Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

International Journal for Parasitology 38 (2008) 1465–1479

www.elsevier.com/locate/ijpara

Geography and host specificity: Two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae)

Wafa Bouzid [a,c,1], Jan Štefka [b,1], Václav Hypša [b,*], Sovan Lek [a], Tomáš Scholz [b], Luc Legal [d], Oum Kalthoum Ben Hassine [c], Géraldine Loot [a]

[a] *Laboratoire Evolution et Diversité Biologique, U.M.R. CNRS-UPS 5174, Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse cedex 4, France*

[b] *Faculty of Science, University of South Bohemia and Biology Centre ASCR, Institute of Parasitology, Branišovská 31, 37005 České Budějovice, Czech Republic*

[c] *Unité de Recherche Biologie, Ecologie et Parasitologie des organismes aquatiques, Faculté des Sciences de Tunis, Tunisia*

[d] *ECOLAB, U.M.R. CNRS-UPS 5245, Université Paul Sabatier, Bâtiment IVR3, 118 Route de Narbonne, F-31062 Toulouse cedex 4, France*

Received 7 January 2008; received in revised form 13 March 2008; accepted 17 March 2008

Abstract

Parasite species with global distributions and complex life cycles offer a rare opportunity to study alternative mechanisms of speciation and evolution in a single model. Here, genealogy and genetic structure, with respect to geography and fish host preference, have been analyzed for *Ligula intestinalis*, a tapeworm affecting freshwater fish. The data analyzed consisted of 109 tapeworms sampled from 13 fish host species in 18 different localities on a macrogeographic scale. Two mitochondrial genes, cytochrome oxidase subunit I and cytochrome *B*, and the nuclear sequence of intergenic transcribed spacer 2 (ITS2) were used for the genetic reconstruction. Different evolutionary patterns were found at the local and at the global geographic scales. On a local scale, the flat genetic structure was mainly attributed to contiguous range expansion. Migrating birds are the most likely cause of the homogenisation of the whole population, preventing the creation of significant genetic barriers. By contrast, on a global scale, genetically distant and well-separated clusters are present in different geographic areas. Reproductive isolation was found even between clades living in sympatry and infecting the same definitive host, suggesting the existence of efficient biologically determined genetic barriers, and thus possibly separate species. Although the ITS2 sequences were found to display considerable intragenomic variability, their relationships were generally in good agreement with the topology derived from mitochondrial genes.

© 2008 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: Phylogeography; Host specificity; Parasite evolution; Cryptic species; Intragenomic variability; *Ligula*

1. Introduction

Inferring the genetic structure of parasites provides great opportunities to elucidate questions about host specificity, co-speciation events and parasite evolution in general. The genetic structure of parasites is becoming accessible to rigorous analysis with the use of tools combining molecular phylogeny and population genetics. Several studies have demonstrated that the genetic structure and the evolution of parasites at the intraspecific level vary considerably among groups of parasites (e.g. Anderson, 2001; McCoy et al., 2001; Johnson et al., 2002; Brant and Ortí, 2003; Wickström et al., 2003).

The large diversity of life-history traits displayed by parasites may explain such variation. For instance, diversity in modes of reproduction (e.g. hermaphroditic, parthenogenetic)

* Corresponding author. Tel.: +420 387 775 409; fax: +420 385 310 388. *E-mail address:* vacatko@paru.cas.cz (V. Hypša).

[1] These authors contributed equally.

0020-7519/$34.00 © 2008 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2008.03.008

and in dispersal capability of the free-living stage is major determinants of the genetic structure of parasites (Johnson et al., 2002; Criscione et al., 2005). In addition, host vagility may also predispose parasites to various genetic structures and ultimately dictate various patterns of speciation (Blouin et al., 1995; McCoy, 2003; Criscione and Blouin, 2004).

Most studies have focused on parasites with a simple life cycle. Examples include ectoparasitic arthropods such as ticks on nesting seabirds (McCoy et al., 2001; McCoy, 2003) and lice on doves (Page et al., 1998; Johnson et al., 2002). In parasitic platyhelminthes, nematodes and acanthocephalans, complex life cycles with one or several intermediate hosts are extremely common (Poulin, 1998). For such parasites, which present complex systems, the genetic structure is controlled by life-history traits in different parasite stages and by the mobility of the hosts (Jarne and Théron, 2003; Prugnolle et al., 2005). Criscione and Blouin (2004) demonstrated that parasites cycling in freshwater hosts only (an autogenic life cycle) had much more highly structured populations and much lower gene flow among sub-populations than parasites cycling through freshwater and terrestrial hosts (an allogenic life cycle).

