



Phylogeographic genetic variation of *Indoplanorbis exustus* (Deshayes, 1834) (Gastropoda: Planorbidae) in South and Southeast Asia

Weerachai Saijuntha^a, Chairat Tantrawatpan^{b,*}, Takeshi Agatsuma^c, R.P. V. Jayanthe Rajapakse^d, K.J.K. Karunathilake^d, Warayutt Pilap^a, Wittaya Tawong^e, Trevor N. Petney^f, Ross H. Andrews^{g,h}

^a Walai Rukhvej Botanical Research Institute, Mahasarakham University, Maha Sarakham 44150, Thailand

^b Division of Cell Biology, Department of Preclinical Sciences, Faculty of Medicine, and Center of Excellence in Stem Cell Research, Thammasat University, Rangsit Campus, Pathum Thani 12120, Thailand

^c Department of Environmental Medicine, Kochi Medical School, Kochi University, Oko, Nankoku, Kochi 783-8505, Japan

^d Department of Veterinary Pathobiology, Faculty of Veterinary Medicine and Animal Science, University of Peradeniya, Peradeniya 20400, Sri Lanka

^e Department of Agricultural Science, Faculty of Agriculture, Natural Resources and Environment, Naresuan University, Phitsanulok 65000, Thailand

^f Department of Zoology and Paleontology and Evolution, State Museum of Natural History Karlsruhe, Erbprinzenstrasse 13, 76133 Karlsruhe, Germany

^g Cholangiocarcinoma Research Institute, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand

^h Faculty of Medicine, Imperial College London, St Mary's Campus, South Wharf Street, London W2 1NY, United Kingdom

ARTICLE INFO

Keywords:

Planorbids
Intermediate host
Genetic structure
Species complex
Mitochondrial DNA

ABSTRACT

The freshwater snail *Indoplanorbis exustus* play an important role as the sole intermediate host of several medically- and economically-important trematodes, especially zoonotic schistosomes and echinostomes, which can infect and cause diseases in livestock and people. This study aims to explore the mitochondrial cytochrome *c* oxidase subunit 1 sequence variation of *I. exustus* collected from new geographical areas; 459 specimens of *I. exustus* were collected from 43 localities in South and Southeast Asia. The 42 haplotypes (Ie1 – Ie42) we detected were classified into haplogroups I – V. Phylogenetic analyses revealed five major clades, A – E, in concordance with all previous studies. Clade E contained two subclades, E1 (haplogroup I) and E2 (haplogroup II). The most widespread genetic group was subclade E1. Clade A, clade B (haplogroup V), and clade C (haplogroup IV) were found only in South Asia, whereas clade D (haplogroup III) was specifically found in Southeast Asia. In Thailand, *I. exustus* showed high genetic divergence with 21 haplotypes. Several isolates showed significant genetic differences from others with unique haplotype(s). Hence, we confidently conclude our findings support all previous studies that *I. exustus* is a species complex with at least four major lineages and five haplogroups. Our additional analyses of 35 samples from Sri Lanka showed these were indeed an independent genetic group as previously found, but they can now be classified as a unique group forming subclade E2 (haplogroup II) of *I. exustus* sensu lato.

1. Introduction

The freshwater snails, *Indoplanorbis exustus* (Deshayes, 1834) are intermediate hosts for several medically-important parasites, such as those responsible for cattle schistosomiasis, especially the *Schistosoma indicum* species group. They are also intermediate hosts for the zoonotic intestinal trematodes *Artyfechinostomum sufrartyfex* and *Echinostoma* sp. [26,29]. In addition, *I. exustus* has been implicated in human cercarial dermatitis outbreaks in India, Lao PDR, Malaysia and Thailand [6,13,21,24]. Cercarial dermatitis is caused by a cutaneous allergic

reaction in people who are exposed to larval schistosomes (cercariae) which are shed by infected snails into freshwater bodies (e.g. ponds, lakes and rice paddy fields). Cercarial infections cause papular pruritic eruptions: severe secondary infections may result from cercarial attempts at infecting non-permissive definitive hosts and subsequently dying in the skin [20].

In addition, *I. exustus* is a hermaphroditic, invasive snail species with high fecundity and self-fertilization ability. It has the ability to colonize habitats containing other well-established pulmonates and prosobranch snails within a year [25]. In optimal conditions, e.g. 30 °C, each snail can

* Corresponding author.

E-mail address: tchairat@tu.ac.th (C. Tantrawatpan).

<https://doi.org/10.1016/j.onehlt.2021.100211>

Received 16 July 2020; Received in revised form 31 December 2020; Accepted 2 January 2021

Available online 6 January 2021

2352-7714/© 2021 The Author(s).

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

lay up to 968 eggs every 23 days [27]. The snails commonly attach to aquatic plants in small ponds, pools, tanks, lakes including stagnant pools of water in rivers and reservoirs, and semi-permanent pools flooding fields, where they can survive buried in mud during the dry season. In the latter case, dispersal may occur in clumps of mud stuck to cattle and water buffalo or disperse via water in debris [17]. Currently, *I. exustus* is known to be a widely distributed species in tropical and subtropical zones throughout South and Southeast Asia, southern China and across Africa to the French West Indies [18,19]. In addition, seasonal rainfall variations have been shown to influence the distribution of *I. exustus* in Assam/Bangladesh and both the Southeast Asian mainland and peninsula. Hence, the habitats of *I. exustus* in these areas are subject to very different ecological conditions [17].

A series of investigation into the genetic and phylogenetic variations in *I. exustus* using mitochondrial DNA sequences has revealed that four to five genetically distinct lineages, clades, were always congruent, suggesting *I. exustus* may comprise more than one species or, in fact, be a species complex [5,10,17,19]. It was initially classified as clades existing in north India (Assam and Bangladesh) and Cratonic India (Sri Lanka), ubiquitous in Thailand, the Philippines, and Nepal, Sundaic in

Indonesia, as well as Arabian and Malaysian clades [17]. Subsequently, these clades were classified and named as clade A to clade E corresponding to their geographical origins. For instance, clade A contained samples from Nepal, clade B from Myanmar and Nepal, clade C from Lao PDR, clade D from India and Nepal, and clade E from a wide distribution covering Sri Lanka, Oman, Nepal, Myanmar, Thailand, Indonesia, Malaysia, and the Philippines [5,10]. More recently, the five major clades were categorized into only four clades. The previously classified clade C and D by Gauffre-Autelin et al. [10] were merged as one clade, following examination of new localities from Africa and the French West Indies [19].

The mitochondrial cytochrome c oxidase subunit 1 (CO1) has been shown to be the best potential marker to elucidate phylogeographic genetic variations in this snail [10,19]. The majority of previous studies have used the CO1 sequence to examine the molecular systematics of *I. exustus*. Unfortunately, previous studies analyzed few geographical localities within the broad-scale distribution clades of *I. exustus* [5,10,17,19]. Thus, our study aims to use the CO1 sequence to examine phylogeographic genetic variations of *I. exustus* collected from novel localities in South and Southeast Asia and in particular, a detailed

Table 1
Sample collection localities and haplotype frequencies examined in this study.

