Geometric morphometric and molecular techniques for discriminating among three cryptic species of the *Anopheles barbirostris* complex (Diptera: Culicidae) in Thailand

Tanawat Chaiphongpachara,a,*, Tanasak Changbunjongb,c, Sedthapong Laojuna, Suchada Sumruayphold, Nantana Suwandittakula, Kewarin Kuntawonga, Siripong Pimsukae

*a* Department of Public Health and Health Promotion, College of Allied Health Sciences, Suan Sunandha Rajabhat University, Bangkok 10300, Thailand  
*b* Department of Pre-Clinic and Applied Animal Science, Faculty of Veterinary Science, Mahidol University, Nakhon Pathom 73170, Thailand  
*c* The Monitoring and Surveillance Center for Zoonotic Diseases in Wildlife and Exotic Animals (MoZWE), Faculty of Veterinary Science, Mahidol University, Nakhon Pathom 73170, Thailand  
*d* Department of Medical Entomology, Faculty of Tropical Medicine, Mahidol University, Bangkok 10400, Thailand  
*e* School of Public Health, Eastern Asia University, Pathumthani 12110, Thailand

**HIGHLIGHTS**

- DNA barcoding based on the cytochrome c oxidase subunit 1 (*COI*) gene is the most reliable identification tool for the *Anopheles barbirostris* complex.
- Analysis of wing size and shape of *Anopheles dissidens*, *An. saeungae* and *An. wejchoochotei* based on geometric morphometrics revealed differences between species (*p* < 0.05).
- The efficacy of wing shape analysis for species identification of the Barbirostris complex was moderate levels of performance (74.29% accuracy score).

**ARTICLE INFO**

**Keywords:**  
*Anopheles barbirostris* complex  
Species identification  
DNA barcoding  
Geometric morphometrics

**ABSTRACT**

*Anopheles* members of the Barbirostris complex are important vectors of malaria in Thailand. However, they are morphologically indistinguishable because they are closely related species. In this study, wing geometric morphometrics (GM) and DNA barcoding based on the cytochrome c oxidase subunit 1 (*COI*) gene were applied to differentiate cryptic species of the Barbirostris complex in Thailand. Three cryptic species of the Barbirostris complex, *Anopheles dissidens* (19.44%), *Anopheles saeungae* (24.54%), and *Anopheles wejchoochotei* (56.02%) were initially identified using the multiplex polymerase chain reaction assay. DNA barcoding analyses showed low intraspecific distances (range, 0.27%–0.63%) and high interspecific distances (range, 1.92%–3.68%), consistent with the phylogenetic analyses that showed clear species groups. While wing size and shape analyses based on landmark-based GM indicated differences between three species (*p* < 0.05). The cross-validated reclassification revealed that the overall efficacy of wing size analysis for species identification of the Barbirostris complex was less than the wing shape analysis (56.43% vs. 74.29% total performance). Therefore, this study’s results are guidelines for applying modern techniques to identify members within the Barbirostris complex, which are still difficult to distinguish by morphology-based identification and contribute to further appropriate malaria control.

1. Introduction

*Anopheles barbirostris* van der Wulp 1884 is a mosquito member within the Myzorhynchus series of the subgenus *Anopheles* Meigen, which was first reported in eastern Java, Indonesia ([Townson et al., 2013](https://doi.org/10.1016/j.jhelminth.2022.e11261)). Previous reports stated that *An. barbirostris* sensu lato (s.l.) is broadly distributed in the Oriental Region, including Timor Island, the Indonesian archipelago, the Philippines, mainland Southeast Asia, southern China, and South Asia ([Sinka et al., 2011](https://doi.org/10.1016/j.jhelminth.2022.e11261)). *Anopheles barbirostris* s.l. has been confirmed as a species complex that is a group of closely related *Anopheles* species, which has been named the *An. barbirostris* complex or Barbirostris complex ([Satoto, 2001](https://doi.org/10.1016/j.jhelminth.2022.e11261)). Members of the Barbirostris complex consist of six sibling
species, *Anopheles barbirostris* sensu stricto (s.s.), *Anopheles campestris*, *Anopheles dissidens*, *Anopheles saeungae*, *Anopheles vanderwelpi*, and *Anopheles wejchoochotei*, which were identified and confirmed by cyto- genetics, cross-mating experiments, and molecular characteristics (Wang et al., 2014; Taai and Harbach, 2015).