The tapeworm *Ligula intestinalis* (L.) is the most common species of the genus *Ligula* (Bloch, 1782). It is widely distributed throughout the whole Holarctic region (Dubinina, 1980) and has recently been also reported from Australasia (Morgan, 2003; Chapman et al., 2006). This cestode presents a complex life cycle with a cyclopoid or diaptomid copepod as first intermediate host and planktivorous fish as second intermediate host. Fish-eating birds serve as the final host in which *L. intestinalis* quickly reaches sexual maturity and releases eggs into the water. The most conspicuous stage within the life cycle is the plerocercoid. It develops in the abdominal cavity of the second intermediate host and has a considerable effect on fish health, fecundity and behaviour. As a result it can cause heavy losses in freshwater pisciculture (Arme and Owen, 1968; Sweeting, 1977; Taylor and Hoole, 1989; Wyatt and Kennedy, 1989; Carter et al., 2005). Although typically reported from cyprinid fish, *L. intestinalis* has been shown to utilize a broad range of hosts, including other fish families such as Catostomidae, Salmonidae or Galaxiidae (e.g. Dubinina, 1980; Bean and Winfield, 1992; Groves and Shields, 2001; Museth, 2001; Barus and Prokes, 2002; Chapman et al., 2006). However, these lists of published records should be considered with caution, as they do not provide an accurate measure of host specificity in a given local population.

Records available on *L. intestinalis* indicate that the host spectra are not uniform across the investigated populations. For instance, analysis of host specificity in natural and experimental fish populations in south-west France has shown that silver bream, *Blicca bjoerkna*, is refractory to *Ligula* infection (Loot et al., 2002). This finding strongly contrasts with other studies identifying bream as a suitable host (Dubinina, 1980; Barus and Prokes, 1994). Loot et al. (2006) suggested that the abundance of potential hosts and the temporal dynamics of their occurrence strongly affect local host specificity. A similar role of host-abundance in the evolution of tapeworm specificity has been suggested for proteocephalids in South America (Hypša et al., 2005).

In addition to the differences in host spectra, Arme (1997) demonstrated that the pathology of *Ligula* infections in two cyprinid species, the roach (*Rutilus rutilus*) and the gudgeon (*Gobio gobio*) differs in several respects. In natural infections of roach, there is pronounced inflammation of host tissue which is never found in natural infections of gudgeon. There are also qualitative and quantitative differences in splenic and pronephric leucocyte counts and in spleen weights between the two hosts. It has been proposed that parasites from roach and gudgeon differ in their ability to stimulate a host response. This could be attributed to the existence of different *Ligula* strains or species (Arme, 1997; Kennedy and Burrough, 1981; McManus, 1985). Recently, Olson et al. (2002) studied *Ligula* tapeworms from three cyprinid fishes, namely gudgeon (*G. gobio*), minnow (*Phoxinus phoxinus*) and roach (*R rutilus*), and provided molecular evidence for separate *Ligula* strains. Luo et al. (2003) found relatively high intraspecific variability between *Ligula* specimens when comparing the genetic sequences of the entire internal transcribed spacer of the ribosomal DNA (ITS rDNA) and the 5′ end of 28S rDNA. Logan et al. (2004), using nucleotide variation of ITS2, showed that samples from Turkey identified as *L. intestinalis* were genetically distinct from European and Chinese *Ligula* isolates from various fish hosts.

A further complication of the *L. intestinalis* phylogenetic picture stems from the indication that *L. intestinalis* may be paraphyletic with at least two species, *Ligula colymbi* and *Digramma interrupta*. In a previous study based on ITS2, these two species clustered within the assemblage of *L. intestinalis* samples (Logan et al., 2004). Whereas the validity of *L. colymbi* might be questioned due to the scarcity and ambiguity of discriminating morphological characters, *D. interrupta* is a morphologically well-defined species which can be reliably distinguished. Taken together, these facts invoke a suspicion that the taxon designated as *L. intestinalis* may in fact be composed of several lineages and/or cryptic species. This hypothesis, emerging from available observations and offering a good explanation of reported discrepancies, has not yet been addressed by rigorous analysis of the genealogy and genetic population structure. The relative significance of geography and host specificity in the speciation process and evolution within the *L. intestinalis* complex thus remains unclear.

The aim of this study is to examine the phylogenetic structure of the common cestode, *L. intestinalis* on a macroevolutionary scale using mtDNA and ITS2 regions. Sequence divergence was used to explore the possible presence of cryptic species and discuss the mechanisms of speciation. On a finer scale, we explored the spatial distribution of genetic variation within the major haplotype cluster (tapeworms from the Europe-Mediterranean region) to discriminate between historic versus geographic events.

2. Materials and methods

2.1. Parasite samples

Tapeworms were collected from 23 localities across a broad geographic range covering 18 countries (see Table 1). The samples were isolated from body cavities of fish intermediate hosts and preserved either by placing them in 96% ethanol or freezing them at −80 °C. The specimens were identified as *L. intestinalis*, *Ligula* sp. or *D. interrupta* using the characters suitable for species identification according to Dubinina (1980), i.e. the shape of the anterior end of the body, the presence/absence of external segmentation (metamerism) of the anterior part of the body and the number of genital complexes visible at fully developed plerocercoids. We also included 11 specimens isolated from definitive hosts (eight from the great crested grebe, *Podiceps cristatus*, two from *Gavia stellata* and one from *Mergus merganser*).

