Phylogenetic relationships and genetic diversity of the *Polypedates leucomystax* complex in Thailand

Kittisak Buddhachat [Corresp., 1] , **Chatmongkon Suwannapoom** [2]

1 Department of Biology, Naresuan University, Phitsanulok, Thailand

2 School of Agriculture and Natural Resources, University of Phayao, Phayao, Thailand

Corresponding Author: Kittisak Buddhachat
Email address: kittisakbu@nu.ac.th

Taxonomic uncertainty of the Asian tree frog *Polypedates leucomystax* complex presents the challenging task of inferring its biogeographical history. Here, we describe its dispersion and the genetic relationships among different populations in Thailand, where we connect the population of the *P. leucomystax* complex of the Sunda Islands to the Indochina (mainland) population based on analyses of 266 sequences of the mitochondrial cytochrome c oxidase subunit I (*COI*) gene. Our maternal genealogy implies that there are four well-supported lineages in Thailand, consisting of Northern A (clade A: *Polypedates* sp.), Nan (clade B: *P.* cf. *impresus*), Southern (clade C: *P.* cf. *leucomystax*) and Northern B (clade D: *P.* cf. *megacephalus*), with Bayesian posterior probability >0.9. Phylogeny and haplotype networks indicate that clades A, B and D are sympatric. In contrast, clade C (*P.* cf. *leucomystax*) and clade D (*P.* cf. *megacephalus*) are genetically divergent due to the geographical barrier of the Isthmus of Kra, resulting in an allopatric distribution. Climatic conditions, in particular differences in rainfall on each side of the Isthmus of Kra, may play an important role in limiting the immigration of both clades. For the within-populations of either clades C or D, there was no significant correlation between geographic and genetic distance by the isolation-by-distance test, indicating intraspecific-dispersal of each clade. Population expansion occurred in clade C, whereas clade D showed a constant population. Taken together, the *P. leucomystax* complex in South East Asia may have diversified under climatic pressure, leading to allopatric and/or sympatric speciation.

1   **Phylogenetic relationships and genetic diversity of the *Polypedates leucomystax* complex in**

2   **Thailand**

3   Kittisak Buddhachat[1,*] and Chatmongkon Suwannapoom[2,*]

4   [1] Department of Biology, Faculty of Science, Naresuan University, Phitsanulok, Thailand
5   [2] School of Agriculture and Natural Resources, University of Phayao, Phayao, Thailand
6   [*] These authors contributed equally to this work.

7   *Corresponding authors
8   Kittisak Buddhachat
9   Email address: kittisakbu@nu.ac.th, k_buddhachat@yahoo.com
10   Chatmongkon Suwannapoom
11   Email address: chatmongkonup@gmail.com

12   **ABSTRACT**

13   Taxonomic uncertainty of the Asian tree frog *Polypedates leucomystax* complex presents the

14   challenging task of inferring its biogeographical history. Here, we describe its dispersion and the

15   genetic relationships among different populations in Thailand, where we connect the population

16   of the *P. leucomystax* complex of the Sunda Islands to the Indochina (mainland) population based

17   on analyses of 266 sequences of the mitochondrial cytochrome c oxidase subunit I (*COI*) gene.

18   Our maternal genealogy implies that there are four well-supported lineages in Thailand,

19   consisting of Northern A (clade A: *Polypedates* sp.), Nan (clade B: *P*. cf. *impresus*), Southern

20   (clade C: *P*. cf. *leucomystax*) and Northern B (clade D: *P*. cf. *megacephalus*), with Bayesian

21   posterior probability >0.9. Phylogeny and haplotype networks indicate that clades A, B and D are

22   sympatric. In contrast, clade C (*P*. cf. *leucomystax*) and clade D (*P*. cf. *megacephalus*) are

23   genetically divergent due to the geographical barrier of the Isthmus of Kra, resulting in an

24   allopatric distribution. Climatic conditions, in particular differences in rainfall on each side of the

25   Isthmus of Kra, may play an important role in limiting the immigration of both clades. For the

26   within-populations of either clades C or D, there was no significant correlation between

27   geographic and genetic distance by the isolation-by-distance test, indicating intraspecific-

28   dispersal of each clade. Population expansion occurred in clade C, whereas clade D showed a

29   constant population. Taken together, the *P. leucomystax* complex in South East Asia may have

30   diversified under climatic pressure, leading to allopatric and/or sympatric speciation.

INTRODUCTION

Southeast Asia contains a substantial genetic diversity of amphibians. Recent molecular phylogenetic analyses have disclosed many anuran lineages that contain cryptic species. Historically, complex changes in the region's geology and climate (e.g., Pleistocene climatic oscillations) altered the topology and environmental conditions, resulting in an initial fragmentation of habitat. These mechanisms generated the high species richness observed today (Hall, 1998; Woodruff, 2010). Of interest to our research were the numerous frog species in Southeast Asia whose taxonomy is still controversial, such as *Microhyla fissipes* (Yuan et al., 2016), *Staurois tuberilinguis* (Matsui et al., 2007), *Microhyla ornata* (Matsui et al., 2005) and *Polypedates leucomystax* (Kuraishi et al., 2013; Rujirawan, Stuart & Aowphol, 2013). Accurate species delimitation is essential to better understand their speciation and diversification and their biogeography for conservation purposes.

The Asian tree frog, the *Polypedates leucomystax* (Gravenhorst, 1829) complex, is an Asian Rhacophoridae frog. These species are widely distributed in Southeast Asia, South China and India. In addition, this species has phenotypic plasticity and exhibits high adaptation to the local environment, leading to its existence in diverse habitats such as forests and even buildings. These high levels of phenotypic plasticity present a great challenge for classification. Phylogenetic and taxonomic relationships of the *P. leucomystax* complex throughout Southeast Asia exhibit adaptive radiation (Kuraishi et al., 2013; Pan et al., 2013; Rujirawan, Stuart & Aowphol, 2013). At least six valid species, including *P. braueri*, *P. leucomystax*, *P. macrotis*, *P. megacephalus*, *P. mutus* and *P. impresus*, have been delimited from the *P. leucomystax* complex based on their morphology, advertisement calls and molecular data (Matsui, Seto & Utsunomiya, 1986; Brown et al., 2010; Kuraishi et al., 2011; Kuraishi et al., 2013; Pan et al., 2013). Five

54   species, *P. leucomystax*, *P. mutus*, *P. macrotis*, *P. colletti* and *P. megacephalus*, can be found in

55   Thailand (Taylor, 1962; Heyer, 1971; Frost, 2013; Kuraishi et al., 2013; Pan et al., 2013;

56   Rujirawan, Stuart & Aowphol, 2013). A study by Brown et al. (2010) indicated that much of the

57   genetic divergence of the *P. leucomystax* complex was discovered in mainland rather than insular

58   populations distributed throughout thousands of islands of the Malay Archipelago, presumably

59   resulting from range expansion mediated by transportation of agricultural products. Recently, a

60   new species, *P. discantus*, belonging to the *P. leucomystax* species complex from southern

61   Thailand was discovered using data on morphological characteristics, advertisement calls and

62   molecular evidence, which showed that this species was highly dissimilar to *P. leucomystax* and

63   *P. megacephalus* (Rujirawan, Stuart & Aowphol, 2013). Several studies have confirmed the

64   existence of highly cryptic species of the *Polypedates leucomystax* complex (Matsui, Seto &

65   Utsunomiya, 1986; Kuraishi et al., 2011; Blair et al., 2013; Kuraishi et al., 2013; Pan et al.,

66   2013).