In Thailand, five sibling species of the Barbirostris complex comprise *An. wejchoochotei* (An. campestris-like), *An. saeungae* (An. barbirostris species A2), *An. dissidens* (An. barbirostris species A1), *An. campestris*, and *An. barbirostris* s.s. (An. barbirostris species A4; Brosseau et al., 2019). Additionally, *An. barbirostris* species A3 reported in Kanchanaburi Province is awaiting formal scientific description (Suwannamit et al., 2009). However, some member species within the Barbirostris complex have been strongly linked with transmitting malaria (Townson et al., 2013) and filarial worms (Brigia timori and Brigia malayi; Atmosoojdjo et al., 1977). *Anopheles wejchoochotei* and *An. campestris* feed mostly on humans than on animals, which is also called anthropophilic behavior (Taai and Harbach, 2015). Previous laboratory studies in Thailand found that *An. wejchoochotei* B and E forms from Chiang Mai were highly susceptible to *Plasmodium vivax* infection (66.67% and 64.29% sporozoite rate, respectively; Thongsahuan et al., 2011), and *An. wejchoochotei* (An. campestris-like) from Sa Kao was less susceptible to *P. vivax* infection (sporozoite rate = 23.80%; Apiwathnasorn et al., 2002). Additionally, *An. wejchoochotei* has been linked with malaria transmission in some parts of Thailand, especially in the Sa Kao Province, southeastern region (Limrat et al., 2001). In contrast, *An. saeungae*, *An. dissidens*, and *An. barbirostris* s.s. reveal no evidence of its relationship with the natural transmission of human malaria (Taai and Harbach, 2015). In addition, *An. saeungae* and *An. dissidens* were zoophilic and had low susceptibility to *P. vivax* (sporozoite rate = 6.67% and 9.09%, respectively), whereas *An. barbirostris* species (s.s. A4) were refractory to *Plasmodium falciparum* and *P. vivax* following laboratory experiments (Thongsahuan et al., 2011). Therefore, vector and non-vector separations are crucial to effectively developing mosquito control strategies in endemic malaria areas. However, several member species within the Barbirostris complex are morphologically indistinguishable because they are very close sibling species (also called cryptic species; Taai and Harbach, 2015).

Generally, morphological methods based on taxonomic keys are the gold standard for identifying mosquito species (Chaiphongpachara et al., 2019). However, this method has several limitations, including the need for entomological expertise and high error rates on morphologically close species or cryptic species (Sumruayphol et al., 2020). Therefore, effective modern alternative approaches should support the accuracy of species identification (Chaiphongpachara and Laojun, 2020). Geometric morphometrics (GM) is a modern technique that plays an important role in identifying insect species based on the shape and size differences of specific organs, especially the wings (Lorenz et al., 2017; Changbunjong et al., 2020). The major advantages of GM are its low cost and fast analytic capability; however, its accuracy is dependent on the morphological differences between the species (Chaiphongpachara and Laojun, 2020; Changbunjong et al., 2021). In Thailand, GM has been applied to morphologically identify similar mosquito vectors, e.g., *Aedes* (Sumruayphol et al., 2016; Chonephetsarah et al., 2021), *Culex* (Chaiphongpachara and Laojun, 2019), *Mansonia* (Ruangsittichai et al., 2011), and *Anopheles* mosquitoes (Champakaew et al., 2021; Chaiphongpachara et al., 2022b). In addition, this technique has currently successfully identified two members of the Minimus complex, *An. minimus* and *An. harrisoni* (Chatipiyaphat et al., 2020), and three members of the Maculatus group, *An. maculatus*, *An. sawadwongporni*, and *An. pseudowillmori* (Chaiphongpachara et al., 2019) as key malaria vectors in Thailand. Ruangsittichai et al. (2011) also reported that GM techniques can be combined with DNA barcoding, increasing mosquito species’ identification efficiency in the field.

Therefore, in this study, landmark-based GM technique and DNA barcoding based on the COI gene were applied to differentiate between the three cryptic species of the Barbirostris complex, i.e., *An. wejchoochotei, An. saeungae*, and *An. dissidens*, in Thailand. However, the true species members of all the samples from the field were identified using multiplex PCR before applying the alternative methods. This study’s results provide procedures for the application of modern alternative techniques, which include modern morphometric and molecular techniques, to identify cryptic species within the Barbirostris complex that are still difficult to distinguish using morphology-based identification. Consequently, they may eventually contribute to further effective control and surveillance of malaria vectors.

