

# DNA Barcoding and Ecology of Myrmecophilous Lepidoptera in Endau-Rompin, Johor

Wong Wai Hong<sup>1</sup>, Yap Jing Wei<sup>1\*</sup>

<sup>1</sup> Department of Technology and Natural Resources, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hab Pendidikan Tinggi Pagoh, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA.

\*Corresponding Author: [jwyap@uthm.edu.my](mailto:jwyap@uthm.edu.my)

DOI: <https://doi.org/10.30880/ekst.2025.05.02.059>

## Article Info

Received: 7 January 2025

Accepted: 18 January 2025

Available online: 19 December 2025

## Keywords

Lepidoptera, Myrmecophilous, DNA Barcoding, Phylogenetic Analysis

## Abstract

The identification of myrmecophilous Lepidoptera, particularly the caterpillar and pupae found inside the ant nests, can be difficult to determine using the traditional morphological identification methods. DNA barcoding technique can help to solve these problems as it is able to identify the species of the specimen whether the specimen is incomplete, damaged, or in the immature stage. In the Endau-Rompin Forest Complex, the samples were manually collected. To ensure reliable species identification, both morphological techniques and DNA barcoding were used. DNA was extracted, and the cytochrome c oxidase subunit 1 (CO1) gene was amplified and sequenced. A total of five Lepidoptera specimens and two host ants were collected in this study. Phylogenetic analysis of the CO1 sequences successfully identified two Lepidoptera samples to the species level, while one was identified to the family level. One Lepidoptera sample failed to amplify during the PCR amplification process. This study had updated Endau-Rompin's checklist by identifying *Anthene emolus*, which had not before been documented in the area. Adult *Spalgis epius* was found living alongside with the ant host *Anoplolepis gracilipes* on the same plants, indicating a potential myrmecophilous association. In addition, the pupa of a species from the family Pyralidae was discovered within the ant nest of *Crematogaster* sp., suggesting a possible link between the host ant and the Pyralidae species.

## 1. Introduction

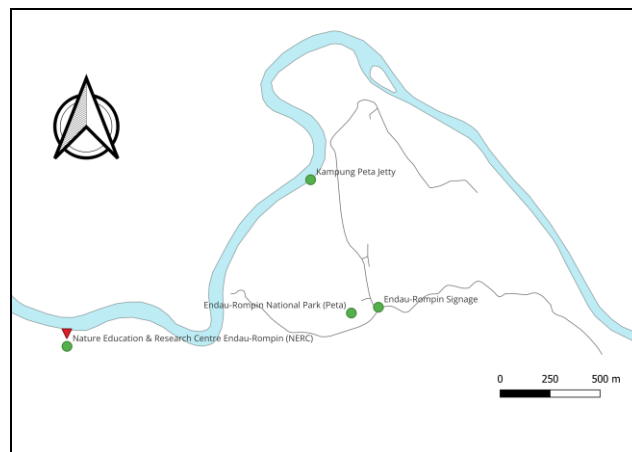
The Endau-Rompin National Park, established in 1993, is one of the largest protected areas in southern Peninsular Malaysia, encompassing 87,685 ha across Johor and Pahang [1][2]. This biodiverse region remains understudied, particularly regarding the myrmecophilous lepidoptera, which exhibit complex ecological interactions with ants [3]. Morphological identification of lepidoptera is often challenging due to incomplete or immature specimens, making traditional methods time-consuming and less effective [4]. DNA barcoding, using markers such as cytochrome oxidase 1 gene (COI) offers a reliable alternative for species identification and can expedite research efforts [5][6]. This study aims to document lepidoptera species and their interactions with ant hosts in Endau-Rompin using DNA barcoding and morphological examination. By providing insights into the ecological relationships and evolutionary significance of myrmecophily, this research will enhance conservation strategies for these butterflies and their habitats, while laying the groundwork for future studies in Johor and Peninsular Malaysia [7][8].

Organisms that depend on ants for at least a portion of their life cycle are known as myrmecophiles [9]. The Lycaenidae family is the second-largest butterfly family, with over 6,000 species and accounting for over 40% of all known butterflies. Some Lycaenidae larvae emit vibrations and low-frequency noises, which are transferred through plants and allow them to communicate with ants. While not all Lycaenidae species depend on ants, over 75% create relationships with them. Certain species of ants safeguard their larvae in return for sugar-rich honeydew, a practice known as myrmecophily [10]. Ants frequently care and protect the larvae of many myrmecophilous lycaenid species, which affect butterfly fitness during growth and developmental period. Many lycaenid larvae have specialized myrmecophilous organs that interact with ants, such as the dorsal nectary organ and tentacle organs [11]. Host ants tending the Lycaenidae are primarily from the Formicinae, Myrmicinae, and Dolichoderinae subfamilies. These interactions are most commonly observed in ants belonging to the genera *Crematogaster* and *Camponotus* [12].