67   Understanding the phylogenetic relationships among species can give insight into how

68   lineages diverged and how new species arose. The process of speciation can be understood based

69   on the geographic overlap of emerging species during divergence. In this study, we investigated

70   the genetic variation, phylogenetic relationships and other relevant factors that limit the dispersal

71   of the *P. leucomystax* complex in Thailand. The present results illustrate the range of distribution

72   of putative *P. leucomystax* and putative *P. megacephalus*, which is influenced by climatic

73   conditions.

74   **MATERIALS AND METHODS**

75   **Sample collection, DNA extraction and sequencing**

76   In this study, a total of 266 adult *Polypedates leucomystax* complex individuals were collected

77   from 15 different localities in Thailand (Table 1). All samples were dissected to obtain the liver,

78   which was then stored in absolute ethanol. Sample collection and euthanization were approved by

79   the Center For Animal Research Naresuan University under project number NU-AE591028.

80   Genomic DNA was extracted from liver tissue using a DNA extraction kit (RBC Bioscience,

81   Singapore) and kept at −20 ℃ for further use. Individual DNA was used as a template for PCR

82   amplification of the mitochondrial COI gene using Taq DNA polymerase in a total volume of 25

83   µL under the following conditions: an initial denaturation at 94 ℃ for 5 min, followed by 35–40

84   cycles at 94 ℃ for 30 s, 50 ℃ for 30 s and 72 ℃ for 1 min, and a final extension step at 72 ℃ for

85   7 min. PCR products were visualized on 1.5% agarose gel under UV illuminator. The expected

86   size of the partial mitochondrial *COI* gene sequence was 688 bp. Subsequently, all PCR products

87   were purified using a QIAquick PCR Purification Kit (Qiagen, Germany) and then sequenced

88   (Macrogen, South Korea).

89   **Phylogeny**

90   Bayesian inference (BI) and maximum likelihood (ML) were independently employed for

91   constructing a phylogenetic tree based on the following partial *COI* sequences of the *Polypedates*

92   *leucomystax* complex that were retrieved from GenBank: *P. impresus*: KP996822 (China),

93   KP996846 (China), KP087862-70 (Laos); *P. leucomystax*: KR087871-2 (Thailand); and *P.*

94   *megacephalus*: KR087879, KR087881 (Thailand). First, the best-fit model of DNA sequence

95   evolution for this locus was identified with the Akaike information criterion (AIC) implemented

96   in MrModeltest v2.3 (Nylander, 2004), resulting in the GTR+I+G model as the best fit with AIC

97   Subsequently, a Bayesian tree was constructed based on the base substitution calculated from the

98   GTR+I+G model through MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003) with two independent

99   searches with random starting trees for 5 million generations, in which the diagnostic was

100   calculated every 1,000 generations and compared using four Markov chain Monte Carlo chains

101   (temp = 0.2). The log-likelihood scores were used for plotting the convergence in Tracer v1.5

102   (Rambaut et al., 2013) and building a consensus tree, which was completed by the removal of the

103   first 25% of the generations from each run. For maximum likelihood analysis, RAxML 7.0.4 was

104   carried out (Stamatakis, Hoover & Rougemont, 2008) using the GTR+I+G model for nucleotide

105   substitution (same as the BI analysis) with 1,000 bootstrap replicates.

106   **Population genetics and structure**

107   A total of 266 sequences of mitochondrial *COI* were aligned using ClustalW (implemented in

108   MEGA 6.0 with default parameters). The number of polymorphic sites, the parsimony-

109   informative sites, singleton sites, the number of haplotypes, haplotype diversity (H_d), and

110   nucleotide diversity for each clade were calculated using DnaSP v5.0 (Librado & Rozas, 2009).

111   Genetic distances among taxa were calculated using the corrected *p*-distance model in MEGA 6.0

112   (Tamura et al., 2011). Furthermore, we detected a boundary line in the genetic landscape between

113   the Northern B clade and the Southern clade using Barrier 2.2 (Manni, Guérard & Heyer, 2004).

114   A minimum spanning network was constructed using PopART (Population Analysis with

115   Reticulate Trees) population genetics software to define the relationships among haplotypes and

116   the distribution of haplotypes in each locality (Bandelt, Forster & Röhl, 1999). To evaluate the

117   effect of geographic distance on the genetic divergence among populations of the Northern B

118   clade and among populations of the Southern clade, a linear regression model was carried out.

119 **Demographic history**

120 To investigate the demographic history of *P. megacephalus* and *P. leucomystax* populations in

121 Thailand, multiple approaches were explored using DnaSP (Librado & Rozas, 2009). Neutrality

122 tests of Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997) for the two species were completed. A

123 significantly positive value indicates a process of subdivision or a recent population bottleneck,

124 whereas a population expansion results in a significantly negative value. Pairwise mismatch

125 distribution was used assuming a constant population size (Roger & Harpendin, 1992).

126 Multimodal mismatch distribution implies stability of the population, while unimodal mismatch

127 distribution reflects an expanding population. In addition to these methods, the raggedness index

128 (r_g) of the observed distribution was calculated (Harpending, 1994). A small r_g indicates a

129 demographic expansion.

130 **RESULTS**

131 **Sequence characteristics**

132 A total of 266 samples of the *P. leucomystax* complex yielded 688 bp fragments of the

133 mitochondrial *COI* gene. All new sequences in this study were deposited in the GenBank

134 database (MG583020-MG583285). After multiple alignment of all *COI* sequences, the sequences

135 were trimmed to the same length (437 bp) before downstream analysis. We observed 82

136 polymorphic sites, which are also 82 parsimony-informative sites without a singleton site,

137 resulting in the acquisition of 15 haplotypes (Table 2). Overall nucleotide and haplotype diversity

138 were 0.0664 and 0.9000, respectively (Table 2).

