Land snails of *Leptopoma* Pfeiffer, 1847 in Sabah, Northern Borneo (Caenogastropoda: Cyclophoridae): an analysis of molecular phylogeny and variations in shell form due to geography

Chee-Chean Phung [1] , **Pooi-San Heng** [1] , **Thor-Seng Liew** [Corresp. 1]

[1] Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia

Corresponding Author: Thor-Seng Liew
Email address: thorsengliew@gmail.com

Leptopoma is a species rich genus with approximately 100 species documented according to shell morphology and animal anatomy. Many of the *Leptopoma* species are described in terms of shell size, shape, sculpture and colour patterns of a small number of examined materials. However, the implications of the inter- and intra-species variations in shell form to the taxonomy of *Leptopoma* species and the congruency of its current shell based taxonomy with its molecular phylogeny are still unclear. Over the last decade, more than 900 collection lots consisting of more than 4000 *Leptopoma* specimens have been obtained in Sabah and deposited in *BORNEENSIS* at Universiti Malaysia Sabah. Access to this collection gave us the opportunity to examine the geographical variations in shell forms and the phylogenetic relationship of *Leptopoma* species in Sabah. The phylogenetic relationship of three *Leptopoma* species was first estimated by performing maximum likelihood and Bayesian analysis based on mitochondrial genes (16S and COI) and nuclear gene (ITS-1). After this, a total of six quantitative shell characters (i.e. shell height, shell width, aperture height, aperture width, shell spire height, and ratio of shell height and width) and three qualitative shell characters (i.e. shell colour patterns, spiral ridges, and dark ring band in aperture) of the specimens were mapped across the phylogenetic tree and tested for phylogenetic signals. Data on shell characters of *Leptopoma sericatum* and *Leptopoma pellucidum* from two different locations (i.e. Balambangan Island and Kinabatangan) where both species occurred sympatrically were then obtained to examine the geographical variations in shell form. The molecular phylogenetic analyses suggested that each of the three *Leptopoma* species was monophyletic and indicated congruence with one of the shell characters (i.e. shell spiral ridges) in the current morphological-based classification. Other qualitative and quantitative shell characters were incongruent with the *Leptopoma* species phylogeny. Although the geographical variation analyses suggested some of the shell characters indicating inter-species differences between the

two *Leptopoma* species, these also pointed to intra-species differences between populations from different locations. This study provides an initiation to resolve the taxonomy conundrum for the remaining 100 little known *Leptopoma* species from other regions and highlights a need to assess variations in shell characters before they could be used in species classification.

1 **Land snails of *Leptopoma* Pfeiffer, 1847 in Sabah, Northern Borneo (Caenogastropoda:**
2 **Cyclophoridae): an analysis of molecular phylogeny and variations in shell form due to**
3 **geography**

4 Chee-Chean Phung, Pooi-San Heng, Thor-Seng Liew

5 *Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Jalan UMS, 88400*
6 *Kota Kinabalu, Malaysia.*

7 Corresponding author: Thor-Seng Liew (thorsengliew@gmail.com)

8 **Abstract.** *Leptopoma* is a species rich genus with approximately 100 species documented
9 according to shell morphology and animal anatomy. Many of the *Leptopoma* species are described
10 in terms of shell size, shape, sculpture and colour patterns of a small number of examined
11 materials. However, the implications of the inter- and intra-species variations in shell form to the
12 taxonomy of *Leptopoma* species and the congruency of its current shell based taxonomy with its
13 molecular phylogeny are still unclear. Over the last decade, more than 900 collection lots
14 consisting of more than 4000 *Leptopoma* specimens have been obtained in Sabah and deposited in
15 *BORNEENSIS* at Universiti Malaysia Sabah. Access to this collection gave us the opportunity to
16 examine the geographical variations in shell forms and the phylogenetic relationship of *Leptopoma*
17 species in Sabah. The phylogenetic relationship of three *Leptopoma* species was first estimated by
18 performing maximum likelihood and Bayesian analysis based on mitochondrial genes (16S and
19 COI) and nuclear gene (ITS-1). After this, a total of six quantitative shell characters (i.e. shell
20 height, shell width, aperture height, aperture width, shell spire height, and ratio of shell height and
21 width) and three qualitative shell characters (i.e. shell colour patterns, spiral ridges, and dark ring
22 band in aperture) of the specimens were mapped across the phylogenetic tree and tested for
23 phylogenetic signals. Data on shell characters of *Leptopoma sericatum* and *Leptopoma pellucidum*
24 from two different locations (i.e. Balambangan Island and Kinabatangan) where both species
25 occurred sympatrically were then obtained to examine the geographical variations in shell form.
26 The molecular phylogenetic analyses suggested that each of the three *Leptopoma* species was
27 monophyletic and indicated congruence with one of the shell characters (i.e. shell spiral ridges) in
28 the current morphological-based classification. Other qualitative and quantitative shell characters
29 were incongruent with the *Leptopoma* species phylogeny. Although the geographical variation
30 analyses suggested some of the shell characters indicating inter-species differences between the
31 two *Leptopoma* species, these also pointed to intra-species differences between populations from
32 different locations. This study provides an initiation to resolve the taxonomy conundrum for the
33 remaining 100 little known *Leptopoma* species from other regions and highlights a need to assess
34 variations in shell characters before they could be used in species classification.
35

36 **Keywords:** Borneo, Cyclophoridae, integrative taxonomy, Malaysia, phylogenetic signals, Sabah,
37 shell morphology variations.

38

INTRODUCTION

39
40 The terrestrial snail genus *Leptopoma* is one of 35 genera in the Cyclophoridae family
41 (Kobelt, 1902) with a wide distribution range that covers Oriental and Australasia zoogeographical
42 regions. An early overview of worldwide *Leptopoma* species classified the genus *Leptopoma* into
43 four subgenera with a total of 105 species (Kobelt, 1902). Several subsequent regional taxonomic
44 reviews of *Leptopoma* were conducted in the Philippines (Zilch, 1956), Ceylon and Burma (Gude,
45 1921), and most recently in Borneo (Vermeulen, 1999). To date, taxonomical works on *Leptopoma*
46 (Kobelt, 1902; Gude, 1921; Zilch, 1954; Vermeulen, 1999) have been mainly based on shell
47 morphology. Besides shell morphology, other anatomical characteristics of the soft body such as
48 radula, operculum, and genital duct have been used limitedly in the species delimitation (Sarasin
49 & Sarasin, 1899; Jonges, 1980). The phylogenetic relationship of *Leptopoma* species *per se* is not
50 known although several species were included in phylogenetic studies of other taxa as outgroup
51 (Colgan *et al*., 2000, 2003, 2007; Lee *et al*. 2008a, 2008b; Nantarat *et al*., 2014a).

52

53 The genus *Leptopoma* is abundant in the Philippines and the adjacent Malaysian state of
54 Sabah located at the northern part of Borneo Island (Godwin-Austen, 1891; Laidlaw, 1937;
55 Vermeulen, 1999; Schilthuizen & Rutjes, 2001; Uchidal *et al*., 2013). Currently, four *Leptopoma*
56 species could be identified from the specimens collected in Sabah with *Leptopoma undatum*
57 (Metcalfe, 1851) distinguished by its uniformed shell colour (translucent when young and white
58 when old) and shell shape (less convex whorl and sharp keep at the last whorl). The other three
59 species – *Leptopoma atricapillum* (Sowerby, 1843), *Leptopoma sericatum* (Pfeiffer, 1851) and
60 *Leptopoma pellucidum* (Grateloup, 1840) are very similar in terms of shell shape with all showing
61 and sharing colour pattern polymorphism. The *Leptopoma sericatum* (Pfeiffer, 1851) differs from
62 *Leptopoma pellucidum* (Grateloup, 1840) due to the presence of stronger spiral thread-like ridges
63 (Vermeulen, 1999). *Leptopoma atricapillum* (Sowerby, 1843) - not included in Vermeulen (1999)
64 - has more pronounced spiral ridges and are hence more easily identified.

65

66 Vermeulen (1999) identified two major challenges when using shell characters as
67 diagnostic indicators for the Borneon *Leptopoma* species. Firstly, the majority of species were
68 similar in shell form thus limiting the number of shell characters that could be used as diagnostic
69 indicators at species level. This problem was noted in Vermeulen's examination of six Bornean
70 species. There is no doubt that this problem would become even more pronounced when examining
71 the other *ca*. 100 species. Secondly, there are intermediate shell forms between *Leptopoma* species
72 which could cause uncertainties in species delimitation. Thus it is clear that to date, the
73 implications of the intra- and inter-species variations in shell form, in terms of shape, size and
74 colour patterns in the taxonomy of *Leptopoma* species have not been studied systematically and
75 comprehensively.

