

RESEARCH ARTICLE

Population structure analysis of the neglected parasite *Thelazia callipaeda* revealed high genetic diversity in Eastern Asia isolates

Xi Zhang¹, Ya Li Shi¹, Lu Lu Han¹, Chen Xiong¹, Shi Qi Yi¹, Peng Jiang¹, Zeng Xian Wang², Ji Long Shen², Jing Cui^{1*}, Zhong Quan Wang^{1*}

1 Department of Parasitology, School of Basic Medical Sciences, Zhengzhou University, Zhengzhou, China, **2** Department of Microbiology and Parasitology, Anhui Medical University, Anhui, China

* wangzq@zzu.edu.cn (ZQW); cuij@zzu.edu.cn (JC)



Abstract

Background

Thelazia callipaeda is the causative agent of thelaziasis in canids, felids and humans. However, the population genetic structure regarding this parasite remains unclear.

Methodology/principal findings

In this study, we first explored the genetic variation of 32 *T. callipaeda* clinical isolates using the following multi-molecular markers: *cox1*, *cytb*, 12S rDNA, ITS1 and 18S rDNA. The isolates were collected from 13 patients from 11 geographical locations in China. Next, the population structure of *T. callipaeda* from Europe and other Asian countries was analyzed using the *cox1* sequences collected during this study and from the GenBank database. In general, the Chinese clinical isolates of *T. callipaeda* expressed high genetic diversity. Based on the *cox1* gene, a total of 21 haplotypes were identified. One only circulated in European countries (Hap1), while the other 20 haplotypes were dispersed in Korea, Japan and China. There were five nucleotide positions in the *cox1* sequences that were confirmed as invariable among individuals from Europe and Asia, but the sequences were distinct between these two regions. Population differences between Europe and Asian countries were greater than those among China, Korea and Japan. The *T. callipaeda* populations from Europe and Asia should be divided into two separate sub-populations. These two groups started to diverge during the middle Pleistocene. Neutrality tests, mismatch distribution and Bayesian skyline plot (BSP) analysis all rejected possible population expansion of *T. callipaeda*.

Conclusions

The Asian population of *T. callipaeda* has a high level of genetic diversity, but further studies should be performed to explore the biology, ecology and epidemiology of *T. callipaeda*.

OPEN ACCESS

Citation: Zhang X, Shi YL, Han LL, Xiong C, Yi SQ, Jiang P, et al. (2018) Population structure analysis of the neglected parasite *Thelazia callipaeda* revealed high genetic diversity in Eastern Asia isolates. PLoS Negl Trop Dis 12(1): e0006165. <https://doi.org/10.1371/journal.pntd.0006165>

Editor: Francesca Tamarozzi, Negrar Hospital, ITALY

Received: September 18, 2017

Accepted: December 13, 2017

Published: January 11, 2018

Copyright: © 2018 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All sequences were deposited in the GenBank database under accession numbers MF795695 – MF795719 (*cytb*), MF795663 – MF795694 (*cox1*), MF795599 – MF795630 (12S rDNA), MF795720 – MF795751 (ITS1) and MF795631 – MF795662 (18S rDNA).

Funding: This work was supported by the National Key Research and Development Program of China (2017YFD0501300), the China Postdoctoral Science Foundation (2017M622367), Outstanding Young Talent Research Fund of Zhengzhou

University (1521328001) and the National Undergraduate Training Programs for Innovation (201710459076). The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Thelazia callipaeda is the causative agent of thelaziasis in canids, felids and humans. Despite the existing threat of thelaziasis in China, the genetic diversity of *T. callipaeda* has not been investigated across its wide geographical distribution in China, yet such information may provide insight into the disease epidemiology and the development of specific control measures. In this study, the genetic variation of 32 *T. callipaeda* clinical isolates collected from 13 patients from 11 geographical locations in China were explored using the following multi-molecular markers: *cox1*, *cytb*, 12S rDNA, ITS1 and 18S rDNA. In addition, the population structure of *T. callipaeda* from Europe and other Asian countries was analyzed using the *cox1* sequences collected during this study and from the GenBank database. In general, the Chinese clinical isolates of *T. callipaeda* demonstrated high genetic diversity. Based on the *cox1* gene, a total of 21 haplotypes were identified, one circulated in European countries (Hap1), while the other 20 haplotypes were dispersed in Korea, Japan and China. There were five nucleotide positions in the *cox1* that were confirmed as invariable among individuals from Europe and Asia, but the sequences were distinct between these two regions. Population differences between Europe and Asian countries were greater than those among China, Korea and Japan, such that the *T. callipaeda* population from Europe and Asia should be divided into two separate sub-populations. These two groups started to diverge during the middle Pleistocene.

Introduction

The spirurid nematode *Thelazia callipaeda* (Spirurida: Thelaziidae) is the major etiological agent of ocular thelaziasis [1]. This worm can parasitize the conjunctival sac of domestic and wild carnivores and humans, causing conjunctivitis, lacrimation and itchiness, and even blindness [2,3]. The parasite is transmitted by a drosophilid insect of the genus *Phortica* (Diptera, Drosophilidae), that feeds on the lachrymal secretions of mammals [4,5]. *T. callipaeda* was previously known as the oriental eyeworm because of its original description in eastern Asia countries [1]. Since the first cases of canine thelaziasis reported in Italy in 1988, the nematode has spread through many southern, central, western and eastern European countries [6]. A broad spectrum of wild carnivores, such as wolves, wildcats, red foxes, badgers, beech martens and even brown hares, plays an important role in maintaining and spreading eyeworm infections to domestic animals and humans [7]. The first human case was described in Beijing, China in 1917 [8]. Although sporadic human thelaziasis has been reported in several European countries, human infections are mainly documented in people living in China, Japan, Korea and India [9]. China is most likely to have the largest number of cases of thelaziasis in the world with more than 600 cases reported to date [10]. Recently, an increase in *T. callipaeda* infections has been reported in animals and humans living in European countries and in China [6,11].

Consequently, *T. callipaeda* poses a serious threat to public health and thelaziasis has even been termed as an emerging zoonotic disease [1,6,9]. Understanding the host specificity, transmission pattern and population genetic characteristics of the parasite are valuable for the prevention and control of thelaziasis in animals and humans [12]. However, our knowledge regarding these issues is still fragmented, and insufficient studies on population genetics of *T. callipaeda* have been carried out. One possible reason is that as a neglected pathogen, *T. callipaeda* does not draw enough attention from most parasitologists. This may also be attributed to the difficulty in collecting *T. callipaeda* isolates from different hosts and distinct geographical locations.

