6-10-2019

Diversification of mitogenomes in three sympatric Altica flea beetles (Insecta, Chrysomelidae)

Rui-E Nie
Chinese Academy of Sciences, Beijing

Jing Wei
Chinese Academy of Sciences, Beijing & University of Chinese Academy of Sciences, Beijing

Shou-Ke Zhang
Chinese Academy of Sciences, Beijing & Chinese Academy of Forestry, Fuyang

Alfried P. Vogler
atural History Museum, London & Campus, Imperial College London, Ascot

Ling Wu
Chinese Academy of Sciences, Beijing & ebei University, Baoding

See next page for additional authors

Received: 27 February 2019 | Revised: 20 May 2019 | Accepted: 10 June 2019

DOI: 10.1111/zsc.12371

ORIGINAL ARTICLE

Zoologica Scripta | WILEY

Diversification of mitogenomes in three sympatric *Altica* flea beetles (Insecta, Chrysomelidae)

Rui-E Nie[1] | **Jing Wei**[1,2] | **Shou-Ke Zhang**[1,3] | **Alfried P. Vogler**[4,5] | **Ling Wu**[1,6] | **Alexander S. Konstantinov**[7] | **Wen-Zhu Li**[1] | **Xing-Ke Yang**[1] | **Huai-Jun Xue**[1]

[1]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Fuyang, China

[4]Department of Life Sciences, Natural History Museum, London, UK

[5]Department of Life Sciences, Silwood Park Campus, Imperial College London, Ascot, UK

[6]College of Life Sciences, Hebei University, Baoding, China

[7]Systematic Entomology Laboratory, USDA, ARS, Washington, DC, USA

Correspondence
Huai-Jun Xue and Xing-Ke Yang, Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, No. 1 Beichen West Road, Chaoyang District, Beijing 100101, China.
Emails: xue@ioz.ac.cn (HJX) and yangxk@ioz.ac.cn (XKY)

Funding information
National Natural Science Foundation of China, Grant/Award Number: 31772496 and 31301900; Key Laboratory of the Zoological Systematics and Evolution of the Chinese Academy of Sciences, Grant/Award Number: Y229YX5105

Abstract

The Asian flea beetles *Altica cirsicola*, *Altica fragariae* and *Altica viridicyanea* are broadly sympatric and morphologically highly similar but feed on distantly related host plants. They have been suggested as a model for ecological speciation studies. However, their phylogeny and species limits remain uncertain. In this study, we added mitochondrial genomes from multiple individuals of each species to the growing database. Phylogenetic analyses based on 15 genes showed clear interspecific divergences of *A. fragariae* from the other species, but *A. cirsicola* and *A. viridicyanea* were not distinguishable by distance-based or tree-based methods of species delimitation due to non-monophyly of mitogenomes relative to the morphologically defined entities, possibly affected by interspecific introgression. This was confirmed by wider sampling of mitochondrial COX1 (58 individuals) and the second internal transcribed spacer of nuclear ribosomal RNA cluster (ITS2; 68 individuals), which showed that ITS2, but not COX1, coincided with the morphological species limits. The full mitochondrial genomes are not able to shed further light on the species status, even with the most sensitive approach based on diagnostic characters, yet the whole mitogenome is useful to get improved estimates of intra- and interspecific variation, not affected by the stochastic error seen in individual genes.

KEYWORDS

Altica, COX1, genetic distance, ITS2, mitogenome, species delimitation

1 | INTRODUCTION

The cosmopolitan genus *Altica* Geoffroy (Coleoptera: Chrysomelidae: Galerucinae) is one of the most taxonomically difficult flea beetle genera with about 235 described species, being referred to as a "taxonomic nightmare" (Reid & Beatson, 2015). Among them, *Altica cirsicola* Ohno, *Altica fragariae* Nakane and *Altica viridicyanea* (Baly) (abbreviated

AC, AF and AV) are highly similar morphologically but can be distinguished reliably in the males by the shape of the median lobe of the aedeagus (Yu, Wang, & Yang, 1996). These species are distributed sympatrically over a wide area of Eastern Asia and are likely to be each others' closest relatives (Zhai, Xue, Wang, & Yang, 2007). Their postmating isolation is incomplete, and they can easily be crossed under laboratory conditions (Xue, Li, Nie, & Yang, 2011; Xue, Li, & Yang, 2009; Xue, Magalhães, Li, & Yang, 2009), further indicating their close affinity. The three species are narrow specialists each with distantly related host plants. While *A. cirsicola* feeds only on *Cirsium* (Asteraceae) and *A. viridicyanea* occurs exclusively on *Geranium nepalens* (Sweet; Geraniaceae; treated as *Geranium wilfordii* in Xue, Li, et al., 2009), *A. fragariae* is an oligophagous species with five recorded host plants in the family Rosaceae (*Duchesnea indica* is the primary one). These differences suggest that speciation may be associated with host switching. Therefore, these three species offer an interesting model system for host specialization and ecological speciation studies (Xue et al., 2011, 2018, 2016; Xue, Li, & Yang, 2014). Their relationship has been explored using small fragments of mitochondrial DNA (the 3′ end of COX1 and the 5′ end of COX2), an internal transcribed spacer (ITS2) of the nuclear ribosomal RNA cluster and a nuclear gene (EF1α; Xue et al., 2011), but different relationships were inferred depending on the gene used. Mitochondrial genes suggested that AF is distantly related to AC and AV, while individuals of the latter could not be clearly separated. The nuclear genes showed different results: the relationship of (AC, (AF, AV)) was supported by an intron of EF1α, while the relationship of (AV, (AC, AF)) was supported by an exon of this gene. The relationships among the three species could also not be resolved confidently based on ITS2. However, in all cases the nuclear markers assigned each individual correctly to the expected species (Xue et al., 2011), unlike the mitochondrial markers which were sometimes inconsistent with the morphological assignments.

Possibly, more individuals from more populations could achieve better resolution. In addition, longer sequence fragments (e.g., complete mitogenomes) may provide greater power to obtain phylogenetic resolution for closely related genotypes (Jacobsen et al., 2012; Logue et al., 2013; Marques et al., 2017; Wielstra & Arntzen, 2011). We therefore sequenced multiple complete mitogenomes for each of the three *Altica* species, specifically to test the performance of whole-mitogenomes in species delimitation and phylogenetic inference, to evaluate if full mitogenomes can overcome the deficiencies (e.g., less resolution power for delimitation of closely related species) of a single mitochondrial gene marker. We also used a wider sampling of populations based on the COX1 and ITS2 markers, as a further possibility to refine the species limits.