2. Materials and methods

2.1. Ethics statement

This study was conducted following the guidelines of animal care and use of Suan Sunandha Rajabhat University, Thailand. The Institutional Animal Care and Use Committee of Suan Sunandha Rajabhat University, Bangkok, Thailand (Approval No. IACUC 64–007/2021) approved animal care and all experimental procedures.

2.2. Anopheles mosquito collection

Adult wild *Anopheles* mosquitoes of the Barbirostris complex were collected from nine Thailand provinces, which include, Chachoengsao (13°30’14.1”N 101°47’09.8”E), Chanthaburi (12°51’29.8”N 102°16’32.0”E), Kanchanaburi (14°07’47.4”N 98°59’46.0”E), Ratchaburi (13°22’36.0”N 99°16’38.0”E), Narathiwat (6°21’19.3”N 101°53’41.6”E), Phetchaburi (12°42’11.9”N 99°38’45.5”E), Samut Songkhram (13°23’18.8”N 99°55’35.4”E), Trat (11°53’39.3”N 102°47’46.2”E), and Ubon Ratchathani (14°26’54.1”N 105°12’33.0”E; Figure 1A), based on previous reports of their distribution (Rattanarithikul et al., 2006; Taai and Harbach, 2015; Brosseau et al., 2019) between June and October 2021. Next, adult-stage mosquitoes were captured using a BG-Pro CDC-style mosquito trap (BioGents, Regensburg, Germany) with a BG-lure cartridge (BioGents) and solid carbon dioxide (dry ice; Figure 1B) at night (6.00 p.m.–6.00 a.m.). Consequently, the mosquito specimens from the field were transported to the laboratory at the College of Allied Health Sciences, Suan Sunandha Rajabhat University, Thailand. Subsequently, an adult female *An. barbirostris* complex was identified by standard taxonomic keys following a morphological characteristic (Rattanarithikul et al., 2006), individually stored in 1.5-ml microcentrifuge tubes, and preserved at −20°C until species identification by multiplex PCR assay.
2.3. Multiplex PCR for species identification

None of the member species within the Barbirostris complex has species-specific morphological traits, making them difficult to differentiate using morphological methods. Multiplex PCR based on the COI gene was used to identify this complex’s true cryptic species before evaluating the effectiveness of other modern methods.

Genomic DNA was extracted from more than two legs of each specimen using a FavorPrep Mini Kit (Favorgen Biotech, Ping-Tung, Taiwan) following the manufacturer’s protocol. Table 1 shows the primers used for the multiplex PCR following the protocol described by Wilai et al. (2020). Briefly, all multiplex PCR reactions were performed in an overall volume of 20 μL per reaction, which contain 1 μL of genomic DNA sample, 1x reaction buffer, 3 mM MgCl₂, 0.2 mM dNTPs, 0.2 μM of each primer, and 0.4 units of Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA, USA), made up to a final volume of 20 μL with distilled water. PCR amplifications were conducted under the following temperature cycling: initial denaturation at 95 °C for 2 min before 40 cycles of amplification at 95 °C for 30 s, 45 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 5 min. Finally, the negative controls were added to each amplification round by replacing the DNA template with distilled water.

Subsequently, all PCR products were run on a 2% agarose gel stained with Midori Green DNA stain (Nippon Gene, Tokyo, Japan) and visualized using an ImageQuant LAS 500 imager (GE Healthcare Japan Corp., Tokyo, Japan). Table 1 presents the species-specific band sizes of the PCR amplicons to identify cryptic species of the Barbirostris complex as described by Wilai et al. (2020).

2.4. DNA barcoding analysis

Ten samples of each species were randomly selected for species identification after the multiplex PCR using the DNA barcoding analysis. The PCR used the COI barcoding region and amplified using universal forward and reverse primers (Table 1). First, each reaction was performed in an overall volume of 25 μL, which contains 4 μL of DNA template, 1x reaction buffer, 3 mM MgCl₂, 0.2 mM dNTPs, 0.2 μM of each primer, and 0.4 units of Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA, USA), made up to a final volume of 20 μL with distilled water. PCR amplifications were conducted under the following temperature cycling: initial denaturation at 95 °C for 2 min before 40 cycles of amplification at 95 °C for 30 s, 45 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 5 min. Finally, the negative controls were added to each amplification round by replacing the DNA template with distilled water.