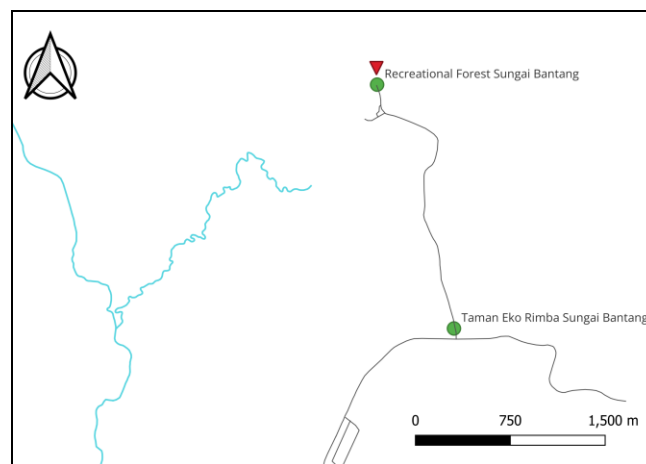
## 2. Methodology

### 2.1 Study Site

The sample collection was conducted at Endau-Rompin Peta, Johor and Nature Education and Research Centre (NERC) (N 02° 31.772, E 103° 24.079) and Taman Eko Rimba Sungai Bantang, Johor (N 02° 20.735, E 103° 9.469). Fig. 1 shows the Map of Endau-Rompin Peta, Johor and Nature Education and Research Centre (NERC) while Fig. 2 Map of Taman Eko Rimba Sungai Bantang, Johor.



**Fig. 1** Map of Endau-Rompin Peta, Johor and Nature Education and Research Centre (NERC)



**Fig. 2** Map of Taman Eko Rimba Sungai Bantang, Johor.

### 2.2 Sample collection

Butterflies displaying characteristics of the Lycaenidae family were captured manually using an aerial net. Ants associated with Lycaenidae butterflies were also collected. Ant nests believed to be associated with Lycaenidae were dug to check for the presence of pupae or larvae tended by the ants. The collected butterflies were

carefully stored in folded envelopes made from tracing paper to prevent damage to the specimens while the host ant were collected using a collection tube and filled in by 70% ethanol.

## 2.3 DNA Extraction

Tissue samples from each Lepidoptera individual were obtained by first removing the wings and then cutting the body into small pieces using a scalpel while holding it with forceps. The tissue samples were then transferred to a 1.5 mL microcentrifuge tube to prepare for the DNA extraction process.

Utilizing the DNeasy® Blood & Tissue Kit (cat. nos. 69504 and 69506) and the included manual paper, DNA was extracted from the tissue samples. The tissue samples prepared during the tissue sampling are placed in a 1.5 mL microcentrifuge tube. Add 180 µL of Buffer ATL and 20 µL of proteinase K. Mix by vortexing and incubate at 56 °C for 2 hours or until completely lysed. Vortexed occasionally during the incubation process.

200 µL of Buffer AL was added to the sample and mix it well using a vortex. The sample was then incubated at 56 °C for 10 minutes. Next, 200 µL of 100% ethanol was added, and the mixture was mixed thoroughly using a vortex. The mixture was transferred into a DNeasy Mini spin column placed in a 2 mL collection tube and centrifuged at 8000 rpm for 1 minute. After centrifugation, the flow-through and the collection tube were discarded, and the spin column was placed into a new 2 mL collection tube.

Then, 500 µL of Buffer AW1 was added to the spin column, and it was centrifuged again at 8000 rpm for 1 minute. The flow-through and the collection tube were discarded, and the spin column was moved to another new 2 mL collection tube. In the next step, 500 µL of Buffer AW2 was added to the spin column, which was centrifuged at 14,000 rpm for 3 minutes. After that, the spin column was placed into a 1.5 mL microcentrifuge tube. To elute the DNA, 50 µL of Buffer AE was added to the center of the spin column membrane. The column was left at room temperature at around 25 °C for 1 minute and then centrifuged at 8000 rpm for 1 minute to collect the DNA. The extracted DNA samples were stored in the refrigerator at a low temperature of 4 °C.

## 2.4 DNA Quantification

The Implen Nanophotometer® was used to quantify the DNA samples extracted during the DNA extraction process. Before starting, the instrument was calibrated using deionized water (diH<sub>2</sub>O) as a blank to ensure accuracy. A small volume of 1 µL of each DNA sample was placed directly onto the lens of the nanophotometer and the concentration and purity of every sample were recorded.

## 2.5 DNA Qualification

The agarose gel electrophoresis (AGE) was used to test the quality of the DNA. The agarose gel was prepared by dissolving 0.6 g of agarose powder in 40 mL of 1× TAE buffer. The mixture was heated in a microwave for 1 minute, then removed and stirred gently. It was heated again for 30 seconds to remove any precipitate. Once the gel mixture cooled to a temperature that was warm to the touch, 1.5 µL of 1<sup>st</sup> base FloroSafe DNA Stain (Catalog No: BIO5170-1n) was added, and the mixture was stirred slowly using a glass rod to avoid bubble formation. The gel was then poured into the casting mold, and a comb was placed to create wells. The mold was covered and left to set for 20 minutes; the comb was then carefully removed from the gel afterward.

