

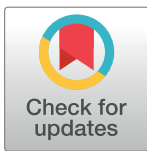
RESEARCH ARTICLE

# Cryptic diversity of the subfamily Calaphidinae (Hemiptera: Aphididae) revealed by comprehensive DNA barcoding

Yerim Lee<sup>1,2</sup>, Wonhoon Lee<sup>3</sup>, Mariusz Kanturski<sup>4</sup>, Robert G. Foottit<sup>5</sup>, Shin-Ichi Akimoto<sup>6</sup>, Seunghwan Lee<sup>1,2\*</sup>

**1** Laboratory of Insect Biosystematics, Department of Agricultural Biotechnology, Seoul National University, Seoul, Republic of Korea, **2** Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea, **3** Department of Plant Medicine and Institute of Agriculture & Life Science, Gyeongsang National University, Jinju, Korea, **4** Department of Zoology, Faculty of Biology and Environmental Protection, University of Silesia, Bankowa, Katowice, Poland, **5** Invertebrate Biodiversity, National Environmental Health Program, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada, **6** Laboratory of Systematic Entomology, Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido University, Kita-ku, Sapporo, Japan

\* [seung@snu.ac.kr](mailto:seung@snu.ac.kr)



**OPEN ACCESS**

**Citation:** Lee Y, Lee W, Kanturski M, Foottit RG, Akimoto S-I, Lee S (2017) Cryptic diversity of the subfamily Calaphidinae (Hemiptera: Aphididae) revealed by comprehensive DNA barcoding. PLoS ONE 12(4): e0176582. <https://doi.org/10.1371/journal.pone.0176582>

**Editor:** Bi-Song Yue, Sichuan University, CHINA

**Received:** January 6, 2017

**Accepted:** April 12, 2017

**Published:** April 27, 2017

**Copyright:** © 2017 Lee et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequence data are available from the Genbank (accession numbers KY306805-KY307305).

**Funding:** This project was funded by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science (NRF-2015R1D1A1A01059031).

**Competing interests:** The authors have declare that no competing interests exist.

## Abstract

Aphids are a species rich group comprising many important pests. However, species identification can be very difficult for aphids due to their morphological ambiguity. DNA barcoding has been widely adopted for rapid and reliable species identification as well as cryptic species detection. In this study, we investigated cryptic diversity in the subfamily Calaphidinae (Hemiptera: Aphididae) based on 899 sequences of *cytochrome c oxidase I (COI)* for 115 morphospecies (78 species collected in this study and sequences of 73 species downloaded from Genbank). Among these 115 morphospecies, DNA barcoding results of 90 (78.3%) species were identical to results of morphological identification. However, 25 (21.7%) morphospecies showed discrepancies between DNA barcoding and traditional taxonomy. Among these 25 discordances, a total of 15 cryptic species were identified from 12 morphospecies. We also found three morphologically distinct species pairs that sharing DNA barcoding. Based on molecular operational taxonomic unit (MOTU) estimation, we discussed on species delimitation threshold value for these taxa. Our findings confirm that Calaphidinae has high cryptic diversity even though aphids are relatively well-studied.

## Introduction

Detecting cryptic species is essential for precise species diversity estimation [1]. With the availability of DNA barcoding methods, recognition of cryptic species has been increased over the past decades [1]. To date, DNA sequences of approximately 10% of all described species (ca. 160,000) are deposited in open access databases such as Barcode of Life Data Systems (BOLD) and the National Center for Biotechnology Information (NCBI) [2–3]. More recently, some researchers have suggested that DNA-sequenced-based taxonomic description alone is not

only enough for species identification but also can be used as an alternative to classical taxonomy [4]. Indeed, DNA barcoding has been contributed to resolve morphological ambiguity in various taxa [5–9] with identification accuracy rate of 97% [10–13]. Therefore, DNA barcoding enables the detection of cryptic species and reassessment of species diversity.

Aphids are small and soft-bodied plant sap sucking insects. Over 5,000 of aphid species have been described worldwide [14], including many important pests in agriculture, forestry, and species of quarantine importance [15–19]. However, the lack of taxonomically informative characters and morphological plasticity of aphids make species recognition challenging [20]. Notably, in the complex life cycle of aphids, multiple morphs within one species can occur due to seasonal changes [21]. Numerous biotic factors such as host-plant relationships [22–23], natural enemies [24], ant attendance [25], maternal effects [26], endosymbionts [27], and infectious microorganisms [28] can affect intraspecific plasticity of aphids. Various abiotic factors, such as climate, temperature and photoperiod [29] also have effects on their intraspecific plasticity. Conversely, extremely similar or even morphologically indistinguishable species can occur among aphids [30–32]. The utility of DNA barcoding for aphid species identification has been demonstrated at family level [7, 13, 30]. It is also useful for some subfamilies such as Eriosomatinae [33], Greenideinae [34], and Lachninae [35]. According to these studies, DNA barcoding can provide rapid and reliable identification results for aphids. However, no assessment has been attempted for the subfamily Calaphidinae.

The subfamily Calaphidinae (Hemiptera: Aphididae) is the second largest subfamily in family Aphididae. About 398 valid species belonging to 59 genera have been described in the world [14, 36–37]. Most calaphidine aphids feed on woody angiosperms belonging to 16 plant families such as Betulaceae, Fagaceae, Juglandaceae, Lythraceae, Myricaceae, and Ulmaceae although some species feed on herbaceous plants belonging to Fabaceae and Poaceae [37–38]. Many aphid species belonging to this subfamily are economically important pests, causing injury and transmitting viral diseases to cultivated plants such as leguminous crops, fruit, and landscape trees [39–40]. For example, *Therioaphis trifolii* (Monell, 1882), *Melanocallis caryefoliae* (Davis, 1910) and *Monellia caryella* (Fitch, 1855) are notorious aphid pest have caused large agricultural economic losses [41–43]. Some aphid pest, such as *Sarucallis kahawaluokalani* (Kirkaldy, 1907), *Shivaphis celti* Das, 1918 and *Tinocallis* spp. have been dispersed from their geographic origins to different continents [18, 44–46]. However, assessing calaphidine species can be difficult and time consuming since their considerable morphological variation based on seasonal changes and various biotic factors [25, 37]. The application of DNA barcoding would assist in rapid and accurate identification of species in this subfamily. It can also aid the detection of cryptic diversity.

In this study, we provided the first comprehensive assessment of DNA barcodes for the subfamily Calaphidinae. A total of 501 *Cytochrome oxidase I* (*COI*) sequences of 78 morphospecies collected in Korea and other countries from 2001 to 2015 were analyzed. The objectives of this study were i) to clarify delimiting species boundaries in morphologically ambiguous taxa, ii) to test the effectiveness of DNA barcoding in this taxa, and ultimately iii) to detect hidden species diversity.

## Materials and methods

### Ethics statement

No permission was required for sampling at the sites studied. No endangered or protected species are included in this study.

## Taxon sampling

A total of 501 aphid individuals of 78 species were collected in Asia: Korea (382 specimens of 52 species), China (20 specimens of 8 species), Japan (23 specimens of 8 species) and Laos (2 specimens of 1 species); Europe: Czech Republic (15 specimens of 7 species), Poland (16 specimens of 7 species) and UK (5 specimens of 1 species); North America: USA (29 specimens of 14 species) and Oceania: New Zealand (9 specimens of 4 species) from 2001 to 2015 (S1 Table). Each specimen was preserved in 95–99% ethanol at -20°C for genomic DNA extraction.

## Species identification

501 individuals were mounted in Canada balsam following the method of Blackman & Eastop [16] and Martin [47]. Measurements for each specimen were taken from digital images by using image analysis software (Active measure ver. 3.0.3 from Mitani Co. Ltd, Japan). Digital images were taken by a digital camera attached to a microscope (Leica 400B, Leica Microsystems, Germany). All slide specimens were deposited in the College of Agriculture and Life sciences, Seoul National University (CAL SNU), the Republic of Korea.

## DNA extraction and DNA barcoding

Genomic DNA was extracted from each sample selected from each colony by using the DNeasy Blood & Tissue kit (Qiagen, Dusseldorf, Germany) according to the modified manufacturer's protocols. To confirm morphological features, we used a nondestructive method: each whole-bodied specimen was put into a mixture of 90 µl of ATL buffer and 10 µl of proteinase K incubated without pulverization. After 24 h incubation, 90 µl AL buffer was added and incubated for another 10 min. The solution was gently pipetted into a mini spin column leaving the cuticle of the specimen which was slide mounted.

A 658 bp of *COI* gene region, generally called as 'barcoding region' was amplified using a universal primer set: LCO1490 5' -GGTCAACAAATCATAAAGATATTGG-3' and HCO2198 5' -TAAACTTCAGGGTGACCAAAAAATCA-3' [48]. Polymerase chain reaction (PCR) was conducted with AccuPower PCR PreMix (Bioneer, Daejeon, Korea) in 20 ml reaction mixtures under the following conditions: initial denaturation at 94°C for 3 min; followed by 35 cycles at 94°C for 30s, an annealing temperature of 45.2°C for 30s, an extension at 72°C for 1min; and the final extension at 72°C for 5min. All PCR products were assessed 1.5% agarose gel electrophoresis. Successfully amplified samples were purified using a QIAquick PCR purification kit (Qiagen, Inc.), and then sequenced directly using an automated sequencer (ABI PrismH 3730 XL DNA Analyzer) at MacroGen Inc. (Seoul, Korea).

## Molecular analyses

All sequences to be analyzed were initially assembled and examined using Seqman pro ver. 7.1.0 (DNA star, Inc., Madison, Wisconsin, USA). Poor quality sequences with ambiguous peaks were removed. We used the molecular identification criteria of putative orthologues and paralogues according to Moulton et al. [49] and Fontaneto et al. [50], to prevent misleading by nuclear mitochondrial pseudogenes (Numts) and heteroplasmy. A total of 501 *COI* sequences of 78 species including previously unknown sequences of 42 species were newly generated for the molecular analyses. Additionally, 398 *COI* sequences of 73 species were downloaded from Genbank using keyword '*COI*' and 'Calaphidinae' (S2 Table). As a result, the final dataset consisted of 899 sequences of 115 species (S1 and S2 Tables).

These data was aligned using online utility MAFFT ver. 7 alignment package [51] and MEGA 6 [52]. In this step, we removed uncertain anterior and posterior regions were removed. Finally,  $\geq 546$  bp was used for analyses. For the aligned dataset, a neighbor-joining analysis was conducted using MEGA 6 based on Kimura-2-Parameter (K2P) model [53], the best for species level analysis, particularly for those with low distances [54]. Intra- and inter-specific distances in different taxonomic levels were calculated using pairwise distance method based on the K2P model [53] using MEGA 6.

To infer species delimitation criteria based on a partial *COI* gene in this subfamily, we performed molecular operational taxonomic units (MOTUs) estimation by using two effective tools to delimit molecular species. First, Automatic Barcode Gap Discovery (ABGD) analysis was conducted to automatically delimit sequences into hypothetical molecular species [55] (<http://wwwabi.snv.jussieu.fr/public/abgd>) by contrasting inter- and intra-specific distances. Standard settings were used with two values of relative gap width ( $X = 1$  and  $X = 1.5$ ) based on Kimura K80 model. Additionally, Bayesian Poisson Tree Processes (bPTP) analysis as implemented on the Exelixis Lab web-server (<http://species.h-its.org/ptp/>) was performed. This method delimits species based on the phylogenetic species concept [56]. Compared to the generalized mixed Yule coalescent (GMYC) model, bPTP model is a more robust and simpler method [57]. The required rooted phylogenetic input tree was drawn using RAxML [58] with GTR+G+I substitution model.

## Results

### Genetic variation of morphospecies

A total of 501 *COI* sequences ( $\geq 546$  bp) from 78 morphospecies belonging to 36 genera of four subtribes, Calaphidina, Monaphidina, Myzocallidina and Panaphidina were newly generated in this study (S1 Table). All sequences are deposited in Genbank (KY306805-KY307305). Results of genetic divergence at different taxonomic levels are summarized in Table 1.

The overall mean distance was 13.2% for the final dataset of 899 sequences of 115 species. The mean interspecific distance ranged from 4.19% to 13.3% at genus level. The mean genetic distance between genera ranged from 11.92% to 14.51% at subtribe level.

Intraspecific genetic distance was calculated for the 100 of 115 morphospecies. For the remaining 15 species, intraspecific distances could not be calculated because there was only one individual representing each species. The mean intraspecific distance ranged from 0 to 6.0% in each species. Among these 100 morphospecies, 78 species showed very low to moderate genetic divergences (below 1.5%, Fig 1). Another 4 species showed ambiguous intraspecific distances ranging from 1.7% to 1.9% (Fig 1). However, the remaining 18 species showed relatively high intraspecific distances ranging from 2.9% to 16.6% (Fig 1, Table 2). Such a high level of intraspecific distance indicate that there might be potential cryptic species and/or misidentified sequences in these of 18 species (Table 2).

**Table 1. Genetic divergences in different taxonomic level within Calaphidinae.**

Comparison within	Mean (%)	Minimum (%)	Maximum (%)
Species	0.7	0	16.6
Genus	21.1	0	9.5
Subtribe	13.4	3.4	22.8
Tribe	13.8	3.4	23.1

<https://doi.org/10.1371/journal.pone.0176582.t001>

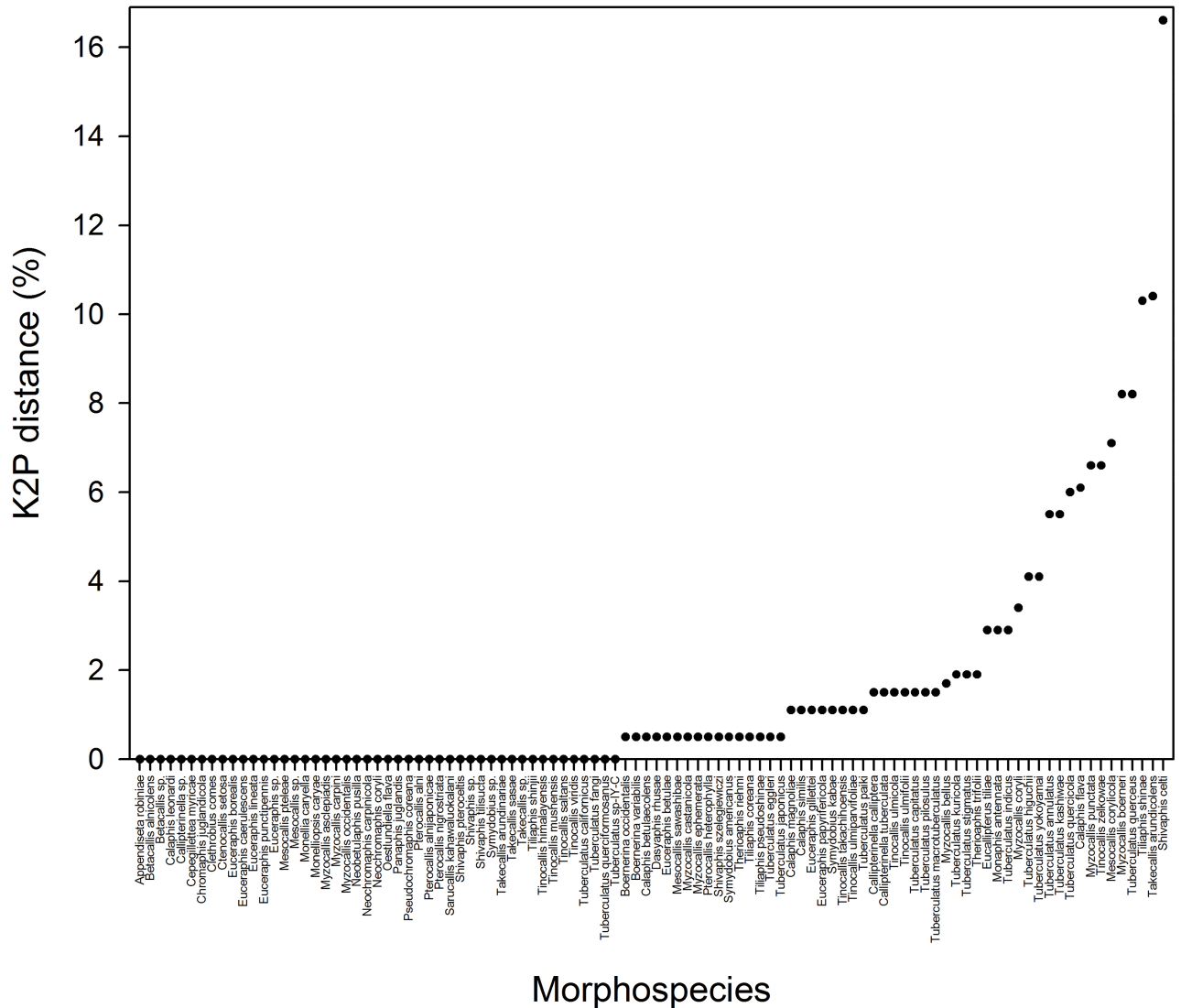


Fig 1. Maximum intraspecific distances (%) based on Kimura-2-parameter (K2P) for 100 morphospecies.