Table 1: Primers used in this study for multiplex PCR and DNA barcoding.

| Anopheles species          | Primer name | Sequence (5’– 3’)                  | Length (bp) | Reference                  |
|----------------------------|-------------|------------------------------------|-------------|----------------------------|
| Multiplex PCR              |             |                                    |             |                            |
| Universal reverse          | HCO2198     | TAA ACT TCA GGG TGA CCA AAA AAT CA | Wilai et al. (2020) |
| An. barbirostris s.s.      | BARA4       | [long GC tail]* AAT AGT AGG AAC TTC TTT ATG A | 706         |                            |
| An. dissidens              | BARA1       | ATT ACT ACT GTT ATT AAT ATA GGA     | 238         |                            |
| An. saeungae               | BARA2       | TTA GGT CAC CCA GGA GCA            | 611         |                            |
| An. wejchoochotei          | BARWEJ      | GAT TTG GAA ACT GAT TAC TG         | 502         |                            |
| An. barbirostris A3        | BARA3       | CGG AAC TGG ATG AAG TGT A          | 365         |                            |
| DNA barcoding              |             |                                    |             |                            |
| Universal forward          | forward     | GGA TTT GGA AAT TGA TTA GTT CCT T  | ~700        | Kumar et al. (2007)        |
| Universal reverse          | reverse     | AAA AAT TTT AAT TCC AGT TGG AAC AGC|             |                            |

* Long GC tail sequence: GCG GGC AGG GCG GCG GCG GGCC GGCC GGCC CC was attached to the BARA4 primer of An. barbirostris s.s. to increase the product band size (Wilai et al., 2020).
buffer, 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.2 μM of each primer, 5% dimethyl sulfoxide, and 1.5 units of Platinum Taq DNA polymerase (Invitrogen), made up to a final volume of 25 μL with distilled water. The temperature cycling for PCR was performed as follows: initial denaturation at 95 ºC for 5 min, followed by five cycles of denaturation at 94 ºC for 40 s, annealing at 45 ºC for 60 s, and extension at 72 ºC for 1 min, 35 cycles of denaturation at 94 ºC for 40 s, annealing at 54 ºC for 60 s, and extension at 72 ºC for 1 min, with a final extension at 72 ºC for 10 min. Finally, the negative controls were applied in all the PCR rounds.

Clear DNA bands at ~700 bp in a volume of 1 μL on a 1.5% agarose gel were used to evaluate the quality and reliability of the PCR products. Consequently, these high-quality products were transported to SolGent Inc (Daejeon, Korea) for nucleotide sequencing in both directions following the Sanger sequencing technique.

The obtained trace files (also known as chromatograms) of COI barcode sequences were visualized and manually edited, and ambiguous sites were removed using BioEdit software (Hall, 1999). Subsequently, consensus sequences were made for each individual using the forward and reverse sequences. Next, Anopheles species were assessed by comparing all final sequences with various reference barcode sequences using the Basic Local Alignment Search Tool (BLAST; https://blast.ncbi.nlm.nih.gov/Blast.cgi) in the GenBank database and the BOLD Identification System (https://www.boldsystems.org/index.php/IDS_OpenIdEngine) in the Barcode of Life Data (BOLD) database. All COI sequences were aligned with the Clustal W (Larkin et al., 2007) option in MEGA X software (Kumar et al., 2018), and intra- and interspecific genetic divergences were calculated based on the Kimura 2-parameter distance model (K2P). Lastly, phylogenetic trees were constructed based on the maximum-likelihood (ML) and neighbor-joining (NJ) methods with bootstrapping (1000 replicates) were built by using MEGA X to examine the genetic relationships among members of the Barbirostris complex for species identification. T92 + G (Tamura 3-parameter with gamma distribution parameter) was the most effective nucleotide substitution model employed in this study for the ML phylogenetic analysis.

2.5. Geometric morphometric analysis

After the morphological identification and molecular analysis, the same set of An. barbirostris complex specimens was employed in the GM analysis. Overall, 140 individuals of the An. barbirostris complex with complete wings (38 An. dissidens, 40 An. saeungae, and 62 An. wejchoochotei) were used for the species investigation by the GM analysis.

2.5.1. Specimen preparation and image manipulation

Anopheles specimens with the completed right wing of each cryptic species were selected and prepared for GM analysis. Specifically, the right-wing female Anopheles mosquitoes were dissected from the thorax and mounted on microscope slides with coverslips using Hoyer’s medium. Next, all wing slides were photographed using a digital camera (Nikon DS-Fi3, Tokyo, Japan) connected to a Nikon SMZ 800 N stereo-microscope (Nikon Corp., Tokyo, Japan), with a size scale of 1 mm in length provided for each image.