The gel was placed into an electrophoresis rig filled with 1× TAE buffer. The first well was loaded with the 1st base 100 bp DNA ladder (Catalog No: BIO-5130) as a molecular size marker to determine the DNA length. The gel electrophoresis was run for 30 minutes at 90 V. After the run, the gel was visualized using the OmniDoc Gel Documentation system. The OmniDoc system was turned on, and the agarose gel was placed onto the UV transilluminator. The UV light source was selected, and adjusting to the proper UV light intensity and exposure time to produce a clear image of the gel and the DNA bands. Images of the gel were captured and being analysed. The concentration and yield of each DNA sample were determined by comparing the intensity of the sample bands to those of the DNA ladder, providing an estimation of DNA quantification. After the analysis, the agarose gel was then removed, and the tray was cleaned with 20% ethanol.

## 2.6 Polymerase Chain Reaction (PCR)

### 2.6.1 Primer

The Cytochrome Oxidase 1 (CO1) regions were used to amplify the DNA sequence, as described in Table 1.

**Table 1** The Primers used in PCR amplification and DNA sequencing of the CO1 [13]

| Targeted gene | Primer Name | Primer sequence                  |
|---------------|-------------|----------------------------------|
| CO1           | Lep-F1      | 5' - ATTCAACCAATCATAAAGATAT - 3' |
| CO1           | Lep-R1      | 5' - TAAACTTCTGGATGTCCAAAAA - 3' |

## 2.6.2 Master Mix Preparation

The recipe for a 100  $\mu\text{L}$  master mix for the primer was prepared as detailed in Table 2.

**Table 2** Master mix recipe

| Component                            | Volume ( $\mu\text{L}$ ) |
|--------------------------------------|--------------------------|
| 5x Green Gotaq buffer (Flexi Buffer) | 50                       |
| MgCl <sub>2</sub>                    | 15                       |
| dNTP                                 | 5                        |
| Lep- F1                              | 15                       |
| Lep- R1                              | 15                       |
| Total                                | 100                      |

## 2.6.3 PCR Amplification

The master mix was then combined with DNA polymerase, DNA template, and deionized water to prepare the 25  $\mu\text{L}$  PCR reaction mixture, as detailed in Table 3. The PCR amplification using the CO1 primer was carried out using the following thermocycling profile, as shown in Table 4 [13].

**Table 3** The recipe for the PCR reaction mixture.

| Component               | Volume ( $\mu\text{L}$ ) |
|-------------------------|--------------------------|
| Master Mix              | 10                       |
| DNA polymerase          | 0.126                    |
| DNA template (2B/3B/4A) | 5                        |
| Deionised water         | 9.874                    |
| Total                   | 25                       |

**Table 4** Thermocycling Profile for PCR Amplification Using Primer CO1

| Step                 | Temperature ( $^{\circ}\text{C}$ ) | Time         |
|----------------------|------------------------------------|--------------|
| 1 Cycle of           |                                    |              |
| Initial Denaturation | 94                                 | 1 min        |
| 6 Cycles of          |                                    |              |
| Denaturation         | 94                                 | 1 min        |
| Annealing            | 45                                 | 1 min 30 sec |
| Extension            | 72                                 | 1 min 15 sec |
| 36 Cycles of         |                                    |              |
| Denaturation         | 94                                 | 1 min        |
| Annealing            | 51                                 | 1 min 30 sec |
| Extension            | 72                                 | 1 min 15 sec |
| 1 Cycle of           |                                    |              |
| Final Extension      | 72                                 | 5 min        |

## 2.6.4 Nested PCR

The samples that failed to amplify during the PCR amplification and do not show distinct band on the agarose gel are then proceed with the next step, nested PCR. The initial PCR products were used as templates for the nested PCR process, following the same thermocycling profile as the initial PCR amplification. Table 5 shows the recipe for the nested PCR reaction mixture.

**Table 5** Recipe for the nested PCR reaction mixture.

| Component       | Volume ( $\mu\text{L}$ ) |
|-----------------|--------------------------|
| Master Mix      | 10                       |
| DNA polymerase  | 0.126                    |
| PCR product     | 1                        |
| Deionised water | 13.874                   |
| Total           | 25                       |

## 2.7 DNA Sequences and Analysis

DNA sequencing was conducted at Apical Scientific Sdn. Bhd. Raw electropherograms were then accessed and viewed using software such as Geneious version 9.1.8 and MEGA-X version 10.2.6. The forward and reverse sequences were edited in Geneious software to reduce overlapping peak and subsequently assembled to create a consensus sequence. The resulting sequence was analysed and compared with reference sequences available in public databases, such as the National Center for Biotechnology Information (NCBI), using the Nucleotide Basic Local Alignment Search Tool (Nucleotide BLAST). BLAST is a tool that identifies areas of similarity between biological sequences by comparing nucleotide or protein sequences to sequence databases and determining their statistical significance. All the nucleotide sequences were transcribed in the FASTA format from NCBI and imported into the Geneious software. Subsequently, multiple sequence alignment was performed in Geneious software using the MUSCLE alignment, the Geneious guidelines describe MUSCLE alignment as a progressive aligner that estimates sequence distance through k-mer counting, performs progressive alignment using a log-expectation score profile function, and refines alignments with tree-dependent restricted sequence partitioning. The aligned sequences are then exported in MEG file format to make it compatible with MEGA-X software for subsequent phylogenetic analysis.