76

77   Hence this study was conducted specifically to (1) estimate the molecular phylogeny of
78   three similar yet polymorphic *Leptopoma* species in Sabah in order to investigate the monophyly
79   of *L. sericatum*, *L. pellucidum* and *L. atricapillum* based on two mitochondrial genes (16S RNA
80   and COI) and a nuclear gene (ITS-1), (2) test the phylogenetic signal of the shell morphological
81   characters in terms of three qualitative shell characters and six quantitative shell measurements
82   across the phylogenetic trees in order to evaluate their reliability as diagnostic characters, and (3)
83   compare the differences in shell characters of two *Leptopoma* species namely *L. sericatum* and *L.*
84   *pellucidum* at two locations where they were abundant and found sympatrically in order to
85   understand the geographical variations in shell form and further assess their reliability as
86   diagnostic characters. The results of this study supported the monophyly of the three *Leptopoma*
87   species in line with the current classifications by Vermeulen (1999) although only the qualitative
88   shell character such as spiral ridges was reliable for species delimitation.
89
90   **MATERIALS AND METHODS**
91   All the *Leptopoma* specimens included in this study were obtained from the BORNEENSIS
92   mollusca collection at the Institute of Tropical Biology and Conservation in Universiti Malaysia
93   Sabah. The collection housed more than 4000 specimens of *Leptopoma* spp. collected since 2000
94   from various locations in Sabah (Fig. 1). From this comprehensive collection, 77 wet specimens
95   of four species (*L. sericatum*, *L. pellucidum*, *L. atricapillum*, *L.undatum*) were selected for
96   molecular analysis. 249 empty shells of adult snails of *L. sericatum* (114) and *L. pellucidum* (135)
97   from Balambangan Island and the Kinabatangan region, where both species existed sympatrically
98   were selected for morphological analysis (Supplemetary File 1). These *Leptopoma* specimens were
99   identified into either *L. pellucidum* or *L. sericatum* based on the spiral ridges on the shell
100   (Vermeulen, 1999).
101
102
103   **Data Collection**
104
105   *Genetic Data*
106   Genomic DNA of 77 selected specimens stored in 70% ethanol was isolated from foot tissue by
107   using DNeasy extraction kit (Qiagen Inc., Hilden, Germany) according to manufacturer
108   instructions. Universal primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and
109   HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') were used to amplify and sequence
110   mitochondrial cytochrome c oxidase subunit 1 (COI) (Folmer *et al.*, 1994). 16s rRNA
111   mitochondrial gene was amplified using primers 16Sar (5'-CGCCTGTTTATCAAAAACAT-3')
112   and 16Sbr (5'-CCGGTCTGAACTCAGATCACGT-3') (Kessing *et al.*, 1989). ITS-1 region was
113   PCR-amplified using the primers 5.8c (5'-GTGCGTTCGAAATGTCGATGTTCAA-3') and 18d
114   (5'-CACACCGCCCGTCGCTACTACCGATTG-3') (Hillis & Dixon, 1991). Thermal cycling
115   was performed with pre-denaturation at 90°C for 2 minutes, denaturation at 94°C for 45 seconds,
116   one minute of annealing at 55°C, 60°C, and 54°C for COI, 16s, and ITS-1 respectively, extension

117 step at 72°C for one minute followed by final extension at 72°C for 5 minutes. Denaturation,
118 annealing and extension steps were repeated for 35 cycles. Positive PCR results were obtained
119 from 17 out of 77 DNA extracts for at least two genes (Table S1 in Supplementary File 2). The
120 PCR products were sequenced at Macrogen, Inc. (Korea). All sequences were subsequently
121 uploaded and stored in Barcoding of Life Database (BOLD, http://www.boldsystems.org,
122 Ratnasingham & Hebert, 2007), under the project title "*Leptopoma* in Sabah" (Code: LEPT).
123
124 *Shell Morphological Characters Data*
125 Quantitative and qualitative shell characters were obtained from the shell aperture view of all 249
126 *Leptopoma* specimens and 14 adult specimens included in the phylogenetic analysis. First, high
127 quality photographs were taken of the aperture of each shell with the aid of a Leica Stereo
128 Microscope M205. Five quantitative linear measurements, namely shell height (SH), shell width
129 (SW), aperture height (AH), aperture width (AW), and shell spire height (SpH) were then taken
130 directly from the photographs by using Leica Application Suite software (Fig. 2A). The sixth
131 quantitative shell character – the ratio between shell height and width – was computed accordingly.
132 Next, the states for the two qualitative shell characters (i.e. the eight types of shell colour patterns
133 and presence of the dark ring band in aperture) were recorded for each of the shells (Figs. 2C, 2D;
134 see Table S2 in Supplementary File 2 for the descriptions of the eight shell colour patterns).
135
136 **Data Analysis**
137
138 *Molecular Phylogenetic Analysis*
139 In addition to the sequences collected from 17 specimens in this study, 16S and COI sequences of
140 *L. tigris*, *L. vitreum* and an outgroup species - *Cyclophorus formosensis* from Lee et al. (2008a)
141 and Nantarat et al. (2014a) - were obtained from GenBank (File S1, Page 1: Table S1 for
142 informations of specimens). All the DNA sequences were aligned and checked manually using
143 Bioedit v7.1.9 (Hall, 1999). In order to find the best-fit model of substitution, jModelTest2
144 (Darriba *et al*., 2012) as implemented in CIPRES portal (Miller *et al*., 2010) was performed based
145 on corrected Akaike Information Criterion (AICc) for ITS-1 sequences, 16S sequences and each
146 of the codon positions of COI sequences. Phylogenetic trees were estimated by using Maximum
147 likelihood (ML) and Bayesian Inference methods (BI) as implemented in CIPRES portal (Miller
148 *et al*., 2010). Maximum likelihood analysis was conducted using Raxml-HPC2 (Stamatakis, 2014)
149 with 100 rapid bootstraps. Bayesian Inference analysis was performed using MrBayes v3.2.3
150 (Huelsen & Ronquist, 2001) which consisted of running four simultaneous chains for 100,000
151 generations and 10 sampling frequency. The first 250 trees were discarded as burn-in, while the
152 rest were used to obtain the final consensus tree.
153
154
155 *Phylogenetic Signal Analysis*

156 Phylogenetic signal analysis was applied to investigate the congruence between phylogeny and
157 morphology with all the analyses done in R statistical environment version 3.1.3 (R Core Team,
158 2015). The tips of juvenile specimens and outgroup taxa in the phylogenetic tree were excluded
159 by using package 'ape' (Paradis *et al*., 2004). The final tree for phylogenetic signal analysis
160 consisted of 14 adults of three *Leptopoma* species. The six quantitative and three qualitative shell
161 characters were mapped onto the tree by utilising package 'phytools'. Phylogenetic signals for
162 each of these nine shell characters were examined using maximum likelihood (λ) (Pagel, 1999)
163 and K (Blomberg *et al*., 2003). The consensus tree was transformed into an ultrametric tree after
164 which a lambda analysis was performed using the 'chronopl' function from the 'ape' package
165 (Paradis *et al*., 2004). As a result, a chronogram was generated using penalised likelihood with an
166 arbitrary lambda value of 0.1, the alternative model. A null model, the *Leptopoma* phylogenetic
167 tree with $\lambda = 0$ (no phylogenetic signal), was generated using the 'rescale' function from the 'geiger'
168 package (Harmon *et al*., 2008). The λ value of each shell character was estimated for both models
169 using the 'fitDiscrete' function for three qualitative shell characters and 'fitContinuous' function for
170 six quantitative shell characters in the 'geiger' package (Harmon *et al*., 2008). Likelihood scores
171 for the alternative and null models were compared by performing a likelihood ratio test in order to
172 examine the phylogenetic signal in each shell character, wherein Blomberg's K was calculated
173 using the 'physig' function from the 'phytool' package (Revell, 2012; R script in Supplementary
174 File 3).

175

176 *Geographical Variation in Shell Morphology Analysis*
177 Two-way ANOVA tests were performed to determine differences in the six quantitative shell
178 characters between (i) the two *Leptopoma* species (*L. pellucidum* and *L. sericatum*), and (ii) the
179 two locations (Balambangan Island and Kinabatangan). In addition, the interaction effects of both
180 factors (species and location) were tested. A Shapiro-Wilk test for normality (Shapiro & Wilk,
181 1965), and a Levene's test (Brown & Forsythe, 1974) for homogeneity of variance, revealed that
182 some datasets were not normally distributed and showed non-homogeneity of variances (Table S3
183 & 4 in Supplementary File 2). Nevertheless, two-way ANOVA tests were still conducted since the
184 deviations of these datasets from the ANOVA assumption were considered not too serious (see
185 boxplots of Fig. 5), and the ANOVA was considered a robust test against the normality assumption
186 (Zar, 1999).