2.5.2. Landmark digitizing

Eighteen landmarks were digitized (Figure 2A) from the wing images. Additionally, the landmark coordinate format was based on previous

![Figure 2. Right wing of the Anopheles barbirostris complex showing the position of 18 landmarks digitized for geometric morphometric analysis (A). Superposition of the mean landmark configurations of An. dissidens (green), An. saeungae (blue), and An. wejchoochotei (red) showing the wing shape difference between species (B).](image-url)
studies, which are effective locations for analyzing mosquito wings using GM (Lorenz et al., 2012; Demari-Silva et al., 2014; Louise et al., 2015; Sauer et al., 2020).

2.5.3. Repeatability and allometry

The repeatability test assessed the digitization accuracy of the coordinates. First, 20 wing images per species were selected randomly and digitized twice by the same user. Next, both image sets were compared and computed using the *repeatability index* as described by Arnqvist and Mårtensson (1998). Additionally, the size-to-shape contribution was evaluated to estimate assignment to its closest group. Furthermore, a hierarchical clustering tree based on Mahalanobis distances was conducted for each sample to estimate assignment to its closest group. Furthermore, a hierarchical clustering tree based on Mahalanobis distances was built to assess the similarity in wing shape among cryptic species within the Barbirostris complex.

2.5.6. Geometric morphometric software

The CLIC package version 99 (freely available at http://mome-clic.com; Dujardin, 2008) and the online XYOM (XY Online Morphometrics) tool (https://xyom.io/; Dujardin and Dujardin, 2019) were used for geometric morphometric analysis in this study. The CLIC software performed landmark digitization, a repeatability test, and an allometric analysis. In contrast, the online XYOM was used to perform wing size and shape analyses, validate classification, and build a hierarchical clustering tree.

3. Results

In this study, 1,431 Anopheles mosquitoes were collected from nine provinces in Thailand. Moreover, 216 specimens were morphologically identified to belong to the An. barbirostris complex. Therefore, all specimens of the Barbirostris complex were examined at the species level using multiplex PCR assays before performing DNA barcoding and geometric morphometric analyses.

3.1. Multiplex PCR

Three cryptic species of the Barbirostris complex, including An. dissidens, An. saeungae, and An. wejchoochotei, were recognized using the multiplex PCR assay. Figure 3 shows the agarose gel electrophoresis of multiplex PCR products according to species-specific differences. The most predominant species of the Barbirostris complex in this survey was An. wejchoochotei (56.02%, n = 121), followed by An. saeungae (24.54%, n = 53) and An. dissidens (19.44%, n = 42; Table 2). The results of the current study also revealed that the three species could be found in the same area (sympatric species) in many provinces, including Kanchana-buri, Narathiwat, and Trat Provinces.

3.2. DNA barcoding

Partial sequences of the COI gene (~707 bp) were amplified from 30 An. barbirostris complex individuals (10 individuals per species) using

---

**Figure 3.** PCR product from the multiplex PCR assay on a 2% agarose gel for species identification of the *Anopheles barbirostris* complex. Lane 1: 3000 bp molecular ladder, lanes 2–3: An. dissidens (238 bp), lanes 4–7: An. saeungae (611 bp), lanes 8–9: An. wejchoochotei (502 bp), and lane 10: negative control.
universal barcode primers. Table 3 presents the samples with unambiguous (no double peak and no noise) sequences submitted to the GenBank database under their accession numbers. Notably, 30 sequences of the An. barbirostris complex were compared with other mitochondrial sequences through BLAST in the GenBank and BOLD databases. All sequences of the current study corresponded with the multiplex PCR-indicated species (>98% similarity with reference sequences in the databases). Furthermore, the average nucleotide composition percentages for 30 sequences of the three cryptic species of the An. barbirostris complex were 39.8% (thymine, T), 30.0% (adenine, A), 15.6% (guanine, G), and 14.6% (cytosine, C). Additionally, the A + T content (69.8%) was higher than the G + C content (30.2%). Table 4 shows the average inter- and intraspecific pairwise divergence under the K2P model of An. dissidens, An. saeungae, and An. wejchoochotei. Each cryptic species showed a low level of average intraspecific genetic distances (range, 0.27%–0.63%) and a high level of average interspecific distances (range, 1.92%–3.68%). Furthermore, the overall mean genetic distance was 2.20%.

