP = 1.00). Within the second clade, a sister relationship between *Panax notoginseng* and the *P. bipinnatifidus* species complex was recovered (BS = 67%, PP = 0.93). The *P. bipinnatifidus* species complex was resolved as two well-supported monophyletic lineages (BS = 100%, PP = 1.00). Relationships within the clade including *Panax zingiberensis*, *P. vietnamensis* and *P. wangianus* were resolved and showed high branch support: *P. vietnamensis* was sister to *P. wangianus* (BS = 93%, PP = 1.00), and these two species, in turn, were sister to *P. zingiberensis* (BS = 100%, PP = 1.00). In the tree topology, the clade comprising *P. bipinnatifidus*, *P. major* and *P. sinensis* was divided into two subclades: one included accessions from central China (east subclade), whereas the other consisted of individuals from southwest China (west subclade). Relationships among *P. bipinnatifidus*, *P. major* and *P. sinensis* remained unresolved.

The discriminatory power of plastomes as DNA barcodes was assessed by investigating the monophyly and branch support recovered (Table 3) in those species with multiple accessions sampled. Among the 10 species with more than one plastome being analysed, seven species (*P. stipuleanatus*, *P. ginseng*, *P. quinquefolius*, *P. notoginseng*, *P. zingiberensis*, *P. vietnamensis* and *P. wangianus*) were recovered as fully supported monophyletic units (BS = 100%, PP = 1.00). In contrast, accessions of *P. bipinnatifidus*, *P. major* and *P. sinensis* were resolved as polyphyletic.

3.3 | Analyses of nrDNA sequences

Our de novo nrDNA assembly covered the 18S, ITS1, 5.8S, ITS2 and 26S regions, and the sequence length ranged from 5,873 to 5,878 bp. Alignment of the nrDNA showed high sequence variation. In total, 138 variable sites were identified, 103 of which were parsimony-informative (Table 2). Phylogenetic analysis based on ML and BI gave the overall same tree topology within *Panax*, although they exhibited minor differences within clades and varying support values (Figure 2).

FIGURE 1 The phylogenetic tree of *Panax*, via maximum likelihood (ML) and Bayesian inference (BI), based on whole plastomes. Numbers at nodes indicate bootstrap percentages (left) from ML analysis and posterior probabilities (right) from BI analysis [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Species-level monophyly of multiple accessions and branch support recovered by analyses of whole plastome, ribosomal DNA sequence and standard DNA barcodes (*rbcL*, *matK*, *trnH-psbA*, ITS)

Taxa	Plastome				Ribosomal DNA				*matK* + *rbcL*+trnH -*psbA*				ITS			
	ML		BI		ML		BI		ML		BI		ML		BI	
	Monophyly	BS%	Monophyly	PP	Monophyly	BS%	Monophyly	PP	Monophyly	BS%	Monophyly	PP	Monophyly	BS%	Monophyly	PP
Panax bipinnatifidus	N	–	N	–	N	–	N	–	N	–	N	–	N	–	N	–
P. ginseng	Y	100	Y	1.00	Y	100	Y	1.00	Y	78	Y	0.85	Y	90	Y	1.00
P. major	N	–	N	–	N	–	N	–	N	–	N	–	N	–	N	–
P. notoginseng	Y	100	Y	1.00	Y	100	Y	1.00	Y	100	Y	1.00	Y	100	Y	1.00
P. quinquefolius	Y	100	Y	1.00	Y	97	Y	0.92	Y	88	Y	1.00	Y	84	Y	0.93
P. sinensis	N	–	N	–	N	–	N	–	N	–	N	–	N	–	N	–
P. stipuleanatus	Y	100	Y	1.00	Y	100	Y	1.00	Y	100	Y	1.00	Y	100	Y	1.00
P. vietnamensis	Y	100	Y	1.00	#	–	#	–	N	–	N	–	#	–	#	–
P. wangianus	Y	100	Y	1.00	Y	91	Y	1.00	N	–	N	–	N	–	N	–
P. zingiberensis	Y	100	Y	1.00	Y	84	Y	0.95	N	–	N	–	N	–	N	–

Abbreviations: BI, Bayesian inference; BS%, bootstrap support percentage; ML, maximum likelihood; N, no; #, only one accession; PP, posterior probability; Y, yes.