187

188   Chi-square two-way contingency table tests were performed to determine whether the
189 types of shell colour patterns and the presence of dark ring bands in the aperture were associated
190 with species identity and location respectively. Prior to the analyses, four two-way contingency
191 tables were produced by summarising the frequency of the categories of (1) shell colour patterns
192 *vs.* species, (2) shell colour patterns *vs.* location, (3) dark ring bands in aperture *vs.* species, and
193 (4) dark ring bands in aperture *vs.* location. Each of the tables was analysed by using Pearson's
194 Chi-squared test. When the expected frequency in the contingency table was less than 5, Fisher
195 exact test was performed instead of Pearson's Chi-squared test (Bower, 2003). All the statistical

196　analyses were performed in R statistical environment version 3.1.3 (R Core Team, 2015) with the
197　significant p-values set at 0.05 (R script in Supplementary File 3).
198
199　**RESULTS**
200
201　**The molecular phylogeny of the *Leptopoma* species in Sabah**
202　A total of 660 nucleotide sites were aligned for the COI gene, 558 nucleotide sites for the 16S gene
203　and 627 nucleotide sites for ITS-1 (Supplementary File 4). The aligned COI dataset consisted of
204　36.9% GC content, 207 (31.4%) parsimony informative, and 253 (38.3%) variable sites. Aligned
205　16s gene had 33.3% GC content with 276 (49.8%) parsimony informative and 406 (73.3%)
206　variable sites. On the other hand, aligned ITS-1 gene had 48.6% GC content, 158 (25.2%)
207　parsimony informative, and 274 (43.7%) variable sites. Phylogenetic analyses were run for four
208　datasets: ITS-1, COI, 16S and concatenated dataset of ITS-1, COI and 16S whereby gaps were
209　treated as missing data. Outgroup *Cyclophorus formensis* was used to root the tree.
210
211　　　The best-fitted models selected based on corrected Akaike Information Criterion (AICc)
212　were TPM3uf+G for ITS-1, TIM3+G for 16S, TIM3ef+G for COI first codon, TPM3uf+I for
213　second codon, and TPM3uf+G for third codon in COI. These models were applied in both ML and
214　BI analyses. Phylogenetic trees produced from both ML and BI based on concatenated dataset
215　showed no conflict in tree topologies. Therefore, only the BI tree based on concatenated dataset
216　was shown (Figure 3). The resultant phylogenetic tree supported the monophyly of three
217　*Leptopoma* species (*L. sericatum*, *L. pellucidum* and *L. atricapillum*) in Sabah. Each major clade
218　formed by each species was supported by significant supporting values (100 PP and ML bootstrap
219　larger than 75%).
220
221
222　**Phylogenetic signals relating to shell characters for the *Leptopoma* species**
223　Figure 4 shows the inter-relation between phylogeny and the quantitative and qualitative shell
224　characters for *L. sericatum*, *L. atricapillum* and *L. pellucidum*. A Phylogenetic signal test based on
225　Pagel's λ and Blomberg's K showed that spiral ridges and presence of dark ring band in aperture
226　represented a strong signal with $\lambda = 1$ and $K > 1$ (K= 4.536 for spiral ridges and K= 1.114 for dark
227　ring band) (see Table 1). However, shell patterns that often used as a diagnostic character in
228　traditional classification indicated a weak phylogenetic signal ($\lambda = 0.997$, K= 0.234). Among the
229　quantitative shell characters, shell height exhibited a strong signal according to Pagel's λ although
230　Blomberg's K indicated a weak phylogenetic signal. The ratio of shell height to width (SH/SW)
231　exhibited the weakest phylogenetic signal among all shell characters ($\lambda = 0$, K= 0.054).
232
233
234　**Geographical Variation in Shell Morphology**

235 Two-way ANOVA showed that all shell quantitative characters (except aperture height) differed
236 between the two locations (Table 2). In addition, all shell quantitative characters except shell width
237 and aperture height also differed between the two species. There was interactive effect of species
238 and location on the aperture height, shell spire height and ratio between shell height and width.
239
240   Chi-square analyses indicated significant association between the frequencies of shell
241 colour patterns and both the factors of species identity (Fisher's exact test: p =0.0000) and location
242 (Fisher's exact test: p=0.0000). Similarly, there was significant association between the
243 frequencies of the presence of the dark ring band in the aperture and both the factors of species
244 identity (Pearson's Chi-Squared with Yates' continuity correction: X^2 (1, N=249) = 4.88,
245 p=0.0271) and location (Pearson's Chi-Squared with Yates' continuity correction: X^2(1, N=249)
246 = 12.910, p=0.0003). Both contingency tables are available in Table S5 & S6 in Supplementary
247 File 2. Overall, the shell characters did not show consistent differences between *L. pellucidum* and
248 *L. sericatum* since the differences in shell form were coupled with geographical variations and
249 interaction effects between geography and species.
250

251 **DISCUSSION**
252 Although Cyclophoridae represents the most diverse family, it is also one of the less taxonomically
253 resolved Caenogastropoda families. The current classification of many Cyclophorids was based
254 solely on morphology characteristics where its reliability remains doubtful as this taxa possesses
255 exceptionally diverse variations in morphology. Past research has shown that molecular
256 phylogenetic analyses could provide insights into the taxonomy of morphologically-ill land snails
257 in this region (Nantarat *et al*., 2014a, 2014b; Liew *et al*., 2009; Liew *et al*., 2014). This study
258 presents the first molecular phylogeny investigation on genus *Leptopoma* in Sabah, one of the
259 understudied taxa within Cyclophoridae, and examines the concordance between morphology and
260 phylogeny as well as geographical variations in shell form.
261

262   All the phylogenetic trees based on different genes were congruent and provided significant
263 support for the monophyly of three morphologically similar *Leptopoma* species in Sabah namely
264 *L. pellucidum*, *L. sericatum* and *L. atricapillum*. The phylogenetic placement of genus *Leptopoma*
265 in Sabah was in concordance with its traditional morphology-based classification. For example,
266 the placement of *Leptopoma pellucidum* 6014 (Fig. 3) in this study, previously assumed as *L.
267 vitreum* due to its white colour shell which differed from other *L. pellucidum*, was revealed as
268 within the *L. pellucidum* clade which supported Vermeulen (1999)'s decision to assign *L. vitreum*
269 as synonymous to *L. pellucidum*. In the case of *L. pellucidum* and *L. sericatum*, Vermeulen
270 separated them into two species provisionally due to the existence of intermediate forms between
271 the two species. In this study, results suggested that the two species could be unequivocally
272 regarded as separate. In short, the findings of this study are in line with past research which
273 proposed that a combination of morphology and molecular approaches could improve taxonomy
274 of land snails.

275

276   A morphological character is assumed to have strong phylogenetic signal when the same
277 character clusters together within closely-related species (Blomberg *et al*., 2003). This could be a
278 useful diagnostic indicator for species delimitation. The phylogenetic signal tests showed that
279 spiral ridges had a significant phylogenetic signal (λ=1, K>1). Distinct spiral ridges were present
280 in *L. sericatum* and *L. atricapillum* while *L. pellucidum* had weak spiral ridges. This indicated that
281 weak spiral ridges might be an automorphy character for *L. pellucidum* which could be useful in
282 discriminating *L. pellucidum* from *L. sericatum* and *L. atricapillum*. This result was in agreement
283 with Vermeulen (1999) where spiral ridges were also used as a key to delimitate between *L.*
284 *pellucidum* and *L. sericatum*.

285

286   The presence of a dark ring band in the aperture of land snails has not been observed in
287 other Cyclophorids and was not mentioned in other revision works of *Leptopoma* species. A
288 phylogenetic signal test showed that the presence of a dark ring band exhibited a significant
289 phylogenetic signal. However, this character was found to be strongly affected by geographical
290 variations when two species from two different locations were compared. All shells with a dark
291 ring band located in the shell aperture were collected from a single location in Kinabatangan, i.e.
292 the Tabin Wildlife Reserve area. The dark ring band was presented in both species with *L.*
293 *pellucidum* showing more instances than *L. sericatum*. The underlying causes of the presence of
294 this shell character remain yet unknown. Compared to results from phylogenetic signal test, the
295 presence of a dark ring band in the shell aperture would not be a reliable character to distinguish
296 between *Leptopoma* species due to geographically-induced morphology variations.