2 | MATERIAL AND METHODS

2.1 | The species identification and DNA extraction of samples

Whereas, identification of females using morphological characters is often problematic for *Altica* species (Konstantinov, 1987), males can be identified by their different genitalia. Subtle but consistent differences of genitalia structure, for example, the shape of ridges on ventral side, were used to distinguish these three species (Warchalowski, 2010; Yu et al., 1996; Figure S1), in combination with the records of host plants from which the specimens were collected, as an important supplementary evidence (Yu et al., 1996). Samples were collected from 2009 to 2015, preserved in 100% ethanol and stored at −30°C until DNA extraction. DNA extraction, sequencing and assembly followed Nie et al. (2018) using TIANprep Midi Plasmid kit (TIANGEN).

2.2 | Sequencing, assembly, annotation and alignment of complete mitogenomes

Mitochondrial genome sequencing was achieved by genome skimming from shotgun sequencing of total DNA. We used the Illumina HiSeq2000 platform with 200 bp insert size and paired-end (2 × 100 bp) sequencing (lib51, lib52, lib53) or Hiseq 2500 with 450 bp insert size and 2 × 250 bp paired-end sequencing for all others (Table S1). The sequence reads were first filtered following Zhou et al. (2013). The remaining high-quality reads were assembled using IDBA-UD (Peng, Leung, Yiu, & Chin, 2012) and SPAdes (Bankevich et al., 2012). Totally, 12 mitochondrial genome sequences were sequenced newly (GenBank accession numbers: MH477591–MH477602). The mitogenome annotation was conducted with Geneious v.8.0.5 (Kearse et al., 2012) using the mitogenome of *Altica ericeti* (GenBank accession number: KX943460) as a reference. The tRNAs were identified by tRNAscan-SE 1.21 (Lowe & Eddy, 1997) with default settings. Some tRNA genes that could not be found by tRNAscan-SE 1.21 were identified by visual comparison to the reference mitogenome.

Protein-coding genes (PCGs) were aligned with TransAlign (Bininda-Emonds, 2005). Sequences of rRNA genes (*12S rRNA*, *16S rRNA*) were aligned separately with Muscle v.3.8.31 (Edgar, 2004) under default parameters. The aligned data from each locus were concatenated using SequenceMatrix v.1.7.8 (Vaidya, Lohman, & Meier, 2011).

2.3 | Polymerase chain reactions (PCR), sequencing and alignment of population samples

To analyse the species boundaries across the wider geographic ranges, COX1 and the unlinked nuclear ITS2 of the rRNA cluster were sequenced from further specimens (Table

S2). Genomic DNA extraction was performed as above. Polymerase chain reactions (PCR) amplified two gene regions: the standard 5′ COX1 barcode region covering 658 bp using primers LCO1490 and HCO2198 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994), and a ~354 bp (340–354 bp) of ITS2 using primers FB5.8SFWD and FB28SREV (Ruhl, Wolf, & Jenkins, 2010). The PCR conditions and sequence methods of COXI and ITS2 were the same as in Xue et al., (2011). COX1 and ITS2 sequences for 46 individuals (AC, $n = 18$; AF, $n = 8$; AV, $n = 20$; GenBank accession numbers: MH477438–MH477483 and MH477484–MH477529) were obtained newly. Gene sequences were edited and aligned using CodonCode Aligner 3.7.1 (CodonCode). No ITS2 sequence was found to be heterozygous, as judged by the lack of double peaks in chromatograms from both directions.

2.4 | Species delimitation and phylogenetic analysis using mitogenomes

Intra- and interspecific sequence divergence of gene sequences was calculated with MEGA 7.0 (Kumar, Stecher, & Tamura, 2016) using Kimura-2-parameter (K2P) distances. Distance-based species delimitation was conducted with the Automatic Barcode Gap Discovery (ABGD) tool which searches for a "barcoding gap" in sequence similarity in all-against-all pairwise comparisons (Puillandre, Lambert, Brouillet, & Achaz, 2012). ABGD was employed to determine the number of operational taxonomic units, and the analysis was performed using the web interface (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) using default parameters. The K2P model was set for distance calculation, and the relative gap width (X) was set to 1.5. The range of prior intraspecific divergence from 0.001 to 0.1 was recorded with 10 steps. Tree-based methods considered a species to be correctly identified if the conspecific sequences were monophyletic (Virgilio, Backeljau, Nevado, & De Meyer, 2010). Each gene individually and the combined PCGs were used to infer Neighbour-joining (NJ) trees (Saitou & Nei, 1987) with MEGA 7.0 based on K2P distances.

Character-based species delimitation was applied using the "diagnostic logic mining method," which extracts nucleotide sites for diagnosis of each species, implemented in BLOG 2.0 (Bertolazzi, Felici, & Weitschek, 2009; Weitschek, van Velzen, Felici, & Bertolazzi, 2013). Default settings were used to find a logic formula of unique nucleotides in the expanded COX1 and ITS2 datasets defining each of the three nominal species. For COX1, the mitogenome sequences and for ITS2, 22 published sequences were used as training sets.

We assessed the best partitioning strategy in PartitionFinder v. 1.1.1 (Lanfear, Calcott, Ho, & Guindon, 2012) for the concatenated mitochondrial matrix, prior to phylogenetic inferences using MrBayes v.3.2 (Ronquist & Huelsenbeck, 2003), and RAxML 8.2.10 (Stamatakis, 2006). For MrBayes analyses, the most appropriate nucleotide substitution model was selected using the Akaike Information Criterion in jModelTest 0.1.1 (Posada, 2008), which was determined for each partition. The MCMC search was conducted for a minimum of 100,000 generations, and sampling was done every 100 generations until the average standard deviation of split frequencies was <0.01. The first 25% of trees were discarded as burn-in and posterior probabilities were estimated for each node. The RAxML tree was calculated with branch support estimated from 1,000 bootstrap replicates. The CAT–GTR model was chosen for the bootstraping phase. The tree was rooted post hoc using *A. ericeti* (GenBank accession number: KX943460) and *Diphaltica* sp. (GenBank accession number: MF351888) as outgroups (Gómez-Rodríguez, Crampton-Platt, Timmermans, Baselga, & Vogler, 2015; Nie et al., 2018) (Table S1).