Our phylogenetic analyses of nrDNA sequences resulted in distinct tree topologies within *Panax*. Although the phylogenetic position of *P. stipuleanatus* exhibited no significant difference between tree topologies of plastome and nrDNA sequences (Figure 3), neither the monophyly of the *P. bipinnatifidus* species complex nor its sister relationship to *P. notoginseng* were recovered by analysing nrDNA sequences. Instead, *P. zingiberensis* and *P. vietnamensis* formed a monophyletic clade (BS = 99%, PP = 1.00), and the clade of *P. ginseng* and *P. quinquefolius* (BS = 87%, PP = 0.99) was grouped with accessions of *P. wangianus*, *P. bipinnatifidus*, *P. sinensis* and *P. major* (BS = 94%, PP = 0.99). For those species with multiple ribosomal DNA sequences available, *P. stipuleanatus* (BS = 100%, PP = 1.00), *P. ginseng* (BS = 100%, PP = 1.00), *P. quinquefolius* (BS = 97%, PP = 0.92), *P. notoginseng* (BS = 100%, PP = 1.00), *P. zingiberensis* (BS = 84%, PP = 0.95) and *P. wangianus* (BS = 91%, PP = 1.00) were resolved as well-supported monophyletic lineages, whereas *P. bipinnatifidus*, *P. major* and *P. sinensis* were not recovered as monophyletic (Table 3).

3.4 | Analyses of standard DNA barcodes

Combined analyses of plastid *rbcL*, *matK* and *trnH-psbA* sequences generated similar tree topologies to the complete plastome data sets (Figure S2). Four major clades were recovered within *Panax*. The relationships among or within these clades were almost identical to that from the plastid genome phylogeny but with lower support values. For those species with multiple individuals, five species, namely *P. stipuleanatus* (BS = 100%, PP = 1.00), *P. ginseng* (BS = 78%, PP = 0.85), *P. quinquefolius* (BS = 88%, PP = 1.00), *P. notoginseng* (BS = 100%, PP = 1.00) and *P. wangianus* (BS = 82%, PP = 0.85), were recovered as monophyletic groups. However, the individuals of the six sampled species belonging to the *P. bipinnatifidus* complex were not resolved as reciprocally monophyletic (Figure S2, Table 3).

The tree topologies of the ITS phylogeny (Figure S3) were almost identical to that of the complete nrDNA sequences, but were largely incongruent with that of the plastid phylogenies (e.g., complete plastomes and the combination of *rbcL*, *matK* and *trnH-psbA*). The ITS phylogenies received relatively low branch support in contrast to that of the full rDNA sequences. Similarly, individuals belonging to the five sampled species from the *P. bipinnatifidus* complex failed to resolve as reciprocally monophyletic (Figure S3, Table 3).

3.5 | Barcoding gaps

Analyses of intra- and interspecific K2P distance identified barcoding gaps that varied between data sets (Figure S4). For complete plastomes, 8/10 species with multiple accessions had distinct barcoding gaps (minimum interspecific distance larger than maximum intraspecific distance), and showed significantly better performance than the combination of *rbcL*, *matK* and *trnH-psbA* sequences (5/10 species displayed DNA barcoding gaps). For nrDNA sequences, 8/9 species with multiple individuals possessed DNA barcoding gaps, and the data set performed better than ITS (6/9 species showed DNA barcoding gaps).