Code	Province	District	Region ^a	Country	Habitat	No.	Haplotype frequency ^b
CMI	Chaing Mai	San Sai	North	Thailand	Canal	10	Ie1(8), Ie2(2)
NAN	Nan	Na Noi	North	Thailand	Canal	10	Ie1(10)
CRI	Chaing Rai	Wiang Pa Pao	North	Thailand	Paddy field	10	Ie1(9), Ie7(1)
LPG	Lampang	Mueang	North	Thailand	Canal	10	Ie1(5), Ie4(2), Ie9(3)
PRE	Prae	Rong Kwang	North	Thailand	Swamp	10	Ie1(6), Ie14(3), Ie15(1)
UTT	Utaradith	Mueang	North	Thailand	Canal	10	Ie1(7), Ie11(3)
PNB	Petchaboon	Lomsak	Central	Thailand	Swamp	10	Ie1(10)
PLK	Pitsanulok	Wang Thong	Central	Thailand	Paddy field	10	Ie1(10)
PCT	Pichit	Sak Lek	Central	Thailand	Canal	10	Ie3(10)
STI	Sukhothai	Mueang	Central	Thailand	Canal	10	Ie1(10)
PTE	Pathumthani	Klong Luang	Central	Thailand	Canal	10	Ie1(8), Ie4(1), Ie5(1)
LRI	Lopburi	Mueang	Central	Thailand	Canal	10	Ie1(10)
AYA	Ayuthaya	Wang Noi	Central	Thailand	Swamp	10	Ie1(3), Ie7(5), Ie19(2)
ATG	Arngthong	Wiset Chai Chan	Central	Thailand	Paddy field	10	Ie1(10)
SPB	Supan Buri	Mueang	Central	Thailand	Swamp	10	Ie16(10)
NYK	Nakhon Nayok	Ongkharak	Central	Thailand	Canal	10	Ie1(8), Ie4(2)
KRI	Kanchanaburi	Mueang	West	Thailand	Canal	10	Ie1(5), Ie20(5)
TAK	Tak	Mae Sot	West	Thailand	Canal	10	Ie1(10)
RBR	Ratchaburi	Suan Phueng	West	Thailand	Swamp	10	Ie1(9), Ie17(1)
CBI	Chonburi	Ban Bueng	East	Thailand	Swamp	10	Ie4(10)
UBN	Ubon Ratchathani	Khong Chiam	Northeast	Thailand	Paddy field	10	Ie1(10)
YST	Yasothon	Mueang	Northeast	Thailand	Canal	10	Ie1(10)
KKN	Khon Kaen	Mueang	Northeast	Thailand	Canal	10	Ie1(7), Ie8(3)
NKI	Nong Khai	Mueang	Northeast	Thailand	Swamp	10	Ie1(5), Ie4(5)
UDN	Udonthani	Mueang	Northeast	Thailand	Swamp	10	Ie1(10)
ACR	Amnatchareon	Mueang	Northeast	Thailand	Canal	10	Ie1(6), Ie18(4)
NBP	Nong Boua Lamphu	Non Sang	Northeast	Thailand	Swamp	10	Ie1(10)
MKM	Maha Sarakham	Kantharawichai	Northeast	Thailand	Paddy field	10	Ie1(3), Ie10(7)
RET	Roi Et	Mueang	Northeast	Thailand	Swamp	10	Ie1(10)
BRM	Buri Ram	Krasang	Northeast	Thailand	Paddy field	10	Ie1(5), Ie4(5)
NRT	Nakhon Ratchasima	Choc Chai	Northeast	Thailand	Paddy field	10	Ie1(6), Ie12(3), Ie13(1)
STN	Satooon	Mueang	South	Thailand	Swamp	10	Ie1(10)
SNI	Surathani	Hat Yai	South	Thailand	Canal	10	Ie21(9), Ie22(1)
LVV	Viantaine	Vang Vieng	-	Lao PDR	Song River	10	Ie1(8), Ie6(2)
LVT	Viantaine	Kampang Nakhon	-	Lao PDR	Canal	10	Ie1(10)
LSV	Savannakhet	Kaisone Phomvihhan	-	Lao PDR	Canal	10	Ie1(10)
CSR	Siem Reab	Siem Reab	-	Cambodia	Canal	10	Ie1(10)
CPS	Pursat	Pursat	-	Cambodia	Swamp	10	Ie26(6), Ie27(4)
CPP	Phnom Penh	Phnom Penh	-	Cambodia	Swamp	10	Ie23(10)
CKC	Kampong Cham	Kampong Cham	-	Cambodia	Swamp	10	Ie24(4), Ie25(6)
PHP	Cagayan	Magrabil	-	Philippines	Canal	13	Ie1(8), Ie42(5)
BAN	Mymensingh	Mymensingh	-	Bangladesh	Canal	11	Ie28(7), Ie29(1), Ie30(1), Ie31(1), Ie32(1)
SRI	Central	Barigama	-	Sri Lanka	Canal	13	Ie33(2), Ie34(4), Ie35(3), Ie36(2), Ie37(1), Ie38(1)
		Hadeniya	-	Sri Lanka	Canal	5	Ie35(4), Ie39(1)
	North Central	Kekirawa	-	Sri Lanka	Canal	9	Ie35(5), Ie36(2), Ie40(2)
		Polonnaruwa	-	Sri Lanka	Canal	8	Ie35(6), Ie41(2)
Total						459	

^a Regions of Thailand.

^b Number in the bracket after haplotype name indicates the observed number of a particular haplotype in each locality.

analysis of samples collected on a fine geographical scale at localities throughout Thailand, including CO1 sequences generated by all previous studies.

2. Materials and methods

2.1. Sample collection

The 459 specimens of *I. exustus* were collected from various natural habitats, e.g. swamps, rivers, canals, and rice paddy fields. Samples were collected from 33 geographical localities in Thailand, three localities in Lao PDR, four localities in Cambodia, four localities in Sri Lanka, one locality in the Philippines, and one locality in Bangladesh (Table 1). After collection, snails were preserved in 80% alcohol before transport to the laboratory. Each snail species was morphologically identified as described by the Danish Bilharziasis Laboratory [4]. The ethics of animal use and biosafety for this research were approved by The Animal Care and Use Committee and the Institute Biosafety Committee of Thammasat University, respectively.