Phylogenetic analyses and haplotype distribution

140 Based on 266 mitochondrial *COI* sequences of the *P. leucomystax* complex, a matrilineal

141 genealogy was generated, and our results indicated that the *P. leucomystax* complex in Thailand

142 could consist of four clades: clade A (Northern A), *Polypedates* sp.; clade B (Nan), *P.* cf.

143 *impresus*; clade C (Southern), *P.* cf. *leucomystax*; and clade D (Northern B), *P.* cf. *megacephalus*

144 (Fig. 1). With respect to phylogenetic inference, clade A was treated as a sister group of clade B,

145 which was found in Nan. Clade A, however, can be seen in genetic samples obtained from the

146 Kanchanaburi (KCB), Mae Hong Son (MHS) and Phetchaburi (PCB) provinces and shared a

147 habitat with clade D, which was recognized as *P.* cf. *megacephalus*; its distribution range was in

148 the far north of the Isthmus of Kra at Chiang Mai (CM), MHS, KCB, PCB, Saraburi (SRB), Loei

149 (LPR), Nakhon Ratchasima (NRS) and Prachuap Khiri Khan (PKK). The dispersal areas of the

150 clade C population, as represented by *P.* cf. *leucomystax*, included Chumphon (CP), Nakhon Si

151 Thammarat (NST), Phuket (PK) and Ranong (RN), which are south of the Isthmus of Kra (Fig.

152 2). Barrier 2.2 was employed to determine a barrier for immigration between clade C and clade D

153 populations based on the dataset of genetic distance (Kimura's two-parameter model). Likely, the

154 Isthmus of Kra (IOK) represents a significant barrier to restrict immigration based on the great

155 genetic distance that was noted around IOK (Fig. 3)

156 A minimum spanning network among the mitochondrial haplotypes was also constructed

157 as shown in Fig. 1. The haplotypes of each group showed unique features, and each group had a

158 different number of haplotypes. Clade D exhibited the highest number of haplotypes at seven (H_d

159 = 0.746), followed by clade C with five haplotypes (H_d = 0.7526) (Table 2). Haplotypes A and B,

160 seen in clade A, and haplotype C found in clade B were unique haplotypes. Haplotypes D–I were

161 noted in clade D, while populations of northern, western and upper southern Thailand (KCB,

162 PCB and PKK, respectively) shared haplotype J. We noted that the haplotypes F and G of clade

163   D1 in NRS were highly divergent from the congeners. Clade C (*P*. cf. *leucomystax*) had high

164   haplotype diversity and contained three unique haplotypes (M, N and O) and two shared

165   haplotypes (K and L).

166   Analyses of the linear regression model between the genetic distance of the mitochondrial

167   *COI* gene sequence and the geographical distance found no significant isolation-by-distance

168   effect among populations of clade C (*P*. cf. *leucomystax*) and clade D (*P*. cf. *megacephalus*) (Fig.

169   4).

170   **Demographic history**

171   When we defined a significant barrier around the Isthmus of Kra (IOK) leading to the genetic

172   divergence between clade C (*P*. cf. *leucomystax*) and clade D (*P*. cf. *magacephalus*), neutrality

173   tests (Tajima's *D* and Fu's *Fs*) of both species were not significantly positive, whereas Fu's *Fs* of

174   clade C was significantly positive (Table 3). Furthermore, the mismatch distribution was tested as

175   a result of a left-skewed multimodal mismatch distribution for clade D with moderate r_g (0.2031)

176   but a unimodal mismatch distribution for clade C with a low r_g (0.0569) (Fig. 5). Overall, these

177   results suggested a constant population size of clade D and a population expansion of clade C.

178   **DISCUSSION**

179   The taxonomy of the Asian tree frog of the *Polypedates leucomystax* complex is contentious due

180   to the species' widespread distribution from Nepal to South East Asia and similar morphologies.

181   To better understand the population structure and biogeography of the *P. leucomystax* complex in

182   Thailand, where there is a substantial area for their genetic dispersal, the *COI* mitochondrial gene

183   sequences of these species were analysed. Our matrilineal genealogy implied four well-supported

184   lineages, consisting of a Northern B clade (clade D), a Southern clade (clade C), a Nan clade

185   (clade B) and a Northern A clade (clade A); based on their sequences and distributions, they

186   might be treated as *P. megacephalus*, *P. leucomystax*, *P. impresus* and *Polypedates* sp.,

187   respectively. However, our study did not perform morphological, call advertisement, ecological

188   habitat, and nuclear DNA analyses for species justification; therefore, it is difficult to confirm the

189   phylogenies with certainty. These would need to be investigated in future work.

190   Based on our phylogenetic tree, Northern A and Nan seemed to share the latest common

191   ancestor with the Southern clade; however, they showed a significant genetic difference against

192   Northern B. The matrilineal haplotype network showed that haplotype C of the Nan clade

193   (putative *P. impresus*) seemed to be related to haplotype N and O of the Southern clade.

194   Moreover, haplotypes A and B within the Northern A clade (*Polypedate* sp.) seem to have

195   evolved from haplotype C. These results indicate that both the Nan clade and the Northern A

196   clade are more closely related to the Southern clade (restricted in south of IOK) than the Northern

197   B clade (restricted in north of IOK) and that the Nan and Northern A clades may even be

198   sympatric. Although mtDNA seemed to offer cleanly split clades with a unique haplotype in each

199   clade, mtDNA introgession/hybridization between the closely related species can occur. This can

200   occur particularly in sympatric populations or at the boundaries of species distributions resulting

201   from incomplete reproductive isolation. Therefore, the use of only mitochondria DNA data is not

202   adequate for species justification. For instance, some populations of European newts *Triturus*

203   *montandoni* have an mtDNA haplotype of *Triturus vulgaris* in an area where their distributions

204   are connected, most likely as a result of historical or ongoing hybridization and multiple

205   introgression of mtDNA from *T. vulgaris* to *T. montandoni* (Babik et al., 2005). Thus, nuclear

206   DNA must be required for species delimitation to explain the lineage history with accuracy and

207   reliability. A recent study by Kuraishi et al., (2013) found one sample of *P. leucomystax* complex

208   from Indochina shared nuclear BDNF and Rag1 haplotypes with one sample from the south of

209   IOK, possibly suggesting deep coalescence or incomplete lineage sorting.

210   Bayesian inference strongly supported the geographic distinction of species in the

211   Northern B clade and the Southern clade. Furthermore, Monmonier's algorithm suggested that

212   the Isthmus of Kra (IOK; located between 11 and 13 °N along the Thai Peninsula) forms a

213   significant region separating them as evidenced by the large genetic divergence between the

214   populations on either side of the isthmus. Kuraishi et al. (2013) explored the population structure

215   of *P. leucomystax* complex in South East Asia and revealed that *P. megacephalus* is restricted to

