Insect Conservation and Diversity

Insect Conservation and Diversity (2020) doi: 10.1111/icad.12462

High diversity and strong variation in host specificity of seed parasitic acorn weevils

QIAN-YA LI,[1†] XING-HUA HU,[1,2†] DE-CHEN LIU,[1] AO OUYANG,[1] XIN TONG,[1] YONG-JIN WANG,[1] RONG WANG[1,3] and XIAO-YONG CHEN[1,3] [1]Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, ECNU-TMS Joint Lab for Biodiversity Conservation, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, 200241, China, [2]Guangxi Institute of Botany, Chinese Academy of Sciences, Guilin, 541006, China and [3]Shanghai Institute of Pollution Control and Ecological Security, Shanghai, 200092, China

Abstract. 1. Host specificity of phytophagous insects plays a critical role in the maintenance of plant diversity. However, a lack of approaches to identify the larvae of phytophagous insects, which usually have similar morphology, can lead to misinterpretation of their host ranges.

2. Acorn weevils are seed-exploiting parasites of Fagaceae species, a dominant plant group in subtropical evergreen broad-leaved forests (EBLFs). Therefore, variation in their host specificity may profoundly influence plant diversity in this ecosystem.

3. Here, we collected larvae of acorn weevils from 36 655 seeds of Fagaceae species in EBLFs in southeastern China and performed multiple operational taxonomic unit (OTU) delimitation analyses using one mitochondrial (*COI*) and two nuclear (*ArgK* and *EF1-α*) genes to reveal species composition of acorn weevils and assess the variation in host specificity among OTUs.

4. We analysed the DNA barcoding sequences from 478 larvae and 4 adults of acorn weevils sampled from 17 common Fagaceae species (five genera) and detected 11 OTUs. Moreover, host range varied drastically among different OTUs, with *Curculio bimaculatus* being the most generalised.

5. Our results revealed high diversity of acorn weevils and that strong variation in host specificity can even exist among OTUs in the same genus, addressing the need to distinguish different *Curculio* species before conducting forest management in EBLFs.

Key words. *Curculio*, DNA barcodes, evergreen broad-leaved forests, Fagaceae, host specificity, OTU diversity.

Introduction

Interactions between phytophagous insects and angiosperms play a crucial role in the maintenance of ecosystem functioning (Forister *et al*., 2012; Burkle *et al*., 2013), due to not only the high diversity of these two groups but also their location in the centre of ecosystem networks, where they are critical contributors to both energy and nutrition flows (Farrell, 1998; Fonseca *et al*., 2005). More importantly, host specificity of phytophagous insects can profoundly affect plant diversity. Host-specific natural enemies of plants, especially local dominant plants, specifically regulate their hosts and benefit rare plant species (Bagchi *et al*., 2014; Wang *et al*., 2019a), whereas generalist phytophagous insects are more likely to be pests that threaten the persistence of many plants (Schmidt *et al*., 2019). Therefore, explicitly delineating the host range of key groups of phytophagous insects provides an important basis for predicting the sustainability of forest ecosystems as well as for restoration and management strategies (Figueroa *et al*., 2018).

To precisely estimate host range, accurate taxonomy is an essential prerequisite. However, biased estimation may occur in the presence of cryptic species and hybrids even within some morphologically well-identified taxa (Hajibabaei *et al*., 2006;

Correspondence: Rong Wang, School of Ecological and Environmental Sciences, East China Normal University, Dongchuan Road 500, Shanghai 200241, China. E-mail: rwang@des.ecnu.edu.cn

[†]These authors are equally contributed to this work.

Smith *et al*., 2006). Moreover, although larvae are more stationary than adults and often live only on their host plants, current morphological identification of insects is based predominantly on adult morphological traits, as the larvae of many species are indistinguishable, even for some highly differentiated species (Marvaldi *et al*., 2002; Kajtoch *et al*., 2018). Therefore, in order to explore the biodiversity of phytophagous insects at the larval stage and improve the estimation of their host ranges, it is necessary to use DNA barcoding techniques to explicitly distinguish different species (Sequeira *et al*., 2008; Bocak *et al*., 2014).

Acorn weevils (*Curculio*, Curculionidae) are a typical group of seed-parasitic insects. Female adults of these species lay eggs in the seeds of nut-bearing trees, and larvae feed exclusively on the tissues inside their host seeds (Bonal & Muñoz, 2009). Their host specificity is related to factors such as host preference, matching of phenology and coevolutionary relationships with potential hosts (Hughes & Vogler, 2004; Arias-Leclaire *et al*., 2018). Compared with insects that attack other plant organs (e.g., leaves), acorn weevil plays a more important role in the seed availability of its host plant (Tong *et al*., 2017), with up to 90% of seeds on a host tree being infested and destroyed by these parasites (Qian-Ya Li, unpublished data). Therefore, acorn weevils often have a serious impact on the regeneration of host plant species (Muñoz *et al*., 2014). Despite the great importance of these seed parasites, little is known about their species diversity and interspecific variation in host specificity (but see Peguero *et al*., 2017).

Southeastern China is one of the most densely populated areas in the world. The zonal vegetation of this region is subtropical evergreen broad-leaved forests (EBLFs), where Fagaceae species (major hosts of acorn weevils) are a dominant group of tree species (Song *et al*., 2005; Tong *et al*., 2017; Wang *et al*., 2020). Therefore, revealing the host specificity of acorn weevils should have important implications for the long-term maintenance of biodiversity in this region. However, the species composition of acorn weevils is still unclear, due to difficulties in collecting ovipositing individuals in the canopy and the fact that different species have similar morphological traits at the larval stage.

In this study, we investigated acorn weevil communities in 17 locations in southeastern China and sequenced mitochondrial (*COI*) and nuclear (*ArgK* and *EF1-α*) genes, to evaluate the species [as represented by operational taxonomic units (OTUs)] diversity within acorn weevils using multiple OTU delimitation approaches. In addition, we recorded the host species for each sampled acorn weevil larva and estimated the host specificity of each acorn weevil OTU. Specifically, we addressed the following questions: (i) How many *Curculio* OTUs are in this region and (ii) is there a large difference in host range among *Curculio* OTUs?

Materials and methods

Sampling sites and focal species

Sampling of acorn weevils was conducted at 17 sites (27.80° −30.33° N, 117.93° −121.80° E) in southeastern China between 2009 and 2016 (Fig. 1; Supporting Information Table S1). This region has a subtropical monsoon climate with hot and rainy summers and mild winters. The annual average temperature and precipitation at different sites range from 15 to 18 °C and from 980 to 2000 mm, respectively, and the altitude of the sampling sites varied from 16 to 869 m. The predominant natural vegetation of this area is EBLFs, with Fagaceae species being a dominant group of tree species (Song *et al*., 2005). These Fagaceae species are the host plants of acorn weevils (Tong *et al*., 2017).

