Molecular Phylogenetics and Evolution 107 (2017) 404–414

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

# Cryptic lineage diversity in the zoonotic pathogen *Angiostrongylus cantonensis*

Sirilak Dusitsittipon [a,1], Charles D. Criscione [b,1,*], Serge Morand [c], Chalit Komalamisra [d], Urusa Thaenkham [a,1,*]

[a] Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand
[b] Department of Biology, Texas A&M University, TX, USA
[c] CNRS-CIRAD, Centre d'Infectiologie Christophe Mérieux du Laos, Sisathanak, Vientiane, Laos
[d] Mahidol-Bangkok School of Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

ARTICLE INFO

*Article history:*
Received 26 August 2016
Revised 2 December 2016
Accepted 6 December 2016
Available online 8 December 2016

*Keywords:*
*Angiostrongylus cantonensis*
Cryptic species
Microsatellite
*Cytochrome b* gene sequences

ABSTRACT

Delimitation of species is still a necessity among parasitic pathogens especially where morphological characters provide limited discernibility. Identification of cryptic lineages (independently evolving lineages that are morphologically similar) is critical as there could be lineage-specific traits that are of epidemiological importance. *Angiostrongylus cantonensis* is a zoonotic pathogen that can cause eosinophilic meningitis in humans. Recent reports of single marker sequence divergence hint at the potential for cryptic diversity in this lungworm. However, to definitively address if single marker divergence corresponds to independent evolving lineages, a multilocus approach is necessary. Using multilocus data, our goal was to determine if there were cryptic lineages within Thailand, a country plagued by several outbreaks and isolated cases of *A. cantonensis* infection. We analyzed the genetic structure of *A. cantonensis* samples collected from snails, *Achatina fulica*, across provinces in Thailand. Multilocus data (mitochondrial sequence data and 12 nuclear microsatellites) and individual based analyses were used to test for cryptic lineages. We found strong linkage disequilibrium patterns between mitochondrial haplotypes and nuclear-identified genetic clusters. There were clearly two divergent and independent clades. Moreover, within each clade, the data suggested additional substructure where individual provinces were likely to harbor unique genetic clusters. The results indicate there are at minimum two and possibly up to eight cryptic lineages within the assumed single species of *A. cantonensis*. Importantly, the two main clades do not show geographic affiliation and can be found in sympatry. With recent studies highlighting *A. cantonensis* strain diversity in pathogenicity and infectivity, it will be important to determine if these critical epidemiological traits are associated with specific lineages.



## 1. Introduction

The delimitation and discovery of species among parasitic pathogens is not simply a taxonomic exercise, but rather a scientific endeavor of utmost epidemiological relevance. The failure to identify cryptic species among medically relevant pathogens could lead to misinterpretation of virulence and pathogenicity as such factors could be lineage-specific (Nadler and Pérez-Ponce de León, 2011; Zingales et al., 2012). Likewise, development of diagnostics, therapeutics, control programs, or surveillance projects could be hampered without proper knowledge of lineage diversity (Nadler and Pérez-Ponce de León, 2011). Here, we define cryptic species as morphologically similar, but genetically distinct lineages with largely independent evolutionary histories. While molecular population genetic and phylogenetic analyses have proven essential for understanding complex evolutionarily epidemiological topics, e.g., parasite transmission among individual hosts, broad scale pathogen dispersal, cross-species transmission, pathogen hybridization, spread of drug resistance (Criscione et al., 2010; Detwiler and Criscione, 2010; Faria et al., 2013; Pybus and Rambaut, 2009; Roper et al., 2004), the basic, but fundamental, application of using genetic markers to identify species is still a necessity among many pathogenic organisms.

The goal of this study was to use molecular population genetic and phylogenetic data to ascertain if there are cryptic lineages/

* Corresponding authors.
*E-mail addresses:* ccriscione@bio.tamu.edu (C.D. Criscione), urusa.tha@mahidol.ac.th (U. Thaenkham).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ympev.2016.12.002
1055-7903/

species within the presumed single species *Angiostrongylus cantonensis*, the rat lungworm. *Angiostrongylus cantonensis* is an emerging zoonotic with global health concerns (Wang et al., 2012). This parasitic nematode is transmitted between rat definitive hosts (where male and female worms reach sexual maturity) and gastropod intermediate hosts, such as slugs or snails, in its natural life cycle. However, if humans ingest the third stage larvae (L3) from a gastropod host or various paratenic hosts (additional hosts such as crustaceans and frogs that may harbor non-developing L3 after the gastropod), the parasite can cause angiostrongyliasis with the main clinical manifestation of eosinophilic meningitis (Wang et al., 2012). Death can occur in severe cases without prompt and proper treatment (Bowden, 1981; Chotmongkol and Sawanyawisuth, 2002; Punyagupta et al., 1990; Yii, 1976). To date, more than 2800 human cases have been documented in approximately 30 countries worldwide (Wang et al., 2008), though many cases could have gone unreported or unrecognized during the past decades (Kliks and Palumbo, 1992). In Thailand alone, at least 1337 cases (47.33% of cases worldwide) of human angiostrongyliasis have been reported (Wang et al., 2008). The ability to infect many gastropod species and rats, a human commensal, has enabled *A. cantonensis* to achieve a global distribution beyond its assumed native range in Southeast Asia and the South Pacific (Kliks and Palumbo, 1992; Lv et al., 2009; Tesana et al., 2009). Coupled with the subtle morphological differences among species in the genus *Angiostrongylus* (Ubelaker, 1986), the broad geographic distribution and wide array of host reports for *A. cantonensis* make it a candidate species that might actually be composed of cryptic lineages (Pérez-Ponce de León and Choudhury, 2010).

Molecular phylogenetic or population genetic studies on the rat lungworm are relatively recent and have largely focused on relationships among species of *Angiostrongylus*. Nonetheless, the available molecular data are suggestive of cryptic species within what is called *A. cantonensis*. For example, Eamsobhana et al. (2010) reported two mitochondrial (mtDNA) clades [cytochrome *c* oxidase subunit I (*CO*1)] up to 3.6% divergent with 5 samples from Thailand in one clade and 7 samples from Hawaii and China in the other (Eamsobhana et al., 2010). *CO*1 haplotypes up to 6.4% different were reported by Monte et al. (2012). Liu et al. (2011) reported 0.8–1.3% divergence in the nuclear internal transcribed spacer 2 (ITS2) between the 1 sample from the Philippines to the 10 samples from China (Liu et al., 2011). While genetic distance itself is not a delimiter of species, these levels of mtDNA or ITS divergence are similar to what is found among morphologically recognized species of helminth parasites and thus, suggest possible cryptic lineages (Blouin, 2002; Vilas et al., 2005). A recent whole mitochondrial comparison of individual nematodes from Taiwan, Thailand and China also revealed large differences among samples (Yong et al., 2015). The above studies, however, used single markers (the lack of recombination renders the mitochondria a single marker even if sequenced across multiple coding genes) often from a limited number of specimens from a given location, thus it is not possible to conclude if these divergent sequences represent independent lineages (i.e., gene trees may not reflect species trees). One study that did use multiple markers, random amplified polymorphic DNA markers (RAPDs), from individual nematodes found two divergent clades of *A. cantonensis* within Thailand (Thaenkham et al., 2012). However, this study used provinces as *a priori* delimitations of populations (Thaenkham et al., 2012). Therefore, if there were co-occurring cryptic lineages within provinces, RAPD allele frequency estimates, and hence estimates of divergence or inference of relationships among population samples would be distorted (i.e., there would be a Wahlund effect). More appropriate and powerful analyses would be based on individual nematodes from which multiple independent markers would be typed.

We conducted an extensive survey of Thailand to address if indeed there were cryptic species within the morphologically recognized species of *A. cantonensis*. Individual based population genetic analyses revealed concordance between mtDNA haplotypes and genetic clusters that were elucidated with nuclear microsatellites. These data strongly support two cryptic lineages among the individuals presumed to be the single species *A. cantonensis*. The data are also suggestive of additional lineages within each of these two major lineages.

## 2. Methods

### 2.1. Sample collection and DNA extraction

L3 of *Angiostrongylus* was collected from the giant African land snail, *Achatina fulica*. Snails were captured by hand between June 2009 and May 2011 from 16 provinces around Thailand (Fig. 1 and Table 1): Nan (NAN), Phitsanulok (PSL), Maha Sarakham (MHS), Lop Buri (LB), Bangkok (BKK), Kanchanaburi (KCB), Prachuap Khiri Khan (PCK), Nakhon Si Thammarat (NST), Chanthaburi (CTI), Chiang Mai (CMI), Nakhon Phanom (NPM), Narathiwat (NWT), Nong Khai (NKI), Phang Nga (PNA), Ranong (RNG), and Khon Kaen (KKN). An average of 10.9 infected snails (range 2–28) per province was used to collect L3. Snails from a given province were collected at the same time and from one village, though possibly more than one spot within a village (Table 1). After removing the hard shell of each snail, the snail's foot was chopped, blended, and then incubated with 1% HCl–pepsin solution at 37 °C for 1 h. L3 were isolated using Baermann's technique and morphologically identified as belonging to the genus *Angiostrongylus* (Alicata and Jindrák, 1970).