297

298   Shell colour patterns are usually used as key determinants to discriminate between species
299 in traditional morphology classification. One of the sister taxa of *Leptopoma*, the species in genus
300 *Cyclophorus*, was distinguished unambiguously based on shell patterns that were also supported
301 by molecular data (Nantarat *et al*., 2014b). Compared to genus *Cyclophorus*, shell colour patterns
302 of the genus *Leptopoma*, particularly in *L. sericatum* and *L. pellucidum*, exhibited a weak
303 phylogenetic signal. This case of shell colour pattern polymorphisms of the two *Leptopoma* species
304 is similar to other well-known land snails namely *Cepaea nemoralis* and *C. hortensis* (Owen &
305 Bengtson, 1972; Ozgo & Schilthuizen, 2012; Cameron & Cook, 2012; Cameran, 2013). However,
306 unlike *Cepaea* land snails that have been studied extensively, the causal mechanism for the
307 *Leptopoma* land snail's diverse shell colour patterns is still unknown. This study also revealed that
308 the *Leptopoma* species exhibits idiosyncratic differences between locations in the degree of shell
309 polymorphisms. For example, the *Leptopoma* population at Balambangan Island has more shell
310 colour patterns as compared to the population at Kinabatangan. As a result, the geographically-
311 induced variations in shell colour patterns and weak phylogenetic signal strongly suggest that shell
312 patterns should not be used as a diagnostic character for the genus *Leptopoma*.

313

314   Significant variations in quantitative shell characters within or between species were often
315 detected in family Cyclophoridae (Lee *et al.*, 2012; Nantarat *et al.*, 2014b) and gastropods
316 (Kameda *et al*, 2007; Desouky & Busais, 2012; Hirano *et al.*, 2014). From the phylogenetic signal
317 test, only shell height produced a significant signal. In the Vermeulen (1999) description of *L.*
318 *sericatum* and *L. pellucidum*, the ratio between shell height and width of *L. sericatum* is slightly
319 smaller than *L. pellucidum*. This study revealed a high degree of geographical variations in the
320 quantitative shell characters; for example, both *Leptopoma* species from Balambangan Island were
321 larger than the same species found in Kinabatangan. Previous studies suggested that land snails
322 found on islands tend to undergo extensive morphological diversification (Johnson & Black, 2000;
323 Stankowski, 2011). In view of this, quantitative shell characters are thus not advisable as a
324 diagnostic indicator for species delimitation due to the strong influence of geographical variations.
325

326   This study has only revealed partial information on the phylogenetic and morphology
327 variations of all *Leptopoma* species in their entire distribution range. However, despite its small
328 geographical scale, the study has resolved taxonomic uncertainties of three *Leptopoma* species and
329 revealed notable variations in both the quantitative and qualitative shell characters for the species.
330 From the findings it is clear that any attempt of the taxonomy works on the rest of *ca.* 100
331 *Leptopoma* species in the future should consider the possible caveats in using the shell characters
332 as the sole evidences and should instead include molecular phylogeny in the study.

333 **CONCLUSION**
334 This study represents the first attempt to conduct phylogenetic investigation into the genus
335 *Leptopoma* and provides phylogenetic assessment of the genus in Sabah. The results
336 unambiguously separate *L. pellucidum*, *L. sericatum* and *L. atricapillum* into three distinct
337 monophyletic groups, and highlight substantial congruence among the traditional morphological
338 classifications based on spiral ridges and molecular phylogeny of the *Leptopoma* species in Sabah.
339 After performing the phylogenetic signal tests, it can be stated that all quantitative and many
340 qualitative shell characters are not reliable diagnostic indicators for discriminating between the
341 *Leptopoma* species due to the considerable geographical variations in shell form. This study
342 represents an attempt to resolve the taxonomy conundrum for the remaining 100 little known
343 *Leptopoma* species from other distribution regions. Further studies that include more samples from
344 a wider geographical reach are recommended.
345

351 **SUPPLEMENTARY FILES**

352 SUPPLEMENTARY FILE 1. Raw data for shell morphological analysis.

353 SUPPLEMENTARY FILE 2. Table S1. Specimens information; Table S2. Shell colour patterns
354 description, Table S3 & S4. Results normality tests and homogeneity of variances tests prior to
355 ANOVA; Table S5 & S6. Frequency data of shell qualitative characters used for chi-square tests.

356 SUPPLEMENTARY FILE 3. R script for shell morphological data and phylogenetic signal
357 analysis.

358 SUPPLEMENTARY FILE 4. FASTA file for DNA sequences alignments for concatenated data
359 of COI, 16S, and ITS-1.

360 **REFERENCES**

361 Blomberg, S. P., Garland, T., & Ives, A. R. 2003. Testing for phylogenetic signal in comparative
362 data: behavioral traits are more labile. Evolution, 57(4): 717-745.

363 Bower, K. M. 2003. When to use Fisher's exact test. In American Society for Quality, Six Sigma
364 Forum Magazine, 2(4): 35-37.

365 Brown, M. B., & Forsythe, A. B. 1974. Robust tests for the equality of variances. Journal of the
366 American Statistical Association, 69(346), 364-367.

367 Cain, A. J., & Sheppard, P. M. 1954. Natural selection in *Cepaea*. Genetics, 39(1): 89.

368 Cameron, R. A., & Cook, L. M. 2012. Habitat and the shell polymorphism of *Cepaea nemoralis*
369 (L.): interrogating the Evolution Megalab database. Journal of Molluscan Studies, eyr052.

370 Cameron, R. A. 2013. The poor relation? Polymorphism in *Cepaea hortensis* (OF Müller) and the
371 Evolution Megalab. Journal of Molluscan Studies, 79(2), 112-117.

372 Colgan, D., Ponder, W. F., & Eggler, P. E. 2000. Gastropod evolutionary rates and phylogenetic
373 relationships assessed using partial 28S rDNA and histone H3sequences. Zoologica Scripta, 29(1):
374 29-63.

375 Colgan, D. J., Ponder, W. F., Beacham, E., & Macaranas, J. M. 2003. Gastropod phylogeny based
376 on six segments from four genes representing coding or non-coding and mitochondrial or nuclear
377 DNA. Molluscan Research, 23(2): 123-148.

378 Colgan, D. J., Ponder, W. F., Beacham, E., & Macaranas, J. 2007. Molecular phylogenetics of
379 Caenogastropoda (Gastropoda: Mollusca). Molecular Phylogenetics and Evolution, 42(3): 717-
380 737.

381 Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. 2012. jModelTest 2: more models, new
382 heuristics and parallel computing. Nature methods, 9(8): 772-772.

383 Desouky, M., & Busais, S. 2012. Phylogenetic relationships of the land snail *Eobania vermiculata*
384 (Müller, 1774) from Egypt and Saudi Arabia. A combined morphological and molecular
385 analysis. The Journal of Basic & Applied Zoology, 65(2): 144-151.

386 Folmer, O., Back, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. 1994. DNA primers for amplification
387 of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar.
388 Biol. Biotech, 3: 294–299.

389 Godwin- Austen, L., & HH, C. 1891. 4. On a collection of land-shells made in Borneo by Mr. A.
390 Everett, with Descriptions of supposed new species. In Proceedings of the Zoological Society of
391 London, 59(1): 22-48.

392 Grateloup, JPS.DE. 1840. Memoire sue plusieurs especes de coquilles nouvellesou peu connues
393 de mollusque exotiques, vivants, terrestres, fluviatiles et marins. Actes de la Societe Linneenne de
394 Bordeaux 11: 389-455.

395 Gude, G. K., 1921.Mollusca. III. Land Operculates (Cyclophoridea, Truncatellidae, Assimineidae,
396 Helicinidae) in The Fauna of British India, London, i-xiv, 1-386, 42figs, 2 pls

397 Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis
398 program for Windows 95/98/NT. In Nucleic acids symposium series, 41: 95-98.

399 Harmon L. J, Jason T. W, Chad D. B, Richard E. G, & Wendell Challenger. 2008. GEIGER:
400 investigating evolutionary radiations. Bioinformatics, 24:129-131.

401 Hillis, D. M., & Dixon, M. T. 1991. Ribosomal DNA: molecular evolution and phylogenetic
402 inference. The Quarterly review of biology, 66(4), 411-453.

403 Hirano, T., Kameda, Y., Kimura, K., & Chiba, S. 2014. Substantial incongruence among the
404 morphology, taxonomy, and molecular phylogeny of the land snails *Aegista*, *Landouria*,
405 *Trishoplita*, and *Pseudobuliminus* (Pulmonata: Bradybaenidae) occurring in East Asia. Molecular
406 phylogenetics and evolution, 70: 171-181.

407 Huelsenbeck, J. P., & Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic
408 trees. Bioinformatics, 17(8): 754-755.