<https://doi.org/10.1371/journal.pone.0176582.g001>

### MOTUs estimation

The number of MOTUs determined by ABGD differed slightly depending on the value of the relative gap width (Fig 2). When *p* value was set at 0.0129, the number of MOTUs was 133 at relative gap width *X* = 1 or *X* = 1.5 (Fig 2). Based on ABGD result, 16 species were divided into 2–4 inner-groups in each species group (Fig 3). Three morphologically distinct species pairs: i) *Pterocallis ahnijaponicae* and *P. nigrostriata*, ii) *Tiliaphis pseudoshinae* and *T. shinae* and iii) *Tuberculatus (Orienttuberculoides) capitatus* and *T. (O.) fangi* were clustered together as a single MOTU, respectively (Table 3).

The bPTP model recognized a total of 136 MOTUs except for outgroup based on the maximum likelihood method. Among 115 morphospecies, 19 species were subdivided into 2–4 inner-groups in each species group (Fig 3). This result was similar not only based on the number of MOTUs, but also based on MOTU compositions obtained from ABGD (Fig 3). Overall, the bPTP model tended to be more sensitive to MOTU delimitation. ABGD and bPTP resulted

**Table 2. 19 Cases of having high intraspecific distances.**

Subtribe	Species	No. of subgroups	Max. Intraspecific distance
Calaphidina	<i>Calaphis flava</i>	3	6.1%
Monaphidina	<i>Monaphis antennata</i>	2	2.9%
Myzocallidina	<i>Myzocallis boernerii</i>	2	8.2%
	<i>Myzocallis coryli</i>	2	3.4%
	<i>Tuberculatus annulatus</i>	2	5.5%
	<i>Tuberculatus higuchii</i>	4	4.1%
	<i>Tuberculatus indicus</i>	2	2.9%
	<i>Tuberculatus kashiwae</i>	2	5.5%
	<i>Tuberculatus punctata</i>	2	6.6%
	<i>Tuberculatus querceus</i>	2	8.2%
	<i>Tuberculatus quercicola</i>	2	6.0%
	<i>Tuberculatus yokoyamai</i>	2	4.1%
	Panaphidina	<i>Eucallipterus tiliae</i>	2
<i>Mesocallis corylicola</i>		2	7.1%
<i>Shivaphis celti</i>		2	16.6%
<i>Takecallis arundicolens</i>		3	10.4%
<i>Therioaphis trifolii</i>		2	1.9%
<i>Tiliaphis shinae</i>		2	10.3%
<i>Tinocallis zerkowae</i>		2	6.6%

<https://doi.org/10.1371/journal.pone.0176582.t002>

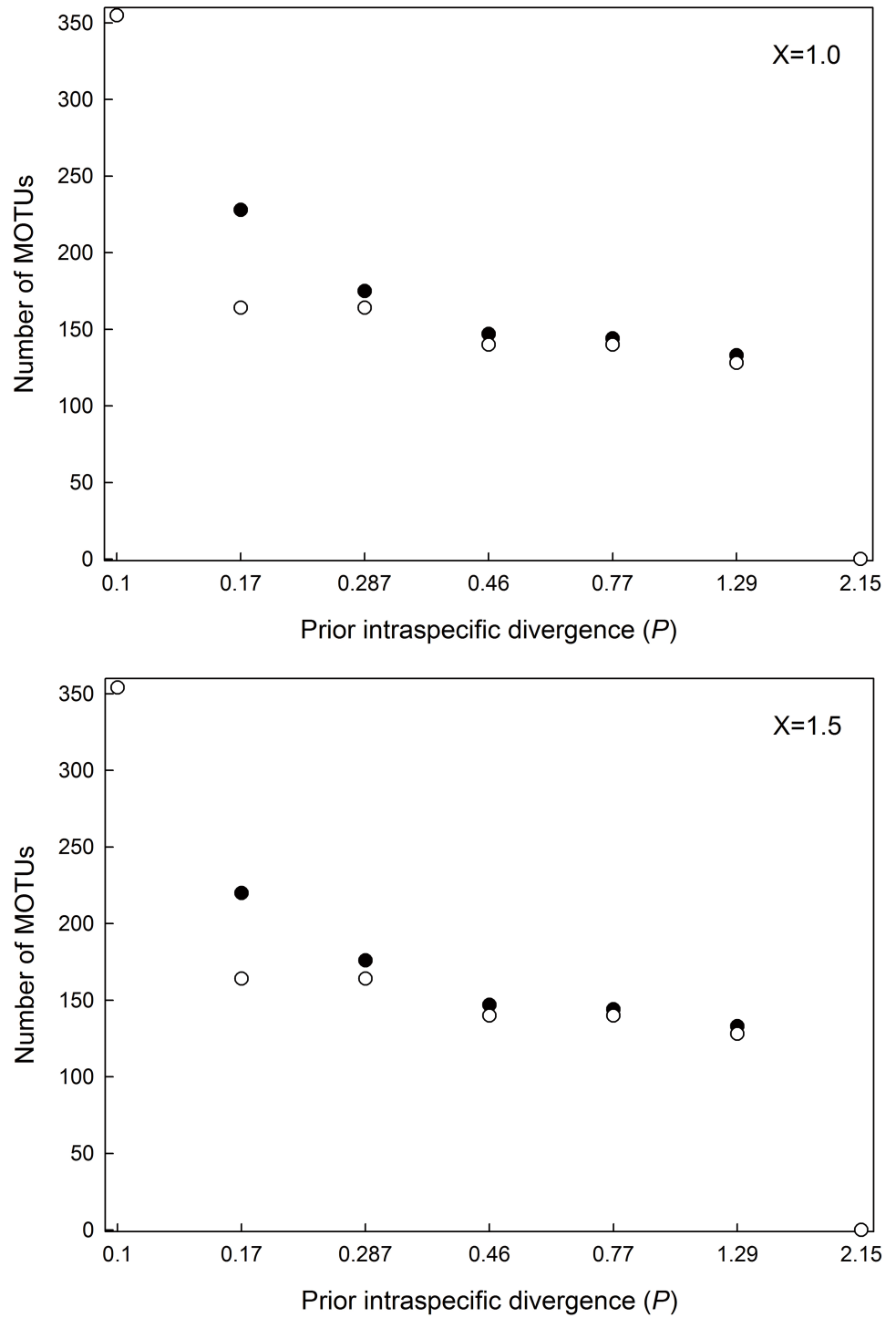
in different estimates for 8 morphospecies (Fig 3). For example, 4 morphospecies, *Takecallis arundinariae*, *Tuberculatus* (*Nippocallis*) *kuricola*, *Tuberculatus* (*Arakawana*) *stigmatus* and *Myzocallis* (*Lineomyzocallis*) *bellus*, were subdivided into two groups each in the bPTP results. However, no subdivision was detected for these 4 species based on ABGD results (Fig 3). On the contrary, *Tiliaphis shinae* and *T. pseudoshinae* pair was clustered together as a single MOTU in the bPTP model (Fig 3).

### Neighbor joining analysis: Case of species delimitation

The Neighbor joining tree (NJ tree) was derived for these 899 COI sequences of the 115 species (S1 Fig). For majority of cases, sequence clusters in the NJ tree showed high congruence with morphological identification results. Practically, 90 species (78.3%) of all species could be clearly identified by COI sequence. However, for the remaining 25 species (21.7%), discrepancies between morphology and barcode based identification were detected. DNA barcoding analyses revealed the following: i) 15 cryptic species from 12 morphospecies, ii) six possible cryptic or potential misidentified sequences in the Genbank, iii) three morphologically distinct species pairs that sharing a single MOTU, and iv) four species with ambiguous inter- and intra-specific distances. Detailed results for these four cases are described as follows.

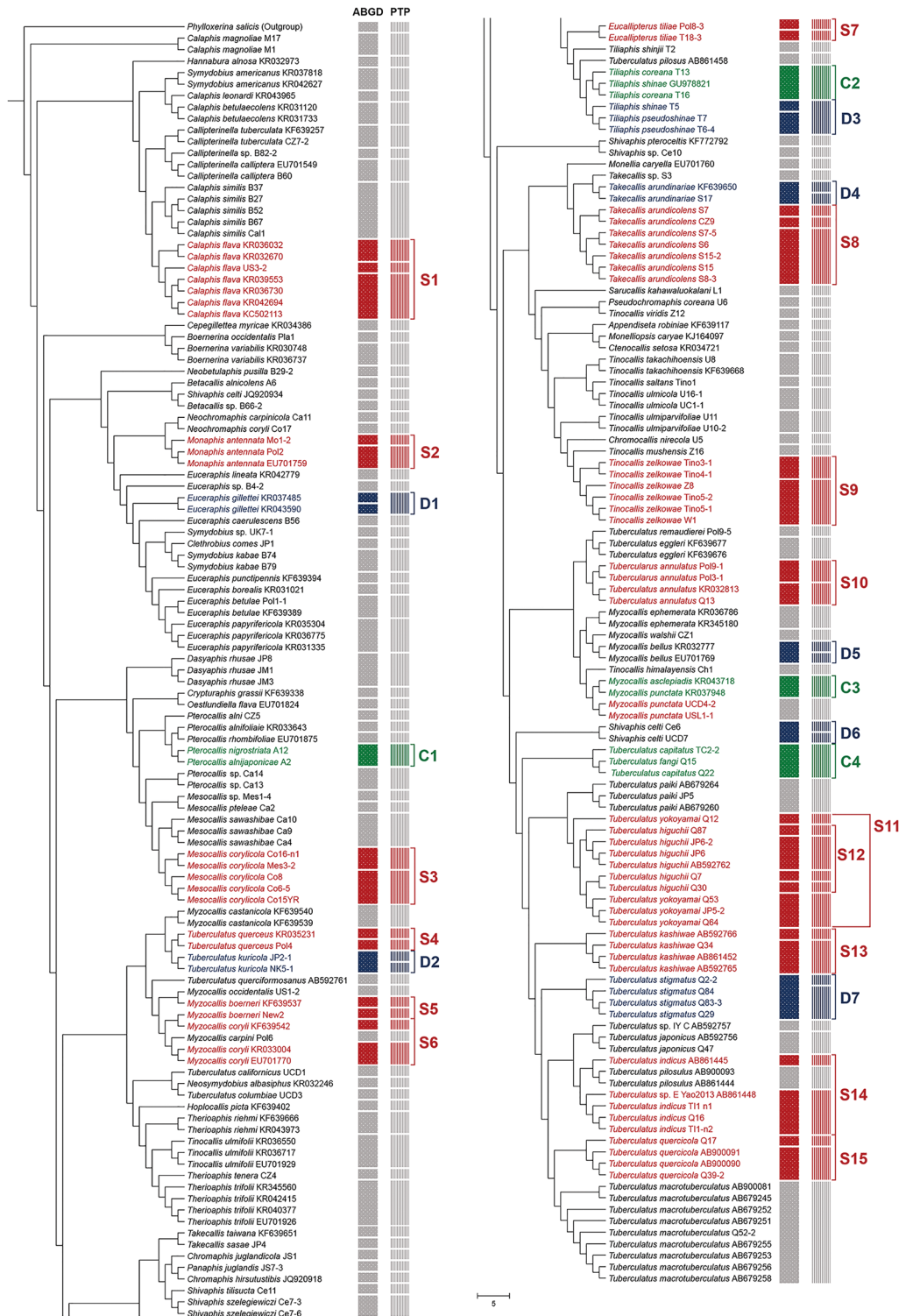
**Case I: Discovering cryptic species.** A total of 15 cryptic species from 12 morphospecies were found based on original description and other information of each morphospecies such as host-plant association and distributional information. Because in any case, it is impossible to identify DNA barcode for type materials. For each case, comparison between original and cryptic species was discussed.

A total of 66 individuals of 5 species belonging to genus *Calaphis* were analyzed in this study. DNA barcoding detected 2 cryptic species in *Calaphis flava* (Fig 4). Between group 1 and group 2, intergroup divergence was 2.9–4.0% (Fig 4). Group 3 consisted of 7 Canadian



**Fig 2. Number of MOTUs by the prior intraspecific divergence using ABGD with two values of relative gap width. (A)  $X=1$ . (B)  $X=1.5$ .**

<https://doi.org/10.1371/journal.pone.0176582.g002>



**Fig 3. Maximum likelihood COI gene tree with delimited MOTUs by ABGD and PTP analyses. (Red S) Subdivided morphospecies. (Green C) combined morphologically different species. (Blue D) discordant between ABGD and PTP results.**

<https://doi.org/10.1371/journal.pone.0176582.g003>



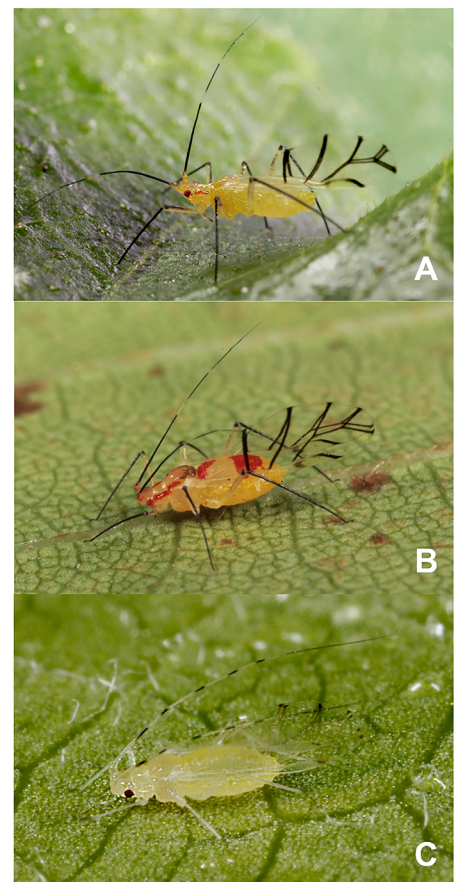
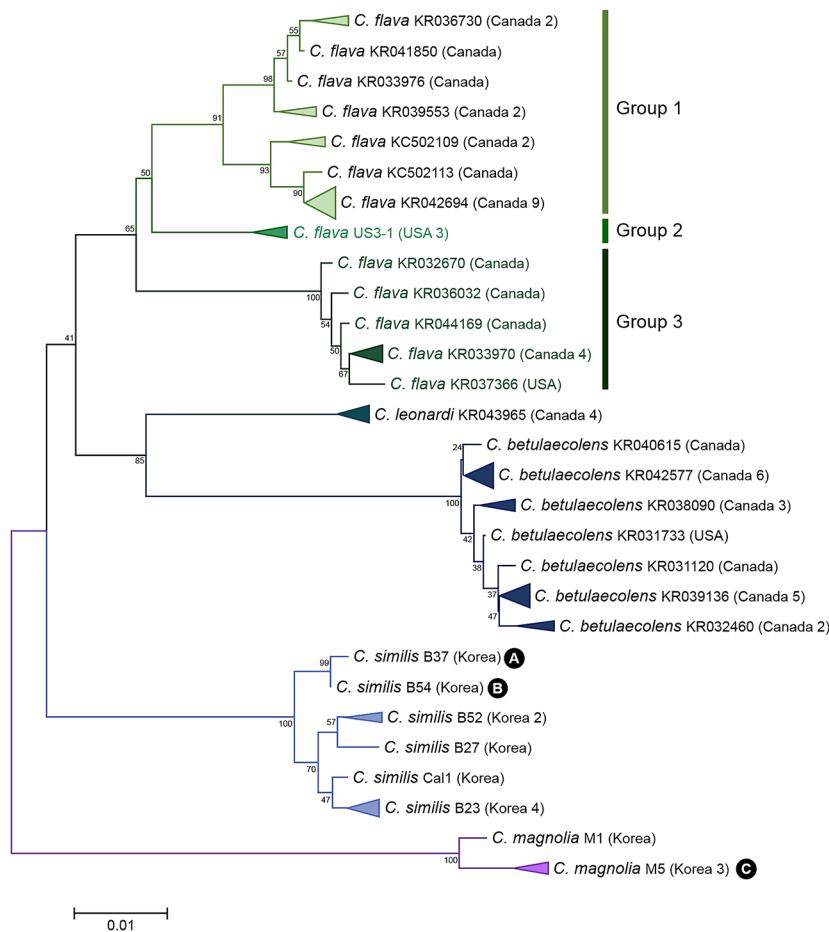
**Table 3. Case of sharing low genetic distances between morphologically distinct species pairs.**

Subtribe	Species 1	Species 2	Genetic distance
Myzocallidina	<i>Tuberculatus capitatus</i>	<i>Tuberculatus fangi</i>	0.5–1.3%
Panaphidina	<i>Pterocallis alnijaponicae</i>	<i>Pterocallis nigrostriata</i>	0.5%
	<i>Tiliaphis pseudoshinae</i>	<i>Tiliaphis shinae</i>	1.7–1.9%

<https://doi.org/10.1371/journal.pone.0176582.t003>

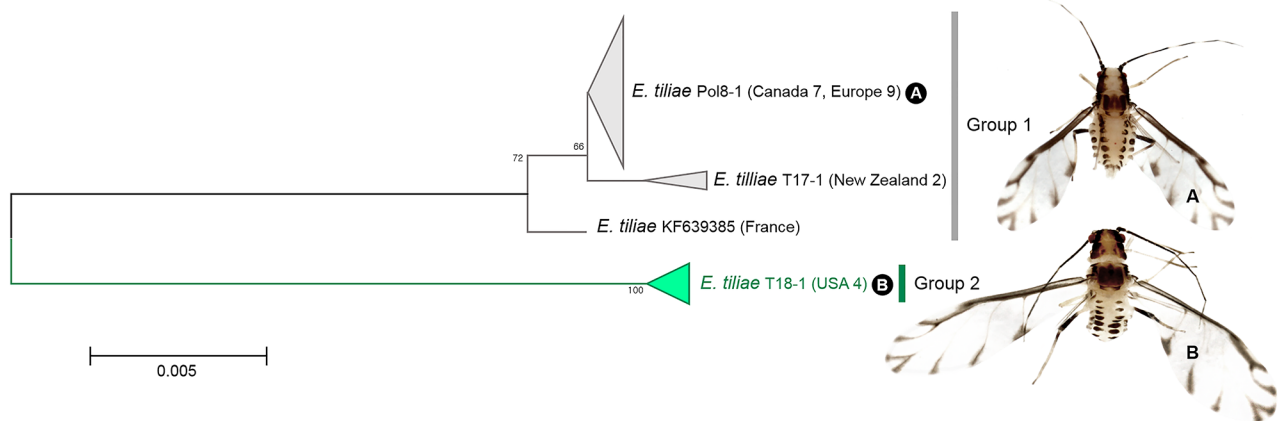
and 1 American individual was distinct from group 1 and group 2 with 4.5–6.1% of genetic divergence (Fig 4). This species was originally distributed throughout Europe and East Asia. Now it is regarded as a widespread species found in South Africa, Australia and North America. Thus, further studies are needed to compare European individuals of *C. flava* in the future study to investigate whether this species is a real cosmopolitan species or a species complex.

