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

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. Species-level identification in this group has been based on shell morphology and colouration, as well as some anatomical features based on small sample sizes. 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. There are four *Leptopoma* species found in Sabah, Borneo, and their taxonomy status remains uncertain due to substantial variation in shell forms. This study focuses on the phylogenetic relationships and geographical variation in shell form of three *Leptopoma* species from Sabah.The phylogenetic relationship of these species was first estimated by performing Maximum Likelihood and Bayesian analysis based on mitochondrial genes (16S and COI) and nuclear gene (ITS-1). Then, 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 to width) and three qualitative shell characters (i.e. shell colour patterns, spiral ridges, and dark apertural band) 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 only one of the shell characters (i.e. shell spiral ridges) in the current morphological-based classification. 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 on *Leptopoma* species is based on small sample size and the findings appear only applicable to *Leptopoma* species in Sabah. Nevertheless, we anticipate this study to be a starting point for more

detailed investigations to include the other still little-known (*ca*. 100) *Leptopoma* species and highlights a need to assess variations in shell characters before they could be used in species classification. we anticipate this study to be a starting point for more detailed investigations to include the other still little-known (*ca*. 100) *Leptopoma* species 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**
2 **(CAENOGASTROPODA: CYCLOPHORIDAE): AN ANALYSIS OF MOLECULAR**
3 **PHYLOGENY AND GEOGRAPHICAL VARIATIONS IN SHELL FORM**

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 Species-level identification in this group has been based on shell morphology and colouration, as
10 well as some anatomical features based on small sample sizes. However, the implications of the
11 inter- and intra-species variations in shell form to the taxonomy of *Leptopoma* species and the
12 congruency of its current shell based taxonomy with its molecular phylogeny are still unclear.
13 There are four *Leptopoma* species found in Sabah, Borneo, and their taxonomy status remains
14 uncertain due to substantial variation in shell forms. This study focuses on the phylogenetic
15 relationships and geographical variation in shell form of three *Leptopoma* species from
16 Sabah.The phylogenetic relationship of these species was first estimated by performing
17 Maximum Likelihood and Bayesian analysis based on mitochondrial genes (16S and COI) and
18 nuclear gene (ITS-1). Then, a total of six quantitative shell characters (i.e. shell height, shell
19 width, aperture height, aperture width, shell spire height, and ratio of shell height to width) and
20 three qualitative shell characters (i.e. shell colour patterns, spiral ridges, and dark apertural band)
21 of the specimens were mapped across the phylogenetic tree and tested for phylogenetic signals.
22 Data on shell characters of *Leptopoma sericatum* and *Leptopoma pellucidum* from two different
23 locations (i.e. Balambangan Island and Kinabatangan) where both species occurred sympatrically
24 were then obtained to examine the geographical variations in shell form. The molecular
25 phylogenetic analyses suggested that each of the three *Leptopoma* species was monophyletic and
26 indicated congruence with only one of the shell characters (i.e. shell spiral ridges) in the current
27 morphological-based classification. Although the geographical variation analyses suggested
28 some of the shell characters indicating inter-species differences between the two *Leptopoma*
29 species, these also pointed to intra-species differences between populations from different
30 locations. This study on *Leptopoma* species is based on small sample size and the findings
31 appear only applicable to *Leptopoma* species in Sabah. Nevertheless, we anticipate this study to
32 be a starting point for more detailed investigations to include the other still little-known (*ca.* 100)
33 *Leptopoma* species and highlights a need to assess variations in shell characters before they
34 could be used in species classification.
35
36

37 **INTRODUCTION**
38 The terrestrial snail genus *Leptopoma* is one of 35 genera in the family Cyclophoridae (Kobelt,
39 1902), which has a wide global distribution range extending across much of the Oriental and

40   Australasia zoogeographical regions. An early global overview of species-level diversity in
41   *Leptopoma* species classified the genus into four subgenera with a total of 105 species (Kobelt,
42   1902). Several subsequent regional taxonomic reviews of *Leptopoma* were conducted for the
43   Philippines (Zilch, 1956), South Asia (Gude, 1921), and most recently for Borneo (Vermeulen,
44   1999). To date, taxonomic works on *Leptopoma* (Kobelt, 1902; Gude, 1921; Zilch, 1954;
45   Vermeulen, 1999) have been based mainly on shell morphology (i.e shell size, shape, colour
46   pattern and sculpture). Besides shell morphology, other characters of the soft body such as radula,
47   operculum, and genital duct have been used from time to time in the species delimitation
48   (Sarasin & Sarasin, 1899; Jonges, 1980). Although several species have been included in
49   phylogenetic studies focussing on other taxa (Colgan *et al*., 2000, 2003, 2007; Lee *et al*. 2008a,
50   2008b; Nantarat *et al*., 2014a), little is known about the relationship within the genus *Leptopoma*
51   itself.
52
53   The genus *Leptopoma* is widespread in the Philippines and the adjacent Malaysian state
54   of Sabah, which is located in the northern part of Borneo (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. Of these, *Leptopoma undatum*
57   (Metcalfe, 1851) and *Leptopoma atricapillum* (Sowerby, 1843) can be found in Borneo and the
58   Philippines (Vermeulen, 1999). *Leptopoma sericatum* (Pfeiffer, 1851) is distributed in Borneo
59   (Vermeulen, 1999). *Leptopoma pellucidum* (Grateloup, 1840) is widely spread in Sabah but the
60   actual global range is unknown due to taxonomy uncertainty (Vermeulen, 1999). Its putative
61   synonym, *L. vitreum* (Lesson, 1830), has a wide range encompassing Taiwan, South Asia, and
62   Papua New Guinea (Vermeulen, 1999).
63
64   *Leptopoma undatum* is readily distinguished from the others by its uniform white shell
65   (translucent when young and opaque when old) and distinctive shell shape (i.e. relatively less
66   convex whorls and sharply keeled at the last whorl). The other three species, *Leptopoma*
67   *atricapillum*, *L. sericatum* and *L. pellucidum*, are very similar in terms of shell shape with all
68   showing and sharing colour pattern polymorphism. *L. atricapillum* and *L. sericatum*, however
69   have strongly-defined spiral ridges on the shell surface, with more pronounced spiral ridges in
70   the former, whereas the spiral ridges of *L. pellucidum* are only weakly defined.
71
72   The Cyclophoridae represents the most diverse Caenogastropoda family, but is remains
73   poorly resolved taxonomically. Delimitation among subgenera and species in Cyclophoridae has
74   long been a conundrum for taxonomists due to exceptionally diverse variation in morphology
75   (e.g. subgenera in genus *Cyclophorus* and *Alycaeus* (Kobelt, 1902; Gude, 1921)). Vermeulen
76   (1999) identified two major challenges with using shell characters to discriminate between six
77   species of Bornean *Leptopoma*. First, the majority of species are similar in shell morphology and
78   this limits the number of shell characters that can be used as diagnostic characters at species level.
79   Second, there appears to be a continuum of variation in some characters (i.e. size, shape, colour

