A peer-reviewed version of this preprint was published in PeerJ on 31 October 2017.

View the peer-reviewed version (peerj.com/articles/3981), which is the preferred citable publication unless you specifically need to cite this preprint.

Phung C, Heng P, Liew T. 2017. Land snails of *Leptopoma* Pfeiffer, 1847 in Sabah, Northern Borneo (Caenogastropoda: Cyclophoridae): an analysis of molecular phylogeny and geographical variations in shell form. PeerJ 5:e3981 https://doi.org/10.7717/peerj.3981

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.

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, Pooi-San Heng, Thor-Seng Liew

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

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

Abstract. *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.

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

INTRODUCTION

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

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

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

Hence this study was conducted specifically to (1) estimate the molecular phylogeny of three similar yet polymorphic *Leptopoma* species in Sabah in order to investigate the monophyly of *L. sericatum*, *L. pellucidum* and *L. atricapillum* based on two mitochondrial genes (16S RNA and COI) and a nuclear gene (ITS-1), (2) test the phylogenetic signal of the shell morphological characters in terms of three qualitative shell characters and six quantitative shell measurements across the phylogenetic trees in order to evaluate their reliability as diagnostic characters, and (3) compare the differences in shell characters of two *Leptopoma* species namely *L. sericatum* and *L. pellucidum* at two locations where they were abundant and found sympatrically in order to understand the geographical variations in shell form and further assess their reliability as diagnostic characters. The results of this study supported the monophyly of the three *Leptopoma* species in line with the current classifications by Vermeulen (1999) although only the qualitative shell character such as spiral ridges was reliable for species delimitation.

MATERIALS AND METHODS

All the *Leptopoma* specimens included in this study were obtained from the BORNEENSIS mollusca collection at the Institute of Tropical Biology and Conservation in Universiti Malaysia Sabah. The collection housed more than 4000 specimens of *Leptopoma* spp. collected since 2000 from various locations in Sabah (Fig. 1). From this comprehensive collection, 77 wet specimens of four species (*L. sericatum*, *L. pellucidum*, *L. atricapillum*, *L.undatum*) were selected for molecular analysis. 249 empty shells of adult snails of *L. sericatum* (114) and *L. pellucidum* (135) from Balambangan Island and the Kinabatangan region, where both species existed sympatrically were selected for morphological analysis (Supplemetary File 1). These *Leptopoma* specimens were identified into either *L. pellucidum* or *L. sericatum* based on the spiral ridges on the shell (Vermeulen, 1999).

Data Collection

Genetic Data

Genomic DNA of 77 selected specimens stored in 70% ethanol was isolated from foot tissue by using DNeasy extraction kit (Qiagen Inc., Hilden, Germany) according to manufacturer instructions. Universal primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') were used to amplify and sequence mitochondrial cytochrome c oxidase subunit 1 (COI) (Folmer *et al*., 1994). 16s rRNA mitochondrial gene was amplified using primers 16Sar (5'-CGCCTGTTTATCAAAAACAT-3') and 16Sbr (5'-CCGGTCTGAACTCAGATCACGT-3') (Kessing *et al*., 1989). ITS-1 region was PCR-amplified using the primers 5.8c (5'-GTGCGTTCGAAATGTCGATGTTCAA-3') and 18d (5'-CACACCGCCCGTCGCTACTACCGATTG-3') (Hillis & Dixon, 1991). Thermal cycling was performed with pre-denaturation at 90°C for 2 minutes, denaturation at 94°C for 45 seconds, one minute of annealing at 55°C, 60°C, and 54°C for COI, 16s, and ITS-1 respectively, extension

step at 72°C for one minute followed by final extension at 72°C for 5 minutes. Denaturation, annealing and extension steps were repeated for 35 cycles. Positive PCR results were obtained from 17 out of 77 DNA extracts for at least two genes (Table S1 in Supplementary File 2). The PCR products were sequenced at Macrogen, Inc. (Korea). All sequences were subsequently uploaded and stored in Barcoding of Life Database (BOLD, http://www.boldsystems.org, Ratnasingham & Hebert, 2007), under the project title "*Leptopoma* in Sabah" (Code: LEPT).