A total of 23 individuals of *Eucallipterus tiliae* were collected from Europe: Czech Republic, France, Italy and Poland, North America: Canada and USA, and Oceania: New Zealand. ABGD, bPTP (Fig 3) and NJ tree revealed 1 cryptic species (group 2) of *E. tiliae* with about 2.9% of intergroup genetic divergence (Fig 5). Morphologically, individuals in group 1 and 2 are very similar. However, cryptic species (group 2) is distinguishable from typical *E. tiliae* (group 1) by having longer length of 3rd–5th antennal segment and 2nd tarsal segment



**Fig 4. Neighbor-joining tree of *COI* partial gene sequences of *Calaphis* spp. (66 sequences of 5 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g004>



**Fig 5. Neighbor-joining tree of *COI* partial gene sequences of *Eucallipterus* spp. (23 sequences).**

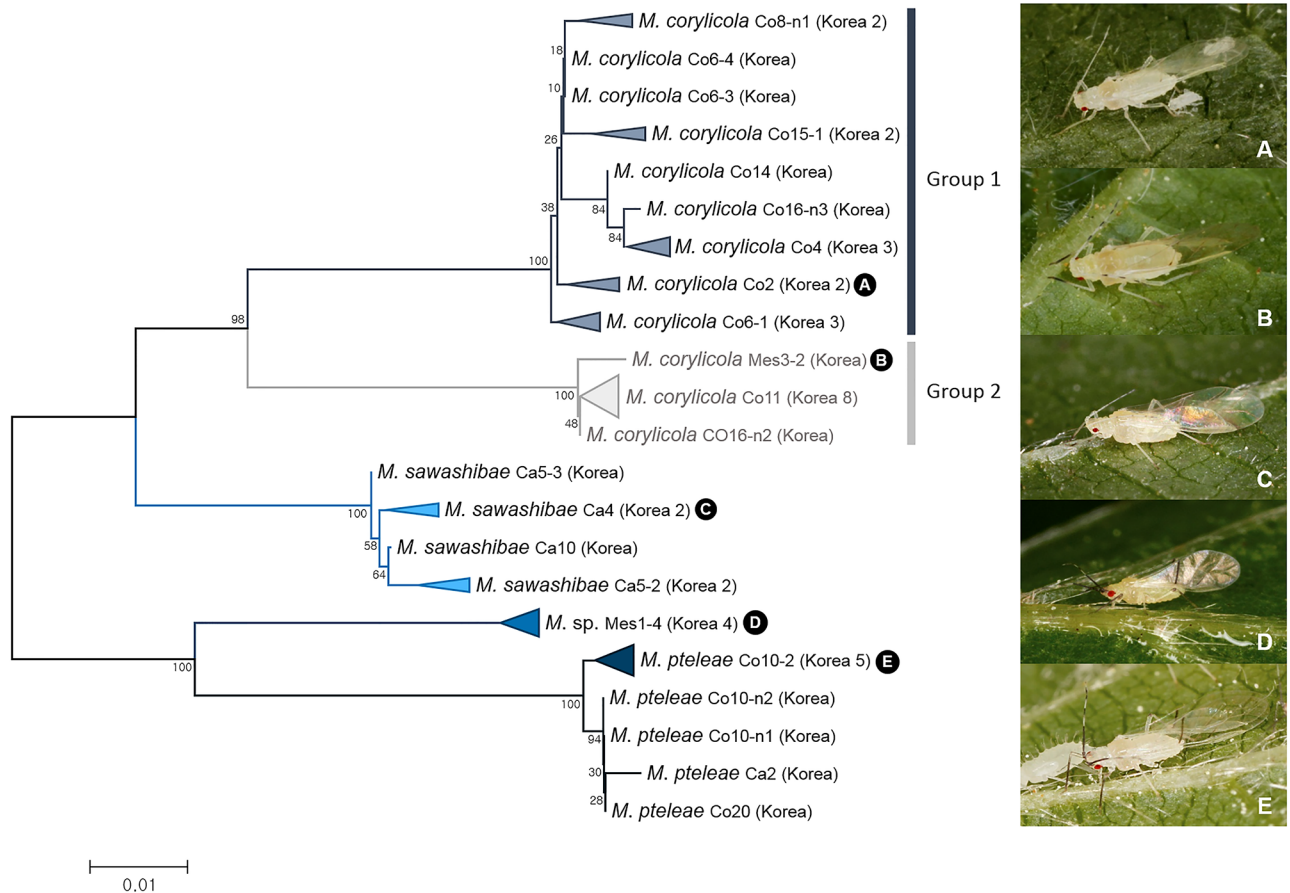
<https://doi.org/10.1371/journal.pone.0176582.g005>

(S2 Fig). In genus *Eucallipterus*, only two species have been described in the world. *E. tiliae* is a common species widely distributed throughout Europe across central Asia, and South-Africa. It has also been introduced into North America and New Zealand [37]. Our results suggest that this species might be a species complex rather than a cosmopolitan species.

ABGD and bPTP analyses on 45 individuals of four Korean species of *Mesocallis* disclosed one cryptic species in *M. corylicola* (Figs 3 and 6). *M. corylicola* (group 1) and cryptic species (group 2) showed 5.5–7.1% of genetic distance (Fig 6). In contrast to such a high level of genetic distance, the cryptic species and *M. corylicola* are superficially similar in morphology. Compared to *M. corylicola*, the cryptic species has a shorter siphunculi and longer length of ultimate rostral segment (S3 Fig). Host-plant preference appeared to differ between the two species. Most *M. corylicola* was collected on *Corylus sieboldiana* while most cryptic species were collected on *C. heterophylla* (S1 Table). According to the original description, original species was collected on *Corylus sieboldiana* [59]. Thus, the two groups have different *COI* sequences, host-plant preference, and morphology.

Eight individuals of *Monaphis antennata* collected from Canada, Korea, and Poland were analyzed with ABGD, bPTP (Fig 3), and NJ tree (Fig 7). Korean individuals formed a group (group 1) distinct from Polish and Canadian individuals with 2.4–2.9% intergroup genetic divergence (Fig 7). *Monaphis* is a monotypic genus originally described from Europe. This species lives solitarily on *Betula* spp. In Korea, *Monaphis* is extremely rare. It has only been collected on *Betula schmidtii*. In Europe and Japan, it has been collected on *B. pendula* [60–61], *B. maximowicziana* and *B. platyphylla* var. *japonica* [59]. Although we could not perform morphological comparisons on subgroups, the Korean cluster seems to be a distinct species based on the molecular divergence level.

We analyzed 58 individuals of five *Takecallis* spp. collected from East Asia: China, Japan, and Korea, Europe: Czech Republic, France, and Italy, and North America: Canada and USA. According to both ABGD and bPTP analyses, *T. arundicolens* were separated into three groups (Figs 3 and 8). Genetic divergence between group 1 and group 2+3 ranged from 7.6% to 10.4%. Genetic divergence between group 2 and group 3 was 2.5%. In each group, morphological differences were only detected between alate nymphs. Alate nymphs in group 1 could be distinguished by a dark colored cauda with short dorsal abdominal setae (S4 Fig). Group 2 and group 3 shared similar morphology. However, group 3 could be distinguished from group 2 by shorter siphunculi with long filiform setae on the body (S6 Fig 4). *T. arundicolens* is one of common bamboo feeding species. Originally, this species was described from East Asia. It

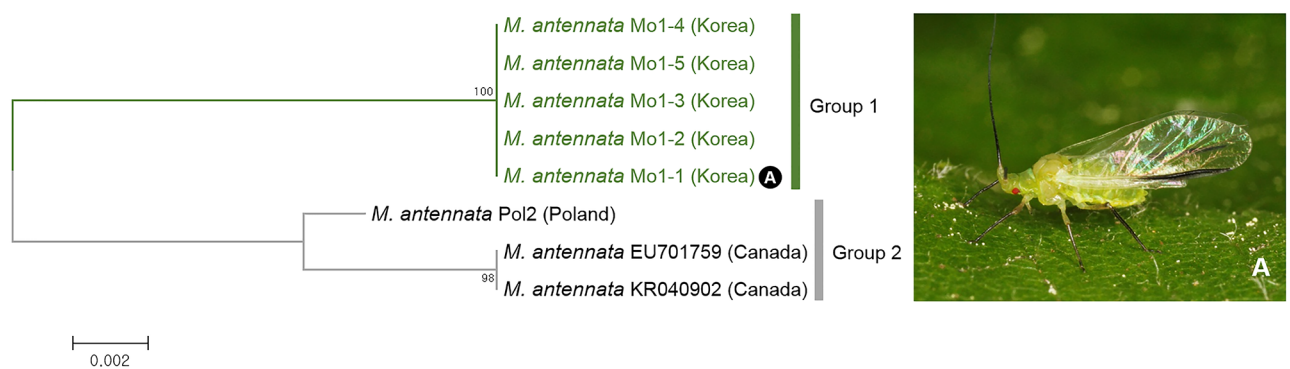


**Fig 6. Neighbor-joining tree of *COI* partial gene sequences of *Mesocallis* spp. (45 sequences of 4 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g006>

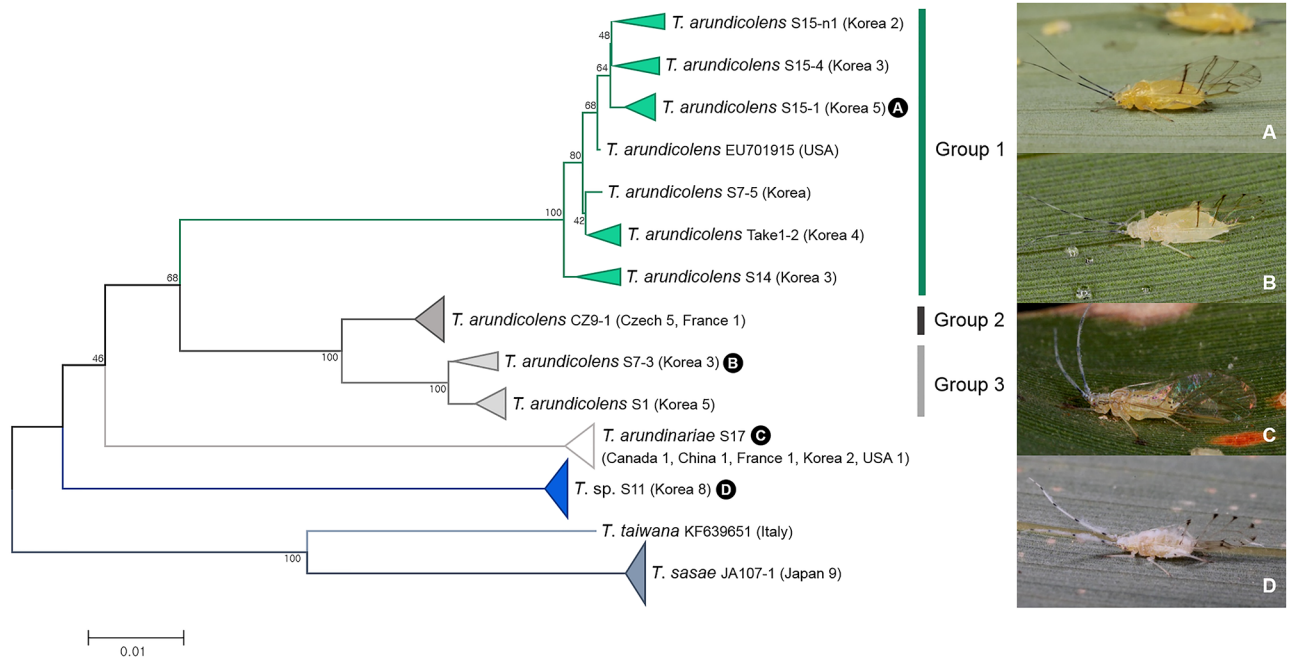
has been introduced into Europe and North America. Our results indicate that European populations (group 2) might be species distinct from Asian species (group 1 and group 3).

A total of 48 individuals of nine *Tinocallis* spp. collected from East Asia: Korea and North America: Canada and USA were analyzed. Both ABGD and bPTP analyses recognized two subgroups (group 1 and group 2) among 22 individuals of *T. zelkowae* collected from Korea (Fig 3). These groups showed genetic divergence of 5.5–6.5% (Fig 9). Morphologically,



**Fig 7. Neighbor-joining tree of *COI* partial gene sequences of *Monaphis antennata* (8 sequences).**

<https://doi.org/10.1371/journal.pone.0176582.g007>



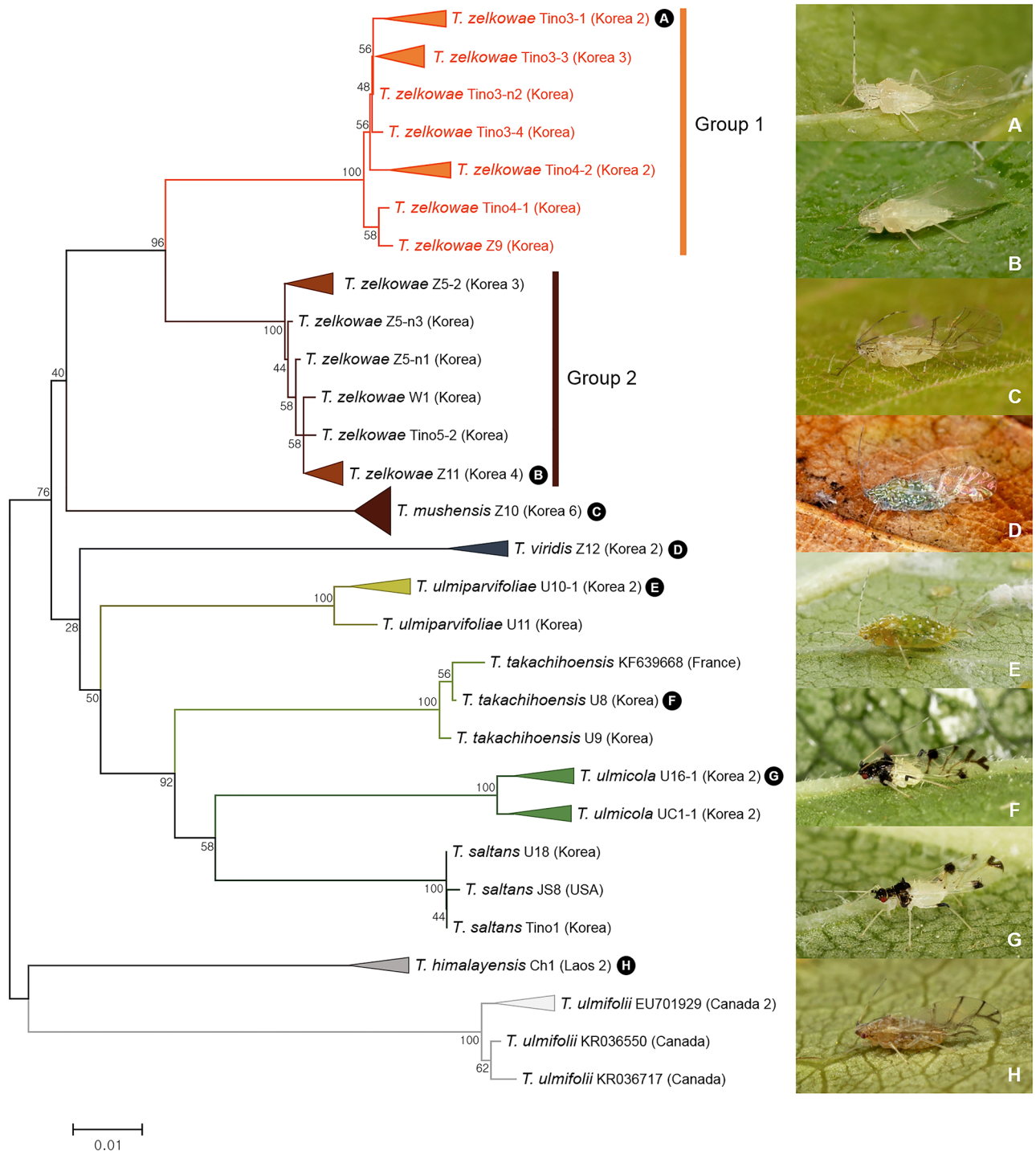
**Fig 8. Neighbor-joining tree of COI partial gene sequences of *Takecallis* spp. (58 sequences of 5 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g008>

individuals of group 1 have longer ultimate rostral segments than those of group 2 with pigmented dorsal abdominal elevations (S5 Fig). Group 1 is relatively rare. It was only collected on native *Zelkova serrata* (var. *latifolia*) growing in Korean mountain areas. *T. zelkowae* of the group 2 is one of the most common species dwelling on *Zelkova* trees in urban area (mostly the Japanese species *Zelkova serrata* (var. *japonica*)). Results of this study indicate that group 1 might be species separated from group 2 with different host-plant association.