2.2. Molecular analyses

Snails were individually used for total genomic DNA extraction using the E.Z.N.A.® Mollusc DNA Kit (Omega Bio-Tek, USA), following manufacturer's instructions. The mitochondrial CO1 gene was amplified using primers LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') [9]. PCR consisted of a 25 µl final reaction volume containing 1 × Ex buffer (Takara, Japan) with 0.2 mM each of dNTP, 0.2 µM of each forward primer, 0.625 U of Ex *Taq* DNA Polymerase (Takara, Japan), and 1 µl (~10–50 ng) of the DNA sample. PCR conditions used for both gene amplifications were: 94°C for 4 min followed by 35 cycles of 1 min each of 94°C, 50°C, and 72°C, final extension at 72°C for 8 min. Amplified products were analyzed using 1.0% agarose gel electrophoresis, which was then cut and purified using an E.Z.N.A.® Gel Purification Kit (Omega Bio-Tek, USA). Purified PCR products were cycle-sequenced at Eurofins Genomics Company, Japan.

2.3. DNA sequence and haplotype analyses

Sequences were manually checked and edited using ABI Sequence Scanner Software v1.0 and BioEdit v7.2.6 [11]. Multiple sequence alignment was performed using ClustalW [15]. Genetic diversity indices, i.e. the segregating sites, haplotype number, haplotype diversity, and nucleotide diversity, were computed and generated by DnaSp v5 [16]. Pairwise values of genetic differentiation (Φ_{ST}) based on the Kimura two-parameter model [12] and genetic structure (Analysis of Molecular Variance: AMOVA) were calculated in Arlequin v.3.5.2.2 [8]. The minimum spanning haplotype network of CO1 and *rrnL* sequences was constructed with Network v.5.0.1.1 (<http://www.fluxus-engineering.com>) based on the median-joining algorithm [1]. Based on haplotype network analysis, the haplogroups of CO1 sequences were defined based on mutational steps (ms) ≥ 10 .

2.4. Phylogenetic analyses

Time-reversible variation data among sites which were generated from a gamma distribution model (GTR + G + I) [22] were selected for analyses by maximum likelihood (ML) and Bayesian inference (BI) methods as determined by the MrModelTest v.2.3 [23]. The ML tree was constructed using MEGA X [14]. The BI was run using MrBayes v.3.2 [28] with 15,000,000 generations using four chains. Trees were sampled every 100th generation, resulting in 150,000 trees. The first 110,000 trees were excluded as burn-in because the average standard deviation of split frequencies (ASDF) was >0.01 . Furthermore, convergence of the run was assessed by checking the potential scale reduction factor (PSRF)

values of each parameter in MrBayes as well as the effective sample size (ESS) values of each parameter in Tracer 1.6 (<http://beast.bio.ed.ac.uk/Tracer>). A value of PSRF ~ 1.00 and of ESS > 200 were considered suggestive of convergence. The remaining 40,000 trees were used to calculate Bayesian posterior probabilities of individual clades. The 98 CO1 sequences of *I. exustus* deposited in GenBank from previous reports, namely HM104223, GU451738 – GU451750 [17], KR811332 – KR811361 [5], KY024420 – KY024472 [10], and MH037074 – MH037075 [19] were included, whereas a CO1 sequence of *Bulinus tropicus* (AY282583) was used as an out-group. The Poisson tree processes model (PTP) for species delimitation and a Bayesian implementation of the PTP (bPTP) were used to infer putative species boundaries on a given phylogenetic input tree [31]. The following parameters were included in PTP and bPTP analyses, namely, MCMC, 500000 generations; thinning, 100; burn-in, 0.1; and seed, 123.

3. Results

3.1. Sequence analyses

The CO1 sequences of 459 individual *I. exustus* from 43 different localities in South and Southeast Asia were determined and deposited in GenBank under accession numbers MT274329 – MT274370. Genetic divergence was detected within all populations, revealed by nucleotide diversity ranging between 0.001 ± 0.001 to 0.043 ± 0.023 , with an average of 0.012 ± 0.001 . Haplotype diversity varied from 0.129 ± 0.079 to 0.831 ± 0.024 , with an average of 0.671 ± 0.087 (Table 2). Population comparisons defined by different countries showed genetic differentiation (Φ_{ST}) was between 0.048 and 0.962, and there were significant differences ($p < 0.05$) between all populations, the exceptions being Thailand vs. Lao PDR, Cambodia vs. Philippines, Lao PDR vs. Philippines (Table 3). Comparisons between 33 different micro-geographical localities (isolates) in Thailand revealed genetic differences (Φ_{ST}) between 0.000 and 1.000. Comparisons between the 33 Thai isolates found that the isolates of Pitchit (PCT), Ayuthaya (AYA), Supan Buri (SPB), Chonburi (CBI), Maha Sarakham (MKM), and Sukhothai (STI) provinces were significantly genetically different (Φ_{ST}) when compared to all other isolates (Supplementary Table 1).

3.2. Haplotype analyses

Based on nucleotide variation of the CO1 sequence, the 459 samples were classified into 42 haplotypes (Ie1 – Ie42) in our study. The most common haplotype detected in Southeast Asia was Ie1. Another common haplotype was Ie4, found in several geographical localities in Thailand, and haplotype Ie7 found in isolates from Chiang Rai (CRI) and Ayutthaya (AYA) provinces (Table 1, Fig. 1). All remaining haplotypes were found to be unique in a particular isolate, such as haplotype Ie3 being found only in the Pichit (PCT) isolate, Ie21 and Ie22 unique to the Satoon (STN) isolate, Ie24 was detected only in Kampong Cham, Cambodia, and Ie42 was present only in Manila, and the Philippine isolate. The South Asian isolates clearly differed from Southeast Asian ones with haplotypes Ie28 – Ie32 seen only in Bangladesh, whereas Ie33

Table 2
Molecular diversity indices of *I. exustus* populations from different countries.

Population	n	S	H	Uh	$\pi \pm SD$	Hd \pm SD
Thailand	330	28	21	20	0.002 ± 0.001	0.507 ± 0.034
Lao PDR	30	8	2	1	0.001 ± 0.001	0.129 ± 0.079
Cambodia	40	62	6	5	0.021 ± 0.011	0.831 ± 0.024
Philippines	13	1	2	1	0.001 ± 0.001	0.513 ± 0.082
Bangladesh	11	67	5	5	0.043 ± 0.023	0.618 ± 0.164
Sri Lanka	35	15	9	9	0.004 ± 0.002	0.717 ± 0.075
All populations	459	126	42	41	0.012 ± 0.001	0.671 ± 0.087

n = number, s = number of polymorphic sites, H = number of haplotypes, Uh = Unique haplotypes, π = nucleotide diversity, Hd = haplotype diversity.

Table 3

Genetic differentiation (Φ_{ST}) (lower triangle) and p -value (upper triangle) among six populations of *I. exustus* from Thailand, Lao PDR, Cambodia, the Philippines, Bangladesh and Sri Lanka examined by CO1 sequence.

