

From nine to two: Ultra-barcoding develops a rigorous species delimitation framework in *Paris* section *Axiparis* (Melanthiaceae)

Yunheng Ji¹, Jin Yang¹, Jacob Landis², Shuying Wang³, Zhenyan Yang¹, Lei Jin¹, Li Li³, Jun-Bo Yang¹, and Ting Shuang Yi⁴

¹Kunming Institute of Botany Chinese Academy of Sciences

²Cornell University

³Kunming Institute of Botany Key Laboratory of Plant Diversity and Biogeography of East Asia

⁴Kunming Institute of Botany, Chinese Academy of Sciences

January 11, 2021

Abstract

Species delimitation is a difficult task in traditional morphology-based taxonomy. The ultra-barcoding approach, which uses whole plastid genomes (plastomes) and nuclear ribosomal DNA (nrDNA) regions as extended DNA barcodes for species identification and delimitation, has been recommended as one of the candidate techniques for plant barcoding 2.0. Yet, the efficacy of this approach in delineating species boundaries remains poorly understood. Here, we attempt to decipher species delimitation in the taxonomically challenging clade, *Paris* L. section *Axiparis* H. Li, using phylogenetic inference and multiple sequence-based species delimitation methods (ABGD, SDP, and mPTP) utilizing complete plastomes and nrDNA clusters from multiple accessions per described species. The results suggest that only two species-level taxonomic units that possess not only morphological uniqueness but also genetic distinctiveness and evolutionary independence can be recognized in *P. sect. Axiparis*. Therefore, previous taxonomic work overemphasized minor intraspecific morphological differences to establish species, and thus resulted in proliferation of as many as seven synonyms in this clade. Inferred from ultra-barcoding analyses, we propose a taxonomic revision of *P. sect. Axiparis* which will help inform future decisions regarding species conservation in the commercially valuable and severely threatened genus *Paris*. This case study indicates that the ultra-barcoding approach has great promise for developing a rigorous species delimitation framework that will facilitate credible taxonomic revision especially in taxonomically difficult plant taxa.

From nine to two: Ultra-barcoding develops a rigorous species delimitation framework in *Paris* section *Axiparis* (Melanthiaceae)

Running title: Species delimitation in *Paris* section *Axiparis*

Yunheng Ji^{1,2*}, Jin Yang^{1,3}, Jacob B. Landis⁴, Shuying Wang^{1,3}, Zhenyan Yang¹, Lei Jin¹, Li Li¹, Junbo Yang^{5*}, Tingshuang Yi^{5*}

¹CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, P. R. China

²Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Population, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, P. R. China

³School of Life Science, Yunnan University, Kunming, Yunnan 650009, P. R. China

⁴School of Integrative Plant Science, Section of Plant Biology and the L. H. Bailey Hortorium, Cornell University, Ithaca, NY14850, USA

⁵Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, P. R. China

***Corresponding author email :** jiyh@mail.kib.ac.cn, jbyang@mail.kib.ac.cn, tingshuangyi@mail.kib.ac.cn

Abstract

Species delimitation is a difficult task in traditional morphology-based taxonomy. The ultra-barcoding approach, which uses whole plastid genomes (plastomes) and nuclear ribosomal DNA (nrDNA) regions as extended DNA barcodes for species identification and delimitation, has been recommended as one of the candidate techniques for plant barcoding 2.0. Yet, the efficacy of this approach in delineating species boundaries remains poorly understood. Here, we attempt to decipher species delimitation in the taxonomically challenging clade, *Paris* L. section *Axiparis* H. Li, using phylogenetic inference and multiple sequence-based species delimitation methods (ABGD, SDP, and mPTP) utilizing complete plastomes and nrDNA clusters from multiple accessions per described species. The results suggest that only two species-level taxonomic units that possess not only morphological uniqueness but also genetic distinctiveness and evolutionary independence can be recognized in *P. sect. Axiparis*. Therefore, previous taxonomic work overemphasized minor intraspecific morphological differences to establish species, and thus resulted in proliferation of as many as seven synonyms in this clade. Inferred from ultra-barcoding analyses, we propose a taxonomic revision of *P. sect. Axiparis* which will help inform future decisions regarding species conservation in the commercially valuable and severely threatened genus *Paris*. This case study indicates that the ultra-barcoding approach has great promise for developing a rigorous species delimitation framework that will facilitate credible taxonomic revision especially in taxonomically difficult plant taxa.

Key words

DNA barcoding, plastome, ribosomal DNA, species boundary, taxonomic revision, synonyms

Introduction

Species delimitation is the crucially important first step for designing research in many fields of biology (Mace, 2004). Traditionally, species boundaries are defined by taxonomists based on the analysis of morphological variation (Lewin, 1981; Henderson et al., 2005), which has resulted in massive disagreements over species identification and delimitation due to either phenotypic plasticity or lack of taxonomically robust morphological characters at the species level (Hebert et al., 2003; de Queiroz, 2007). To compensate, additional data types, such as molecular and ecological profiles, are needed to explicitly decipher species boundaries (e.g., Sites and Crandall, 1997; Sites and Marshall, 2003; Ellis et al., 2006; Bickford et al., 2007; Eisenring et al., 2016) especially for taxonomically challenging taxa (e.g., Duminil and Di Michele, 2009; Su et al., 2015; Lambert et al., 2017; Cheng et al., 2020; Ma et al., 2020).

Analysis of DNA sequence variation can provide useful information for identifying and delineating species (Brower et al., 1996; Hebert et al., 2003; Pons et al., 2006; Duminil et al., 2012; Puillandre et al., 2012). DNA barcoding, a technique that involves the standardized utilization of one or a few sequence regions for species identification and discrimination (Hebert et al., 2003; Kress et al., 2005; Hollingsworth, 2011; Hollingsworth et al., 2009, 2011, 2016), has been widely used as a molecular tag for species diagnosis and delimitation (e.g., Hebert et al., 2004; Liu et al., 2011; Kekkonen and Hebert, 2014; Mutanen et al., 2015). Despite the mitochondrial *COI* gene acting as the standard DNA barcodes in animals (Hebert et al., 2003, 2004; Kekkonen and Hebert, 2014; Mutanen et al., 2015), reliable and accurate species identification and delimitation based on standard DNA barcodes (i.e. *rbcL*, *matK*, *trnH-psbA*, and ITS) remains problematic in plants, especially in lineages that have experienced rapid radiations or complicated evolutionary histories (Hollingsworth et al., 2009, 2011, 2016; Hollingsworth, 2011; Coissac et al., 2016).