80 patterns), particularly between *L. pellucidum* and *L. sericatum* and this contributes to uncertainty
81 when delimiting species. Thus it is clear that to date, the implications of the intra- and inter-
82 specific variation in shell morphology have not been studied systematically and comprehensively
83 in the context of the taxonomy of this genus.

84

85   This study had three aims. (1) To estimate the molecular phylogenetic relationship of
86 three similar yet polymorphic *Leptopoma* species in Sabah to investigate the monophyly of *L.*
87 *sericatum*, *L. pellucidum* and *L. atricapillum* based on two mitochondrial genes (16S RNA and
88 COI) and nuclear gene (ITS-1). (2) To test the phylogenetic signal of three qualitative and six
89 quantitative shell characters to evaluate their reliability as diagnostic characters. (3) To compare
90 the differences in shell characters of two *Leptopoma* species, *L. sericatum* and *L. pellucidum*, at
91 two locations, where they co-occur and are abundant to understand the geographical variations in
92 shell characters under consideration and further assess their reliability as diagnostic characters.

93

94 **MATERIALS AND METHODS**
95 All the *Leptopoma* specimens included in this study were obtained from the *BORNEENSIS*
96 Mollusca collection at the Institute of Tropical Biology and Conservation in Universiti Malaysia
97 Sabah. The collection houses more than 4000 specimens of *Leptopoma* spp. collected since 2000
98 from various locations in Sabah (Figure 1). From this comprehensive collection, 77 alcohol-
99 preserved specimens of four species (*L. sericatum*, *L. pellucidum*, *L. atricapillum*, *L. undatum*)
100 were selected for molecular analysis. 249 empty shells of adult snails of *L. sericatum* (114) and *L.*
101 *pellucidum* (135) from Balambangan Island and the Kinabatangan region, where both species
102 exist sympatrically were selected for morphological analysis (Supplementary File 1). These
103 *Leptopoma* specimens were identified as either *L. pellucidum* or *L. sericatum* on the basis of the
104 presence/absence of distinct spiral ridges on the shell (Vermeulen, 1999).

105

106 **Data Collection**

107

108 *Genetic Data*
109 Genomic DNA was extracted from a total of 77 individuals preserved in 70% ethanol, but
110 sequence data for at least two genes could only generated for 17 of these (Table S1 in
111 Supplementary File 2). DNA was extracted from foot tissue by using DNeasy extraction kit
112 (Qiagen Inc., Hilden, Germany) according to manufacturer instructions. We used two
113 mitochondrial genes, the protein coding COI (cytochrome oxidase 1) and the non-coding 16S
114 rRNA, and one nuclear gene ITS-1 (internal transcribed spacer 1). Universal primers LCO1490
115 (5'-GGTCAACAAATCATAAAGATATTGG-3')   and   HCO2198   (5'-
116 TAAACTTCAGGGTGACCAAAAAATCA-3') were used to amplify and sequence COI
117 (Folmer *et al*., 1994). The 16s rRNA region was amplified using primers 16Sar (5'-
118 CGCCTGTTTATCAAAAACAT-3') and 16Sbr (5'-CCGGTCTGAACTCAGATCACGT-3')
119 (Kessing *et al*., 1989). ITS-1 region was PCR-amplified using the primers 5.8c (5'-

120   GTGCGTTCGAAATGTCGATGTTCAA-3')   and   18d   (5'-
121   CACACCGCCCGTCGCTACTACCGATTG-3') (Hillis & Dixon, 1991). Thermal cycling was
122   performed with pre-denaturation at 90°C for 2 minutes, denaturation at 94°C for 45 seconds, one
123   minute of annealing at 55°C, 60°C, and 54°C for COI, 16s, and ITS-1 respectively, extension at
124   72°C for one minute followed by final extension at 72°C for 5 minutes. The denaturation,
125   annealing and extension steps were repeated for 35 cycles. The PCR products were sequenced at
126   Macrogen, Inc. (Korea). All sequences were subsequently uploaded and stored in Barcoding of
127   Life Database (BOLD, http://www.boldsystems.org, Ratnasingham & Hebert, 2007), under the
128   project title "*Leptopoma* in Sabah" (Code: LEPT).

129

130   *Shell Morphological Characters Data*
131   Shell form in this study included both quantitative (i.e size and shape) and qualitative (i.e colour
132   patterns, spiral ridges and presence/absence dark apertural band) shell characters. These
133   morphological characters were evaluated from the apertural view of the 264 shells examined in
134   total (i.e. the 249 dry specimens and 14 adult specimens included in the phylogenetic analysis).
135   First, high quality photographs were taken of the aperture of each shell with the aid of a Leica
136   Stereo Microscope M205. Five quantitative linear measurements, namely shell height (SH), shell
137   width (SW), aperture height (AH), aperture width (AW), and shell spire height (SpH) were then
138   measured directly from the photographs using Leica Application Suite software (Fig. 2A). A
139   sixth quantitative shell character – the ratio of shell height to width – was computed. Next, the
140   states for two qualitative shell characters (i.e. eight distinct shell colour patterns and
141   presence/absence of the dark apertural band) were recorded for each of the shells (Figs. 2C, 2D;
142   see Table S2 in Supplementary File 2 for the descriptions of the eight shell colour patterns).