To confirm that the collected larvae belonged to the genus *Angiostrongylus*, 32 experimental rats (*Rattus norvegicus* obtained from National Laboratory Animal Center Mahidol University; NLAC-MU) were infected with the isolated larvae (two rats per province) and then euthanized after 3 months. Adult worms were isolated from the rats' pulmonary arteries and identified by morphological characters as *Angiostrongylus cantonensis* (Bhaibulaya, 1979). The identified adult worms were preserved separately by locality in absolute ethanol and stored at −20 °C until use. DNA extractions were performed following the protocol of the Genomic DNA Mini Kit (Tissue) (Geneaid Biotech Ltd., New Taipei City, Taiwan). The genomic DNA was stored at 4 °C until PCR.

Animal studies were carried out in strict accordance with the Guidelines for Ethical Principles and Guidelines for the Use of Animals (National Research Council of Thailand). The animal experiments were conducted in strict compliance with animal husbandry and welfare regulations. Use of laboratory reared Wistar rats, which are devoid of natural *Angiostrongylus* infections, was approved by National Laboratory Animal Center Mahidol University. The protocol was approved by the Animal Care and Use Committee of Faculty of Tropical Medicine – Mahidol University, Bangkok (No. FTM-ACUC 002/2013). After infection by the L3, the rats were cared for regularly and monitored for balanced nutrition and environmental enrichment by the Animal Care and Use Committee of Faculty of Tropical Medicine – Mahidol University, Bangkok. After 3 months infection, the rats were euthanized by a high concentration of ether.

### 2.2. Microsatellite genotyping and mtDNA sequencing

Methods describing the microsatellite library construction and sequencing are given in Andres and Bogdanowicz (2011). Library construction and sequencing were conducted by S. Bogdanowicz at the Evolutionary Genetics Core Facility at Cornell University,

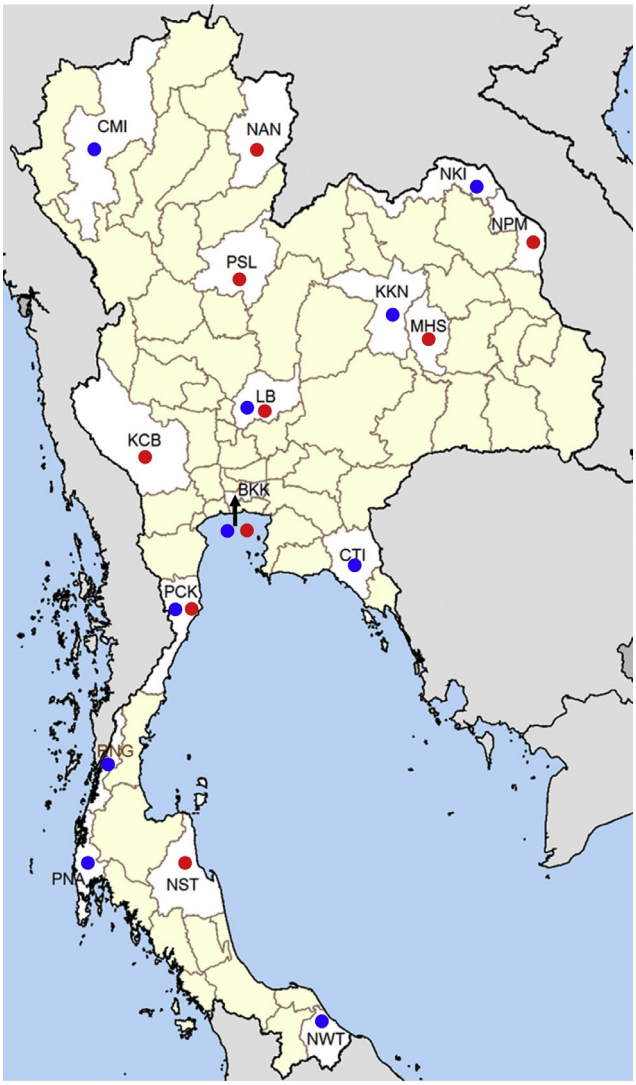


**Fig. 1.** Map of the sixteen collection localities of *Achatina fulica* in Thailand. Province abbreviations are as follows: Nan (NAN), Phitsanulok (PSL), Maha Sarakham (MHS), Lop Buri (LB), Bangkok (BKK), Kanchanaburi (KCB), Prachuap Khiri Khan (PCK), Nakhon Si Thammarat (NST), Chanthaburi (CTI), Chiang Mai (CMI), Nakhon Phanom (NPM), Narathiwat (NWT), Nong Khai (NKI), Phang Nga (PNA), Ranong (RNG), and Khon Kaen (KKN). Red and blue dots indicate that Clade I and Clade II, respectively, were detected in the province. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

USA. To design primers that flanked microsatellites, we blasted our microsatellite enriched sequences against the available draft genome of *A. cantonensis* (v1.5.4 available from the Welcome Trust Sanger Institute, http://www.sanger.ac.uk/resources/downloads/helminths/, or WormBase ParaSite PRJEB493). To help reduce the possibility of downstream null amplification, we designed primers where the sequences of the flanking regions from our independently constructed microsatellite library completely matched the sequences of the draft genome. Primer design, PCR, and genotyping methods of microsatellites were the same as those described in a prior publication (Detwiler and Criscione, 2011). Briefly, the M13 method for genotyping was used where a M13 oligonucleotide (TGTAAAACGACGGCCAGT) was added to the 5′-end of the forward primer (Schuelke, 2000). To reduce polyadenylation, a 5′ sequence tag (GTTTCTT) was added to the reverse primer (Brownstein et al., 1996). Supplementary Table S1 provides the primers for the 12 microsatellite loci used in this study.

Partial mitochondrial *CYTB* sequences were amplified using the PCR conditions and specific primers cytb-F, 5′-TGAATAGACA GAATTTTAAGAG-3′ and cytb-R, 5′-ATCAACTTAACATTACAGAAAC-3′ described in a previous study (Dusitsittipon et al., 2015). With these primers, some samples either failed to amplify or some samples produced sequences that translated with premature stop-codons (based on translation table 5, the invertebrate mtDNA code in GenBank). The latter indicates the potential presence of a pseudogene, but these data will be presented in a separate publication. In order to obtain sequences from the above problem samples, we designed new primers that enabled us to obtain *CYTB* sequences with uninterrupted translations. These new primers, cytb-F3 5′-A TGTTATCTTGATAAGGTAGAG-3′ and cytb-R2 5′-AACCTAAATTCAA TATACATACTAT-3′, were designed by looking for conserved regions that flanked the *CYTB* gene in an alignment using the complete mitochondrial *A. cantonensis* genome samples in GenBank, NC_013065 and KT186242, which originated from China and Thailand, respectively (Lv et al., 2012; Yong et al., 2015). Primer cytb-F3 is in the tRNA-Isoleucine, which is 5′ of *CYTB*, and primer cytb-R2 is in the tRNA-Leucine, which is 3′ of *CYTB*. PCR conditions were as described in a previous study (Dusitsittipon et al., 2015). Each PCR contained about 1 ng of genomic DNA from a single worm, 1.5 mM $Mg_2Cl$, 0.2 mM of each dNTP, 0.4 μM of primers, and 1.25 U of GoTaq DNA Polymerase (Promega Corporation, Madison, WI, USA). Ultra-pure water was added to a 50-μL final volume. PCR products were run on 2.0% agarose gel with ethidium bromide staining and visualized on a UV illuminator. PCR product purification was conducted using the UltraClean PCR Clean-Up Kit (MO BIO Laboratories, Inc., Carlsbad, CA, USA), and the eluted PCR products were then quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, NC, USA) prior to DNA sequencing. All purified PCR products were sequenced by Big-Dye Terminator chemistry with the specific PCR primers described above using an Applied Biosystems 3730xl DNA Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA).

### 2.3. Microsatellite analyses

Initial inspection of the genotype data showed that many individuals had identical multilocus genotypes (MLGs) and that several of these MLGs were often homozygous across all or most of the 12 microsatellite loci (Supplementary Fig. S1A). We returned to this unexpected pattern in the Results and Discussion. However, we noted these results here because it influenced the analyses we used.