More than 10 years ago, Otranto and Traversa [13] performed the first genetic variance analysis among *Thelazia* species by using the first internal transcribed spacer (ITS1) ribosomal DNA sequence, and concluded that the ITS1 sequence was a useful genetic marker for the molecular identification of *Thelazia* spp. In 2005, Otranto et al. [12] investigated the genetic variability among 50 individual adult specimens of *T. callipaeda* from Europe and Asia based on the mitochondrial cytochrome *c* oxidase subunit 1 gene (*cox1*). Recently, although various publications of the *T. callipaeda* infections in animals and humans have been reported [3,14–18], no subsequent related studies about its population genetics have been reported. In this study, we explored the genetic variability within human *T. callipaeda* isolates collected from different geographical locations in China by using three mitochondrial genes and two nuclear ribosomal DNA sequences as follows: mitochondrial cytochrome *b* (*cytb*), the small subunit of ribosomal DNA gene (12S rDNA) and *cox1*; ITS1 and the small subunit of nuclear ribosomal DNA (18S rDNA). These molecular markers were selected because they are suitable for inferring population differences and conducting phylogenetic analysis at different taxonomic levels [12,13,19–21]. Using *cox1*, we also performed a genetic variability comparative analysis on clinical *T. callipaeda* isolates collected in China and from previous publications. Additionally, the presumed transmission pattern of *T. callipaeda* investigated here relied on phylogeny and molecular dating methods.

Materials and methods

Ethics statement

This study was approved by the Life Science Ethics Committee of Zhengzhou University (No. 2017–0006). The protocol and written informed consent form were approved by the Human Ethics Committees of the Zhengzhou University. All subjects older than eighteen years old provided written informed consent; in the case of children, they provided written informed assent, and their parents/guardians provided written consent for them. All worms were collected from patients to treat their thelaziasis and not expressly for the purpose of the present study.

Parasite collection and identification

A total of 32 worms were harvested from 13 patients from 11 distinct geographical locations in China from September 2007 to July 2016 (Table 1). All nematodes were removed from the eyes of patients with intraocular forceps, while the patients were anesthetized with oxybuprocaine. The collected nematodes were transferred to Petri dishes containing physiological saline (0.9% NaCl). These eyeworms were identified as *T. callipaeda* according to the morphological characteristics (e.g., shape of the buccal capsule, presence of transversally striated cuticle and cloacal papillae, morphology of the spicules in males and the position of the vulva in females) described in Otranto et al. [22].

Sequencing of target genes and alignment

Total genomic DNA was extracted from individual specimens using the EasyPure Genomic DNA Kit (Transgen, China) following the manufacturer's protocol. Five molecular markers, viz. *cytb*, *cox1*, 12S rDNA, ITS1 and 18S rDNA, were amplified to explore the genetic diversity of *T. callipaeda*. For 12S, ITS1 and *cox1*, the amplifications were obtained using primer combinations described in Casiraghi et al. (2004) [23], Otranto and Traversa (2004) [13], and Otranto et al. (2005) [12], respectively. The forward and reverse primers used for amplifying the *cytb* and 18S markers were designed as follows: cobF 5'-TGATTGGTGGTTTTGGTAA-3';

Table 1. Geographical and host origins of *Thelazia callipaeda* specimens studied herein.

Geographical origin	Host origin	SS ^a	Haplotypes ^b	References
Italy (Piemonte)	<i>Canis familiaris</i>	5	h1	Otranto et al. (2005)
Italy (Piemonte)	<i>Vulpes vulpes</i>	2	h1	Otranto et al. (2005)
Italy (Basilicata)	<i>Felis catus</i>	5	h1	Otranto et al. (2005)
Italy (Basilicata)	<i>C. familiaris</i>	11	h1	Otranto et al. (2005)
Italy (Basilicata)	<i>V. vulpes</i>	2	h1	Otranto et al. (2005)
Italy (Calabria)	<i>C. familiaris</i>	5	h1	Otranto et al. (2005)
Germany	<i>C. familiaris</i>	6	h1	Otranto et al. (2005)
The Netherlands	<i>C. familiaris</i>	1	h1	Otranto et al. (2005)
Portugal (Vinhais)	<i>Oryctolagus cuniculus</i>	4	h1	Gama et al. (2016)
Serbia (Baljevac)	<i>C. familiaris</i>	1	h1	Gajic et al. (2014) ^c
Serbia (Baljevac)	<i>F. catus</i>	1	h1	Gajic et al. (2014) ^c
Romania (Oradea, Bihor)	<i>C. familiaris</i>	3	h1	Mihalca et al. (2015)
Romania	<i>C. familiaris</i>	2	h1	Ionita et al. (2015) ^c
Slovakia	<i>C. familiaris</i>	1	h1	Cabanova (2017) ^c
Hungary	<i>F. catus</i>	1	h1	Takacs and Farkas (2016) ^c
China (Anhui)	<i>C. familiaris</i>	5	h2 (1), h3 (1), h4 (2), h5 (1)	Otranto et al. (2005)
Korea (Kou Kuk)	<i>C. familiaris</i>	8	h5 (1), h6 (1), h7 (2), h8 (4)	Otranto et al. (2005)
Japan (Okayama)	<i>Homo sapiens</i>	9	h9	Kumase et al (2009) ^c
Japan (Saitama)	<i>C. familiaris</i>	1	h10	Yoshikawa (2013) ^c
Japan (Tokyo)	<i>C. familiaris</i>	7	h10 (5), h11 (1), h12 (1)	Yoshikawa (2013) ^c
China (Anhui, Hefei)	<i>H. sapiens</i>	1	h13	this study
China (Anhui, Liuan)	<i>H. sapiens</i>	6	h13	this study
China (Liaoning, Dandong)	<i>H. sapiens</i>	2	h14	this study
China (Shaanxi, Shangluo)	<i>H. sapiens</i>	3	h15	this study
China (Shaanxi, Tongchuan)	<i>H. sapiens</i>	3	h16	this study
China (Hubei, Huanggang)	<i>H. sapiens</i>	2	h17	this study
China (Hubei, Wuhan)	<i>H. sapiens</i>	1	h18	this study
China (Henan, Pingdingshan)	<i>H. sapiens</i>	8	h15 (4), h19 (4)	this study
China (Henan, Jiaozuo)	<i>H. sapiens</i>	2	h20	this study
China (Henan, Luoyang)	<i>H. sapiens</i>	2	h15	this study
China (Henan, Zhengzhou)	<i>H. sapiens</i>	2	h15 (1), h21 (1)	this study

a, sampling size

b, haplotypes identified based on the *cox1*

c, direct submitted sequences.