Shell Morphological Characters Data

Quantitative and qualitative shell characters were obtained from the shell aperture view of all 249 *Leptopoma* specimens and 14 adult specimens included in the phylogenetic analysis. First, high quality photographs were taken of the aperture of each shell with the aid of a Leica Stereo Microscope M205. Five quantitative linear measurements, namely shell height (SH), shell width (SW), aperture height (AH), aperture width (AW), and shell spire height (SpH) were then taken directly from the photographs by using Leica Application Suite software (Fig. 2A). The sixth quantitative shell character – the ratio between shell height and width – was computed accordingly. Next, the states for the two qualitative shell characters (i.e. the eight types of shell colour patterns and presence of the dark ring band in aperture) were recorded for each of the shells (Figs. 2C, 2D; see Table S2 in Supplementary File 2 for the descriptions of the eight shell colour patterns).

Data Analysis

Molecular Phylogenetic Analysis

In addition to the sequences collected from 17 specimens in this study, 16S and COI sequences of *L. tigris*, *L. vitreum* and an outgroup species - *Cyclophorus formosensis* from Lee et al. (2008a) and Nantarat et al. (2014a) - were obtained from GenBank (File S1, Page 1: Table S1 for informations of specimens). All the DNA sequences were aligned and checked manually using Bioedit v7.1.9 (Hall, 1999). In order to find the best-fit model of substitution, jModelTest2 (Darriba *et al*., 2012) as implemented in CIPRES portal (Miller *et al*., 2010) was performed based on corrected Akaike Information Criterion (AICc) for ITS-1 sequences, 16S sequences and each of the codon positions of COI sequences. Phylogenetic trees were estimated by using Maximum likelihood (ML) and Bayesian Inference methods (BI) as implemented in CIPRES portal (Miller *et al*., 2010). Maximum likelihood analysis was conducted using Raxml-HPC2 (Stamatakis, 2014) with 100 rapid bootstraps. Bayesian Inference analysis was performed using MrBayes v3.2.3 (Huelsen & Ronquist, 2001) which consisted of running four simultaneous chains for 100,000 generations and 10 sampling frequency. The first 250 trees were discarded as burn-in, while the rest were used to obtain the final consensus tree.

Phylogenetic Signal Analysis

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

Geographical Variation in Shell Morphology Analysis

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

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

analyses were performed in R statistical environment version 3.1.3 (R Core Team, 2015) with the significant p-values set at 0.05 (R script in Supplementary File 3).

RESULTS

The molecular phylogeny of the *Leptopoma* species in Sabah

A total of 660 nucleotide sites were aligned for the COI gene, 558 nucleotide sites for the 16S gene and 627 nucleotide sites for ITS-1 (Supplementary File 4). The aligned COI dataset consisted of 36.9% GC content, 207 (31.4%) parsimony informative, and 253 (38.3%) variable sites. Aligned 16s gene had 33.3% GC content with 276 (49.8%) parsimony informative and 406 (73.3%) variable sites. On the other hand, aligned ITS-1 gene had 48.6% GC content, 158 (25.2%) parsimony informative, and 274 (43.7%) variable sites. Phylogenetic analyses were run for four datasets: ITS-1, COI, 16S and concatenated dataset of ITS-1, COI and 16S whereby gaps were treated as missing data. Outgroup *Cyclophorus formensis* was used to root the tree.

The best-fitted models selected based on corrected Akaike Information Criterion (AICc) were TPM3uf+G for ITS-1, TIM3+G for 16S, TIM3ef+G for COI first codon, TPM3uf+I for second codon, and TPM3uf+G for third codon in COI. These models were applied in both ML and BI analyses. Phylogenetic trees produced from both ML and BI based on concatenated dataset showed no conflict in tree topologies. Therefore, only the BI tree based on concatenated dataset was shown (Figure 3). The resultant phylogenetic tree supported the monophyly of three *Leptopoma* species (*L. sericatum*, *L. pellucidum* and *L. atricapillum*) in Sabah. Each major clade formed by each species was supported by significant supporting values (100 PP and ML bootstrap larger than 75%).

Phylogenetic signals relating to shell characters for the *Leptopoma* species

Figure 4 shows the inter-relation between phylogeny and the quantitative and qualitative shell characters for *L. sericatum*, *L. atricapillum* and *L. pellucidum*. A Phylogenetic signal test based on Pagel's λ and Blomberg's K showed that spiral ridges and presence of dark ring band in aperture represented a strong signal with $\lambda = 1$ and $K > 1$ ($K= 4.536$ for spiral ridges and $K= 1.114$ for dark ring band) (see Table 1). However, shell patterns that often used as a diagnostic character in traditional classification indicated a weak phylogenetic signal ($\lambda=0.997$, $K= 0.234$). Among the quantitative shell characters, shell height exhibited a strong signal according to Pagel's λ although Blomberg's K indicated a weak phylogenetic signal. The ratio of shell height to width (SH/SW) exhibited the weakest phylogenetic signal among all shell characters ($\lambda= 0$, $K= 0.054$).