In genus *Tuberculatus*, a total of seven cryptic species were discovered. Detailed results for each subgenus are provided as follows. A total of 69 individuals of two species belonging to the subgenus *Acanthocallis* collected from Japan and Korea were analyzed. Of 21 Korean individuals of *Tuberculatus* (*Acanthocallis*) *quercicola*, 16 individuals formed group 3, distinct from Japanese and remaining 5 Korean individuals of *T. (A.) quercicola* (group 2) with about 5.2% of intergroup genetic divergence (Figs 3 and 10). Individuals of group 3 were different from group 2 only by having shorter setae on 3rd antennal segment with more setae on 4th–5th antennal segments (S6 Fig). Watanabe et al. [62] have reported that *Acanthocallis* species tend to have high host specificity. It has been shown that Japanese *T. (A.) quercicola* and *T. (A.) macrotuberculatus* have distinct host-plant associations with *Quercus mongolica* spp. *crispula* and *Q. dentate*, respectively [62]. Likewise, three Korean species: *T. (A.) quercicola* in group 2, *T. (A.) macrotuberculatus* in group 1, and cryptic species in group 3 showed distinct host-plant association with *Q. mongolica*, *Q. dentate*, and *Q. aliena*, respectively. Therefore, different host associations between species can be used for species identification in this group.

A total of 26 individuals of four species belonging to subgenus *Acanthotuberculatus* were analyzed. As shown in Figs 3 and 11, 14 specimens of *Tuberculatus* (*Acanthotuberculatus*) *indicus* were split into two subgroups with intergroup genetic divergence ranging from 2.4% to 2.9%. Nine specimens collected from Korea were identical to undescribed species *Tuberculatus* sp. E (Yao, unpublished, COI sequence Genbank accession no. AB861448). *Tuberculatus* (*A.*)

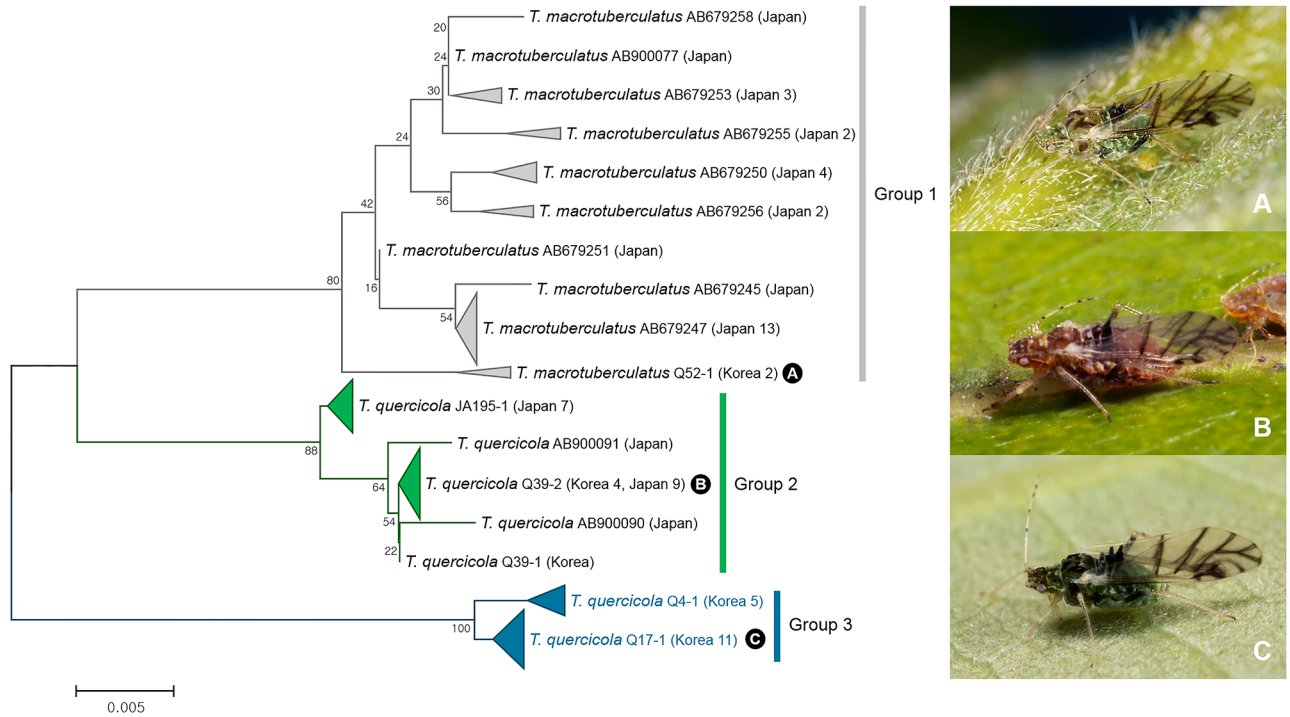


**Fig 9. Neighbor-joining tree of COI partial gene sequences of *Tinocallis* spp. (48 sequences of 9 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g009>

*japonicus* and undescribed species, *Tuberculatus* sp. IY-C [63], showed inter-specific distance of 4.5% to 5.1%. No morphological comparisons was undertaken.

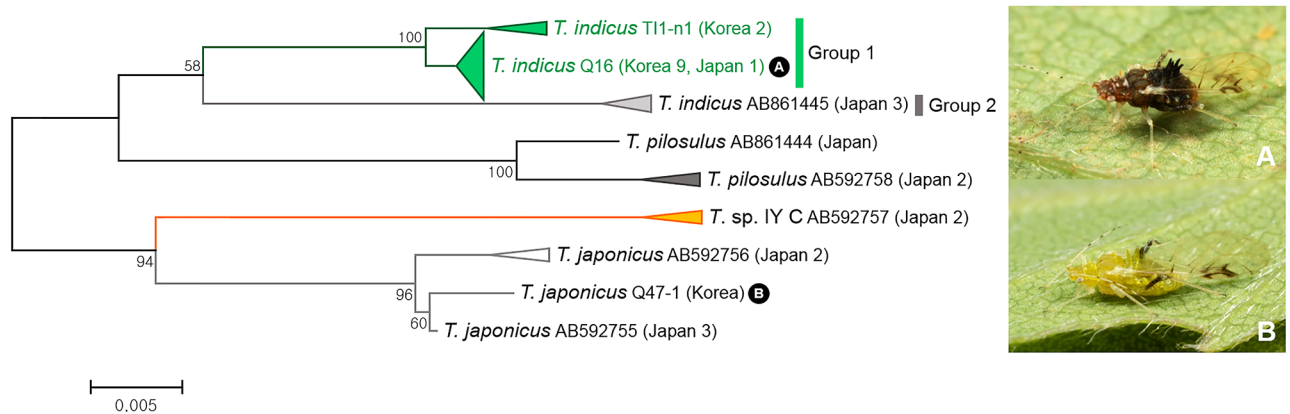
DNA barcoding uncovered five cryptic species among 86 specimens of seven species in subgenus *Orienttuberculoides* (Fig 3). *Tuberculatus* (*Orienttuberculoides*) *higuchii* was subdivided



**Fig 10. Neighbor-joining tree of *COI* partial gene sequences of subgenus *Acanthocallis* spp. (69 sequences of 2 morphospecies).**

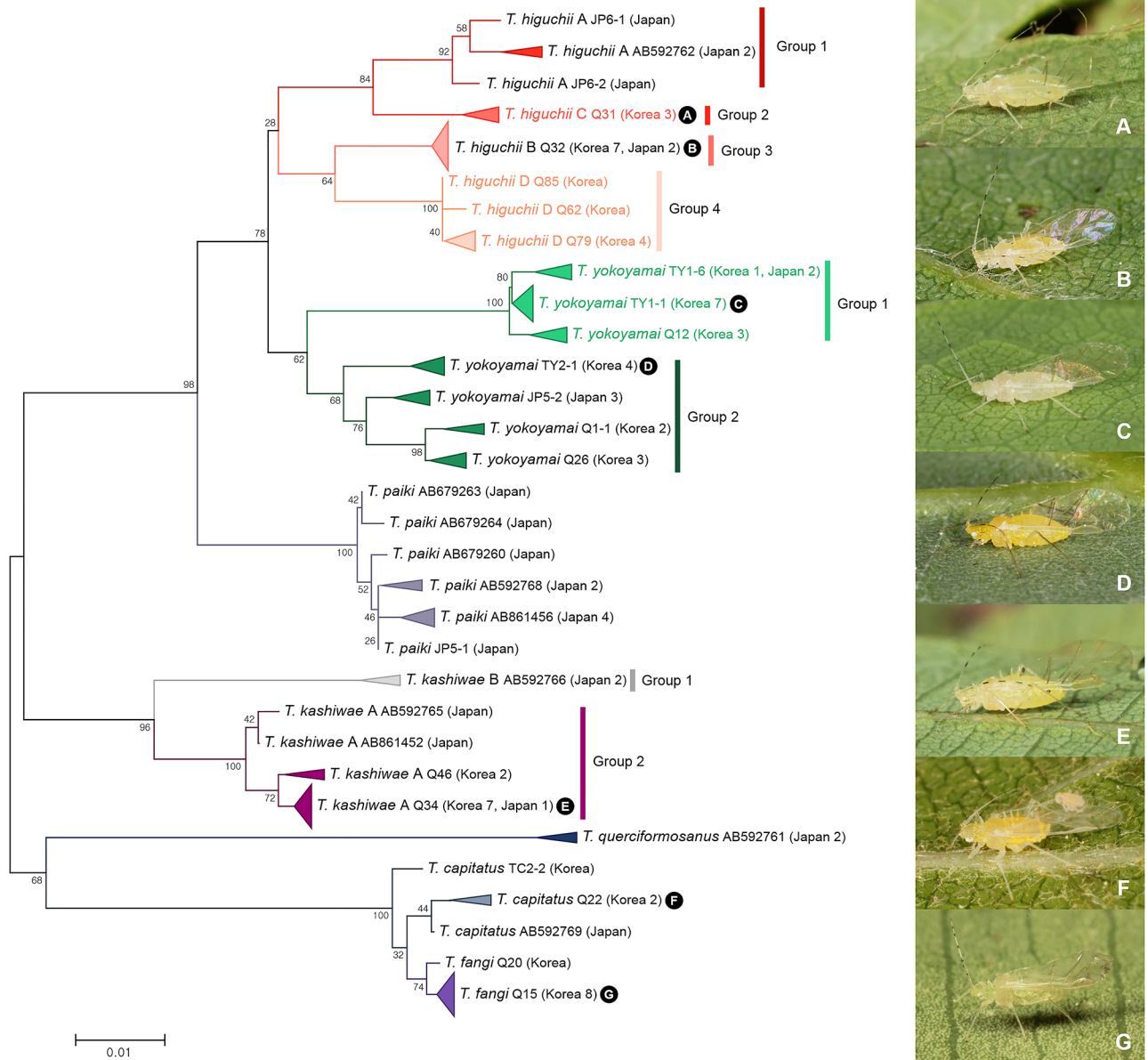
<https://doi.org/10.1371/journal.pone.0176582.g010>

into four subgroups, including previously detected species *T. (O.) higuchii* A (group 1) and *T. (O.) higuchii* B (group 3) [63] and two newly detected subgroups (Fig 12). For convenience, these newly detected subgroups were named as *T. (O.) higuchii* C (group 2) and *T. (O.) higuchii* D (group 4). Among those four subgroups, intergroup genetic divergence of 2.8% to 4.1% was observed (Fig 12). *T. (O.) higuchii* A has shorter 2nd–4th antennal segments in comparison with other subgroups (S7 Fig). *T. (O.) higuchii* B is distinct from other subgroups by having shorter siphunculi (S7 Fig). *T. (O.) higuchii* C has slightly longer second tarsal segments and cauda (S7 Fig). *T. (O.) higuchii* D is distinguished from others by having more secondary sensoria on the 3rd antennal segment (S7 Fig).



**Fig 11. Neighbor-joining tree of *COI* partial gene sequences of subgenus *Acanthotuberculatus* spp. (26 sequences of 4 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g011>

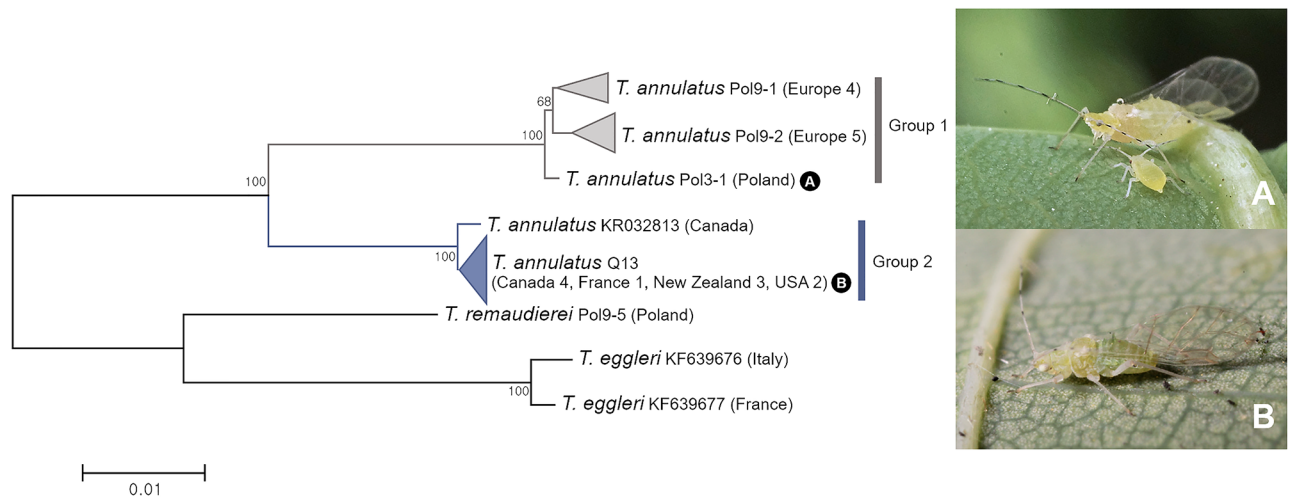


**Fig 12. Neighbor-joining tree of COI partial gene sequences of subgenus *Orienttuberculoides* spp. (86 sequences of 7 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g012>

Fourteen specimens of *Tuberculatus* (*Orienttuberculoides*) *kashiwae* collected from Japan and Korea were divided into two subgroups with intergroup genetic divergence of 5.1% to 5.5% (Fig 12), confirming earlier suggestion by Yao [63]. All Korean specimens were clustered together with *T. (O.) kashiwae* B [63]. Only two sequences from Japan formed group 1 (*T. (O.) kashiwae* A).