Adult acorn weevils usually emerge from the surface soil in summer (June to August) after pupation (Higaki, 2016; Qian-Ya Li, unpublished data). They then fly up to the canopy and feed on leaves and young shoots. Copulation and oviposition take place in autumn (August to mid-October). During oviposition, adult females excavate a hole in the coat of the host seed using their specialised rostrum and generally only lay a single egg into the developing seed (Desouhant *et al*., 2000). However, an adult female sometimes can lay several eggs in an oviposition, and a seed can be parasitised by multiple females, leading to more than one larva within an infested seed (Bonal & Muñoz, 2008). The larvae of acorn weevils live exclusively on cotyledons (Bonal & Muñoz, 2009), and leave their host seeds in winter (from November to January next year), burrowing underground for diapause (lasting for 1–3 years) and becoming adults after pupation (Espelta *et al*., 2009). Therefore, host plants of acorn weevils can be accurately identified only at the larval stage. However, given that the larvae of most acorn weevil species share similar morphological traits (Tong *et al*., 2017), morphological identification is difficult, and it is necessary to use DNA barcoding techniques.

Field sampling and laboratory rearing

To sample acorn weevil larvae for identification using a DNA barcoding technique and to record host plants, we collected seeds from all Fagaceae species found at the 17 sites in late autumn (from Mid-October to early November). Except for the Zhoushan Archipelago and Thousand-Island Lake, where sampling was only conducted twice (2015 and 2016), we carried out sampling at other sites for four times between 2009 and 2012 (Supporting Information Table S1). At each site, at least 20 Fagaceae trees were surveyed, and a total of 36 655 seeds were collected during our sampling period (Supporting Information Table S1). Each collected seed was stored separately, and its maternal tree was recorded.

All collected seeds were checked daily until acorn weevil larvae emerged. Approximately 10–20 larvae per Fagaceae species were collected from each site and were preserved in pure ethanol at 4 °C for DNA extraction. Other larvae were reared for morphological identification.

We placed the reared larvae into heat-sterilised soil and maintained humidity at c. 70% and indoor temperature at 20–25 °C to facilitate pupation and emergence of adults. All emerged adults were identified to morphospecies based on the description by Pelsue and Zhang (2000). In each morphospecies, we chose one adult as a specimen. After photographing, its hind legs were

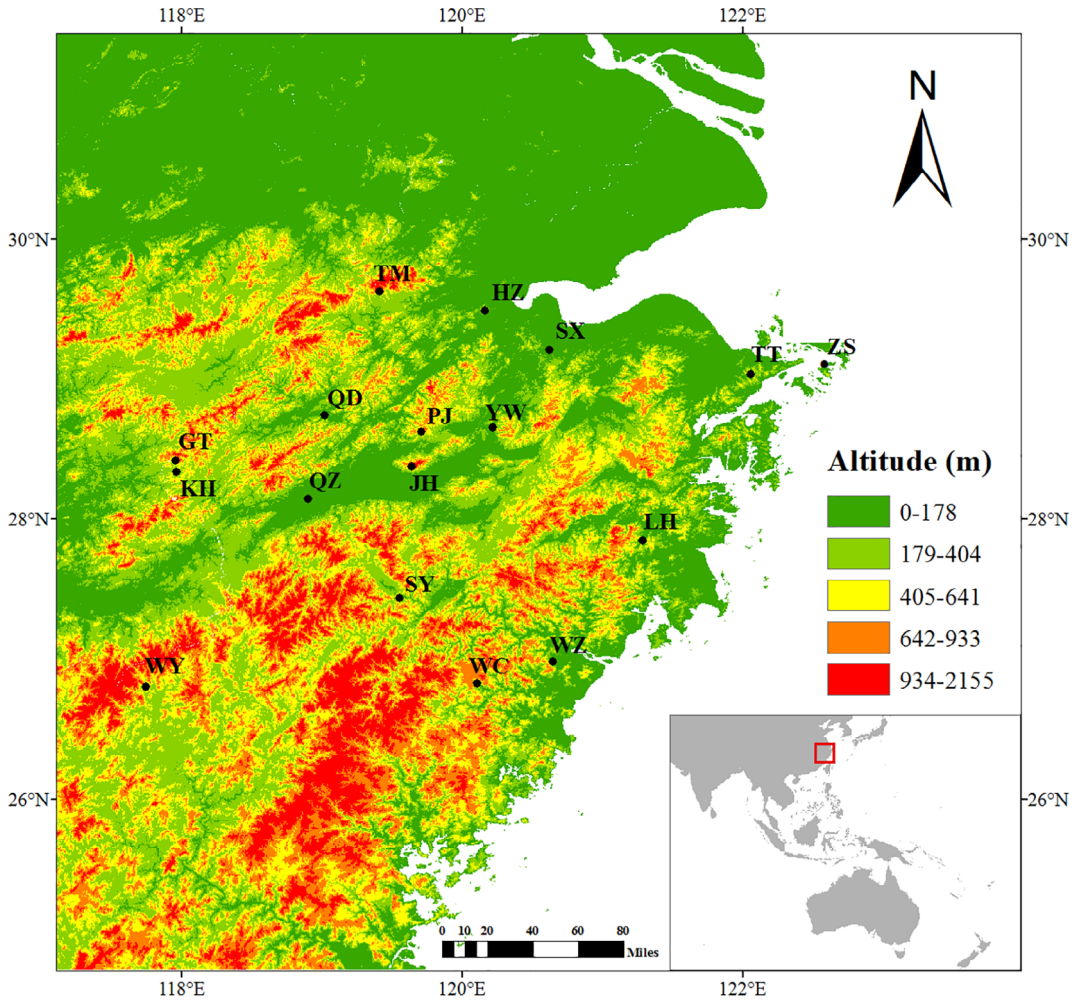

Fig 1. Distribution of sampling sites in southeastern China. The full name of each sampling site is listed in the Supporting Information Table S1. [Color figure can be viewed at wileyonlinelibrary.com]

used for DNA extraction (retaining the rest of body as the morphological reference).

DNA extraction, amplification and sequencing

We extracted genomic DNA from the entire body of each collected acorn weevil larva and the hind legs of each emerged adult. DNA extraction was carried out using a micro-tissue extraction kit (TIANGEN, Shanghai) following the manufacturer's instructions.