216   Indochina (the northern Thailand, Laos, Vietnam and southern China) (Kuraishi et al., 2013; Pan

217   et al., 2013), while the expansion range of *P. leucomystax* is restricted to the Sunda regions

218   (southern Thailand and Malay Archipelago) (Brown et al., 2010; Kuraishi et al., 2013). Based on

219   their distribution, we assumed that the northern B clade and the Southern clade might be *P.*

220   *megacephalus* and *P. leucomystax*, respectively. The Isthmus of Kra has been well-characterized

221   as a remarkable biogeographic boundary of both faunal and floral assemblages that are limited on

222   either side. Bruyn et al (2005) used genetic evidence from *Macrobrachium rosenbergii*

223   (freshwater shrimp with a broad distribution) to suggest a hypothesis that the existence of a

224   seaway across the IOK over 1 Myr contributed to its geographically discontinuous distribution. In

225   contrast, Hughes et al. (2011), based on marine fossils and geological data, proposed that the

226   occurrence of species with distinct distributions in the north and south of the IOK is related to the

227   influence of the climatic zone rather than either marine inundation or breaches at any point near

228   the IOK. In addition, flying animals such as birds and bats also have a restricted species

229   distribution which is bounded by the IOK (Hughes et al.,2011). This indicates that the

230   geophysical barriers at the IOK, such as the waterway, are unlikely to be significant

231   biogeographical boundaries causing vicariance of various species at the IOK. A study by Hughes

232   et al. (2011) also determined the climatic zone in Indochina and the Malay Peninsula is the result

233   of four climatic zonations, which are displayed as similar climatic zones found in each time

234   period from the Last Interglacial Period to the current conditions. Of interest is the climatic

235   divergence in precipitation, temperature and seasonality between the central zone (north of IOK)

236   and the southern zone (south of IOK). Additionally, data on rainfall in Thailand by the Thai

237   Meteorological Department indicates a difference in the amount of rainfall between areas to the

238   north and south of the isthmus (Fig. 3). Based on this, we believe that climate may be a

239   significant factor in shaping the spatial distribution of the genetic lineage between the Northern

240   and Southern clades. Moreover, the matrilineal haplotype network provides strong evidence for

241   the suppression of migration between the Northern and Southern clades (Fig. 1). The haplotype

242   network also shows that the Northern B clade and the Southern clade have the highest genetic

243   diversity and a wide range distribution in Thailand, although the most recent divergence time

244   between *P. megacephalus* (putative Northern B clade) and *P. leucomystax* (putative Southern

245   clade) was estimated to be in the late Pliocene or early Pleistocene (1.4–4.0 MYBP). We assumed

246   that the populations of the Southern clade (putative *P. leucomystax*) expanded and colonized the

247   northern part of Indochina, leading to the diversification of the species. This hypothesis may be

248   partially supported by the results of the unimodal pairwise difference and the small value of the

249   raggedness index in the Southern clade (Fig. 4).

250   A previous study revealed the range expansion of insular populations of *P. leucomystax*,

251   characterized by small genetic distances among islands of the Malay Archipelago, as they can

252   migrate from one island to another through anthropogenic means, particularly through

253   transportation of agricultural products (Brown et al., 2010). Therefore, it was postulated that

254   climatic conditions might be the key barrier restricting the distribution range of the modern frog

255   populations. One hypothesis to support the limit of both Northern A and southern populations

256   may result from the adaptation in biological features such as behaviour, reproductive timing or

257   specific niche for a specific climatic zone having a notable difference in different zones,

258   according to Hughes et al. (2011). A previous study revealed that the variation of breeding season

259   length of the *P. leucomystax* complex of Thailand (putative *P. megacephalus*; 6 months a year for

260   breeding) and Singapore (putative *P. leucomystax*; every month) led to an adaptation to shortened

261   breeding season length by increasing clutch size (Sheridan, 2009). It is clear that climate can

262   influence the alteration in phenotypes of frogs (Sheridan, 2009; Todd et al., 2011).

263   Within the population of the northern B clade, maternal genealogy demonstrated that the

264   genetic samples from Nakhon Ratchasima province (NRS) seemed to represent a naturally

265   occurring divergence because of the emergence of endemic haplotypes; however, it was a low-

266   supported lineage, with 0.7 Bayesian posterior probability (BPP). When we considered the

267   topography of this region, the population of NRS as clade D1 is partitioned from the other

268   populations within clade D by the two mountain ranges, Dong Phaya Yen and Sankamphaeng.

269   This may be a possible barrier to gene flow among the modern populations of Northern B

270   between eastern (clade D1) and western (clade D2; NRS) Thailand. Unfortunately, only a limited

271   number of populations from the east of Thailand were investigated in this study. Further work for

272   validating whether Dong Phaya Yen and Sankamphaeng are a great barrier to gene flow in *P.*

273   *megacephalus* is required. According to the demographic history, the population of Northern B

274   was a stable population, but the population of Southern clade showed a relatively similar

275   unimodal distribution with a small raggedness index, possibly indicating a population expansion.

276   This result was similar to that observed in the population of *P. leucomystax* in the northern

277   Philippines (Brown et al., 2010). This scenario implied a genetically homogenous population,

278   especially in the population of Phuket Island, which shared a haplotype with NST, probably

279 caused by a recent population expansion due to the founder effect. Although the population of the

280 Northern B clade expanded, it was limited to localities south of the Isthmus of Kra.

281 **CONCLUSIONS**

282 Our matrilineal genealogy of the *Polypedates leucomystax* complex in Thailand suggested four

283 lineages, i.e., Nan (putative *P. impresus*), Northern B (putative *P. megacephalus*), Southern

284 (putative *P. leucomystax*) and Northern A (*Polypedates* sp.) clades. We noted that the populations

285 of the Northern B, Nan and *Polypedates* sp. clades are in sympatry, while their distributions are

286 allopatric to the southern clade (*P. leucomystax*) due to the separation by the Isthmus of Kra.

287 Climatic conditions may be a major contributor to limited migration of the current populations of

288 both clades, but climatic oscillation in the Pliocene and Pleistocene is a highly possible scenario

289 that drove speciation resulting in diversification of the *P. leucomystax* complex in Southeast Asia

290 and China, which includes the divergence of the southern and northern clades in Thailand.

291 **ACKNOWLEDGEMENTS**

292 Many thanks to Mr. Roy Morien of the Naresuan University Language Centre for his editing

293 assistance and advice on English usage.

294 **REFERENCES**

295   **Bandelt H-J, Forster P, Röhl A. 1999.** Median-joining networks for inferring intraspecific

296   phylogenies. *Molecular Biology and Evolution* **16:**37–48.