2.2. PCR amplification and DNA sequencing

Genomic DNA was extracted from small pieces of tapeworm tissue (10–20 mg) using a Qiagen Tissue extraction kit. Three DNA regions were amplified; the partial sequence of nuclear ribosomal ITS2, the 5.8S partial sequence rDNA and two mitochondrial genes, cytochrome

Table 1
Geographic origin of the samples sequenced here

Geographic origin		Symbol	Host species	Species	No. of samples analysed per gene		
					ITS2	COI	COB
Algeria	Oued Hamiz Reservoir	Al BS	*Barbus* sp.	*L.* sp.	4	3 (1,2,3)	3
Australia	Goodga River	AU Gt	*Galaxias truttaceus* (Osmeriformes)	*L.* sp.	1	1 (4)	1
	Moates Lake	Gm	*Galaxias maculatus* (Osmeriformes)	*L.* sp.	1	1 (5)	1
Canada	Dalpec Lake	CA Cp	*Coulsius plumbeus*	*L.* sp.	1	1 (8)	1
	Lac Dumbo	Sa	*Semotilus atromaculatus*	*L.* sp.	8	4 (6,7)	4
China	Dong Tink Lake	CN Hb	*Hemiculter bleekeri*	*D. i.*	2	2 (9)	2
	Zhanghe reservoir	Nt	*Neosalanx taihuensis* (Osmeriformes)	*L.* sp.	7	7 (10,11)	3
Czech Rep.	Lipno Reservoir	CZ1Rr	*Rutilus rutilus*	*L. i.*	8	9 (14,15,16,17,18)	4
	Zelivka Reservoir	CZ2Rr	*Rutilus rutilus*	*L. i.*	–	1 (19)	1
		Aa	*Alburnus alburnus*	*L. i.*	1	1 (14)	1
	Nove Mlyny Reservoir	CZ3Ab	*Abramis brama*	*L. i.*	13	12 (14,20,21,22)	6
	Tlumacov ponds	CZ4Pc	*Podiceps cristatus* – bird host	*L.* sp.	4	7 (14,23,24,25,26)	8
		Mm	*Mergus merganser* – bird host	*L.*sp.	–	1 (14)	1
Estonia	Peipsi Lake	EE Ab	*Abramis brama*	*L. i.*	5	3 (14, 27)	3
Ethiopia	Tana Lake	ET Bh	*Barbus humilis*	*L.* sp.	2	1 (28)	1
		Bt	*Barbus tsanensis*	*L.* sp.	1	1 (29)	1
		Bi	*Barbus intermedius*	*L.* sp.	1	1 (30)	1
		Bb	*Barbus brevicephalus*	*L.* sp.	1	1 (31)	-
France	Pareloup Reservoir	FR1Rr	*Rutilus rutilus*	*L. i.*	7	1 (14)	1
		Aa	*Alburnus alburnus*	*L. i.*	1	1 (14)	1
	Lavernose and Muret	FR2Rr	*Rutilus rutilus*	*L. i.*	4	–	-
		Bb	*Blicca bjoerkna*	*L. i.*	2	1 (14)	1
	Creteil Lake	FR3Rr	*Rutilus rutilus*	*L. i.*	7	3 (14, 32, 33)	3
Germany	Müggelsee Lake	DE Rr	*Rutilus rutilus*	*L .i.*	4	4 (34)	4
Great Britain	Scotland, River Gryfe	GB Rr	*Rutilus rutilus*	*L .i.*	–	1 (35)	1
		Pp	*Phoxinus phoxinus*	*L. i.*	–	1 (36)	1
	Wales, Aberystwyth	Pp	*Phoxinus phoxinus*	*L.* sp.	–	1 (14)	1
Mexico	Tonatzahua lake	MXGm	*Girardinichthys multiradiatus* (Cyprinodontiformes)	*L.* sp.	2	2 (38,39)	[a]
N. Ireland	Lough Neagh	IE Rr	*Rutilus rutilus*	*L.* sp.	–	3 (14)	3
		Gg	*Gobio gobio*	*L.* sp.	–	3 (25, 37)	3
Poland	Gdansk	PL Gs	*Gavia stellata* – bird host	*L.* sp.	2	1 (40)	–
	Wloclawski reservoir	PL Ra	*Rhodeus amarus*	*L.*sp.	1	1 (25)	1
Russia	Khanka Lake, Far East	RU Hl	*Hemiculter lucidus*	*D. i.*	–	1 (43)	1
	Rybinsk Reservoir	RU Ab	*Abramis brama*	*L. i.*	9	9 (14, 41,42)	5
Tunisia	Sidi Salem Reservoir	TN Rb	*Rutilus rubilio*	*L. i.*	6	5 (14,44)	4
		Se	*Scardinius erythrophthalmus*	*L. i.*	1	1 (14)	1
Ukraine	Dniester River	UA Rr	*Rutilus rutilus*	*L. i.*	1	1 (46)	1
		Aa	*Alburnus alburnus*	*L. i.*	1	1 (45)	1
		Cc	*Carassius carassius*	*L. i.*	1	1 (14)	1
Totals					109	99 (46)	76

L. i., *Ligula intestinalis*; *L.* sp.: sample could not be unequivocally determined (*L. intestinalis* versus *Ligula colymbi*; see text for details); *D. i.*, *Digramma interrupta*.
Numbers in parentheses refer to haplotype numbers used in Fig. 1.
[a] No PCR product was obtained for this gene. Sequences obtained from GenBank are listed in Table 3.