While the broader geographic sampling increased our chances of detecting cryptic lineages across Thailand, it came at the expense of smaller sample sizes within some provinces (Table 1). Thus, we utilized individual based clustering analyses to test for overarching patterns of structure within the data set. Because of the repeated MLGs, we first reduced the complexity of the complete microsatellite data set ($n = 182$) by using only one copy per unique MLG. We emphasize that qualitatively, clustering results were the same when repeated copies of MLGs were included. Repeated MLGs were identified in Genalex v6.501 (Peakall and Smouse, 2012). Moreover, it was clear that one individual, BKK-5, was a F1 hybrid between two cryptic species (details given in Results and Discussion). Thus this individual was also removed as we did not want this individual to misrepresent genetic distance relationships among clusters or individuals (neighbor-joining trees cannot deal with reticulations, which hybridization causes). The reduced data set of unique MLGs ($n = 82$) was analyzed two ways. First, we used the Bayesian genetic clustering algorithm in Baps v6.0 (Corander and Marttinen, 2006; Corrander et al., 2006). In Baps, the mixture analysis was based on the 'clustering of individuals' option. We ran 10 replicates of $k$ (the number of possible

**Table 1**
Samples of *Angiostrongylus* spp. ($n$ = 192) used in the *CYTB* sequencing ($n$ = 177) and microsatellite ($n$ = 182) analyses.

| Locality (n) Year | Provinces | Co-ordinates | Sequence *CYTB* only F&R[a] | Microsatellite loci only | Both *CYTB* and Microsatellite loci | |
|---|---|---|---|---|---|---|
| | | | | | F&R[a] | F3&R2[a] |
| BKK (12) 2009 | Bangkok | 13°45′0″N 100°31′1.20″E | – | – | BKK 1–5 | BKK 6–9, 11, 12, 14 |
| KCB (19) 2009 | Kanchanaburi | 14°6′56″N 99°8′40″E | KCB 1, 10 (F&R) | KCBD 5 | KCB 2–9, 11–14, KCBD 6–8, 13 | – |
| LB (21) 2011 | Lop Buri | 14°51′21″N 100°59′24″E | LB 1–6 (F&R) | LB 17, 18 | LB 7–16 | LB 19–21 |
| MHS (14) 2010 | Maha Sarakham | 16°2′18″N 103°7′9″E | – | – | MHS 1–14 | – |
| NAN (12) 2010 | Nan | 19°23′14″N 100°52′33″E | NAN 6 (F&R) | – | NAN 1–5, 7–11, 13 | – |
| NST (14) 2010 | Nakhon Si Thammarat | 8°40′0″N 99°55′54″E | – | – | NST 4–7, 9–13, 16–20 | – |
| PCK (19) 2010 | Prachuap Khiri Khan | 12°23′42″N 99°55′0″E | – | PCK 9, 15–19 | PCK 1, 3, 4, 8, 11, 14 | PCK 2, 5–7, 10, 12–13 |
| PSL (11) 2011 | Phitsanulok | 16°49′29″N 100°15′34″E | – | – | PSL 1–11 | – |
| CTI (3) 2011 | Chanthaburi | 12°36′3″N 102°16′14″E | – | – | – | CTI 1–2, 16 |
| CMI (24) 2011 | Chiang Mai | 18°47′43″N 98°59′55″E | – | CMI 7, 13–15 | – | CMI 2, 3, 5, 6, 8–12, 16–26 |
| NPM (13) 2010 | Nakhon Phanom | 17°24′25″N 104°46′51″E | NPM 1 (F&R) | – | NPM 2–11, 13, 14 | – |
| NWT (2) 2011 | Narathiwat | 6°25′N 101°49′E | – | – | – | NWT 7, 9 |
| NKI (6) 2009 | Nong Khai | 17°52′5″N 102°44′40″E | – | – | – | NKI 1–3, 6, 7, 10 |
| PNA (14) 2010 | Phang Nga | 8°27′52″N 98°31′54″E | – | PNA 8 | – | PNA 2, 4–7, 9, 10, PNAs 1–6 |
| RNG (7) 2011 | Ranong | 9°57′43″N 98°38′20″E | – | RNG 8 | – | RNG 1–5, 7 |
| KKN (1) 2009 | Khon Kaen | 16°26′N 102°50′E | – | – | – | KKN 6 |
| Total (192) | | | 10 | 15 | 99 | 68 |

[a] F&R and F3&R2 indicate the primer pairs used to amplify the *CYTB* gene.

clusters) ranging from 2 to 20 then constructed a neighbor-joining tree with the Kullback–Leibler divergence matrix, which is provided as output with Baps. This matrix can be used as a measure of relative genetic distance between the Baps-identified clusters. The Baps algorithm makes assumptions of Hardy-Weinberg equilibrium (HWE) and linkage equilibrium to perform the clustering, both of which are likely violated in our dataset. Thus, a second clustering analysis was used that makes no assumptions of HWE or linkage equilibrium. The second analysis entailed the construction of a neighbor-joining tree of unique MLGs ($n$ = 82) where genetic distance among MLGs was measured as 1 minus the proportion of shared alleles (*Dps*). The distance measure was calculated in Msa v 4.05 (Dieringer and Schlötterer, 2003). To test the robustness of clades, 1000 bootstraps over loci were conducted. Neighbor-joining trees were constructed in PHYLIP v3.695 (Felsenstein, 2005).

Upon identifying clusters, we ran post hoc tests of HWE and genotypic disequilibrium across all or subsets of the data set. These analyses excluded the potential hybrid individual BKK-5. The purpose of these post hoc analyses was to provide a description of the structure among lineages. Weir and Cockerham's estimator of $F_{IS}$ (per locus and multilocus) was computed with Spagedi 1.3 (Hardy and Vekemans, 2002). Deviations from HWE (per locus and multilocus estimates of $F_{IS}$) were tested by permutating alleles among individuals 10,000 times in Spagedi. Genotypic disequilibrium between pairs of loci was tested with Genepop 4.3 with Markov chain parameters of 5000 dememorizations, 5000 batches and 5000 iterations (Rousset, 2008). We also tested HWE and genotypic equilibrium with the CMI samples because this province had a reasonable sample size ($n$ = 24) and because all individuals had a unique MLG (Supplementary Fig. S1).

### 2.4. Phylogenetic analyses of mtDNA

In the study by Dusitsittipon et al. (2015), they obtained 91 *CYTB* sequences from samples across Thailand. These samples are included in this study (GenBank numbers given in Supplementary Table S2) because many of these same samples were also genotyped at the microsatellite loci mentioned above. In addition, we obtained *CYTB* sequences from 86 new samples for the current study (Supplementary Table S2). These 177 sequences were aligned using ClustalX 2.1 (Thompson et al., 1997). To verify that we obtained the coding gene as opposed to a pseudogene, we translated the DNA sequences into amino acids. As additional verification, we compared our sequences to *CYTB* from the complete mitochondrial genome samples in GenBank, NC_013065 and KT186242. For descriptive purposes, pairwise nucleotide and amino acid distances among *CYTB* haplotypes or clades were estimated with uncorrected *p*-distance using Mega6 (Tamura et al., 2013).

Phylogenetic relationships among the unique *CYTB* haplotypes were assessed using Bayesian inference as implemented in MrBayes (Ronquist et al., 2012). In addition to our collected samples, we included the *A. cantonensis* samples NC_013065 and KT186242 from Genbank. For outgroups, we included NC_013067 (*Angiostrongylus costaricensis*) and NC_018602 (*Angiostrongylus vasorum*). Before constructing a phylogenetic tree, PartitionFinder v1.1.1 (Lanfear et al., 2012) was used to select the best-fit partitioning scheme and models of molecular evolution implemented

in MrBayes. The three data blocks in PartitionFinder corresponded to the three codon positions. The BIC (Bayesian Information Criterion) metric was used for model selection across all possible partitioning schemes. Based on the results of PartitionFinder, the first and second codon positions were analyzed with the HKY + I model and the third position was analyzed with the HKY model in MrBayes. The relevant nucleotide model substitution parameters (stationary nucleotide frequencies, proportion of invariable sites, and transition-transversion rate ratio) were set to unlinked across the three character sets (i.e., three codon positions). The character sets were allowed to evolve under different rates (ratepr = variable). Two parallel runs each with 4 chains were set for 2,000,000 generations with sampling every 1000th generation. The first 500 samples (25%) were discarded as burn-in. The run length was deemed appropriate based on the convergence of the parallel runs (standard deviation of split frequencies < 0.01). Also, for all model parameters, the potential scale reduction factors approximated 1, thus indicating a good sample from the posterior probability distribution. We also did phylogenetic analyses using maximum likelihood in Mega6 using the HKY + G model as selected in MEGA6. One thousand bootstraps were run to assess node support.

## 3. Results

### 3.1. Microsatellite results

Excluding BKK-5, there were 82 unique MLGs among the 181 individuals genotyped at the 12 loci; 22 MLGs had more than one copy. MLGs tended to be homozygous across loci; 71% (128/181) of the individuals were homozygous for 11 or 12 of the 12 loci (Supplementary Fig. S1). For example, all individuals genotyped from the KCB province ($n = 17$; indicated by the MLG labeled KCB14 in Fig. 2) had the same MLG and were homozygous at all loci.