Countries	Thailand	Cambodia	Lao PDR	Philippines	Bangladesh	Sri Lanka
Thailand	–	<0.001	0.694	0.036	<0.001	<0.001
Cambodia	0.294	–	<0.001	0.063	<0.001	<0.001
Lao PDR	0.048	0.113	–	0.063	<0.001	<0.001
Philippines	0.087	0.100	0.133	–	<0.001	<0.001
Bangladesh	0.962	0.746	0.877	0.814	–	<0.001
Sri Lanka	0.910	0.633	0.903	0.898	0.863	–

Bold indicated statistically significant differences p -value.

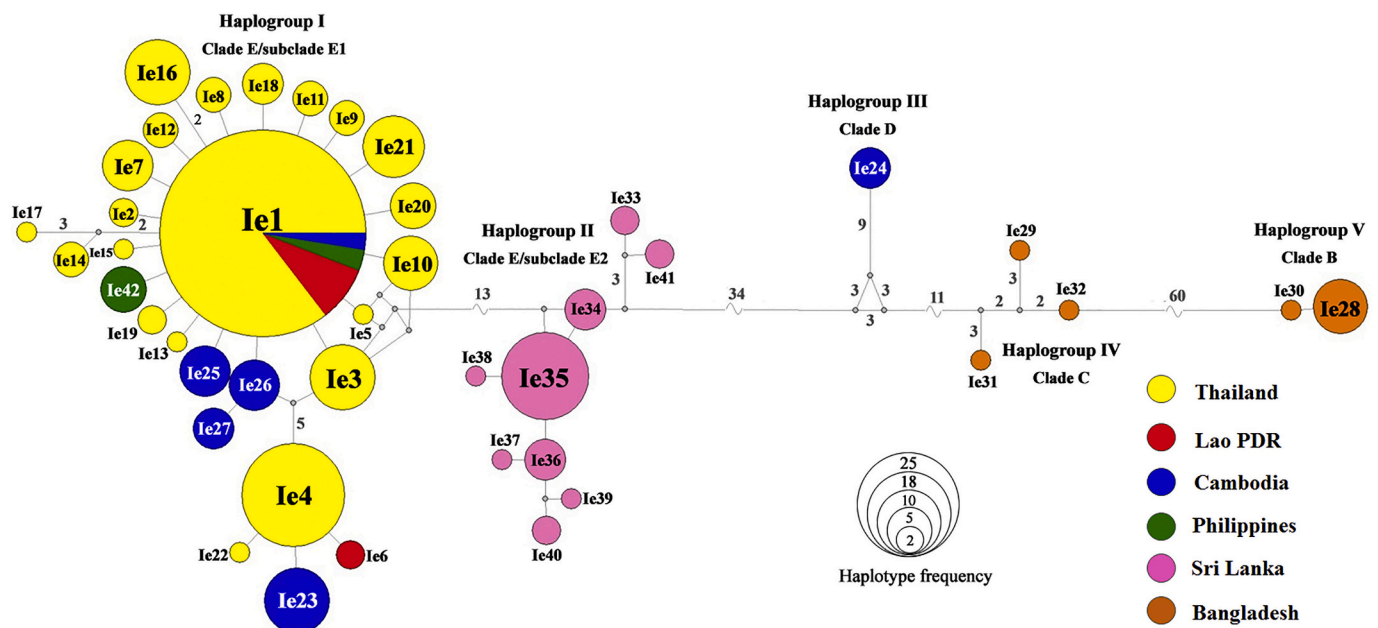


Fig. 1. Minimum spanning haplotype network of *Indoplanorbis exustus* sensu lato generated based on partial CO1 sequence classified into 42 haplotypes (Ie1 – Ie42), which corresponds to their geographical localities, separated into seven countries, represented in different color. The area of the circles represents the proportion of specimens found in each haplotype. Number in each branch represent mutational step number, while the branch with no number represents one mutation step. Small gray dots are median vectors. Haplogroups I – V were classified based on mutational steps >10 .

– Ie41 were detected only in Sri Lanka (Table 1, Fig. 1).

The haplotype network analysis of the 459 CO1 sequences (42 haplotypes) were placed into five haplogroups, I – V based on mutation steps (ms) ≥ 10 . Haplogroup I contained 27 haplotypes (Ie1 – Ie23, Ie25 – Ie27 and Ie42) from all isolates in Southeast Asia, haplogroup II contained nine haplotypes (Ie33 – Ie41) from Sri Lanka, haplogroup III contained haplotype Ie24 from four samples in Kampong Cham, Cambodia, haplogroup IV consisted of three haplotypes (Ie29, Ie31 and Ie32) from Bangladesh, and haplogroup V had haplotypes Ie28 and Ie30 from Bangladesh (Fig. 1). All haplotypes generated in this study combined with all retrieved sequences from GenBank were also used to construct a haplotype network, in which the genetic grouping was similar as indicated in Fig. 1 (Figure Supplementary Fig. 1).

3.3. Phylogenetic analyses

Based on previously studies, five clades (A – E) were classified and sequences were deposited in GenBank, i.e. clade A consisted of sequences from Nepal, clade B from Myanmar and Nepal, clade C from Lao PDR, clade D from India and Nepal, and clade E from wide range covering Sri Lanka, Oman, Nepal, Myanmar, Thailand, Indonesia, Malaysia, and the Philippines. Current phylogenetic tree analyses, combining our 42 haplotypes with 98 sequences from previous studies retrieved from GenBank, demonstrated there were five major lineages, clades A – E, supported by high bootstrap values and BI posterior

probabilities (Fig. 2). Importantly, the PTP/bPTP analyses also supported the presence of five separate cryptic genetic groups corresponding to the five clades (A – E) classification in the phylogenetic tree (Supplementary Fig. 1). This was based on an acceptance rate of 0.61454; merge rate of 250,058; and split rate of 249,942. The estimated number of species was between 4 and 104 with a mean of 53.0 species. We subdivided clade E into two subclades, namely subclades E1 and E2 (Fig. 2). Of these subclades, E1 was the most common and widespread genetic group worldwide, whereas subclade E2 was specifically detected in Sri Lanka and more distantly in the Caribbean (Fig. 3). Our study revealed for the first time that several samples from Bangladesh belonged to clade C together with the sequences from Nepal and India. Almost all of the samples from Bangladesh in our study were found to belong to clade B together with the sequences from Myanmar and Nepal. Clade D was rarely found; only three deposited sequences from Lao PDR and four sequences recently examined from Cambodia clustered in this genetic group. We found that none of the sequences generated in our study belonged to clade A (Fig. 3).