oxidase subunit I (COI) and cytochrome *B* (COB). A new primer pair COIA2–COIB2 was designed on the basis of the primers COIA/COIB (Li et al., 2000) and cestode COI sequences available in GenBank. The COBA–COBB primers were constructed using the COB sequences available in the GenBank mitochondrion database (Accession No. AF314223). Primer sequences are given in Table 2. PCRs were carried out in a 20 μl vol. using 1 μl of extracted DNA solution (approximately 20 ng), 5 pM of each primer, 15 mM $MgCl_2$, 10 mM of each dNTP and 0.25 U of *Taq* polymerase. Amplifications were performed using methods modified from Logan et al. (2004), Li et al. (2000) or von Nickisch-Rosenegk et al. (2001). Purified DNA (20 ng/μl) was either sequenced directly with ABI BigDye chemistry using the amplification primers, or cloned into *Escherichia coli* (Invitrogen) using pGEM-Teasy vector (Promega). Plasmids were sequenced using the universal primers M13 or Sp6 and T7. Sequencing products were alcohol-precipitated and separated on ABI 3100 or 3130 automatic sequencers (Applied Biosciences).

2.3. Phylogenetic analysis

To prepare alignments, new sequences were supplemented with data from GenBank (Table 3). Mitochondrial genes were aligned with BioEdit 7.05 (Hall, 1999) and the program Collapse 1.2 (Posada, 2004, Collapse. A tool for collapsing sequences to haplotypes, version 1.2, http://darwin.uvigo.es) was used to retrieve individual haplotypes. The ITS2 sequences could not be unequivocally aligned due to considerable length variation, particularly within the regions containing microsatellite repeats. The sequences were prealigned using Megalign (DNASTAR, Inc.) and the alignment was manually adjusted in Bioedit 7.05 (Hall, 1999). The following procedure was then applied to extract phylogenetic characters from the ITS2 sequences. The microsatellite regions were removed altogether since they could not be reliably aligned. Within the rest of the alignment, each indel (usually 2–5 bp long) was treated as a single evolutionary event. To achieve this, the first position of each deletion was coded as a gap, while the rest was considered missing data. In this way, the weight of all indels was identical regardless of their actual lengths.

Sequences of mitochondrial genes were used in several subsequent steps. First, the relationships among the haplotypes on a global scale were inferred using the COI alignment. The COB sequences, displaying a higher degree of variability, were then explored to establish the intrapopulational relationships within the two European lineages (clades A and B; see Section 3). Finally, a concatenated matrix of both genes was used to evaluate the genealogical structure of clade A by nested clade phylogeographic analysis (NCPA) (Panchal, 2007).

Maximum parsimony (MP) using the matrix of COI haplotypes was performed in PAUP* (Swofford, 1998. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods), version 4.0. Sinauer Associates, Sunderland, MA) with Ts:Tv weights set to 1:1, 1:2 and 1:3 using the TBR algorithm with 50 replicates of random sequence addition. Calculation of bootstrap support was executed with 1000 replications of TBR search with Ts:Tv set to 1:1. The maximum likelihood (ML) tree was constructed in Phyml (Guindon et al., 2005) with a GTR+γ+Inv model and the parameters estimated from data, bootstrap support was obtained by 1000 replications. COI sequences of the diphyllobothridean tapeworm *Diphyllobothrium latum* (NC008945) and two cyclophyllidean tapeworms, *Taenia crassiceps* (NC002547) and *Hymenolepis diminuta* (AF314223), were used as outgroups. MP analysis of ITS sequences, together with *D. latum* (DQ768176.1) as an outgroup, was performed using the TNT program (Giribet, 2005) with a TBR algorithm with 500 replications.

Several methods were employed to evaluate the demographic information on the European and North African populations of *Ligula* samples (clades A and B, Fig. 3). Neutrality tests were performed separately for the COI and COB haplotypes; 95% and 99% confidence intervals were calculated by 10,000 coalescence simulations using DNASP 4.10 (Rozas et al., 2003). Mismatch distribution of pairwise substitutions between COB haplotypes were calculated and compared with the Poisson model using DNASP 4.10. The observed patterns of population growth/decline were tested using the Lamarc 2.0.2 software package (Kuhner, 2006) under a model of molecular evolution identified in Modeltest 3 (Posada and Crandall, 1998). Lamarc was run in a likelihood mode and the search strategy consisted of three replications of 10 short initial chains followed by two long final chains. The initial chains were run with 500 samples and a sampling interval of 20 (10,000 samples); burn-in was set to 1000 samples for each chain. The same parameters were used for the final chains performed with 10,000 samples. Values of population size (*Theta*) and population growth (*g*) were estimated and confidence intervals calculated using the percentile approach.