The two methods of clustering (Baps and 1- *Dps*) produced identical results with regards to the placement of unique MLGs ($n = 82$) into clusters. To reiterate, clustering results were qualitatively the same if repeated copies of MLGs were included ($n = 181$; results not shown). The Baps analysis had a posterior probability of 0.99 for 9 clusters. Fig. 2 shows the neighbor-joining tree based on 1- *Dps* where the 9 black nodes represent the same cluster assignments of unique MLGs as that found in the optimal partition of the Baps analysis. Bootstrap support of these 9 nodes was typically strong (>75%) for all clusters with the exception of Cluster 1 (31% support) and Cluster 7 (55% support). Also, with the exception of the placement of Clusters 1 and 2, relationships among the clusters were similar between the Baps and 1- *Dps* analyses (Fig. 2). In particular, both methods showed a divergent split between two groupings (Clusters 1–6 vs 7–9) (Fig. 2). We refer to these two groupings as Clade I and Clade II as they are 100% concordant with two divergent and well supported mtDNA clades (discussed below). For 8 of the 12 microsatellite loci, there were no alleles shared between these two clades. We also note that loci 17B19952 and 26B46312 were likely duplicated, but only from samples that belonged to Clade II. Because no alleles were shared between the two clades for loci 17B19952 and 26B46312, it was possible to score these two loci as single locus genotypes. For example, at locus 17B19952, all individuals from Clade I had a 148/148 genotype. In Clade II, individuals always had 2 or 3 alleles where the 2nd or 3rd allele was always allele 287. We assumed this locus was duplicated and scored individuals ignoring the 287 allele. Thus, an individual with peaks showing at 174 and 287 were scored as 174/174 and an individual with peaks showing at 162, 174, and 287 were scored as 162/174. At locus 26B46312, all individuals from Clade I had a 172/172 genotype whereas all individuals from Clade II had 182/184. We scored the Clade II individuals as 182/182 assuming this was a duplicated locus within Clade II. Levels of allelic and gene diversity by cluster are given in Supplementary Table S3.

There is no apparent geographic association of the two clades in Thailand (Fig. 1). In fact, the two clades are sympatric within 3 provinces, LB, BKK, and PCK. Interestingly, the putative hybrid, BKK-5, was from one of these provinces where there was sympatry of the two clades. Visual inspection of the data revealed that BKK-5 was heterozygous at every locus with alleles combined from Cluster 2 (Clade I) and Cluster 8 (Clade II). In the Discussion, we present evidence that Clades I and II represent genetically distinct species. Thus, BKK-5 would be classified as a F1 hybrid between these two species (genotype data are given in Supplementary Table S4).

Although there was no geographic pattern associated with the two clades, clusters within the clades were largely restricted to individual provinces (Fig. 2). In contrast to the other clusters, Cluster 8 contained individuals from 7 provinces. Specific patterns for the other clusters are as follows: In Clade I, samples from PSL and NPM were exclusive to and the only samples that made up Clusters 3 and 4, respectively. Cluster 6 consisted of all the individuals from NST and one individual from NAN. Cluster 1 was composed of samples from PCK and KCB where all 17 of the KCB samples had the same MLG. In Cluster 5, all MHS samples can be found. In addition, Cluster 5 had one MLG (which fell outside the cluster of MHS samples) with 20 copies that was found in NAN and LB. In Clade II, Cluster 7 is made up of all and only the CMI samples and Cluster 9 is composed of all NKI samples, 1 from NWT and 1 from KKN.

Cluster 7, which is exclusive to the CMI province, is the only cluster where every individual had a unique MLG (Supplementary Fig. S1H). In this population sample, 10 of 12 loci were polymorphic (Supplementary Table S1). Of the 10 polymorphic loci, only 1 locus (locus 7R5254) had a deviation from HWE ($p = 0.014$, $F_{IS} = 0.57$, Supplementary Table S1). However, this locus was not significant with a Bonferroni correction. The multilocus $F_{IS}$ was 0.037 and did not deviate from HWE ($p = 0.47$). There was significant genotypic disequilibrium within the CMI sample where 9 out of 45 possible pairwise comparisons had $p < 0.05$ (after a sequential Bonferroni correction, 2 pairwise comparisons remained significant).

### 3.2. mtDNA results

We obtained 852 nucleotides of partial *CYTB* sequences that could be aligned with no gaps among the 177 sequenced worms. Among these 177 Thailand samples, there were 135 (15.8%) variable sites (142 total observed substitutions: 115 synonymous changes and 27 nonsynonymous changes) from 21 unique *CYTB* haplotypes. The two GenBank samples, NC_013065 and KT186242 also constituted different haplotypes. Two strongly supported (posterior probabilities = 99 and 100%) and divergent clades of *A. cantonensis* were revealed by the Bayesian phylogenetic analysis; Clade I consisted of haplotypes H1–H12 and NC_013065 whereas Clade II had haplotypes H13–H21 and KT186242 (Fig. 3). The maximum likelihood analysis qualitatively yielded a similar topology; bootstrap support for nodes are superimposed on the Bayesian-based tree (Fig. 3). Average *p*-distance between Clades I and II was 11.4 ± 1.0% (range: 10.8–12.1%) (averaged over all haplotype pairs between the two clades ± standard error from 1000 bootstrap replicates). Average *p*-distance within Clade I was 2.6 ± 0.3% (range: 0.1–6.1%) and within Clade II was 0.5 ± 0.1% (range: 0.1–1.2%). The evolutionary distance between Clade I and Clade II is further emphasized when looking at amino acid differences (283 amino acids). Average amino acid *p*-distance between