Geographical Variation in Shell Morphology

Two-way ANOVA showed that all shell quantitative characters (except aperture height) differed between the two locations (Table 2). 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 between shell height and width.

Chi-square analyses indicated significant association between the frequencies of shell colour patterns and both the factors of species identity (Fisher's exact test: p =0.0000) and location (Fisher's exact test: p=0.0000). Similarly, there was significant association between the frequencies of the presence of the dark ring band in the aperture and both the factors of species identity (Pearson's Chi-Squared with Yates' continuity correction: X^2 (1, N=249) = 4.88, p=0.0271) and location (Pearson's Chi-Squared with Yates' continuity correction: X^2 (1, N=249) = 12.910, p=0.0003). Both contingency tables are available in Table S5 & S6 in Supplementary File 2. Overall, the shell characters did not show consistent differences between *L. pellucidum* and *L. sericatum* since the differences in shell form were coupled with geographical variations and interaction effects between geography and species.

DISCUSSION

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

All the phylogenetic trees based on different genes were congruent and provided significant support for the monophyly of three morphologically similar *Leptopoma* species in Sabah namely *L. pellucidum*, *L. sericatum* and *L. atricapillum*. The phylogenetic placement of genus *Leptopoma* in Sabah was in concordance with its traditional morphology-based classification. For example, the placement of *Leptopoma pellucidum* 6014 (Fig. 3) in this study, previously assumed as *L. vitreum* due to its white colour shell which differed from other *L. pellucidum*, was revealed as within the *L. pellucidum* clade which supported Vermeulen (1999)'s decision to assign *L. vitreum* as synonymous to *L. pellucidum*. In the case of *L. pellucidum* and *L. sericatum*, Vermeulen separated them into two species provisionally due to the existence of intermediate forms between the two species. 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 which proposed that a combination of morphology and molecular approaches could improve taxonomy of land snails.

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

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

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

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

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

CONCLUSION

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

ACKNOWLEDGEMENTS

We would like to thank Cornelius Peter for his help in DNA works and Joumin Rangkasan who assisted in fieldwork. This project was carried by CCP and PSH as part of their undergraduate honours degree dissertation. This research is supported by Universiti Malaysia Sabah through grant no. SLB0107-STWN-2015 to TSL and MyBrainSc scholarship to CCP.

SUPPLEMENTARY FILES

SUPPLEMENTARY FILE 1. Raw data for shell morphological analysis.

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

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

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

REFERENCES

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

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

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

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

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

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

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

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

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

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. 2012. jModelTest 2: more models, new 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*

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

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

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

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

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

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

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

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

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

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

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

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

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

Owen, D. F., & Bengtson, S. A. 1972. Polymorphism in the land snail *Cepaea hortensis* in Iceland. 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.