After DNA extraction, polymerase chain reactions (PCRs) were performed using universal DNA barcoding primers for one mitochondrial DNA (mtDNA) (*COI*) (Simon *et al*., 1994) and two nuclear DNA (nDNA) (*ArgK* and *EF1-α*) (Sequeira *et al*., 2008) (Supporitng Information Table S2). A 20-μl PCR system was used, and PCRs were performed following the procedure: initial denaturation at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 45 s; annealing at 50–54 °C (*COI* and *ArgK*) and 46.5–52 °C (*EF1-α*) for 45 s; extension at 72 °C for 1 min and a final extension at 72 °C for 10 min. Bidirectional Sanger sequencing of the amplified products was implemented using an ABI 3730 Genetic Analyser (Applied Biosystems™, Foster City, California).

Sequence alignment, concatenation and haplotype statistics

Each DNA sequence was assembled using SeqMan version 7.1 in software DNASTAR (Lasergene, Madison, USA) and then aligned using program MUSCLE in MEGA version 5.0 (Tamura *et al*., 2011). In addition, to avoid overestimating the OTU diversity of acorn weevils due to the presence of pseudogenes (NUMTs), we screened sequences based on divergence from orthologous sequences, indels and in-frame stop codons (Song *et al*., 2008). We used Sequence Matrix version 1.7.8 (Vaidya *et al*., 2011) to compile a combined matrix of 2135 bp from three gene regions (642 bp for *COI*, 684 bp for *ArgK* and

809 bp for *EF1-α*) to improve the reliability of OTU delimitation analysis (Sequeira *et al*., 2008). The incongruence length difference (ILD) test using PAUP version 4.0b10 (Swofford, 2002) showed homogenous base substitution rates in the combined sequences of three genes (ILD test, $P = 0.25$). Prior to OTU delimitation, we identified the haplotypes from the dataset of the combined three genes and from the dataset of each of the three genes using DnaSP version 5.10.01 (Librado & Rozas, 2009).

OTU delimitation of acorn weevils

To precisely explore OTU richness of acorn weevils in our study region, we used three OTU delimitation approaches including the priori distance threshold approach (PDTA), a distance-based model [Automatic Barcode Gap Discovery (ABGD)], and a tree-based model [Generalised Mixed Yule Coalescent model (GMYC)].

The PDTA calculates the genetic distance between sequences using Kimura's 2-parameter distance (K2P or K80) and sets a cut-off standard value as the 'barcode gap' of 3% divergence (Smith *et al*., 2005) or the 10× rule (Hebert *et al*., 2004). Therefore, we assessed the intra/inter-clade variation of unique haplotypes of the combined three genes through K2P model in MEGA 5.0 to determine the putative species/OTUs and validated the reliability of this delimitation using three methods (bpNewTraining, fuzzyId and Bayesian) with package BarcodingR version 1.0-3 (Zhang *et al*., 2017) in R version 3.6.3 (R Development Core Team, 2020).

ABGD has a clear definition of the barcode gap and can partition data into candidate species/OTUs even if overlapping of inter- and intra-clade genetic distances exists (Puillandre *et al*., 2012). This approach delimits OTUs by calculating a model-based one-sided confidence limit for the intra-specific divergence using a priori knowledge from the data. The first significant barcode gap detected beyond this limit is used to split molecular operational taxonomic units (MOTUs) (initial division), and then recursive repartition of groups is performed to further delimit OTUs (recursive division). Reliable results can be obtained when initial and recursive divisions converge in a large range of intra-specific divergence. ABGD for the combined three genes were conducted in the website (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) setting the following parameters: X (relative gap width) = 1.5, Steps = 20, Nb bins = 20, $P_{min} = 0.0003$ and $P_{max} = 0.1$ (minimum and maximum priori values of intra-specific divergence) with the Kimura (K80) TS/TV model. We also tried a range of different values for each parameter, but the results of OTUs delimitations remained the same.

We also used GMYC to establish divergence threshold (T, which represents the time of transition from coalescence to speciation) for dividing species/OTUs in a phylogenetic tree. This method relies on an ultrametric phylogenetic tree to identify the inflection points, i.e., branching events where intraspecific coalescent processes turn to be interspecific Yule processes) using maximum likelihood method (Fujisawa & Barraclough, 2013). We used the unique haplotype sequences from the combined three genes to construct the ultrametric tree in BEAST version 1.10.4 software (Drummond & Rambaut, 2007), using settings for the Yule speciation model, uncorrelated Relaxed Clock with Lognormal distribution and Markov chain Monte Carlo (MCMC) with the length of 100 million generations. Three insect species, *Curculio pyrrhoceras* (Hughes & Vogler, 2004), *Naupactus xanthographus* (Lanteri & del Río, 2017) and *Cylas formicarius* (Bentidae), were used as the outgroups (see Supporting Information Table S3 for details). Best partitioning scheme and evolutionary models for three pre-defined partitions (the optimal base substitution model for *COI*, *ArgK* and *EF1-α* was GTR (general time reversible) + I (invariant sites) + G (Gamma distributed), SYM (symmetrical model) + G and GTR + I + G, respectively) were selected by greedy algorithm and Akaike's Information Criterion corrected for small sample size (AIC_c) in PartitionFinder2 (Lanfear *et al*., 2017) in PhyloSuite version 1.2.1 (Zhang *et al*., 2020). For the mtDNA and the two nDNA sequences, we also set different codon partitions (GTR + I + G for *COI*; GTR + I + G for *ArgK*_codon 1 and *EF1-α*_codon 1; GTR + G for *ArgK*_codon 2; HKY (Hasegawa-Kishino-Yano) + I + G for *EF1-α*_codon 2; HKY + G for *ArgK*_codon 3 and *EF1-α*_codon 3). The convergence of MCMC chain was examined in TRACER version 1.6 (Rambaut *et al*., 2014), and effective sample sizes (ESS) values >200 indicated adequate posterior sampling. TreeAnnotator version 1.10.4 in BEAST was applied to create a maximum clade credibility tree, setting the burn-in percentage of 10% and the posterior probability (PP) threshold of 0.5. Single threshold model was set in GMYC (Pons *et al*., 2006), and likelihood ratio test (LRT) was used to examine the difference between the null and the GMYC model for evaluating the significance of inflection points for OTU delimitation using R package splits version 1.0-11 (Ezard *et al*., 2009).

The above three OTU delimitation approaches were also performed for each single gene, with the same parameter settings and outgroup species as those used in the combined three genes.