409 Johnson, M.S. & Black, R. 2000. Associations with habitat versus geographic cohesiveness: size
410 and shape of *Bembicium vittatum* Philippi (Gastropoda: Littorinidae) in the Houtman Abrolhos
411 Islands. Biological Journal of the Linnean Society. 71:563-580.

412 Jonges, K. 1980. Genital ducts in several species of the cyclophorid genus *Leptopoma* (Mollusca,
413 Mesogastropoda). Bijdragen tot de Dierkunde, 50: 292-302.

414 Kameda, Y., Kawakita, A., & Kato, M. 2007. Cryptic genetic divergence and associated
415 morphological differentiation in the arboreal land snail *Satsuma* (*Luchuhadra*) *largillierti*

416 (Camaenidae) endemic to the Ryukyu Archipelago, Japan. Molecular phylogenetics and
417 evolution, 45(2): 519-533.

418 Kessing, B., Croom, H., Martin, A., McIntosh, C., McMillan, W. O., & Palumbi, S. 1989. The
419 simple fool's guide to PCR. Department of Zoology, University of Hawaii, Honolulu, 1–23.

420 Kobelt, W. 1902. Cyclophoridae. Das Tierreich 16: 1–662.

421 Lee, Y. C., Lue, K. Y., & Wu, W. L. 2008a. A molecular phylogenetic investigation of
422 *Cyathopoma* (Prosobranchia: Cyclophoridae) in East Asia. Zoological Studies, 47(5): 591-604.

423 Lee, Y. C., Lue, K. Y., & Wu, W. L. 2008b. Molecular evidence for a polyphyletic genus *Japonia*
424 (Architaenioglossa: Cyclophoridae) and with the description of a new genus and two new
425 species. Zootaxa. 1792: 22-38.

426 Lee, Y. C., Lue, K. Y., & Wu, W. L. 2012. The Phylogeny and Morphological Adaptations of
427 *Cyclotus taivanus* ssp.(Gastropoda: Cyclophoridae).Malacologia, 55(1), 91-105.

428 Liew, T. S., Schilthuizen, M., & Vermeulen, J. J. 2009. Systematic revision of the genus *Everettia*
429 Godwin-Austen, 1891 (Mollusca: Gastropoda: Dyakiidae) in Sabah, northen Borneo. Zoological
430 Journal of the Linnean Society, 157(3): 515-550.

431 Liew, T. S., Vermeulen, J. J., Marzuki, M. E., & Schilthuizen, M. 2014. A cybertaxonomic revision
432 of the micro-land snail genus *Plectostoma Adam* (Mollusca, Caenogastropoda, Diplommatinidae),
433 from Peninsular Malaysia, Sumatra and Indochina. ZooKeys, 393: 1-107.

434 Laidlaw, F. F. 1937. Notes on the land Mollusca from Mount Kinabalu, British North Borneo. The
435 Bulletin of the Raffles Museum, 13: 177-189.

436 Metcalfe W., 1851. An enumeration of species of recent shells, received by W.J. Hamilton. Esq.,
437 from Borneo, in November 1850. With descriptions of new species. Proceedings of the Zoological
438 Society of London 19: 70-74.

439 Miller, M.A., Pfeiffer, W., & Schwartz, T. 2010. Creating the CIPRES Science Gateway for
440 inference of large phylogenetic trees. In, Proceedings of the Gateway Computing Environments
441 Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1-8.

442 Nantarat, N., Tongkerd, P., Sutcharit, C., Wade, C. M., Naggs, F., & Panha, S. 2014a. Phylogenetic
443 relationships of the operculate land snail genus *Cyclophorus*   Montfort,   1810   in
444 Thailand. Molecular Phylogenetics and Evolution, 70: 99-111.

445 Nantarat, N., Wade, C. M., Jeratthitikul, E., Sutcharit, C., & Panha, S. 2014b. Molecular evidence
446 for cryptic speciation in the *Cyclophorus fulguratus* (Pfeiffer, 1854) species complex
447 (Caenogastropoda: Cyclophoridae) with description of new species. PloS one, 9(10), e109785.

448 Owen, D. F., & Bengtson, S. A. 1972. Polymorphism in the land snail *Cepaea hortensis* in Iceland.
449 Oikos, 218-225.

450 Ożgo, M., & Schilthuizen, M. 2012. Evolutionary change in *Cepaea nemoralis* shell colour over
451 43 years. Global Change Biology, 18(1), 74-81.

452 Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of
453 discrete characters on phylogenies. Systematic biology, 612-622.

454 Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R
455 language. Bioinformatics, 20: 289-290.

456 Pfeiffer, L., 1851. Descriptions of forty-three new species of Cyclostomatacea, from collection of
457 H. Cumming, Esq. Proceedings of Zoological Society of London, 19: 242-251.

458 R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for
459 Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

460 Ratnasingham, S., & Hebert, P. D. 2007. BOLD: The Barcode of Life Data System (http://www.
461 barcodinglife. org). Molecular Ecology Notes, 7(3): 355-364.

462 Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other
463 things). Methods Ecol. Evol, 3: 217-223.

464 Sarasin P, Sarasin F. 1899. Die Land-mollusken von Celebes. Wiesbaden: Kreidel. 15-25.

465 Schilthuizen, M., & Rutjes, H. A. 2001. Land snail diversity in a square kilometre of tropical
466 rainforest in Sabah, Malaysian Borneo. Journal of Molluscan Studies, 67(4): 417-423.

467 Shapiro, S. S., & Wilk, M. B. 1965. An analysis of variance test for normality (complete
468 samples). Biometrika, 52(3-4), 591-611.

469 Stamatakis, A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of
470 Large Phylogenies. Bioinformatics 10.1093/bioinformatics/btu033.

471 Stankowski, S. 2011. Extreme, continuous variation in an island snail: local diversification and
472 association of shell form with the current environment. Biological Journal of the Linnean Society,
473 104(4): 756-769.

474 Uchidal, R., Yahya, B. E., Ahmad, A. H., Sapaat, A., & Tomiyamal, K. 2013. Land snail fauna of
475 Sabah, Borneo, Malaysia. Venus. 71(1/2): 49-59.

476 Vermeulen, J. J. 1999. Notes on the non-marine molluscs of the island of Borneo 9. The genera
477 *Cyclophorus*, *Leptopoma*, and *Craspedotropis* (Gastropoda Prosobranchia:
478 Cyclophoridae). BASTERIA-LISSE-. 63(4/6): 139-163.

479 Zar, J.H .1999. Biostatistical Analysis, 4th ed., New Jersey: Prentice-Hall Inc

480 Zilch, A., 1954. Die Typen und Typoide des Natur-Museums Senckenberg, 13 Mollusca,
481 Cyclophoridae, Cyclophorinae-Cyclophoreae (1).Arch. Molluskenk. 83:141-157, 4 pl

Figure 1(on next page)

Distribution map of four *Leptopoma* species in Sabah based on the records from BORNEENSIS Mollusca collection, Institute of Tropical Biology and Conservation, Universiti Malaysia Sabah.

Each circle represents a collection lot of the *Leptopoma* species and the size of circles increase indicates the number of specimens in the lot. The inset (A) and (B) show the sympatric species of *L. sericatum* and *L. pellucidum* in Balambangan Island and in Kinabatangan that used for shell morphological analysis.

Legend

Leptopoma species

- • Leptopoma atricapillum (Sowerby, 1843)
- • Leptopoma pellucidum (Grateloup, 1840)
- • Leptopoma sericatum (Pfeiffer, 1851)
- • Leptopoma undatum (Metcalfe, 1851)
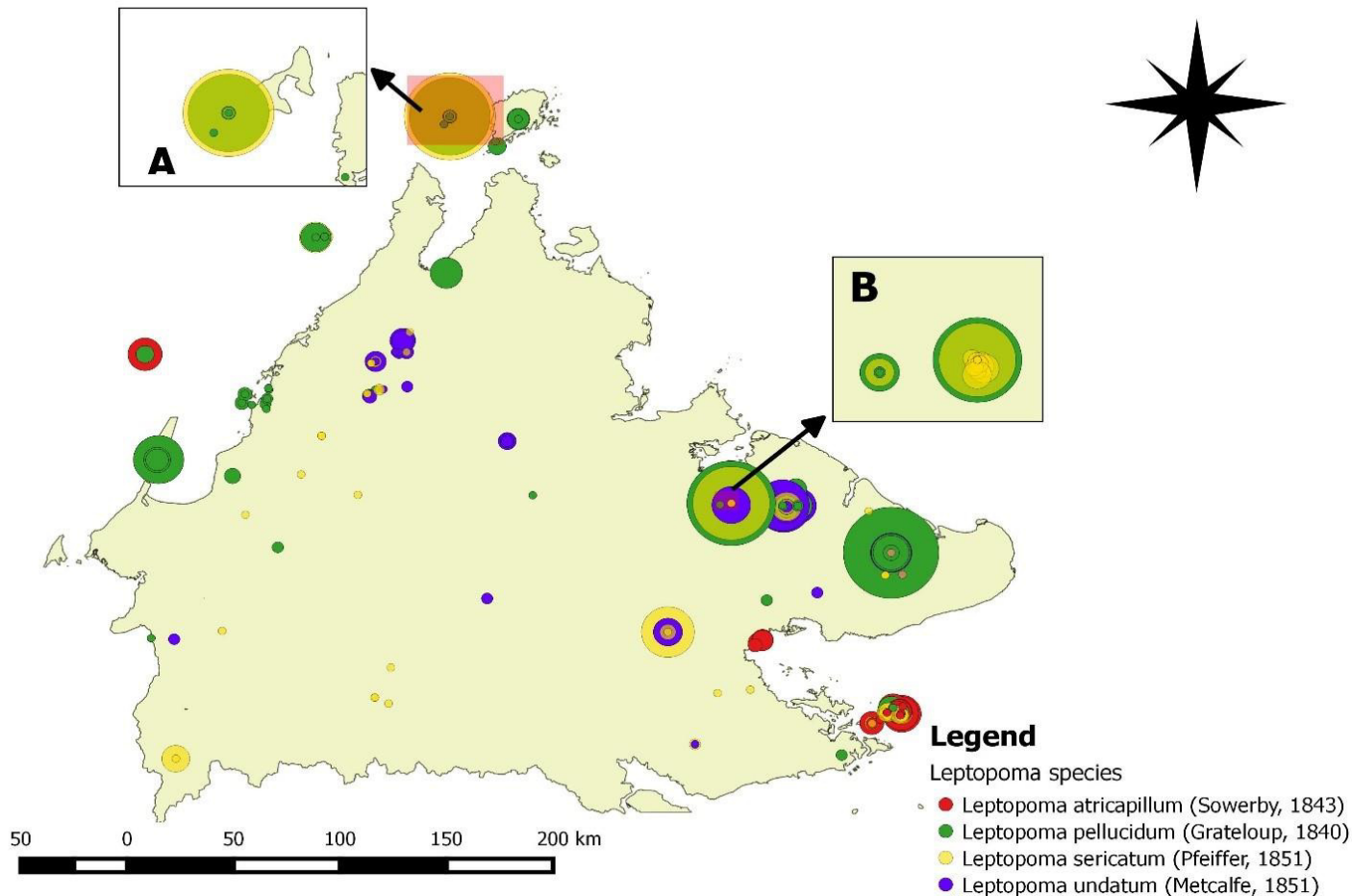
Figure 2(on next page)

Qualitative and quantitative shell traits obtained from shell aperture view.

(A) The five shell quantitative measurements: SH - Shell height, SW - Shell width, AH - Aperture height, AW - Aperture width, SpH - Shell spire height. (B) Spiral ridges: Left – Strong, Right – Weak. (C) The eight types of shell colour patterns. (D) Dark ring band in aperture: Left – Presence, Right – Absence.

(A) Shell measurements
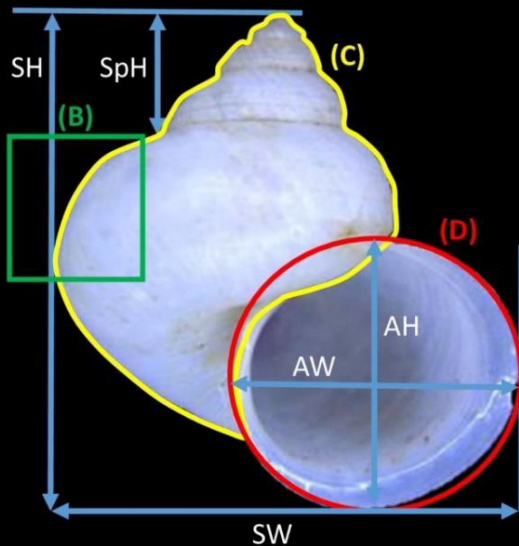
(B) Spiral ridges

(C) Shell colour patterns

P1 P2 P3

P4 P5 P6

P7 P8

5 mm
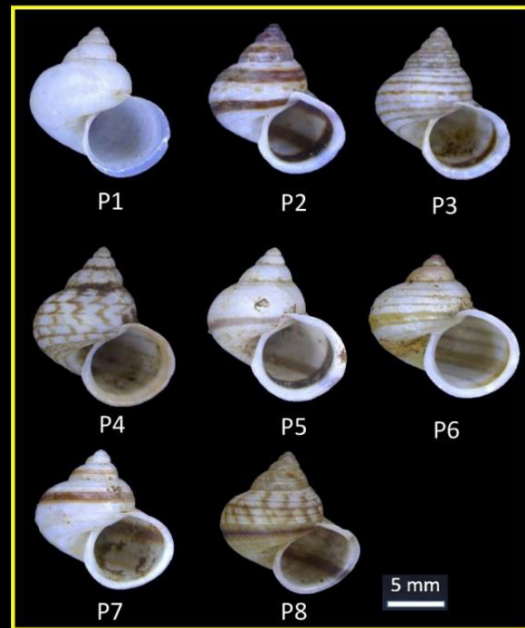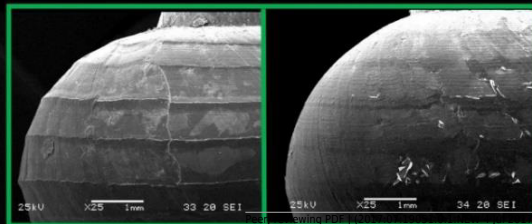
(D) Dark ring band in aperture
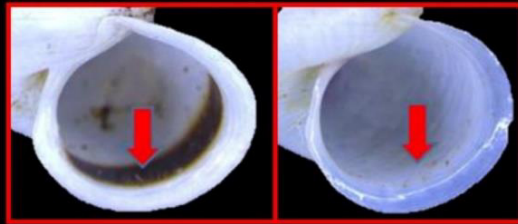
Figure 3(on next page)

Bayesian inference tree of *Leptopoma* spp. based on concatenated dataset of 16S, COI and ITS-1.

Support values on branches indicate Bayesian posterior probability (BI)/ maximum likelihood bootstrap value (ML). Internal branches with ML bootstrap value = 100% and PP value =100 were not represents in the figure. Number behind each specimens of Sabah *Leptopoma* species refer to specimen number as in (Table S1 in Supplementary File 2), and the specimens with asterisk are non-Sabah's *Leptopoma* species. The monophyly of three *Leptopoma* species in Sabah were supported as shown in clade A, B, and C. Scale bar for branch length = 0.1 substitutions per site. Asterisk marked the taxa obtained from genbank.
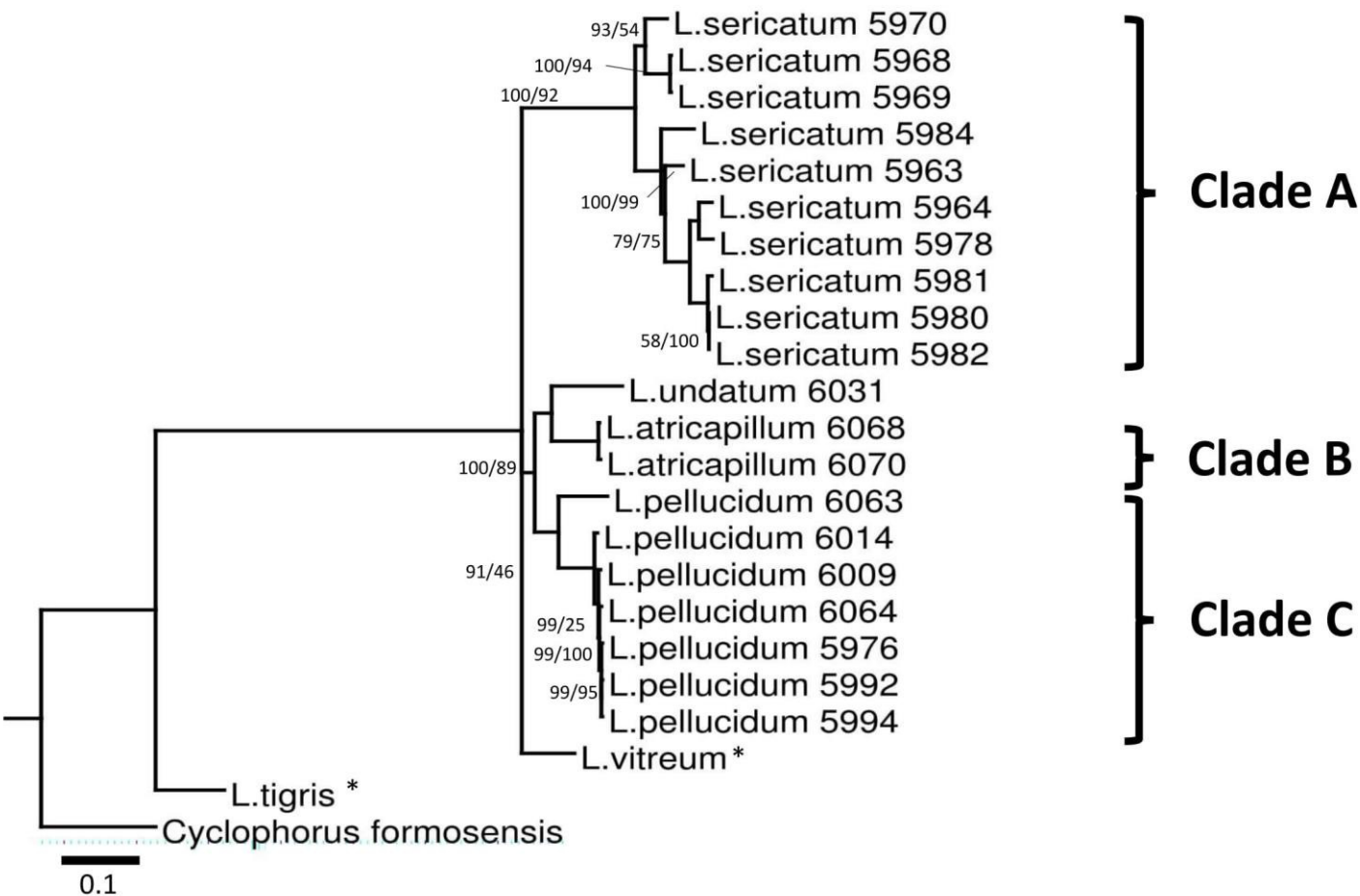
Figure 4(on next page)