4. Discussion

Our current study was conducted to examine the phylogenetic relationships of *I. exustus* that had been proposed by previous studies [5,10,17,19]. The phylogenetic analyses herein were based on more samples than previous studies, which also covered novel localities with a

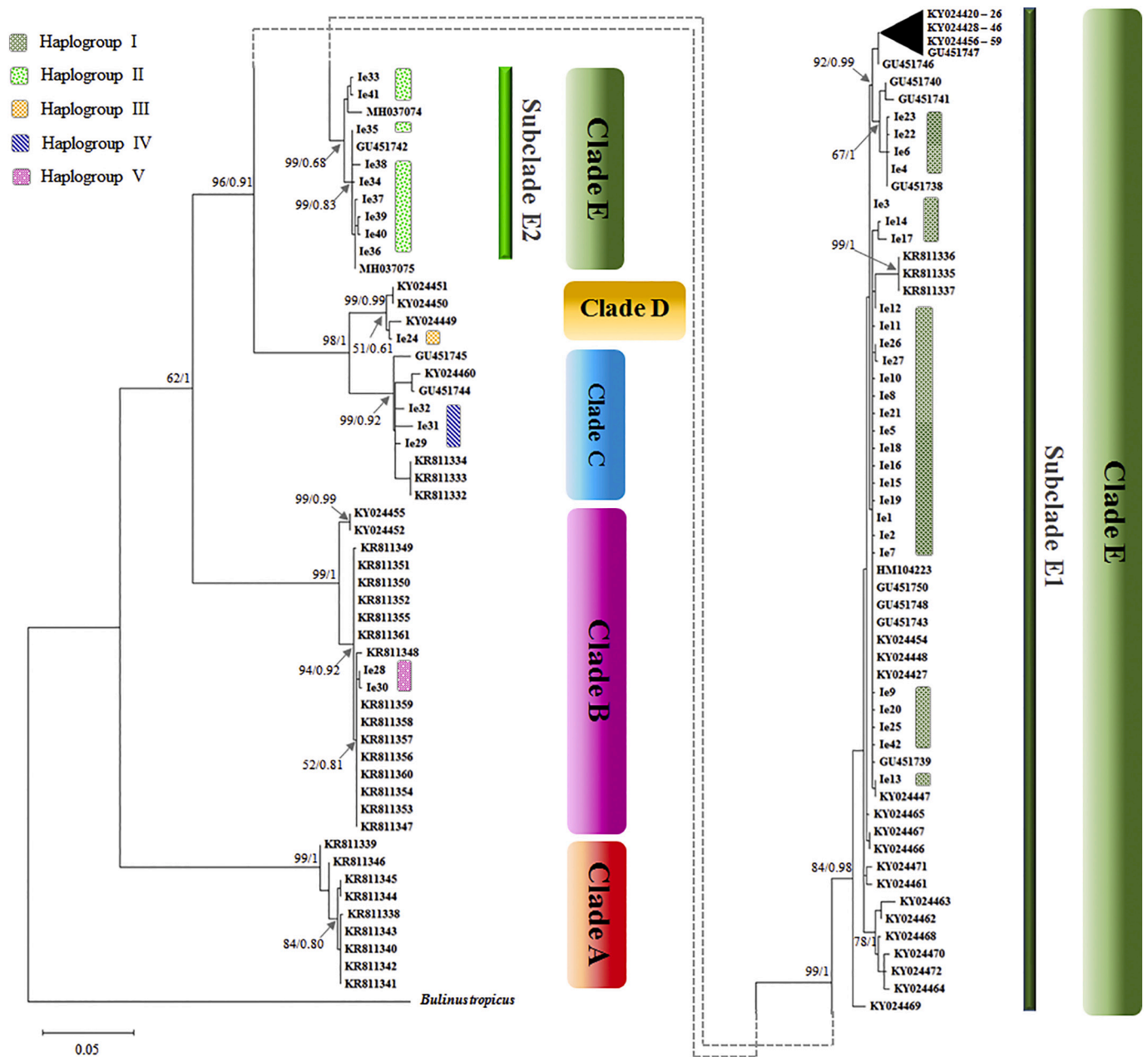


Fig. 2. Phylogenetic tree analyses of CO1 sequences of 42 haplotypes of *Indoplanorbis exustus* sensu lato (sl), which were classified into five haplogroups (I – V) generated in this study, including 98 CO1 sequences of *I. exustus* deposited in GenBank under accession number HM104223, GU451738 – GU451750, KR811332 – KR811361, KY024420 – KY024472, and MH037074 – MH037075 based on maximum likelihood (ML) and Bayesian inference (BI). The tree of *I. exustus* were classified into five major lineages, i.e. clade A – E, whereas clade E were separated into subclades E1 and E2. Nodal supports are sequential values of bootstrap value generated by ML, and Bayesian posterior probability generated by BI. A CO1 sequence of *Bulinus tropicus*, accession number AY282583, was used as an out-group.

specific focus in South and Southeast Asia, especially Thailand. Genetic groups classified in our study were in concordance to all previous studies [10,17,19]. Clades A – E were initially classified corresponding to their geographical origin by Gauffre-Autelin et al. [10], clade A was found specifically in Nepal, clade B occurred in Myanmar and Nepal, clade C was specifically found in Lao PDR, clade D was found in India and Nepal, clade E occurred throughout a wide geographical range covering Sri Lanka, Oman, Nepal, Myanmar, Thailand, Indonesia, Malaysia, and the Philippines. Subsequent analyses proposed that the five major phylogenetic lineages should be classified into only four major lineages, namely clades A – D by the merged of clade C and D as a monophyletic group [19].

However, our study was based on haplotype networks and

phylogenetic analyses, and it showed that *I. exustus* examined in our study can be classified into five haplogroups, belonging to four major evolutionary lineages being clade B – E. We found no samples belonged to clade A, which is congruent with previous report by Gauffre-Autelin et al. [10]. Furthermore, our data showed that clade E were subdivided into two subclades, namely E1 and E2, corresponding to a report by Gauffre-Autelin et al. [10]. All 35 samples from Sri Lanka examined in this study were grouped into subclade E2 (haplogroup II) together with a sequence from Sri Lanka [17] and two sequences from the French West Indies [19]. Our results showed that subclade E2 can confidently be classified as a novel cryptic genetic group, whereas, subclade E1 is the most common genetic group with a worldwide distribution. Furthermore, we discovered two and three novel haplotypes from Bangladesh