Table 2. Number of the Anopheles barbirostris complex collected in this study based on the multiplex PCR assay.

| Province         | Total number of the Anopheles barbirostris complex |
|------------------|---------------------------------|
|                  | An. dissidens | An. saeungae | An. wejchoochotei |
| Chachoengsao     | 12             | -             | 11                |
| Chanthaburi      | -              | -             | 22                |
| Kanchanaburi     | 8              | 8             | 30                |
| Narathiwat       | 7              | 9             | 6                 |
| Phetchaburi      | -              | -             | 10                |
| Ratchaburi       | -              | 8             | 26                |
| Samut Songkhram  | -              | -             | 12                |
| Trat             | 15             | 6             | 14                |
| Ubon Ratchathani | -              | 12            | -                 |
| Total            | 42             | 53            | 121               |

Both phylogenetic analyses, including ML and NJ methods, revealed similar results of genetic relationships among three cryptic species of the An. barbirostris complex (Figure 4). Phylogenetic analyses demonstrated a clear-cut separation between the three species with bootstrap support values between 75% and 92% based on the ML method and 91% and 99% based on the NJ method. Additionally, the phylogenetic relationship revealed that the An. saeungae cluster was more closely related to the An. wejchoochotei cluster than that of An. dissidens, whereas Anopheles nigerrimus as an outgroup species that belongs to the Nigerrimus subgroup of the Hyrcanus group was separated from clusters of the Barbirostris complex.

3.3. Geometric morphometrics

3.3.1. Repeatability and allometry

The precision of determining landmark coordinates can be estimated from the repeatability scores. Additionally, repeated measurements between tested and compared image sets showed high repeatability scores in both wing size (99.6%) and shape (97.2%). Furthermore, allometric examination indicated that wing size affects the wing shape (Figure 5). The influence of size on the discriminant factor derived from the Procrustes residuals was 8% on the first discriminant factor (DF1; p < 0.05).

3.3.2. Wing size variation

The wing CS variation of An. dissidens, An. saeungae, and An. wejchoochotei is illustrated in Figure 6. The analysis of the wing size indicated that An. wejchoochotei had the highest wing CS (average CS = 3.43 mm), followed by An. saeungae (3.22 mm) and An. dissidens (3.10 mm). Significant wing CS differences appeared in all pair species (p < 0.05, Table 5).

3.3.3. Wing shape variation

The graphic wireframe is based on the superposition of the mean landmark configurations of An. dissidens, An. saeungae, and An. wejchoochotei revealed a slight difference between species (Figure 2B). The factor map based on discriminant analysis displayed an overlapping distribution between the three species (Figure 7). The comparison of pairwise Mahalanobis distance values between species indicated that wing shape significantly differed among the three species (p < 0.05, Table 6). A hierarchical clustering tree based on Mahalanobis distances of An. dissidens, An. saeungae, and An. wejchoochotei are shown in Figure 8. The wing shape of An. dissidens was distinct from the other two taxa, whereas An. saeungae and An. wejchoochotei were close to each other.

3.3.4. Species validation and phenetic wing morphology relationships

A cross-validated reclassification procedure to verify the correct assignment based on the wing size of An. dissidens, An. saeungae, and An. wejchoochotei provided low percentages (56.43% accuracy score, Table 7), which ranged from 40% to 69.35%. The verification of accuracy in the correct assignment based on wing shape showed moderate levels of performance (74.29% accuracy score) ranging from 55% to 83.87%.