Table 2
Primers used for PCR

Target genes	Primers Forward Reverse	Sequence (5′–3′ direction)	Reference
ITS2 (rDNA)	Flo1	CGGTGGATCACTCGGCTC	Logan et al. (2004)
	ITS2	TCCTCCGCTTATTGATATGC	
COI (mtDNA)	COIA2	CATATGTTTTGATTTTTTGG	This study
	COIB2	AKAACATAATGAAAATGAGC	Modified from Li et al. (2000)
COB (mtDNA)	COBA	GTATGTGGCTGATTCAGAGTTGAGC	This study
	COBB	TTCGAGCCCGAAGAATGCAAGTAG	

Table 3
Sequences downloaded from GenBank

Geographic origin		Symbol	Host species	Species	No. of sequences analysed per gene		
					ITS2	COI	Accession no.
China	Qinghai	CN Gp	*Gymnocypris przewalskii*	*L.* sp.	2	–	AY121751/52[a]
	Jiangxi	CN Hb2	*Hemiculter bleekeri*	*D. i.*	1	–	AF354293[a]
	Hubei	CN Cd	*Culter dabryi*	*D. i.*	2	–	AF354291/92[a]
	Hubei	CN Ca	*Carassius auratus*		1	–	AY12175[a]
	various localities	CN XL	various Cyprinidae and *Protosalanx hyalocranius* (Osmeriformes)	*L.* sp.	–	1 (12)	AF524041[b]
	various localities	CN XD	various cyprinidae	*D. i.*	–	1 (13)	AF153910[c]
Czech Rep.	Lipno Reservoir	CZ1Rr	*Rutilus rutilus*	*L. i.*	1	–	AY549520[d]
	Nove Mlyny Reservoir	CZ3Ab	*Abramis brama*	*L. i.*	1	–	AY549519[d]
	Tlumacov Ponds	CZ4Pc	*Podiceps cristatus*	*L. c.*	2	–	AY549510/11[d]
Great Britain	Scotland, River Gryfe	GB Rr	*Rutilus rutilus*	*L. i.*	1	–	AY549512[d]
		Pp	*Phoxinus phoxinus*	*L. i.*	1	–	AY549513[d]
	Wales, Aberystwyth	Pp	*Phoxinus phoxinus*	*L.* sp.	1	–	AF385760[e]
	N. Ireland, Lough Neagh	IE Rr	*Rutilus rutilus*	*L.* sp.	3	–	AF385761/63/64[e]
		Gg	*Gobio gobio*	*L.* sp.	3	–	AF385765/67/69[e]
Russia	Khanka Lake, Far East	RU Hl	*Hemiculter lucidus*	*L. i.*	1	–	AY549506[d]
Turkey	Iznik Lake	TR CS	*Chalcaburnus sp.*	*L. i.*	1	[f]	AY549517[d]
		Sg	*Silurus glanis* (Siluriformes)	*L. i.*	1	[f]	AY549516[d]
Totals					22	2 (2)	

Tapeworm species determination was adapted from the original studies: [a]Luo et al., 2003; [b]Li and Liao, 2003; [c]Li et al., 2000; [d]Logan et al., 2004; [e]Olson et al., 2002. Numbers in parentheses refer to haplotype numbers used in Fig. 1.
[f]No PCR product was obtained for this gene.

The genealogical structure of clade A populations was constructed in program TCS 1.21 (Clement et al., 2000) using a concatenated matrix of COI and COB, which produced 37 haplotypes. Information about the biogeographical history of European populations (clade A) was inferred using software for NCPA. This program implements TCS and GeoDis algorithms (Clement et al., 2000; Posada et al., 2000) testing the congruence between network structure and geography, and an inference key of Templeton (2004) evaluating possible historical and geographical events. The probability of the null hypothesis (no association of genetic structure with geography) was estimated by 1 million permutations. The geographic coordinates of the localities were used to describe their distribution; the radius of the water body was used as a size estimate for each sampled area. To allow for possible long distance migration and/or population fragmentation between the European continent and its surroundings, the North Sea and the Mediterranean were entered as regions uninhabitable for *Ligula*. Following the suggestions of Posada et al. (2000) and Panchal (2007), three regions with excessively large gaps between the sampled populations were indicated to prevent a false inference of isolation by distance. These regions cover unsampled areas in southern Germany and Austria, two areas in Eastern Europe (corresponding approximately to Poland and Byelorussia), and the western part of Russia.

3. Results

3.1. Sequences and alignments

We obtained partial sequences of three DNA regions from 109 *L. intestinalis*, *Ligula* sp. and *D. interrupta* samples: two mitochondrial genes COI (396 bp) and COB (404 bp), and the non-coding nuclear sequence of ITS2 varying from 610 to 650 bp in size.

3.1.1. Mitochondrial genes

In the 101 samples sequenced for COI (including two sequences retrieved from GenBank) and 76 samples sequenced for COB, we identified 46 (COI) and 44 (COB) different haplotypes. Distribution of COI haplotypes among the samples was extremely uneven. While the majority of the haplotypes were only represented by a single *Ligula* specimen, the most frequent haplotype (H14) was retrieved from 39 European samples. The set of COB haplotypes was more evenly distributed, with the two most abundant haplotypes containing eight and seven samples. The complete list and distribution of COI haplotypes is provided in Tables 1 and 3. The levels of genetic variability were similar for both genes. Without outgroups, the 396 positions long COI matrix contained 275 (70%) constant and 89 (22%) parsimony informative characters. For COB, the proportion of constant sites was 278 out of 405 (67%) and the number of parsimony informative sites was 91 (22%). Alignments of mitochondrial genes were unequivocal and did not require any indel adjustments. The observed variability consisted almost exclusively of synonymous mutations at the third codon position.