A total of 25 individuals of *Tuberculatus* (*Orienttuberculoides*) *yokoyamai* collected from Japan and Korea were analyzed. Results are shown in Fig 12. ABGD and bPTP analyses detected two cryptic species within *T. (O.) yokoyamai*, showing genetic divergence of 3.5% to 4.1% (Fig 12). Thirteen specimens collected from Japan and Korea formed group 1 together



**Fig 13. Neighbor-joining tree of COI partial gene sequences of subgenus *Tuberculoides* spp. (19 sequences of 3 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g013>

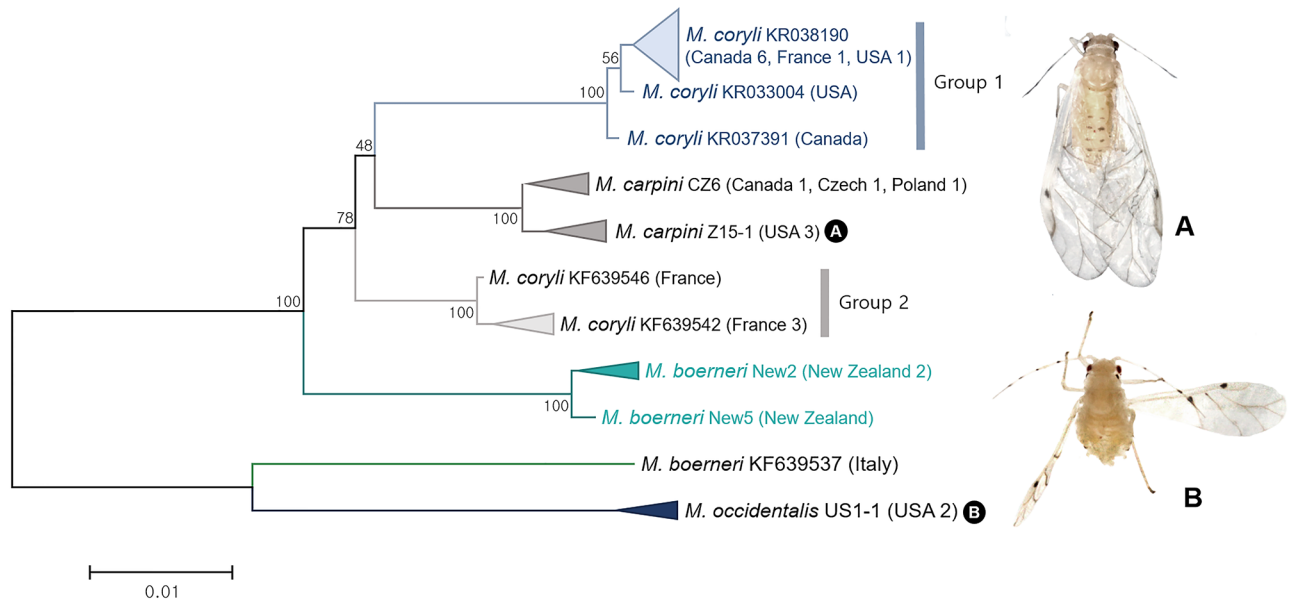
with *Tuberculatus* sp. D Yao (unpublished, COI sequence Genbank accession no. AB861455). Morphologically, group 1 was distinguished from group 2 by having longer length of 5th antennal segment, siphunculi, and cauda (S8 Fig). There was no host-plant difference between the two groups.

COI sequences of 19 specimens of three species in the subgenus *Tuberculoides* revealed two groups within *Tuberculatus* (*Tuberculoides*) *annulatus* (Figs 3 and 13) with genetic divergence of 5.1–5.5%. All 10 individuals of group 1 were collected from Europe: France, Poland and UK. Group 2 comprised of 7 North American specimens, 3 New Zealand specimens, and one French specimen. One French specimen in group 2 has been regarded as an outlier in a previous study of Coeur d’Acier et al. [64]. However, this haplotype was present in North America and New Zealand. Morphological features of the two subgroups (group 1 and group 2) of *T. (T.) annulatus* were compared. Group 2 is distinguishable from group 1 by shorter antennae and smaller cauda knob with longer setae (S9 Fig). Based on molecular and morphological differences, group 1 and group 2 might be distinct species. It has been widely assumed that *T. (T.) annulatus* is introduced into North America, South America, Australia, and New Zealand from Europe [38]. However, such assumption need to be examined in future studies.

**Case II: Discovering possible cryptic or misidentified sequences in NCBI.** A total of 398 COI sequences of 73 species were downloaded from Genbank and used in this study. The following statements are only suggestions since specimen morphology could not be examined in this study.

In genus *Myzocallis*, a total of three possible discrepancies compared to current species concepts were identified. Detailed results of each subgenus are described as follows. We analyzed 26 sequences of four species belonging to subgenus *Myzocallis* collected from Europe: Czech Republic, France, Italy and Poland; North America: Canada and USA; and Oceania: New Zealand. ABGD and bPTP analyses revealed two subgroups in *Myzocallis* (*Myzocallis*) *coryli*. Group 1 and group 2 showed intergroup genetic divergence of 2.9–3.4% (Fig 14). Group 1 of *M. (M.) coryli* formed a sister group of *M. (M.) carpini* with genetic divergence of 2.7–3.1%. Group 1 mostly comprised of North American specimens except for one French specimen (Genabank accession no. KF639545) while all specimens of group 2 were from France. Coeur d’Acier et al. [64] have reported that *M. (M.) coryli* show exceptionally high intraspecific divergence due to the outlier. However, the outlier was the most common haplotype of *M. (M.)*





**Fig 14. Neighbor-joining tree of COI partial gene sequences of subgenus *Myzocallis* spp. (26 sequences of 4 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g014>

*coryli* in our dataset. *M. (M.) coryli* is widely known as a cosmopolitan species. However, our results suggested that this species might be a species complex.

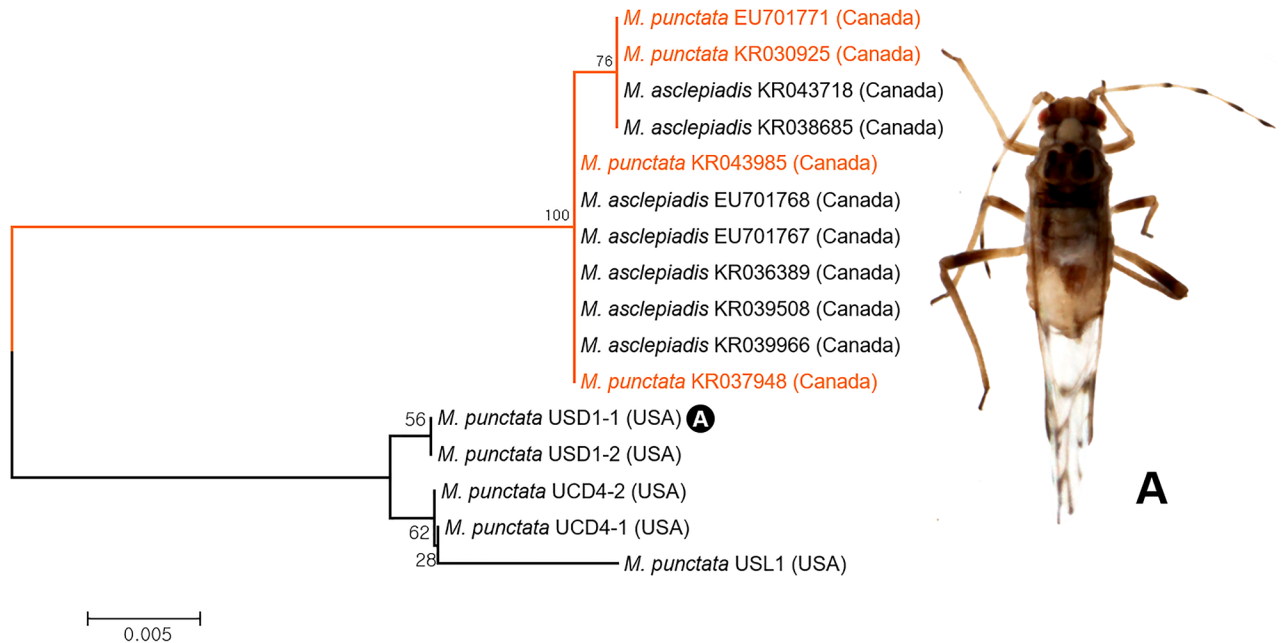
*Myzocallis (Myzocallis) boernerii* specimens collected from New Zealand and Italy showed genetic divergence of 8.4% (Fig 14). This species was originally described from Europe and introduced into New Zealand [65]. Assuming species identification of Italian *M. (M.) boernerii* is correct. Our results suggested that European and New Zealand populations are distinct species.

COI sequences of two *Neosymydobius* species, *Myzocallis (Neosymydobius) asclepiadis* and *M. (N.) punctata*, revealed possible misidentified sequences of *M. (N.) punctata* (Fig 15). As a result of ABGD and bPTP analyses, *M. (N.) punctata* was subdivided into two subgroups (Fig 3). Four sequences of *M. (N.) punctata* (Genbank accession no. EU701771, KR030925, KR043984, and KR037948) were completely identical to those of *M. (N.) asclepiadis* rather than to *M. (N.) punctata* with 100% support value (Fig 15). Our results suggested that 4 individuals (Genbank accession no. EU701771, KR030925, KR043984, and KR037948) of *M. (N.) punctata* might be misidentification of *M. (N.) asclepiadis*.

Among 22 specimens of five *Shivaphis* spp., 7 individuals of *Shivaphis celti* showed extremely high levels of intraspecific distance (16.1% to 16.6%) due to a single sequence (Genbank accession no. JQ920934) (Fig 16). Such a high intraspecific distance suggests that this specimen might be a misidentification or a cryptic species. Subsequent morphological re-examination of the voucher specimen for this sequence (Genbank accession no. JQ920934) is needed.

A total of 20 specimens of four *Tiliaphis* species were analyzed. We found a possible misidentified sequence which was identified as *Tiliaphis shinae* (Genbank accession no. GU978821) (Fig 17). Subsequently, we re-examined the voucher specimen of this sequence and found that this specimen was in fact *T. coreana*.

Three sequences of *Tuberculatus (Tuberculatus) querceus* were divided into two subgroups with divergence of 8.2% (Fig 18). This species is distributed through Europe to South-western

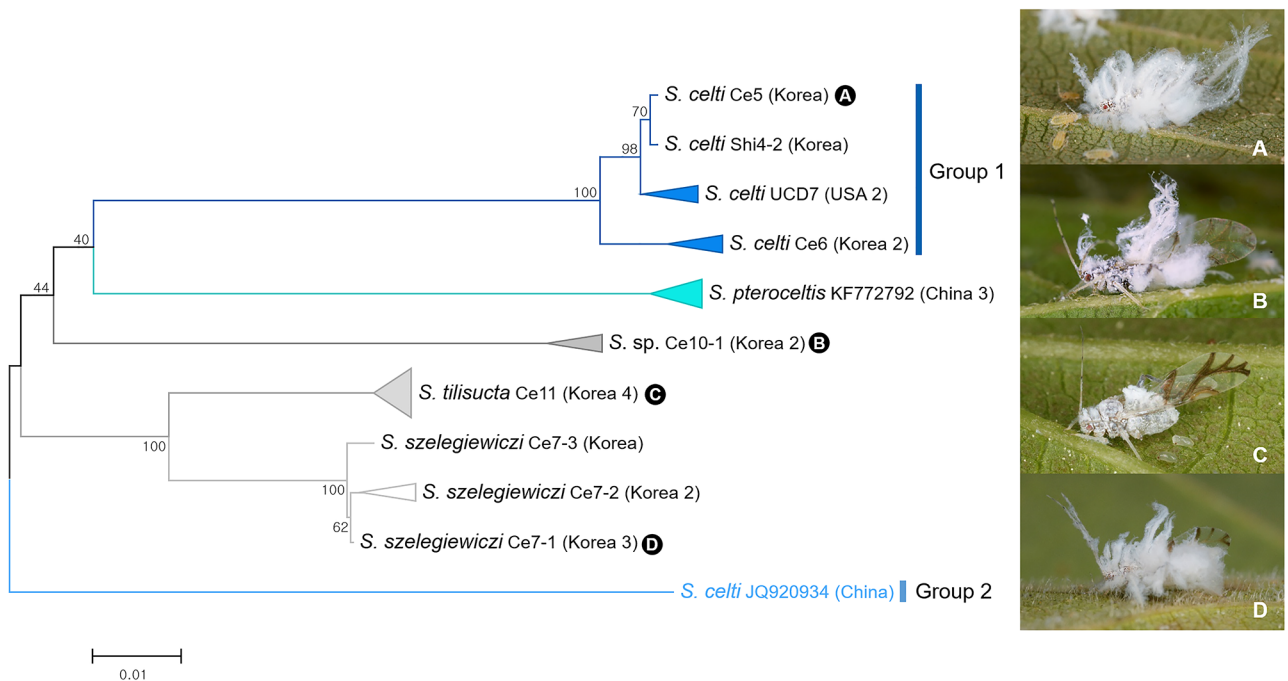


**Fig 15. Neighbor-joining tree of *COI* partial gene sequences of subgenus *Neomyzocallis* spp. (16 sequences of 2 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g015>

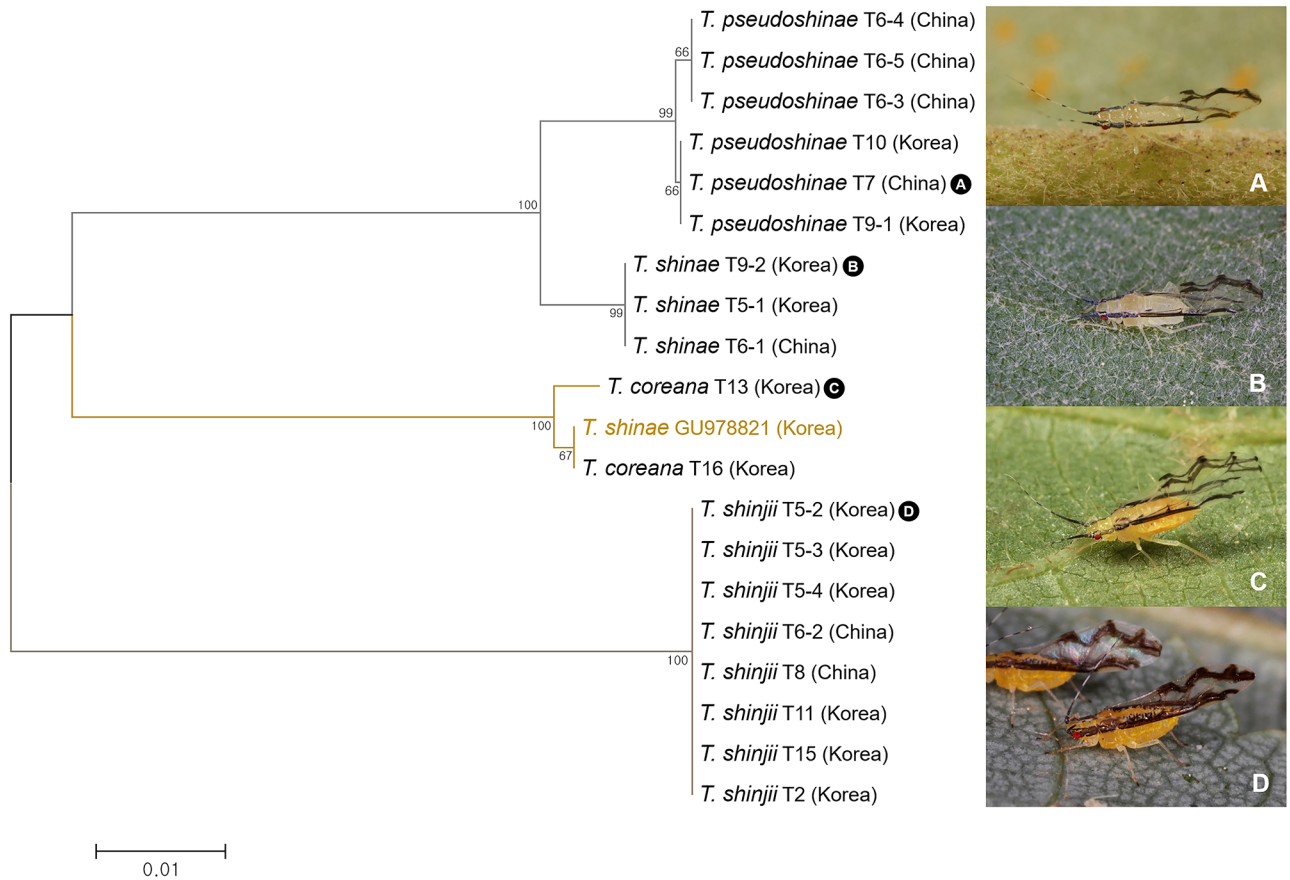
Asia. It is also introduced into Canada [30, 38]. However, our results suggested that European and Canadian specimens might be distinct species.