4. Discussion

The An. barbirostris complex is a group comprising closely related and morphologically indistinguishable mosquito species. In addition, An.
Figure 4. Maximum-likelihood phylogenetic analysis based on COI gene sequences of *Anopheles dissidens*, *An. saeungae*, and *An. wejchoochotae* (n = 30, black labels) from the present study and reference species from GenBank (n = 6, red labels). One *An. nigerrimus* sequence (n = 1, red labels) was used as an outgroup to root the tree. Bootstrap values (1,000 replicates) of maximum likelihood (green) and neighbor-joining (blue) appeared near the branches. Only bootstrap values ≥50% are shown.
barbirostris s.l. is a potential malaria vector in Thailand (Taai and Harbach, 2015). The reports of natural infections with Plasmodium parasites, particularly P. vivax (Thongsahuan et al., 2011) in An. barbirostris s.l. were observed in the various provinces of Thailand, e.g., Mae Hong Son (Somboon et al., 1994) and Tak (Sriwichai et al., 2016). Multiplex PCR based on the COI gene was recently developed to identify members of this complex using allele-specific primers (Wilai et al., 2020), which has the advantage of being faster and less expensive than the multiplex PCR assay based on the internal transcribed spacer 2 sequences, which required two PCR steps (Brosseau et al., 2019). Therefore, this technique was applied to initially screen the cryptic species within the Barbirostris complex from the nine provinces in Thailand and discovered that An. wejchoochotei was the most abundant, followed by An. saeungae and An. dissidens. Similar to the results of previous surveys, An. wejchoochotei is widely distributed in Thailand, confirmed by molecular biology techniques, with the following provinces reporting the distribution of An. wejchoochotei: Ayutthaya, Chanthaburi, Chiang Mai, Chaiyaphum, Chumphon, Kamphaeng Phet, Khon Kaen, Mahasarakham, Mukdahan, Prachuap Khiri Khan, Sa Kaeo, and Udon Thani (Taai and Harbach, 2015). Anopheles saeungae can be found in many provinces in Thailand, e.g., Lampang, Phetchaburi, Ratchaburi, Sa Kaeo, Trat, Ubon Ratchathani, Ratchaburi, and Udon Thani (Taai and Harbach, 2015). However, An. dissidens is mostly found in mountainous areas with provinces of reported distributions, including Chiang Mai, Mae Hong Son, Sa Kaeo, Tak, and Trat. Unfortunately, An. campestris and An. barbirostris s.s. were not observed in this study, which is consistent with previous reports that both species are rare in Thailand (Thongsahuan et al., 2011; Taai and Harbach, 2015). Anopheles campestris is found only in the lowland areas of the central and eastern regions and is not distributed in the foothills. However, An. barbirostris s.s. was reported to be found in a high mountain area in Chiang Mai (Thongsahuan et al., 2011; Taai and Harbach, 2015). Additionally, the current survey showed that An. dissidens, An. saeungae, and An. wejchoochotei were classified as sympatric species in Kanchanaburi, Narathiwat, and Trat Provinces. These findings suggested a high tendency for species misidentification because they have not yet been described morphologically in species-specific characters (Taai and Harbach, 2015; Sriwichai et al., 2016).

DNA barcoding is a molecular technique for taxonomic identification based on universal DNA barcodes to amplify standardized short DNA regions (400–800 bp; Hebert et al., 2003). Although unknown species based on damaged specimens are not suitable for detection by multiplex PCR, DNA barcoding can identify them. The results of genetic distances (low level of intraspecific distances [range, 0.27%–0.63%] and high level of interspecific distances [range, 1.92%–3.68%]) indicated a barcoding gap. The barcoding gap, which can be examined from intraspecific genetic distances lower than interspecific distances, confirms the efficiency of species classification by this molecular technique (Meyer and Paulay, 2005; Justi et al., 2014). This result is consistent with this study’s phylogenetic analyses showing clear groupings among the species. Thus, DNA barcoding is an effective tool for identifying unknown species within the Barbirostris complex. However, several reports explained that DNA barcoding was not effective enough to identify some Anopheles species genetically very close, e.g., Anopheles epiroticus, to other members within the Anopheles sundis complex (Syafruddinid et al., 2020) and both Anopheles albertyi and Anopheles strodei within the An. strodei subgroup.

### Table 5. Average wing centroid size (CS) and statistical differences of An. dissidens, An. saeungae and An. wejchoochotei.

| Anopheles species | Average wing CS (mm) | S.D. | S.E. | Min-Max |
|-------------------|----------------------|------|------|---------|
| An. dissidens     | 3.10<sup>a</sup>     | 0.14 | 0.02 | 2.82–3.39 |
| An. saeungae      | 3.22<sup>b</sup>     | 0.26 | 0.04 | 2.82–3.85 |
| An. wejchoochotei | 3.43<sup>c</sup>     | 0.25 | 0.03 | 2.68–3.85 |

Different superscript letters show significant differences between species at p-value < 0.05.
study, the specimens of the Barbirostris complex had relatively low effects of wing size on the shape but found a relationship between size and shape. However, this relationship usually does not affect inter-species investigations (Lorenz et al., 2017). In comparison, wing size and shape analyses based on landmark-based GM revealed differences between species \((p < 0.05)\). However, the accuracy of the species identification of members within the Barbirostris complex showed different results between wing size and shape. Overall, the efficacy for species identification of three cryptic species within the Barbirostris complex based on wing shape obtained from cross-validated reclassification showed that 74.29% were correctly classified. In contrast, the efficacy for species identification based on wing size was 56.43%. Therefore, wing size is an unsuitable variable for member identification of the Barbirostris complex because environmental conditions easily affect the size more than shape variables (Lorenz et al., 2017; Phanitchat et al., 2019).