The MEG file was then opened using MEGA-X software, and phylogenetic tree was constructed using the Maximum Likelihood and Maximum Parsimony method. The Maximum Likelihood tree was generated with the following settings: the substitution model used was the General Time Reversible (GTR) model with rates among sites modelled as Gamma Distributed with Invariant Sites (GTR+G+I) with a bootstrap of 10,000 replicates was performed to ensure the robustness of the tree. The Maximum Parsimony (MP) tree was constructed using Tree-Bisection-Reconnection (TBR) as the MP search methods for the tree inference option and with a bootstrap 10,000 replicates.

## 3. Results and Discussion

A total of five Lepidoptera samples were collected during the study. Two samples were obtained from Nature Education & Research Centre (NERC), Endau-Rompin Johor National Park and three were collected from Taman Eko Rimba Sungai Bantang. One of the Lepidoptera samples was a pupa found inside an ant nest. The pupa was stored in a container to allow it to undergo metamorphosis into its adult stage, making it easier to identify through morphological analysis.

In addition, two host ant samples were collected: one from Kampung Peta NERC and another from Sungai Bantang. These ants were collected as host samples because they showed ecological relationships with the Lepidoptera species. The table 6 below provides detailed information on the collected samples, including their names, locations, and the method of identification.

**Table 6** Information of the collected samples

| Sample Name | Location Collected | Date Collected  | Notes            | Identification Method           |
|-------------|--------------------|-----------------|------------------|---------------------------------|
| 1A          | Sungai Bantang     | 7 April 2024    | Lepidoptera Pupa | Morphological and DNA Barcoding |
| A1          | Sungai Bantang     | 7 April 2024    | Host Ant         | Morphological                   |
| 4A          | Kampung Peta, NERC | 29 April 2024   | Lycaenidae       | DNA Barcoding                   |
| 4B          | Kampung Peta, NERC | 29 April 2024   | Lycaenidae       | Morphological                   |
| A2          | Kampung Peta, NERC | 29 April 2024   | Host Ant         | Morphological                   |
| 2B          | Sungai Bantang     | 6 November 2024 | Lycaenidae       | Morphological and DNA Barcoding |

|    |                |                 |            |                                 |
|----|----------------|-----------------|------------|---------------------------------|
| 3B | Sungai Bantang | 6 November 2024 | Lycaenidae | Morphological and DNA Barcoding |
|----|----------------|-----------------|------------|---------------------------------|

### 3.1 Morphological Identification







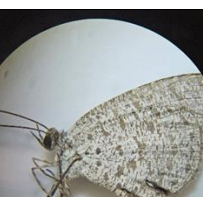

The morphological identification for the collected samples during the study was performed using a microscope at the Makmal Gunasama Fizik (MGF Lab). *Google Lens* and *Picture Insect* were primarily used for preliminary identification, with their results cross-checked against morphological keys and expert guides for accuracy. Additionally, websites such as MyBIS and iNaturalist were consulted to confirm the species' distribution and their presence in Malaysia.

#### 3.1.1 Lepidoptera

Identification of the Lepidoptera samples were based on keys by [14]. Table 7 presents the collected samples of Lepidoptera, along with their corresponding species names. The pupa sample 1A was successfully developed into an adult moth. However, due to improper preservation, the specimen was severely damaged, making it difficult to identify morphologically. Sample 2B was identified with the aid of *Picture Insect* shown as *Anthene emolus*. The Sample 2B displayed dark marginal spots in spaces 1b, 2, and 3 on the hindwing. The underside was pale and earthy brown, with catenulate bands defined by whitish stripes. Additionally, the hindwing near space 2 featured an orange-crowned black spot, matching the description of *Anthene emolus*.

Sample 3B was identified as *Prosotas dubiosa* using *Picture Insect*. According to the referral guidebook, *Prosotas dubiosa* featured a small black tornal patch at the end of vein 1b. The species had a greyish underside and no apparent subapical mark in place 6 on the hindwings. The species is tailless which matches the description of the Sample 3B. Sample 4B was identified using *Google Lens* and *Picture Insect* as *Spalgis epius*. Reference to the guidebook, the morphological features of *Spalgis epius*, including smooth eyes, short antennae with a gradually incrassate club. The larvae feed on scale insects, bodies are coated white waxy substances secreted by the scale insects, providing protection against ants. Sample 4B was found on a tree infested with scale insects, cohabiting with ants. This ecological context further supported the identification.