**Case III: Low genetic distances between morphologically distinct species.** Morphologically, *Pterocallis alnijaponicae* and *P. nigrostriata* are easily distinguishable species (Fig 19). As



**Fig 16. Neighbor-joining tree of *COI* partial gene sequences of *Shivaphis* spp. (22 sequences of 5 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g016>



**Fig 17. Neighbor-joining tree of *COI* partial gene sequences of *Tiliaphis* spp. (20 sequences of 4 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g017>

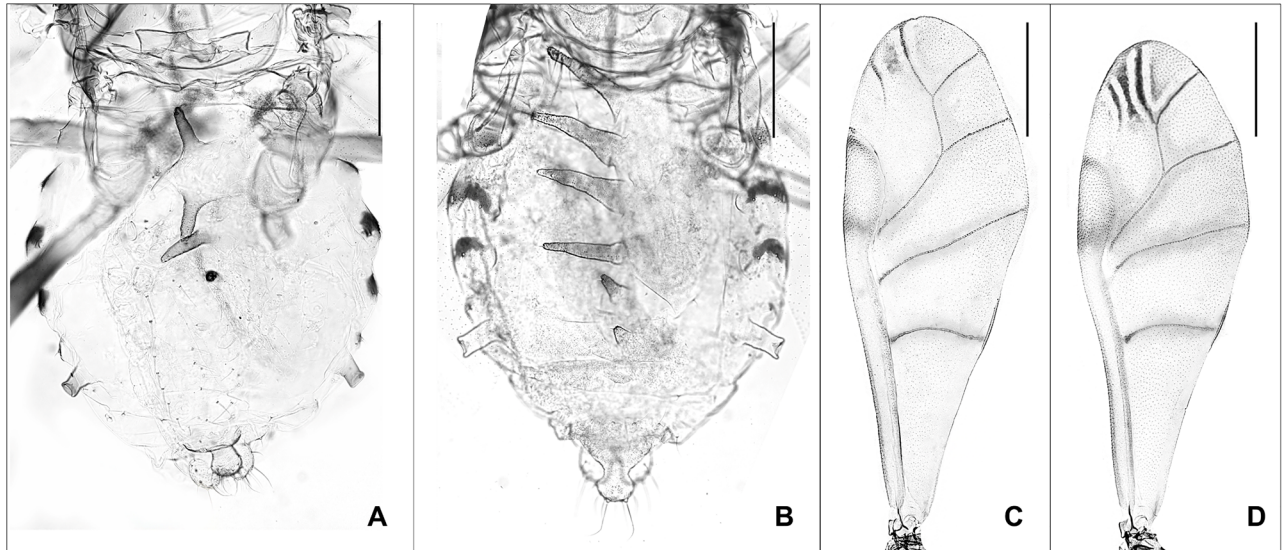
shown in Fig 19, *P. nigrostriata* has 5–6 long and conspicuous dorsal abdominal tubercles and forewing with unique marginal patch while *P. alnijaponicae* has 3–4 short dorsal abdominal tubercles and forewing without marginal patches. Unexpectedly, low genetic divergence (0.5%) between *P. alnijaponicae* and *P. nigrostriata* was found (Fig 20). ABGD and bPTP analyses also supported these results by combining *P. alnijaponicae* and *P. nigrostriata* as a single MOTU (Fig 3). Such discrepancies between morphological and DNA barcoding results raise a question about the validity of these species.

*Tiliaphis pseudoshinae* and *T. shinae* are morphologically similar. However, *T. pseudoshinae* can be distinguished from *T. shinae* by having longer ultimate rostral segment, shorter



**Fig 18. Neighbor-joining tree of *COI* partial gene sequences of *Tuberculatus querceus* (3 sequences).**

<https://doi.org/10.1371/journal.pone.0176582.g018>

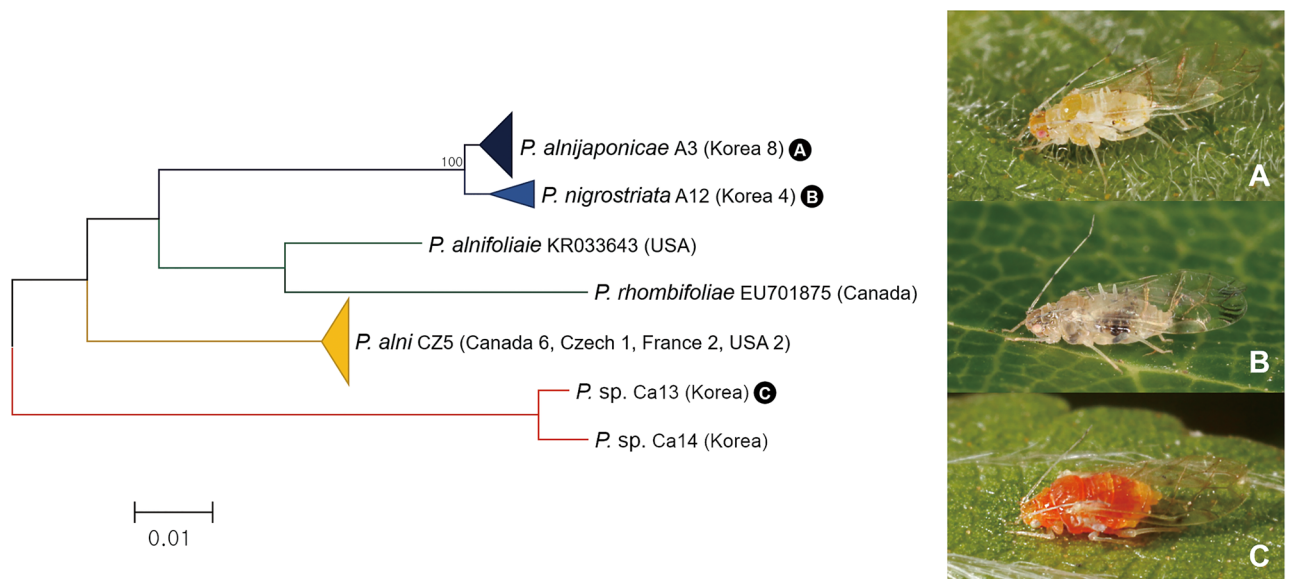


**Fig 19. Alate vivipara of *Pterocallis alnijaponicae* (A, C) and *P. nigrostriata* (B, D).** (A-B) abdomen. (C-D) forewing (scale bars 0.5mm).

<https://doi.org/10.1371/journal.pone.0176582.g019>

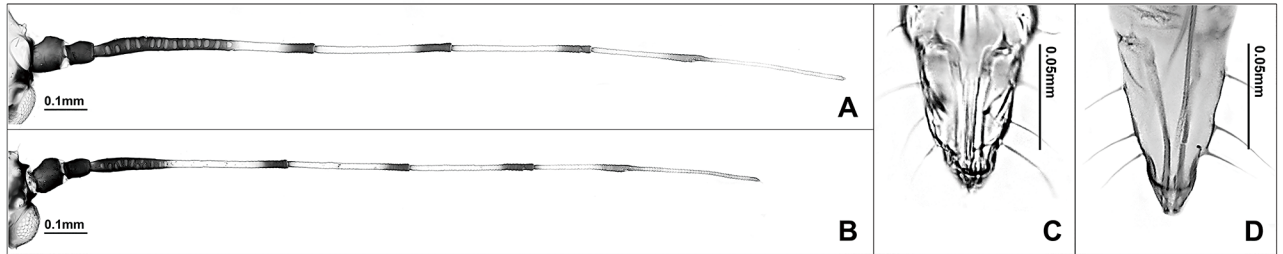
antenna, and less secondary sensoria on the 3rd antennal segment (Fig 21). Between these two species, sequence divergence of 2.1% to 2.5% was observed (Fig 17). Based on ABGD analysis, *T. pseudoshinae* and *T. shinae* were separated into distinct MOTUs. However, bPTP analysis grouped these two species as a single MOTU (Fig 3). Thus, comparing more specimens with multiple generic markers is needed for better delimitation of these species.

Four individuals of *Tuberculatulus (Orientuberculoides) capitatus* and nine individuals of *T. (O.) fangi* were analyzed. *T. (O.) capitatus* and *T. (O.) fangi* can be distinguished by having different shapes and lengths of setae (Fig 22). However, their genetic divergence was only 0.5% to 1.5% (Fig 12). In fact, different shapes, arrangement, and lengths of hairs are often can be



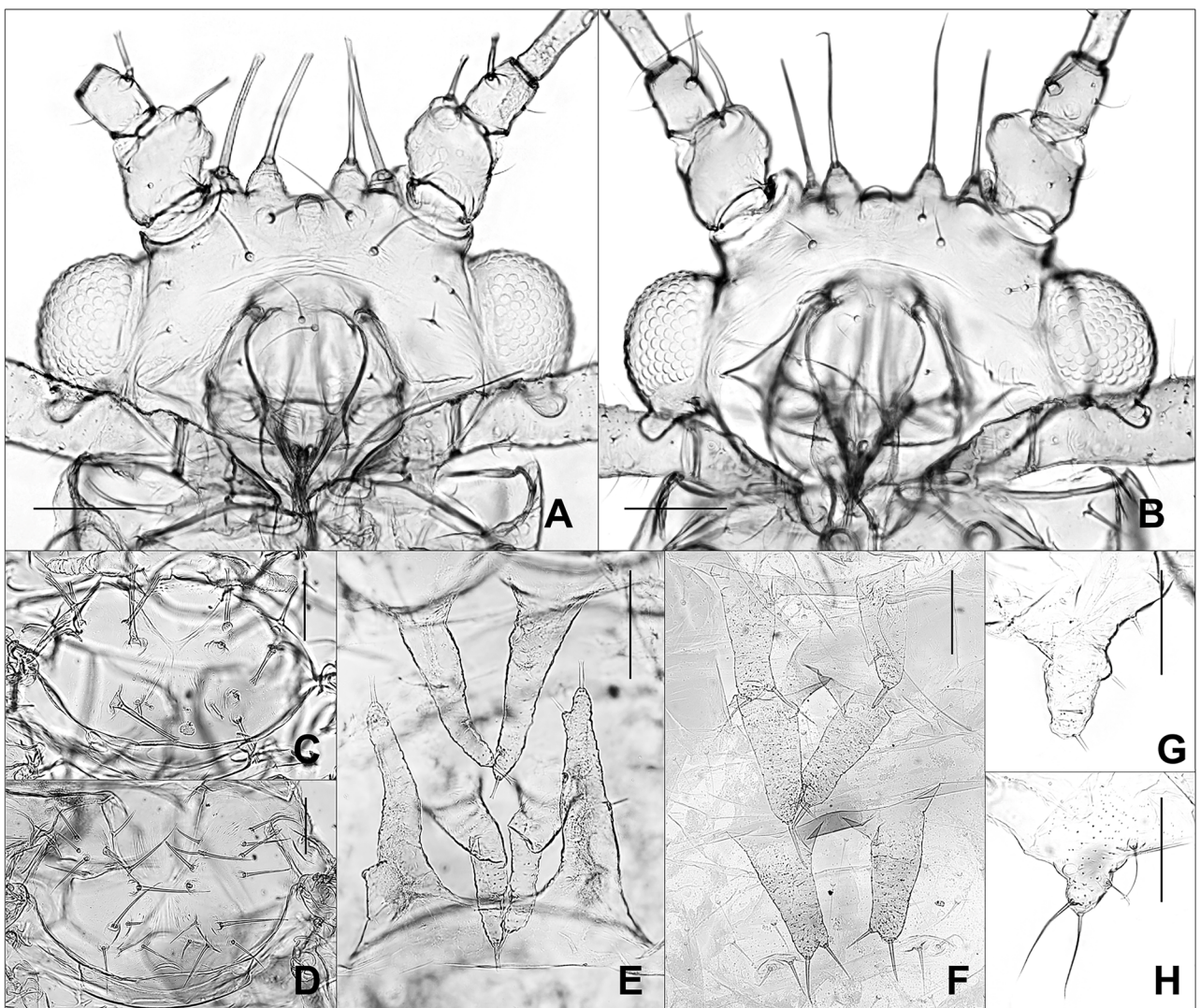
**Fig 20. Neighbor-joining tree of *COI* partial gene sequences of *Pterocallis* spp. (27 sequences of 6 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g020>



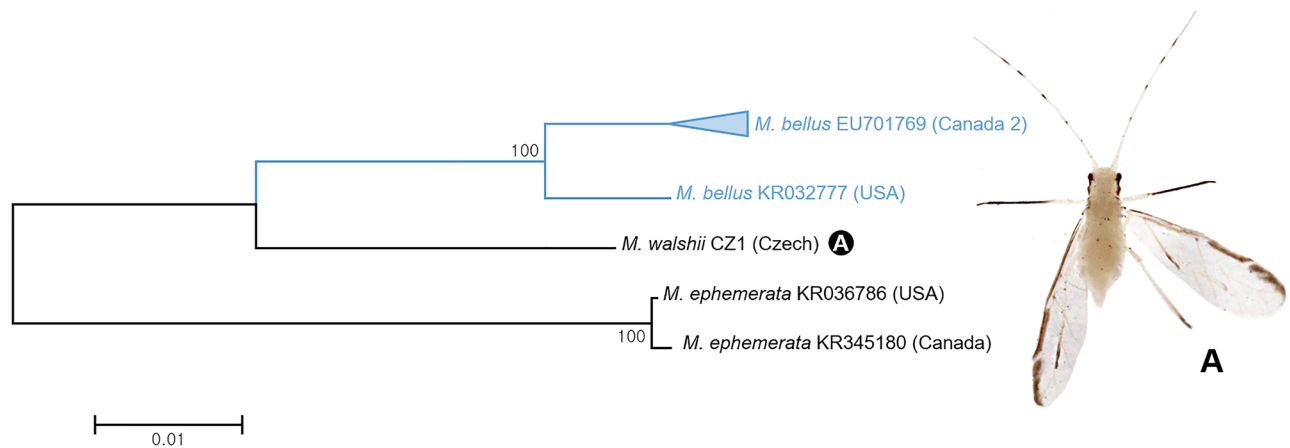
**Fig 21.** Alate vivipara of *Tiliaphis pseudoshinae* (A, C) and *T. shinae* (B, D). (A-B) antenna. (C-D) ultimate rostral segment (scale bars 0.1mm).

<https://doi.org/10.1371/journal.pone.0176582.g021>



**Fig 22.** Alate vivipara of *Tuberculaphis (Orienttuberculoides) capitatus* (A, C, E, G) and *T. (O.) fangi* (B, D, F, H). (A-B) head. (C-D) setae on thorax. (E-F) abdominal dorsal tubercles. (G-H) 4th abdominal marginal tubercle (scale bars, 0.1mm).

<https://doi.org/10.1371/journal.pone.0176582.g022>



**Fig 23. Neighbor-joining tree of *COI* partial gene sequences of subgenus *Lineomyzocallis* spp. (6 sequences).**

<https://doi.org/10.1371/journal.pone.0176582.g023>

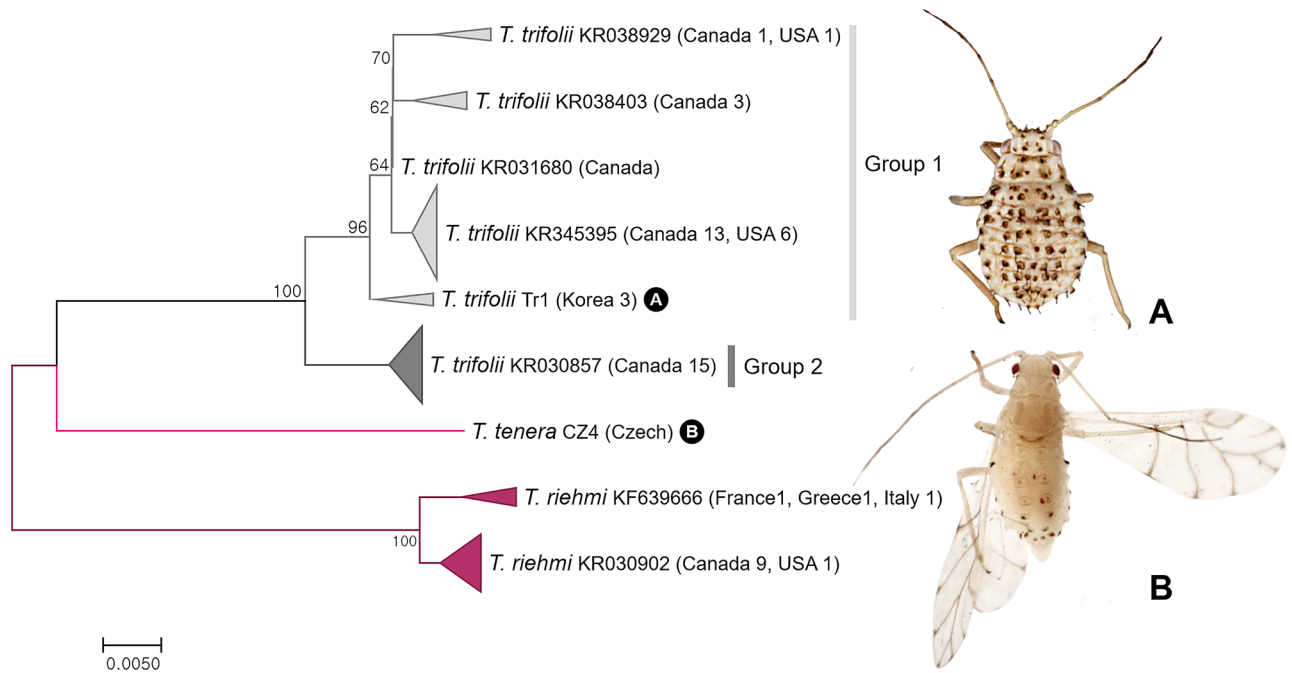
critical characteristics for aphid species delimitation. A review of the variations within and between these species is still required.