143

144   **Data Analysis**

145

146   *Molecular Phylogenetic Analysis*
147   In addition to the sequences collected from 17 specimens in this study, 16S and COI sequences
148   of *L. tigris*, *L. vitreum* and an outgroup species - *Cyclophorus formosensis* (Nevill, 1882) from
149   Lee *et al*. (2008a) and Nantarat *et al*. (2014a) - were obtained from GenBank (Supplementary
150   File S1, Page 1: Table S1 for informations of specimens). All the DNA sequences were aligned
151   and checked manually using Bioedit v7.1.9 (Hall, 1999). In order to find the best-fit model of
152   substitution, jModelTest2 (Darriba *et al*., 2012) as implemented in CIPRES portal (Miller *et al*.,
153   2010) was performed based on corrected Akaike Information Criterion (AICc) for ITS-1
154   sequences, 16S sequences and each of the codon positions of COI sequences. Phylogenetic trees
155   were estimated using Maximum likelihood (ML) and Bayesian Inference methods (BI) as
156   implemented in CIPRES portal (Miller *et al*., 2010). ML analysis was conducted using Raxml-
157   HPC2 (Stamatakis, 2014) with 1000 rapid bootstraps. BI analysis was performed using MrBayes
158   v3.2.3 (Huelsen & Ronquist, 2001). This consisted of running four simultaneous chains for

159   100,000 generations and 10 sampling frequency. The first 250 trees were discarded as burn-in,
160   while the rest were used to obtain the final consensus tree.

161

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

183

184   *Geographical Variation in Shell Morphology Analysis*
185   Two-way ANOVA tests were performed to determine if there were differences in the six
186   quantitative shell characters between: (i) the two *Leptopoma* species (*L. pellucidum* and *L.
187   sericatum*), and (ii) the two locations (Balambangan Island and Kinabatangan). In addition, the
188   interaction effects of both factors (species and location) were tested. A Shapiro-Wilk test for
189   normality (Shapiro & Wilk, 1965), and a Levene's test (Brown & Forsythe, 1974) for
190   homogeneity of variance, revealed that some datasets were not normally distributed and showed
191   non-homogeneity of variances (Table S3 & S4 in Supplementary File 2). Nevertheless, two-way
192   ANOVA tests were still conducted since the deviations of these datasets from the ANOVA
193   assumption were considered not too serious, and the ANOVA was considered a robust test
194   against the normality assumption (Zar, 1999).

195

196   Chi-square two-way contingency table tests were performed to determine whether the
197   shell colour patterns and the presence/absence of a dark apertural band were associated with
198   species identity and location respectively. Prior to the analyses, four two-way contingency tables

199 were produced. These summarised frequencies as follows: (1) shell colour patterns *vs.* species, (2)
200 shell colour patterns *vs.* location, (3) dark apertural band *vs.* species, and (4) dark apertural band
201 *vs.* location. Each of the tables was analysed by using Pearson's Chi-squared test. When the
202 expected frequency in the contingency table was less than 5, a Fisher exact test was performed
203 instead of the Pearson's Chi-squared test (Bower, 2003). All the statistical analyses were
204 performed in the R statistical environment version 3.1.3 (R Core Team, 2015) with the
205 significant p-values set at 0.05. (R script in Supplementary File 3).

206
207 **RESULTS**

208
209 **The molecular phylogeny of the *Leptopoma* species in Sabah**
210 A total of 660 nucleotide sites were aligned for the COI gene, 558 nucleotide sites for the 16S
211 gene and 627 nucleotide sites for ITS-1 (Supplementary File 4). Gaps were treated as missing
212 data and were retained in all phylogenetic analyses. The aligned COI dataset consisted of 36.9%
213 GC content, 207 (31.4%) parsimony informative, and 253 (38.3%) variable sites. The aligned
214 16s gene had 33.3% GC content with 276 (49.8%) parsimony informative and 406 (73.3%)
215 variable sites. On the other hand, aligned ITS-1 gene had 48.6% GC content, 158 (25.2%)
216 parsimony informative, and 274 (43.7%) variable sites. Phylogenetic analyses were run for four
217 datasets: three single-gene datasets (ITS-1, COI, 16S) and a concatenated dataset of the three
218 genes. The tree was rooted on the outgroup *Cyclophorus formosensis* (Fig. 3).

219
220   Best-fitted models were selected based on the corrected Akaike Information Criterion
221 (AICc); the models were TPM3uf+G for ITS-1, TIM3+G for 16S, TIM3ef+G, TPM3uf+I, and
222 TPM3uf+G respectively for first, second and third codons of COI. These models were used in
223 both ML and BI analyses. Phylogenetic trees from the ML and BI analyses of the concatenated
224 dataset showed no conflict in tree topology. The monophyly of the three *Leptopoma* species (*L.
225 sericatum*, *L. pellucidum* and *L. atricapillum*) in Sabah was consistently strongly supported
226 (posterior probability of 100% and bootstrap support greater than 75).

227
228 **Phylogenetic signals in shell characters of the *Leptopoma* species**
229 Figure 4 shows the correlation between phylogeny and the quantitative and qualitative shell
230 characters for *L. sericatum*, *L. atricapillum* and *L. pellucidum*. A phylogenetic signal test based
231 on Pagel's λ and Blomberg's K showed that spiral ridges and the presence/absence of dark
232 apertural band represented a strong signal with $\lambda = 1$ and $K > 1$ (K= 4.536 for spiral ridges and
233 K= 1.114 for dark apertural band) (Table 1). However, shell colour pattern, a character that is
234 often used as a diagnostic character in traditional classification, showed a weak phylogenetic
235 signal (λ=0.997, K= 0.234). Among the quantitative shell characters, shell height exhibited a
236 strong signal according to Pagel's λ although Blomberg's K gave a weak phylogenetic signal.
237 The ratio of shell height to width (SH/SW) exhibited the weakest phylogenetic signal among all
238 the shell characters (λ= 0, K= 0.054).

239
240

Geographical Variation in Shell Forms

Two-way ANOVA showed that all shell quantitative characters (except aperture height) differed between the two locations (Table 2, Fig. 5). In addition, all shell quantitative characters except shell width and aperture height also differed between the two species. There was interactive effect of species and location on the aperture height, shell spire height and ratio of shell height to width.