3.1.2. ITS2

In total 188 ITS2 sequences from 131 specimens were obtained (Tables 1 and 3). Screening for ITS2 intragenomic variability confirmed that different forms of ITS2 can occur together in a single genome. Eight sequences contained 70–

430 bp deletions extending into the 3′ end of 5.8S rDNA. These sequences were considered pseudogenes and were excluded from further analyses. Among the presumably functional sequences, two types of intragenomic variability were observed. The first was only minor. In phylogenetic analyses these variants clustered as closely related sequences within the same COI-derived clade. This variability was mainly due to singletons, numbers of short repeats and several short indels. The second type was represented by highly divergent paralogues; in the phylogenetic tree, these paralogues clustered in different lineages. In contrast to the *Ligula* samples, a low level of variability was observed in *D. interrupta* (Ru Hl). The five haplotypes identified only differed by a few single-nucleotide mutations and in one case by a microsatellite region. The intragenomic variability is summarized in Table 4. The sequences were deposited in GenBank (NCBI) under Accession Nos. EU240978–EU241317. The numbers of samples sequenced for particular genes are reported in Table 1.

3.2. Phylogenetic analysis

3.2.1. COI and COB

The MP analysis of COI performed under three different Ts/Tv ratios yielded 286 trees. Their strict consensus was entirely compatible with the ML tree and contained several well-supported clades (Fig. 1). Distribution of haplotypes among the clades was consistent with their geographic origin. Whereas some of the geographic areas were represented by a single monophyletic clade (e.g. Canada, Mexico and Ethiopia), the samples from the Euro-Mediterranean area were split into two distinct lineages (the clades A and B). Clade A was comprised of samples from Europe and Tunisia, and clade B contained, apart from European and Algerian samples, the samples from China and Australia. The split of European populations into two clades reflected the difference in host range observed in these two groups of parasites. The European and Tunisian samples of clade A were recovered exclusively from the phylogenetically derived cyprinid fish (*Abramis*, *Alburnus*, *Phoxinus*, *Rutilus*, *Scardinius*), whereas the European and Algerian samples of clade B were found to parasitize the basal cyprinid species (*Barbus*, *Gobio*, *Rhodeus*). Further geographically dependent diversification was seen within clade B. The Algerian lineage was composed of haplotypes 1 and 2, the mostly European lineage (Haps 25, 26, 37 and 40) also included a single Algerian haplotype, Hap 3. The Chinese and Australian samples again formed two distinct monophyletic branches, represented by haplotypes 10–11 and 4–5, respectively. The isolates collected from *P. cristatus* (Czech Republic) were scattered over clades A (haplotypes 14, 23, 24) and B (haplotypes 25, 26). The COB sequences provided less robust topology, but they supported the basic geographically-dependent pattern, including the split of European samples into two lineages (not shown). On the other hand, the higher degree of COB variability resulted in better resolution within clade A (26 haplotypes out of 51 samples compared with 22 haplotypes out of 66 COI haplotypes). This information was used for population structure analysis based on alignments of both mitochondrial genes (see below).

The samples determined as *D. interrupta* clustered at two different positions. The Russian and Chinese samples sequenced in this study (H9 and H43) formed a monophyletic clade separated from Chinese and European *Ligula* of clades A and B. By contrast, haplotypes H12 and H13, corresponding to the Chinese *Ligula* and *Digramma* samples sequenced by Li and Liao (2003) and Li et al. (2000), branched together with clades A and B.

3.2.2. ITS2

The ITS2 sequences provided a weaker phylogenetic signal resulting in lower topological resolution. Despite this limitation, the strict consensus of 286 MP trees was surpris-

Table 4
Intragenomic variability of ITS2 sequences

Sample-Voucher No.	COI Haplotype No.	No. of plasmids sequenced	No. of haplotypic sequences		
			Total	Clustering within the COI clade	Clustering outside the COI clade, or pseudogene
CHNt8	10	20	9	7 (clade B2)	2 (pseudogenes)
CZ1Rr15	14	20	11	9 (clade A)	2 (pesudogenes)
RUAb2	14	20	9	7 (clade A)	2 (unknown ITS only lineage)
RUHl	43	15	5	5 (*Digramma*)	0
CNNt6	10	2	2	1 (China)	1 (pseudogene)
EEAb1	14	1	1	0 (clade A)	1 (Digramma)
EEAb3	14	2	2	1 (clade A)	1 (pseudogene)
CZ4Pc107	14	2	2	0 (clade A)	2 (unknown ITS only lineage)
PLGs1	40	2	2	0 (clade B1)	1 (clade A) 1 (clade B2)
CZ3Ab26	14	2	2	1 (clade A)	1 (pseudogene)
CZ4Pc2	26	[a]	1	0 (clade B1)	1[a] (unknown ITS only lineage)
IERr6	14	2	3	1 (clade A) 1[a] (clade A)	1 (clade B1)
MXGm2	39	1	1	0 (Mexico)	1 (clade A)

"Clustering within/outside the COI clade" refers to the ITS sequence(s) position in Fig. 2.
[a] Sequences obtained from GenBank.