297   **Babik W, Branicki W, Crnobrnja-Isailović J, Cogălniceanu D, Sas I, Olgun K, Poyarkov**

298   **NA. Garcia-Paris M, Arntzen, JW. 2005.** Phylogeography of two European newt

299   species—discordance between mtDNA and morphology. Molecular Ecology **14:** 2475-

300   2491.

301   **Blair C, Davy CM, Ngo A, Orlov NL, Shi HT, Lu SQ, Gao L, Rao DQ, Murphy RW. 2013.**

302   Genealogy and demographic history of a widespread amphibian throughout Indochina.

303   *Journal of Heredity* **104:**72–85.

304   **Brown RM, Linkem CW, Siler CD, Sukumaran J, Esselstyn JA, Diesmos AC, Iskandar DT,**

305   **Bickford D, Evans BJ, McGuire JA, Grismer L, Supriatna J, Andayani N. 2010.**

306   Phylogeography and historical demography of *Polypedates leucomystax* in the islands of

307   Indonesia and the Philippines: evidence for recent human-mediated range expansion?

308   *Molecular Phylogenetics and Evolution* **57:**598–619.

309   **de Bruyn M, Nugroho E, Hossain MM, Wilson JC, Mather PB. 2005.** Phylogeographic

310   evidence for the existence of an ancient biogeographic barrier: the Isthmus of Kra

311   Seaway. *Heredity* **94:** 370-378.

312   **Frost DR. 2013.** Amphibian species of the world: an online reference. Version 6.0. New York:

313   American Museum of Natural History. [electronic database] *Available at*

314   *http://research.amnh.org/herpetology/amphibia/index.html* (accessed 22 June 2017).

315   **Fu YX. 1997.** Statistical tests of neutrality of mutations against population growth, hitchhiking

316   and background selection. *Genetics* **147**: 915-925.

317   **Hall R. 1998.** The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. In:

318   Hall R, Holloway JD, eds. *Biogeography and geological evolution of SE Asia*. Leiden,

319   The Netherlands: Backhuys Publishers, 99–131.

320   **Harpending HC. 1994.** Signature of ancient population growth in a low-resolution

321   mitochondrial DNA mismatch distribution. *Human Biology* 591-600.

322   **Heyer WR. 1971.** *Mating calls of some frogs from Thailand*. Chicago: Field Museum of Natural

323   History.

324   **Hughes AC, Satasook C, Bates PJ, Bumrungsri S, Jones G. 2011.** Explaining the causes of the

325   zoogeographic transition around the Isthmus of Kra: using bats as a case study. *Journal of*

326   *Biogeography* **38:** 2362-2372.

327   **Kuraishi N, Matsui M, Hamidy A, Belabut DM, Ahmad N, Panha S, Sudin A, Yong HS,**

328   **Jiang JP, Ota H, Thong HT, Nishikawa K. 2013.** Phylogenetic and taxonomic

329   relationships of the *Polypedates leucomystax* complex (Amphibia). *Zoologica Scripta*

330   **42:**54–70.

331   **Kuraishi N, Matsui M, Ota H, Chen S-L. 2011.** Specific separation of *Polypedates braueri*

332   (Vogt, 1911) from *P. megacephalus* (Hallowell, 1861) (Amphibia: Anura:

333   Rhacophoridae). *Zootaxa* **2744:**53–61.

334   **Librado P, Rozas J. 2009.** DnaSP v5: a software for comprehensive analysis of DNA

335   polymorphism data. *Bioinformatics* **25:**1451–1452.

336   **Manni F, Guérard E, Heyer E. 2004.** Geographic patterns of (genetic, morphologic, linguistic)

337   variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology*

338   **76:**173–190.

339   **Matsui M, Ito H, Shimada T, Ota H, Saidapur SK, Khonsue W, Tanaka-Ueno T, Wu GF.**

340   **2005.** Taxonomic relationships within the pan-oriental narrow-mouth toad *Microhyla*

341   *ornata* as revealed by mtDNA analysis (Amphibia, Anura, Microhylidae). *Zoological*

342   *Science* **22:**489–495.

343   **Matsui M, Mohamed M, Shimada T, Sudin A. 2007.** Resurrection of *Staurois parvus* from *S.*

344   *tuberilinguis* from Borneo (Amphibia, Ranidae). *Zoological Science* **24:**101–106.

345   **Matsui M, Seto T, Utsunomiya T. 1986.** Acoustic and karyotypic evidence for specific

346   separation of *Polypedates megacephalus* from *P. leucomystax*. *Journal of Herpetology*

347   **20:**483–489.

348   **Nylander JA. 2004.** MrModeltest v2.3. [program distributed by the author] Uppsala, Sweden:

349   Evolutionary Biology Centre, Uppsala University.

350   **Pan S, Dang N, Wang J, Zheng Y, Rao D, Li J. 2013.** Molecular phylogeny supports the

351   validity of *Polypedates impresus* Yang 2008. *Asian Herpetological Research* **4:**124–133.

352   **Rambaut A, Suchard MA, Xie D, Drummond AJ. 2013.** Tracer v1.5. *Available at*

353   *http://beast.bio.ed.ac.uk/Tracer* (accessed 22 June 2017).

354   **Rogers AR, Harpending H. 1992.** Population growth makes waves in the distribution of

355   pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552-569.

356   **Ronquist F, Huelsenbeck JP. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed

357   models. *Bioinformatics* **19:**1572–1574.

358   **Rujirawan A, Stuart BL, Aowphol A. 2013.** A new tree frog in the genus *Polypedates* (Anura:

359   Rhacophoridae) from southern Thailand. *Zootaxa* **3702:**545–565.

360   **Sheridan JA. 2009.** Reproductive variation corresponding to breeding season length in three

361   tropical frog species. *Journal of Tropical Ecology* **25:**583–592.

362   **Stamatakis A, Hoover P, Rougemont J. 2008.** A rapid bootstrap algorithm for the RAxML web

363   servers. *Systematic Biology* **57:**758–771.

364   **Tajima F. 1989.** Statistical method for testing the neutral mutation hypothesis by DNA

365   polymorphism. *Genetics* **123**: 585-595.

366   **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: Molecular

367   Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30:**2725–

368   2729.

369   **Taylor EH. 1962.** The amphibian fauna of Thailand. *University of Kansas Science Bulletin*

370   **43:**265–599.

371   **Todd BD, Scott DE, Pechmann JH, Gibbons JW. 2011.** Climate change correlates with rapid

372   delays and advancements in reproductive timing in an amphibian community.

373   *Proceedings of the Royal Society B: Biological Sciences* **278:**2191–2197.

374   **Woodruff DS. 2010.** Biogeography and conservation in Southeast Asia: how 2.7 million years of

375   repeated environmental fluctuations affect today's patterns and the future of the remaining

376   refugial-phase biodiversity. *Biodiversity and Conservation* **19:**919–941.