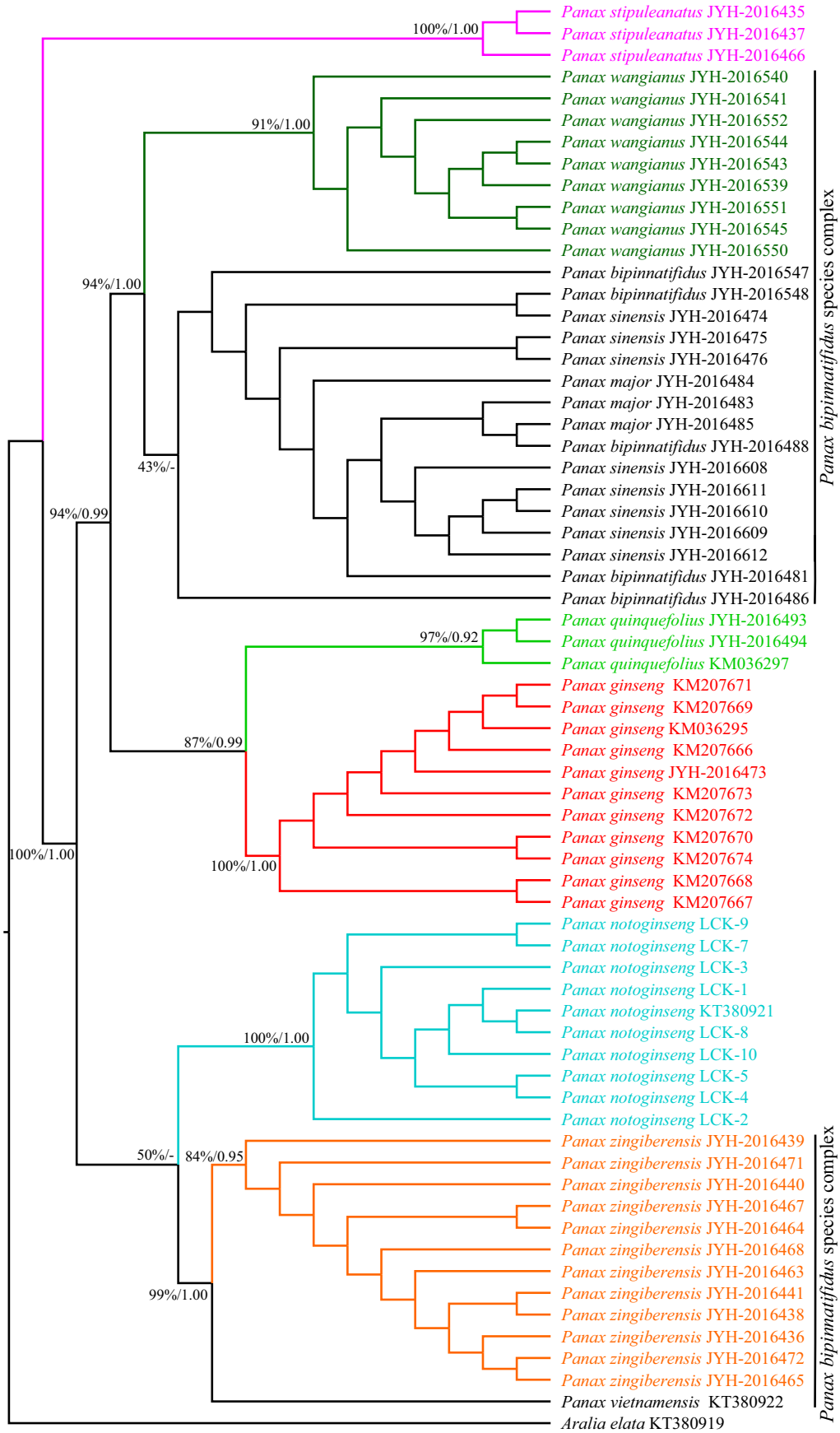

FIGURE 2 The phylogenetic tree of *Panax*, via maximum likelihood (ML) and Bayesian inference (BI), based on complete ribosomal DNA sequences. Numbers at nodes indicate bootstrap percentages (left) from ML analysis and posterior probabilities (right) from BI analysis [Colour figure can be viewed at wileyonlinelibrary.com]

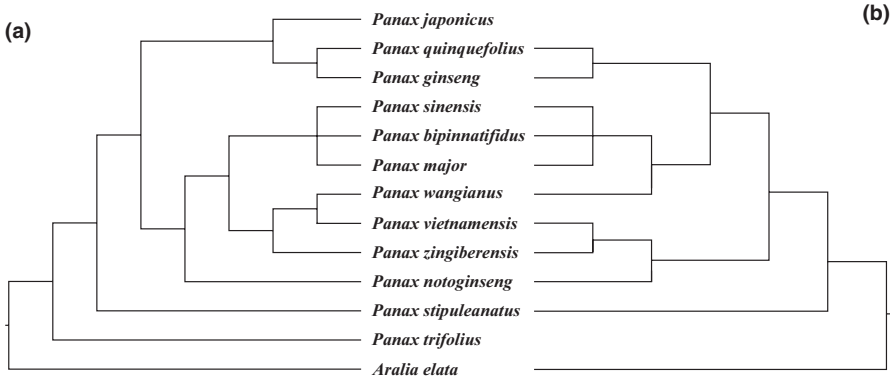

FIGURE 3 Topological differences between plastid (a) and nuclear (b) phylogenies

4 | DISCUSSION

4.1 | Phylogenetic resolution

Analyses of complete plastomes recovered four major clades within *Panax*. These clades were also identified by the current study using standard DNA barcodes (*rbcL*, *matK* and *trnH-psbA*), by Shi et al. (2015) using four plastid DNA regions, by Zuo et al. (2017) combining three plastid sequences and a mitochondrial DNA region, and by Jiang et al. (2018), Manzanilla et al. (2018) and Liu, Yang, Yang, Yang, and Ji (2018) using complete plastomes. However, with longer sequence length and more variable characters, or a much larger taxon sample than has previously been available, our data give the highest resolution to the relationships among and within these clades.