Chi-square analyses indicated a significant association between the frequencies of shell colour patterns and both species identity (Fisher's exact test: p =0.0000) and location (Fisher's exact test: p=0.0000). On the other hand, there was a fairly significant association between the frequency of presence/absence of dark apertural band and species identity (Pearson's Chi-Squared with Yates' continuity correction: X^2 (1, N=249)=4.019, p=0.0449) but not with location (Pearson's Chi-Squared with Yates' continuity correction: X^2 (1, N=249)=1.5505, p=0.2131). Both contingency tables are available in Table S5 & S6 in Supplementary File 2. Overall, the shell characters considered in this study did not show consistent differences between *L. pellucidum* and *L. sericatum* since the differences in shell forms were coupled with geographical variations and interaction effects between geography and species.

DISCUSSION

This study presents the first molecular phylogeny investigation on genus *Leptopoma* in Sabah, one of the many understudied genera within the Cyclophoridae, and examines the concordance between morphology and phylogeny as well as geographical variation in shell form.

We found consistently significant support for the monophyly of the three morphologically similar Sabah *Leptopoma* species, *L. pellucidum*, *L. sericatum* and *L. atricapillum*. This finding is in concordance with the existing traditional morphology-based classification. For example, take the placement of *Leptopoma pellucidum* 6014 (Fig. 3) in our phylogeny, this population was previously assumed to be *L. vitreum* (Lesson, 1830) due to its uniformly white shell, but we have shown that it falls within the *L. pellucidum* clade, and this provides support to Vermeulen (1999)'s decision to assign *L. vitreum* from Sabah to *L. pellucidum*. In the case of *L. pellucidum* and *L. sericatum*, the presence of intermediate forms between the two species led Vermeulen to recognise them as being distinct on a purely provisional basis. In this study, results suggested that the two species could be unequivocally regarded as separate. In short, the findings of this study are in line with past research in this region which proposed that a combination of morphology and molecular approaches could improve taxonomy of land snails (Nantarat *et al*., 2014a, 2014b; Liew *et al*., 2009; Liew *et al*.,

277 2014). However, we aware that this findings still need to be verified with more specimens of
278 both species from their entire distribution range outside of Sabah.

279

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

289

290   The presence of a dark apertural band has not been observed in other cyclophorids and
291 was not mentioned in other revisionary works of *Leptopoma* species. Results showed that the
292 presence/absence of a dark apertural band exhibited a significant phylogenetic signal and does
293 not associate with geography for two of the species. Nevertheless, association of dark apertural
294 band with species identity was fairly significant. Our results showed that dark apertural band
295 only present in some *L. sericatum,* with this character being observed at just two sites (i.e.
296 Kinabatangan and the Tabin Wildlife Reserve area). However, when observing all the specimens
297 in the *BORNEENSIS* collection, we found that such character is actually present in both *L.*
298 *pellucidum* and *L. sericatum*. The apertural band of *L. pellucidum* was darker than *L. sericatum*.
299 We also observed that dark apertural band is generally present in shells with thickened outer lip
300 (gerontic shell) and a high abundance of shells that such character can be found in the Tabin
301 Wildlife Reserve area. This might indicate a longer life span of *Leptopoma* species in the area,
302 though this requires proper investigation. Overall, our findings here suggest that although this
303 character shows a strong phylogenetic signal, it is not an appropriate character for species-level
304 identification in *Leptopoma*.

305

306   Shell colour patterns are usually used as one of the key determinants to discriminate
307 between species in traditional taxonomic classification. One of the sister taxa of *Leptopoma*, the
308 species in genus *Cyclophorus*, was distinguished unambiguously based on shell patterns that
309 were also supported by molecular data (Nantarat *et al*., 2014b). Compared to genus *Cyclophorus*,
310 shell colour patterns of the *L. sericatum* and *L. pellucidum*, exhibited a weak phylogenetic signal.
311 This case of shell colour pattern polymorphisms of the two *Leptopoma* species is similar to other
312 well-known land snails such as *Cepaea nemoralis* and *C. hortensis* (Owen & Bengtson, 1972;
313 Ozgo & Schilthuizen, 2012; Cameron & Cook, 2012; Cameran, 2013). However, unlike *Cepaea*
314 land snails that have been studied extensively, the causal mechanism for the *Leptopoma* land
315 snail's diverse shell colour patterns is still unknown. This study also revealed that the *Leptopoma*
316 species in Sabah exhibits idiosyncratic differences between locations in the degree of shell

317 colour patterns polymorphism. For example, the *Leptopoma* population at Balambangan Island
318 has more shell colour patterns as compared to the population at Kinabatangan. As a result, the
319 geographically-associated variations in shell colour patterns and weak phylogenetic signal
320 strongly indicate that this character is unreliable as a diagnostic character for the identification of
321 the *Leptopoma* species considered in this study.

322

323   Significant inter- and intra-specific variations in quantitative shell characters have been
324 noted both in family Cyclophoridae (Lee *et al*., 2012; Nantarat *et al*., 2014b) and in other
325 gastropods (Kameda *et al*, 2007; Desouky & Busais, 2012; Hirano *et al.*, 2014). In Vermeulen's
326 (1999) descriptions of *L. sericatum* and *L. pellucidum*, the ratio of shell height to width of *L.*
327 *sericatum* is slightly smaller than *L. pellucidum*. However, intermediate in ratio of shell height to
328 width between the two species occur and lead to weak phylogenetic signal in this character.
329 From the phylogenetic signal test, only shell height produced a significant signal. Nevertheless,
330 this study revealed a high degree of geographical variations in the quantitative shell characters;
331 for example, both *Leptopoma* species from Balambangan Island were larger than the same
332 species found in Kinabatangan. Previous studies suggested that land snails found on islands tend
333 to undergo extensive morphological diversification (Johnson & Black, 2000; Stankowski, 2011).
334 In view of this, quantitative shell characters are thus not advisable as a diagnostic indicator to
335 delimit among these two *Leptopoma* species in Sabah due to the strong influence of geographical
336 variations.

337

338   While our findings reflected the reliability of 'hard' character (i.e. shell sculpture) over
339 colour patterns and banding, and revealed considerable geographical variation in some shell
340 characters, the findings from this study were based exclusively on *Leptopoma* species from
341 Sabah. For future work to improve the taxonomy of this genus, the study needs to be extended to
342 include larger numbers of specimens from a larger geographical area. In addition, more genetic
343 markers and examination of reproductive characters are required to elucidate comprehensive
344 phylogeny and morphological variation among the *Leptopoma* species

345

CONCLUSION

346
347 This study has revealed partial information on the phylogeny and morphology variations of all
348 *Leptopoma* species in their entire distribution range. Despite its small geographical scale, the
349 study has resolved taxonomic uncertainties of three *Leptopoma* species in Sabah and revealed
350 notable variations in both the quantitative and qualitative shell characters for the species. From
351 the findings, it is suggest that any future revisionary attempt of the taxonomy on the rest of *ca*.
352 100 *Leptopoma* species should consider the possible caveats in using the shell characters as the
353 sole evidence and should include molecular markers in the study. Further studies that include
354 more samples from a wider geographical reach are strongly recommended.