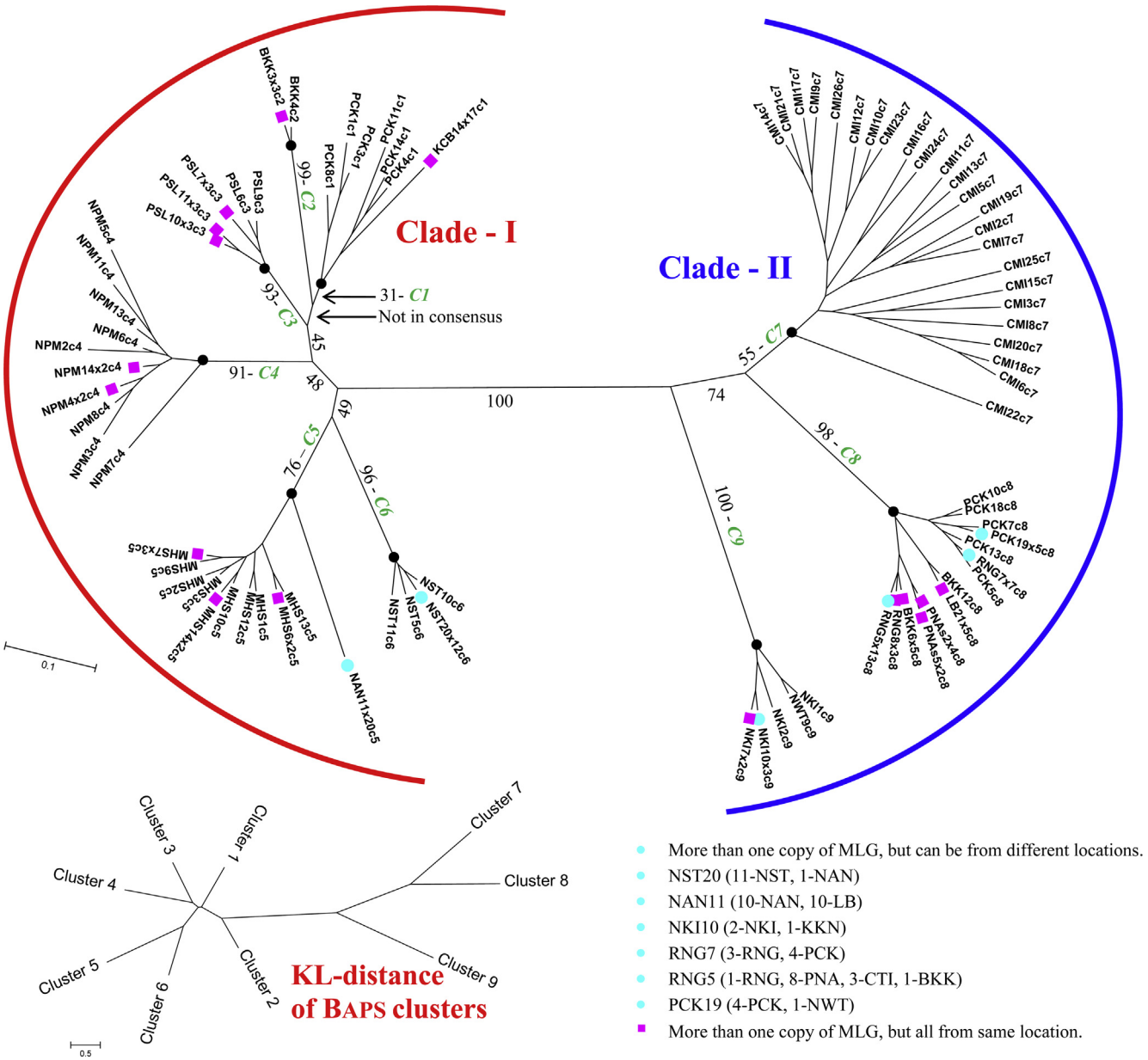


**Fig. 2.** Genetic clusters based on microsatellite analyses. The small inset figure on the bottom left shows the unrooted relationships among the BAPS identified clusters. The tree is based on a Neighbor-Joining analysis of the Kullback-Leibler divergence among clusters. The large figure is an unrooted Neighbor-Joining tree based on the distance measure 1-*Dps*. Taxa labels are as follows: the first letters followed by a number give the location and individual worm identification. If the label is followed by "x#", this indicates the number of individual worms that had the same MLG. The "c#" indicates the cluster that individual was assigned to in the BAPS analysis. For example, MHS13c5 is worm 13 from Maha Sarakham and this worm was assigned to Cluster 5 based on the BAPS analysis. For LB21x5c8, there were 5 worms that had this MLG including worm 21 from Lop Buri; this MLG was assigned to Cluster 8 based on the BAPS analysis. The pink squares indicate MLGs with more than one copy and where all of these copies originated from the same province. The teal circles also indicate MLGs with more than one copy; however, the copies could be found in different provinces. The small text inset on the bottom right indicates these locations. For example, 3 copies of RNG7x7c8 are found in RNG and 4 copies are found in PCK. Notice the formation of clades based on the proportion of shared alleles is 100% congruent with the BAPS identified clusters. The 9 black circles on the tree indicate the clade nodes that correspond to the BAPS identified clusters. Percent bootstrap support for these nodes followed by the cluster number "C#" is given on the branch leading to the node. Bootstrap support values for all branches that were internal to these nodes are also given. As an example, all samples from NPM were assigned to Cluster 4 in the BAPS analysis and these samples all form a clade with 91% bootstrap support. Lastly, the 1-*Dps* Neighbor-Joining tree shows the major clade designations (I and II) that correspond to the 2 major *CYTB* mtDNA clades. That is, for all samples for which we genotyped at microsatellites and also sequenced the *CYTB*, we found perfect congruence where all individuals from Clusters 1–6 had a mtDNA haplotype from Clade I and all individuals from Clusters 7–9 had a haplotype from Clade II. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Clade I and II was 7.4 ± 1.5%. Average within *p*-distances was 0.9 ± 0.3 for Clade I and 0 for Clade II.

Internal nodes were well supported with posterior probabilities generally being greater than 95% (Fig. 3). The one exception was the placement of haplotype H12 and NC_013065 with a posterior probability of only 53%. Many of the tip relationships are polytomies reflecting the fact that the haplotypes were separated by only a few mutations at the tips (Fig. 4). For example, within Clade II, a statistical parsimony network (constructed with TCS v1.21; (Clement et al., 2000)) shows most haplotypes only differ by 1–4 mutations (Fig. 4B). The phylogeny largely remained the same if we allowed MrBayes to integrate over evolutionary model uncertainty (i.e., sampling occurred across the entire general time reversible model space). The only difference was that haplotype H12

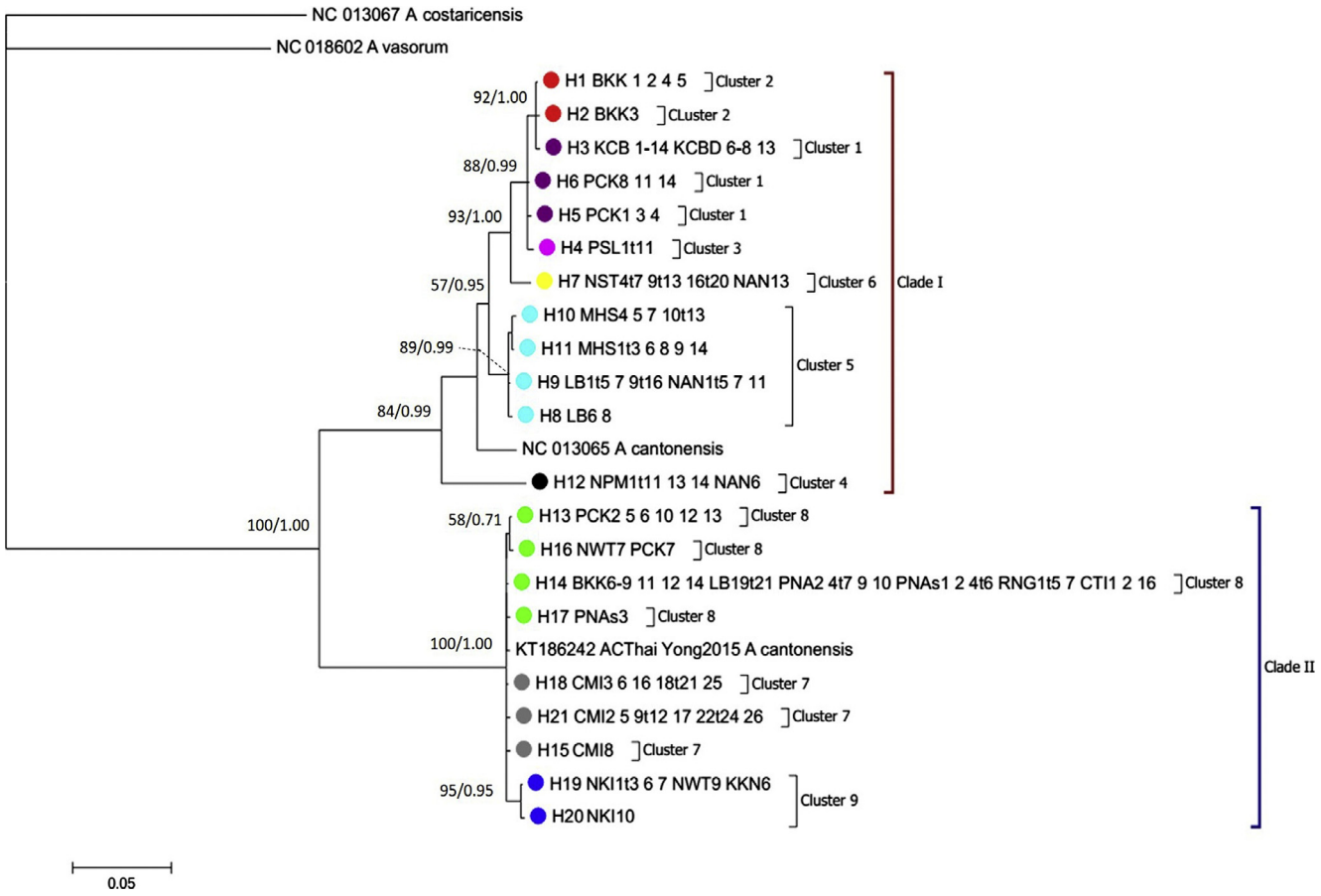


**Fig. 3.** Bayesian phylogeny of the unique *CYTB* mtDNA haplotypes. Numbers at branch nodes indicate bootstrap support/posterior probabilities. Taxa labels H1-H21 indicate the 21 unique haplotypes observed in the study. Next to the haplotype identifiers is a list of the sample identifications (province code followed by individual worm number within province) indicating which individual worms had which haplotype. Haplotype identifiers are also color coded to indicate which microsatellite cluster the haplotype is associated with (see Fig. 4 for the key). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and NC_013065 switched places (as was also the case in the maximum-likelihood analysis), but again support for this node was low with a posterior probability of only 60%.

### 3.3. Combined mtDNA and microsatellite results

We ended up with 167 individuals from which we both genotyped the microsatellites and sequenced partial *CYTB*. An intriguing result was the strong concordance between the mtDNA haplotypes and microsatellite clusters. In other words, if more than one individual had a given haplotype, those individuals were almost always assigned to the same microsatellite cluster (Fig. 4). All individuals with haplotypes H1-H12 were found in Clusters 1–6 and all individuals with haplotypes H13-H21 were found in Clusters 7–9 (Figs. 3 and 4). Thus, the deepest divergence in the mtDNA was completely congruent with the deepest divergence in the microsatellite analyses.

Within Clades I and II, there was further evidence of non-random association patterns between mtDNA haplotypes and nuclear microsatellite clusters. In Clade I there are 4 clear disequilibrium patterns: (1) H4 with Cluster 3, (2) H7 with Cluster 6, (3) H12 with Cluster 4, and (4) H8-H11 with Cluster 5 (Fig. 4A). We note that within Cluster 5 there were 20 individuals (10 from NAN and 10 from LB) that had the same MLG (indicated by NAN11 in Fig. 2) and these individuals only had haplotypes H8 or H9. The remaining 14 individuals in Cluster 5 came from MHS and they all had haplotypes H10 or H11 (Fig. 4). Thus, there exists possible subdivision within Cluster 5. The patterns of association between Clusters 1 and 2 with haplotypes H1-H3, H5, and H6 are not completely concordant and bootstrap support is low for Cluster 1 (Fig. 2). Thus, a conservative interpretation of the data would indicate that H1-H3, H5, and H6 with Clusters 1 and 2 represent a single, and thus, fifth association within Clade I.