With next-generation DNA sequencing (NGS) technologies, the concept of DNA barcoding for plant species

has been extended from a single or few sequence regions to genome-wide sequence variations (Li et al., 2015; Coissac et al., 2016; Hollingsworth et al., 2016). The development of extended DNA barcoding approaches for plant species discrimination has been referred to as ‘plant barcoding 2.0’ (Hollingsworth et al., 2016). The ultra-barcoding approach, which employs both whole plastid genomes (plastomes) and nuclear ribosomal DNA (nrDNA) clusters as extended DNA barcodes for resolution of species identification and delimitation (Nock et al., 2011), is one of the recommended techniques for plant barcoding 2.0 (Hollingsworth et al., 2016). This approach yields a tremendous amount of sequence variation, allowing far more sensitivity than standard DNA-barcoding (Nock et al., 2011; Kane et al., 2012; Dodsworth, 2015; Li et al., 2015; Hollingsworth et al., 2016), and thus providing far more genetic information than standard DNA barcodes to improve resolution in species identification (Ruhsam et al., 2015; Firetti et al., 2017; Fu et al., 2019; Ji et al., 2019a, 2020; Knope et al., 2020; Šlipiko et al., 2020). Compared with Restriction-site Associated DNA sequencing (Baird et al., 2008), another recommended approach for plant barcoding 2.0 (Hollingsworth et al., 2016) that has been extensively used to generate genomic data for species delimitation in recent years (e.g., Wu et al., 2018; Donkpegan et al., 2020; Li et al., 2020; Ma et al., 2020; Zhou et al., 2020), the above mentioned ultra-barcoding approach is seldomly applied to investigate species boundaries. As a result, its efficacy in delineating species is still poorly understood.

This study focuses on *Paris* Linn. section *Axiparis* H. Li (Melanthiaceae), a well-supported monophyletic clade including nine described species restricted to southwest China and the Himalayas (Li, 1998; Ji et al., 2006, 2019b; Huang et al., 2016). Discovery of the first species (*Paris vaniotii*) of this group, dating back to 1906, was based on collections from Guizhou, southwestern China (Léveillé, 1906), with the remaining species have been described since the 1980s (Takhtajan, 1983; Li, 1984, 1992; He, 1990; Ji et al., 2017; Xu et al., 2019; Yang et al., 2019). The rapidly increasing species count in the last 40 years led to considerable taxonomic confusion in *P. sect. Axiparis*. After critical examination of types and specimens assigned to these species, it was found that, except for slight differences in leaf shape and size (Figure 1), they exhibit high levels of similarity in flower, fruit, and seed morphology. Additionally, leaf shape and size of foliar displays have high levels of intraspecific variations that sometimes far exceed the divergence used for species diagnosis. High variability of key diagnostic morphological characters makes species delimitation and correct species identification a particularly difficult task in this clade.

Due to the above-mentioned aspects, *P. sect. Axiparis* is in great need of taxonomic revision, and thus provides an ideal model to evaluate the usefulness of ultra-barcoding approach for species delimitation. In this study, we attempt to use the ultra-barcoding approach to establish a basic understanding of species boundaries in this taxonomically challenging lineage. Specifically, we employed a genome skimming method (Straub et al., 2012) to recover complete plastomes and nuclear rDNA sequences by sampling multiple accessions per species within *P. sect. Axiparis*. Based on phylogenetic inference and multiple sequence-based species delimitation methods, we test whether the ultra-barcoding approach could develop a rigorous species delimitation framework in *P. sect. Axiparis*. Inferred from the current study, we discuss the prospects on the ultra-barcoding approach as a tool to explore species boundaries in taxonomically challenging plant groups.

Materials and methods

Plant samples, shotgun sequencing, assembly and annotation

In total, 51 individual plants from all nine described species of *P. sect. Axiparis* were sampled, specifically *Paris axialis* (4 accessions), *P. dulongensis* (8 accessions), *P. forrestii* (10 accessions), *P. guizhouensis* (3 accessions), *P. lihengiana* (3 accessions), *P. rugosa* (8 accessions), *P. tunchongensis* (7 accessions), *P. vaniotii* (5 accessions), and *P. variabilis* (3 accessions) (Table S1). Plant material was collected from the field, and voucher specimens were deposited at the herbarium of Kunming Institute of Botany, Chinese Academy of Science (KUN). For each species, multiple individuals within a species representing different localities were included. The sampling strategy allows testing for species-level monophyly and to explore species boundaries.

Total genomic DNA for each accession was isolated from ~10 mg silica gel dried leaf tissues collected in

the field using CTAB method (Doyle and Doyle, 1987). Approximately 5 µg of purified genomic DNA was used to construct shotgun libraries with a TruSeq DNA Sample Prep Kit (Illumina, Inc., USA) following the manufacturer's instructions. Low coverage genome sequencing was performed on the Illumina HiSeq 2500 system using paired end (2×150) reads. Raw reads were filtered to remove adaptors and low quality read using the NGS QC Toolkit (Patel and Jain, 2012) with default parameters.

We used a reference-guided method for plastome and nrDNA assembly. The published complete plastome (MN125565) and nrDNA sequence (MN174877) of *P. forrestii* was used as the reference to assembly newly sequenced plastomes and nrDNA sequences using GetOrganelle, a pipeline developed by Jin et al. (2020). Briefly, the pipeline consists of a customized python script, which conducts BLAST to search read mapping to the reference plastome, was used to enrich plastid-like reads, and assembled them into contigs with Bowtie v2.2.6 (Langmead et al., 2012), with default parameters and preset options. Contigs were iteratively extended to recover whole plastomes in SPAdes v3.10.1 (Bankevich et al., 2012). For nrDNA assembly, we first excluded all plastid-like reads, and used the steps described above to generate the nrDNA sequence (including 26S, 18S and 5.8S ribosomal RNA genes and the internal transcribed spacers) of each accession.

The newly assembled plastomes were annotated with the Dual Organellar Genome Annotator database (Wyman et al., 2004). The annotation of protein-coding genes was further confirmed with a BLAST search against the NCBI protein database. Genes putatively annotated as transfer RNA (tRNA) were further verified by tRNAscan-SE 1.21 (Schattner et al., 2005) with default parameters. The boundary of the large-single copy (LSC), small-single copy (SSC), and inverted-repeat (IR) regions for each plastome were visually examined and manually adjusted with Geneious V10.2.3 (Kearse et al., 2012). For nrDNA annotation, the ribosomal RNA genes and boundaries with the intergenic transcribed spacer (ITS) regions were annotated and defined by comparison with the reference in Geneious v10.2.3 (Kearse et al., 2012).

Phylogenetic analyses

Using inferred phylogenies, we tested whether the nine nominal species within *P. sect. Axiparis* were monophyletic units. The complete plastome and nrDNA sequences were aligned using MAFFT v7.450 (Katoh and Standley, 2013). For each dataset, phylogenetic relationships were inferred using both maximum likelihood (ML) and Bayesian inference (BI) methods. For each analysis, *Paris thibetica* representing *P. section Thibeticae*, the closest relative of *P. sect. Axiparis* (Ji et al., 2019b), was selected as the outgroup. The ML analysis was performed with RAxML-HPG BlackBox v8.1.24 (Stamatakis, 2006), using the GTRCAT sequence substitution model. The best-scoring ML tree for each dataset was produced with 1,000 bootstrap (BS) replicates to provide support values for each node. The BI tree was inferred using MrBayes v3.2 (Ronquist and Huelsenbeck, 2003), with the sequence substitution model GTR + G (plastomes) and GTR+I+G (nrDNA), which were recommended by Modeltest v3.7 (Posada and Crandall, 1998) using the Akaike information criterion (Posada and Buckley, 2004). Two independent Markov Chain Monte Carlo (MCMC) runs were performed with one million generations, sampling every 100 generations, with the initial 25% of the sampled trees as burn-in. Posterior probability (PP) values were computed based on the remaining trees.