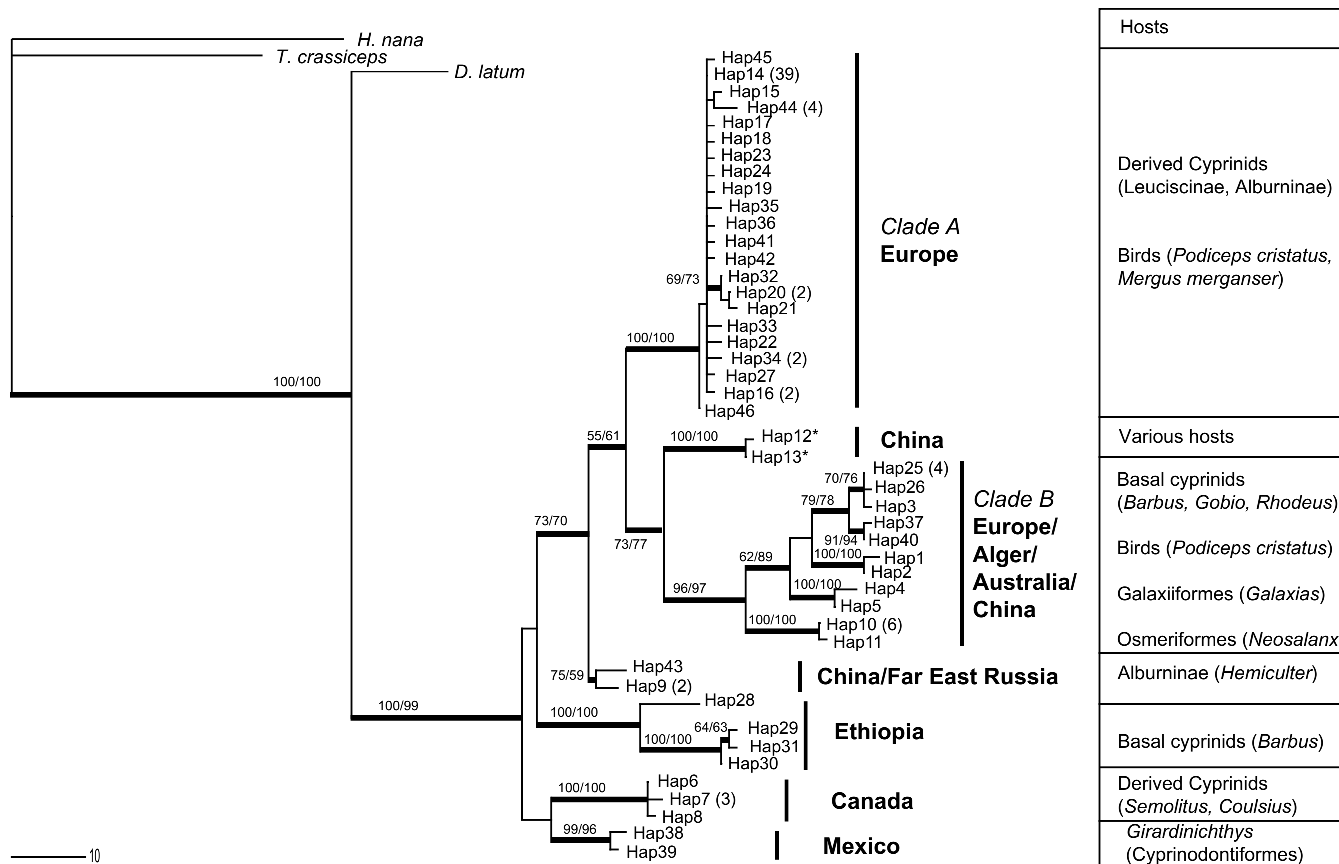

Fig. 1. One of the 288 maximum parsimony trees provided assuming a transition/transversion ratio of 1, 2 and 3; based on sequences of cytochrome oxidase subunit I. The numbers at the nodes indicate bootstrap support values higher than 50% (maximum parsimony/maximum likelihood, 1000 replications). Strict consensus in all trees is depicted in bold lines. Haplotype numbers refer to the numbers listed in Tables 1 and 3. Asterisks indicate sequences obtained from GenBank. Numbers in parentheses are the number of samples grouped within a haplotype. Phylogenetic position of cyprinids (basal/derived) was based on Briolay et al. (1998), Gilles et al. (2001), Cunha et al. (2002), Durand et al. (2002), Liu and Chen (2003) and Saitoh et al. (2006).

ingly compatible with the main geographically-specific lineages inferred from the mitochondrial genes (Fig. 2). Out of the 180 ITS2 sequences analysed, the positions of 170 sequences were consistent with COI topology. Only occasionally were some of the ITS2 paralogues placed into two or more different clades (Table 4).