3 | RESULTS

3.1 | Mitogenome organization and gene content

The sequencing effort resulted in twelve complete mitogenomes, two of which lacked the rRNA genes (Table S3). The mitogenomes in all three species were composed of 37 genes (13 protein- coding, 22 tRNAs and two rRNAs) usually present in insects, in addition to a large non-coding control region, and were arranged in the presumed ancestral insect gene order. Sequence length in *A. cirsicola*, *A. fragariae* and *A. viridicyanea* comprised 14,535, 14,533 and 14,537 bps, respectively, excluding the variable AT-rich region (Table S3). The length of the tRNAs ranged from 56 to 72 bp (Table S3). All anticodons of tRNAs were conservative except for a UCU to UUU base change in the anticodon of $tRNA^{Lys}$ already described for Chrysomelidae (Nie et al., 2018; Timmermans et al., 2016). All PCGs started with ATN start codons except ND1 with TTG, consistent with the typical coleopteran mitogenomes (Sheffield, Song, Cameron, & Whiting, 2008). The termination codons of 13 PCGs were TAA/TAG or truncated termination codons TA (ATP6 and ND4) or T (ND5), which was presumed to be completed via posttranscriptional polyadenylation (Montagna et al., 2012; Sheffield et al., 2008). The presence of partial stop codons is well documented in insects (Beard, Hamm, & Collins, 1993; Castro, Ruberu, & Dowton, 2006; Coates, Sumerford, Hellmich, & Lewis, 2005; Sheffield et al., 2008).

3.2 | Species delimitation and phylogenetic relationships using mitogenomes

Accepting the morphologically defined units, the intraspecific nucleotide divergence was generally below 1%, but differed among the three species and 13 PCGs (Table 1).

TABLE 1 Intraspecific and interspecific mean nucleotide divergence (Kimura-2-parameter) of *Altica* species inferred from single mitochondrial genes and combined 13 PCGs (mitogenome sampling)

	Intraspecific distance ± *SD* (%)			Interspecific distance ± *SD* (%)		
Gene name	**AC**	**AF**	**AV**	**AC-AF**	**AC-AV**	**AF-AV**
ND2	0.60 ± 0.20	0.07 ± 0.05	0.30 ± 0.13	3.65 ± 0.05	0.65 ± 0.15	3.97 ± 0.18
COX1	0.55 ± 0.18	0.10 ± 0.06	0.29 ± 0.11	3.11 ± 0.19	0.55 ± 0.16	3.31 ± 0.15
COX2	0.87 ± 0.17	0.65 ± 0.27	0.36 ± 0.18	2.31 ± 0.55	0.67 ± 0.20	2.12 ± 0.33
ATP8	0	0	0.66 ± 0.38	3.34 ± 0	0.33 ± 0.33	3.34 ± 0.48
ATP6	0.40 ± 0.30	0.30 ± 0.30	0.07 ± 0.07	2.95 ± 0.09	0.26 ± 0.23	3.01 ± 0.07
COX3	0.51 ± 0.21	0.34 ± 0.26	0.60 ± 0.44	2.95 ± 0.34	0.63 ± 0.40	2.89 ± 0.44
ND3	0.96 ± 0.49	0.29 ± 0.29	0.43 ± 0.28	2.03 ± 0.72	0.72 ± 0.56	2.55 ± 0.41
ND5	0.91 ± 0.23	0.45 ± 0.19	1.35 ± 0.57	2.50 ± 0.33	1.15 ± 0.36	2.04 ± 0.44
ND4	0.41 ± 0.10	0.36 ± 0.26	1.04 ± 0.24	1.91 ± 0.15	0.86 ± 0.19	1.61 ± 0.52
ND4L	0.73 ± 0.73	0.24 ± 0.17	0	2.19 ± 0.37	0.36 ± 0.63	2.37 ± 0.19
ND6	0.51 ± 0.10	1.03 ± 0.40	0.20 ± 0.12	2.77 ± 0.60	0.36 ± 0.14	2.74 ± 0.54
CYTB	0.66 ± 0.27	0.26 ± 0.21	0.47 ± 0.30	2.99 ± 0.27	0.86 ± 0.19	2.51 ± 0.35
ND1	0.34 ± 0.11	0.32 ± 0.32	0.16 ± 0.10	2.65 ± 0.20	0.37 ± 0.13	2.72 ± 0.11
13 PCGs	0.60 ± 0.07	0.33 ± 0.08	0.56 ± 0.12	2.73 ± 0.18	0.69 ± 0.06	2.64 ± 0.23

Abbreviations: AC, *Altica cirsicola*; AF, *Altica fragariae*; AV, *Altica viridicyanea*; PCGs, protein-coding genes.

Intraspecific variation calculated for the entire set of PCGs was lowest in AF at 0.33%, and approximately twice as high in the other two species, but these values subsume great differences in the estimates of variation for the individual genes, including some cases of no variation at all (ATP8 in AC and AF and ND4L in AV). The interspecific divergences were 2.73% (2.43%–3.10%), 0.69% (0.55%–0.80%) and 2.64% (2.27%–3.08%) between AC-AF, AC-AV and AF-AV, confirming the distant position of AF, and the recognition of AC + AV as a single unit based on pairwise sequence divergences of mitogenomes.

The ABGD analyses largely supported the recognition of two units based on the full-length mitogenomes, but some variability in inferred species numbers was observed when using single genes only (Table S4). Only the ND4L locus supported three genetically distinct entities, but they did not fully correspond to the three species as defined by morphology.

The phylogenetic analyses of combined genes (Figures 1 and S2) and most of the single genes and control region (Table S5) supported the relationship of AF as sister to the AC + AV complex, with a few exceptions of unresolved relationships in particular loci, but the reciprocal monophyly of AC and AV was not supported by any tree (Figures 1 and S2; Table S5).

3.3 | Species delimitation of COX1/ITS2 dataset

For the wider population sampling, the COXI matrix included 58 sequences (46 sequences from PCR amplification plus 12 mitogenomes), with *A. ericeti* (KX943460) and *Altica koreana* (MK138539) as outgroups. Intraspecific nucleotide divergence was 0.58%, 0.36% and 0.55% in AC, AF and AV, respectively; the interspecific nucleotide divergence was 3.50%, 0.77% and 3.36% in AC-AF, AC-AV and AF-AV, which was slightly higher than in the respective COX1 sequences of the mitogenomes. The ITS2 matrix included 68 sequences (46 new sequences and 22 downloaded from GenBank, KJ803199, KJ803200, JN903083–JN903102), and *A. ericeti* (KF163167) and *A. koreana* (JN903103) as outgroups. The intraspecific nucleotide divergence of ITS2 was 0.10%, 0.06% and 0.02% in AC, AF and AV, respectively; the interspecific nucleotide divergence was 1.09%, 1.75% and 0.64% in AC-AF, AC-AV and AF-AV (Table S6). For the COX1 barcode region, 1, 2, 4 and 29 operational taxonomic units (OTUs) were produced by the ABGD analysis according to different a priori threshold values of 0.0077, 0.0046, 0.0017 (or 0.0028) and 0.0010, respectively, while for ITS2, 5 (priori threshold value: 0.0046) and 6 (priori threshold values: 0.0028, 0.0017, 0.0010) OTUs were suggested.