<https://doi.org/10.1371/journal.pntd.0006165.t001>

and *cobR* 5'-ATAAGTACGAGTATCAATATC-3'; and 18SF 5'-CTCATAAAATAATTGG TGAATCTGAATAGC-3' and 18SR 5'-ATAACTTTTCAGCAATGGTTACAG-3', respectively. PCR products were purified using the EasyPure PCR Purification Kit (Transgen, China) and sequenced in both directions at the Genewiz Company (Beijing, China). All sequences were deposited in the GenBank database (S1 Table). The sequenced genes were initially aligned using the default settings in the program Clustal X v.2.0 [24] and adjusted in MEGA v.6.06 [25] according to the corresponding amino acid sequences of protein-coding genes and secondary structure of ribosomal DNA sequences. The nucleotide composition, conserved sites, variable sites, parsimony-informative sites, and singleton sites were estimated using MEGA v.6.06.

Genetic diversity analysis of *T. callipaeda* isolates from China

The program DnaSP v5.10 [26] was employed to analyze the number of haplotypes, haplotype diversity (Hd), and nucleotide diversity (Pi) of each molecular marker. Network v5.0 [27] was used to draw a median-joining network to analyze the relationships among the detected haplotypes. Analysis of molecular variance (AMOVA) was computed in Arlequin v.3.5 [28] with non-parametric permutations of 1,000 times ($p = 0.05$) to detect the partitions of genetic diversity within and among populations. Pairwise F_{ST} values between populations were performed for all datasets in Arlequin to explore levels of genetic differentiation among the populations. The significance of F_{ST} values evaluated was based on 1000 random permutations. Demographic changes were also estimated using mismatch distributions in Arlequin with 1000 simulations, under a scenario of no recombination. The validity of the expansion model was tested by using the sum of squared deviations (SSD) and raggedness index (RI) between observed and expected mismatches. The neutrality tests using Tajima's D [29] and Fu's F_s [30] were also applied through Arlequin as an assessment of possible population expansion.

Genetic variability comparative analysis of *T. callipaeda* isolates using *cox1*

To make a worldwide genetic variability comparative analysis of *T. callipaeda*, all available sequences of *cox1* in the GenBank database were included (S1 Table). In addition, we performed a Bayesian skyline plot analysis (BSP) implemented in BEAST v1.8.2 [31] to estimate the change in population size over time, and the time to the most recent common ancestor (tMRCA) for each *T. callipaeda* haplotype. A piecewise-constant skyline model was selected. The molecular evolutionary rate of *cox1* was fixed at 0.01 substitutions per site per million years ago (Mya) according to the substitution rate for nematode mtDNA [32]. Tracer v1.5 [33] was used to reconstruct the demographic history over time.

Phylogenetic diversity analysis

The phylogenetic pattern of all *cox1* haplotypes was estimated through maximum parsimony (MP) and Bayesian inference (BI). MP analysis was performed in MEGA v.6.06. Confidence in each node was assessed by boot-strapping (1000 pseudo-replicates). Bayesian inference was performed in MrBayes v.3.2 [34], after determining the appropriate substitution model by applying the Akaike information criterion (AIC) in jModelTest 2 [35]. The analysis consisted of two runs, each with four MCMC chains running for 5,000,000 generations, and sampling every 100th generation. Stationarity was assessed using a convergence diagnostic. An average standard deviation of the split frequencies (ASDSF) < 0.01 was used as criteria for convergence between both runs. The consensus tree was drawn after removing the first 10,000 trees (20%) as the burn-in phase. The approximate divergence time was estimated using an uncorrelated log-normal relaxed molecular-clock model in the BEAST v1.8.2 program. The substitution model was assigned following model selection by jModelTest 2. For the earlier tree, a basic coalescent model was chosen, assuming a constant population size over the given time period considered. Two replicate MCMC runs were performed, with the tree and parameter values sampled every 1,000 steps over a total of 1×10^8 steps.

Results

Genetic diversity of Chinese *T. callipaeda* isolates

The sequence alignments for *cytb*, *cox1*, 12S, ITS1 and 18S were 1035 bp, 660 bp, 453 bp, 772 bp and 1201 bp, respectively. In 32 isolates from eleven localities, the genetic markers of *cox1*, 12S, ITS1 and 18S were used to identify 9, 8, 8 and 7 new haplotypes (Haps). For *cytb*, 8 Haps

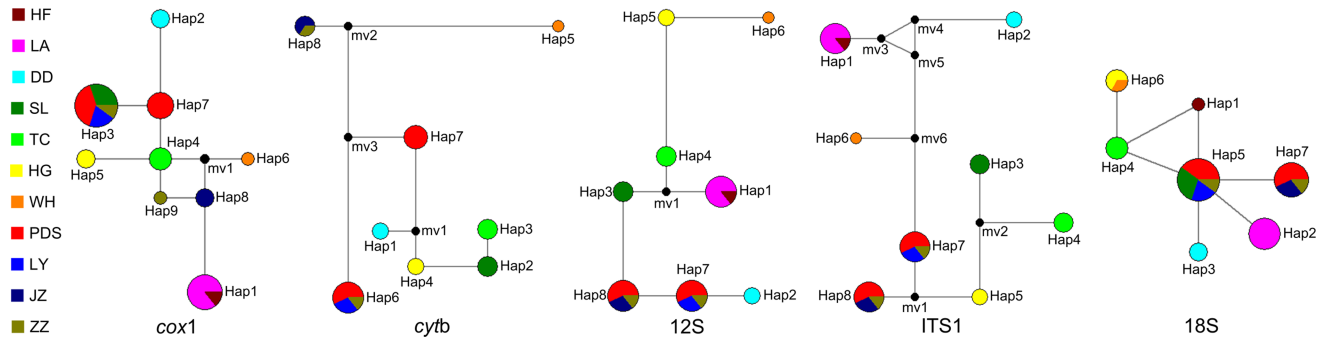


Fig 1. *Thelazia callipaeda* haplotype networks by using *cox1*, *cytb*, 12S rDNA, ITS1 and 18S rDNA. Each haplotype is represented by a circle, with the area of the circle proportional to its frequency. The median vector is indicated by a solid black circle. Geographic regions are designated as follows for China: Hefei (HF), Liuan (LA), Dandong (DD), Shangluo (SL), Tongchuan (TC), Huanggang (HG), Wuhan (WH), Pingdingshan (PDS), Luoyang (LY), Jiaozuo (JZ) and Zhengzhou (ZZ).