Within Clade II, there are 3 non-random associations of mtDNA haplotypes and microsatellite clusters: (1) H13, H14, H16, and H17 with Cluster 8, (2) H15, H18, and H21 with Cluster 7, and (3) H19 and H20 with Cluster 9. We recognize there are few nucleotide differences among haplotypes within Clade II. For example, there is only a single mutation between H14 and H18. However, the relationships among haplotypes in the network show a non-random pattern with the microsatellite clustering where the branch between H14 and H18 is a clear demarcation of this disequilibrium (Fig. 4). The low level of mtDNA differences may simply reflect a recent splitting among the clusters or possible maintenance of similar mtDNA haplotypes via selection.

As a point of reference, we note that the mtDNA sequence NC_013065, which originated from a lab strain in Lijiang, China (Lv et al., 2012), falls within Clade I and the mtDNA sample KT186242, which originated from a lab strain in Thailand (Yong et al., 2015), falls within Clade II. Also, a qualitative assessment of microsatellite allele sizes suggests that the sample used to generate our microsatellite library belongs in Clade II, whereas the draft genome sample, which originated from Taiwan (PRJEB493 WormBase ParaSite), falls in Clade I. The latter is confirmed by

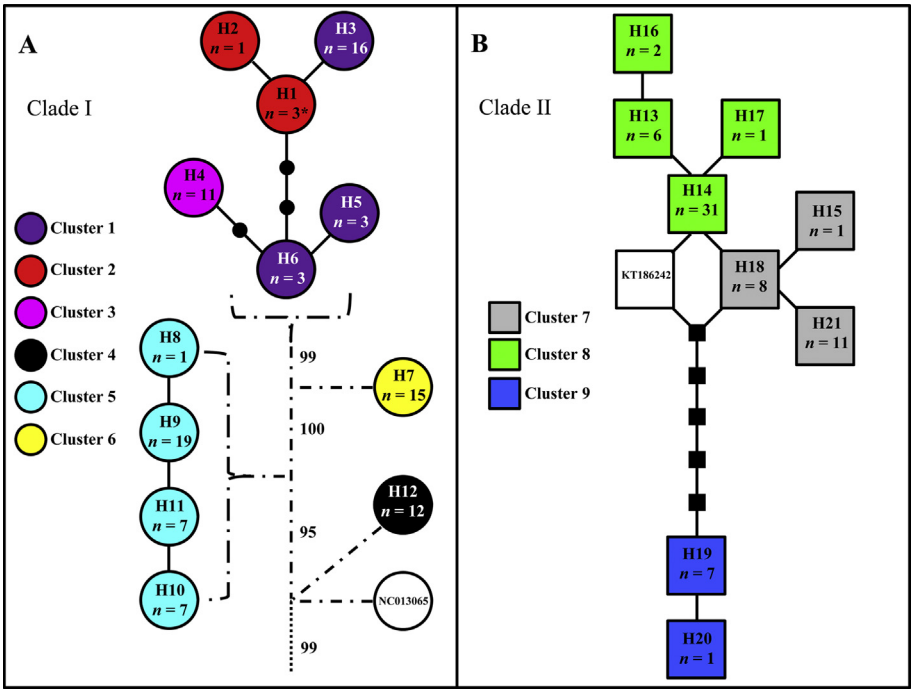


**Fig. 4.** Networks of *CYTB* haplotypes within A) Clade I and B) Clade II indicating the microsatellite cluster assignment of individuals that were both sequenced and genotyped ($n$ = 167). Sampled haplotypes are indicated by colored circles (Clade I) or squares (Clade II) where haplotype number is given by H#. The colors and shapes represent the microsatellite cluster that the individuals were assigned to. For example, all $n$ = 19 individuals that had haplotype H9 were assigned to Cluster 5 based on the microsatellite data. Solid line connections indicate a single mutational step and black circles or squares represent inferred haplotypes. Within each of the clades, haplotype networks were constructed with statistical parsimony (see text). In Clade I (A), haplotypes H7, H12 and NC_013065 did not connect due to the large number of nucleotide differences to the other haplotypes. Similarly, haplotypes H8-H11 and H1-H6 came out as separate networks. The dotted-dashed lines indicate the hypothesized relationships (posterior probabilities given along the branches) among the haplotype networks based on the Bayesian phylogenetic analysis in Fig. 3. The dotted line is the branch connection to the overall phylogeny given in Fig. 3. The asterisk in H1 indicates that a 4th individual (BKK-5) had this haplotype and that this individual was found to have a hybrid genotype between Clusters 2 and 8. In Clade II (B), a single network was constructed with haplotypes H13-H19 and KT186242. Ancestral connection to the network is not depicted as to better reflect the polytomy shown in Fig. 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

assessment of the mtDNA *CYTB* sequence in this draft genome (PRJEB493 WormBase ParaSite) (to be addressed in a future manuscript).

## 4. Discussion

There is strong support for the existence of two cryptic species from the genetic results based on samples of *A. cantonensis* collected across Thailand. First, we found two mtDNA lineages with an average of 11% $p$-distance, which is a level of mtDNA divergence commonly found among morphological recognized distinct species of nematodes (Blouin, 2002) and other helminth parasites (Vilas et al., 2005). We recognize that genetic distance at a single locus is in itself not a criterion to delimit species. However, we have a set of 12 nuclear markers that collectively show the same pattern. At the broadest level, the 100% concordance between the mtDNA and nuclear-identified clusters makes it clear that at minimum, Clades I and II represent distinct species. Among the microsatellites themselves there is significant genotypic disequilibrium (all of the 66 pairwise loci associations were significant with $p < 0.05$ in both the reduced and full data sets). Indeed 8 of 12 loci (all assumed to be neutral) do not share alleles between these two clades. The only way for this complete linkage disequilibrium between the mtDNA and nuclear markers to be maintained is if there is no genetic exchange (e.g., sexual reproduction). Thus, under any species concept (e.g., biological or phylogenetic), we conclude there are two independently evolving lineages (i.e., two distinct species). These two lineages do not appear to have any geographic association and can be found in sympatry in 3 provinces (Fig. 1).

The finding of two divergent clades corroborates the results based on RAPD data in Thaenkham et al. (2012). Indeed, the pattern of population-based clustering seen in their study, i.e., whole province samples KCB and LB were in one group and CMI, CTI, KKN, NKI, and NWT were in a second group (Thaenkham et al., 2012), was consistent with our analyses in that individuals from these provinces clustered into Clades I and II, respectively (Figs. 1 and 3). Though, we did find some LB samples that fell out in Clade II. In addition, they (Thaenkham et al., 2012) noted that RAPD bands were largely monomorphic among samples within provinces, but polymorphic between provinces (i.e., province samples had distinctive RAPD banding patterns). We observed a similar pattern with the microsatellite data. Individual clusters, which represent distinctive allelic profiles, had low genetic diversity (mean gene diversity and mean number of alleles per locus per cluster was 0.14 and 1.6, respectively; Supplementary Table S3) and were largely restricted to a single province (Fig. 2). The main exception to the restricted distributions of clusters was Cluster 8, which was found in 7 provinces extending from LB to the southernmost NWT (Figs. 1–3).

Although we found patterns similar to that of Thaenkham et al. (2012), our analyses at the individual level elucidate several findings that are not possible when basing analyses on *a priori* grouped samples (i.e., provinces in the case of their study). First, they indicated that the province sample PCK had banding patterns that were intermediate between the two groups they identified (Thaenkham et al., 2012). In our individual based analyzes, PCK was one of the provinces that had individuals from both Clade I and Clade II (Fig. 2). Thus, the intermediate genetic divergence pattern of PCK found in Thaenkham et al. (2012) is likely explained by

the fact that their PCK sample was composed of individuals belonging to both clades.

Second, individual based analyses allow for the identification of hybrids as admixed genotypes are readily identified among divergent genotypes. Such was the case for sample BKK-5, which was heterozygous across all 12 microsatellite loci. Because the microsatellite clusters had distinct allelic profiles, it was possible to determine that BKK-5 was a F1 hybrid between a Cluster 2 female parent (as indicated by the H1 mtDNA haplotype) and Cluster 8 male parent. The finding of this hybrid in BKK makes sense given individuals from Clusters 2 and 8 were found in BKK. Although the finding of a hybrid suggests there has been sexual reproduction between Clades I and II, clearly the concordance between microsatellite and mtDNA markers is maintained even in sympatry. Thus, the finding of this hybrid does not negate the species status of Clades I and II. Moreover, the finding of a F1 hybrid does not necessitate introgression between the Clades as the hybrid may have low fitness or may not even be fertile.