The NJ tree based on COX1 ($n = 58$) showed that individuals of AF and AC + AV formed two well supported clusters except for one individual of AV (SX1314) that fell into the cluster of AF (Figure 2). Again, AC and AV were not recognized as distinct lineages, as AV was nested within AC. Furthermore, the AV cluster was weakly supported (43%) with one individual morphologically assigned to AV (SX1320) not part of this cluster (Figure 2). The Bayesian tree and ML tree showed similar topology to the NJ tree but with the individual SX1320 clustered to AC + AV. In the NJ, Bayesian and ML trees from

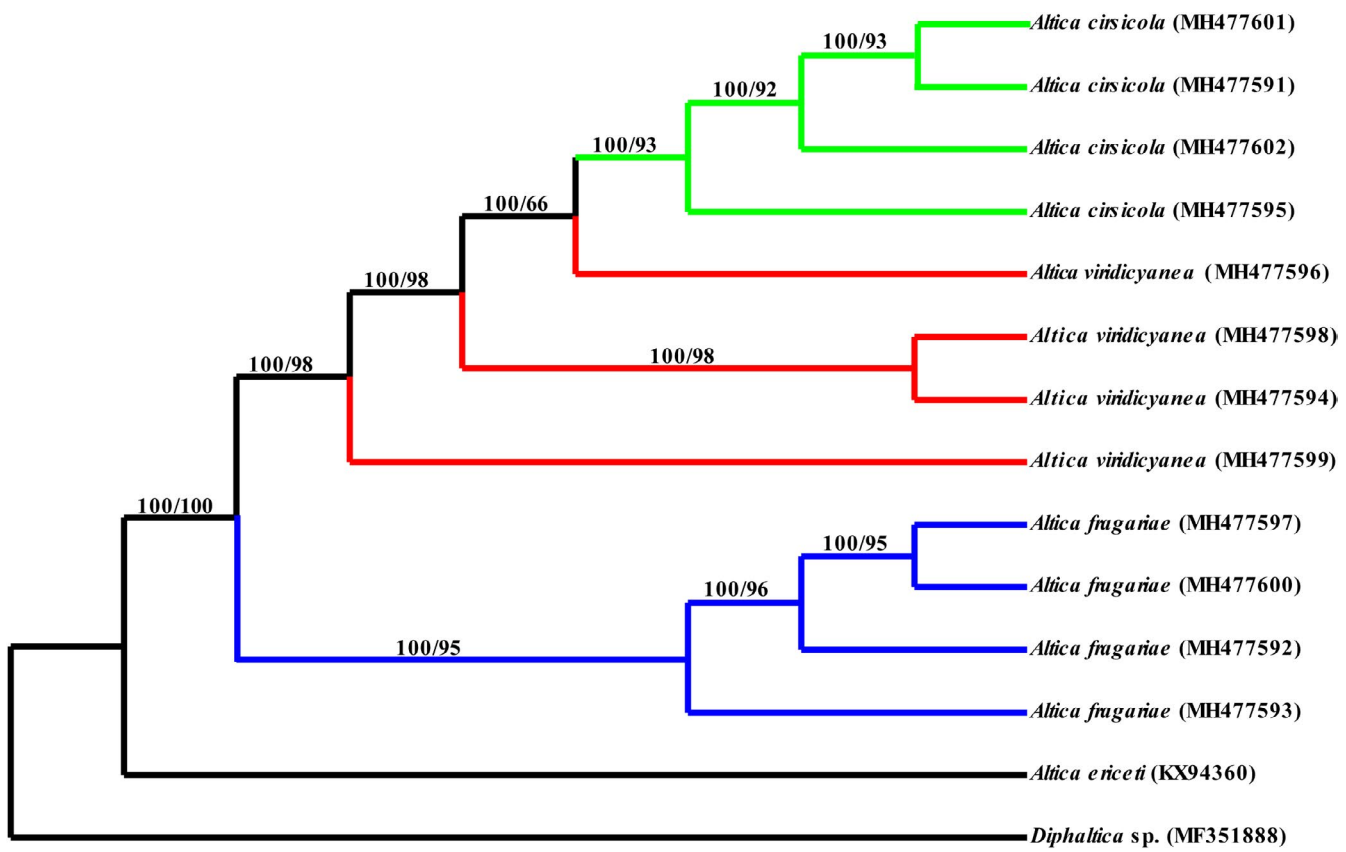

FIGURE 1 Phylogenetic tree based on the combined data of 13 protein-coding genes performed using MrBayes v.3.2 and RAxML 8.2.10. Bayesian analysis posterior probabilities (first number) and RAxML bootstrap support values (second number) >50% are shown on the branches [Colour figure can be viewed at wileyonlinelibrary.com]

ITS2 sequences, AC and AV were recovered as monophyletic (n = 68) although monophyly of the assemblage AF was not supported (Figure 3).

Compared with the distance-based and tree-based methods, character-based DNA analysis using BLOG supported clear genetic disparity of the three species. Following initial training runs (see section 2), the three *Altica* species were defined by the following nucleotide formulas by BLOG: for COX1: AC: position 59 = T and position 205 = A; AF: position 49 = C; AV: position 205 = C. These diagnostic sites achieved a correct classification for all individuals of the three species, except for one individual of AV (SX1314) that was wrongly classified to AF. For ITS2, AC: position 303 = A; AF: position 143 = C and position 303 = G; AV: position 143 = T achieved a correct classification for all individuals of the three *Altica* species.

4 | DISCUSSION

The mitogenome sequencing confirmed earlier studies (Xue et al., 2011; Zhai et al., 2007) establishing the close relationship of AC and AV relative to the more distant AF, based on similar divergence pattern in most PCG or in the 13 combined PCGs (Table 1). Furthermore, compared with the interspecific divergence between them, AC and AV showed high intraspecific variation which led to the absence of a clear-cut gap. Accordingly, in most cases, neither distance-based nor tree-based methods can assign the individuals to AC or AV based on mitochondrial genomes (Tables S4 and S5). Only the character-based analysis using BLOG on the wider dataset was able to classify most individuals correctly based on COX1, but this is a post hoc analysis that selects among numerous potential nucleotide positions based on a training set of specimens allocated to predefined species, that is the method does not provide de novo species delimitations. However, the fact that virtually all newly sequenced specimens could be assigned to these preexisting species definitions in the training set indicates a certain degree of subdivision recognizable from the COX1 marker, even if it is not sufficient to reach reciprocal monophyly, as required for the tree-based species definition, or to reach significant distance thresholds in distance-based methods, which require a larger number of character changes to detect species divergence.