Sequence-based species delimitation

Based on complete plastomes and nrDNA sequences, we used three different approaches to explore putative species boundaries within *P. sect. Axiparis*: the distance-based method automatic barcode gap discovery (ABGD; Puillandre et al., 2012), the coalescence-based method species delimitation plugin (SDP; Masters et al., 2011) in Geneious v10.2.3 (Kearse et al., 2012), and the tree-based method multi-rate Poisson tree processes model (mPTP; Kapli et al., 2017). The ABGD analysis was conducted with the online server (<http://www.wabi.snv.jussieu.fr/public/abgd/abgdweb.html>) with default settings (Pmin=0.001, Pmax=0.1, Steps=10, X=1.5, Nb=20). All three genetic distance models (JC69, K2P, and uncorrected *P*-distances) specified by the program were used. All assignments for intraspecific divergence (*P*) values between 0.001 and 0.100 were recorded. Next, we used the mPTP v0.2.3 algorithm (Kapli et al., 2017), an improvement to PTP (Zhang et al., 2013), to estimate Bayesian support for the putative species boundaries. The mPTP analysis was performed on the web server (<http://species.h-its.org/ptp/>) with standard default settings, using

the inferred ML trees of plastome and nrDNA as inputs since the branch lengths of ML tree represent number of sequence mutations. For each analysis, the MCMC algorithm was conducted for 50 million generations, thinning every 100 generations and a 20% burn-in. Lastly, we performed SDP analyses to test whether species-like taxonomic units proposed by ABGD and mPTP analyses are evolutionarily distinct species. Based on ML and BI trees of plastome and nrDNA datasets, the distinctiveness of these candidate species was estimated using Rosenberg's $P_{(AB)}$: the probability of reciprocal monophyly under a random coalescent model (Rosenberg, 2007), and Rodrigo's $P_{(RD)}$: the probability that a clade has the observed degree of distinctiveness due to a random coalescent process (Rodrigo et al., 2008).

Results

Low coverage genome sequencing and assembly

The summary of low coverage genome sequencing is shown in Table S2, which produced 5.785 to 33.107 million paired-end clean reads per accession. Of these, 70,765 to 1,044,267 and 3,875 to 66,322 were mapped to the reference plastome and nrDNA, respectively. Based on these reads, we assembled the complete plastomes and nrDNA for each sample, with an average sequencing coverage ranging from $67.67\times$ to $1,711.16\times$ and $92.58\times$ to $1,710.43\times$, respectively. The reference-guided assembly generated the complete plastomes of all samples, which possess a typical quadripartite structure, with the sequence length varying from 156,061 to 157,653 bp. The plastomes contain 114 genes, including 80 protein-coding genes, 30 tRNA genes, and four plastid rRNA genes. In addition, the assembly of nrDNA entirely covered 18S, ITS1, 5.8S, ITS2 and 26S regions in all accessions, with the sequence length being either 5,851 or 5,852 bp.

Species-level monophyly

Alignment of the plastome sequences yielded a matrix of 160,681 positions, in which we identified 1,724 variable sites (1.07%) with 1,192 (0.74%) being parsimoniously informative (Table S3). The ML and BI analyses of complete plastomes produced highly similar tree topologies (Figure 2). Two diverging clades were recovered in *P. sect. Axiparis*. Clade I (BS=100%, PP=1.00) comprises accessions of *P. dulongensis*, *P. forrestii*, *P. rugosa*, and *P. tengchongensis*. These species share the morphological similarity of stamens numbering twice as many as the sepal number (two-whorled stamens), and being geographically distributed from the Hengduan Mountains (southwestern China) to the Himalayas (Figure 3). Comparatively, Clade II (BS=100%, PP=1.00) includes accessions of *P. axialis*, *P. guizhouensis*, *P. lihengiana*, *P. vaniotii*, and *P. variabilis*, whose stamen numbers are three times as many as the sepal number (three-whorled stamen). These species are distributed from central China to the Wumeng Mountains in southwestern China (Figure 3). Strikingly, none of the nine nominal species within *P. sect. Axiparis* were recovered as a monophyletic unit in either ML or BI phylogeny.

The alignment of the nrDNA sequences produced 193 variable sites (3.30%) with 112 (1.91%) being parsimoniously informative (Table S3). ML and BI analyses of nrDNA sequences (Figure 4) recovered two major clades within *P. sect. Axiparis*, corresponding to those from the plastome reconstruction but with slightly lower support values (clade I: BS=98%, PP=1.00; clade II: BS=97%, PP=1.00). Similar to the plastome phylogeny, analyses of nrDNA sequences failed to resolve any of the nine nominal species within *P. sect. Axiparis* as a monophyletic unit.

Species delimitation scheme

Complete plastomes and nrDNA sequences of 51 individuals were used for species delimitation analyses. Overall, ABGD, SDP, and mPTP analyses of plastid and nuclear datasets produced highly congruent results that are reflected in the phylogenetic trees (Figure 2 & 4). The ABGD analyses of complete plastomes (Table S4) and nrDNA sequences (Table S5) resulted in a stable count ($n = 2$) of species division with a range of prior intraspecific values (plastomes: $P = 0.001000$ – 0.004642 ; nrDNA: $P = 0.002000$ – 0.007573) with JC69, K2P and P -distances initial and recursive partitions. One proposed species includes individuals of *P. dulongensis*, *P. forrestii*, *P. rugosa*, and *P. tengchongensis*, while the other comprises individuals of *P. axialis*, *P. guizhouensis*, *P. lihengiana*, *P. vaniotii*, and *P. variabilis* (Figure 2 & 4). The mPTP

analyses yielded the same delimitation scheme as ABGD: all individuals were grouped into two species-like entities that coincide with the two clades recovered by our phylogenetic analyses (Figure 2 & 4), and both received strong posterior support (Table S6). Similarly, the SDA analyses showed that either plastome or nrDNA dataset, excluding the outgroup, comprised two putative species (Figure 2 & 4). For each dataset, both proposed species (corresponding to Clade I and Clade II in the inferred phylogenetic tree) were distinct with $P_{(RD)} < 0.05$, and significant with $P_{(AB)} < 10^{-5}$ (Table S7).