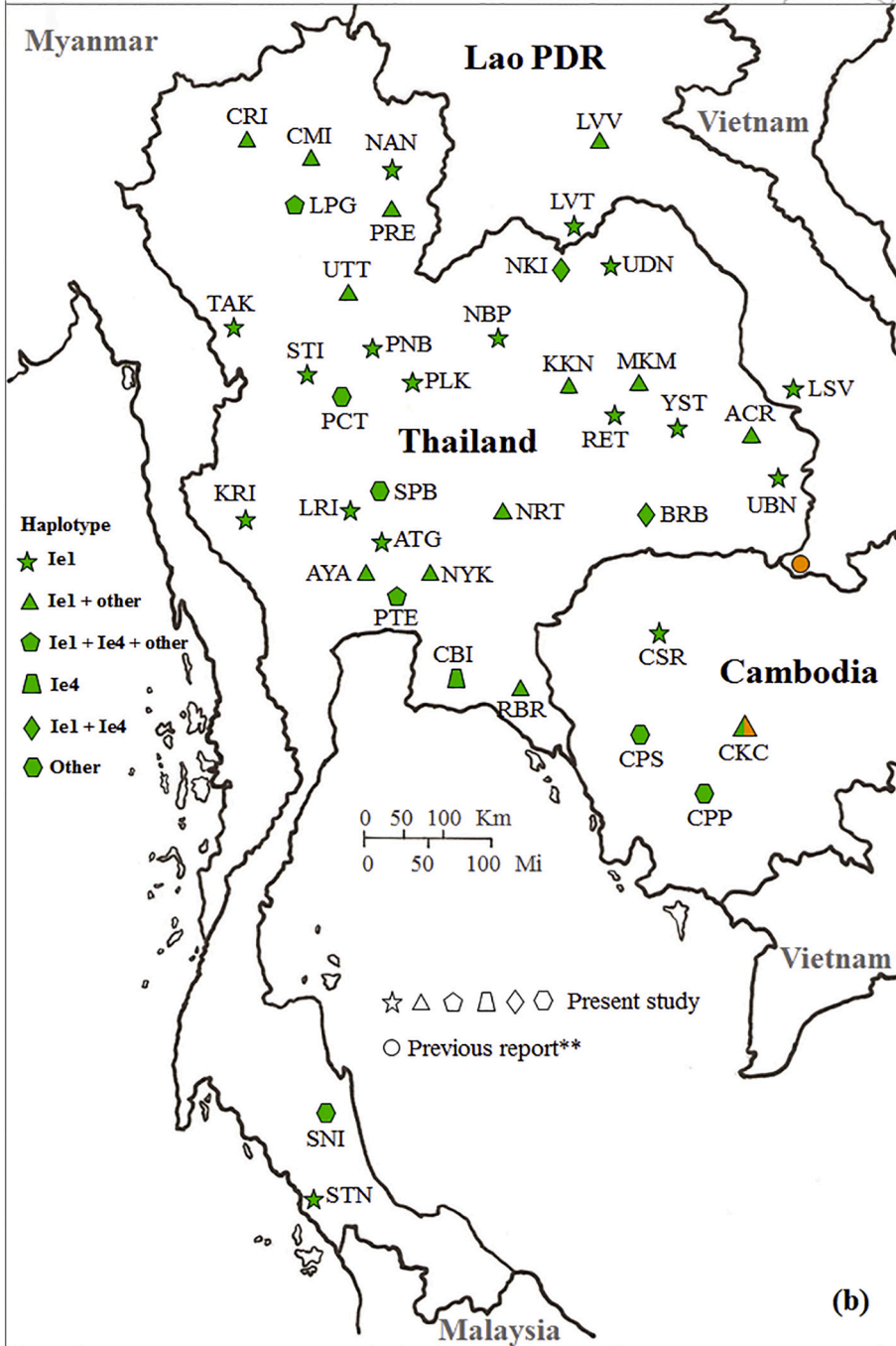
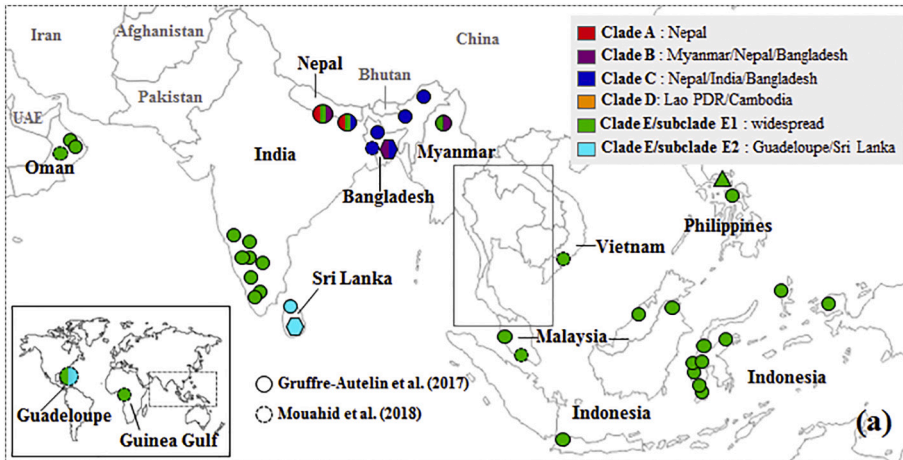


Fig. 3. Distribution of the major clades for *Indoplanorbis exustus* sensu lato. (a) Five main phylogenetic clades (A – E) together with two minor subclades (E1 – E2) were identified. The circle were reported by two previous studies [10,19] represent by different periphery line, whereas the other geometric shapes represent sampling localities in this study. Colors referred to the main clades/subclades based on a CO1 sequence. Mixed colors in each circle indicate co-occurrence of distinct clades/subclades at the same locality. (b) In Southeast Asia, different geometric shape represent the most common haplotypes, Ie1 and Ie4 found in a particular locality, while the “other” represent other haplotypes apart from either Ie1 and Ie4 (see more detail in Table 1).

grouped within clade B and C, respectively.

Previously samples from Lao PDR have been classified as an independent clade D [10]. This was, however, was based on a low sample size of three sequences from Lao PDR (rare genotypes) which may not reflect actual phylogenetic relationships. Subsequently, further analyses by Mouahid et al. [19] of samples from Bangladesh, India and Nepal provided evidence to merge Clade D from Lao PDR into Clade C. Our results now provide the basis to separate clade C [19] into two different clades, i.e. clades C and D corresponding to the previous classification of Gauffre-Autelin et al. [10]. Clades C and D classification was also supported by the PTP species delimitation test. Furthermore, we discovered that a unique haplotype of four snails from a novel place in Cambodia clustered within clade D together with previously generated sequences from Lao PDR. This suggests that this genetic group may be specifically genetically adapted and restricted within these geographical areas. Additionally, *I. exustus* from Philippines were sampled from limited localities, hence, only genotypes that belong to a common genetic group were detected, for instance, subclade E1. Other cryptic genetic groups/genotypes may also exist. Therefore, it is important to examine the extent of genetic variation at micro-geographical scales, covering a far broader range between these areas than has been done to date. Therefore, it is important to examine the extent of genetic variation at micro-geographical scales, covering a far broader range between these areas than has been done previously.