355 **ACKNOWLEDGEMENTS**

356 We would like to thank Cornelius Peter for his help with molecular lab work and Joumin
357 Rangkasan who assisted with fieldwork. This project was carried out by CCP and PSH as part of
358 their undergraduate honours degree dissertation. We thank Rudiger Bieler, Siong Kiat Tan and
359 an anonymous reviewer for their constructive comments.

360

SUPPLEMENTARY FILES

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

363 SUPPLEMENTARY FILE 2. Table S1. Specimen information; Table S2. Description of shell
364 colour patterns, Table S3 & S4. Results of the normality tests and homogeneity of variances tests
365 prior to ANOVA; Table S5 & S6. Frequency data for shell qualitative characters used in chi-
366 square tests.

367 SUPPLEMENTARY FILE 3. R script for shell morphological data and phylogenetic signal
368 analysis.

369 SUPPLEMENTARY FILE 4. FASTA file for the concatenated DNA sequences dataset for COI,
370 16S, and ITS-1.

REFERENCES

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

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

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

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

379 Cameron, R. A., & Cook, L. M. 2012. Habitat and the shell polymorphism of *Cepaea nemoralis*
380 (L.): interrogating the Evolution Megalab database. Journal of Molluscan Studies, 78(2), 179-
381 184.

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

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

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

390 Colgan, D. J., Ponder, W. F., Beacham, E., & Macaranas, J. 2007. Molecular phylogenetics of
391 Caenogastropoda (Gastropoda: Mollusca). Molecular Phylogenetics and Evolution, 42(3): 717-
392 737.

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

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

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

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

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

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

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

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

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

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

419 Huelsenbeck, J.P., Ronquist, F.R., Nielsen, R., Bollback, J.P., 2001. Bayesian inference of
420 phylogeny and its impact on evolutionary biology. Science 294, 2310–2314.

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

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

426 Kameda, Y., Kawakita, A., & Kato, M. 2007. Cryptic genetic divergence and associated
427 morphological differentiation in the arboreal land snail *Satsuma* (*Luchuhadra*) *largillierti*
428 (Camaenidae) endemic to the Ryukyu Archipelago, Japan. Molecular phylogenetics and
429 evolution, 45(2): 519-533.

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

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

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

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

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

440 Lesson, R. P. 1831. Mollusca. – In: voyage autour du monde, éxécuté par ordre du Roi, sur la
441 corvette de Sa Majesté, La Coquille, pendant les années 1822-1825. 2(1): 1-471. Paris.

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

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

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

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

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

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

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

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

464   Ożgo, M., & Schilthuizen, M. 2012. Evolutionary change in *Cepaea nemoralis* shell colour over
465   43 years. Global Change Biology, 18(1), 74-81.

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

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

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

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

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

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

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

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

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

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

485 Stankowski, S. 2011. Extreme, continuous variation in an island snail: local diversification and
486 association of shell form with the current environment. Biological Journal of the Linnean Society,
487 104(4): 756-769.

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

490 Vermeulen, J. J. 1999. Notes on the non-marine molluscs of the island of Borneo 9. The genera
491 *Cyclophorus*, *Leptopoma*, and *Craspedotropis* (Gastropoda Prosobranchia:
492 Cyclophoridae). Basteria, 63(4/6): 139-163.

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

494 Zilch, A., 1954. Die Typen und Typoide des Natur-Museums Senckenberg, 13 Mollusca,
495 Cyclophoridae, Cyclophorinae-Cyclophoreae (1). Archiv für Molluskenkunde 83:141-157.

496

497

Figure 1

Distribution map of four *Leptopoma* species in Sabah based on the records from *BORNEENSIS* Mollusca collection, Universiti Malaysia Sabah and localities of molecular samples.

Each circle represents a collection lot for a single *Leptopoma* species with the size of circle indicating the number of specimens in the lot (smallest circle = 1 individuals; largest circles = 140 individuals). The insets (A) and (B) show the sympatric populations of *L. sericatum* and *L. pellucidum* (i.e. on Balambangan Island and in Kinabatangan), which were used for shell morphological analysis. Orange triangles show localities of specimens for molecular study.