Phylogenetic incongruence between the maternally inherited plastid and biparentally inherited nuclear DNA makers has been observed in many plant groups (e.g., Folk, Soltis, Soltis, & Guralnick, 2018; Soltis & Kuzoff, 1995; Vargas, Ortiz, & Simpson, 2017). The infrageneric relationships within *Panax* reconstructed by nrDNA sequences are largely congruent with the analyses of multilocus nuclear regions (Shi et al., 2015) and ITS sequences (Zuo et al., 2017; current study), but conflict with those of plastid phylogenies (Figure 3). The cytonuclear discordances are possibly caused by "chloroplast capture" as a result of hybridization events followed by introgression (e.g., Acosta & Premoli, 2010; Rieseberg & Soltis, 1991; Soltis, Johnson, & Looney, 1996). The well-supported sister relationship between the clade comprising *Panax japonicus*, *P. ginseng* and *P. quinquefolius* and the clade including *P. notoginseng* and the *P. bipinnatifidus* species complex in the plastid phylogeny suggest that they may have originated from a common maternal ancestor. In addition, the distinct placement of the *P. ginseng* and *P. quinquefolius* clade on the tree topologies of nuclear sequences implies that hybridization or introgression might have occurred between the common ancestor of *P. ginseng* and *P. quinquefolius* and the common ancestor of *P. wangianus*, *P. sinensis*, *P. major* and *P. bipinnatifidus*.

A particularly difficult problem in the phylogenetics of *Panax* is that the relationships within the *P. bipinnatifidus* species complex remain unresolved (Shi et al., 2015; Zuo et al., 2015, 2017). Analyses of complete plastomes recovered the *P. bipinnatifidus* species complex as a fully supported monophyletic group, and split it into two well-supported subclades. Although well-supported relationships among *P. zingiberensis*, *P. vietnamensis* and *P. wangianus* have been reconstructed by plastome phylogenies, the relationships within the subclade including accessions of *P. sinensis*, *P. major* and *P. bipinnatifidus* remain unresolved. This suggests that neither whole plastomes nor nrDNA sequences are powerful enough to reconstruct robust phylogenetic relationships among some closely related species, as previous studies have indicated (Barrett, Davis, Leebens-Mack, Conran, & Stevenson, 2013; Turner et al., 2016).

4.2 | Monophyly of species

Within the ginseng genus, 10 species with more than one individual with complete plastome sequences (from three to 15 accessions per species) and nine species with more than one individual with complete ribosomal DNA sequences (from three to 12 accessions per species) were sampled, including four species (*P. stipuleanatus*, *P. ginseng*, *P. quinquefolius* and *P. notoginseng*) which were well recognized by previous investigations using standard DNA barcodes (Zhu et al., 2003; Zuo et al., 2011). Among the six species in the *P. bipinnatifidus* species complex, the monophyly of which has not been resolved by previous studies using standard DNA barcodes, *P. zingiberensis* and *P. wangianus* were resolved as monophyletic by both data sets, and the monophyly of *P. vietnamensis* was recovered by plastome sequencing (only a single individual was sampled for complete nrDNA, precluding a test for monophyly). Neither complete plastome sequences nor complete nrDNA sequences supported the monophyly of *P. bipinnatifidus*, *P. major* and *P. sinensis*, despite the fact that these species are morphologically distinct from each other (Table S5).

The absence of monophyly in *P. bipinnatifidus*, *P. major* and *P. sinensis* can be contributed to their young evolutionary ages. Results from Zuo et al. (2015, 2017) indicated that the *P. bipinnatifidus* species complex may have experienced a dramatic evolutionary radiation since the early Miocene. Therefore, there may have been insufficient time to accumulate enough species-specific sequence variations, or for complete sorting of ancestral polymorphisms among these closely related species (Barrett et al., 2013; Parks, Cronn, & Liston, 2009; Ruhsam et al., 2015).

FIGURE 4 Morphological intermediates of leaves in the genus *Panax*: (a) leaves of *P. major*, (b) leaves of putative hybrid between *P. major* and *P. bipinatifidus*, and (c) leaves of *P. bipinnatifidus* [Colour figure can be viewed at wileyonlinelibrary.com]

Another possible explanation for the nonmonophyly of *P. bipinnatifidus*, *P. major* and *P. sinensis* accessions is interspecific hybridization. In most localities across their ranges, these species are sympatric and flower in synchrony (Xiang & Lowry, 2007). This implies that natural hybridization between sympatric species is possible if their pollination mechanisms are compatible. Morphological intermediates of leaves between *P. bipinnatifidus* and *P. major* have been observed in the field (Figure 4), suggesting that natural hybridization or introgression may occur among at least some of these species.