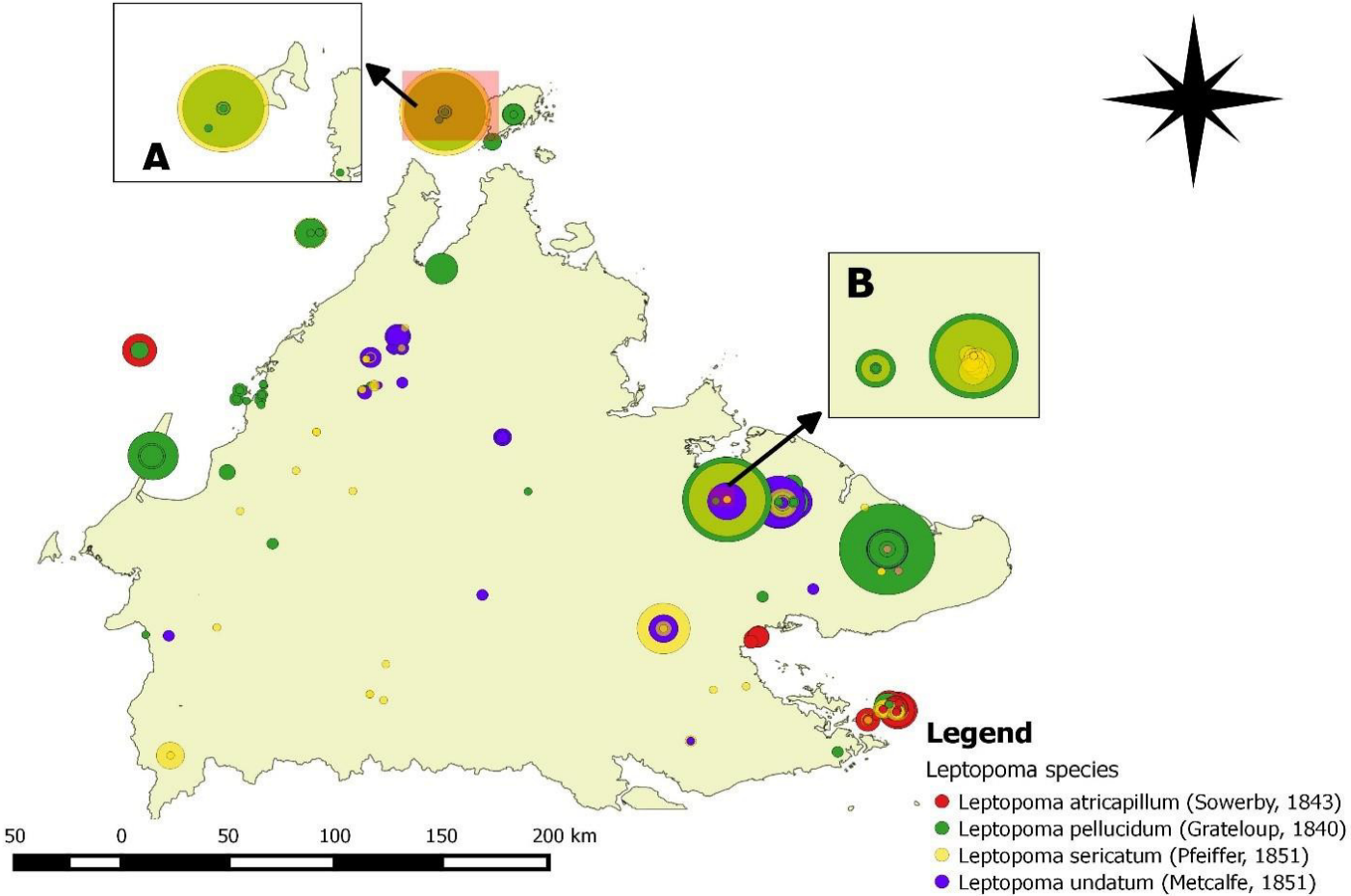
479 Zar, J.H .1999. Biostatistical Analysis, 4th ed., New Jersey: Prentice-Hall Inc

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.



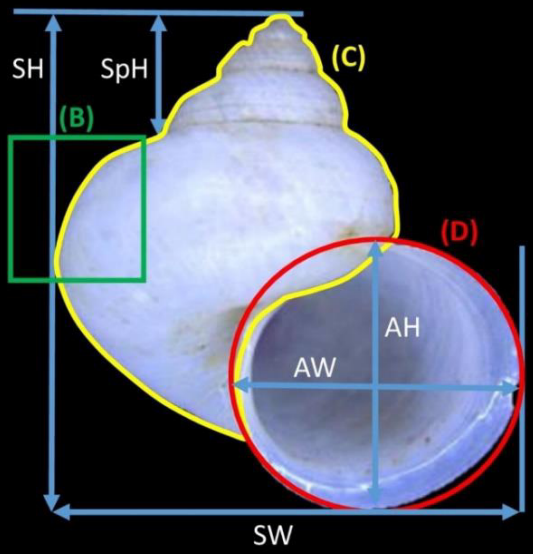
Peerj Preprints | https://doi.org/10.7287/peerj.preprints.3078v1 | CC BY 4.0 Open Access | rec: 8 Jul 2017, publ: 8 Jul 2017

Figure 2(on next page)