3.3. Population structure and genetics

Neutrality tests, originally developed to examine selection neutrality of mutations, proved to be a useful tool for testing population expansion. For COI and COB matrices, the neutrality tests provided mainly insignificant results; the only exception was found for COI sequences in clade A (Table 5). This finding corresponds to the low variability of COI haplotypes observed within this clade and indicates possible changes in population size in clade A. A similar trend, although not significantly supported, was identified in the COB matrix (Table 5). Furthermore, for clade A the mismatch distribution of COB haplotypes created a bell-shaped pattern indicating rapid population growth (Fig. 3), while the multimodal pattern corresponding to a stable non-expanding population was found in clade B. Similar differences were detected by coalescence analysis; both *Theta* and *g* were much higher in clade A than in clade B (Table 6). *Theta* is a joint estimator of effective population size and mutation rate. Since there is no reason to expect any considerable difference in mutation rate between the two lineages, we conclude that the higher value of *Theta* for the clade A populations reflects a larger and expanding population (larger *g* than in clade B). In concordance with these results, NCPA based on a concatenated alignment of both genes identified two cases of contiguous population expansion in European populations of clade A. The contiguous expansion was particularly significant for clade 3-2 and the tests were close to significant values for the entire cladogram (Fig. 4 and Table 7). In addition, restricted gene flow with isolation by distance (IBD) was suggested for clade 2-3. In clade 3-3, due to insufficient genetic resolution, no decision was made to distinguish restricted gene flow with IBD from long distance range expansion. The general lack of geographically dependent structure was particularly well documented in several mixed populations (e.g. RuAb, TuRb and CZ1Rr), containing haplotypes separated by many mutational steps and distributed across distant clades of the network (Fig. 4).

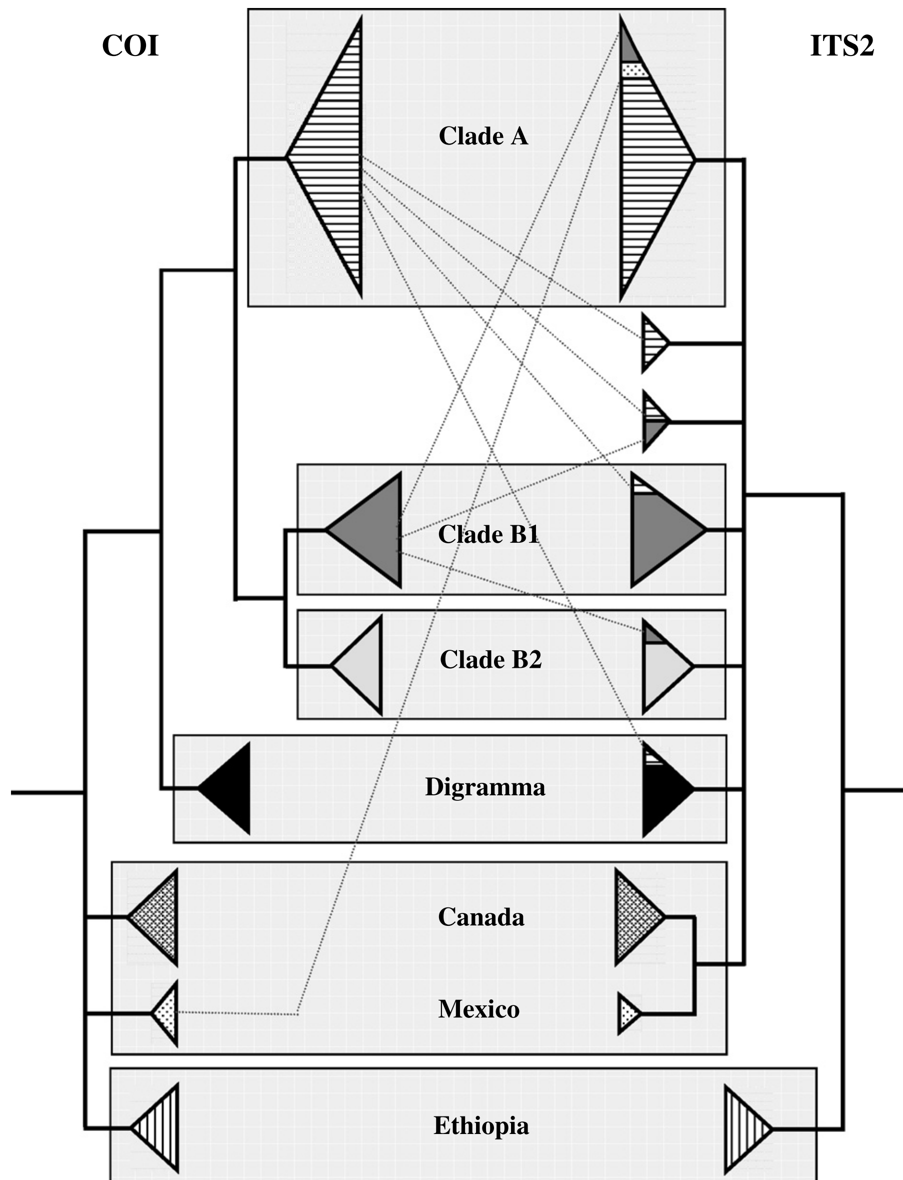

Fig. 2. Comparison of cytochrome oxidase subunit I (COI) and intergenic transcribed spacer 2 (ITS2) phylogenies. The COI tree corresponds to the topology shown in Fig. 1 with haplotypes 12 and 13 removed. The ITS2 tree is a strict consensus of 286 maximum parsimony trees. