Results

1. Statistic result under dataset partitioning strategy I

In total 60 data matrices were obtained from published journal articles and individual authors (Table 1). Thirty of them were assorted to category A, 7 were assorted to category B, 7 were placed in category C, 7 were sorted in category D and 5 were placed in category E. The proportions of colour/pattern versus non-colour/pattern characters were significant in three of the five categories (category A: $P < 0.0001$**, r = 0.3271, number of pairs = 29; category B: $P = 0.0039$**, r = - 0.4937, number of pairs = 9; category C: $P = 0.0625$, r = 0.1539, number of pairs = 5; category D: $P = 0.0420$*, r = 0.2597, number of pairs = 11 and category E: $P = 0.6250$, r = - 0.800, number of pairs = 5). There is no significant difference in proportion of the studies having higher colour/pattern RI values between each category ($\chi^2 = 6.025$, d.f. = 4, $P = 0.1973$) (Table 2). In Areekul & Quicke's (2005) study, the proportion between colour/pattern characters and non-colour/pattern characters (termed A/M and morphology characters in their study) is not significant either in their "A/M" or "morphology" partition.

Differential weighting of colour/pattern and non-colour/pattern characters in the category A (without aposematic and mimic attributions) displayed a notable asymmetry in the ratio of taxa remained in ASTs, and the colour/pattern partition seemed to not offer strong signals (Fig. 3), that might generate conflicting phylogenetic signals with those based on the non-colour/pattern characters. The asymmetry in category A was

significant when the weighted value was over 5 x. Besides category A, the other asymmetrical pattern only occurred in category D at the weighted value with 1.5x, but not appearing in higher weighing scheme (Table 3). In category B, C and E, there was no significant difference between up-weighting colour/pattern and morphological characters even with the most extreme weighting (1000x) (Fig. 3, Table 3).

Comparing the proportions of taxa retained in the agreement subtrees between five categories at each relative weighting scheme, there was significant different when up-weighting non-colour/pattern or colour/pattern characters higher than 5x (see Table 4), and the *P* value of up-weighting non-colour/pattern characters at 2x was relatively small (*P* = 0.0667) even it was not quite significant. The groups with significant differences were further examined by Dunn's multiple comparison test as a post-test to find out the exactly two-category combinations which caused the significant differences. The result showed that the combination of "category A versus B" and "B versus C" were significant combinations while the *P* value is smaller than 0.05 in the comparison at non-colour/pattern up-weighting schemes, whereas "B versus E" was the only significant combination while the *P* value was smaller than 0.05 in the comparison at colour/pattern up-weighting schemes.

2. Statistic result under dataset partitioning strategy II

In this treatment, the proportions of colour/pattern versus non-colour/pattern characters was significant in category C+D+E (*P* =

0.0101*, r = 0.2700, # of pairs = 21). Ratios of the taxa remained in agreement subtrees from equally and differentially weighted analyses for category C+D+E is showed in figure 5. The plotting results of category C+D+E (Fig. 5) showed a more symmetric and gradual structure than that of category A (see Fig. 4(a)). There was no significant difference in the proportions of taxa retained in agreement subtrees with different weighting between non-colour/pattern and colour/pattern characters within category C+D+E (see Table 5). However, significant difference of what was detected when comparing the proportions of taxa retained in agreement subtrees between 3 categories at each relative weighting scheme with up-weighting non-colour/pattern characters from more than 1.5x to 5x (see Table 6).