377   **Yuan Z-Y, Suwannapoom C, Yan F, Poyarkov NA Jr, Nguyen SN, Chen H-m, Chomdej S,**

378   **Murphy RW, Che J. 2016.** Red River barrier and Pleistocene climatic fluctuations

379   shaped the genetic structure of *Microhyla fissipes* complex (Anura: Microhylidae) in

380   southern China and Indochina. *Current Zoology* **62:**531–543.

Figure 1

Phylogeographic relationships and a minimum spinning network of all haplotype of the *Polypedates leucomystax* complex among populations in Thailand.

It consists of clade A as *Polypedates* sp., clade B as *P.* cf. *impresus*, clade C as *P.* cf. *leucomystax*, and clade D as *P.* cf. *megacephalus*, as well as outgroups (KR087858, KP996762 = *P. braueri*) inferred from Bayesian analysis of mitochondrial COI gene sequences. Bayesian posterior probability values are expressed above internodes. The asterisks above branches represent bootstrap support for Bayesian posterior probabilities and maximum likelihood (>95%). Scale bar represents 0.5 nucleotide substitutions per site. IOK represents the Isthmus of Kra.

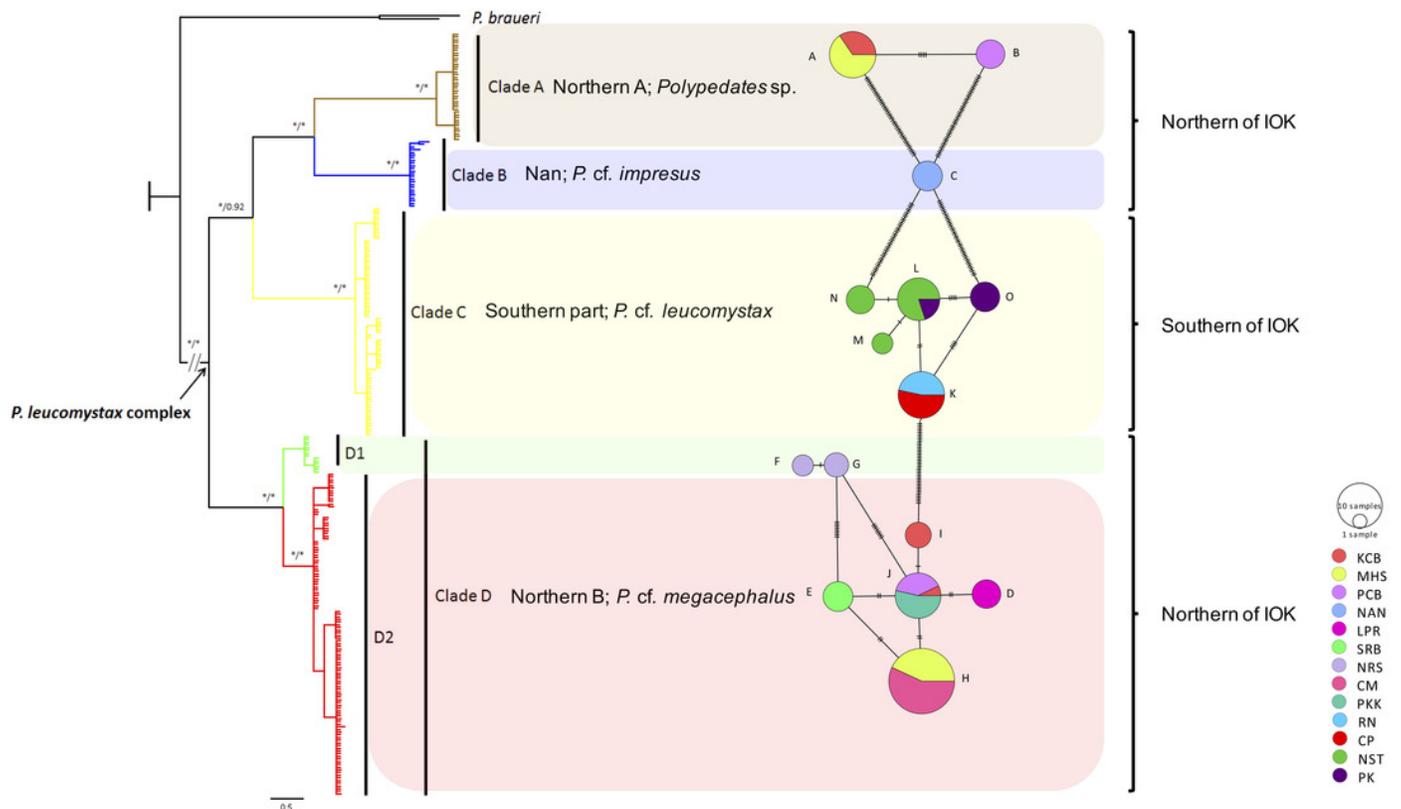



Figure 2

Clade distribution of the *Polypedates leucomystax* complex throughout Thailand.

The abbreviations for each locality are given in Table 1. Different colors represent the different clades.