Discussion

Species delimitation scenario

The alpha taxonomy of *P. sect. Axiparis* is still not well understood, since nearly all nominal species show subtle morphological differences. Both evolutionary and genetic profiles should be integrated into species delimitation and taxonomic revision (de Queiroz, 1998; Wiley and Mayden, 2000; Sites and Marshall, 2003). With that in mind, this study aims to test whether the ultra-barcoding approach could develop a rigorous species delimitation framework in *P. sect. Axiparis*. To achieve this goal, we not only investigated the species-level monophyly of each described species by sampling multiple individuals per species, but also explore putative species boundaries using multiple sequence-based species delimitation methods.

None of the nine described species within *P. sect. Axiparis* were resolved as a monophyletic unit, suggesting that genetic differentiation is low. Thus, their taxonomic status as separate species needs to be reassessed. By contrast, both plastome and nrDNA phylogenies recovered two clades that possess fairly distinct morphological traits (Figure 2 & 4) and distribution ranges (Figure 3). Notably, the ABGD, SDP, and mPTP analyses consistently recognized the two clades recovered by phylogenetic analyses as species-like taxonomic units (Figure 2 & 4). Briefly, the ABGD analyses partitioned all the samples into two clusters comprised of individuals having two-whorled and (Clade I) three-whorled stamens (Clade II). This implies that there may be two distinct species with significant genetic gaps between them (Puillandre et al., 2012). On the other hand, the mPTP analyses grouped all accessions into two putative species with high posterior delimitation probability, coinciding with the results found in the ABGD analyses. The putative species boundaries proposed by ABGD and mPTP analyses are further validated by the SDP analyses: with $P_{(RD)}$ value < 0.05 and the $P_{(AB)}$ value $< 10^{-5}$, both of the two candidate species can be recognized as evolutionarily distinctive entities (Rosenberg, 2007; Rodrigo et al., 2008). These reinforcing results suggest that only two species-level taxonomic units can be recognized in *P. sect. Axiparis*.

Notably, the two putative species possess not only distinct morphologies but also allopatric distribution ranges, between which gene flow is restricted. Therefore, it is reasonable to recognize them as distinct species under the unified species concept (de Queiroz, 2007), given that the species boundary proposed by our data reflects the unity of morphological uniqueness, genetic distinctiveness, and evolutionary independence. The findings suggest that previous morphology-based taxonomic studies (Li, 1984; He, 1990; Li, 1992; Ji et al., 2017; Xu et al., 2019; Yang et al., 2019) overemphasize intraspecific and minor morphological differences to establish species, therefore leading to the proliferation of synonyms. As a result, it is necessary to reduce *P. axialis*, *P. guizhouensis*, *P. lihengiana*, and *P. variabilis* as the synonyms of *P. vaniotii*, and to merge *P. dulongensis*, *P. rugosa* and *P. tengchongensis* into *P. forrestii*. The taxonomic revision proposed here reduces the number of species in *P. sect. Axiparis* from nine to two.

Conservation Implications

Taxonomy and species conservation are often assumed to be completely interdependent practices (Mace, 2004). Estimating the number of species under threat is an essential step in setting conservation priorities (May, 1988, 1990; Margules and Pressey, 2000; Dirzo and Raven, 2003; Mace et al., 2003). Prioritizing species for natural conservation heavily depends on accurate delineation of species boundaries: well delimited species are fundamental to allow adequate conservation and biodiversity management (Hopkins and Freckleton, 2002; Sites and Marshall, 2003; Mace, 2004). On the contrary, poorly delineated species boundaries usually make it difficult to properly address conservation issues (Rojas, 1992; Sites and Crandall, 1997; Prance et al., 2000; Mace, 2004).

Paris is a plant genus of great conservation concern with the majority of species in the genus facing varying degrees of extinction risk due to overharvesting of wild populations for medicinal purposes (Ji, 2020). For instance, among the nine nominal species of *P. sect. Axiparis*, *P. dulongensis* and *P. lihengiana* are categorized as critically endangered (Qin et al., 2017; Xu et al., 2019); *P. forrestii*, *P. guizhouensis*, *P. rugosa*, *P. tengchongensis* and *P. vaniotii* are classified as endangered (Ji et al., 2017; Qin et al., 2017); *P. axialis* and *P. variabilis* are identified as Vulnerable (Qin et al., 2017; Yang et al., 2019). As suggested by our ultra-barcoding analyses, prior taxonomic studies generally over-split species, and thus proliferated as many as seven synonyms in *P. sect. Axiparis*. In conservation practice, the poorly delineated species boundaries within *P. sect. Axiparis* may lead to the misallocation of limited conservation resource to widespread species that are not at high risk of extinction (Agapow et al., 2004; Mace, 2004; Joppa et al., 2011). In this study, the ultra-barcoding approach depicts a more accurate picture of the species diversity of *P. sect. Axiparis*, which will help to better inform future decisions regarding the species conservation in the genus *Paris*.

On the other hand, commercial collection of rhizomes, seeds and seedlings from wild populations pose the most serious threat to the survival and future conservation status of *Paris* species. With the depletion of wild resources, some of the rare and endangered species are getting collected as commercial products (Ji, 2020). Although rarity alone does not mean a species is at a high risk of extinction, species with small populations are undoubtedly more prone to extinct when heavily disturbed by anthropogenic activities (Burgman, 2002). To protect these species, it is necessary to take an effective way to end entirely the collection of threatened *Paris* species from the wild. However, the rhizome, seed and seedling morphologies exhibit high levels of interspecific similarity, and thus it is often difficult to distinguish these products harvested from different *Paris* species. This makes the commercial collection and trade of *Paris* products difficult to monitor and manage.

As revealed by previous studies (e.g., Kane et al., 2012; Ji et al., 2019a, 2020; Knope et al., 2020) and this study, the ultra-barcoding approach performs well in discriminating species. More importantly, with advances in NGS technologies, it has become much easier to generate complete plastomes and nrDNA sequences even using trace and highly degraded genomic DNA to construct sequencing libraries (Zeng et al., 2018). Therefore, ultra-barcoding is likely an effective tool for the identification of plant products to control the commercial collection and trade of rare and endangered *Paris* species.

Prospects on ultra-barcoding approach for species delimitation

The advancement of molecular-based approaches brought about great progress in species identification and delimitation (Wiley and Mayden, 2000; Sites and Marshall, 2003; de Queiroz, 2007; Hollingsworth et al., 2016). Concomitantly, a wide variety of species delimitation methods have been developed, especially over the past 15 years (e.g., Rannala and Yang, 2003; Yang and Rannala, 2010; Puillandre et al., 2012; Ence and Carstens, 2011; Masters et al., 2011; Zhang et al., 2013). As a result, it has been widely accepted that multiple methods should be simultaneously employed to develop a robust species delimitation framework, given that using different delimitation approaches allows accommodations for the weaknesses of each approach (e.g., Hebert et al., 2004; Aguilar et al., 2013; Kekkonen and Hebert, 2014; Mutanen et al., 2015; Fujita et al., 2012). Unfortunately, previous studies using a single or few sequence regions (e.g. the standard DNA barcodes) indicated that employing multiple methods on the same data always produced incongruent delimitation proposals (e.g., Camargo et al., 2012; Mutanen et al., 2015; Jacob et al., 2018; Jirapatrasilp et al., 2019) mainly due to the inadequacies of available genetic information to properly delineate species boundaries (Carstens et al., 2013; Jacob et al., 2018; Sukumaran and Knowles, 2017). Therefore, recent studies have attempted to use complete plastomes and nrDNA sequences for species delimitation (e.g. Ruhsam et al., 2015; Ji et al., 2020; Knope et al., 2020). However, such studies did not employ multiple delimitation approaches, but inferred tree topology alone and under the premise of reciprocal monophyly to explore species boundaries. Factors such as recent diversification, radiative speciation, and restricted intraspecific gene flow may result in the absence of reciprocal monophyly among closely related species, and thus not tracking species boundaries (Ruhsam et al., 2015; Hollingsworth et al., 2016), which bias delimitation schemes if only phylogenetic inference is consulted. Therefore, these studies may provide little precise information to evaluate the efficacy

of either plastome or nrDNA barcoding approach for species delimitation.

Using multiple sequence-based species delimitation methods (ABGD, SDP, and mPTP), here we not only investigate the performance of the ultra-barcoding approach in delineating species boundaries but also examine to what extent these methods generate congruent delimitation schemes. Although members of *P. sect. Axiparis* represent recently diverged entities because their origins have been estimated to no earlier than the late Miocene (Ji et al., 2019b), these delimitation analyses generated highly congruent schemes in delineating species boundaries (Figure 2 & 4). Based on the results, the proposed taxonomic revision establishes a clear-cut species boundary in *P. sect. Axiparis*. The findings suggest that both complete plastomes and nrDNA can provide more genetic information than standard DNA barcodes to robustly delineate species boundaries. This study indicates that the ultra-barcoding approach has great promise for developing a credible species delimitation framework in taxonomically difficult groups to provide guidance for taxonomic revision. In addition, the approach offers great advantages in reliable species identification and delimitation, which are important for formulating adequate conservation and biodiversity management for those groups with great conservation concerns. As discussed above, relying on reciprocal monophyly alone to delineate species boundaries likely produces bias delimitation schemes. The development of some coalescence-based species delimitation methods, such as General Mixed Yule-Coalescent (GYMC; Fujisawa and Barraclough, 2013), species tree estimation using DNA sequences from multiple loci (STACEY; Jones, 2017), Bayesian Phylogenetics and Phylogeography (BPP; Rannala and Yang, 2013), and spedeSTEM (Ence and Carstens, 2011), may provide the solution to this issue since these methods do not require reciprocal monophyly of any species in focus within gene tree. Unfortunately, we tried to use GMYC and STACEY to analyze our data, but found that the BEAST runs using either complete plastomes or nrDNA sequence to generate tree outputs ran for several weeks but failed to produce results. In addition, both BPP and spedeSTEM are only suitable for processing genomic segments no longer than 500 or 1,000 bp (Ence and Carstens, 2011; Rannala and Yang, 2013), and thus cannot be used to analyze ultra-barcoding data. Accordingly, there are still some technical limitations to applying the ultra-barcoding approach to species delimitation on a large scale. We look forward to the early development of such coalescence-based method that enable to delineate species boundaries with ultra-barcoding data in a time-saving manner.

Acknowledgements

We are grateful to Chengjin Yang, Guohua Zhou, Ren Zhao, Tingzhou Zhao, Zhangming Wang, Fajian Yang, Juxiang Li, Changming Yang, Zhiping Zhang, Yongxing Xiong, Guohai He, Yulong Li, Fengtian Lin and Qiang Zhao for their helps in sample collection. This study is supported by the National Natural Science Foundation of China (31872673), the NSFC-Joint Foundation of Yunnan Province (U1802287), and a grant of the Large-scale Scientific Facilities of the Chinese Academy of Sciences (No. 2017-LSF-GBOWS-02).

References

- Agapow, P. M., Bininda-Emonds, O. R., Crandall, K. A., Gittleman, J. L., Mace, G. M., Marshall, J. C., & Purvis, A. (2004). The impact of species concept on biodiversity studies. *The Quarterly Review of Biology*, 79, 161–179. doi: 10.1086/383542
- Aguilar, C., Wood, P., Cusi, J. C., Guzman, A., Huari, F., Lundberg, M., . . . Sites, J. W., Jr. (2013). Integrative taxonomy and preliminary assessment of species limits in *Liolaemus walker* complex (Squamata, Liolaemidae) with descriptions of three new species from Peru. *ZooKeys*, 364, 47–91. doi: 10.3897/zookeys.364.6109
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., . . . Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, 3, e3376. doi: 10.1371/journal.pone.0003376
- 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. doi: 10.1089/cmb.2012.0021

- Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., ... Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* , 22 , 148–155. doi: 10.1016/j.tree.2006.11.004
- Brower, A. V. Z., DeSalle, R., & Vogler, A. (1996). Gene trees, species trees and systematics: a cladistic perspective. *Annual Review of Ecology and Systematics* , 27 , 423–450. doi: 10.1146/annurev.ecolsys.27.1.423
- Burgman, M. A. (2002). Are list threatened plant species actually at risk? *Australian Journal of Botany* , 50 , 275. doi: 10.1071/BT01052_CO
- Cai, C.-N., Ma, H., Ci, X.-Q., Conran, J. G. & Li, J. (2020). Comparative phylogenetic analyses of Chinese *Horsfieldia* (Myristicaceae) using complete chloroplast genome sequences. *Journal of Systematics and Evolution* , doi: 10.1111/jse.12556
- Camargo, A., Morando, M., Avila, L. J., & Sites, J. W., Jr. (2012). Species delimitation with ABC and other coalescent-based methods: A test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwini* complex (Squamata: Liolaemidae). *Evolution* , 66 , 2834–2849. doi: 10.1111/j.1558-5646.2012.01640.x
- Carstens, B. C., & Satler, J. D. (2013). The carnivorous plant described as *Sarracenia alata* contains two cryptic species. *Biological Journal of the Linnean Society* , 109 , 737–746. doi: 10.1111/bij.12093
- Cheng, Z., Shu, H., Zhang, S., Luo, B.-S., Gu, R.-H., Zhang, R.-F., ... Long, C.-L. (2020). From folk taxonomy to species confirmation of *Acorus* (Acoraceae): evidences based on phylogenetic and metabolomic analyses. *Frontiers in Plant Science* , 11 , 965. doi: 10.3389/fpls.2020.00965
- 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. doi: 10.1111/mec.13549
- de Queiroz, K. (1998). The General Lineage Concept of Species, Species Criteria, and the Process of Speciation. In *Endless Forms: Species and Speciation* , 57–75. Oxford University Press.
- de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology* , 56 , 879–886. doi: 10.1080/10635150701701083
- Dirzo, R., & Raven, P. H. (2003). Global state of biodiversity and loss. *Annual Review of Environment and Resources* , 28 , 137–167. doi: 10.1146/annurev.energy.28.050302.105532
- Dodsworth, S. (2015). Genome skimming for next-generation biodiversity analysis. *Trends Plant Science* , 20 , 525–527. doi: 10.1016/j.tplants.2015.06.012
- Donkpegan, A. S. L., Doucet, J. L., Hardy, O. J., Heuertz, M., & Pineiro, R. (2020). Miocene Diversification in the Savannas Precedes Tetraploid Rainforest Radiation in the African Tree Genus *Afzelia* (Detarioideae, Fabaceae). *Frontiers in Plant Science* , 11 , 798. doi: 10.3389/fpls.2020.00798
- Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* , 19 , 11–15.
- Duminil, J., & Di Michele, M. (2009). Plant species delimitation: a comparison of morphological and molecular markers. *Plant Biosystems* , 143 , 528–542. doi: 10.1080/11263500902722964
- Duminil, J., Kenfack, D., Viscosi, V., Grumiau, L., & Hardy, O. J. (2012). Testing species delimitation in sympatric species complexes: The case of an African tropical tree, *Carapa* spp. (Meliaceae). *Molecular Phylogenetics and Evolution* , 62 , 275–285. doi: 10.1016/j.ympev.2011.09.020
- Eisenring, M., Altermatt, F., Westram, A. M., & Jokela, J. (2016). Habitat requirements and ecological niche of two cryptic amphipod species at landscape and local scales. *Ecosphere* . 7 , 1–13. doi: 10.1002/ecs2.1319