**Case IV: Ambiguous genetic distances.** In this study, six individuals of three species in subgenus *Lineomyzocallis* belonging to genus *Myzocallis* were analyzed. DNA barcoding showed 1.7% intraspecific distance between two Canadian and one American specimen of *Myzocallis* (*L.*) *bellus* (Fig 23). ABGD analysis detected them as a single MOTU. However, bPTP analysis separated these Canadian and American specimens as distinct MOTUs. This situation requires further morphological and molecular analysis.

A total of 57 specimens belonging to 3 species of *Therioaphis* were analyzed in this study. As shown in Fig 24, 43 specimens of *Therioaphis* (*Pterocallidium*) *trifolii* were subdivided into two subgroups with intergroup barcode divergence of 1.9% (Fig 24). Group 1 comprised of American, Canadian, and Korean specimens while group 2 only contained Canadian specimens. However, both groups formed a single MOTU according to both ABGD and bPTP analyses. There are four subspecies, *T. (P.) trifolii albae*, *T. (P.) trifolii maculata*, *T. (P.) trifolii trifolii* and *T. (P.) trifolii ventromaculata* in of *T. (P.) trifolii* [14, 38]. Detailed molecular studies have not been conducted on this group of taxa. Thus, comparing worldwide samples of *T. (P.) trifolii* spp. is needed in future studies.

Within *Tuberculatus* (*Arakawana*) *stigmatus*, two distinct subgroups were detected based on NJ tree and bPTP analyses with genetic divergence of 1.5% to 1.9% (Figs 3 and 25). Individuals in group 1 and group 2 showed distinct morphological features (S10 Fig). Group 2 has shorter siphunculi and small abdominal marginal tubercles without marginal tubercle on the 5th abdominal segment (S10 Fig). Based on their morphological differences, group 1 and group 2 might be distinct species despite the relatively low divergence level between these two subgroups.

Among 38 individuals of *Tuberculatus* (*Nippocallis*) *kuricola*, two subgroups (group 1 and group 2) were detected based on NJ tree and bPTP analysis (Fig 3). Genetic divergence between the two groups was 1.9% (Fig 26). Morphologically, the two groups are superficially similar. However, group 2 can be recognized by having 3–4 setae on each dorsal abdominal tubercle (rather than 2 in group 1) with a shorter ultimate rostral segment (S11 Fig). Takahashi (1936) has described subspecies *T. (N.) kuricola cantoensis*. However, it is currently unclear whether *T. (N.) kuricola cantoensis* and *T. (N.) kuricola kuricola* correspond to these two genetically divided subgroups.

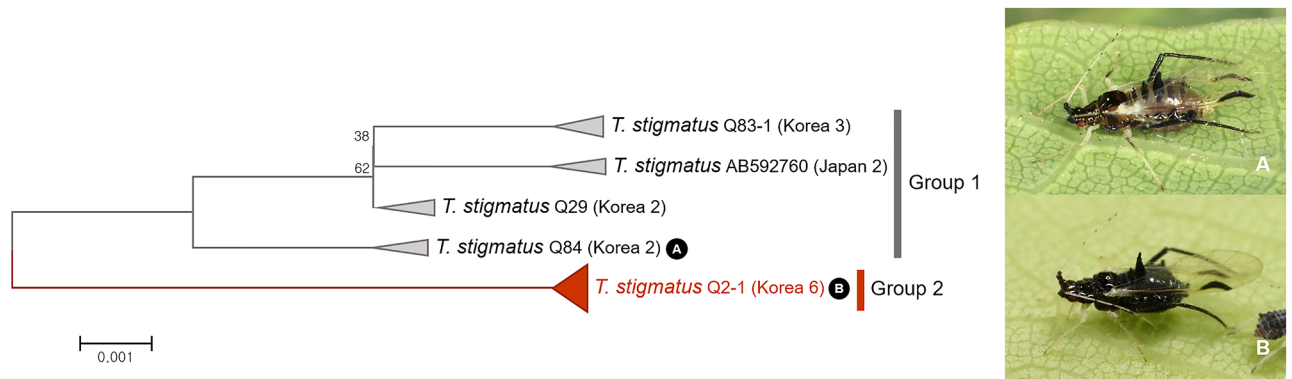


**Fig 24. Neighbor-joining tree of *COI* partial gene sequences of genus *Therioaphis* spp. (57 sequences of 3 morphospecies).**

<https://doi.org/10.1371/journal.pone.0176582.g024>

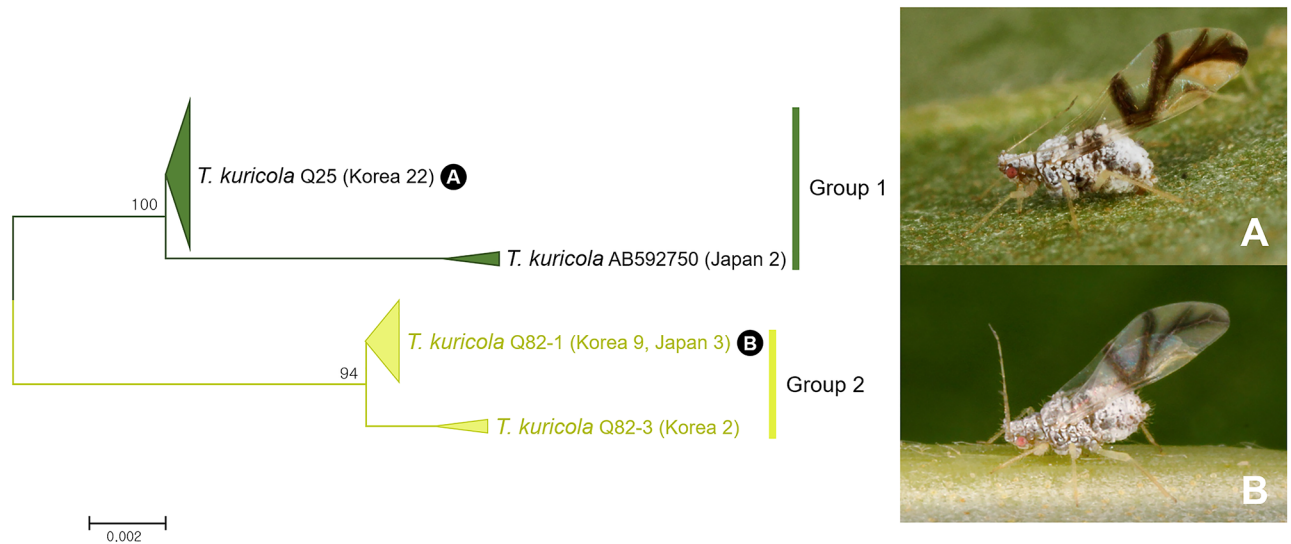
## Discussion

This study demonstrates that DNA barcoding can be used to reliably identify aphid species in the subfamily Calaphidinae. DNA barcoding of 115 morphospecies (899 sequences) revealed 25 discordances between DNA barcoding results and morphology results. These conflicts involved 18 cases of exceptionally high intraspecific distances, three morphologically distinct species pairs with low genetic distances and four cases of ambiguous intraspecific distances. Except for four undeterminable cases, a total of 15 cryptic species were identified from 12 morphospecies. Among these cases, slight morphological differences were detected in seven species complexes. In the most cases, morphological differences were due to different lengths of various body parts rather than a different shapes or numerical characters. Slight length differences



**Fig 25. Neighbor-joining tree of *COI* partial gene sequences of *Tuberculatus (Arakawana) stigmatus* (15 sequences).**

<https://doi.org/10.1371/journal.pone.0176582.g025>



**Fig 26. Neighbor-joining tree of *COI* partial gene sequences of *Tuberculatus (Nippocallis) kuricola* (38 sequences).**

<https://doi.org/10.1371/journal.pone.0176582.g026>

are easy to be overlooked. They can be obscured by intraspecific variations. From these results, we can infer the presence of cryptic diversity in Calaphidinae.

In this study, distinct host-plant associations were revealed in three species complexes: *Mesocallis corylicola*, *Tinocallis (Tinocallis) zekowae* and *Tuberculatus (Acanthocallis) quercicola*. Differences in host-plant associations are often critical for aphid species identification. Host-plant shifts and subsequent genetic differentiation can lead to speciation (ecological species concept) [66–69]. These ecological differences between cryptic species emphasize the necessity to reexamine different host-plant associated populations in a single species.

We also confirmed that species with a wide distribution could have more possibilities to include cryptic species. Some species assessed in this study demonstrated this assumption. For example, *Calaphis flava*, which has been known as invasive species with cosmopolitan distribution, included subgroups with each restricted collection area. However, most previous DNA barcoding studies in aphids targeted samples only collected within restricted area [7, 35, 64]. Our results suggest that multiregional sampling would be very important to recognize cryptic diversity.

However, we also found deep intraspecific divergence within species complexes without obvious host-plant and geographic difference. In fact, groups within *Takecallis arundicolens* and *Tuberculatus (Nippocallis) kuricola* were found together within the same individual colony, respectively. In particular, viviparous females of *T. arundicolens* (group 1 and group 3) were not distinguishable by morphology. Considering that both bamboo and chestnut trees are urban landscape plants, the phenomenon (genetically distinct groups occur in the same colony) might indicate that human mediated transportation might have played a role in subsequent colony merging.

We found three morphologically distinct species pairs with relatively low genetic distances: *Pterocallis alnijaponicae* and *P. nigrostriata*, *Tiliaphis pseudoshinae* and *T. shinae* and *Tuberculatus (Orientatuberculoides) capitatus* and *T. (O.) fangi*. This could be due to many reasons such as rapid radiation, balanced polymorphism and introgressive hybridization [70]. In the three DNA barcode sharing species pairs, it was common that each pair of species was collected together on the same host-plants and geographical regions. Considering that, there can be opportunities of hybridization between species pair during sexual reproduction on the same



host-plant. Such phenomenon has been frequently reported in several taxa of aphids [71–73]. However, conducting additional ecological and molecular research is needed to verify whether barcode sharing is due to hybridization.

In insects, species delimitation threshold value from 2 to 5% of divergence, depending on the group, is generally accepted as a means to estimate species boundaries. For example, 2% intraspecific divergence may indicate the existence of hidden species in Lepidoptera [11, 74]. However, within Diptera, this level of divergence only represents an intraspecific difference [75–77]. In the suborder Heteroptera, threshold value of 2.2% has been applied for DNA barcode based species delimitation [78–79]. In this study, about 2.5% of species delimitation value could be applied for most of species. However, there is clear limits applying certain species delimitation threshold value in every case. In this study, we found highly varied interspecific genetic variation for *COI*, from 0.5% to 20.1% between congeneric species of Calaphidinae. These results suggested that applying a single threshold value might not work for all members in this group. To get more accurate species delimitation results, combination of additional information such as different genetic markers, morphological characters, and ecological differences is required. For example, in the present study, morphological re-examination of two species, *Tuberculatus (Arakawana) stigmatus* and *Tuberculatus (Nippocallis) kuricola*, with intraspecific divergence of 1.9% resulted in the detection of morphological differences, suggesting that additional species are present.

Some aphid species are important pest of various crop and ornamental plants. They play a critical role in ecosystems. Although DNA barcoding has been carried out several times for aphids before, most studies have been mainly focused on the largest subfamily Aphidinae with less emphasis on other subfamilies [7, 13, 30, 64]. DNA barcodes produced in this study are of value in aiding in the identification of species of Calaphidinae. Further, remarkable cryptic diversity and suspicious cases such as barcode sharing species pairs are detected in this study. Our findings suggest that many more cryptic diversity are not yet been uncovered in aphids. For more accurate and higher resolution of possibly overlooked species diversity investigation, future studies should focus on well-designed sampling plan to reflect morphological, ecological and distributional diversities within species.

## Supporting information

**S1 Fig. Neighbor-joinig tree for the 899 individuals of 115 morphospecies based on *COI* barcoding region.**

(TIF)

**S2 Fig. Alate vivipara of 2 subgroups of *Eucallipterus tiliae* group 1 (A-B) and group 2 (C-D).** (A, C) antenna. (B, D) 2nd segment of hind tarsi (scale bars, 0.1mm).

(TIF)

**S3 Fig. Alate vivipara of 2 subgroups of *Mesocallis corylicola* group 1 (A, C) and group 2 (B, D).** (A-B) siphunculi. (C-D) ultimate rostral segment (scale bars, 0.05mm).

(TIF)

**S4 Fig. Alatoid nymph of 3 subgroups of *Takecallis arundicolens* group 1 (A), group 2 (B) and group 3 (C).** (A-C) body (scale bars, 0.5mm).

(TIF)

**S5 Fig. Alate vivipara of 2 subgroups of *Tinocallis zekowae* group 1 (A, C) and group 2 (B, D).** (A-B) ultimate rostral segment. (C-D) abdomen.

(TIF)

**S6 Fig. Alate vivipara of 2 subgroups of *Tuberculatus (Acanthocallis) quercicola* group 2 (A) and group 3 (B).** (A-B) antenna (scale bars 0.1mm).

(TIF)

**S7 Fig. Alate vivipara of 4 subgroups of *Tuberculatus (Orientubercoloides) higuchii* group 1 (A, E, I, M), group 2 (B, F, J, N), group 3 (C, G, K, O) and group 4 (D, H, L, P).** (A-D) cauda. (E-H) siphunculi. (I-L) 2nd segment of hind tarsi. (M-P) antenna (scale bars 0.1mm).

(TIF)

**S8 Fig. Alate vivipara of 2 subgroups of *Tuberculatus (Orientubercoloides) yokoyamai* group 1 (A, C, E) and group 2 (B, D, F).** (A-B) antenna. (C-D) cauda. (E-F) siphunculi (scale bars 0.1mm).

(TIF)

**S9 Fig. Alate vivipara of 2 subgroups of *Tuberculatus (Tuberculooides) annulatus* group 1 (A, C) and group 2 (B, D).** (A-B) antenna. (C-D) cauda (scale bars, 0.1mm).

(TIF)

**S10 Fig. Alate vivipara of 2 subgroups of *Tuberculatus (Arakawana) stigmatus*, group 1 (A, C, E) and group 2 (B, D, F).** (A-B) abdomen. (C-D) siphunculi. (E-F) antenna.

(TIF)

**S11 Fig. Alate vivipara of 2 subgroups of *Tuberculatus (Nippocallis) kuricola*, group 1 (A, C) and group 2 (B, D).** (A-B) ultimate rostral segment. (C-D) abdomen.

(TIF)

**S1 Table. Detailed collection information and Genbank accession numbers of species used in this study.**

(XLSX)

**S2 Table. Detailed collection information and Genbank accession numbers of sequences downloaded from Genbank.**

(XLSX)

## Acknowledgments

We are grateful to Hyoseok Lee for his great support and help particularly collecting samples with Yerim Lee. We thanks to Dr. Hyojoong Kim, Dr. Jaroslav Holman, Dr. Jong-Kook Jung, Dr. Lukasz Depa and Dr. Ram Duwal Keshari. This project was funded by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science (NRF-2015R1D1A1A01059031).