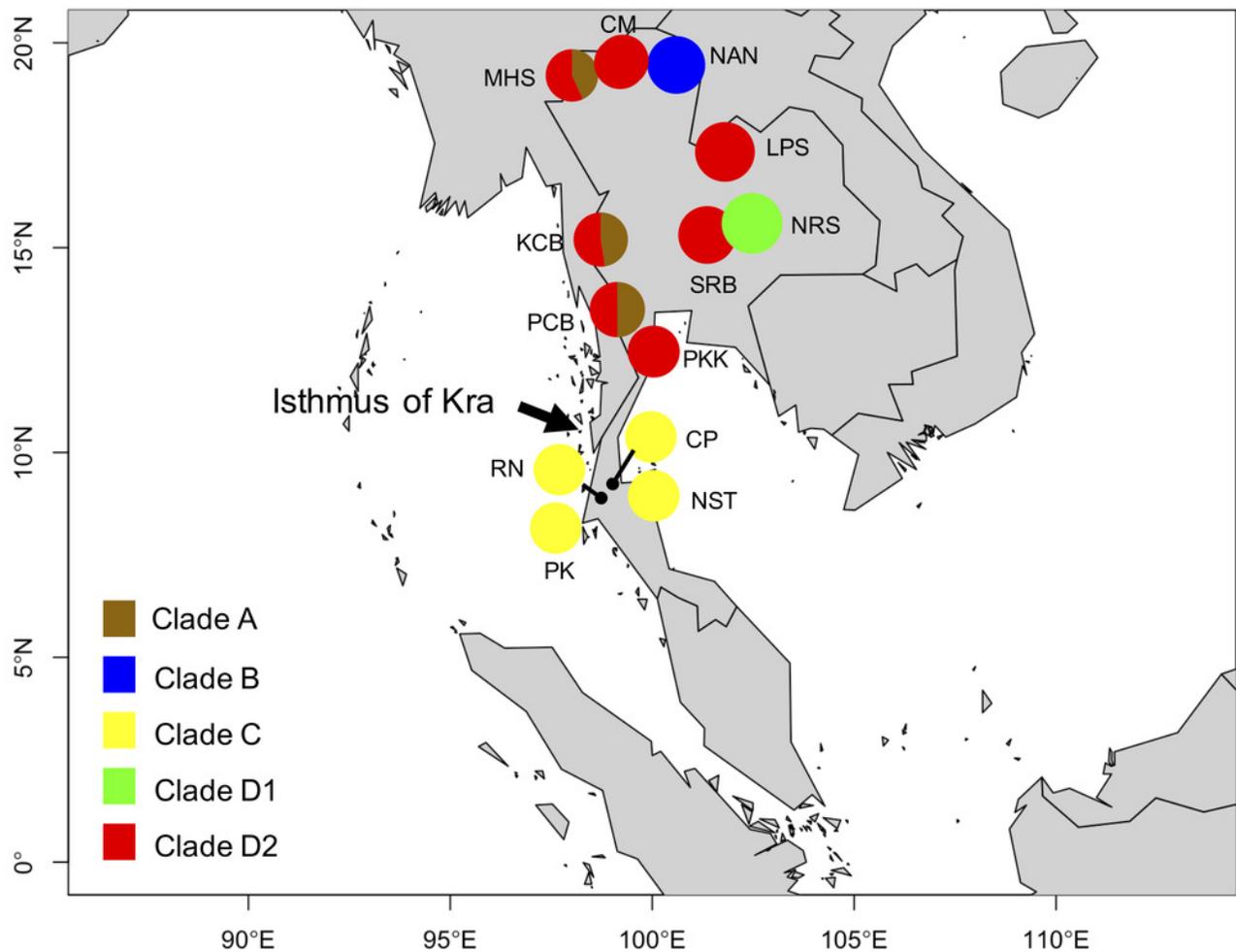



Figure 3

The relationship between climatic condition and genetic differentiation across the Northern B clade (putative *Polypedates megacephalus*) and the Southern clade (putative *Polypedates leucomystax*).

(A) annual rainfall (in mm) in Thailand for 2015 (Image credit: Thai Meteorological Department); and (B) a significant barrier to partition the distribution of the Northern B (putative *P. megacephalus*) and Southern (putative *P. leucomystax*) clades, by Barrier version 2.2.

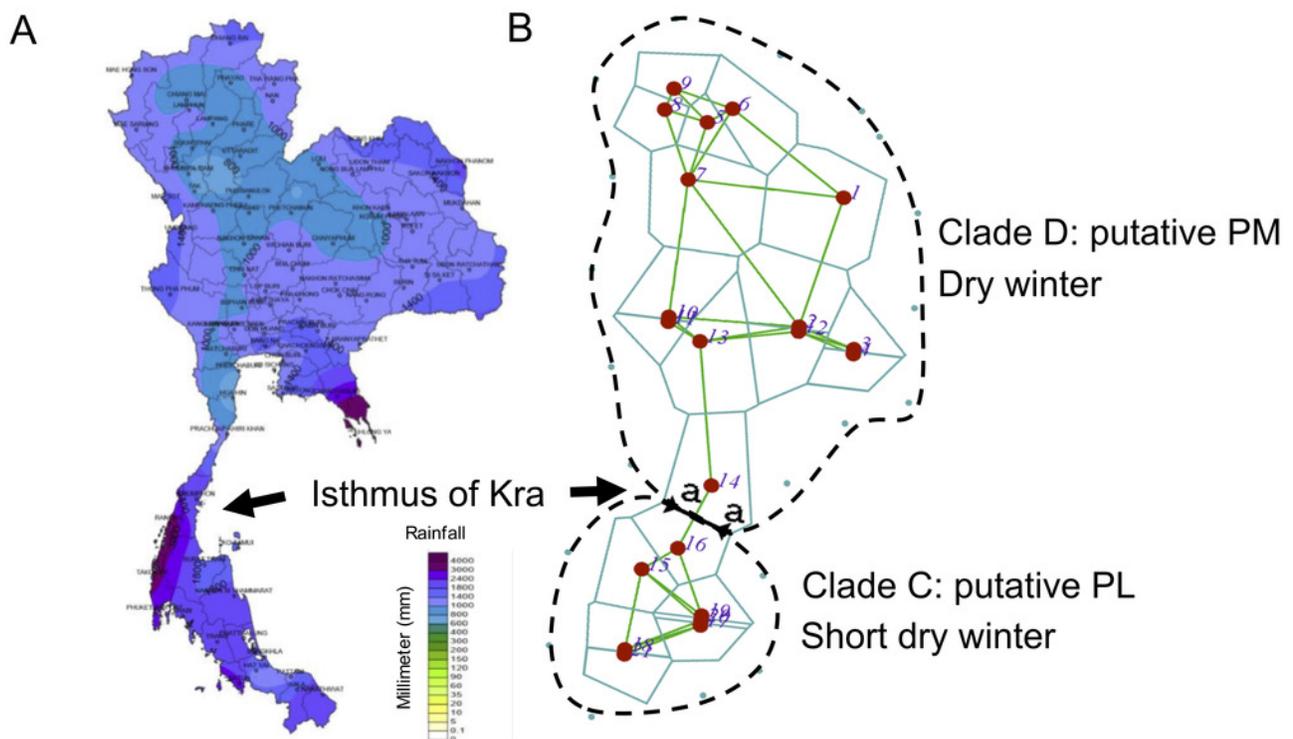



Figure 4

The correlation of genetic distance and linear geographic distance (km) for (A) the Northern B clade (putative *Polypedates megacephalus*) and (B) the Southern clade (putative *Polypedates leucomystax*).