- Ellis, J., Knight, M. E., Carvell, C. & Goulson, D. (2006). Cryptic species identification: A simple diagnostic tool for discriminating between two problematic bumblebee species. *Molecular Ecology Notes* , 6 , 540–542. doi: 10.1111/j.1471-8286.2006.01231.x
- Ence, D. D., & Carstens, B. C. (2011). SpedeSTEM: A rapid and accurate method for species delimitation. *Molecular Ecology Resources* ,11 , 473–480. doi: 10.1111/j.1755-0998.2010.02947.x
- Firetti-Leggieri, F., Zuntini, A. R., Gaiarsa, J., Oliveira, R. S., Lohmann, L. G., & Van Sluys, M.-A. (2017). Complete chloroplast genome sequences contribute to plant species delimitation: A case study of the *Anemopaegma* species complex. *American Journal of Botany* ,104 , 1493–1509. doi: 10.3732/ajb.1700302
- 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. doi: 10.1038/s41598-019-39161-x
- Fujisawa, T. G., Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed Yule coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology* , 62 , 707–724.
- Fujita, M. K., Leache, A. D., Burbrink, F. T., McGuire, J. A. & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* , 27 , 480–488. doi: 10.1016/j.tree.2012.04.012
- He, S.-Z. (1990). A new medicinal species of *Paris* from Guizhou province, China. *Guizhou Science* , 3, 16.
- Hebert, P. D. N., Cywinska, A., Ball, S. L., & De-Waard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* , 270, 313–322. doi: 10.1098/rspb.2002.2218
- Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* , 101 , 14812–14817. doi: 10.1073/pnas.0406166101
- Henderson, A. (2005). The methods of herbarium taxonomy. *Systematic Botany* , 30 , 456–459. doi: 10.1600/0363644054223701
- Hollingsworth, P. M. (2011). Refining the DNA barcode for land plants. *Proceedings of the National Academy Sciences of the United States of America* , 108 , 19451–19452. doi: 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 Sciences of the United States of America* , 106 , 12794–12797. doi: 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. doi: 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. doi: 10.1098/rstb.2015.0338
- Hopkins, G. W. & Freckleton, R. P. (2002). Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation* , 5 , 245–249. doi: 10.1017/S1367943002002299
- Huang, Y.-L., Li, X.-J., Yang, Z.-Y., Yang, C.-J., Yang, J.-B., & Ji, Y. (2016). Analysis of complete chloroplast genome sequences improves phylogenetic resolution of *Paris* (Melanthiaceae). *Frontiers in Plant Sciences* , 7 , 1797. doi:10.3389/fpls.2016.01797
- Jacobs, S. J., Kristofferson, C., Uribe-Convers, S., Latvis, M., & Tank, D. C. (2018). Incongruence in molecular species delimitation schemes: What to do when adding more data is difficult. *Molecular Ecology*

, 27 , 2397–2413. doi: 10.1111/mec.14590

Ji, Y. (2020). A Monograph of *Paris* (Melanthaceae). Beijing: Science Press; Singapore: Springer.

Ji, Y., Fritsch, P. W., Li, H., Xiao, T., & Zhou, Z. (2006). Phylogeny and classification of *Paris* (Melanthiaceae) inferred from DNA sequence data. *Annals of Botany* , 98 , 245–256. doi: 10.1093/aob/mcl095

Ji, Y., Yang, C.-J., & Huang, Y.-L. (2017). A new species of *Paris* sect. *Axiparis* (Melanthiaceae) from Yunnan, China. *Phytotaxa* , 306 , 234–236. doi: 10.11646/phytotaxa.306.3.6

Ji, Y., Liu, C.-K., Yang, Z.-Y., Yang, L.-F., He, Z.-S., Wang, H.-C., ... Yi, T.-S. (2019a). Testing and using complete plastomes and ribosomal DNA sequences as the next generation DNA barcodes in *Panax* (Araliaceae). *Molecular Ecology Resources* , 19 , 1333–1345. doi: 10.1111/1755-0998.13050

Ji, Y., Yang, L.-F., Chase, M. W., Liu, C.-K., Yang, Z.-Y., Yang, J., ... Yi, T.-S. (2019b). Plastome phylogenomics, biogeography, and clade diversification of *Paris* (Melanthiaceae). *BMC Plant Biology* , 19 , 543. doi: 10.1186/s12870-019-2147-6

Ji, Y., Liu, C.K., Yang, J., Jin, L., Yang, Z.-Y., & Yang, J.-B. (2020). Ultra-barcoding discovers a cryptic species in *Paris yunnanensis* (Melanthiaceae), a medicinally important plant. *Frontiers in Plant Sciences* , 11 , 411. doi: 10.3389/fpls.2020.00411

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., Yi, T.-S., & Li, D.-Z. (2018). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* , 21 , 241. doi:10.1186/s13059-020-02154-5).