4.3 | Implications for DNA barcoding

Our results indicated that, for those species (*P. stipuleanatus*, *P. ginseng*, *P. quinquefolius* and *P. notoginseng*) whose monophyly has been recognized by previous studies (Zuo et al., 2011, 2017) or the standard DNA barcodes subsampled from our data sets, analyses of complete plastomes and ribosomal DNA sequences similarly recovered them as monophyletic groups, but with significantly increased support values. Furthermore, within the *P. bipinnatifidus* species complex, although none of the six species was resolved as monophyletic by standard DNA barcodes, gains in discrimination were made using complete plastome and nrDNA sequences. Thus, three additional species were discriminated using complete plastome sequences (*P. vietnamensis*, *P. zingiberensis* and *P. wangianus*) and two species using nrDNA sequences (*P. zingiberensis* and *P. wangianus*; with only one accession of *P. vietnamensis* precluding a test for monophyly). These findings indicate that next-generation DNA barcodes improve the resolution of species identification in *Panax* compared with standard DNA barcodes. However, as previously discussed, *P. bipinnatifidus*, *P. major* and *P. sinensis* were not resolved as monophyletic in any of the data sets. Similarly, analysis of barcoding gaps indicated that complete plastome and nrDNA sequences have better performance than standard DNA barcodes, but display overlap for more or fewer species. The case study on *Panax* suggests that both complete plastomes and nrDNA sequences, to some extent, increase the resolution of species discrimination. Nevertheless, the gains in discriminatory power do not reach full species identification in the genus.

Complete plastomes and nrDNA sequences have been proposed as candidates for the next generation of DNA barcodes for plant species identification (Coissac et al., 2016; Hollingsworth et al., 2016; Kan & Cronk, 2008; Kane et al., 2012; Li et al., 2015; Nock et al., 2011; Tontifilippini et al., 2017), although the efficiency of this approach still lacks comprehensive evaluation. Furthermore,

the discriminatory power revealed by previous studies sampling multiple intraspecific individuals from multiple congeneric species is largely incongruent. For instance, either complete plastomes or nrDNA sequences were shown to successfully distinguish closely related species in *Theobroma* (Kane et al., 2012) and *Taxus* (Fu et al., 2019). However, they failed to fully identify species in *Panax* (the current study) and New Caledonian *Araucaria* (Ruhsam et al., 2015), even though there are some gains in discriminatory power. Taken together, these case studies suggest that using complete plastomes and nrDNA sequences in DNA barcoding substantially improve resolution in species identification, and even reach full species discrimination in some plant groups. However, the resolution of disagreement between species and gene trees is the fundamental challenge for DNA barcodes (Hollingsworth et al., 2009, 2011, 2016). Factors such as hybridization, recent diversification or rapid radiations, slow sequence mutation rates and restricted infraspecies gene flow may result in plant barcodes being commonly shared among closely related species, and thus not tracking species boundaries (Hollingsworth et al., 2016; Ruhsam et al., 2015). Given that complete plastomes and nrDNA sequences are a linkage group of plastid and ribosomal genes, they only represent at best two large single-locus DNA regions. Using these two data set, it is also difficult to fully address the problems that standard DNA barcodes have not resolved. For those evolutionarily young or complex plant groups, to realize the goal of telling all species apart, large numbers of unlinked nuclear markers are likely to be required as the complement to the next generation of DNA barcodes (Hollingsworth et al., 2016; Ruhsam et al., 2015).

There is a feasible routine to access substantial numbers of nuclear genes in *Panax*. To date, the whole nuclear genome of *P. ginseng* and *P. notoginseng* have been sequenced (Chen et al., 2017; Kim et al., 2018). Using them as references, it is now possible to assemble and annotate nuclear genes from the low-coverage genome sequencing data. A future study will screen the loci possessing high sequence variation to assess their discriminatory power in *Panax*.

5 | CONCLUSIONS

In this study, we have comprehensively evaluated the resolution of complete plastomes and nrDNA sequences in phylogenetic reconstruction and species discrimination in *Panax*, by including multiple accessions for 10/12 sampled species. Our data produced strongly supported and highly resolved phylogenies in this economically important but phylogenetically and taxonomically difficult plant group. In comparison with standard DNA barcodes, analyses of whole plastomes and nrDNA sequences improved resolution in species discrimination. Nevertheless, despite longer sequence length and more informative characters, the gene trees from both data sets do not fully match the species boundaries in *Panax*. This case study suggests that complete plastomes and nrDNA can substantially increase species discriminatory power in plants, but they are not are powerful enough to reach a full species identification, especially in evolutionarily young and complex plant groups. To achieve the ultimate goal of DNA barcodes (telling all plant species apart), a substantial number of nuclear markers is likely to be required as the complement to the next generation of DNA barcodes.

ACKNOWLEDGEMENTS

We thank Dr Peter M. Hollingsworth (Royal Botanic Garden Edinburgh) and two anonymous reviewers for their insightful comments in improving the manuscript. We are also grateful to Guohua Zhou, Ren Zhao, Zhangming Wang, Guohai He, Chengjin Yang, Bao Su, Yulong Li, Fengtian Lin, Qiang Zhao, Mingfu Yang, Liangxue Sun, Sirong Yi, Qiuping Tan and Qiliang Gan for help with sample collection. This study was supported by a grant of the Large-scale Scientific Facilities of the Chinese Academy of Sciences (No. 2017-LSF-GBOWS-02), the Major Program of National Natural Science Foundation of China (31590820, 31590823), the NSFC-Joint Foundation of Yunnan Province (U1802287), and the National Natural Science Foundation of China (31872673).