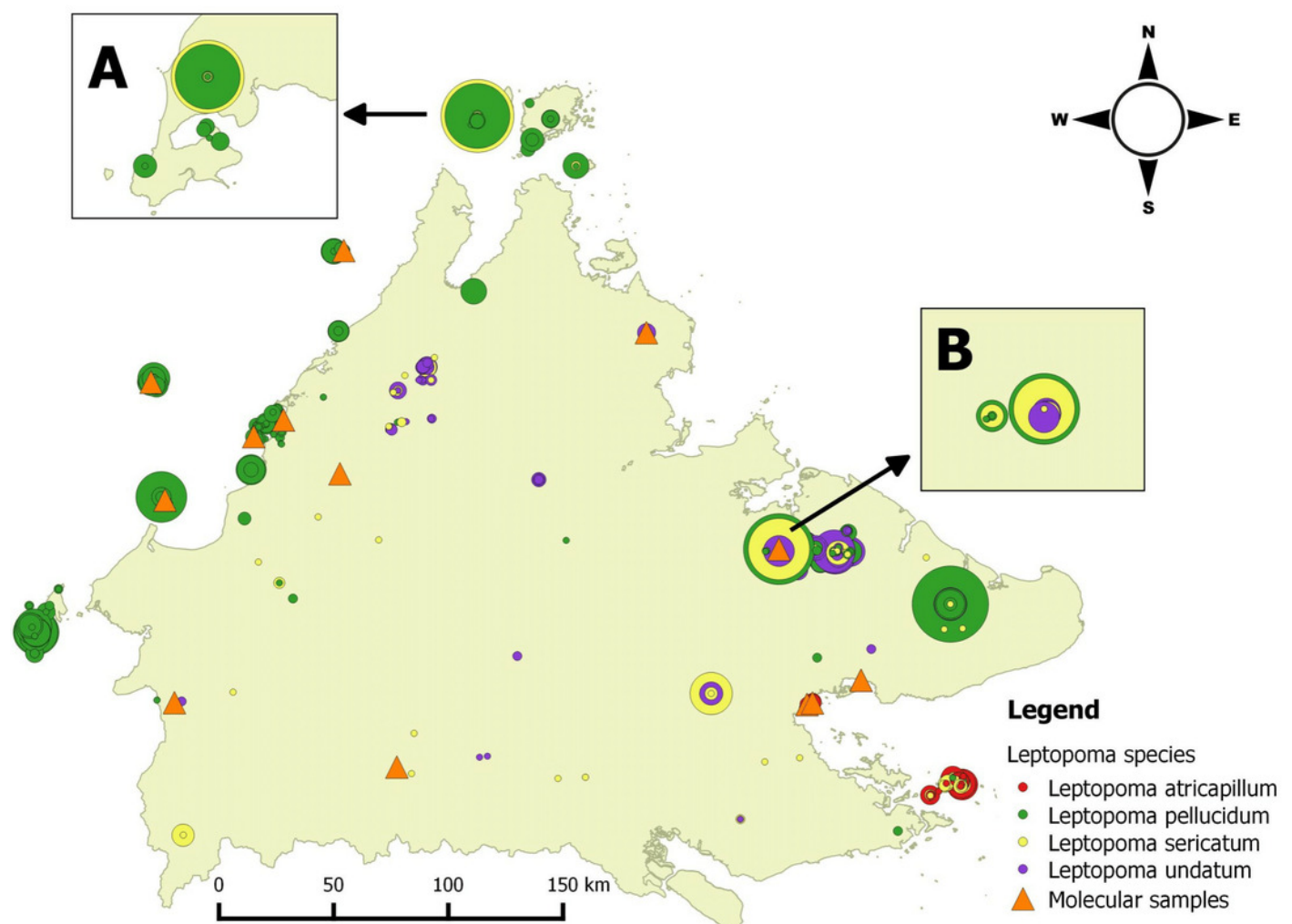



Figure 2

Qualitative and quantitative shell characters included in the study were assessed on the basis of the shell apertural 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 distinct shell colour patterns identified in the study. (D) Dark apertural band: Left – Presence, Right – Absence.

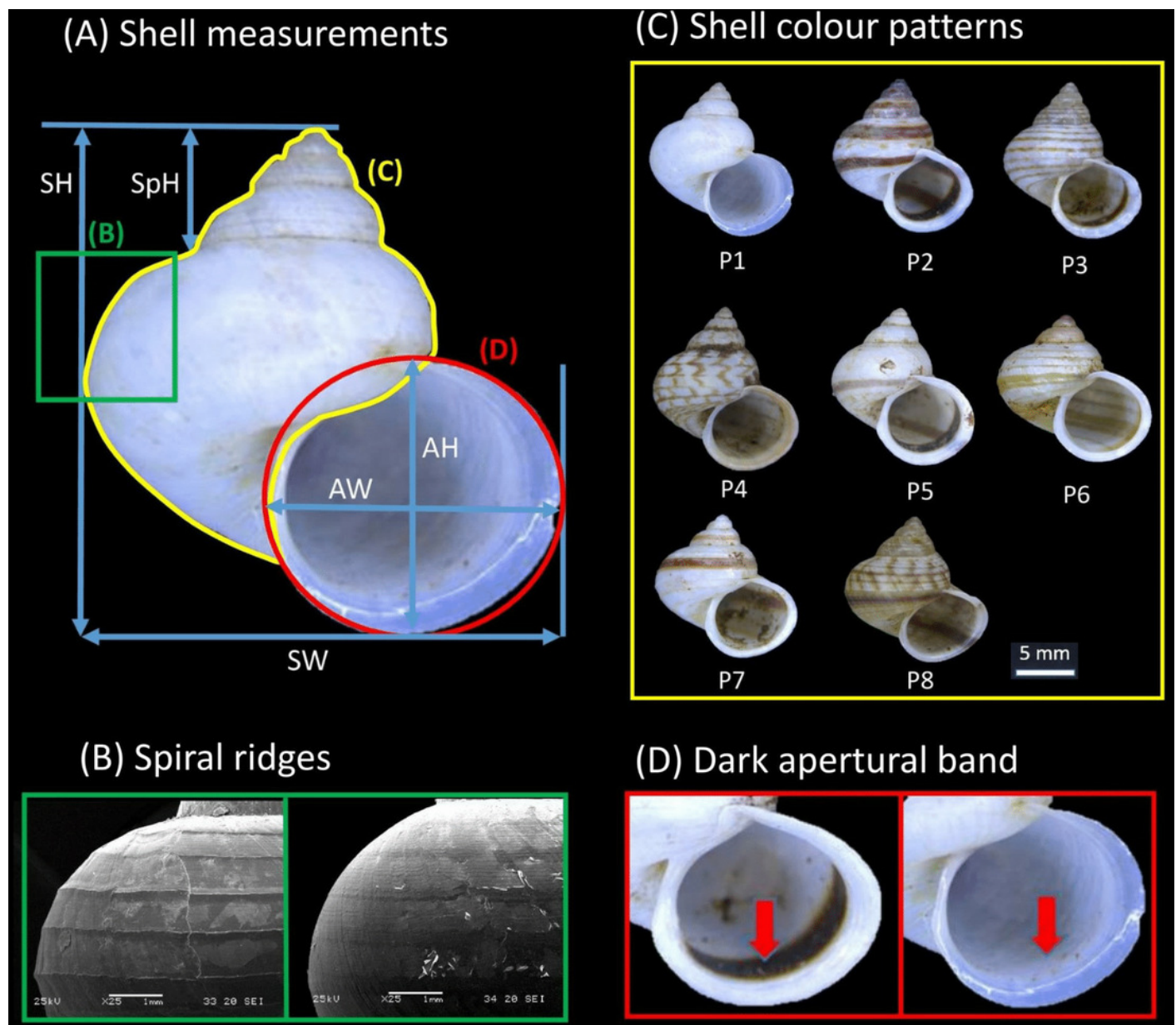



Figure 3

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

Support values on branches are Bayesian posterior probability (BI) followed by maximum likelihood (ML) bootstrap value. Internal branches with ML bootstrap value = 100% and PP value =100 were not represents in the figure. The number shown beside each specimen of Sabah *Leptopoma* is the relevant specimen number (Table S1 in Supplementary File 2), and the specimens with asterisk are non-Borneo's *Leptopoma* species (i.e sequences obtained from Genbank). Clades A, B and C indicate the three Sabah species of focal interest. Scale bar for branch length = 0.1 substitutions per site.