**Table 7** Collected samples of Lepidoptera, along with their corresponding species names.



| Sample | Photo   |   | Species                 |
|--------|---|---|-------------------------|
| 1A     |  |  | -                       |
| 2B     |  |  | <i>Anthene emolus</i>   |
| 3B     |  |  | <i>Prosotas dubiosa</i> |
| 4B     |  |  | <i>Spalgis epius</i>    |

### 3.1.2 Host Ant Species

The host ant samples were examined under a microscope at the MGF laboratory. Identification was performed based on the keys by [15]. Table 8 provides details of the collected host ant samples, including pictures, species names, and observational notes recorded during their collection, focusing on their interactions with butterfly species.

Referring to the *Identification Guide to Ant Genera of Borneo*, the key to Bornean genera of the subfamily Myrmicinae states that a postpetiole attached to the upper surface of the gaster indicates the genus *Crematogaster*. The postpetiole's dorsal connection to the fourth abdominal segment unique morphological feature of the genus *Crematogaster* [16]. These two characteristics suggest that the sample host ant 1 belongs to the genus *Crematogaster* sp. *Anoplolepis gracilipes* is easily identified by its large legs, very long antennal scapes, and distinctive yellowish coloration [17]. These characteristics are similar to those of host ant 2. It is then cross check with the Picture Insect software, further supporting the species as *Anoplolepis gracilipes*.

**Table 8** Collected samples of host ant, along with their corresponding species names.

| Sample Name | Photo  | Species Name                  | Observation Notes.  |
|-------------|--|-------------------------------|---|
| Host Ant 1  |   | <i>Crematogaster</i> sp.      | The pupa sample 1A was discovered within the nest of Host Ant 1 |
| Host Ant 2  |  | <i>Anoplolepis gracilipes</i> | Ants were observed living with sample 4B, along with cocoidea   |

### 3.2 DNA Extraction Results

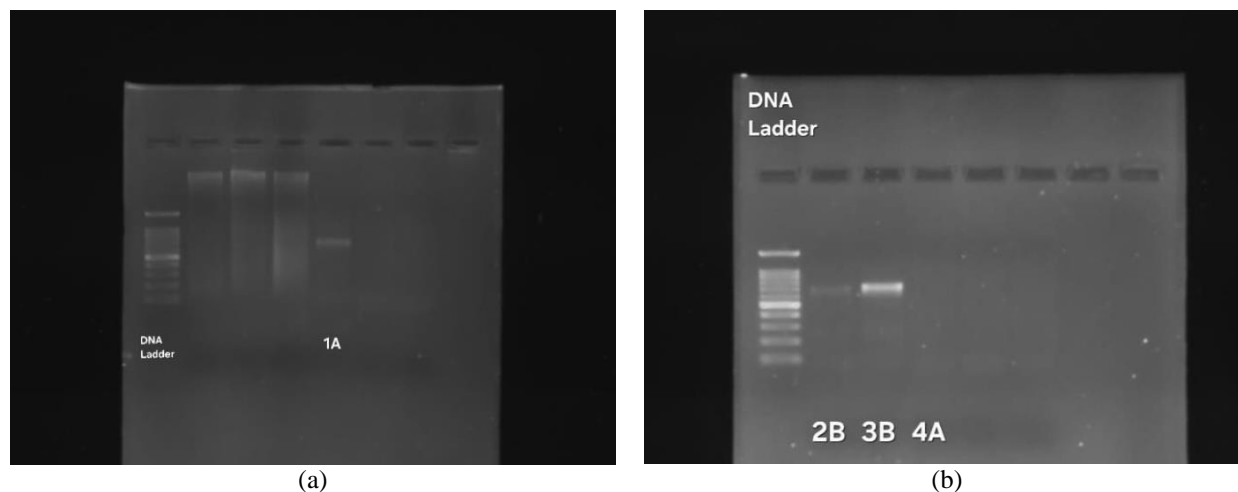
The Implen Nanophotometer® was used to quantify the DNA samples extracted during the DNA extraction process with the result recorded in Table 9. To determine samples purity, the A260/A280 and A260/A230 ratios were compared together, with an ideal A260/A280 value of 1.8 and an expected A260/A230 range of 2.0-2.2. Sample 1A had an A260/A280 ratio of 1.475, which is much lower than the optimal value and the A260/A230 ratio of 1.047, was much lower than the predicted range, indicating possible contamination. Sample 2B had an A260/A280 ratio of 1.747, which is near to the ideal value, but the A260/A230 ratio of 0.680 was lower than predicted, indicating the presence of impurities. Similarly, sample 3B had an A260/A280 ratio of 1.734, but the A260/A230 ratio of 1.476, indicating contamination. Sample 4A had an A260/A280 ratio of 1.714, but its A260/A230 ratio was lower than predicted at 0.691, indicating possible contamination. The agarose gel electrophoresis includes DNA ladder, samples 1A, 2B, 3B and 4B. Bands are observed in the agarose gel, indicating the presence of genomic DNA. The intensity of the bands suggests sufficient DNA quantity for downstream applications, such as PCR.