AUTHOR CONTRIBUTIONS

Y.J., J.Y. and T.Y. designed the study. Y.J., C.L. and J.Y. collected the data. Z.Y., L.Y., and Z.H. analysed the data. Y.J. wrote the manuscript. H.W. and T.Y. discussed the results and revised the manuscript.

DATA AVAILABILITY

All sequence data generated in current study have been deposited with the NCBI GenBank database (accession number in Table 1).

ORCID

Yunheng Ji https://orcid.org/0000-0002-1116-7485
Tingshuang Yi https://orcid.org/0000-0001-7093-9564

REFERENCES

Acosta, M. C., & Premoli, A. C. (2010). Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae). *Molecular Phylogenetics and Evolution*, *54*, 235–242. https://doi.org/10.1016/j.ympev.2009.08.008

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). Spades: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*, 455–477. https://doi.org/10.1089/cmb.2012.0021

Barrett, C. F., Davis, J. I., Leebens-Mack, J., Conran, J. G., & Stevenson, D. W. (2013). Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics*, *29*, 65–87. https://doi.org/10.1111/j.1096-0031.2012.00418.x

Chen, W., Kui, L., Zhang, G., Zhu, S., Zhang, J., Wang, X., … Yang, S. (2017). Whole-genome sequencing and analysis of the Chinese herbal plant *Panax notoginseng*. *Molecular Plant*, *10*, 899–902. https://doi.org/10.1016/j.molp.2017.02.010

Choi, H. K., & Wen, J. (2000). A phylogenetic analysis of *Panax* (Araliaceae): Integrating cpDNA restriction site and nuclear rDNA ITS sequence data. *Plant Systematics & Evolution*, *224*, 109–120. https://doi.org/10.1007/BF00985269

Coissac, E., Hollingsworth, P. M., Lavergne, S., & Taberlet, P. (2016). From barcodes to genomes: Extending the concept of DNA barcoding. *Molecular Ecology*, *25*, 1423–1428. https://doi.org/10.1111/mec.13549

Collins, R., & Cruickshank, R. (2013). The seven deadly sins of DNA barcoding. *Molecular Ecology Resources*, *13*, 969–975.

Costa, F. O., Dewaard, J. R., Boutillier, J., Ratnasingham, S., Dooh, R. T., Hajibabaei, M., & Hebert, P. D. N. (2007). Biological identifications through DNA barcodes: The case of the Crustacea. *Canadian Journal of Fisheries & Aquatic Sciences*, *64*, 272–295. https://doi.org/10.1139/f07-008

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin*, *19*, 11–15.

Folk, R. A., Soltis, P. S., Soltis, D. E., & Guralnick, R. (2018). New prospects in the detection and comparative analysis of hybridization in the tree of life. *American Journal of Botany*, *105*, 364–375. https://doi.org/10.1002/ajb2.1018

Fu, C.-N., Wu, C.-S., Ye, L.-J., Mo, Z.-Q., Liu, J., Chang, Y.-W., … Gao, L.-M. (2019). Prevalence of isomeric plastomes and effectiveness of plastome super-barcodes in yews (*Taxus*) worldwide. *Scientific Reports*, *9*, 2773. https://doi.org/10.1038/s41598-019-39161-x

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., & Hebert, P. D. N. (2006). DNA barcodes distinguish species of tropical *Lepidoptera*. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 968–971. https://doi.org/10.1073/pnas.0510466103

Hollingsworth, P. M. (2011). Refining the DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19451–19452. https://doi.org/10.1073/pnas.1116812108

Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., … Little, D. P. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 12794–12797. https://doi.org/10.1073/pnas.0905845106

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS ONE*, *6*, e19254. https://doi.org/10.1371/journal.pone.0019254

Hollingsworth, P. M., Li, D. Z., Michelle, V. D. B., & Twyford, A. D. (2016). Telling plant species apart with DNA: From barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*, 20150338. https://doi.org/10.1098/rstb.2015.0338

Jiang, P., Shi, F. X., Li, M. R., Liu, B., Wen, J., Xiao, H. X., & Li, L. F. (2018). Positive selection driving cytoplasmic genome evolution of the medicinally important ginseng plant genus *Panax*. *Frontiers in Plant Science*, *9*, 359. https://doi.org/10.3389/fpls.2018.00359