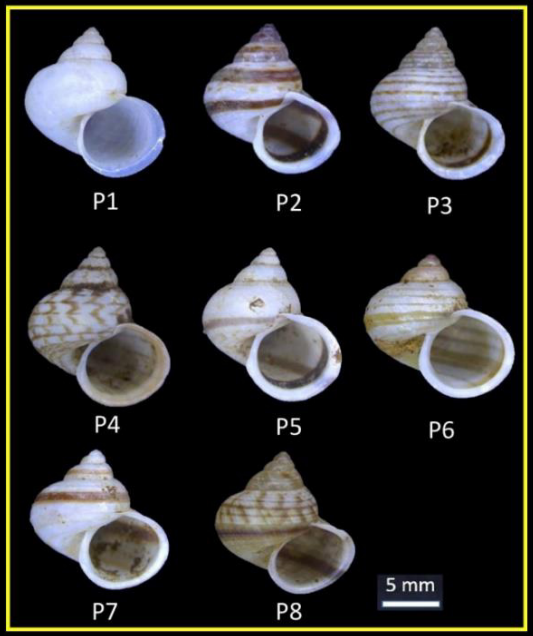
Qualitative and quantitative shell traits obtained from shell aperture view.

(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.

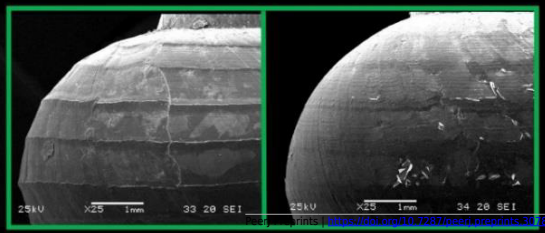
(A) Shell measurements



(C) Shell colour patterns



(B) Spiral ridges

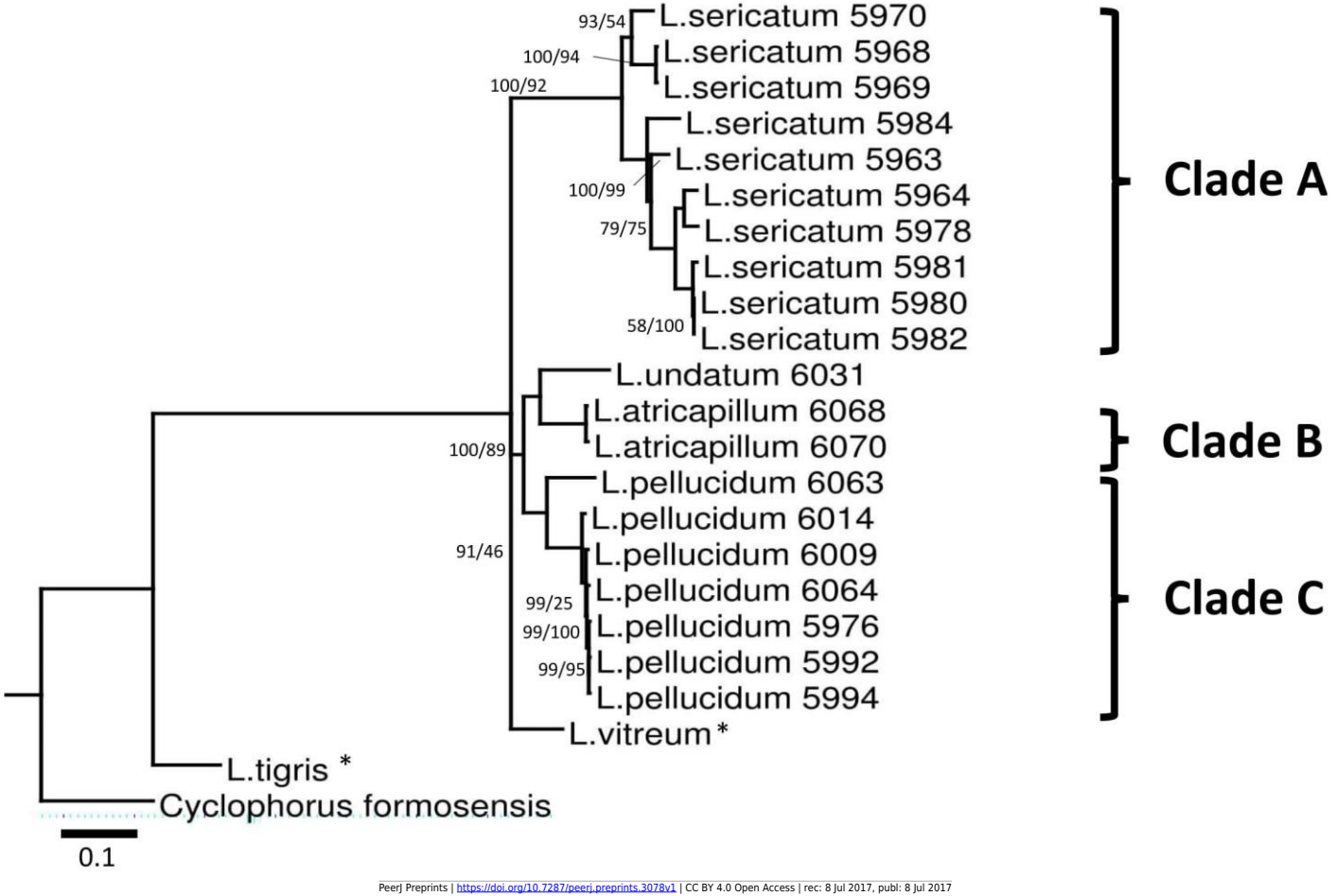


(D) Dark ring band in aperture

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.

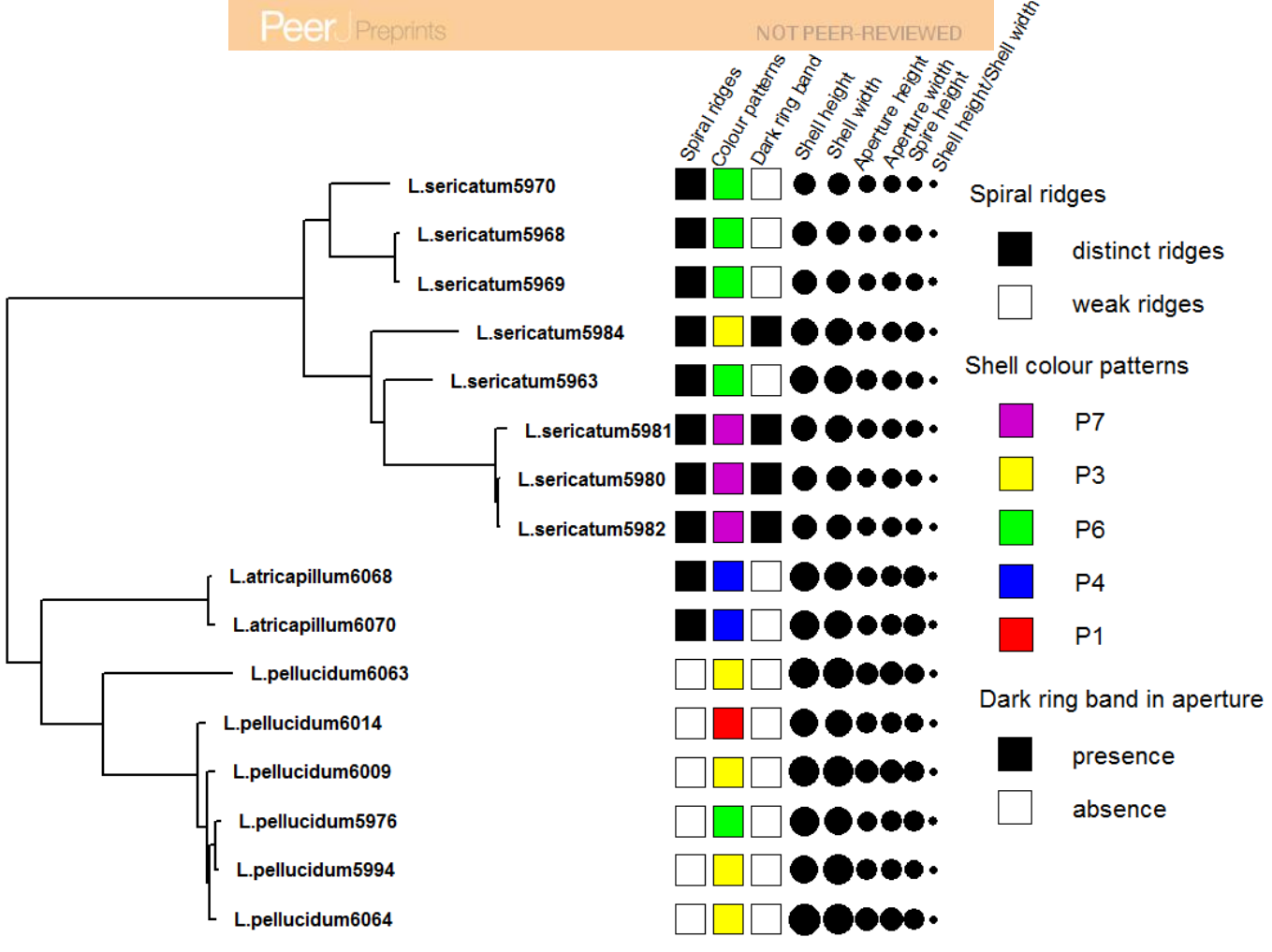


Peerj Preprints | https://doi.org/10.7287/peerj.preprints.3078v1 | CC BY 4.0 Open Access | rec: 8 Jul 2017, publ: 8 Jul 2017

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.

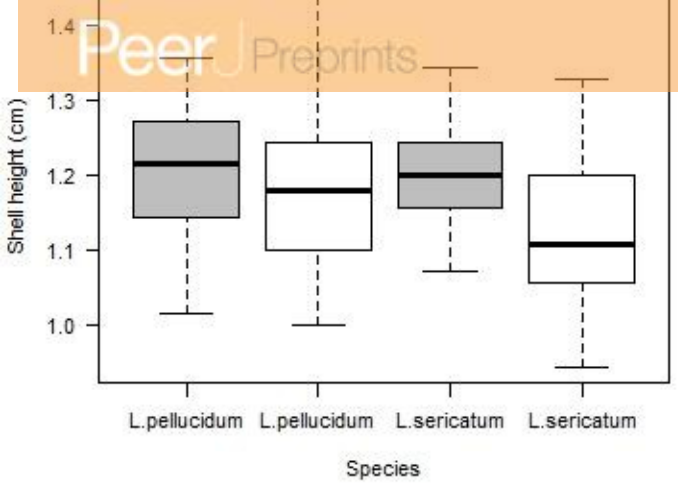


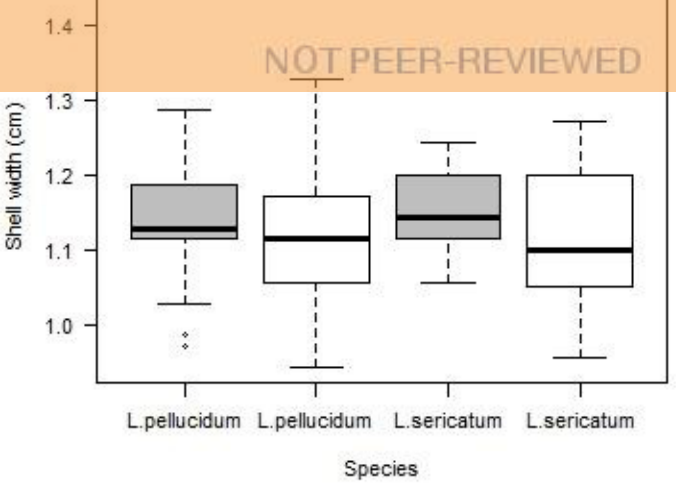