Figure 4 Shell quantitative and qualitative shell traits were mapped for the phylogenetic tree.

Tree as in Figure 3, which only 14 adult of the three *Leptopoma* species were included whereas the juvenile specimens together with other outgroup taxa were dropped from the tree. Different categories of the three qualitative shell characters: spiral ridges, shell colour patterns and dark ring band in aperture (referred to Figure 1A, 1B, 1C respectively) were represented by different colour of the squares; and the six shell quantitative measurements: shell height, shell width, aperture height, aperture width, shell spire height and ratio of shell height and width were represented by the size of the circle.
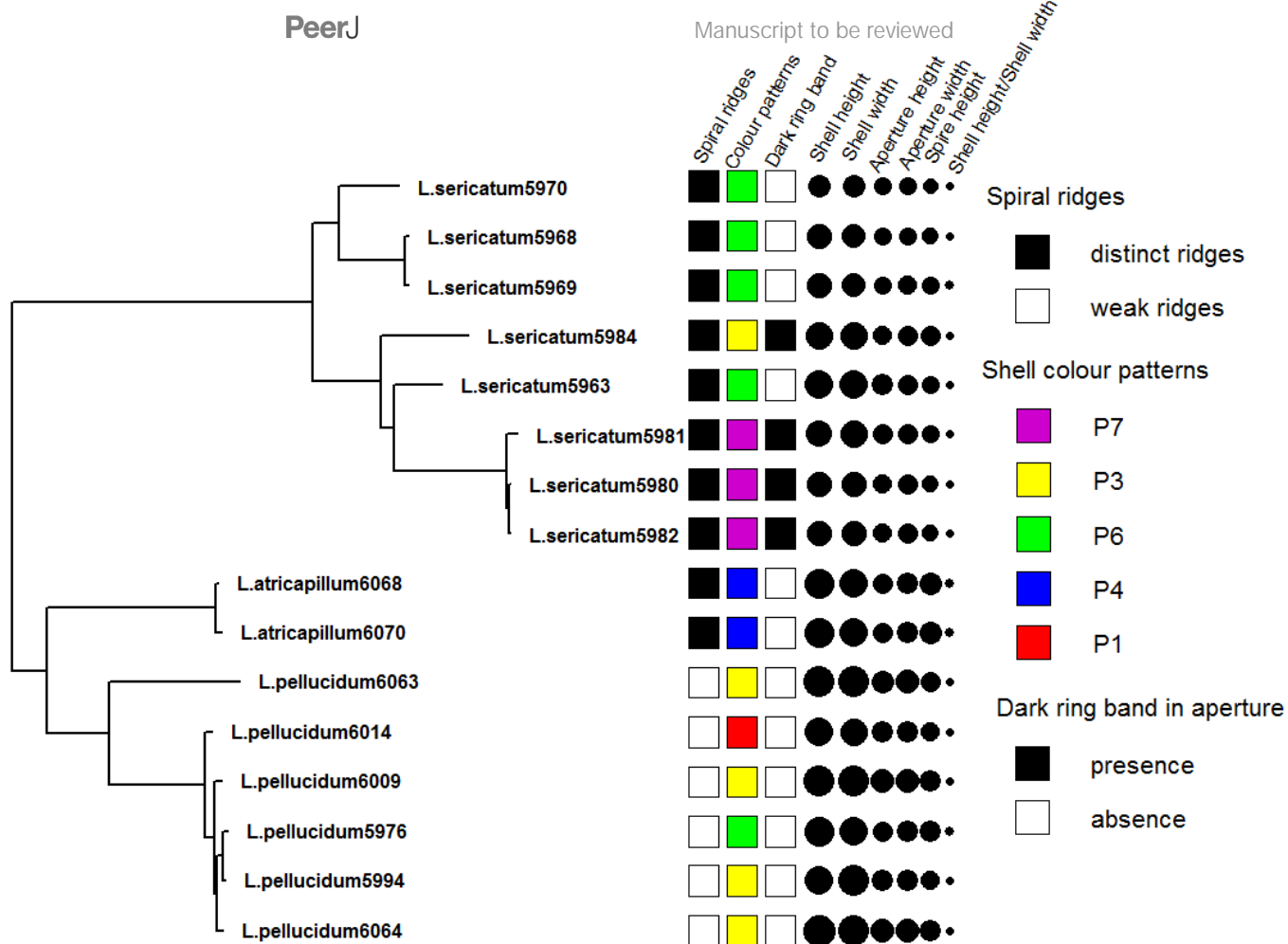
Figure 5(on next page)

Boxplots show the differences of the six quantitative measurements of shell for the *Leptopoma pellucidum* and *L. sericatum* in each of the two locations (Balambangan Island and Kinabatangan region).