Kan, N. C., & Cronk, Q. (2008). Botany without borders, barcoding in focus. *Molecular Ecology*, *17*, 5175–5176. https://doi.org/10.1111/j.1365-294X.2008.03972.x

Kane, N., Sveinsson, S., Dempewolf, H., Yang, J. Y., Zhang, D., Engels, J. M. M., & Cronk, Q. (2012). Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. *American Journal of Botany*, *99*, 320–329. https://doi.org/10.3732/ajb.1100570

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kim, D.-W., Yoo, W. G., Park, H. C., Yoo, H. S., Kang, D. W., Jin, S. D., … Lim, J. (2012). DNA barcoding of fish, insects, and shellfish in Korea. *Genomics & Informatics*, *10*, 206–211. https://doi.org/10.5808/GI.2012.10.3.206

Kim, N. H., Jayakodi, M., Lee, S. C., Choi, B. S., Jang, W., Lee, J., … Yang, T. J. (2018). Genome and evolution of the shade-requiring medicinal herb *Panax ginseng*. *Plant Biotechnology Journal*, *16*, 1904–1917. https://doi.org/10.1111/pbi.12926

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lee, C., & Wen, J. (2004). Phylogeny of *Panax* using chloroplast *trnC-trnD* intergenic region and the utility of *trnC-trnD* in interspecific studies of plants. *Molecular Phylogenetics & Evolution*, *31*, 894–903. https://doi.org/10.1016/j.ympev.2003.10.009

Li, D.-Z., Gao, L.-M., Li, H.-T., Wang, H., Ge, X.-J., Liu, J.-Q., … Duan, G.-W. (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19641–19646. https://doi.org/10.1073/pnas.1104551108

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., & Chen, S. (2015). Plant DNA barcoding: From gene to genome. *Biological Reviews*, *90*(1), 157–166. https://doi.org/10.1111/brv.12104

Liu, C., Yang, Z., Yang, L., Yang, J., & Ji, Y. (2018). The complete plastome of *Panax stipuleanatus*: Comparative and phylogenetic analyses of the genus *Panax* (Araliaceae). *Plant Diversity*, *40*, 265–276. https://doi.org/10.1016/j.pld.2018.11.001

Liu, J., Shi, L., Han, J., Li, G., Lu, H., Hou, J., … Downie, S. R. (2014). Identification of species in the angiosperm family Apiaceae using DNA barcodes. *Molecular Ecology Resources*, *14*, 1231–1238. https://doi.org/10.1111/1755-0998.12262

Manzanilla, V., Kool, A., Nhat, L. N., Van, H. N., Thu, H. L. T., & de Boer, H. J. (2018). Phylogenomics and barcoding of *Panax*: Toward the identification of ginseng species. *BMC Evolutionary Biology*, *18*, 44. https://doi.org/10.1186/s12862-018-1160-y

Nock, C. J., Waters, D. L. E., Edwards, M. A., Bowen, S. G., Rice, N., Cordeiro, G. M., & Henry, R. J. (2011). Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal*, *9*, 328–333. https://doi.org/10.1111/j.1467-7652.2010.00558.x

Parks, M., Cronn, R., & Liston, A. (2009). Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology*, *7*, 84. https://doi.org/10.1186/1741-7007-7-84

Patel, R. K., & Jain, M. (2012). NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS ONE*, *7*, e30619. https://doi.org/10.1371/journal.pone.0030619

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., … Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*, 595–609. https://doi.org/10.1080/10635150600852011

Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, *53*, 793–808. https://doi.org/10.1080/10635150490522304

Posada, D., & Crandall, K. A. (1998). MODELTEST: Testing the model of DNA substitution. *Bioinformatics*, *14*, 817–818. https://doi.org/10.1093/bioinformatics/14.9.817

Rieseberg, L. H., & Soltis, D. E. (1991). Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants*, *5*, 65–84. https://doi.org/10.1007/BF00021248

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Ruhsam, M., Rai, H. S., Mathews, S., Ross, G., Graham, S. W., Raubeson, L. A., … Hollingsworth, P. M. (2015). Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in *Araucaria*? *Molecular Ecology Resources*, *15*, 1067–1078. https://doi.org/10.1111/1755-0998.12375

Schattner, P., Brooks, A. N., & Lowe, T. M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research*, *33*, W686–W689. https://doi.org/10.1093/nar/gki366

Shi, F.-X., Li, M.-R., Li, Y.-L., Jiang, P., Zhang, C., Pan, Y.-Z., … Li, L.-F. (2015). The impacts of polyploidy, geographic and ecological isolations on the diversification of *Panax* (Araliaceae). *BMC Plant Biology*, *15*, 297. https://doi.org/10.1186/s12870-015-0669-0