<https://doi.org/10.1371/journal.pntd.0006165.g001>

were identified in only 25 isolates because the sequences of isolates from Hefei (HF) and Liuan (LA) of Anhui province were not amplified (S2 Table). With the exception of isolates from Pingdingshan (PDS) and Zhengzhou (ZZ) of Henan province, which shared two haplotypes, each geographical population shared only a single haplotype. Analyses of the median-joining networks (Fig 1) showed that samples from HF and LA shared a haplotype (Hap1) when using *cox1*, 12S and ITS1, but under the analysis with 18S, these samples identified two Haps (Hap1 and Hap2). Clinical eyeworms from Dandong (DD) and Tongchuan (TC) shared geographical specific Haps using each of the selected markers. Isolates from Huanggang (HG) and Wuhan (WH) also revealed specific Haps when using *cytb*, *cox1*, 12S and ITS1; however, these samples shared a single haplotype using 18S. Using *cytb*, 12S and ITS1, a single geographical specific haplotype was identified in *T. callipaeda* from Shangluo (SL). Interestingly, using *cox1* (Hap3) and 18S (Hap5), these isolates from SL shared the same haplotype with all samples from Luoyang (LY), 4 samples from Pingdingshan (PDS) and one from Zhengzhou (ZZ). The remaining isolates from PDS and ZZ shared the same haplotype using 12S, ITS1 and 18S. However, when using *cytb* and *cox1*, the remaining isolates possessed two distinct Haps. Using *cytb*, 12S, ITS1 and 18S, worms from Jiaozuo (JZ) shared the same haplotype with the second sample from ZZ. Only with *cox1*, did the JZ isolates identify as a distinct haplotype. The results from the analysis of molecular variance (AMOVA) showed that much more genetic variance lay among the populations than within the populations for all datasets, more specifically, *cytb*: 69.77% vs. 30.23%; *cox1*: 91.32% vs. 8.68%; 12S: 94.26% vs. 5.74%; ITS1: 91.69% vs. 8.31%; 18S: 83.71% vs. 16.29% (Table 2). The pairwise fixation index (F_{ST}) values between specified

Table 2. Analysis of molecular variance (AMOVA) of the populations of *Thelazia callipaeda* from China.

Markers	Source of variation							
	Among populations				Within populations			
	d.f	Sum of squares	Variance components	Percentage of variation	d.f	Sum of squares	Variance components	Percentage of variation
<i>cytb</i>	8	105.360	4.32759	69.77%	16	30.000	1.87500	30.23%
<i>cox1</i>	10	57.250	2.00345	91.32%	21	4.000	0.19048	8.68%
12S	10	55.156	1.95351	94.26%	21	2.500	0.11905	5.74%
ITS1	10	112.406	3.93972	91.69%	21	7.500	0.35714	8.31%
18S	10	18.094	0.61188	83.71%	21	2.500	0.11905	16.29%

d.f = degrees of freedom.

<https://doi.org/10.1371/journal.pntd.0006165.t002>

Table 3. Mismatch and neutrality test results of the *Thelazia callipaeda* population from China.

Markers	Neutrality Tests		Mismatch	
	Fu's F_S	Tajima's D	SSD	RI
cytb	5.59212 (0.974)	0.95309 (0.864)	0.02703 (0.432)	0.05787 (0.403)
cox1	0.44453 (0.596)	1.04997 (0.891)	0.02243 (0.409)	0.07200 (0.433)
12S	0.97235 (0.707)	0.79949 (0.815)	0.02074 (0.459)	0.03965 (0.716)
ITS1	4.53458 (0.934)	1.91959 (0.974)	0.04715 (0.021)	0.11855 (0.044)
18S	-1.65282 (0.164)	0.18826 (0.618)	0.01384 (0.056)	0.12259 (0.068)

SSD = Sum of squared deviation, RI = Raggedness index. Number in parentheses is the P value.

<https://doi.org/10.1371/journal.pntd.0006165.t003>

geographical regions were estimated for all molecular markers used to measure the population differentiation (S3 Table). With the exception of F_{ST} values between PDS and the remaining populations, and ZZ and the remaining populations, most of the F_{ST} values reached 1.00. No estimated pairwise F_{ST} values were statistically significant besides those between LA and the remaining populations, and PDS and the remaining populations. Based on all genes, the neutrality tests of Tajima's D and Fu's F_S for the total population showed non-significant positive values, except for the negative value of Fu's F_S (-1.65282, $p = 0.164$), using the 18S gene (Table 3). Using all markers, mismatch distribution analyses revealed multi-modal frequency distributions for the total population, rejecting possible population expansion (S1 Fig). In addition, low values were found for the sum of squared deviation (SSD) and raggedness index (RI) under the demographic expansion model (Table 3).

Compare the genetic variability between Chinese isolates and isolates from European and other Eastern Asia countries

A total 21 haplotypes were identified within *T. callipaeda* isolates from 25 localities in 12 countries (Table 1). Hap1 was only found in *T. callipaeda* from European countries (Italy, Germany, The Netherlands, Portugal, Serbia, Romania, Slovakia, Hungary, etc). The remaining haplotypes were shared by samples from Eastern Asia countries (Korea, Japan and China). Haps2–8 and haps10–12 were shared by samples isolated from domestic dogs (*Canis familiaris*). Hap9 and haps13–21 were identified in eyeworms collected from humans (*Homo sapiens*). The alignment of 21 haplotypes revealed nucleotide variations (36 transitions and 1 transversion) at 37 alignment positions (S2 Fig). The majority of the nucleotide variability was at the third codon position ($n = 31$; 83.8%), whereas the remainder was at the first codon position ($n = 6$; 16.2%). Of the 37 variable positions, only two sites generated nonsynonymous mutations. One was at alignment position 189 (A—G) of Hap4, which changed the amino acid from Methionine to Valine. The other, also located in Hap4 (position 252: C—T), changed Leucine to Phenylalanine. Five nucleotides (i.e., G—A at alignment positions 89, 149, 206 and 257; and C—T at position 539) were invariable among all individuals from Europe and among all individuals from Eastern Asia (Korea, Japan and China), but they were different between Europe and Asia. The pairwise comparisons among 21 haplotypes ranged from 0.15 to 2.86% (Table 4). Within each country studied, the intraspecific divergences in 4 haplotypes from Korea, 4 haplotypes from Japan and 13 haplotypes from China were 0.31–1.09, 0.15–2.55 and 0.15–1.73%, respectively. The AMOVA results showed that much more genetic variance lay among the populations (74.79%) than within the populations (25.21%). Between geographical regions, the population differences (F_{ST} value) were 0.961 for Europe vs. Korea, 0.866 for Europe vs. Japan, 0.827 for Europe vs. China, 0.444 for Korea vs. Japan, 0.275 for Korea vs. China, and 0.430 for Japan vs. China (S4 Table).