**Table 9** The reading of Nanodrop spectrophotometry

| Sample | Concentration of DNA (ng/μl) | A260/A280 | A260/A230 | A260 (10mm path) |
|--------|------------------------------|-----------|-----------|------------------|
| 1A     | 4.5000                       | 1.475     | 1.047     | 0.084            |
| 2B     | 6.9000                       | 1.747     | 0.680     | 0.133            |
| 3B     | 12.4000                      | 1.734     | 1.476     | 0.243            |
| 4A     | 6.6000                       | 1.714     | 0.691     | 0.132            |

### 3.3 PCR Results

In the gel electrophoresis process, a 1st Base 100 bp DNA Ladder (Catalog No: BIO-5130) was used, providing size markers ranging from 100 bp to 1500 bp. Fig. 3(a) and Fig. 3(b) show the PCR amplification results that were obtained by agarose gel electrophoresis. Samples 2B, 3B and 4A were nested PCR products. Sample 3B exhibited multiple bands around the 700 bp region rather than a single, well-defined band at the target DNA regions, COI at 648bp [13]. This may be due to nonspecific amplification, which is a common occurrence in nested PCR [18]. Sample 2B also showed multiple bands, but they were faint, suggesting low DNA concentration and nonspecific amplification. Although the band for sample 1A is less obvious and pale, it still has the band at the expected size at around 700bp, which means the success of DNA amplification at the region COI. However, sample 4A did not show any band after several attempts.



**Fig. 3** PCR product results (a) Sample 1A; (b) Sample 2B, 3B and 4A

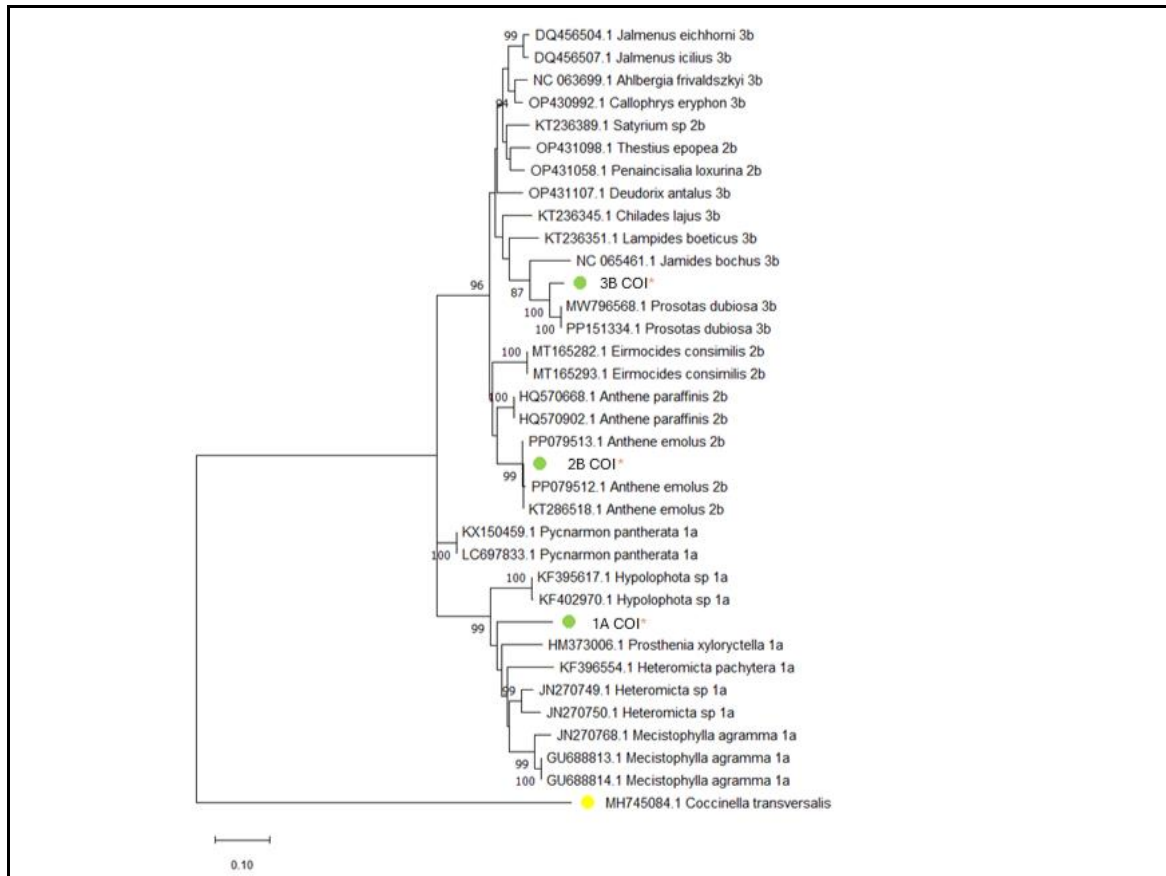
### 3.4 DNA Sequencing

All the selected PCR Products are selected from the gel electrophoresis result. The PCR products were then outsourced to Apical Scientific Sdn. Bhd. for DNA sequencing. A total of 3 Sequence was produced from the sequencing using the COI region. The sequences of 1A, 2B, 3B were analyzed and assembled using the De Novo Assembly in the Geneious v9.1.8 software to produce consensus sequences. The quality of the sequences that were produced from the sequencing were good, with clear distinct peak only little overlapping. The overlapping peaks at the beginning and the end of the sequence were trimmed using the Geneious software.