Grey boxplot indicated sample from Balambangan Island (BI) and white boxplot indicated sample from Kinabatangan (K). Sample sizes for each dataset were: BI-pellucidum (n=45); K-pellucidum (n=90); BI-sericatum (n=46); K-sericatum (n=68).
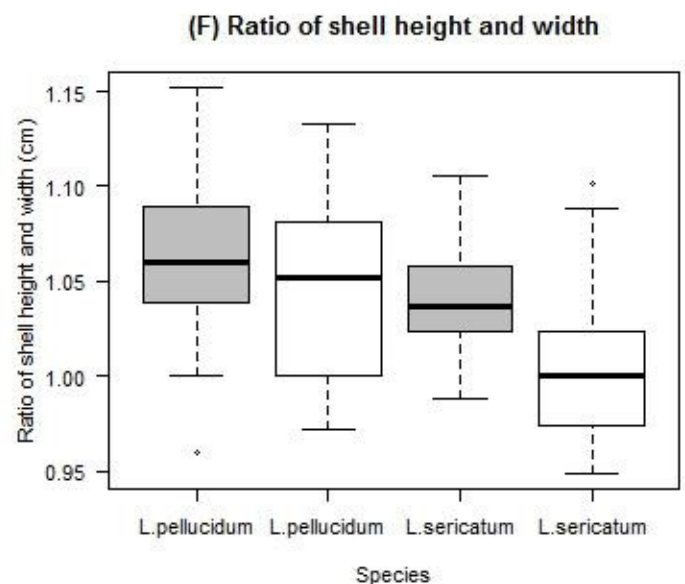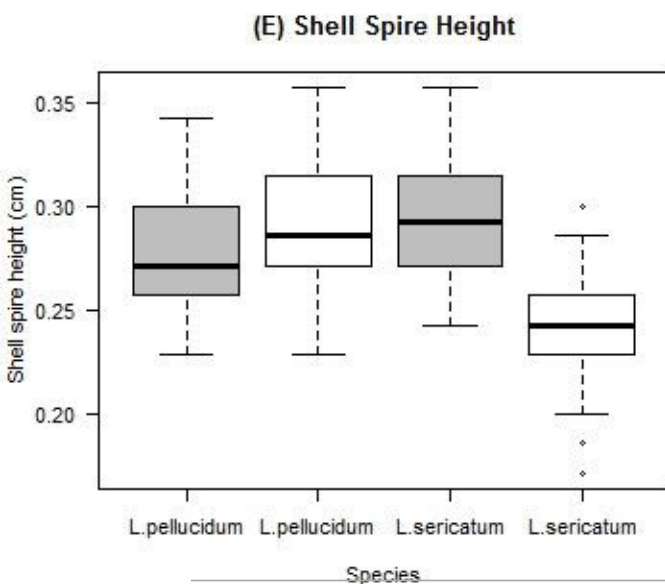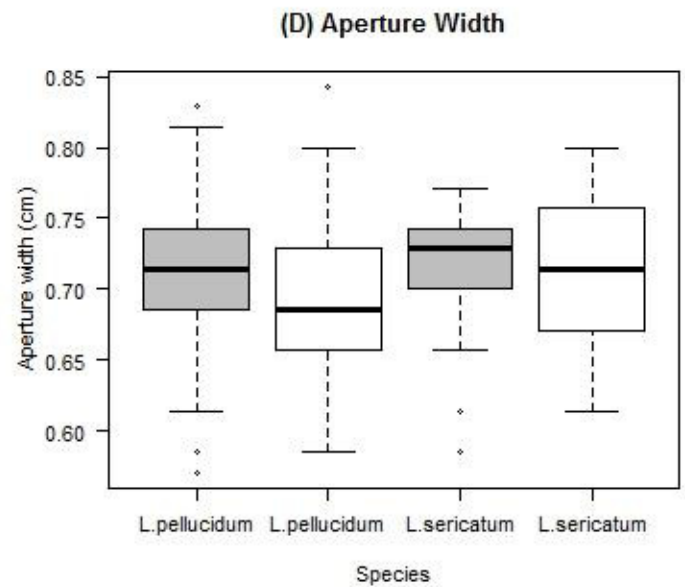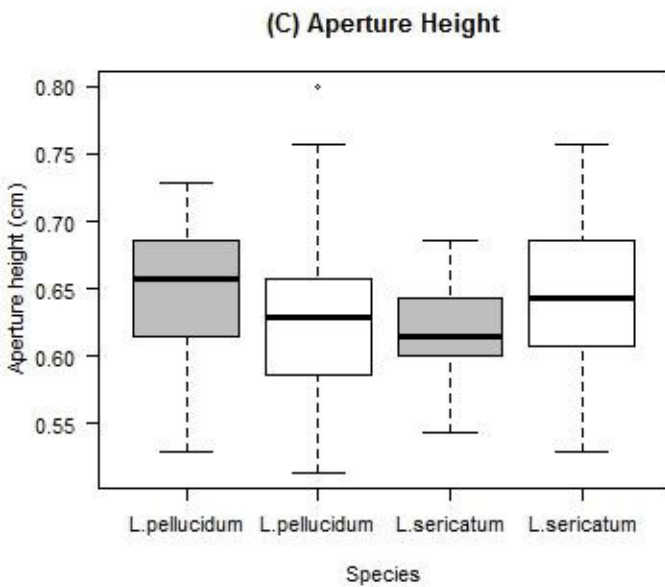
(C) Aperture Height

(D) Aperture Width

(E) Shell Spire Height
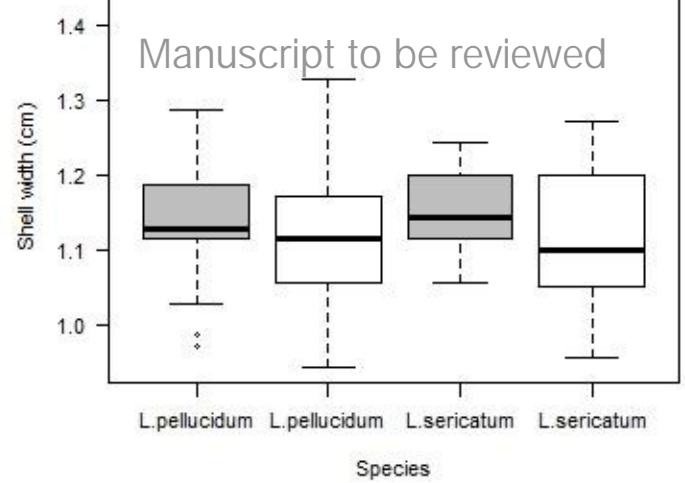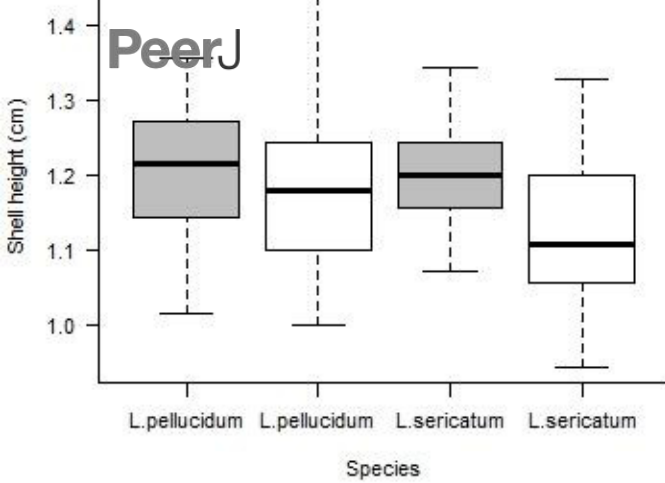
(F) Ratio of shell height and width

Table 1(on next page)

Phylogenetic signal test result acquired from Pagel's λ method and Blomberg's K method. Values equal to 1 or more than 1 were bolded.

Abbreviations: SH, shell height; SW, shell width; AH, aperture height; AW, aperture width;

SpH, shell spire height; SH/SW ratio of shell height and width.

1 **Table 1.** Phylogenetic signal test result acquired from Pagel's λ method and Blomberg's K method.
2 Values equal to 1 or more than 1 were bolded.

Shell traits	Lambda (λ)	likelihood score (alternative model)	likelihood score (null model, λ=0)	p-value	K	P
Patterns	0.997	-17.986	-21.906	0.005	0.234	0.014
Spiral ridges	**1.000**	-3.654	-9.704	0.0005	**4.490**	0.001
Dark ring band	**1.000**	-4.418	-7.274	0.017	**1.317**	0.007
AH	0.998	-15.969	-21.266	0.001	0.518	0.001
AW	0.866	-16.641	-21.395	0.002	0.437	0.001
SpH	0.894	-17.850	-22.426	0.002	0.444	0.003
SH	**1.000**	-24.197	-29.651	0.0007	0.567	0.001
SW	0.829	-24.040	-29.651	0.001	0.442	0.001
SH/SW	0	17.147	17.147	1	0.056	0.320

3 Abbreviations: SH, shell height; SW, shell width; AH, aperture height; AW, aperture width; SpH,
4 shell spire height; SH/SW ratio of shell height and width.

5

Table 2(on next page)

Two-way ANOVA for the effect of geographical variation and species identity on six quantitative shell traits. Significant p-values were bolded.

Abbreviations: SH, shell height; SW, shell width; AH, aperture height; AW, aperture width;

SpH, shell spire height; SH/SW, ratio of shell height and width.

1 **Table 2.** Two-way ANOVA for the effect of geographical variation and species identity on six q
2 uantitative shell traits. Significant p-values were bolded.
3

	Geographical region			Species identity			Geographical* Species		
	df	F	P-value	df	F	P-value	df	F	P-value
SH	1	18.88	**2.03e-05**	1	12.763	**0.0004**	1	3.551	0.0607
SW	1	5.376	**0.0212**	1	0.104	0.7473	1	0.586	0.4447
AH	1	0.086	0.770	1	0.000	0.987	1	16.185	**7.66e-05**
AW	1	4.235	**0.0407**	1	4.399	**0.0370**	1	1.994	0.1592
SpH	1	24.92	**1.14e-06**	1	36.33	**6.08e-09**	1	80.01	**< 2e-16**
SH/SW	1	17.36	**4.29e-05**	1	62.10	**1.07e-13**	1	5.53	**0.0195**

4 Abbreviations: SH, shell height; SW, shell width; AH, aperture height; AW, aperture width; SpH,
5 shell spire height; SH/SW, ratio of shell height and width.

6