Table 4. Pairwise comparison of sequence differences (%) among the 21 *cox1* haplotypes, representing *Thelazia callipaeda* from different geographical origins.

Hap	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10	h11	h12	h13	h14	h15	h16	h17	h18	h19	h20	h21	
h1	–																					
h2	1.73	–																				
h3	1.57	0.47	–																			
h4	1.25	1.09	0.94	–																		
h5	2.06	0.94	0.78	1.42	–																	
h6	1.89	0.78	0.62	1.25	1.10	–																
h7	1.57	0.47	0.31	0.94	0.78	0.62	–															
h8	1.89	0.78	0.62	1.25	1.10	0.31	0.62	–														
h9	2.54	1.73	1.57	1.89	2.06	1.89	1.57	1.89	–													
h10	1.90	1.10	0.94	1.57	0.47	1.26	0.94	1.26	2.22	–												
h11	2.70	1.89	1.73	2.05	2.22	2.05	1.73	2.05	0.15	2.38	–											
h12	2.86	2.05	1.89	2.21	2.38	2.21	1.89	2.21	0.31	2.54	0.15	–										
h13	1.41	0.62	0.78	0.78	1.25	1.09	0.78	1.09	1.73	1.41	1.89	2.05	–									
h14	2.38	1.26	1.10	1.73	0.94	0.78	1.10	0.78	2.38	0.78	2.54	2.71	1.57	–								
h15	2.06	0.94	0.78	1.42	0.62	1.10	0.78	1.10	2.06	0.47	2.22	2.38	1.25	0.62	–							
h16	1.57	0.47	0.31	0.94	0.47	0.62	0.31	0.62	1.57	0.62	1.73	1.89	0.78	0.78	0.47	–						
h17	1.57	0.78	0.62	0.94	0.78	0.94	0.62	0.94	1.89	0.94	2.06	2.22	0.78	1.09	0.47	0.31	–					
h18	1.57	0.47	0.31	0.94	0.78	0.62	0.31	0.62	1.57	0.94	1.41	1.57	0.78	1.10	0.78	0.31	0.62	–				
h19	1.89	0.78	0.62	1.26	0.47	0.94	0.62	0.94	1.89	0.31	2.06	2.22	1.09	0.47	0.15	0.31	0.62	0.62	–			
h20	1.57	0.47	0.31	0.94	0.78	0.62	0.31	0.62	1.57	0.94	1.73	1.89	0.47	1.10	0.78	0.31	0.62	0.31	0.62	–		
h21	1.73	0.62	0.47	1.10	0.62	0.78	0.47	0.78	1.73	0.78	1.89	2.05	0.62	0.94	0.62	0.15	0.47	0.47	0.47	0.15	–	

<https://doi.org/10.1371/journal.pntd.0006165.t004>

Phylogenetic diversity and demographic estimation

The likelihood models identified by the jModelTest (AIC) suggested that the HKY+G model was most suitable for *cox1* haplotypes. Both maximum parsimony and Bayesian analyses generated consistent tree topologies (Fig 2A and S3 Fig). Among the tested haplotypes, Hap1 and Hap4 composed a single clade (clade I), and the remaining haplotypes made up another clade (clade II). Within clade II, the earliest diversifications gave rise to Hap13, then to Hap9, Hap11 and Hap12 (Japanese haplotypes without Hap10). The next diversification event separated the remaining haplotypes. The molecular dating analysis suggested that the two clades began to diverge during the middle Pleistocene (Fig 2A). The time of origin of clade I is estimated to be approximately 0.58 Mya (late Pleistocene) with a 95% highest posterior density (HPD) of 0.23–1.01 Mya. Clade II started to develop in the middle Pleistocene (0.78 Mya) with a 95% HPD of 0.47–1.17 Mya. The early branching of the Japanese haplotypes (without Hap10) started in the late Pleistocene (0.14 Mya, with a 95% HPD of 0.02–0.32 Mya). Neutrality tests of Tajima’s *D* and Fu’s *F_S* for the total population showed non-significant positive values, thereby rejecting possible population expansion (S5 Table). Mismatch distribution analyses revealed multi-modal frequency distributions (Fig 2B). In addition, low values for the sum of squared deviation and raggedness index under the demographic expansion model were found (S5 Table). The result of Bayesian skyline plot analysis of the *T. callipaeda* population revealed a gradual expansion trend, but also rejected sudden population expansion (Fig 2C).

Discussion

Although more than 10 species of the spirurid genus *Thelazia* can cause veterinary or medical problems in many parts of the world [13], only *T. callipaeda* affects humans in China, causing mild to severe clinical symptom [36]. It is still unknown whether there are other human