Host range and host specificity of acorn weevils

To demonstrate the host ranges of acorn weevils, we constructed phylogenetic trees of the 17 host plant species using maximum likelihood (ML) in IQ-Tree version 1.6.8 (Nguyen *et al*., 2015) and Bayesian approaches in MrBayes version 3.2.6 (Ronquist *et al*., 2012) in PhyloSuite. A combined sequence matrix of 1708 bp was compiled from one nuclear gene (*ITS/5.8S*, 392 bp) and two chloroplast genes (*rbcL* and *matK*, 528 bp and 788 bp). The optimal base substitution model for each codon partition was determined by AIC_c criterion and greedy algorithm in PartitionFinder2 (GTR + G for *ITS/5.8S*_ codon 1, 2 and 3; HKY + G for *rbcL*_codon 1; HKY + I + G for *rbcL*_codon 2 and GTR for *rbcL*_codon 3 and *matK*_codon 1, 2 and 3). We downloaded the sequences of the three genes for these species and two outgroup species [*Elaeocarpus hookerianus* (Elaeocarpaceae) and *Poliothyrsis sinensis* (Flacourtiaceae)] from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) (Supporting Information Table S3). When constructing the ML trees, 10 000 ultrafast bootstraps were set under

Edge-linked partition model (Minh *et al*., 2013) in PhyloSuite, and bootstrap value (BS) was calculated for each node. The Bayesian trees were constructed using MCMC approach with two parallel runs setting four chains (three cold and one heated) and 2 000 000 generations (burn-in percentage of 25%), and PP was calculated for each node. The convergence of MCMC chains was monitored using Tracer version 1.6 with ESS > 200.

To quantify the host specificity of each delimitated OTU, host specificity index (*S*) was then calculated as the ratio of the actual number of host plants sampled to the number of potential hosts in the study area (Kavanagh & Burns, 2012), using the formulas: $S = \left(\frac{1}{e}\right)^{K_{QG} \times K_{QS}}$, $K_{QG} = \frac{\beta_G}{\alpha_G + 1}$, and $K_{QS} = \frac{\beta_S}{\alpha_S + 1}$, where S is the overall host specificity at both the genus and species levels, ranging from 0 (general) to 1 (specific); K_{QG} and K_{QS} are host range diversity at the genus and species levels, respectively; β_G and β_S are the minimum host ranges at the genus and species levels, respectively; and α_G and α_S are the number of potential host plants at the genus (number of genera minus

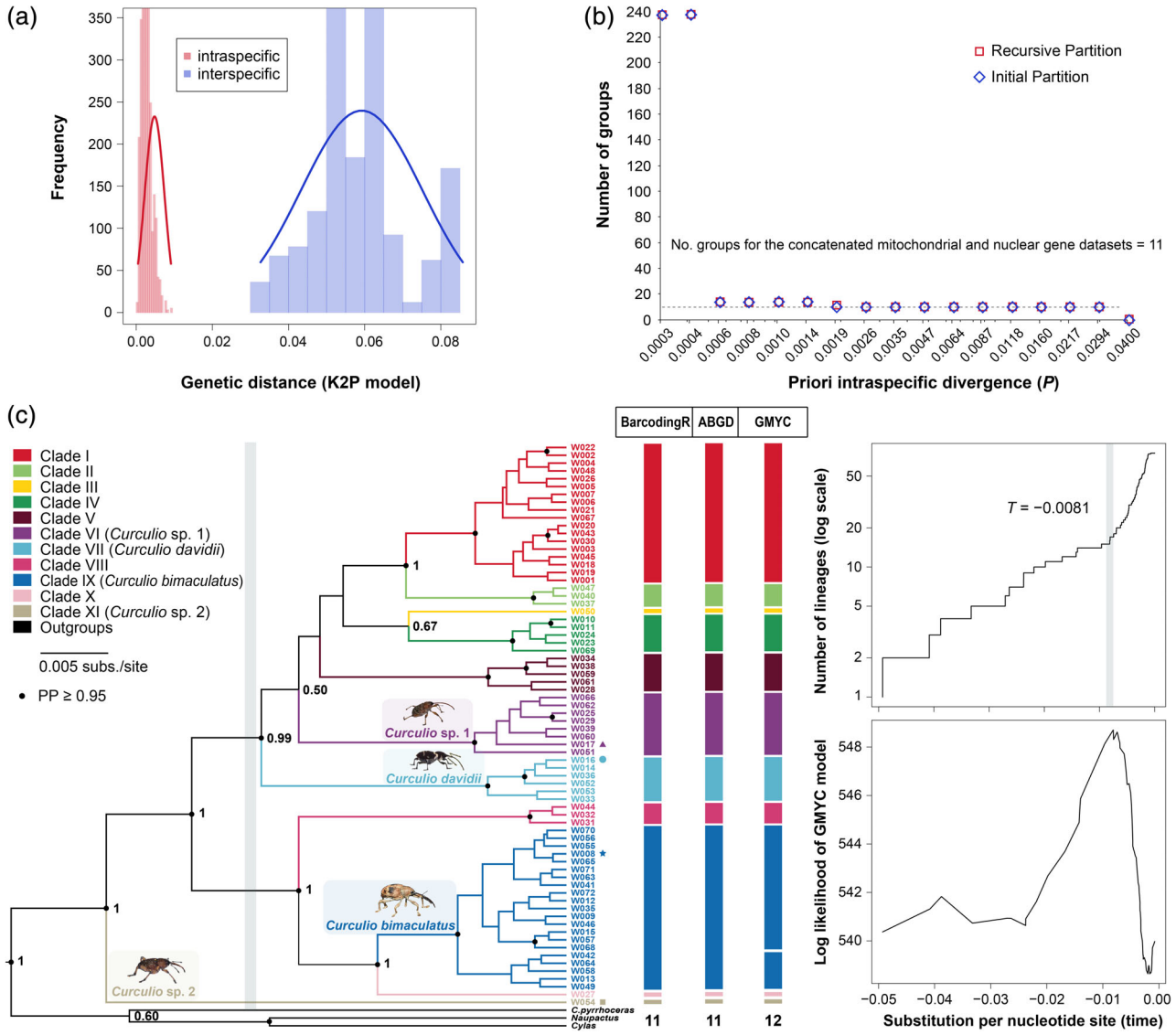

Fig 2. OTU delimitation using the combined mtDNA and nDNA genes (*COI* + *ArgK* + *EF1-α*). (a) The barcoding gap of haplotypes. The intra- and inter-specific K2P distance distributions are shown with red and blue histograms respectively. (b) OTU delimitation using ABGD. Blue diamonds are initial division, and red squares represent recursive division. (c) Summary of OTU delimitation. Bayesian ultrametric tree used in the single threshold GMYC model is shown on the left. Solid tree nodes have posterior probabilities (PP) greater than 0.95. The unique haplotypes belonging to the same OTU are grouped by the same colour in the bars, and the different bars represent the results of different OTU delimitation approaches. The number of delimitated OTUs by each approach is shown at the bottom of each bar. The single threshold (*T*) of transition from coalescence (colour branches) to speciation (black branches) is shown by the vertical grey bar on the tree. The changing trends of the lineage number and log likelihood of the single threshold GMYC model with increasing value of substitutions per nucleotide site are shown on the right. [Color figure can be viewed at wileyonlinelibrary.com]

β_G) and species (number of species minus β_S) levels, respectively.