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3078v1 | CC BY 4.0 Open Access | rec: 8 Jul 2017, publ: 8 Jul 2017

Figure 5(on next page)

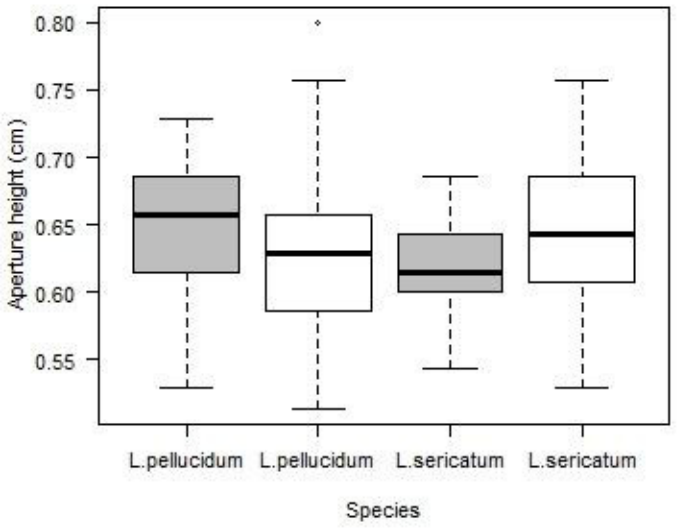
Boxplots show the differences of the six quantitative measurements of shell for the *Leptopoma pellucidum* and *L. sericatum* in each of the two locations (Balambangan Island and Kinabatangan region).