Smith, M. A., Poyarkov, N. A., & Hebert, P. D. N. (2008). DNA barcoding: CO1 DNA barcoding amphibians: Take the chance, meet the challenge. *Molecular Ecology Resources*, *8*, 235–246. https://doi.org/10.1111/j.1471-8286.2007.01964.x

Soltis, D. E., Johnson, L. A., & Looney, C. (1996). Discordance between ITS and chloroplast topologies in the *Boykinia* group (Saxifragaceae). *Systematic Botany*, *21*, 169–176. https://doi.org/10.2307/2419746

Soltis, D. E., & Kuzoff, R. K. (1995). Discordance between nuclear and chloroplast phylogenies in the heuchera group (Saxifragaceae). *Evolution*, *49*(4), 727–742. https://doi.org/10.2307/2410326

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analysis with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Straub, S. C. K., Fishbein, M., Cronn, R., Liston, A., Parks, M. M., & Weitemier, K. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany*, *99*(2), 349–364. https://doi.org/10.3732/ajb.1100335

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729. https://doi.org/10.1093/molbev/mst197

Tontifilippini, J., Nevill, P. G., Dixon, K., & Small, I. (2017). What can we do with 1000 plastid genomes? *The Plant Journal*, *90*, 808–818. https://doi.org/10.1111/tpj.13491

Turner, B., Paun, O., Munzinger, J., Chase, M. W., & Samuel, R. (2016). Sequencing of whole plastid genomes and nuclear ribosomal DNA of *Diospyros* species (Ebenaceae) endemic to new Caledonia: Many species, little divergence. *Annals of Botany*, *117*, 1175–1185. https://doi.org/10.1093/aob/mcw060

Vargas, O. M., Ortiz, E. M., & Simpson, B. B. (2017). Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-andean diversification (asteraceae: Astereae: *Diplostephium*). *The New Phytologist*, *214*, 1736–1750. https://doi.org/10.1111/nph.14530

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*, 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Wen, J., & Zimmer, E. A. (1996). Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): Inferences from ITS sequences of nuclear ribosomal DNA. *Molecular Phylogenetics & Evolution*, *6*, 167–177. https://doi.org/10.1006/mpev.1996.0069

Wyman, S. K., Jansen, R. K., & Boore, J. L. (2004). Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*, *20*, 3252–3255. https://doi.org/10.1093/bioinformatics/bth352

Xiang, Q. B., & Lowry, P. P. (2007). Panax. In Z. Y. Wu, & P. H. Raven (Eds.), *Flora of China* (Vol. 13, pp. 489–491). Beijing, China: Science Press.

Yang, J. B., Tang, M., Li, H. T., Zhang, Z. R., & Li, D. Z. (2013). Complete chloroplast genome of the genus *Cymbidium*: Lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evolutionary Biology*, *13*, 84. https://doi.org/10.1186/1471-2148-13-84

Zeng, C. X., Hollingsworth, P. M., Yang, J., He, Z. S., Zhang, Z. R., Li, D. Z., & Yang, J. B. (2018). Genome skimming herbarium specimens for DNA barcoding and phylogenomics. *Plant Methods*, *14*, 43. https://doi.org/10.1186/s13007-018-0300-0

Zhu, S., Fushimi, H., Cai, S., & Komatsu, K. (2003). Phylogenetic relationship in the genus *Panax*: Inferred from chloroplast *trnK* gene and nuclear 18S rRNA gene sequences. *Planta Medica*, *69*, 647–653. https://doi.org/10.1055/s-2003-41117

Zuo, Y., Chen, Z., Kondo, K., Funamoto, T., Wen, J., & Zhou, S. (2011). DNA barcoding of *Panax* species. *Planta Medica*, *77*, 182–187. https://doi.org/10.1055/s-0030-1250166

Zuo, Y., Wen, J., Ma, J., & Zhou, S. (2015). Evolutionary radiation of the *Panax bipinnatifidus* species complex (Araliaceae) in the Sino-Himalayan region of eastern Asia as inferred from AFLP analysis. *Journal of Systematics and Evolution*, *53*, 210–220. https://doi.org/10.1111/jse.12119

Zuo, Y., Wen, J., & Zhou, S. (2017). Intercontinental and intracontinental biogeography of the eastern Asian–eastern North American disjunct *Panax* (the ginseng genus, Araliaceae), emphasizing its diversification processes in eastern Asia. *Molecular Phylogenetics and Evolution*, *117*, 60–74. https://doi.org/10.1016/j.ympev.2017.06.016

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Ji Y, Liu C, Yang Z, et al. Testing and using complete plastomes and ribosomal DNA sequences as the next generation DNA barcodes in *Panax* (Araliaceae). *Mol Ecol Resour*. 2019;19:1333–1345. https://doi.org/10.1111/1755-0998.13050