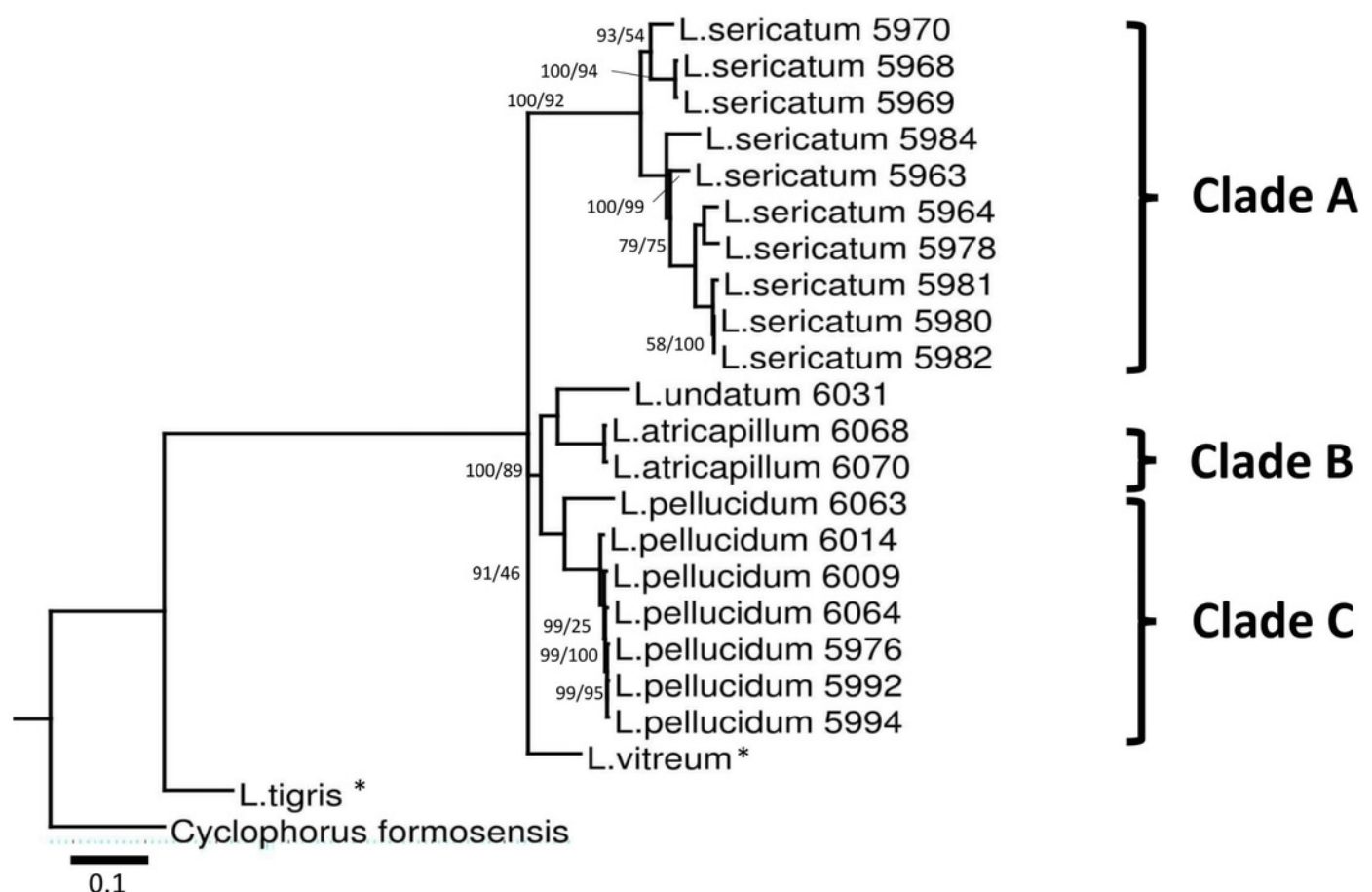



Figure 4

Shell quantitative and qualitative shell characters as mapped on to the phylogenetic tree.

Tree as in Figure 3 with the juvenile specimens were dropped from the tree and only the 14 adult specimens of the three *Leptopoma* species were retained. Different categories of the three qualitative shell characters: spiral ridges, shell colour patterns and dark apertural band (Figures 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 to width were represented by the size of the circle.