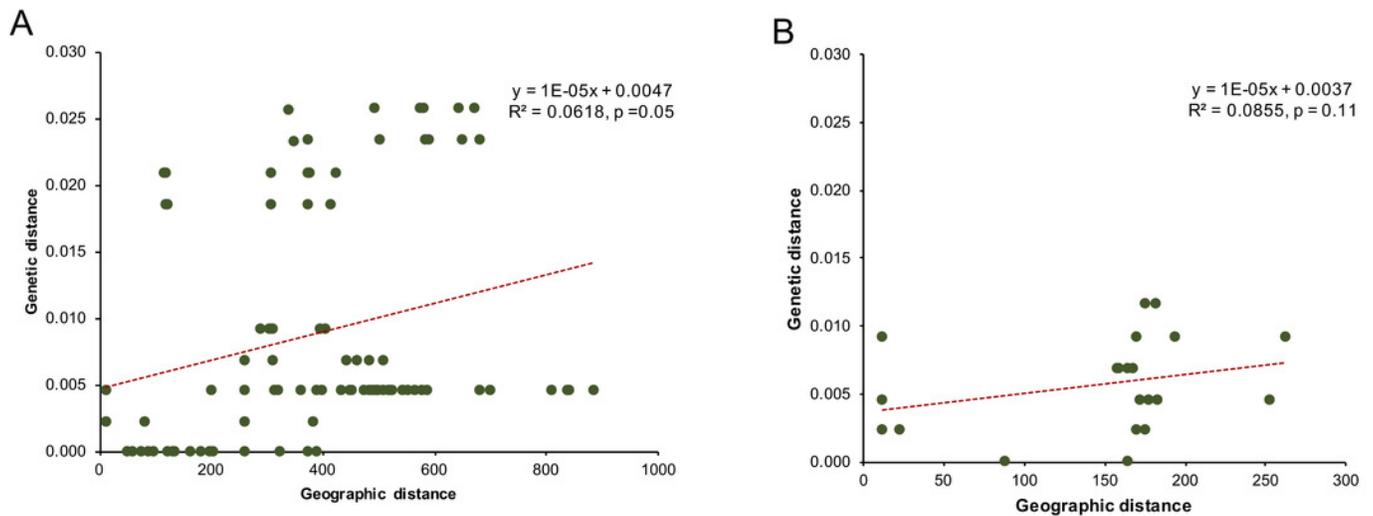
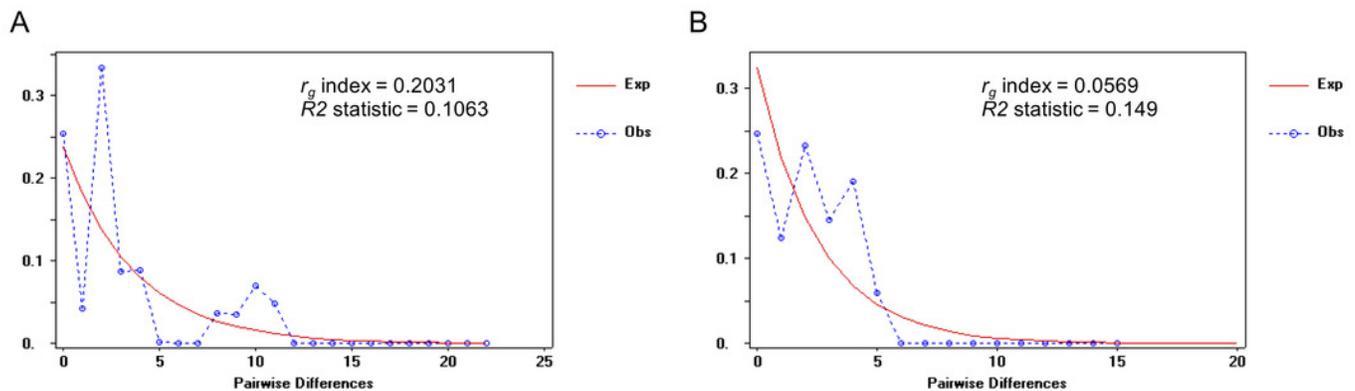



Figure 5

Mismatch distribution of the mitochondrial COI gene in (A) the Northern B clade (putative *Polypedates megacephalus*) and (B) the Southern clade (putative *Polypedates leucomystax*).

The raggedness (rg) index is calculated to evaluate the population expansion of each species. Ramos-Onsins and Rozas's R2 statistic represents the population growth.

Localities of sample collection for the *Polypedates leucomystax* complex in Thailand.

1   **Table 1** Localities of sample collection for *Polypedates leucomystax* complex in Thailand.

Locality	Abbreviation	Number	Altitude (m above sea level)	Longitude	Latitude
Nan province	NAN	12	665	18.980974	101.182594
Kanchanaburi province	KCB	20	917	14.69329	98.40535
Loei province: Phu Ruea	LPR	11	939	17.48193	101.34982
Nakhon Ratchasima province	NRS	14	865	14.49336	101.87364
Chiang Mai province: Mae Wang	CM	7	678	18.657305	98.681831
Chiang Mai province: Doi Saket	CM	13	402	18.98777	99.11455
Chiang Mai province: Omkoi	CM	13	460	17.47137	98.45785
Mae Hong Son province	MHS	44	396	19.24797	97.99542
Saraburi province	SRB	12	105	14.70993	100.81819
Phetchaburi province	PCB	22	329	14.70993	100.81819
Prachuap Khiri Khan province	PKK	15	23	11.43678	99.56011
Ranong province	RN	14	18	9.6052	98.4669
Nakhon Si Thammarat province	NST	37	97	8.76902	99.80349
Phuket province: Thalang	PK	17	31	7.96804	98.33589
Chumphon province	CP	15	103	10.110278	99.082778

Summary of the *P. leucomystax* complex in Thailand

major lineages clades, putative scientific name, number of individuals (N), number of mtDNA haplotypes (n), number of polymorphic sites (P), parsimony-informative sites (PI) and singleton sites (S), haplotype diversity (H_d) and nucleotide diversity (π).

Table 2 Summary of the *P. leucomystax* complex in Thailand: major lineages clades, putative scientific name, number of individuals (*N*), number of mtDNA haplotypes (*n*), number of polymorphic sites (P), parsimony-informative sites (PI) and singleton sites (S), haplotype diversity (H_d) and nucleotide diversity (π).

Clade	Putative species	*N*	*n*	π	H_d	P	S	PI
A (the Northern A)	*Polypedates* sp.	40	2	0.0037	0.4089	4	0	4
B (Nan)	*P. impresus*	12	1	0	0	0	0	0
C (the Southern)	*P. megacephalus*	131	7	0.0048	0.746	15	1	14
D (the Northern B)	*P. leucomystax*	83	5	0.0073	0.7526	7	0	7
Total		266	15	0.0664	0.9	82	0	82

Summary of statistics used to compute the demographic history of populations of the Northern B clade (putative *P. megacephalus*) and the Southern clade (putative *P. leucomystax*).

Table 3 Summary of statistics used to compute the demographic history of populations of the Northern B clade (putative *P. megacephalus*) and the Southern clade (putative *P. leucomystax*).

Clade	Tajima's *D*		Fu's *Fs*	
	D	*P* value	*Fs*	*P* value
The Northern B	0.439	>0.1	3.213	0.045
The Southern	1.176	>0.1	3.031	0.071