## Author Contributions

**Conceptualization:** YL.

**Data curation:** YL.

**Formal analysis:** YL.

**Funding acquisition:** YL SL WL.

**Investigation:** YL.

**Project administration:** YL.

**Resources:** YL SL MK SIA.

**Supervision:** YL.

**Validation:** YL.

**Visualization:** YL.

**Writing – original draft:** YL.

**Writing – review & editing:** YL SL MK RGF.

## References

1. Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, et al. Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*. 2007; 22(3): 148–155.
2. Ratnasingham S, Hebert PD. BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*. 2007; 7(3): 355–64. <https://doi.org/10.1111/j.1471-8286.2007.01678.x> PMID: 18784790
3. Federhen S. The NCBI taxonomy database. *Nucleic Acids Research*. 2012; 40(D1): D136–D143.
4. Cook LG, Edwards RD, Crisp MD, Hardy NB. Need morphology always be required for new species descriptions? *Invertebrate Systematics*. 2010; 24(3): 322–326.
5. Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrartes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(41): 14812–14817. <https://doi.org/10.1073/pnas.0406166101> PMID: 15465915
6. Stuart BL, Inger RF, Voris HK. High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology Letters*. 2006; 2(3): 470–474. <https://doi.org/10.1098/rsbl.2006.0505> PMID: 17148433
7. Lee W, Kim H, Lim J, Choi HR, Kim Y, Kim YS, et al. Barcoding aphids (Hemiptera: Aphididae) of the Korean Peninsula: updating the global data set. *Molecular Ecology Resources*. 2011; 11(1), 32–7. <https://doi.org/10.1111/j.1755-0998.2010.02877.x> PMID: 21429098
8. Saitoh T, Sugita N, Someya S, Iwami Y, Kobayashi S, Kamigaichi H, et al. DNA barcoding reveals 24 distinct lineages as cryptic bird species candidates in and around the Japanese Archipelago. *Molecular Ecology Resources*. 2015; 15(1): 177–186. <https://doi.org/10.1111/1755-0998.12282> PMID: 24835119
9. Han T, Lee W, Lee S, Park IG, Park H. Reassessment of Species Diversity of the Subfamily Denticolliinae (Coleoptera: Elateridae) through DNA Barcoding. *PloS One*. 2016; 11(2): e0148602. <https://doi.org/10.1371/journal.pone.0148602> PMID: 26848744
10. Hebert PD, Stoeckle MY, Zemlak TS, Francis CM. Identification of birds through DNA barcodes. *PloS Biology*, 2004; 2(10): e312. <https://doi.org/10.1371/journal.pbio.0020312> PMID: 15455034
11. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PD. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(4): 968–971. <https://doi.org/10.1073/pnas.0510466103> PMID: 16418261
12. Hajibabaei M, Singer GA, Hebert PD, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*. 2007; 23(4): 167–172. <https://doi.org/10.1016/j.tig.2007.02.001> PMID: 17316886
13. Footitt RG, Maw HE, Von Dohlen CD, Hebert PD. Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. *Molecular Ecology Resources*. 2008; 8(6): 1189–1201. <https://doi.org/10.1111/j.1755-0998.2008.02297.x> PMID: 21586006
14. Favret C. Aphid species file. Version, 5 (5.0). <http://Aphid.SpeciesFile.org>. (last accessed 3 Sep 2016).
15. Eastop VF. Worldwide importance of aphids as virus vectors. *Aphids as virus vectors*, London. 1977. pp. 3–62.
16. Blackman RL, Eastop VF. *Aphids on the World's Crops: Identification and information Guide* 2nd ed. John Wiley & Sons, England. 2000. pp. 2–466.
17. Teulon DAJ, Stufkens MAW. Biosecurity and aphids in New Zealand. *New Zealand Plant Protection*. 2002. pp. 12–17.
18. Footitt RG, Halbert SE, Miller GL, Maw E, Russell LM. Adventive aphids (Hemiptera: Aphididae) of America North of Mexico. *Proceedings of the Entomological Society of Washington*. 2006; 108(3): 583–610.

19. Messing RH, Tremblay MN, Mondor EB, Footitt RG, Pike KS. Invasive aphids attack native Hawaiian plants. *Biological Invasions*. 2007; 9(5): 601–607.
20. Hille Ris Lambers D. Polymorphism in Aphididae. *Annual review of entomology*. 1966; 11(1): 47–78.
21. Miyazaki M. Forms and Morphs of aphids. In: Minks AK, Harrewijn P (eds). *Aphids their Biology, Natural enemies and control*. Elsevier, Amsterdam. 1987. pp. 27–47.
22. Wool D, Hales DF. Phenotypic plasticity in Australian cotton aphid (Homoptera: Aphididae): host-plant effects on morphological variation. *Annals of the Entomological Society of America*. 1997; 90(3): 316–328. <http://dx.doi.org/10.1093/aesa/90.3.316>.
23. Margaritopoulos JT, Tsitsipis JA, Zintzaras E, Blackman RL. Host-correlated morphological variation of *Myzus persicae* (Hemiptera: Aphididae) populations in Greece. *Bulletin of Entomological Research*. 2000; 90(3): 233–244. <https://doi.org/10.1017/S000748530000353>. PMID: 10996864
24. Weisser WW, Braendle C, Minoretti N. Predator-induced morphological shift in the pea aphid. *Proceedings of the Royal Society of London B: Biological Sciences*. 1999; 266(1424): 1175–1181.
25. Yao I. Ant attendance reduces flight muscle and wing size in the aphid *Tuberculatus quercicola*. *Biology Letters*. 2012; 8: 624–627. <https://doi.org/10.1098/rsbl.2012.0014> PMID: 22378740
26. Hardie J, Lees AD. 13 Endocrine Control of Polymorphism and Polyphenism. In: Kerkut GA, Gilbert LI (eds). *Comprehensive insect physiology biochemistry and Pharmacology, Endocrinology II*. Pergamon Press. 2013. pp. 441.
27. Tsuchida T, Koga R, Horikawa M, Tsunoda T, Maoka T, Matsumoto S, et al. Symbiotic bacterium modifies aphid body color. *Science*. 2010; 330(6007): 1102–1104. <https://doi.org/10.1126/science.1195463> PMID: 21097935
28. Johnson SN, Douglas AE, Woodward S, Hartley SE. Microbial impacts on plant-herbivore interactions: the indirect effects of a birch pathogen on a birch aphid. *Oecologia*. 2003; 134(3): 388–396. <https://doi.org/10.1007/s00442-002-1139-6> PMID: 12647147
29. Blackman RL, Spence JM. The effects of temperature on aphid morphology, using a multivariate approach. *European Journal of Entomology*. 1994; 91: 7–22.
30. Footitt RG, Maw HE, Pike KS. DNA barcodes to explore diversity in aphids (Hemiptera Aphididae and Adelgidae). *Redia*. 2009; 92: 87–91.
31. Massimino Cocuzza GE, Cavalieri V. Identification of aphids of *Aphis frangulae*-group living on *Lamiaceae* species through DNA barcode. *Molecular Ecology Resources*. 2014; 14(3): 447–457. <https://doi.org/10.1111/1755-0998.12199> PMID: 24188728
32. Lee Y, Lee W, Lee S, Kim H. A cryptic species of *Aphis gossypii* (Hemiptera: Aphididae) complex revealed by genetic divergence and different host-plant association. *Bulletin of Entomological Research*. 2015; 105(1): 40–51. <https://doi.org/10.1017/S0007485314000704> PMID: 25413997
33. Lee W, Akimoto SI. Development of new barcoding loci in gall-forming aphids (Eriosomatinae: Eriosomatini): Comparing three mitochondrial genes, *ATP6*, *ATP8*, and *COI*. *Journal of Asia-Pacific Entomology*. 2015; 18(2): 267–275. <http://dx.doi.org/10.1016/j.aspen.2015.03.003>.
34. Liu QH, Jiang LY, Qiao GX. DNA barcoding of Greenideinae (Hemiptera: Aphididae) with resolving taxonomy problems. *Invertebrate Systematics*. 2013; 27(4): 428–438. <http://dx.doi.org/10.1071/IS13014>.
35. Chen R, Jiang LY, Qiao GX. The effectiveness of three regions in mitochondrial genome for aphid DNA barcoding: a case in Lachininae. *PLoS One*. 2012; 7(10): e46190. <https://doi.org/10.1371/journal.pone.0046190> PMID: 23056258
36. Quednau FW. Atlas of the Drepanosiphine aphids of the world. I. Panaphidini Oestlund, 1922—Myzocalidina Börner, 1942 (1930) (Hemiptera: Aphididae: Calaphidinae). American Entomological Institute. 1999. pp. 218.
37. Quednau FW. Atlas of the Drepanosiphine aphid Part II: Panaphidini Oestlund, 1923—Panaphidina Oestlund, 1923 (Hemiptera: Aphididae: Calaphidinae). American Entomological Institute. 2003. pp. 301.
38. Blackman RL, Eastop VF. Aphids on the world's plants: An online identification and information guide. <http://www.aphidwordsplants.info/> (last accessed 3 Sep 2016).
39. Nebreda M, Moreno A, Pérez N, Palacios I, Seco-Fernández V, Ferreres A. Activity of aphids associated with lettuce and broccoli in Spain and their efficiency as vectors of Lettuce mosaic virus. *Virus Research*. 2004; 100(1): 83–88. <https://doi.org/10.1016/j.virusres.2003.12.016> PMID: 15036838
40. Herbert JJ, Mizell RF. Crapemyrtle Aphid, *Sarucallis kahawaluokalani* (Kirkaldy) (Hemiptera: Aphididae). *Encyclopedia of Entomology*. 2008. pp. 1103–1105.
41. Stern VM, Smith RF, Van den Bosch R, Hagen KS. The integration of chemical and biological control of the spotted alfalfa aphid. The integrated control concept. *Hilgardia*. 1959; 29(2): 81–101.

42. Cottrell TE, Wood BW, Ni X. Application of plant growth regulators mitigates chlorotic foliar injury by the black pecan aphid (Hemiptera: Aphididae). *Pest Management Science*. 2010; 66(11): 1236–1242. <https://doi.org/10.1002/ps.2000> PMID: 20715019
43. Wood BW, Tedders WL, Taylor J. Control of pecan aphids with an organosilicone surfactant. *HortScience*. 1997; 32(6): 1074–1076.
44. Halbert SE, Pike KS. An Asian elm aphid (Homoptera: Aphididae) new to North America. *Proceedings of the Entomological Society of Washington*. 1990; 92(4): 672–674.
45. Halbert SE, Choate PM. An Asian Woolly Hackberry Aphid *Shivaphis celti* Das: (Homoptera: Aphididae). Florida Department Agriculture & Consumer Services, Division of Plant Industry. 1998. pp. 2.
46. Kondo T, Cortés RS. *Sarucallis kahawaluokalani* (Kirkaldy) (Hemiptera: Aphididae), a new invasive aphid on San Andres island and mainland Colombia, with notes on other adventive species. *Insecta Mundi*. 2014; 0362: 1–10.
47. Martin JH. The identification of common aphid pests of tropical agriculture. *International Journal of Pest Management*. 1983; 29(4): 395–411.
48. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial *cytochrome c oxidase subunit I* from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*. 1994; 3(5): 294–299. PMID: 7881515
49. Moulton MJ, Song H, Whiting M. Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from Orthoptera (Arthropoda: Insecta). *Molecular Ecology Resources*. 2010; 10: 615–627. <https://doi.org/10.1111/j.1755-0998.2009.02823.x> PMID: 21565066
50. Fontaneto D, Flot J-F, Tang CQ. Guidelines for DNA taxonomy, with a focus on the meiofauna. *Marine Biodiversity*. 2015; 45(3): 433–451.
51. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*. 2013; 30(4): 772–780. <https://doi.org/10.1093/molbev/mst010> PMID: 23329690
52. Tamura K, Stecher G, Peterson D, FilipSKI A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*. 2013; 30(12): 2725–2729. <https://doi.org/10.1093/molbev/mst197> PMID: 24132122
53. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*. 1980; 16(2): 111–120. PMID: 7463489
54. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*. 2003; 270(1512): 313–21.
55. Puillandre N, Lambert A, Brouillet S, Achaz G. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, 2012; 21(8): 1864–1877. <https://doi.org/10.1111/j.1365-294X.2011.05239.x> PMID: 21883587
56. Zhang J, Kapli P, Pavlidis P, Stamatakis AA. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*. 2013; 29(22): 2869–2876. <https://doi.org/10.1093/bioinformatics/btt499> PMID: 23990417
57. Tang CQ, Humphreys AM, Fontaneto D, Barraclough TG. Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. *Methods in Ecology and Evolution*. 2014; 5(10): 1086–94. <https://doi.org/10.1111/2041-210X.12246> PMID: 25821577
58. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006; 22(21): 2688–90. <https://doi.org/10.1093/bioinformatics/btl446> PMID: 16928733
59. Higuchi H. A taxonomic study of the subfamily Callipterinae in Japan (Homoptera: Aphididae). *Insecta Matsumurana*. 1972; 35(2): 19–126.
60. Hopkins GW, Dixon AF. Feeding site location in birch aphids (Sternorrhyncha: Aphididae): The simplicity and reliability of cues. *European Journal of Entomology*. 2000; 97(2): 279–80.
61. Borowiak-Sobkowiak BE, Wilkaniec BA. Occurrence of aphids (Hemiptera, Aphidoidea) on tree and shrubs in Cytadela Park in Pozna. *Aphids and Other Hemipterous Insects*. 2010; 16:2735.
62. Watanabe T, Yao I, Akimoto SI. Taxonomic revision of the *Tuberculatus quercicola* group (Hemiptera: Aphididae: Drepanosiphinae), myrmecophilous aphids associated with *Quercus* species, based on morphometric and molecular phylogenetic studies. *Entomological Science*. 2015; 18(1): 55–65.
63. Yao I. Phylogenetic comparative methods reveal higher wing loading in ant-attended *Tuberculatus* aphids (Hemiptera: Aphididae). *The Canadian Entomologist*. 2011; 143(1): 35–43. <https://doi.org/10.4039/n10-050>.

64. Coeur d'Acier A, Cruaud A, Artige E, Genson G, Clamens AL, Pierre E, et al. DNA barcoding and the associated PhylAphidB@ se website for the identification of European aphids (Insecta: Hemiptera: Aphididae). *PloS One*. 2014; 9(6): e97620. <https://doi.org/10.1371/journal.pone.0097620> PMID: 24896814
65. Rohitha BH. *Myzocallis borneri* (Hemiptera: Aphididae) in New Zealand. *New Zealand Entomologist*. 1982; 7(3): 317–318. <http://dx.doi.org/10.1080/00779962.1982.9722409>.
66. Via S, Bouck AC, Skillman S. Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution*. 2000; 54(5): 1626–1637. PMID: 11108590
67. Drès M, Mallet J. Host races in plant–feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2002; 357(1420): 471–492.
68. Peccoud J, Ollivier A, Plantegenest M, Simon JC. A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(18): 7495–7500. <https://doi.org/10.1073/pnas.0811117106> PMID: 19380742
69. Schluter D. Evidence for ecological speciation and its alternative. *Science*. 2009; 323(5915): 737–741. <https://doi.org/10.1126/science.1160006> PMID: 19197053
70. Sturmbauer C, Meyer A. Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature*. 1992; 358(6387): 578–81. <https://doi.org/10.1038/358578a0> PMID: 1501712
71. Guldmond JA, Dixon AFG, Tigges WT. Mate recognition in *Cryptomyzus* aphids: copulation and insemination. *Entomologia experimentalis et applicata*. 1994; 73(1): 67–75.
72. Thieme T, Dixon AFG. Mate recognition in the *Aphis fabae* complex: daily rhythm of release and specificity of sex pheromones. *Entomologia experimentalis et applicata*. 1996; 79(1): 85–89.
73. Delmotte F, Sabater-Muñoz B, Prunier-Leterme N, Latorre A, Sunnucks P, Rispé C, Simon JC. Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species. *Evolution*. 2003; 57(6): 1291–1303. PMID: 12894937
74. Zahiri R, Lafontaine JD, Schmidt BC, Zakharov EV, Hebert PD. A transcontinental challenge—a test of DNA barcode performance for 1,541 species of Canadian Noctuoidea (Lepidoptera). *Plos One*. 2014; 9(3): e92797. <https://doi.org/10.1371/journal.pone.0092797> PMID: 24667847
75. Meier R, Shiyang K, Vaidya G, Peter PK. DNA barcoding and taxonomy in Diptera: a tale of high intra-specific variability and low identification success. *Systematic Biology*. 2006; 55(5): 715–728. <https://doi.org/10.1080/10635150600969864> PMID: 17060194
76. Renaud AK, Savage J, Adamowicz SJ. DNA barcoding of Northern Nearctic Muscidae (Diptera) reveals high correspondence between morphological and molecular species limits. *BMC Ecology*. 2012; 12(1): 24.
77. Nzelu CO, Cáceres AG, Arrunátegui-Jiménez MJ, Lafías-Rosas MF, Yañez-Trujillano HH, Luna-Caipo DV, et al. DNA barcoding for identification of sand fly species (Diptera: Psychodidae) from leishmaniasis-endemic areas of Peru. *Acta Tropica*. 2015; 145: 45–51. <https://doi.org/10.1016/j.actatropica.2015.02.003> PMID: 25697864
78. Park DS, Foottit R, Maw HE, Hebert PD. Barcoding bugs: DNA-based identification of the true bugs (Insecta: Hemiptera: Heteroptera). *Plos One*. 2011; 6(4): e18749. <https://doi.org/10.1371/journal.pone.0018749> PMID: 21526211
79. Raupach MJ, Hendrich L, Kuchler SM, Deister F, Morinière J, Gossner MM. Building-up of a DNA barcode library for true bugs (Insecta: Hemiptera: Heteroptera) of Germany reveals taxonomic uncertainties and surprises. *PloS One*. 2014; 9(9): e106940. <https://doi.org/10.1371/journal.pone.0106940> PMID: 25203616