As it stands, the lack of a "barcode gap" and shallow divergence of mitogenomes make it difficult to separate these species with the standard COX1 marker, and the full mitochondrial genomes added here are not able to shed further

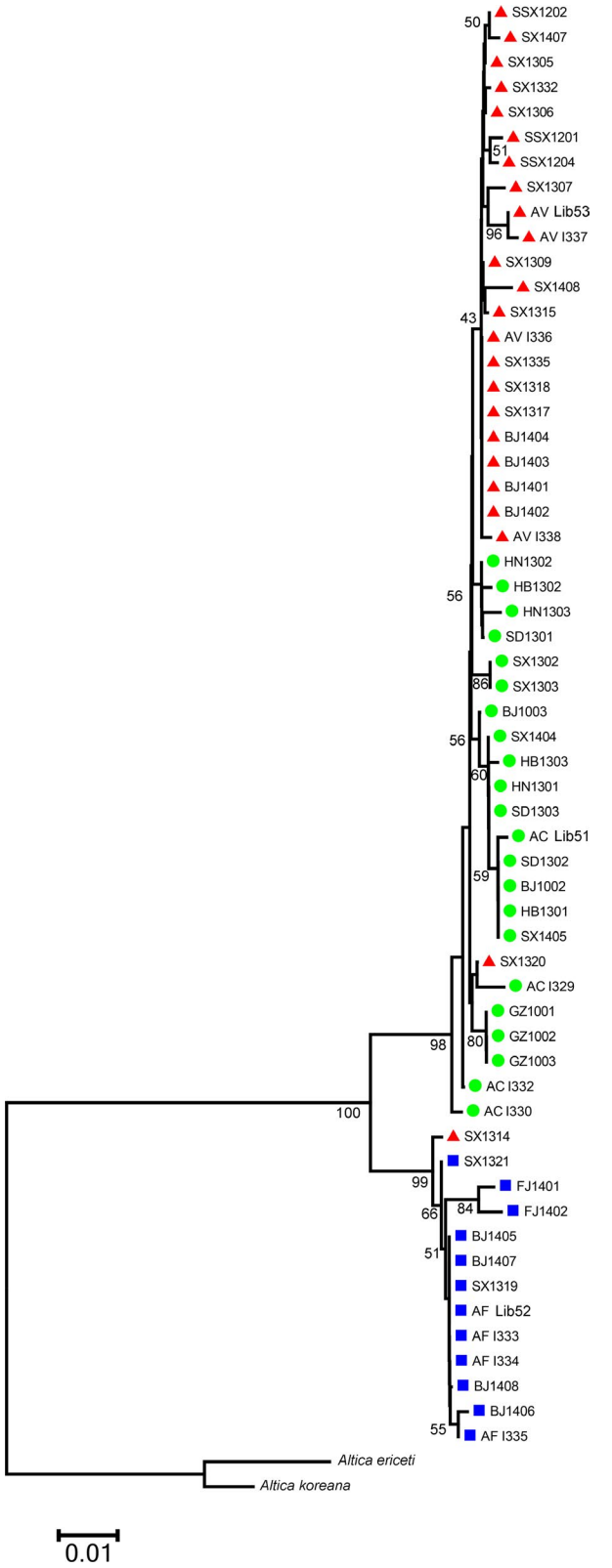

FIGURE 2 Neighbour-joining tree based on Kimura-2-parameter model for COX1 barcoding fragment (658 bp, 5′ end to middle) from *Altica cirsicola* ($n = 22$), *Altica fragariae* ($n = 12$), *Altica viridicyanea* ($n = 24$) and two outgroups. Nodal values of bootstrap support are shown. Green circle: AC; blue square: AF; red triangular: AV [Colour figure can be viewed at wileyonlinelibrary.com]

light on the species status, even with the most sensitive approach based on diagnostic characters. Individual markers, including COX1, mostly produce the same outcome as whole mitogenomes, but not in all cases. However, the whole mitogenome is useful to get improved estimates of intra- and interspecific variation, not affected by the stochastic error seen in individual genes. Further, COX1 sequences essentially confirm the mitogenome analysis from the smaller subsample, which were obtained from a single site for each species. As expected (Bergsten et al., 2012), the sampling of a wider geographic range increased the intraspecific variation (Tables 1 and S6). COX1 is therefore a good marker to capture the diversity of mitogenomes without the need for sequencing the whole mitogenome. The ITS2 has been proven to be useful for phylogenetic analyses at the species level and also has unveiled a different ability to identify closely related species in various groups (Coleman, 2003; Yao et al., 2010). Here, it was proven as a more suitable genetic marker for species delimitation in the *Altica* system when we expanded taxon sampling.

Although the taxonomic study of *Altica* is notoriously difficult (Jäkel, Mora, & Dobler, 2013; Magoga, Sahin, Fontaneto, & Montagna, 2018; Reid & Beatson, 2015), several types of data can be used for identification of those three species, for example, morphology of male genitalia and distinct host plants. So the possibility of imperfect taxonomy could be excluded. Paralogous pseudogenes are equally unlikely because abnormal variants (e.g., disrupted reading frames, stop codons) were not detected (Jäkel et al., 2013). For recently diverged species, incomplete lineage sorting may be invoked, which might explain the non-monophyly of those three *Altica* species (Figures 2 and S2). However, the shallow divergence between AC and AV inferring from mitochondrial sequences was not supported by ITS2 data and, on the contrary, AC displayed the largest distance of ITS2 to AV (Table S6). We reanalysed the EF1α dataset of previous studies (Xue et al., 2011, 2014), which showed that the largest interspecific divergence of exon sequences is between AC and AV (0.43%; consistent with ITS2), while the interspecific divergence of intron is between AF and AV (3.01%). In fact, the nuclear markers in particular support the morphospecies because of perfect congruence. Although the results inferred from nuclear markers were not consistent either, they still implied that mitochondrial genes may not unravel the species genealogy.