We also evaluated our sampling completeness using rarefaction curves of the number of acorn weevil OTUs versus increasing overall sample size and the sample size in each host plant genus using aRarefactWin software (Holland, 1998), with bootstrap resampling of sequences to calculate standard deviations.

Results

OTU diversity of acorn weevils

Of the 2088 reared acorn weevil larvae, 545 (an emergence rate of 26.1%) belonging to four morphospecies of *Curculio* became adults. Among them, *Curculio bimaculatus* was the predominant morphospecies with a relative abundance of 89.9% (490 of 545 adults), whereas 42 adults were *Curculio davidii*. In addition, there were two unidentified morphospecies (12 of *Curculio* sp. 1 and 1 of *Curculio* sp. 2). For OTU delimitation and host specificity analysis, we screened four adult individuals (one for each morphospecies) and 478 acorn weevil larvae, all of which had a complete set of mtDNA (*COI*) and nDNA (*ArgK* and *EF1-α*) sequences. After comparing the sequences of our selected larvae with those of other known acorn weevils, we found that all of them belonged to genus *Curculio*.

Using the combined sequences of these three genes, PDTA delimited a total of 11 OTUs with no instances of overlap found between the intra- and inter-clade distance distributions, and haplotypes of the adults of the four morphospecies were assigned into distinct OTUs (Fig. 2a; Supporting Information Table S4). This result was further supported by the three validation methods (vote = 3 for all OTUs). ABGD approach also divided all acorn weevil individuals into 11 OTUs that had exactly the same grouping pattern as revealed by PDTA (Fig. 2b, c). However, GMYC found 12 OTUs from 72 unique haplotypes from our samples with a 95% confidence interval ranging from 11 to 18 (Fig. 2c) and the single threshold T being −0.0081, and this GMYC model had a significantly better fit for branching process than the null model with no coalescence-speciation transitions (LRT: LR = 16.67, $P < 0.001$). The allocation of haplotypes into 10 of the 12 OTUs in GMYC model was completely consistent with that detected by the other two approaches, and the only difference was that the GMYC model splits the clade containing the *C. bimaculatus* adult into two OTUs (which was considered as a single OTU in the other two approaches) (Fig. 2c). Given that we failed to find any differences in morphological characters among the reared *C. bimaculatus* adults, we therefore adopted the conservative delimitation of 11 OTUs in our sampled acorn weevils.

In addition, the results from OTU delimitation analysis using *COI* gene also showed 11 delimited OTUs, with an identical grouping pattern to that revealed by the combined three genes (Fig. S1). However, using *ArgK* gene alone, there were 15 OTUs detected by both ABGD and GMYC (Supporting Information Fig. S2), while a total of 13 OTUs were found in these two approaches only using *EF1-α* gene (Supporting Information Fig. S3). Such a discrepancy is mainly caused by the oversplitting in some clades, e.g., the clade containing the adult *Curculio bimaculatus* (Supporting Information Fig. S2) and that with the adult *Curculio* sp. 1 (Supporting Information Fig. S3), probably due to fewer polymorphic sites in the two nuclear genes (95 for *ArgK* and 107 for *EF1-α*) compared to *COI* gene (208).

Curculio bimaculatus had the highest haplotype diversity, comprising 29.2% of the total unique haplotypes and was the most dominant OTU with the relative abundance of 54.0% (258 of 478 individuals) (Fig. 2c; Table 1). In addition, six *Curculio* OTUs were rare, with very low haplotype diversity (<5 unique haplotypes) and abundance (<8 individuals) in each of these OTUs (Fig. 2c; Table 1).

Host range and host specificity of acorn weevils

The dominant acorn weevil OTU (*C. bimaculatus*) had the widest distribution (found in 12 of 17 sampling sites), but other OTUs were found at only a few sites, with five OTUs being detected only at one site (Table 1). The observed host ranges varied drastically among acorn weevil OTUs, with the dominant OTU having the widest range (10 Fagaceae species belonging to 5 genera) but two OTUs (OTU 3 and 10) being found in only

Table 1. Host range and host specificity of delimitated *Curculio* OTUs.

Curculio OTUs	*N* larvae	*N* sites	*N* haplotypes	*N* host genera	*N* host species	α_G	α_S	K_{QG}	K_{QS}	S
1	71	8	18	3	7	2	10	1.0	0.64	0.527
2	4	1	3	2	3	3	14	0.5	0.20	0.905
3	5	1	1	1	1	4	16	0.2	0.06	0.988
4	59	8	5	3	4	2	13	1.0	0.29	0.748
5	7	2	5	2	2	3	15	0.5	0.13	0.937
6 (*Curculio* sp. 1)	50	5	8	3	6	2	11	1.0	0.50	0.607
7 (*Curculio davidii*)	12	2	6	3	5	2	12	1.0	0.38	0.684
8	8	1	3	1	2	4	15	0.2	0.13	0.974
9 (*Curculio bimaculatus*)	257	12	21	5	10	0	7	5.0	1.25	0.002
10	5	1	1	1	1	4	16	0.2	0.06	0.988

We do not include OTU 11 (*Curculio* sp. 2) because it only contains a single individual. α_G and α_S represent the total number of genera and species of potential host plants for each *Curculio* OTU. K_{QG} and K_{QS} are host range diversity at the genus and species levels. The host specificity index (S) ranges from 0 (generality) to 1 (specificity).