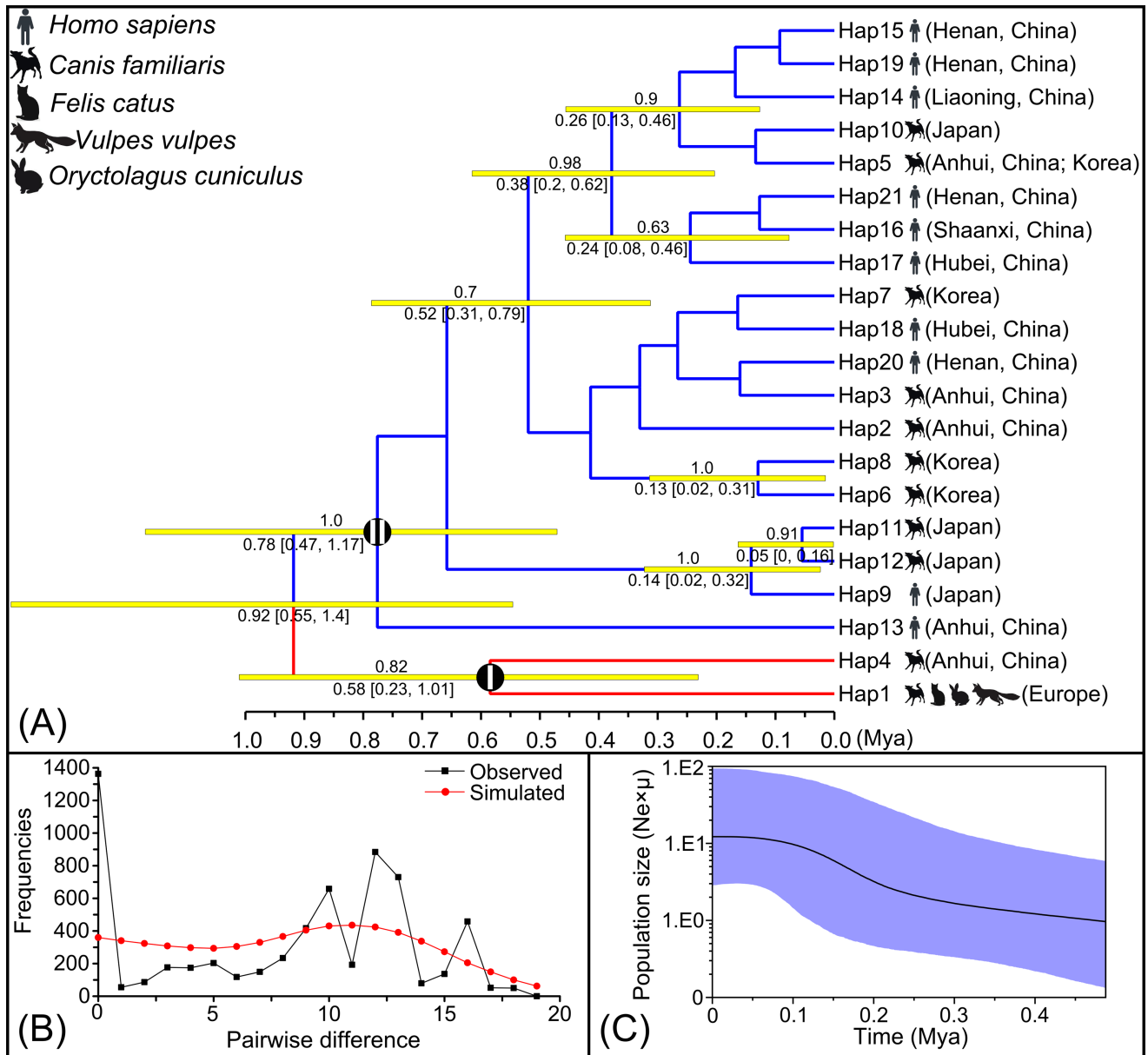


Fig 2. (A) Phylogram of *Thelazia callipaeda* isolates from Europe and Asia with divergence time estimation. Yellow bars at each node show the 95% highest posterior density interval for the main nodes. Numbers above branches represent the Bayesian posterior probabilities. Only posterior probabilities above 0.6 are shown. Numbers below branches indicate the estimated age and 95% confidence intervals (shown in square brackets). (B) Mismatch distribution analysis for the *T. callipaeda* population. The line charts represent the observed frequencies of pairwise differences among haplotypes. (C) Bayesian skyline plot calculates for the *T. callipaeda* population. The X-axis is in units of million years ago (Mya) and the Y-axis is $N \times \mu$ (effective population size \times mutation rate per site per generation). The median estimates are shown as thick solid lines, and the 95% HPD limits are shown by the colored areas.

<https://doi.org/10.1371/journal.pntd.0006165.g002>

infective species or genotypes of *T. callipaeda* in China. In this study, we collected 32 eye worms from 13 patients from 11 distinct geographical locations in China over a period of 10 years, and performed a genetic variation analysis of these isolates using a multi-gene approach to investigate the population structure of *T. callipaeda*.

Haplotypes of the Chinese *T. callipaeda* population identified by each molecular marker were generally consistent. Of all the geographic populations, only PDS and ZZ isolates revealed

two different haplotypes. Worms from ZZ were harvested in two different patients. However, interestingly, all PDS isolates that were collected from a single child also had two haplotypes, indicating the patient probably experienced two or more infections of drosophilid flies. Each geographic population and the total population demonstrated high haplotype diversity (Hd) values (nearly 1.0) by all markers, however, the nucleotide diversity (Pi) values of each gene were below 0.01, indicating that multiple haplotypes were differentiated by few nucleotide mutations. Correspondingly, the AMOVA results also showed that much more genetic variance exists among the populations than within the populations [19]. The pairwise fixation index (F_{ST}) was used to measure the population differentiation. Most of the F_{ST} values between specified geographical regions were far more than 0.25, indicating very high genetic differentiation and a long-term interruption of gene flow among the geographical populations [37].

Based on the partial *cox1* sequence, Otranto et al. [12] performed the first genetic variability investigation of the *T. callipaeda* population in 2005. Since then, sequences of *T. callipaeda* from different geographical locations and different hosts have been published [3, 16–18]. In this study, based on previous research by Otranto et al. [12], we added new data from Chinese clinical isolates and other published data to generate a comprehensive genetic variation analysis of the *T. callipaeda* population. Some interesting discoveries were made: (1) Thirteen novel genotypes of *T. callipaeda* were identified; four from Japan (Haps9–12), and another nine from Chinese clinical isolates (Haps13–21); within the 4 Japanese genotypes, one was from a human, and the other 3 were from *Canis familiaris*. (2) In *cox1*, Otranto et al. [12] found six nucleotide positions (positions 89, 149, 206, 257, 539 and 608) that were invariable among individuals from Europe and Asia, but they were distinct between Europe and Asia. We confirmed five of them (positions 89, 149, 206, 257, 539 and 608) using an enlarged dataset in this study. (3) Population differences between Europe and Asian countries were larger than those among China, Korea and Japan. Within Asian countries, the genetic difference between China and Korea was the smallest, while the genetic difference between Korea and Japan was the largest. Only within human isolates were the genetic differences among Chinese samples smaller than those between China and Japan. This phenomenon may be attributed to the geographical distance's influence on gene flow among species [38], and possibly to the geographical origin, which needs to be analyzed further in the future. (4) The *T. callipaeda* population from Europe and Asia should be divided into the following two subgroups: one group (clade I) comprising worms from European countries (Hap1) and a sample isolated from a dog in Anhui of China (Hap4); the other group (clade II) comprising *T. callipaeda* from Korea, Japan and China. Individuals in clade II diverged earlier than those in clade I. Within the Asian haplotypes, Hap13 (clinical isolates of Anhui) was the ancestral haplotype; then, it was transmitted to Japan, Korea and other regions of China. (5) All of the neutrality tests, mismatch distribution and BSP analyses rejected possible population expansion of *T. callipaeda*, indicating this nematode population was in the stable phase [37].