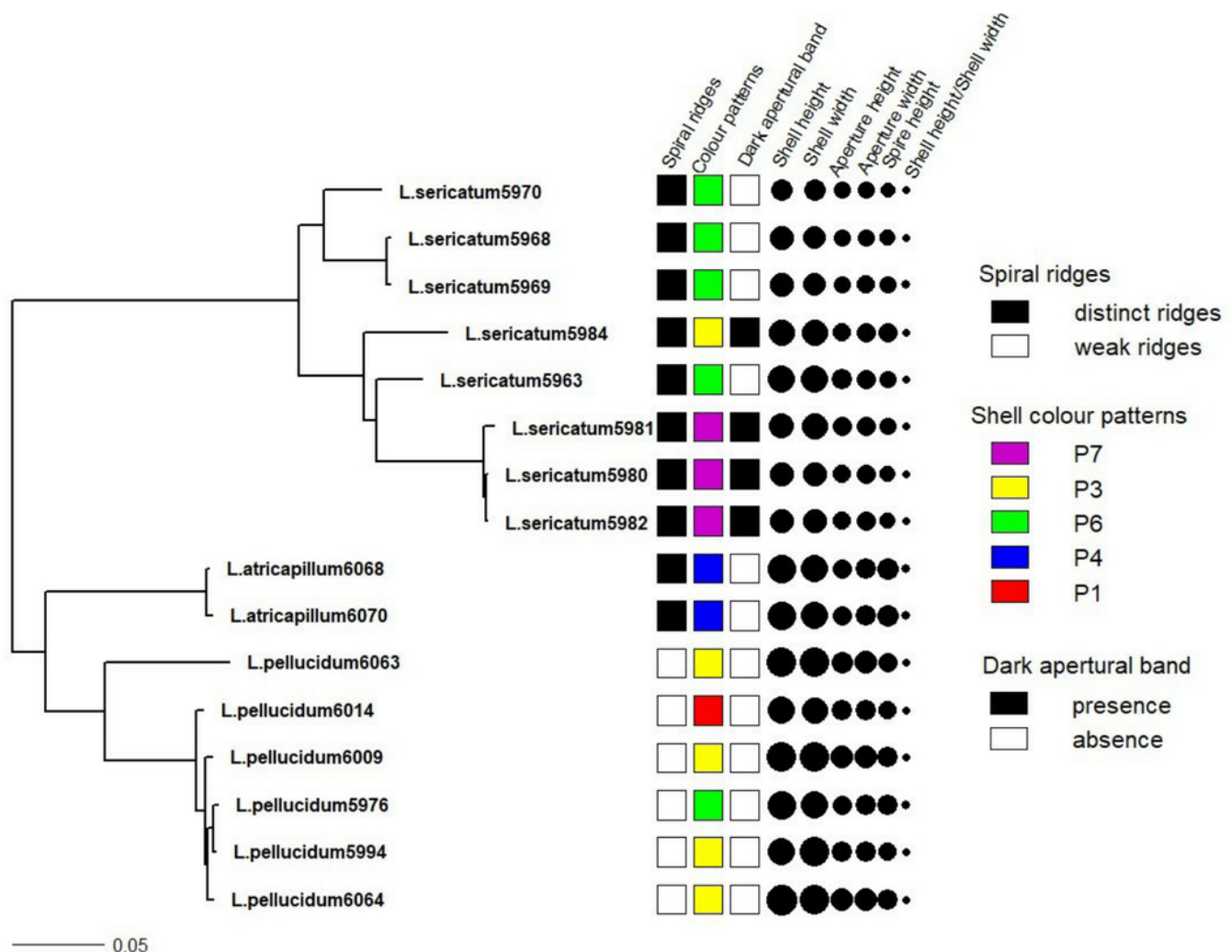
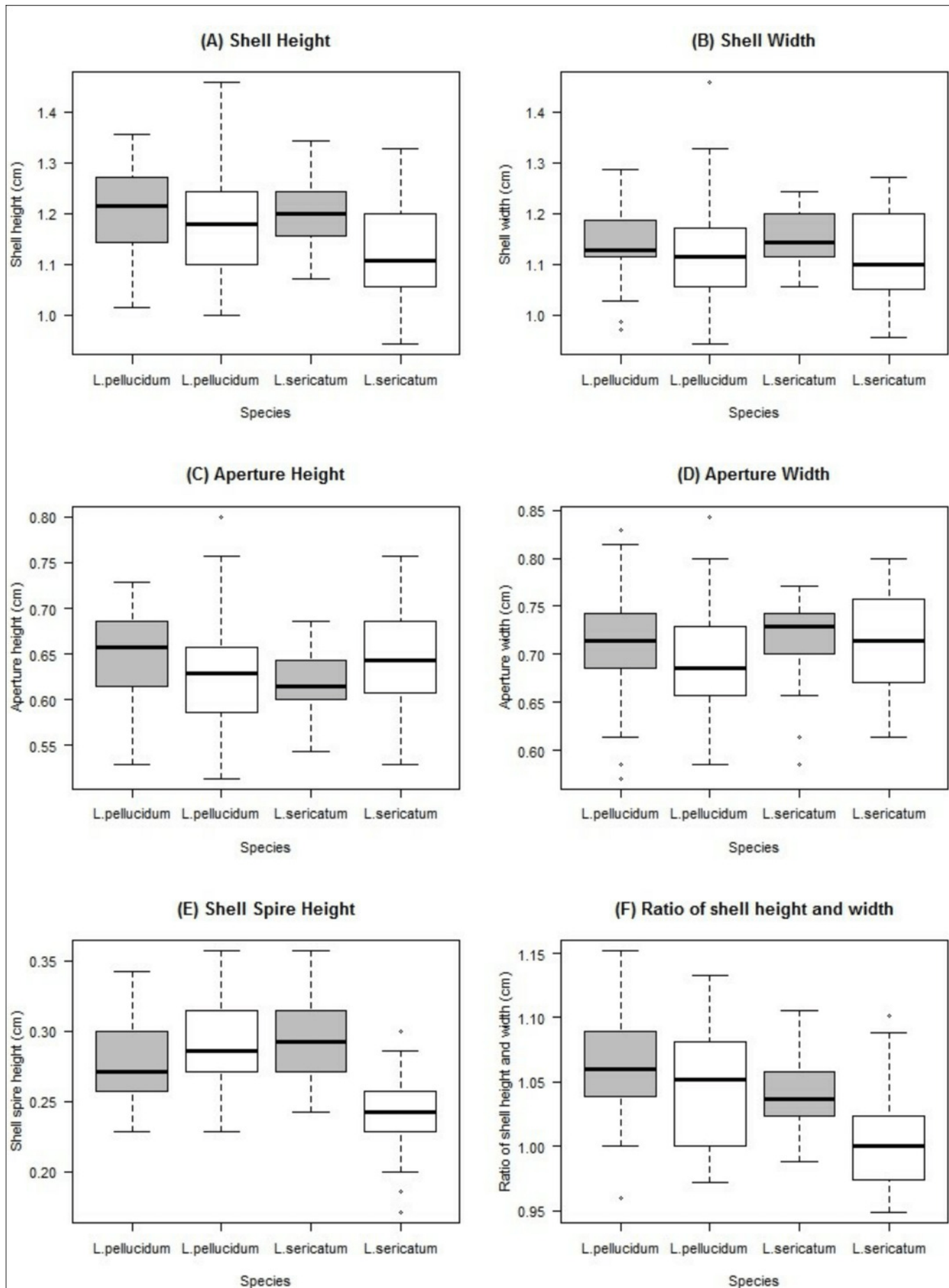



Figure 5

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).

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.

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.

	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.