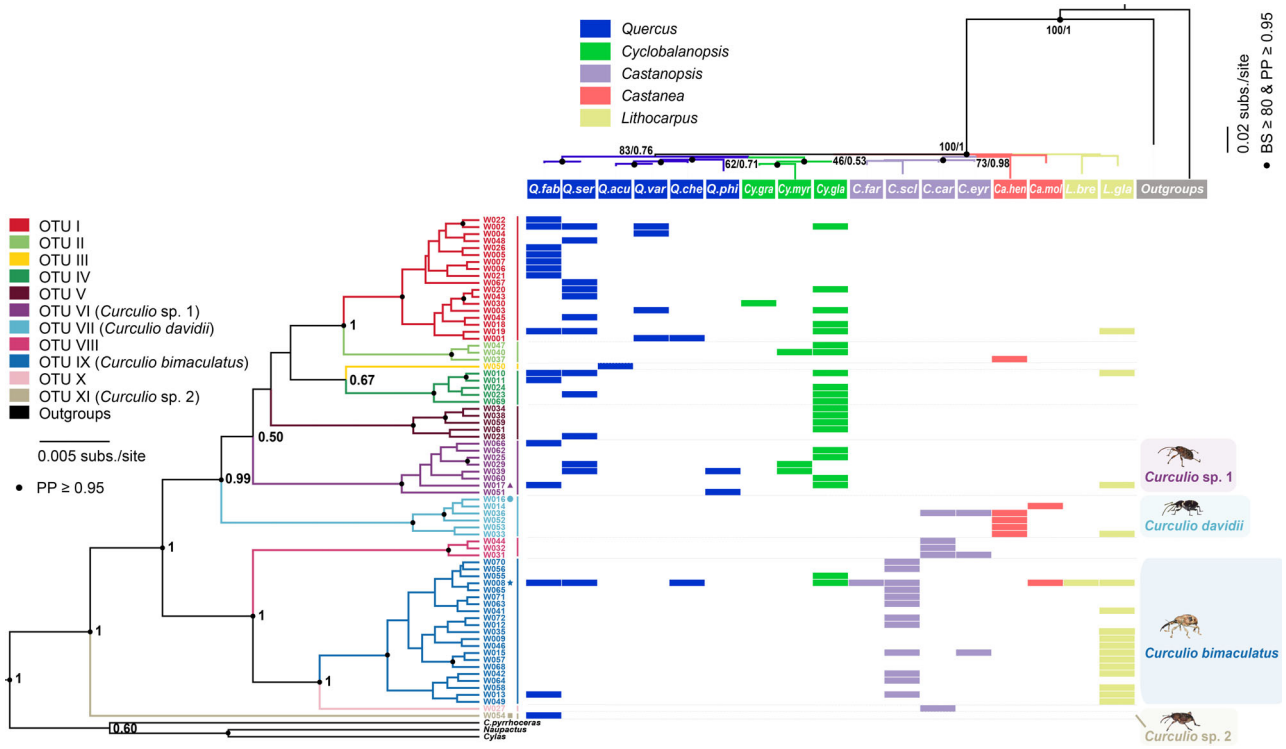

Fig 3. Host ranges of delimitated *Curculio* OTUs. Host plant species of the sampled acorn weevil larvae belonging to each haplotype of the combined mtDNA and nDNA genes are shown, and each *Curculio* OTU and Fagaceae genus is represented by a different colour. The Bayesian ultrametric tree of acorn weevils used in OTU delimitation (GMYC) is located on the left, and the phylogenetic tree based on the combined nuclear and chloroplast dataset (*ITS*/*5.8S*, *rbcL* and *matK*) from the 17 host plant species is located on the top right. The full names of host plants are listed in the Supporting Information Table S3. [Color figure can be viewed at wileyonlinelibrary.com]

one Fagaceae species (Fig. 3; Table 1). Most of the other acorn weevil OTUs parasitised seeds of Fagaceae species in the same genus or closely related genera (*Curculio davidii*, *Curculio* sp. 1 and OTU 1, 4, 5, 8), and only one *Curculio* OTU (OTU 2) exploited seeds from distantly related plant genera (Fig. 3; Table 1). These results suggested that most acorn weevil OTUs had a certain extent of host specificity.

Host specificity index (*S*) varied from 0.002 (*C. bimaculatus*) to 0.988 (OTU 3 and 10). With the exception of *C. bimaculatus*, all other acorn weevil OTUs displayed relatively high host specificity, with *S* values greater than 0.527 (Table 1). In addition, rarefaction curves showed sufficient sample sizes in all host plant genera except genera *Lithocarpus* and *Castanea* (Supporting Information Fig. S4).

Discussion

A clear description of the biodiversity and host ranges of phytophagous insects provides an essential basis for understanding the mechanisms that stabilise ecosystem functioning (Scriber, 2010; Burkle *et al*., 2013). Based on both mtDNA and nDNA, we detected high OTU diversity of acorn weevils in our study area. In addition, given the same OTU delimitation results from the combined three genes and *COI* gene alone, acorn weevil larvae can be identified only using this mtDNA locus. More importantly, host specificity varied across these *Curculio* OTUs, indicating the necessity to distinguish closely related OTUs before conducting pest control programs.

The high OTU diversity of acorn weevils may originate from their worldwide distribution and high diversity of host tree species. *Curculio* species are prevalent in tropical, subtropical and temperate forests (Pelsue & Zhang, 2000; Peguero *et al*., 2017), and thus allopatric speciation caused by biogeographic factors like the emergence of geographic barriers (e.g., Chen *et al*., 2012) must play an important role in the diversification of this genus. Alternatively, the species-rich pool of host trees also allows deep niche divergence within a *Curculio* species/OTU, likely facilitating adaptive radiations and sympatric speciation (e.g., Savolainen *et al*., 2006). However, to explore the relative contribution of the above two mechanisms of speciation, a continent-level sampling is necessary to track the evolutionary trajectories of *Curculio* OTUs.

Though we conservatively considered the clade containing the adult *C. bimaculatus* as a single OTU based on morphological characters (this approach had been reported in other studies; e.g., Blair & Bryson, 2017), it is still likely that there are two OTUs (i.e., cryptic species) within this clade (see results of GMYC in Fig. 2c). However, delimitation of cryptic species needs further validation, e.g., mating experiments between two

putative cryptic species (Hernández-Vera *et al*., 2010). In addition, insufficient sample size in two host genera (Supporting Information Fig. S4) suggested the probability of existence of some extra *Curculio* OTUs. Low density of *Castanea* species and low parasitism of their seeds in our study sites (only 231 seeds from 7 trees were collected with just 4.3% of parasitism) led to limited sample size of acorn weevil larvae from this host, while the lack of saturation in the rarefaction curve in *Lithocarpus* was largely due to the presence of many rare *Curculio* OTUs (four OTUs had only one individual). In future sampling, we will pay more attention to these two host genera in case that we missed some rare *Curculio* OTUs.

One hypothesis is that host specificity is created by long-term coevolution between phytophagous insects and their host plants (Musser *et al*., 2002; Habermannova *et al*., 2013). In our case, the high level of host specificity in most acorn weevil OTUs likely reflects the role of phylogenetic relatedness of their host Fagaceae species, because most specialist *Curculio* OTU prefer to attack plant species in the same genus or closely related genera (Fig. 3; Table 1). Given the strong impact of coevolution, differentiation among acorn weevils may be at least partly driven by variation in chemical (e.g., concentration of tannins) and physical (e.g., thickness and hardness of seed coasts) defences of host species. The adaptive radiation of acorn weevils therefore may be related to detoxification genes (Dai *et al*., 2019) and body size, which is considered to be associated with some important adaptive traits like rostrum length and is substantially controlled by size of host seeds (Hughes & Vogler, 2004).