Generally, the population structures of parasites are associated with large genetic differentiation among populations from different geographical regions and/or hosts and low intra-population genetic variability [39,40]. In this study, there was a relatively large genetic differentiation (above 0.82) between the European and Asian populations; however, the genetic variability among Asian countries was small (below 0.44). In addition, phylogenetic analyses supported the conclusion that isolates from Europe and Asia (excluding the Hap4) are separate populations. However, the genetic differentiation of *T. callipaeda* from distinct hosts (dogs, foxes, cats and human) was inconspicuous, suggesting a lack of connection between eyeworms and host species [12]. That being said, an important factor to consider is the abundance of genetic diversity in the Asian isolates. A possible explanation for this genetic diversity is that *T. callipaeda* populations are tightly linked to the intermediate host (vector). The drosophilid

fly *Phortica variegata* was the vector of *T. callipaeda* in Europe, whereas the *T. callipaeda* in Anhui of China was transmitted by *P. okadai* [9,12]. Although it is still unknown whether other vector species exist in Asian countries, considering the high genetic diversity of *T. callipaeda* in Korea, Japan and China, it is very likely that some novel species or genotypes of *Phortica* flies exist in these areas. Hence, further studies should be launched to explore the biology, population genetics, ecology and epidemiology of *T. callipaeda* in the future.

In summary, the Chinese clinical isolates of *T. callipaeda* expressed a high level of genetic diversity. Using the *cox1* gene, thirteen new genotypes were identified, four from Japan (Haps9–12), and nine from China (Haps13–21). There were five positions in the *cox1* sequences (89, 149, 206, 257, 539 and 608) that were confirmed invariable among individuals from Europe and Asia, but the sequences were distinct between these two regions. Population differences between Europe and Asian countries were greater than those among China, Korea and Japan, and the *T. callipaeda* populations from Europe and Asia should be divided into two separate sub-populations. These two groups started to diverge during the middle Pleistocene.

Supporting information

S1 Table. *Thelazia callipaeda* sampling and data summary for this study.

(DOC)

S2 Table. Genetic diversity in the *cytb*, *cox1*, 12S, ITS1, and 18S sequences and the concatenated sequences in the populations of *Thelazia callipaeda* from China.

(DOC)

S3 Table. Estimated pairwise F_{ST} values of sequences between *Thelazia callipaeda* populations.

(DOC)

S4 Table. Estimated pairwise F_{ST} values of *cox1* sequences between *Thelazia callipaeda* populations from Europe and Asia.

(DOC)

S5 Table. Mismatch and neutrality tests results of *Thelazia callipaeda* populations from Europe and Asia.

(DOC)

S1 Fig. Mismatch distribution analyses of Chinese clinical isolates for *cytb*, *cox1*, 12S, ITS1 and 18S.

(DOC)

S2 Fig. Alignment of the twenty one *cox1* haplotypes (Hap1–Hap21) representing *Thelazia callipaeda* from Europe and Asia.

(DOC)

S3 Fig. Maximum parsimony (MP) and Bayesian phylogenetic trees of *Thelazia callipaeda* from Europe and Asia based on the analysis of the *cox1* gene.

(DOC)

Acknowledgments

We are grateful to a large number of friends and colleagues who assisted with the collection of field samples, in particular to Prof. Ya Zhao, Jiahui Lei, Yanbin Cheng, and Jijia Shen.

Author Contributions

Conceptualization: Jing Cui.

Data curation: Xi Zhang, Peng Jiang.

Funding acquisition: Xi Zhang, Chen Xiong, Shi Qi Yi.

Methodology: Xi Zhang, Ya Li Shi, Lu Lu Han, Jing Cui, Zhong Quan Wang.

Resources: Chen Xiong, Shi Qi Yi, Zeng Xian Wang, Ji Long Shen, Jing Cui, Zhong Quan Wang.

Software: Xi Zhang, Ya Li Shi, Lu Lu Han.

Writing – original draft: Xi Zhang, Jing Cui, Zhong Quan Wang.

References

- Otranto D, Eberhard ML (2011) Zoonotic helminths affecting the human eye. *Parasit Vectors* 4: 41. <https://doi.org/10.1186/1756-3305-4-41> PMID: 21429191
- Mihalca AD, Ionică AM, D'Amico G, Daskalaki AA, Deak G, et al. (2016) *Thelazia callipaeda* in wild carnivores from Romania: new host and geographical records. *Parasit Vectors* 9: 350. <https://doi.org/10.1186/s13071-016-1628-9> PMID: 27316971
- Otranto D, Lia RP, Buono V, Traversa D, Giangaspero A (2004) Biology of *Thelazia callipaeda* (Spirurida, Thelaziidae) eyeworms in naturally infected definitive hosts. *Parasitology* 129 (Pt 5): 627–633. PMID: 15552407
- Otranto D, Brianti E, Cantacessi C, Lia RP, Máca J (2006) The zoophilic fruitfly *Phortica variegata*: morphology, ecology and biological niche. *Med Vet Entomol* 20: 358–364. <https://doi.org/10.1111/j.1365-2915.2006.00643.x> PMID: 17199746
- Otranto D, Lia RP, Cantacessi C, Testini G, Troccoli A, et al. (2005) Nematode biology and larval development of *Thelazia callipaeda* (Spirurida, Thelaziidae) in the drosophilid intermediate host in Europe and China. *Parasitology* 131 (Pt 6): 847–855. <https://doi.org/10.1017/S0031182005008395> PMID: 16336738
- Colella V, Kirkova Z, Fok É, Mihalca AD, Tasić-Otašević S, et al. (2016) Increase in Eyeworm Infections in Eastern Europe. *Emerg Infect Dis* 22: 1513–1515. <https://doi.org/10.3201/eid2208.160792> PMID: 27434135
- Otranto D, Dantas-Torres F, Mallia E, DiGeronimo PM, Brianti E, et al. (2009) *Thelazia callipaeda* (Spirurida, Thelaziidae) in wild animals: report of new host species and ecological implications. *Vet Parasitol* 166: 262–267. <https://doi.org/10.1016/j.vetpar.2009.08.027> PMID: 19782474
- Stuckey E (1917) Circumocular filariasis. *British J Ophthalmol* 1: 542.
- Papadopoulos E, Komnenou A, Thomas A, Ioannidou E, Colella V, et al. (2017) Spreading of *Thelazia callipaeda* in Greece. *Transbound Emerg Dis* <https://doi.org/10.1111/tbed.12626> PMID: 28239956
- Wang XL, Guo JY, Wang XL, Ma XL, Wang Y, An CL (2014) Two cases of human thelaziasis as confirmed by mitochondrial *cox1* sequencing in China. *Pathog Glob Health* 108: 298–301. <https://doi.org/10.1179/2047773214Y.0000000153> PMID: 25253041
- Zhang X, Shi YL, Wang ZQ, Duan JY, Jiang P, Liu RD, Cui J (2017) Morphological and mitochondrial genomic characterization of eyeworms (*Thelazia callipaeda*) from clinical cases in central China. *Front Microbiol* 8: 1335. <https://doi.org/10.3389/fmicb.2017.01335> PMID: 28751889
- Otranto D, Testini G, De Luca F, Hu M, Shamsi S, Gasser RB (2005) Analysis of genetic variability within *Thelazia callipaeda* (Nematoda: Thelazioidea) from Europe and Asia by sequencing and mutation scanning of the mitochondrial cytochrome *c* oxidase subunit 1 gene. *Mol Cell Probes* 19: 306–313. <https://doi.org/10.1016/j.mcp.2005.05.001> PMID: 16084062
- Otranto D, Traversa D (2004) Molecular characterization of the first internal transcribed spacer of ribosomal DNA of the most common species of eyeworms (Thelazioidea: *Thelazia*). *J Parasitol* 90: 185–188. <https://doi.org/10.1645/GE-115R1> PMID: 15040692
- Blouin MS (2002) Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer. *Int J Parasitol* 32: 527–531. PMID: 11943225
- Lin H, Liang X (2012) Eye can see a nest of worms! *Lancet* 379: e42. [https://doi.org/10.1016/S0140-6736\(11\)61455-4](https://doi.org/10.1016/S0140-6736(11)61455-4) PMID: 22244656