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).

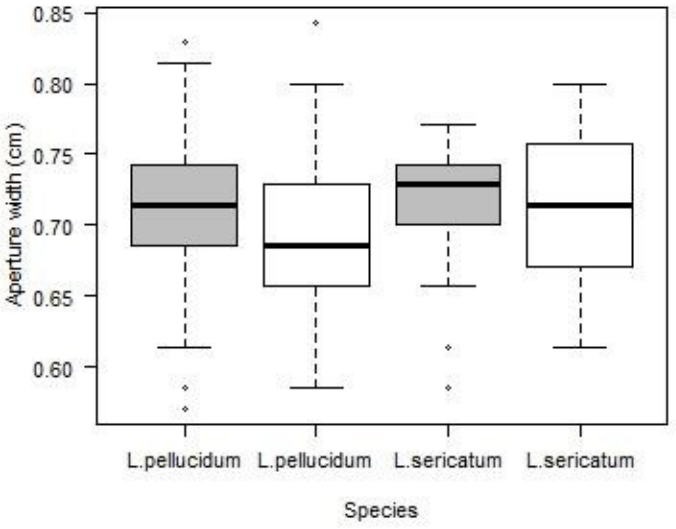




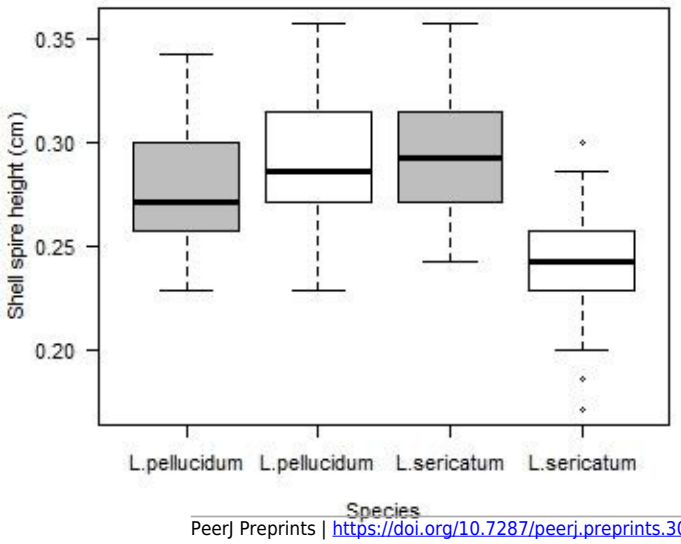
(C) Aperture Height



(D) Aperture Width



(E) Shell Spire Height



(F) Ratio of shell height and width

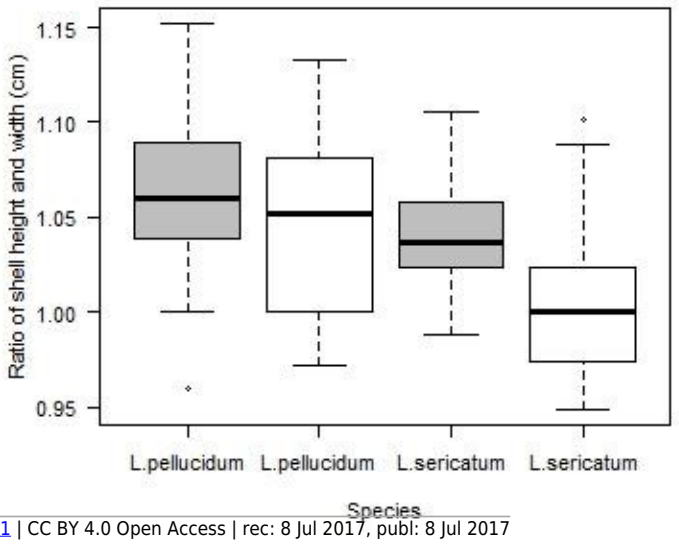


Table 1(on next page)

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.

Table 1. 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.

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

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.

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.

Table 2. Two-way ANOVA for the effect of geographical variation and species identity on six quantitative 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**

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.