Based on the validated reclassification scores of the current study, the GM performance is imperfect (74.29% accuracy score for wing shape analysis) and cannot replace molecular biological methods in identifying cryptic species within the Barbirostris complex. These results agreed with previous studies reporting that genetic markers provide more sensitivity and accuracy in detecting targeted population variability than wing shape analysis (Carvajal et al., 2021). However, this modern morphometrics is suitable for the initial species assessment of the Barbirostris complex in the field in the event of budget constraints and a lack of advanced equipment.

Interestingly, the hierarchical clustering tree pattern based on wing shape between species was consistent with ML and NJ phylogenetic analyses based on the COI gene. *Anopheles saeungae* was more closely related to *An. wejchoochotei* than *An. dissidens*. Several previous reports have explained that variations in wing shape are related to the genetic background (Vicente et al., 2011; Gómez and Correa, 2017; Carvajal et al., 2021). Similar to the current study, applying the GM technique to 19 mosquito species in Germany revealed that both wing morphometric analyses were consistent with the results of DNA barcoding (Sauer et al., 2020).

In summary, DNA barcoding showed satisfactory results, whereas wing GM showed moderate performance levels. Correct species identification contributes to a better understanding of the Barbirostris complex’s cryptic species-specific bionomics and distributions (Udom et al., 2021). Nevertheless, further studies on many specimens and adding to the missing rare species (*An. campestris* and *An. barbirostris* s.s.) are needed to make the guidelines more complete for identifying the cryptic members within the *An. barbirostris* complex in Thailand by DNA barcoding and GM.

### 5. Conclusions

This study is the first attempt to investigate the effectiveness of wing GM combined with molecular techniques to differentiate the three cryptic species of the Barbirostris complex. The current study revealed
that DNA barcoding and wing GM could distinguish three cryptic species, *An. disdens*, *An. saeungae*, and *An. wejchoochoei*, within the *Barbirostris* complex in Thailand. This information can guide the adoption of alternative methods for identifying cryptic *An. barbirostris* complex species in the field. Multiplex PCR based on species-specific COI primers is proposed to be the correct approach for use with specimens identified as members within the *An. barbirostris* complex, but DNA barcoding based on universal COI primers should be used in cases of unknown specimens caused by damaged morphology from the field collection. Conversely, the GM technique may be an effective option for use in remote areas that lack molecular biology equipment or have a limited budget. However, the application of the GM technique should be revalidated in different target locations and environments before species investigation to avoid errors from morphological variations of specimens.

**Declarations**

**Author contribution statement**

Tanawat Chaiphongpachara: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Nanta Suwandittakul, Kewarin Kuntawong, Sedthapong Laojun, Siripong Pimsuka: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

**Data availability statement**

Data included in article supplementary material/referenced in article.

**Declaration of interests statement**

The authors declare no conflict of interest.

**Additional information**

No additional information is available for this paper.

**Acknowledgements**

The authors would like to thank Mr. Teerayoot Nutepsua and Mr. Pipat piyapongdecha for mosquito collection in the field.

**References**

Apiwathanasorn, C., Prommongkol, S., Samung, Y., Limrat, D., Rojruthai, B., 2002. Potential for *Anopheles campestris* (Diptera: Culicidae) to transmit malaria parasites in Pa Rai subdistrict (Aranyaprathet, Sa Kaeo province), Thailand. J. Med. Entomol. 39 (4), 583–586.

Arnaqvist, G., Mårtensson, T., 1998. Measurement error in geometric morphometrics: empirical strategies to assess and its impact on measures of shape. Acta Zool. Acad. Sci. Hungar. 44, 73–96.

Attems, J., 1907. *Anopheles barbirostris* (Diptera: Culicidae) as a vector of the *Timor filaria* on Flores island: preliminary observations. J. Med. Entomol. 13 (4–5), 611–613.

Bookstein, F.L., 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med. Image Anal. 1 (3), 225–243.