### 3.5 Phylogenetic Analysis

In the Fig. 4, bootstrap value lower than 70 percent was hidden, the sample 1A, 2B and 3B were highlighted in green and outgroup, *Coccinella transversalis* highlighted in yellow. The COI Maximum Likelihood phylogenetic tree shows the species 3B and 2B were found resolving with *Prosotas dubiosa* (MW796568.1 and PP151334.1) and *Athene emolus* (PP079512.1 and KT286518.1) respectively with the support bootstrap of 10000 replicates. However, the exact identity of the species 1A is unclear, it resolves on a clade containing *Hypolophota* sp (KF395617.1 and KF402970.1) and *Prosthenia xyloryctella* (HM373006.1), thus indicating that it belongs to a species in the Pyrrallidae family. *Athene emolus* has not been previously recorded from Endau-Rompin [19].





**Fig. 4** The evolutionary history was inferred by using the Maximum Likelihood method and General Time Reversible model.

The tree with the highest log likelihood (-6380.31) is shown in Fig. 4. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.3649)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 36.86% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 35 nucleotide sequences. There were a total of 1495 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.

### 3.6 Discussion

In discussion, this study demonstrated the successful application of DNA barcoding alongside morphological techniques in identifying myrmecophilous Lepidoptera and revealing its ecological interaction with ant hosts. The ecological interactions discovered in this study provide insight on the complex relationships between myrmecophilous Lepidoptera and their related host ant species. The discovery of *Athene emolus*, a new record for Endau-Rompin, broadens the species' known in the region and suggests that the region might contain a greater diversity of Lycaenidae than previously documented. The identification of a Pyralidae pupa within a *Crematogaster* sp. nest provides evidence that certain Pyralidae species exhibit myrmecophilous interactions with the ant genera *Crematogaster*. Furthermore, the ecological interaction between a Lycaenid butterfly and its host ant is highlighted by the myrmecophilous association between Lycaenid *Spalgis epius* and its host ant, *Anoplolepis gracilipes*.

Throughout this study, a number of unexpected findings were discovered, especially in the DNA amplification and identification processes. Instead of a single, distinct band at the expected COI target region of 648 bp, two of the samples showed several bands around the 700 bp region. This is probably because of nested PCR, which amplified unwanted DNA fragments as well as the results of nonspecific amplification. Furthermore, one sample failed to amplify in this study, which could have been caused by DNA degradation from prolonged storage. Successful amplification depends on the quality of the DNA, unsuccessful PCR results are frequently due

to damaged DNA. Another challenge was that the collected pupa could only be identified until the family level. The limitation may have been caused by an insufficient reference library in the DNA database, which would have prevented an accurate species match, or by DNA degradation, which decreases the accuracy of sequencing.

#### 4. Conclusion

This study has successfully used DNA barcoding in combination with morphological techniques to identify myrmecophilous Lepidoptera and uncover new information on their ant hosts. Three myrmecophilous Lepidoptera to the species level, including *Athene emolus*, which represents a new record for Endau-Rompin. One pupa from *Crematogaster* nest was also identified as a member of Pyralidae. In addition, this study reports documented *Spalgis epius* associating with *Anoplolepis gracilipes*. These findings have expanded our knowledge on the diversity and ecology of myrmecophilous Lepidoptera and their ant hosts in Peninsular Malaysia and demonstrated that CO1 can be successfully used for aiding the identification of a morphologically challenging group.

There were challenges with morphological identification in this study. Sample 1A was difficult to identify due to improper preservation causing damage to the sample. To ensure that the morphological identity can be observed for the purpose of morphological identification, taking photos from all directions after the sample was collected or to preserve the collected sample in a proper method. Because certain species are only active during particular seasons and host-ant interactions change depending on the stage of the lepidoptera life cycle, season-dependent sampling presents a challenge. Researchers can address this by carrying out long-term studies that sample at various times during the year. Working along with local ecologists who are knowledgeable about the variations in species activity.

#### Acknowledgement

The authors would like to thank to Dr. Aqilah, providing knowledge and information in identify host ant species and butterflies. Thanks also go to the staff of UTHM for their assistance and support during the field trips to Endau-Rompin. The Forestry Department of Johor's contributions throughout this study is deeply appreciated, especially in obtaining permits for conducting fieldwork in the area. Special thanks go to Jeremiah Sia Yiao Rong for his help with laboratory work, such as DNA extraction and PCR amplification, as well as for offering helpful direction throughout these procedures. Gratitude is also extended to Liyana Nasir for her assistance with sample collection and to Calvin Tsz Kin Leung for his support in the identification of host ant samples.

#### Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of the paper.

#### Author Contribution

The authors confirm their contribution to the paper as follows: **study conception and design:** Wong Wai Hong, Yap Jing Wei; **data collection:** Wong Wai Hong; **analysis and interpretation of results:** Wong Wai Hong, Yap Jing Wei; **draft manuscript preparation:** Wong Wai Hong, Yap Jing Wei. All authors reviewed the results and approved the final version of the manuscript.