Third, the individual based analyses of the microsatellites along with the mtDNA sequencing of the same individuals indicated that there are additional non-random associations between mtDNA haplotypes and nuclear clusters within each of the two clades. Overall, there appears to be 8 groupings (5 within Clade I and 3 within Clade 2) that may indicate independent lineages (Fig. 4). In the microsatellite data set alone there is evidence of additional substructure (i.e., the Wahlund effect) within each of the clades. For example, post hoc analyses showed that linkage disequilibrium among the microsatellites remained high within both clades. Within Clade I, 26 of 36 (72%) and 35 of 36 (97%) possible pairwise comparisons between loci showed significant ($p < 0.05$) genotypic disequilibrium when using only unique MLGs ($n = 39$) or including repeated MLGs ($n = 99$), respectively. In Clade II, 45 out of 45 possible pairwise comparisons had significant genotypic disequilibrium whether using a reduced data set ($n = 43$) or one with repeated MLGs ($n = 82$). There were also significant deviations from HWE within each clade. Within Clade I, multilocus $F_{IS}$ was 0.61 in the reduced data set ($n = 39$) and increased to 0.81 if repeated MLGs were included ($n = 99$). In Clade II, multilocus $F_{IS}$ was 0.5 in the reduced data set ($n = 43$) and 0.64 with repeated MLGs ($n = 82$) (Supplementary Table S5).

Do these 8 additional groups with mtDNA-nuclear associations also indicate independent lineages? It might be argued that few parasite families were sampled from each province, (i.e., all the samples from a location were descended from 1 or very few sets of parents), thus resulting in apparent mtDNA-microsatellite cluster concordance. Sampling a limited number of families per province would also explain the low levels of microsatellite diversity (Supplementary Table S3). Family-restricted sampling would have required the sampled snails to have acquired parasite larvae from a single rat host and that rat host to have few breeding worms. Unfortunately, such details on transmission are currently unknown. However, we sampled an average of 10.9 infected snails per province, thus we believe our sampling would not be limited to just one to three families (as suggested by number of mtDNA haplotypes within a location) of parasites per provinces. Moreover, there are 3 lines of evidence that suggest sampling in the provinces was not limited to a few families. First, in 3 provinces (BKK, LB, and PCK) we were able to sample individuals from both clades, which necessarily implies more than one family was sampled. Second, local family-based sampling does not explain why 6 MLGs can be found in more than one province (Fig. 2) and third, Cluster 8 is widely dispersed across 7 provinces. Despite the wide distribution of Cluster 8, there is a clear mtDNA-nuclear cluster association (Fig. 4) and there remains limited microsatellite diversity within this cluster leading to MLGs that are predominately homozygous (Supplementary Fig. S1). Thus, additional mechanisms beyond a family-limited sampling still need to be invoked to explain the limited microsatellite diversity within clusters and the cluster associations with mtDNA haplotypes.

Assuming then that family-limited sampling does not explain the substructure within Clades I and II, what can explain the mtDNA-nuclear cluster associations (i.e., lineage diversity), repeated MLGs across locations, and largely homozygous MLGs? The patterns of strong linkage disequilibrium and propagation of the same or largely unchanged MLGs are particularly peculiar given that *A. cantonensis* is a dioecious species with presumed sexual reproduction. Indeed, these patterns of repeated MLGs and linkage disequilibrium define a clonal-like structure, which is commonly observed among several parasitic protozoa that are capable of asexual propagation (Tibayrenc and Ayala, 2012). More intensive sampling on local scales is needed as our current data set has small province sample sizes (in part due to the cryptic lineages found in sympatry) and too little diversity within or across loci (Supplementary Table S3) to provide meaningful tests of HWE or linkage disequilibrium within clusters. Here, we provide three possible (not necessarily mutually exclusive) hypotheses to explore in future studies. Local populations may have extremely small effective sizes, which is in part supported by the excess genotypic disequilibrium, 9 of 45 pairwise comparisons, observed in the CMI sample. However, drift alone would not explain the repeated MLGs across provinces or the continuity of Cluster 8 across provinces. Second, although it would require a consistent rate of a high proportion of brother-sister mating, there could be extreme inbreeding. If a single mating pair colonizes a new location, repeated or similar MLGs could be propagated across locations. Such patterns have been reported in inbred organisms, though typically in self-compatible hermaphrodites (e.g., Bomblies et al., 2010). Third, we do not think asexual propagation is likely as this should lead to more fixed heterozygote MLGs (De Meeûs et al., 2006). However, there are forms of parthenogenetic reproduction (descent without the need for male gametes) that can lead to instantaneous homozygosity across all or large portions of the genome (e.g., see Fig. 3a in De Meeûs et al., 2007). Pseudogamy (sperm required to stimulate oocyte development, but sperm nucleus subsequently degenerates) followed by parthenogenesis has been reported in free-living nematodes of the *Rhabditis* group (Poinar and Hansen, 1983). *Angiostrongylus* belongs to the *Rhabditis* group (van Megen et al., 2009); however, to our knowledge, parthenogenetic reproduction has not been reported for *Angiostrongylus* spp.

In conclusion, regardless of the mechanism(s), it is clear that there are 2 and possibly up to 8 cryptic lineages within what was previously regarded as the single species of *A. cantonensis*. This cryptic lineage diversity has important epidemiological implications. In particular, two recent studies have highlighted how different isolates (defined by *CO*1 haplotypes) of *A. cantonensis* can vary in reproductive capacity, pathogenicity, or infectivity (Lee et al., 2014; Monte et al., 2014). Moreover, differences in severity and pathogenicity of the disease have been found among patients with angiostrongyliasis in Taiwan (Hwang et al., 1994; Tsai et al., 2001, 2004). Thus, major questions that need to be addressed include the following. Do the lineages we identified vary in important epidemiological parameters or pathogenicity? Are some or all of the lineages we identified responsible for human infections? If so and as the 2 major clades can occur in sympatry, it will be important to develop diagnostic or monitoring methods to accurately discriminate among lineages because the morphology may not allow accurate delimitation. The identification of a putative hybrid is also of epidemiological relevance as parasite hybridization could allow introgression of genes related to parasite infectivity, virulence, transmission, or host specificity (Detwiler and Criscione, 2010). Additional studies are necessary to confirm and assess the frequency of hybridization. Lastly, from an evolutionary perspec-

tive, the clonal-like structure is an intriguing pattern for a dioecious species. We have highlighted several hypotheses that warrant testing to elucidate the mechanism(s) that have generated this pattern.

## Conflict of interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This work was supported by Agricultural Research Development Agency (CRP5705020410); and Development and Promotion of Science and Technology Talents Project (DPST; Royal Government of Thailand scholarship). We thank our colleagues in Criscione Lab (Isabel Caballero, Emily Kasl, and Andrew Sakla) at the Department of Biology, Texas A&M University, TX, USA for providing helpful suggestions and technical support, particularly for the microsatellite work. We also thank the Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Thailand for technical support, particularly for part of the DNA analysis.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2016.12.002.

## References

Alicata, J.E., Jindrák, K., 1970. Angiostrongylosis in the Pacific and Southeast Asia. C. C. Thomas.

Andres, J.A., Bogdanowicz, S.M., 2011. Isolating microsatellite loci: looking back, looking ahead. Mol. Methods Evol. Genet. 772, 211–232.

Bhaibulaya, M., 1979. Morphology and taxonomy of major *Angiostrongylus* species of Eastern Asia and Australia. In: Studies on *Angiostrongylus* in Eastern Asia and Australia Taipei, Taiwan: US Naval Medical Research Unit, vol. 2, pp. 4–13.

Blouin, M.S., 2002. Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer. Int. J. Parasitol. 32, 527–531.

Bomblies, K., Yant, L., Laitinen, R.A., Kim, S.T., Hollister, J.D., Warthmann, N., et al., 2010. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. PLoS Genet. 6, e1000890.

Bowden, D.K., 1981. Eosinophilic meningitis in the New Hebrides: two outbreaks and two deaths. Am. J. Trop. Med. Hyg. 30, 1141–1143.

Brownstein, M.J., Carpten, J.D., Smith, J.R., 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20 (1004–1006), 1008–1010.

Chotmongkol, V., Sawanyawisuth, K., 2002. Clinical manifestations and outcome of patients with severe eosinophilic meningoencephalitis presumably caused by *Angiostrongylus cantonensis*. Southeast Asian J. Trop. Med. Public Health 33, 231–234.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659.

Corander, J., Marttinen, P., 2006. Bayesian identification of admixture events using multilocus molecular markers. Mol. Ecol. 15, 2833–2843.

Corrander, J., Marttinen, P., Mäntyniemi, S., 2006. A Bayesian method for identification of stock mixtures from molecular marker data. Fish. Bull. 104, 550–558.