16. Hodžić A, Latrofa MS, Annoscia G, Alić A, Beck R, et al. (2014) The spread of zoonotic *Thelazia callipaeda* in the Balkan area. *Parasit Vectors* 7: 352. <https://doi.org/10.1186/1756-3305-7-352> PMID: 25078286
17. Mihalca AD, D'Amico G, Scurtu I, Chirilă R, Matei IA, et al. 2015. Further spreading of canine oriental eyeworm in Europe: first report of *Thelazia callipaeda* in Romania. *Parasit Vectors* 8: 48. <https://doi.org/10.1186/s13071-015-0663-2> PMID: 25623605
18. Gama A, Pires I, Canado M, Coutinho T, Lopes AP, et al. (2016) First report of *Thelazia callipaeda* infection in wild European rabbits (*Oryctolagus cuniculus*) in Portugal. *Parasit Vectors* 9: 236. <https://doi.org/10.1186/s13071-016-1526-1> PMID: 27160400
19. Zhang X, Wang H, Cui J, Jiang P, Fu GM, et al. (2015) Characterisation of the relationship between *Spirometra erinaceieuropaei* and *Diphyllobothrium* species using complete *cytb* and *cox1* genes. *Infect Genet Evol* 35: 1–8. <https://doi.org/10.1016/j.meegid.2015.07.025> PMID: 26205687
20. Zhang X, Wang H, Cui J, Jiang P, Lin ML, et al. (2016) The phylogenetic diversity of *Spirometra erinaceieuropaei* isolates from southwest China revealed by multi genes. *Acta Trop* 156: 108–114. <https://doi.org/10.1016/j.actatropica.2016.01.012> PMID: 26774686
21. Zhang X, Duan JY, Shi YL, Jiang P, Zeng DJ, et al. (2017) Comparative mitochondrial genomics among *Spirometra* (Cestoda: Diphyllbothriidae) and the molecular phylogeny of related tapeworms. *Mol Phylogenet Evol* 117: 75–82. <https://doi.org/10.1016/j.ympev.2017.06.003> PMID: 28606444
22. Otranto D, Lia RP, Traversa D, Giannetto S (2003) *Thelazia callipaeda* (Spirurida, Thelaziidae) of carnivores and humans: morphological study by light and scanning electron microscopy. *Parassitologia* 45: 125–133. PMID: 15267100
23. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. *Int J Parasitol* 34: 191–203. <https://doi.org/10.1016/j.ijpara.2003.10.004> PMID: 15037105
24. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
25. Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30: 2725–2729. <https://doi.org/10.1093/molbev/mst197> PMID: 24132122
26. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187> PMID: 19346325
27. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036> PMID: 10331250
28. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x> PMID: 21565059
29. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595. PMID: 2513255
30. Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitch-hiking, and background selection. *Genetics* 147: 915–925. PMID: 9335623
31. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29: 1969–1973. <https://doi.org/10.1093/molbev/mss075> PMID: 22367748
32. Zarlenga DS, Rosenthal BM, La Rosa G, Pozio E, Hoberg EP (2006) Post-miocene expansion, colonization, and host switching drove speciation among extant nematodes of the archaic genus *Trichinella*. *Proc Natl Acad Sci USA* 103: 7354–7359. <https://doi.org/10.1073/pnas.0602466103> PMID: 16651518
33. Rambaut A, Drummond AJ (2007) Tracer v1.5. Available at: <http://beast.bio.ed.ac.uk/> Tracer.
34. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61: 539–542. <https://doi.org/10.1093/sysbio/sys029> PMID: 22357727
35. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9: 772.
36. Shen J, Gasser RB, Chu D, Wang ZX, Yuan X, et al. (2006) Human thelaziosis—a neglected parasitic disease of the eye. *J Parasitol* 92: 872–875. <https://doi.org/10.1645/GE-823R.1> PMID: 16995411
37. Héritier L, Badets M, Du Preez LH, Aisien MSO, Lixian F, et al. (2015) Evolutionary processes involved in the diversification of chelonian and mammal polystomatid parasites (Platyhelminthes, Monogenea, Polystomatidae) revealed by palaeoecology of their hosts. *Mol Phylogenet Evol* 92: 1–10. <https://doi.org/10.1016/j.ympev.2015.05.026> PMID: 26072314

38. Yilmaz E, Fritzenwanker M, Pantchev N, Lendner M, Wongkamchai S, et al. (2016) The Mitochondrial Genomes of the Zoonotic Canine Filarial Parasites *Dirofilaria (Nochtiella) repens* and Candidatus *Dirofilaria (Nochtiella) Honkongensis* Provide Evidence for Presence of Cryptic Species. PLoS Negl Trop Dis 10: e0005028. <https://doi.org/10.1371/journal.pntd.0005028> PMID: [27727270](https://pubmed.ncbi.nlm.nih.gov/27727270/)
39. Blouin MS, Yowell CA, Courtney CH, Dame JB (1995) Host movement and the genetic structure of populations of parasitic nematodes. Genetics 141: 1007–1014. PMID: [8582607](https://pubmed.ncbi.nlm.nih.gov/8582607/)
40. Blouin MS, Liu J, Berry RE (1999) Life cycle variation and the genetic structure of nematode populations. Heredity 83: 253–259. PMID: [10504422](https://pubmed.ncbi.nlm.nih.gov/10504422/)