#### References

- [1] Kiew, G.W.H. Davison, & R. Kiew. (1986). The Malaysian Heritage and Scientific Expedition, Endau-Rompin 1985. *Malayan Naturalist*, 39(3-4), 3-58. <https://eurekamag.com/research/021/877/021877142.php>
- [2] Department of Wildlife and National Park. (1996). Sumantran riitnoceros in endau-rompin, malaysia: their plight and fate. Department of Wildlife and National Park.
- [3] Pierce, N. E., Braby, M. F., Heath, A., Lohman, D. J., Mathew, J., Rand, D. B., & Travassos, M. A. (2002). The Ecology and Evolution of Ant Association in the Lycaenidae (Lepidoptera). *Annual Review of Entomology*, 47(1), 733-771. <https://doi.org/10.1146/annurev.ento.47.091201.145257>
- [4] Elias, M., Hill, R. I., Willmott, K. R., Dasmahapatra, K. K., Brower, A. V. Z., Mallet, J., & Jiggins, C. D. (2007). Limited performance of DNA barcoding in a diverse community of tropical butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 274(1627), 2881-2889. <https://doi.org/10.1098/rspb.2007.1035>
- [5] Meyer Christopher P, & Paulay Gustav. (2005). DNA Barcoding: Error Rates Based on Comprehensive Sampling. *PLoS Biology*, 3(12), e422. <https://doi.org/10.1371/journal.pbio.0030422>
- [6] Gong, L., Zhang, D., Ding, X., Huang, J., Guan, W., Qiu, X., & Huang, Z. (2021). DNA barcode reference library construction and genetic diversity and structure analysis of *Amomum villosum* Lour. (Zingiberaceae) populations in Guangdong Province. *PeerJ*, 9. <https://doi.org/10.7717/peerj.12325>

- [7] Carvalho, A. M., & Frazão-Moreira, A. (2011). Importance of local knowledge in plant resources management and conservation in two protected areas from Trás-os-Montes, Portugal. *Journal of Ethnobiology and Ethnomedicine*, 7(1), 36. <https://doi.org/10.1186/1746-4269-7-36>
- [8] Hoffmann, S. (2021). Challenges and opportunities of area-based conservation in reaching biodiversity and sustainability goals. *Biodiversity and Conservation*, 31. <https://doi.org/10.1007/4b0531-021-02340-2>
- [9] Kronauer, D. J., & Pierce, N. E. (2011). Myrmecophiles. *Current biology : CB*, 21(6), R208–R209. <https://doi.org/10.1016/j.cub.2011.01.050>
- [10] Khyade, V. B., Pawar, R. S., & Khilare, A. D. (2018). Description of Lycaenidae Butterflies. *International Academic Journal of Innovative Research*, 05(02), 9–25. <https://doi.org/10.9756/iajir/v5i1/1810012>
- [11] Mizuno, T., Hagiwara, Y., & Akino, T. (2019). Varied effects of tending ant species on the development of facultatively myrmecophilous lycaenid butterfly larvae. *Insects*, 10(8), 234. <https://doi.org/10.3390/insects10080234>
- [12] Fiedler, K. (2021). The ant associates of Lycaenidae Butterfly Caterpillars – Revisited. *Nota Lepidopterologica*, 44, 159–174. <https://doi.org/10.3897/nl.44.68993>
- [13] Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrartes fulgurator*. *Proceedings of the National Academy of Sciences of the United States of America*, 101(41), 14812–14817. <https://doi.org/10.1073/pnas.0406166101>
- [14] Corbet, A. S., Pendlebury, H. M., Van der Poorten, G. M., & Van der Poorten, N. E. (2020). The butterflies of the Malay Peninsula. Malaysian Nature Society.
- [15] Hashimoto, Y. (2003). Identification guide to ant genera of Borneo.
- [16] Blaimer, B. (2010). Taxonomy and natural history of the *Crematogaster* (Decacrema)-group (Hymenoptera: Formicidae) in Madagascar. *Zootaxa*, 2714(1), 1–39. <https://doi.org/10.11646/zootaxa.2714.1.1>
- [17] Nur Zati Akma Mustafa & Ong Su Ping (2017, April). *Anoplolepis gracilipes*. Malaysia Biodiversity Information System (MyBIS). Retrieved December 30, 2024, from <https://www.mybis.gov.my/art/171>.
- [18] AAT Bioquest. (n.d.). Advantages & limits of nested PCR vs. standard PCR. AAT Bioquest. Retrieved January 25, 2025, from <https://www.aatbio.com/resources/application-notes/advantages-limits-of-nested-pcr-vs-standard-pcr>
- [19] Awg Abdul Rahman, A., & Mohamed, M. (2019). Checklist of butterflies (Lepidoptera: Papilionoidea) of Sg. Bantang, Labis Forest Reserve, Johor. *IOP Conference Series: Earth and Environmental Science*, 269(1), 012042. <https://doi.org/10.1088/1755-1315/269/1/012042>