The most likely reason for non-monophyly (Figure 2) inferred from mitochondrial genes is interspecific introgression. It is widely accepted that mitochondrial genomes can be replaced by that of another species without leaving any trace in the nuclear genome because of historical genome introgression (Liu et al., 2010; Melo-Ferreira, Boursot, Suchentrunk, Ferrand, & Alves, 2005; Wilson & Bernatchez, 1998). In a previous more extensive study of AF and AV, even in these two

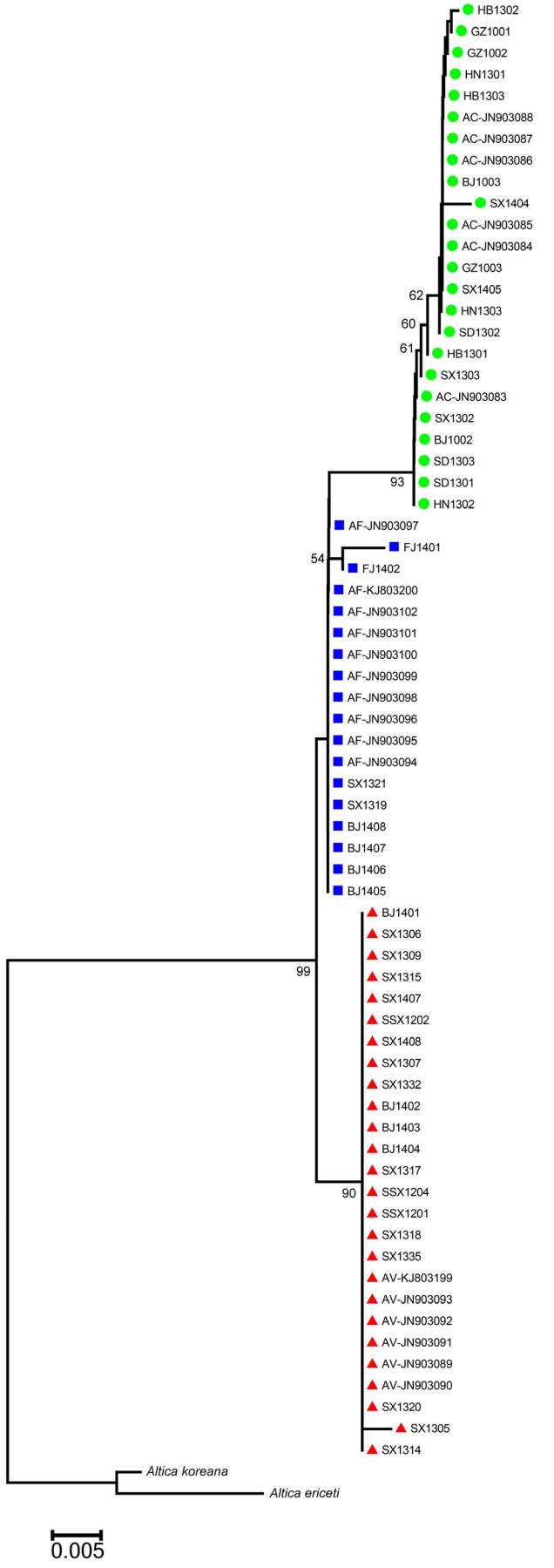

FIGURE 3 Neighbour-joining tree based on Kimura-2-parameter model for ITS2 sequence from *Altica cirsicola* ($n = 24$), *Altica fragariae* ($n = 18$), *Altica viridicyanea* ($n = 26$) and two outgroups. Nodal values of bootstrap support are shown. Green circle: AC; blue square: AF; red triangular: AV [Colour figure can be viewed at wileyonlinelibrary.com]

species mitochondrial markers indicated historical gene flow in an area of sympatry, albeit to a very limited degree, unlike the nuclear markers in these populations (Xue et al., 2014). In fact, the introgression could be prevalent more generally in *Altica*. The polyphyly of *Altica lythri* inferred from mitochondrial data was interpreted as introgression by interspecies hybridization and subsequent backcrossing (Jäkel et al., 2013). In another study of 14 *Altica* morphospecies from the Euro-Mediterranean area, seven shared haplotypes with at least one other species (Magoga et al., 2018). In the present study, the more likely scenario is the recent introgression of mitochondrial haplotypes from AC to AV or vice versa, at a recent point in time, after which only minimal divergence of these two populations has occurred. It is interesting that character-based species delimitation is effective in the face of mitochondrial introgression or even complete replacement of one species' original mitochondrial genes. We speculate the diagnostic substitutions must have arisen and come to fixation after the mtDNA swap.

The three *Altica* species can be crossed with one another under laboratary conditions: For AF-AC and AF-AV combinations, high hatch rates were achieved when AF is the male parent; for AC-AV combination, the hatch rate is considerable in both cross directions. Furthermore, some backcrosses and F_2 are also viable (Xue et al., 2011, 2014, 2018, 2016). The shallow divergence between AC and AV inferred from mitochondrial sequences is consistent with this observation. In the current study, although each *Altica* species with a larger sampling and more extensive mitochondrial sequencing was not supported to be monophyletic, the character-based BLOG analysis, which is based on species specific nucleotide sites, supported a considerable genetic disparity even for the AC-AV split. What is more important, the analysis based on nuclear markers always suggest their distinct genetic lineages (Xue et al., 2011; Xue et al., 2014; present study). To sum up, the molecular genetic analyses gave enough evidence to support existence of three separate species.

It was suggested that at least 10% of animal species are involved in interspecific hybridization in the field (Abbott et al., 2013; Mallet, 2005) and even between non-sister species (Dasmahapatra, Silva-Vásquez, Chung, & Mallet, 2007). Therefore, neither incomplete postmating isolation estimated in the laboratory nor limited interspecific gene flow detected in field samples in the *Altica* system can be treated as solid evidence for incomplete speciation. For sympatric populations, the key mechanism ensuring reproductive isolation is assortative mating (Kondrashov & Shpak, 1998; Malausa et al., 2005)

which may occur due to ecological, temporal or behavioural isolation (Funk, Filchak, & Feder, 2002). In this *Altica* system, the strong assortative mating achieved by a combination of reduced encounter rates due to differential host preference and strong sexual isolation supports the recognition of separate species (Xue et al., 2014). Furthermore, behavioural isolation among these species is dominated by species-specific cuticular hydrocarbon profiles (Xue et al., 2016; H. J. Xue, unpublished data), which also can be treated as a diagnostic trait in insect taxonomy (Bagnères & Wicker-Thomas, 2010; Kather & Martin, 2012; Pokorny, Lunau, Quezada-Euan, & Eltz, 2014). The species status was also supported by the morphological characters, which are the basis for most of the above, because samples were assigned to species according to these characters, although female identification by morphological characters in *Altica* is often problematic (Konstantinov, 1987). The subtle but consistent difference of genitalia in the males (Jäkel et al., 2013; Reid & Beatson, 2015), in particular the different shape of ridges on the ventral side thus, is confirmed by the molecular data (the ITS2 in particular) as a reliable diagnostic feature to distinguish these three species effectively (Yu et al., 1996; Figure S1).