The ongoing ecological changes may also strongly affect plant–insect interactions and host specificity (Wang *et al*., 2019b). EBLFs in our study area are facing severe deforestation and fragmentation, and such serious land use changes can rapidly reduce biodiversity (Figueroa *et al*., 2018; Guo *et al*., 2018), possibly leading to local extinction of certain host plants and narrowing the observed host range of some rare acorn weevil OTUs. This was indicated in our study by some *Curculio* OTUs utilising seeds from distantly related host genera (e.g., OTU 3 and 10; Fig. 3), and the true host ranges of these OTUs might be underestimated. Nevertheless, given that most host plant species were detected in each study site (Qian-Ya Li, personal observations), isolation by evolutionary constraints and mismatching phenology are still the most likely factors resulting in high host specificity of rare acorn weevil OTUs. To further explore these potential factors forming host specificity, we will strengthen sampling efforts in future studies.

Curculio OTUs with different host specificity may play distinct roles in the maintenance of plant diversity in EBLFs. Most *Curculio* OTUs identified in our study act as host-specific natural enemies of local dominant plants and, therefore, are likely to contribute to the density-dependent regulation of dominant Fagaceae species and facilitate the coexistence of plant species (Tong *et al*., 2017; Wang *et al*., 2019a). In contrast, generalists, especially *C. bimaculatus*, appear to be pests that drastically restrict the regeneration of most Fagaceae species and threaten the long-term persistence of EBLFs. We therefore raise the potential importance of controlling the density of generalist acorn weevils.

Acknowledgements

The authors are sincerely grateful to Jian-Hui Li, Kai Jiang, Min Liu, Yang Yang, Fang-Lu Wei, Zhe-Ting Cai, and Xue-Ke Du for their contributions in collecting the Fagaceae seeds and molecular experiments. The authors also are appreciative of the editor and the anonymous reviewers for comments on this manuscript. This study was funded by the National Key Research and Development Program (2016YFC0503102), the National Natural Science Foundation of China (31800446), and China Postdoctoral Science Foundation (2018M641965).

Data availability statement

All sequencing data used in this study have been deposited in NCBI database (https://www.ncbi.nlm.nih.gov/), with accession numbers of MN942048−MN942259.

Supporting information

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

Appendix S1. Supporting Information.

References

Arias-Leclaire, H., Bonal, R., García-López, D. & Espelta, J.M. (2018) Role of seed size, phenology, oogenesis and host distribution in the specificity and genetic structure of seed weevils (*Curculio* spp.) in mixed forests. *Integrative Zoology*, **13**, 267–279.

Bagchi, R., Gallery, R.E., Gripenberg, S., Gurr, S.J., Narayan, L., Addis, C.E., Freckleton, R.P. & Lewis, O.T. (2014) Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature*, **506**, 85–88.

Blair, C. & Bryson, R.W. (2017) Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: *Phrynosoma*). *Molecular Ecology Resources*, **17**, 1168–1182.

Bocak, L., Barton, C., Crampton-Platt, A., Chesters, D., Ahrens, D. & Vogler, A.P. (2014) Building the Coleoptera tree-of-life for > 8000 species: composition of public DNA data and fit with Linnaean classification. *Systematic Entomology*, **39**, 97–110.

Bonal, R. & Muñoz, A. (2008) Seed growth suppression constrains the growth of seed parasites: premature acorn abscission reduces *Curculio elephas* larval size. *Ecological Entomology*, **33**, 31–36.

Bonal, R. & Muñoz, A. (2009) Seed weevils living on the edge: pressures and conflicts over body size in the endoparasitic *Curculio* larvae. *Ecological Entomology*, **34**, 304–309.

Burkle, L.A., Marlin, J.C. & Knight, T.M. (2013) Plant-pollinator interactions over 120 years: loss of species, co-occurrence, and function. *Science*, **339**, 1611–1615.

Chen, Y., Compton, S.G., Liu, M. & Chen, X.-Y. (2012) Fig trees at the northern limit of their range: the distributions of cryptic pollinators indicate multiple glacial refugia. *Molecular Ecology*, **21**, 1687–1701.

Dai, L.-L., Gao, H.-M., Ye, J.-Q., Fu, D.-Y., Sun, Y.-Y. & Chen, H. (2019) Isolation of *CarE* genes from the Chinese white pine beetle

Dendroctonus armandi (Curculionidae: Scolytinae) and their response to host chemical defense. *Pest Management Science*, **75**, 986–997.

Desouhant, E., Debouzie, D., Ploye, H. & Menu, F. (2000) Clutch size manipulations in the chestnut weevil, *Curculio elephas*: fitness of oviposition strategies. *Oecologia*, **122**, 493–499.

Drummond, A.J. & Rambaut, A. (2007) BEAST: bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology*, **7**, 214.

Espelta, J.M., Bonal, R. & Sánchez-Humanes, B. (2009) Pre-dispersal acorn predation in mixed oak forests: interspecific differences are driven by the interplay among seed phenology, seed size and predator size. *Journal of Ecology*, **97**, 1416–1423.

Ezard, T., Fujisawa, T., & Barraclough, T. (2009) Splits: species' limits by threshold statistics. R package version 1.0-11.

Farrell, B.D. (1998) "Inordinate fondness" explained: why are there so many beetles? *Science*, **281**, 555–559.

Figueroa, C.C., Niemeyer, H.M., Cabrera-Brandt, M., Briones, L.M., Lavandero, B., Zuniga-Reinoso, A. & Ramirez, C.C. (2018) Forest fragmentation may endanger a plant-insect interaction: the case of the highly specialist native aphid *Neuquenaphis staryi* in Chile. *Insect Conservation and Diversity*, **11**, 352–362.

Fonseca, C.R., Prado, P.I., Almeida-Neto, M., Kubota, U. & Lewinsohn, T.M. (2005) Flower-heads, herbivores, and their parasitoids: food web structure along a fertility gradient. *Ecological Entomology*, **30**, 36–46.

Forister, M.L., Dyer, L.A., Singer, M.S., Stireman, J.O. & Lill, J.T. (2012) Revisiting the evolution of ecological specialization, with emphasis on insect-plant interactions. *Ecology*, **93**, 981–991.

Fujisawa, T. & 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.

Guo, Q.-F., Riitters, K.H. & Potter, K.M. (2018) A subcontinental analysis of forest fragmentation effects on insect and disease invasion. *Forests*, **9**, 744.

Habermannova, J., Bogusch, P. & Straka, J. (2013) Flexible host choice and common host switches in the evolution of generalist and specialist cuckoo bees (Anthophila: *Sphecodes*). *PLoS One*, **8**, e64537.