Criscione, C.D., Anderson, J.D., Sudimack, D., Subedi, J., Upadhayay, R.P., Jha, B., et al., 2010. Landscape genetics reveals focal transmission of a human macroparasite. PLoS Negl. Trop. Dis. 4, e665.

De Meeûs, T., Lehmann, L., Balloux, F., 2006. Molecular epidemiology of clonal diploids: a quick overview and a short DIY (do it yourself) notice. Infect., Genet. Evol. 6, 163–170.

De Meeûs, T., Prugnolle, F., Agnew, P., 2007. Asexual reproduction: genetics and evolutionary aspects. Cell. Mol. Life Sci. 64, 1355–1372.

Detwiler, J.T., Criscione, C.D., 2010. An infectious topic in reticulate evolution: introgression and hybridization in animal parasites. Genes 1, 102–123.

Detwiler, J.T., Criscione, C.D., 2011. Testing Mendelian inheritance from field-collected parasites: revealing duplicated loci enables correct inference of reproductive mode and mating system. Int. J. Parasitol. 41, 1185–1195.

Dieringer, D., Schlötterer, C., 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes 3, 167–169.

Dusitsittipon, S., Thaenkham, U., Watthanakulpanich, D., Adisakwattana, P., Komalamisra, C., 2015. Genetic differences in the rat lungworm, *Angiostrongylus cantonensis* (Nematoda: Angiostrongylidae), in Thailand. J. Helminthol. 89, 545–551.

Eamsobhana, P., Lim, P.E., Solano, G., Zhang, H., Gan, X., Yong, H.S., 2010. Molecular differentiation of *Angiostrongylus* taxa (Nematoda: Angiostrongylidae) by *cytochrome c* oxidase subunit I (*COI*) gene sequences. Acta Trop. 116, 152–156.

Faria, N.R., Suchard, M.A., Rambaut, A., Streicker, D.G., Lemey, P., 2013. Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 368, 20120196.

Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package) Version 3.695. Distributed by the author Department of Genome Sciences, University of Washington Seattle.

Hardy, O.J., Vekemans, X., 2002. SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes 2, 618–620.

Hwang, K.P., Chen, E.R., Chen, T.S., 1994. Eosinophilic meningitis and meningoencephalitis in children. Zhonghua Min. Guo Xiao Er Ke Yi Xue Hui Za Zhi. 35, 124–135.

Kliks, M.M., Palumbo, N.E., 1992. Eosinophilic meningitis beyond the Pacific Basin: the global dispersal of a peridomestic zoonosis caused by *Angiostrongylus cantonensis*, the nematode lungworm of rats. Soc. Sci. Med. 34, 199–212.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.

Lee, J.D., Chung, L.Y., Wang, L.C., Lin, R.J., Wang, J.J., Tu, H.P., et al., 2014. Sequence analysis in partial genes of five isolates of *Angiostrongylus cantonensis* from Taiwan and biological comparison in infectivity and pathogenicity between two strains. Acta Trop. 133, 26–34.

Liu, C.Y., Zhang, R.L., Chen, M.X., Li, J., Ai, L., Wu, C.Y., et al., 2011. Characterisation of *Angiostrongylus cantonensis* isolates from China by sequences of internal transcribed spacers of nuclear ribosomal DNA. J. Anim. Vet. Adv. 10, 593–596.

Lv, S., Zhang, Y., Liu, H.X., Hu, L., Yang, K., Steinmann, P., et al., 2009. Invasive snails and an emerging infectious disease: results from the first national survey on *Angiostrongylus cantonensis* in China. PLoS Negl. Trop. Dis. 3, e368.

Lv, S., Zhang, Y., Zhang, L., Liu, Q., Liu, H.X., Hu, L., et al., 2012. The complete mitochondrial genome of the rodent intra-arterial nematodes *Angiostrongylus cantonensis* and *Angiostrongylus costaricensis*. Parasitol. Res. 111, 115–123.

Monte, T.C., Simoes, R.O., Oliveira, A.P., Novaes, C.F., Thiengo, S.C., Silva, A.J., et al., 2012. Phylogenetic relationship of the Brazilian isolates of the rat lungworm *Angiostrongylus cantonensis* (Nematoda: Metastrongylidae) employing mitochondrial COI gene sequence data. Parasit. Vectors 5, 248.

Monte, T.C.C., Gentile, R., Garcia, J., Mota, E., Santos, J.N., Maldonado, A., 2014. Brazilian *Angiostrongylus cantonensis* haplotypes, ac8 and ac9, have two different biological and morphological profiles. Mem. Inst. Oswaldo Cruz 109, 1057–1063.

Nadler, S.A., Pérez-Ponce de León, G., 2011. Integrating molecular and morphological approaches for characterizing parasite cryptic species: implications for parasitology. Parasitology 138, 1688–1709.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28, 2537–2539.

Pérez-Ponce de León, G., Choudhury, A., 2010. Parasite inventories and DNA-based taxonomy: lessons from helminths of freshwater fishes in a megadiverse country. J. Parasitol. 96, 236–244.

Poinar, G., Hansen, E., 1983. Sex and reproductive modifications in nematodes. Helminthol. Abstr. Series B, 145–163.

Punyagupta, S., Bunnag, T., Juttijudata, P., 1990. Eosinophilic meningitis in Thailand. Clinical and epidemiological characteristics of 162 patients with myeloencephalitis probably caused by *Gnathostoma spinigerum*. J. Neurol. Sci. 96, 241–256.

Pybus, O.G., Rambaut, A., 2009. Evolutionary analysis of the dynamics of viral infectious disease. Nat. Rev. Genet. 10, 540–550.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Roper, C., Pearce, R., Nair, S., Sharp, B., Nosten, F., Anderson, T., 2004. Intercontinental spread of pyrimethamine-resistant malaria. Science 305, 1124.

Rousset, F., 2008. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol. Ecol. Resour. 8, 103–106.

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 18, 233–234.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tesana, S., Srisawangwong, T., Sithithaworn, P., Laha, T., Andrews, R., 2009. Prevalence and intensity of infection with third stage larvae of *Angiostrongylus cantonensis* in mollusks from Northeast Thailand. Am. J. Trop. Med. Hyg. 80, 983–987.

Thaenkham, U., Pakdee, W., Nuamtanong, S., Maipanich, W., Pubampen, S., Sa-Nguankiat, S., et al., 2012. Population structure of *Angiostrongylus cantonensis* (Nematoda: Metastrongylidae) in Thailand based on PCR-RAPD markers. Southeast Asian J. Trop. Med. Public Health 43, 567–573.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Tibayrenc, M., Ayala, F.J., 2012. Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. Proc. Natl. Acad. Sci. USA 109, E3305–E3313.

Tsai, H.-C., Liu, Y.-C., Kunin, C.M., Lee, S.S.-J., Chen, Y.-S., Lin, H.-H., et al., 2001. Eosinophilic meningitis caused by *Angiostrongylus cantonensis*: report of 17 cases. Am. J. Med. 111, 109–114.

Tsai, H.C., Lee, S.S., Huang, C.K., Yen, C.M., Chen, E.R., Liu, Y.C., 2004. Outbreak of eosinophilic meningitis associated with drinking raw vegetable juice in southern Taiwan. Am. J. Trop. Med. Hyg. 71, 222–226.

Ubelaker, J.E., 1986. Systematics of species referred to the genus *Angiostrongylus*. J. Parasitol. 72, 237–244.

van Megen, H., Holovachov, O., Bongers, T., Bakker, J., Helder, J., van den Elsen, S., et al., 2009. A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. Nematology 11, 927–950.

Vilas, R., Criscione, C.D., Blouin, M.S., 2005. A comparison between mitochondrial DNA and the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. Parasitology 131, 839–846.

Wang, Q.P., Lai, D.H., Zhu, X.Q., Chen, X.G., Lun, Z.R., 2008. Human angiostrongyliasis. Lancet Infect. Dis. 8, 621–630.

Wang, Q.P., Wu, Z.D., Wei, J., Owen, R.L., Lun, Z.R., 2012. Human *Angiostrongylus cantonensis*: an update. Eur. J. Clin. Microbiol. Infect. Dis. 31, 389–395.

Yii, C.Y., 1976. Clinical observations on eosinophilic meningitis and meningoencephalitis caused by *Angiostrongylus cantonensis* on Taiwan. Am. J. Trop. Med. Hyg. 25, 233–249.

Yong, H.S., Song, S.L., Eamsobhana, P., Goh, S.Y., Lim, P.E., 2015. Complete mitochondrial genome reveals genetic diversity of *Angiostrongylus cantonensis* (Nematoda: Angiostrongylidae). Acta Trop. 152, 157–164.

Zingales, B., Miles, M.A., Campbell, D.A., Tibayrenc, M., Macedo, A.M., Teixeira, M.M., et al., 2012. The revised *Trypanosoma cruzi* subspecific nomenclature: rationale, epidemiological relevance and research applications. Infect., Genet. Evol. 12, 240–253.