ACKNOWLEDGEMENTS

We thank Min Tang and Xin Zhou (China National GeneBank-Shenzhen, BGI-Shenzhen), Berry Genomics for help with sequencing and Gowri Prakash for help with sequence submission. The project was supported by grants from the National Science Foundation of China (No. 31772496, 31301900) and a grant (No. Y229YX5105) from the Key Laboratory of the Zoological Systematics and Evolution of the Chinese Academy of Sciences.

ORCID

Shou-Ke Zhang https://orcid.org/0000-0002-7941-2029
Huai-Jun Xue https://orcid.org/0000-0002-7714-7532

REFERENCES

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Bagnères, A. G., & Wicker-Thomas, C. (2010). Chemical taxonomy with hydrocarbons. In G. J. Blomquist & A. G. Bagnères (Eds.), *Insect hydrocarbons* (pp. 121–162). Cambridge, UK: Biology, Biochemistry and Chemical Ecology, Cambridge University Press.

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

Beard, C. B., Hamm, D. M., & Collins, F. H. (1993). The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. *Insect Molecular Biology*, *2*, 103–124. https://doi.org/10.1111/j.1365-2583.1993.tb00131.x

Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., … Vogler, A. P. (2012). The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology*, *61*, 851–869. https://doi.org/10.1093/sysbio/sys037

Bertolazzi, P., Felici, G., & Weitschek, E. (2009). Learning to classify species with barcodes. *BMC Bioinformatics*, *10*(Suppl. 14), S7. https://doi.org/10.1186/1471-2105-10-S14-S7

Bininda-Emonds, O. R. P. (2005). TransAlign: Using amino acids to facilitate the multiple alignment of protein-coding DNA sequences. *BMC Bioinformatics*, *6*, 156.

Castro, L. R., Ruberu, K., & Dowton, M. (2006). Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: Vanhorniidae) and *Primeuchroeus* spp. (Aculeata: Chrysididae): Evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera). *Genome*, *49*, 752–766.

Coates, B. S., Sumerford, D. V., Hellmich, R. L., & Lewis, L. C. (2005). Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*. *International Journal Biological Sciences*, *1*, 13–18.

Coleman, A. W. (2003). ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends in Genetics*, *19*, 370–375. https://doi.org/10.1016/S0168-9525(03)00118-5

Dasmahapatra, K. K., Silva-Vásquez, A., Chung, J. W., & Mallet, J. (2007). Genetic analysis of a wild-caught hybrid between non-sister *Heliconius* butterfly species. *Biology Letters*, *3*, 660–663.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*, 294–299.

Funk, D. J., Filchak, K. E., & Feder, J. L. (2002). Herbivorous insects: Model systems for the comparative study of speciation ecology. *Genetica*, *116*, 251–267.

Gómez-Rodríguez, C., Crampton-Platt, A., Timmermans, M. J. T. N., Baselga, A., & Vogler, A. P. (2015). Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution*, *6*, 883–894.

Jacobsen, M. W., Hansen, M. M., Orlando, L., Bekkevold, D., Bernatchez, L., Willerslev, E., & Gilbert, M. T. P. (2012). Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and ecologically distinct European whitefish (*Coregonus* spp.). *Molecular Ecology*, *21*, 2727–2742.

Jäkel, R., Mora, D., & Dobler, S. (2013). Evidence for selective sweeps by *Wolbachia* infections: Phylogeny of *Altica* leaf beetles and their reproductive parasites. *Molecular Ecology*, *22*, 4241–4255.

Kather, R., & Martin, S. J. (2012). Cuticular hydrocarbon profiles as a taxonomic tool: Advantages, limitations and technical aspects. *Physiological Entomology*, *37*, 25–32. https://doi.org/10.1111/j.1365-3032.2011.00826.x

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious basic: An

integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kondrashov, A. S., & Shpak, M. (1998). On the origin of species by means of assortative mating. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*, 2273–2278.

Konstantinov, A. S. (1987). On the morphological structures used for identification of females in flea-beetles of the genus *Altica* (Coleoptera, Chrysomelidae, Alticinae). *Zoological Journal*, *54*, 42–50.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054

Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*, 1695–1701. https://doi.org/10.1093/molbev/mss020

Liu, K., Wang, F., Chen, W., Tu, L. H., Min, M. S., Bi, K., & Fu, J. Z. (2010). Rampant historical mitochondrial genome introgression between two species of green pond frogs, *Pelophylax nigromaculatus* and *P. plancyi*. *BMC Evolutionary Biology*, *10*, 201.

Logue, K., Chan, E. R., Phipps, T., Small, S. T., Reimer, L., Henry-Halldin, C., … Serre, D. (2013). Mitochondrial genome sequences reveal deep divergences among *Anopheles punctulatus* sibling species in Papua New Guinea. *Malaria Journal*, *12*, 64. https://doi.org/10.1186/1475-2875-12-64

Lowe, T. M., & Eddy, S. R. (1997). tRNAscan–SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, *25*, 955–964. https://doi.org/10.1093/nar/25.5.955

Magoga, G., Sahin, D. C., Fontaneto, D., & Montagna, M. (2018). Barcoding of chrysomelidae of Euro-Mediterranean area: Efficiency and problematic species. *Scientific Reports*, *8*, 13398. https://doi.org/10.1038/s41598-018-31545-9

Malausa, T., Bethenod, M., Bontemps, A., Bourguet, D., Cornuet, J., & Ponsard, S. (2005). Assortative mating in sympatric host races of the European corn borer. *Science*, *308*, 258–260. https://doi.org/10.1126/science.1107577

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, *20*, 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Marques, J. P., Sotelo, G., Larsson, T., Johannesson, K., Panova, M., & Faria, R. (2017). Comparative mitogenomic analysis of three species of periwinkles: *Littorina fabalis*, *L. obtusata* and *L. saxatilis*. *Marine Genomics*, *32*, 41–47. https://doi.org/10.1016/j.margen.2016.10.006

Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N., & Alves, P. C. (2005). Invasion from the cold past: Extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology*, *14*, 2459–2464. https://doi.org/10.1111/j.1365-294X.2005.02599.x

Montagna, M., Sassera, D., Griggio, F., Epis, S., Bandi, C., & Gissi, C. (2012). Tick-Box for end formation of mitochondrial transcripts in ixodida, basal Chelicerates and *Drosophila*. *PLoS ONE*, *7*, e47538.