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.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S. & Francis, C.M. (2004) Identification of birds through DNA barcodes. *PLoS Biology*, **2**, e312.

Hernández-Vera, G., Mitrović, M., Jović, J., Toševski, I., Caldara, R., Gassmann, A. & Emerson, B.C. (2010) Host-associated genetic differentiation in a seed parasitic weevil *Rhinusa antirrhini* (Coleoptera: Curculionidae) revealed by mitochondrial and nuclear sequence data. *Molecular Ecology*, **19**, 2286–2300.

Higaki, M. (2016) Prolonged diapause and seed predation by the acorn weevil, *Curculio robustus*, in relation to masting of the deciduous oak *Quercus acutissima*. *Entomologia Experimentalis Et Applicata*, **159**, 338–346.

Holland, S. (1998) *aRarefactWin*. University of Georgia, Athens.

Hughes, J. & Vogler, A.P. (2004) Ecomorphological adaptation of acorn weevils to their oviposition site. *Evolution*, **58**, 1971–1983.

Kajtoch, Ł., Montagna, M. & Wanat, M. (2018) Species delimitation within the *Bothryorrhynchapion* weevils: multiple evidence from genetics, morphology and ecological associations. *Molecular Phylogenetics and Evolution*, **120**, 354–363.

Kavanagh, P.H. & Burns, K.C. (2012) Mistletoe macroecology: spatial patterns in species diversity and host use across Australia. *Biological Journal of the Linnean Society*, **106**, 459–468.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, **34**, 772–773.

Lanteri, A.A. & del Río, M.G. (2017) *Naupactus xanthographus* (Germar) species group (Curculionidae: Entiminae: Naupactini): a comprehensive taxonomic treatment. *Journal of Natural History*, **51**, 1557–1587.

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

Marvaldi, A.E., Sequeira, A.S., O'Brien, C.W. & Farrell, B.D. (2002) Molecular and morphological phylogenetics of weevils (Coleoptera, Curculionoidea): do niche shifts accompany diversification? *Systematic Biology*, **51**, 761–785.

Minh, B.Q., Nguyen, M.A.T. & von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, **30**, 1188–1195.

Muñoz, A., Bonal, R. & Espelta, J.M. (2014) Acorn-weevil interactions in a mixed-oak forest: outcomes for larval growth and plant recruitment. *Forest Ecology and Management*, **322**, 98–105.

Musser, R.O., Hum-Musser, S.M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J.B. & Felton, G.W. (2002) Herbivory: caterpillar saliva beats plant defences - A new weapon emerges in the evolutionary arms race between plants and herbivores. *Nature*, **416**, 599–600.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274.

Peguero, G., Bonal, R., Sol, D., Muñoz, A., Sork, V.L. & Espelta, J.M. (2017) Tropical insect diversity: evidence of greater host specialization in seed-feeding weevils. *Ecology*, **98**, 2180–2190.

Pelsue, F.W., Jr. & Zhang, R.Z. (2000) A review of the genus *Curculio* L. from China with descriptions of new taxa. Part I (Coleoptera: Curculionidae: Curculioninae: Curculionini). *Coleopterists Bulletin*, **54**, 125–142.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D. & Vogler, A.P. (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595–609.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. (2012) ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

R Development Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Rambaut, A., Suchard, M., Xie, W. & Drummond, A. (2014) *Tracer v. 1.6*. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012) MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Savolainen, V., Anstett, M.C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N. & Baker, W. J. (2006) Sympatric speciation in palms on an oceanic Island. *Nature*, **441**, 210–213.

Schmidt, J.M., Whitehouse, T.S., Green, K., Krehenwinkel, H., Schmidt-Jeffris, R. & Sial, A.A. (2019) Local and landscape-scale heterogeneity shape spotted wing drosophila (*Drosophila suzukii*) activity and natural enemy abundance: implications for trophic interactions. *Agriculture Ecosystems & Environment*, **272**, 86–94.

Scriber, J.M. (2010) Integrating ancient patterns and current dynamics of insect-plant interactions: taxonomic and geographic variation in herbivore specialization. *Insect Science*, **17**, 471–507.

Sequeira, A.S., Sijapati, M., Lanteri, A.A. & Albelo, L.R. (2008) Nuclear and mitochondrial sequences confirm complex colonization patterns and clear species boundaries for flightless weevils in the Galápagos archipelago. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **363**, 3439–3451.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the entomological Society of America*, **87**, 651–701.

Smith, M.A., Fisher, B.L. & Hebert, P.D.N. (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **360**, 1825–1834.

Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W. & Hebert, P. D.N. (2006) DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 3657–3662.

Song, H., Buhay, J.E., Whiting, M.F. & Crandall, K.A. (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13486–13491.

Song, Y.-C., Chen, X.-Y. & Wang, X.-H. (2005) Studies on evergreen broad-leaved forests of China: a retrospect and prospect. *Journal of East China Normal University (Natural Science)*, **1**, 1–8.

Swofford, D.L. (2002) *PAUP. Phylogenetic analysis using parsimony (and other methods), version 4.0b10*. Sinauer Associates, Sunderland, Massachusetts.

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

Tong, X., Zhang, Y.-X., Wang, R., Inbar, M. & Chen, X.-Y. (2017) Habitat fragmentation alters predator satiation of acorns. *Journal of Plant Ecology*, **10**, 67–73.

Vaidya, G., Lohman, D.J. & Meier, R. (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Wang, R., Chen, X.-Y., Chen, Y., Wang, G., Dunn, D.W., Quinnell, R. J. & Compton, S.G. (2019b) Loss of top-down biotic interactions changes the relative benefits for obligate mutualists. *Proceedings of the Royal Society B-Biological Sciences*, **286**, 20182501.

Wang, R., Shi, Y.-S., Zhang, Y.-X., Xu, G.-F., Shen, G.-C. & Chen, X.-Y. (2019a) Distance-dependent seed–seedling transition in the tree *Castanopsis sclerophylla* is altered by fragment size. *Communications Biology*, **2**, 277.

Wang, R., Zhang, X., Shi, Y.-S., Li, Y.-Y., Wu, J.-G., He, F.-L. & Chen, X.-Y. (2020) Habitat fragmentation changes top-down and bottom-up controls of food webs. *Ecology*, **101**, e03062.

Zhang, A.-B., Hao, M.-D., Yang, C.-Q. & Shi, Z.-Y. (2017) BarcodingR: an integrated R package for species identification using DNA barcodes. *Methods in Ecology and Evolution*, **8**, 627–634.

Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W.-X. & Wang, G.-T. (2020) PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources*, **20**, 348–355.

Accepted 5 November 2020

Editor: Yves Basset; Associate Editor: Matthew Forister