Jirapatrasilp, P., Backeljau, T., Prasankok, P., Chanabun, R., & Panha, S. (2019). Untangling a mess of worms: Species delimitations reveal morphological crypsis and variability in Southeast Asian semi-aquatic earthworms (Almidae, *Glyphidrilus*). *Molecular Phylogenetics and Evolution* , 139 , 106531. doi: 10.1016/j.ympev.2019.106531

Jones, G. L. (2017). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematic Biololgy* , 74 , 447–467. doi: 10.1007/s00285-016-1034-0

Joppa, L. N., Roberts, D. L., & Pimm, S. L. (2011). How many species of flowering plants are there? *Proceedings of the Royal Society B: Biological Sciences* , 278 , 554–559. doi: 10.1098/rspb.2010.1004

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

Kane, N., Sveinsson, S., Dempewolf, H., Yang, J.-Y., Zhang, D.-P., Engels, J. 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. doi: 10.3732/ajb.1100570

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* , 33 , 1630–1638. doi: 10.1093/bioinformatics/btx025

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. doi: 10.1093/molbev/mst010

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

Kekkonen, M., & Hebert, P. D. N. (2014). DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Molecular Ecology Resources* , 14 , 706–715. doi: 10.1111/1755-0998.12233

Knope, M. L., Bellinger, M. R., Datlof, E. M., Gallaher, T. J., & Johnson, M. A. (2020). Insights into the

- Evolutionary History of the Hawaiian *Bidens* (Asteraceae) Adaptive Radiation Revealed Through Phylogenomics. *Journal of Heredity* , 111 , 119–137. doi: 10.1093/jhered/esz066
- Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weight, L. A., & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* , 102 , 8369–8374. doi: 10.1073/pnas.0503123102
- Lambert, F., Clark, J. L., & Joly, S. (2017). Species delimitation in the Caribbean *Gesneria viridiflora* complex (Gesneriaceae) reveals unsuspected endemism. *Taxon* , 66 , 1171–1183. doi: 10.12705/665.10
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie2. *Nature Methods* , 9 , 357–359. doi: 10.1038/nmeth.1923
- Leveille, H. (1906). *Memorie della Pontificia Accademia Romana dei Nuovi Lincei* , 24 , 21, 354, 355.
- Lewin, R. A. (1981). Three species concepts. *Taxon* , 30 , 609–613. doi: 10.2307/1219942
- Li, H. (1984). Two new species of the genus *Paris* . *Acta Botanica Yunnanica* , 6 , 273–276.
- Li, H. (1992). New taxa of the genus *Paris* L. *Acta Botanica Yunnanica* , Suppl 5 , 13–18.
- Li, H. (1998). The genus *Paris* (Trilliaceae). Beijing: Science Press.
- Li, L., Zhang, J., Lu, Z.-Q., Zhao, J.-L., & Li, Q.-J. (2020). Genomic data reveal two distinct species from the widespread alpine ginger *Roscoea tibetica* Batalin (Zingiberaceae). *Journal of Systematics and Evolution* . doi: 10.1111/jse.12596
- Li, X.-W., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y.-T., & Chen, S.-L. (2015). Plant DNA barcoding: from gene to genome. *Biological Reviews* , 90 , 157–166. doi: 10.1111/brv.12104
- Liu, J., Moller, M., Gao, L.-M., Zhang, D.-Q., & Li, D.-Z. (2011). DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species. *Molecular Ecology Resources* , 11 , 89–100. doi: 10.1111/j.1755-0998.2010.02907.x
- Ma, Z.-Y., Wen, J., Tian, J.-P., Gui, L.-L., & Liu, X.-Q. (2020). Testing morphological trait evolution and assessing species delimitations in the grape genus using a phylogenomic framework. *Molecular Phylogenetics and Evolution* , 148 , 106809. doi: 10.1016/j.ympev.2020.106809
- Mace, G. M. (2004). The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society B : Biological Sciences* , 359 , 711–719. doi: 10.1098/rstb.2003.1454
- Mace, G. M., Gittleman, J. L., & Purvis, A. (2003). Preserving the Tree of Life. *Science* , 300 , 1707–1709. doi: 10.1126/science.1085510
- Margules, C. R. & Pressey, R. L. (2000). Systematic conservation planning. *Nature* , 405 , 243–253. doi: 10.1038/35012251
- Masters, B. C., Fan, V., & Ross, H. A. (2011). Species delimitation – a geneious plugin for the exploration of species boundaries. *Molecular Ecology Resources* , 11 , 154–157. doi: 10.1111/j.1755-0998.2010.02896.x
- May, R. M. (1988). How many species are there on earth? *Science* , 241 , 1441–1449. doi: 10.1126/science.241.4872.1441
- May, R. M. (1990). How many species? *Philosophical Transactions of the Royal Society B : Biological Sciences* , 330 , 293–304. doi: 10.1098/rstb.1990.0200
- Mishler, B. D. & Donoghue, M. J. (1982). ‘Species Concepts: A Case for Pluralism’, *Systematic Biology* , 31 , 491–503. doi: 10.1093/sysbio/31.4.491

- Mutanen, M., Kekkonen, M., Prosser, S. W. J., Hebert, D. N., & Kaila, L. (2015). One species in eight: DNA barcodes from type specimens resolve a taxonomic quagmire. *Molecular Ecology Resources* , 15 , 967–984. doi: 10.1111/1755-0998.12361
- 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. doi: 10.1111/j.1467-7652.2010.00558.x
- Patel, R. K. & Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS ONE* , 7 , e30619. doi: 10.1371/journal.pone.0030619
- 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. doi: 10.2307/4135365
- Posada, D., & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* , 14 , 817–818. doi: 10.1093/bioinformatics/14.9.817
- 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. doi: 10.1080/10635150600852011
- Prance, G. T., Beentje, H., Dransfield, J., Johns, R., & Johns, R. (2000). The tropical flora remains undercollected. *Annals of the Missouri Botanical Garden* , 67–71. doi: 10.2307/2666209
- Puillandre, N., Modica, M. V., Zhang, Y., Sirovich, L., Boisselier, M. C., Cruaud, C., ... Samadi, S. (2012). Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology* , 21 , 2671–2691. doi: 10.1111/j.1365-294X.2012.05559.x
- Qin, H.-N., Yang, Y., Dong, S.-Y., He, Q., Jia, Y., Zhao, L., ... Xue, N.-X. (2017). Threatened species list of China's higher plants. *Biodiversity Science* , 25 , 696–744. doi: 10.17520/biods.2017144
- Rannala, B., & Yang, Z. (2013). Improved reversible jump algorithms for Bayesian species delimitation. *Genetics* , 194 , 245–253. doi: 10.1534/genetics.112.149039
- Rodrigo, A., Bertels, F., Heled, J., Noder, R., Shearman, H., & Tsai, P. (2008). The perils of plenty: what are we going to do with all these genes? *Philosophical Transactions of the Royal Society B: Biological Sciences* , 363 , 3893. doi: 10.1098/rstb.2008.0173
- Rojas, M. (1992). The species problem and conservation: what are we protecting? *Conservation Biology* , 6 , 170–178. doi: 10.2307/2386238
- Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* , 19 , 1572–1574. doi: 10.1093/bioinformatics/btg180
- Rosenberg, N. A. (2007). Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution* , 61 , 317–323. doi: 10.1111/j.1558-5646.2007.00023.x
- Ruhsam, M., Rai, H. S., Mathews, S., Ross, T. 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. doi: 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. doi: 10.1093/nar/gki366
- Sites, J. W., Jr. & Crandall, K. A. (1997). Testing species boundaries in biodiversity studies. *Conservation Biology* , 11 , 1289–1297. doi: 10.2307/2387356
- Sites, J. W., Jr. & Marshall, J. C. (2003). Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution* , 18 , 462–470. doi: 10.1016/S0169-5347(03)00184-8