Nie, R.-E., Breeschoten, T., Timmermans, M. J. T. N., Nadein, K., Xue, H.-J., Bai, M., … Vogler, A. P. (2018). The phylogeny of Galerucinae (Coleoptera: Chrysomelidae) and the performance of mitochondrial genomes in phylogenetic inference compared to nuclear rRNA genes. *Cladistics*, *34*, 113–130. https://doi.org/10.1111/cla.12196

Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, *28*, 1420–1428.

Pokorny, T., Lunau, K., Quezada-Euan, J. J. G., & Eltz, T. (2014). Cuticular hydrocarbons distinguish cryptic sibling species in *Euglossa* orchid bees. *Apidologie*, *45*, 276–283. https://doi.org/10.1007/s13592-013-0250-5

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, *25*, 1253–1256. https://doi.org/10.1093/molbev/msn083

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, *21*, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Reid, C. A. M., & Beatson, M. (2015). Disentangling a taxonomic nightmare: A revision of the Australian, Indomalayan and Pacific species of *Altica* Geoffroy, 1762 (Coleoptera: Chrysomelidae: Galerucinae). *Zootaxa*, *3918*, 503–551.

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

Ruhl, M. W., Wolf, M., & Jenkins, T. M. (2010). Compensatory base changes illuminate morphologically difficult taxonomy. *Molecular Phylogenetics and Evolution*, *54*, 664–669. https://doi.org/10.1016/j.ympev.2009.07.036

Saitou, N., & Nei, M. (1987). The neighbor joining method - a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*, 406–425.

Sheffield, N. C., Song, H., Cameron, S. L., & Whiting, M. F. (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Molecular Biology and Evolution*, *25*, 2499–2509. https://doi.org/10.1093/molbev/msn198

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

Timmermans, M. J. T. N., Barton, C., Haran, J., Ahrens, D., Culverwell, C. L., Ollikainen, A., … Vogler, A. P. (2016). Family-level sampling of mitochondrial genomes in Coleoptera: Compositional heterogeneity and phylogenetics. *Genome Biology and Evolution*, *8*, 161–175. https://doi.org/10.1093/gbe/evv241

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. https://doi.org/10.1111/j.1096-0031.2010.00329.x

Virgilio, M., Backeljau, T., Nevado, B., & De Meyer, M. (2010). Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics*, *11*, 206. https://doi.org/10.1186/1471-2105-11-206

Warchalowski, A. (2010). *The Palaearctic Chrysomelidae. identification keys*. Warszawa, Poland: Natura Optima Dux Foundation.

Weitschek, E., van Velzen, R., Felici, G., & Bertolazzi, P. (2013). BLOG 2.0: A software system for character-based species classification with DNA Barcode sequences. What it does, how to use it. *Molecular Ecology Resources*, *13*, 1043–1046. https://doi.org/10.1111/1755-0998.12073

Wielstra, B., & Arntzen, J. W. (2011). Unraveling the rapid radiation of crested newts (*Triturus cristatus* superspecies) using complete

mitogenomic sequences. *BMC Evolutionary Biology*, *11*, 162. https://doi.org/10.1186/1471-2148-11-162

Wilson, C. C., & Bernatchez, L. (1998). The ghost of hybrids past: Fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Molecular Ecology*, *7*, 127–132.

Xue, H. J., Li, W. Z., Nie, R. E., & Yang, X. K. (2011). Recent speciation in three closely related sympatric specialists: Inferences using multi-locus sequence, post-mating isolation and endosymbiont data. *PLoS ONE*, *6*, e27834. https://doi.org/10.1371/journal.pone.0027834

Xue, H. J., Li, W. Z., & Yang, X. K. (2009). Genetic analysis of feeding preference in two related species of *Altica* (Coleoptera: Chrysomelidae: Alticinae). *Ecological Entomology*, *34*, 74–80.

Xue, H. J., Li, W. Z., & Yang, X. K. (2014). Assortative mating between two sympatric closely-related specialists: Inferred from molecular phylogenetic analysis and behavioral data. *Scientific Reports*, *4*, 5436. https://doi.org/10.1038/srep05436

Xue, H. J., Magalhães, S., Li, W. Z., & Yang, X. K. (2009). Reproductive barriers between two sympatric beetle species specialized on different host plants. *Journal of Evolutionary Biology*, *22*, 2258–2266. https://doi.org/10.1111/j.1420-9101.2009.01841.x

Xue, H. J., Segraves, K. A., Wei, J., Zhang, B., Nie, R. E., Li, W. Z., & Yang, X. K. (2018). Chemically mediated sexual signals restrict hybrid speciation in a flea beetle. *Behavioral Ecology*, *29*, 1462–1471. https://doi.org/10.1093/beheco/ary105

Xue, H. J., Wei, J. N., Magalhães, S., Zhang, B., Song, K. Q., Liu, J., … Yang, X. K. (2016). Contact pheromones of 2 sympatric beetle species are modified by the host plant and affect mating. *Behavioral Ecology*, *27*, 895–902.

Yao, H., Song, J. Y., Liu, C., Luo, K., Han, J. P., Li, Y., … Chen, S. L. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS ONE*, *5*, e13102.

Yu, P. Y., Wang, S. Y., & Yang, X. K. (1996). *Ecological insect fauna of China. Fasc. 54 Coleoptera: Chrysomeloidea (II)*. Beijing, China: Science Press.

Zhai, Z. Z., Xue, H. J., Wang, S. Y., & Yang, X. K. (2007). Molecular phylogeny of the sympatric species of *Altica* (Coleoptera: Chrysomelidae: Alticinae) with references to their host plant relationship. *Acta Zootaxonomica Sinica*, *32*, 137–142.

Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X. U., Zhou, L., … Huang, Q. (2013). Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *Gigascience*, *2*, 4. https://doi.org/10.1186/2047-217X-2-4

SUPPORTING INFORMATION

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

How to cite this article: Nie R-E, Wei J, Zhang S-K, et al. Diversification of mitogenomes in three sympatric *Altica* flea beetles (Insecta, Chrysomelidae). *Zool Scr*. 2019;48:657–666. https://doi.org/10.1111/zsc.12371