Hence, we also examined the genetic variation of *I. exustus* on a micro-geographical scale covering 33 localities within Thailand, together with three localities in Lao PDR and four in Cambodia. In Thailand, we found that all 21 haplotypes belonged to subclade E1, the most common genetic group distributed worldwide. However, there were several unique haplotypes found in particular localities, leading to significant genetic differences between many localities. The high self-fertilization capacity of *I. exustus* [2,7] may have facilitated clonal production of new genetic types adapted to each environment, leading to the unique haplotypes found in several different areas evidenced in our study. The different catchments in Thailand may also influence genetic differences and genetic structure of *I. exustus*, as water flow is limited or not possible between different catchments. This would lead to limited migration and low gene flow between the *I. exustus* populations across these catchments. For instance, in northeastern Thailand, the sampling locality Nong Khai (NKI) is in the Lower Mekong catchment, Maha Sarakham (MKM) is located in the Chi catchment, and Buri Ram (BRI) is located in Mun catchment. Our results showed that there were significant genetic differences between isolates from these different catchments in northeast Thailand compared to other isolates examined. Similar patterns of genetic differences have previously reported in the other freshwater species of snails, e.g. *Bithynia siamensis* [30] and *Hydrobioides nassa* [3].

The differences and variation in seasonal rainfall patterns and different ecological conditions between different regions of Thailand may also influence the genetic differences we have detected in *I. exustus*.

For instance, significant genetic differences were detected when isolates from the following geographical localities in Thailand were compared to all other isolates examined, namely, isolates from Pichit (PCT), Ayutthaya (AYA), and Supan Buri (SPB) located in the alluvial plain in the central region of Thailand, Lampang (LPG) from the large mountain ranges in northern Thailand, Surathani (SNI) in the peninsular of Thailand, Kanchanaburi (KRI) located in steep river valleys in western, Chonburi (CBI) nearby the Gulf of Thailand in the eastern region, and NKI, MKM and BRI are localities in the Khorat Plateau in northeast Thailand. Interestingly, we found several isolates from geographically distant localities within a defined catchment and isolates from different catchments which showed low or no genetic differences (i.e. a low F_{ST} value). These results may have been due to human activities of trading aquatic plants on which *I. exustus* can be attached. This is one of main sources of migration and gene flow of *I. exustus* in Thailand and worldwide.

Commercial activities are possibly linked to the main route of migration of *I. exustus* from its original geographic region [25]. Genetic examination of *I. exustus* from Africa and French West Indies indicate that these populations originated from the geographical expansion of clade E. Subclade E1, the most common genetic group, has been found in both Africa and French West Indies, whereas subclade E2, is predominantly found in Sri Lanka, as well as being detected in the French West Indies [19]. There are many endemic areas of *I. exustus* worldwide that have not been surveyed to date nor have the snails been genotyped. The natural parasites of *I. exustus* include many medically and veterinary important trematodes, especially schistosomes and echinostomes. Detailed knowledge of co-evolution or co-speciation between *I. exustus* and their parasites is still scarce. Therefore, it is very important to explore their phylogeographic genetic variation in a global context to provide a comprehensive understanding of their distribution, adaptation and host-parasites selection.

In conclusion, our mitochondrial CO1 sequence analyses of the phylogeographic genetic variation of *I. exustus*, which was based on a large sample size, validated the worldwide distribution determined by previous studies [10,17,19]. Currently, at least six genetic groups have been classified based on haplotype networks and phylogenetic analyses. Subclade E1, or haplogroup I, is the most common genetic group distributed globally. Subclade E2 or haplogroup II is a novel cryptic group identified and defined in our study, specifically found in Sri Lanka and recently shown to exist in the French West Indies [19].

Furthermore, we found that clade D occurs in southern Lao PDR and as well as Cambodia. We also found high genetic variation and genetic differences between *I. exustus* collected from different micro-geographical localities in Thailand. The uncovering of unique subclades and haplotypes confirms that *I. exustus*, found worldwide, is a species complex and should hence be referred to as “*I. exustus* sensu lato”.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.onehlt.2021.100211>.

Authorship contributions

Conceptualization; W. Saijuntha, C. Tantrawatpan.
 Data curation; W. Saijuntha, T. Agatsuma, W. Tawong.
 Formal analysis; W. Saijuntha, C. Tantrawatpan.
 Funding acquisition; C. Tantrawatpan, R.H. Andrews.
 Investigation; W. Saijuntha, C. Tantrawatpan.
 Methodology; W. Saijuntha, C. Tantrawatpan, W. Pilap.
 Project administration; W. Saijuntha, C. Tantrawatpan.
 Resources; T. Agatsuma, R.P.V.J. Rajapakse, K.J.K. Karunathilake.
 Writing-original draft; W. Saijuntha.
 Wrintint-review & editing; W. Saijuntha, C. Tantrawatpan, T. Agatsuma, W. Tawong, R.H. Andrews, T.N. Petney.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgements

This research was supported by Thammasat University, Thailand, in a grant (GE.2-10/2557) to C Tantrawatpan. We wish to acknowledge the support of the Japan Society for the Promotion of Science (JSPS) for research in Japan (grant no. S20082) awarded to Prof R H Andrews hosted by Prof. T. Agatsuma at the Department of Environmental Medicine, Kochi Medical School, Japan. We would also like to thank Prof. Mohammed MH Mondal for providing snail samples from Bangladesh.