- Ślipiko, M., Myszczyński, K., Buczkowska, K., Baczkiewicz, A., Szczecińska, M., & Sawicki, J. (2020). Molecular delimitation of European leafy liverworts of the genus *Calypogeia* based on plastid super-barcodes. *BMC Plant Biology* , 20 , 243. doi: 10.1186/s12870-020-02435-y
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analysis with thousands of taxa and mixed models. *Bioinformatics* , 22 , 2688–2690. doi: 10.1093/bioinformatics/btl446
- Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., & Liston, A. (2012). Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* , 99 , 349–364. doi: 10.3732/ajb.1100335
- Su, X., Wu, G.-L., Li, L.-L., & Liu, J.-Q. (2015). Species delimitation in plants using the Qinghai-Tibet plateau endemic *Orinus* (Poaceae: Tridentinae) as an example. *Annals of Botany* , 116 , 35–48. doi: 10.1093/aob/mcv062
- Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America* , 114 , 1607–1612. doi: 10.1073/pnas.1607921114
- Takhtajan, A. (1983). A revision of *Daiswa* (Trilliaceae). *Brittonia* , 35 , 255–270. doi: 10.2307/2806025
- Wiley, E.O. & Mayden, R. (2000). The evolutionary species concept. In *Species Concepts and Phylogenetic Theory: A Debate* (Wheeler, Q.D. and Meier, R., eds), pp. 70–89, Columbia University Press.
- Wu, W., Ning, W.-L., Yang, J.-X., Li, W.-M., & Ge, X.-J. (2018). High cryptic species diversity is revealed by genome-wide polymorphisms in a wild relative of banana, *Musa itinerans* , and implications for its conservation in subtropical China. *BMC Plant Biology* , 18 , 194. doi: 10.1186/s12870-018-1410-6
- Wyman, S. K., Jansen, R. K., & Boore, J. L. (2004). Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* , 20 , 3252–3255. doi: 10.1093/bioinformatics/bth352
- Xu, Z., Wei, N., Tan, Y., Peng, S. Ngumbau, V. M., Hu, G.-W., & Wang, Q.-F. (2019). *Paris lihengiana* (Melanthiaceae: Parideae), a new species from Yunnan, China. *Phytotaxa* , 392 , 045–053. doi: 10.11646/phytotaxa.392.1.4
- Yang, J., Wang, Y.-H., & Li, H. (2017). *Paris qiliangiana* (Melanthiaceae), a new species from Hubei, China. *Phytotaxa* , 329 , 193–196. doi: 10.11646/phytotaxa.329.2.13
- Yang, Z.-H., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America* , 107 , 9264–9269. doi: 10.1073/pnas.0913022107
- Yang, Z.-Y., Yang, C.-J., & Ji, Y.-H. (2019). *Paris variabilis* (Melanthiaceae), A new species from southwestern China. *Phytotaxa* , 401 , 190–198. doi: 10.11646/phytotaxa.401.3.4
- 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. doi: 10.1186/s13007-018-0300-0
- Zhang, J.-J., Kapli, P., Pavlidis, P., & Stamatakis, A., (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* , 29 , 2869–2876. doi: 10.1093/bioinformatics/btt499
- Zhou, M.-M., Yang, G.-Q., Sun, G.-L., Guo, Z.-H., & Pan, Y.-Z. (2020). Resolving complication relationships of the *Panax bipinnatifidus* complex in southwestern by RAD-seq data. *Molecular Phylogenetics and Evolution* , 149 , 106851. doi: 10.1016/j.ympev.2020.106851

Data Accessibility Statement

All sequences data generated in current study have been deposited into NCBI GenBank database (accession number are presented in **Table S1**).

Author Contributions

YJ, JBY, and TSY designed the study. YJ, JBY, ZY, JY, and SW collected data. YJ, JY, SW, ZY, LJ, and LL analyzed data. YJ wrote the manuscript. JBL and TSY discussed the results and revised the manuscript.

Figure legends

Figure 1. Nine nominal species within *Paris* section *Axiparis* , showing their leaf size and shape variations. A, *Paris axialis* ; B, *P. dulongensis* (8 accessions); C, *P. forrestii* ; D, *P. guizhouensis* ; E, *P. lihengiana* ; F, *P. rugosa* ; G, *P. tenchongensis* ; H, *P. vaniotii* ; I, *P. variabilis* .

Figure 2. Phylogenetic tree of *Paris* section *Axiparis* based on Maximum likelihood (ML) and Bayesian inference (BI) analyses of complete plastomes. Well-supported nodes (BS > 90% and PP > 0.95) are indicated with a star. Species delimitation schemes proposed by automatic barcode gap discovery (ABGD), multi-rate Poisson tree processes model (mPTP), and species delimitation plugin (SDP) of Geneious are reflected on the tree topology.

Figure 3. Distribution of the nine nominal species within *Paris* section *Axiparis*.

Figure 4. Phylogenetic tree of *Paris* section *Axiparis* based on Maximum likelihood (ML) and Bayesian inference (BI) analyses of nuclear ribosomal DNA (nrDNA). Well-supported nodes (BS > 90% and PP > 0.95) are indicated with a star. Species delimitation schemes proposed by automatic barcode gap discovery (ABGD), multi-rate Poisson tree processes model (mPTP), and species delimitation plugin (SDP) of Geneious are reflected on the tree topology.

Supporting Information

Table S1. Samples used in this study with locality, voucher and GenBank accession numbers.

Table S2. Summary of Illumina sequencing.

Table S3. Characteristics of complete plastome and nuclear ribosomal DNA datasets in *Paris* section *Axiparis* .

Table S4. The number of putative species recognized by Automatic Barcode Gap Discovery (ABGD) analyses among 51 complete plastomes using three distance metrics.

Table S5. The number of putative species recognized by Automatic Barcode Gap Discovery (ABGD) analyses among 51 nuclear ribosomal DNA (nrDNA) sequences using three distance metrics.

Table S6. Posterior delimitation probability of two putative species proposed by mPTP analyses of complete plastomes and nuclear ribosomal DNA (nrDNA) sequences.

Table S7. Species delimitation plugin (SDP) analyses show the distinctiveness (Rodrigo's $P_{(RD)}$) and significance (Rosenberg's $P_{(AB)}$) of the two clades recovered by phylogenetic analyses of complete plastomes and nuclear ribosomal DNA (nrDNA) as species-level taxonomic units.

Hosted file

Figure 1.pdf available at <https://authorea.com/users/388751/articles/503504-from-nine-to-two-ultra-barcoding-develops-a-rigorous-species-delimitation-framework-in-paris-section-axiparis-melanthiaceae>