References

- H.J. Bandelt, P. Forster, A. Rohlf, Median-joining networks for inferring intraspecific phylogenies, *Mol. Biol. Evol.* 16 (1999) 37–48, <https://doi.org/10.1093/oxfordjournals.molbev.a026036>.
- K.Y. Bony, K.F. Konan, O.E. Edia, C. Kouassi N'Gouan, D. Diomande, A. Ouattara, Anatomie et stratégies de reproduction de *Indoplanorbis exustus* (Deshayes, 1834), un mollusque invasif d'eau douce en Côte d'Ivoire (Afrique de l'Ouest), *J. Appl. Biosci.* 71 (2013) 5742–5752, <https://doi.org/10.4314/jab.v71i11.98822>.
- N. Bunchom, W. Saijuntha, K. Vaisusuk, W. Pilap, W. Suksavate, N. Suganuma, T. Agatsuma, T.N. Petney, C. Tantrawatpan, Genetic variation of a freshwater snail, *Hydrobioides nassa* (Gastropoda: Bithyniidae) in Thailand examined by mitochondrial DNA sequences, *Hydrobiologia* (2019), <https://doi.org/10.1007/s10750-019-04013-2> (in press).
- Danish Bilharziasis Laboratory, *A field guide to freshwater snails in countries of the WHO eastern Mediterranean region*, in: Copenhagen, 1983.
- R. Devkota, S.V. Brant, E.S. Loker, The *Schistosoma indicum* species group in Nepal: presence of a new lineage of schistosome and use of the *Indoplanorbis exustus* species complex of snail hosts, *Int. J. Parasitol.* 45 (2015) 857–870, <https://doi.org/10.1016/j.ijpara.2015.07.008>.
- O. Ditrach, V. Nasincova, T. Scholz, M. Giboda, Larval stages of medically important flukes (Trematoda) from Vientiane province, Laos. Part II. Cercariae, *Ann. Parasitol. Hum. Comp.* 67 (1992) 75–81, <https://doi.org/10.1051/parasite/199267375>.
- J.S. Escobar, J.R. Auld, A.C. Correa, J.M. Alonso, Y.K. Bony, M.A. Coutellec, J. M. Koene, J.P. Pointier, P. Jarne, P. David, Patterns of mating-system evolution in hermaphroditic animals: correlations among selfing rate, inbreeding depression, and the timing of reproduction, *Evolution* 65 (2011) 1233–1253, <https://doi.org/10.1111/j.1558-5646.2011.01218.x>.
- L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, *Mol. Ecol. Resour.* 10 (2010) 564–567, <https://doi.org/10.1111/j.1755-0998.2010.02847.x>.
- O. Folmer, M. Black, W. Hoeh, R. Lutz, R. Vrijenhoek, DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates, *Mol. Mar. Biol. Biotechnol.* 3 (1994) 294–299.
- P. Gauffre-Autelin, T. von Rintelen, B. Stelbrink, C. Albrecht, Recent range expansion of an intermediate host for animal schistosome parasites in the Indo-Australian Archipelago: phylogeography of the freshwater gastropod *Indoplanorbis exustus* in south and Southeast Asia, *Parasit. Vectors* 10 (2017) 126, <https://doi.org/10.1186/s13071-017-2043-6>.
- T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucl. Acids Symp.* 41 (1999) 95–98.
- M. Kimura, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, *J. Mol. Evol.* 16 (1980) 111–120, <https://doi.org/10.1007/BF01731581>.
- P. Kullavanijaya, H. Wongwaisayawan, Outbreak of cercarial dermatitis in Thailand, *Int. J. Dermatol.* 32 (1993) 113–115, <https://doi.org/10.1111/j.1365-4362.1993.tb01448.x>.
- S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: molecular evolutionary genetics analysis across computing platforms, *Mol. Biol. Evol.* 35 (2018) 1547–1549, <https://doi.org/10.1093/molbev/msy096>.
- M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T. J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0, *Bioinformatics* 23 (2007) 2947–2948, <https://doi.org/10.1093/bioinformatics/btm404>.
- P. Librado, J. Rozas, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, *Bioinformatics* 25 (2009) 1451–1452, <https://doi.org/10.1093/bioinformatics/btp187>.
- L. Liu, M.M. Mondal, M.A. Idris, H.S. Lokman, P.J. Rajapakse, F. Satrija, J.L. Diaz, E.S. Upatham, S.W. Attwood, The phylogeography of *Indoplanorbis exustus* (Gastropoda: Planorbidae) in Asia, *Parasit. Vectors* 3 (2010) 57, <https://doi.org/10.1186/1756-3305-3-57>.
- J.A. Morgan, R.J. DeJong, Y. Jung, K. Khallaayoune, S. Kock, G.M. Mkoji, E. S. Loker, A phylogeny of planorbid snails, with implications for the evolution of *Schistosoma* parasites, *Mol. Phylogenet. Evol.* 25 (2002) 477–488, [https://doi.org/10.1016/s1055-7903\(02\)00280-4](https://doi.org/10.1016/s1055-7903(02)00280-4).
- G. Mouahid, C. Clerissi, J.F. Allienne, C. Chaparro, S.A. Yafae, R.M. Nguema, M. Ibikounlé, H. Moné, The phylogeny of the genus *Indoplanorbis* (Gastropoda, Planorbidae) from Africa and the French West Indies, *Zool. Scr.* 47 (2018) 558–564, <https://doi.org/10.1111/zsc.12297>.
- C.A. Mulvihill, J.W. Burnett, Swimmer's itch: a cercarial dermatitis, *Cutis* 46 (1990) 211–213, <https://doi.org/10.1186/1471-2458-3-35>.
- K. Narain, J. Mahanta, Dermatitis associated with paddy field environment in Assam, India: a review, *Man-Environ. Relationship* 9 (2000) 213–220.
- M. Nei, S. Kumar, *Molecular Evolution and Phylogenetics*, Oxford University Press, New York, 2000.
- J.A.A. Nylander, MrModeltest v2.3 Program Distributed by the Author, Evolutionary Biology Centre. Uppsala University, 2008.
- J.R. Palmieri, J.T. Sullivan, C.K. Ow-Yang, A survey of snail hosts and larval trematodes collected by peninsular Malaysia and Singapore from 1972 to 1977, *Southeast Asian J. Trop. Med. Public Health* 8 (1977) 275–277.
- J.P. Pointier, P. David, P. Jarne, Biological invasions: the case of planorbid snails, *J. Helminthol.* 79 (2005) 249–256, <https://doi.org/10.1079/joh.2005292>.
- P. Radomyos, A. Tangtrongchitr, S. Krudsood, P. Wilairatana, S. Looareesuwan, *Atlas of Medical Parasitology*, Medical Media Co.Ltd, Bangkok, 2004.
- S.K. Raut, M.S. Rahman, S.K. Samanta, Influence of temperature on survival, growth and fecundity of the freshwater snail *Indoplanorbis exustus* (Deshayes), *Mem. Inst. Oswaldo Cruz* 87 (1992) 15–19, <https://doi.org/10.1590/s0074-02761992000100003>.
- F. Ronquist, J.P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models, *Bioinformatics* 19 (2003) 1572–1574, <https://doi.org/10.1093/bioinformatics/btg180>.
- C. Tantrawatpan, W. Saijuntha, P. Sithithaworn, R.H. Andrews, T.N. Petney, Genetic differentiation of *Artyfechinostomum malayanum* and *A. sufarartefex* (Trematoda: Echinostomatidae) based on internal transcribed spacer sequences, *Parasitol. Res.* 112 (2013) 437–441, <https://doi.org/10.1007/s00436-012-3065-9>.
- C. Tantrawatpan, W. Saijuntha, N. Bunchom, W. Suksavate, W. Pilap, T. Walalite, T. Agatsuma, W. Tawong, P. Sithithaworn, R.H. Andrews, T.N. Petney, Genetic structure and geographical variation of *Bithynia siamensis goniomphalos* sensu lato (Gastropoda: Bithyniidae), the snail intermediate host of *Opisthorchis viverrini* sensu lato (Digenea: Opisthorchiidae) in the lower Mekong Basin revealed by mitochondrial DNA sequences, *Int. J. Parasitol.* 50 (2020) 55–62, <https://doi.org/10.1016/j.ijpara.2019.10.007>.
- J. Zhang, P. Kapli, P. Pavlids, A. Stamatakis, A general species delimitation method with applications to phylogenetic placements, *Bioinformatics* 29 (2013) 2869–2876, <https://doi.org/10.1093/bioinformatics/btt499>.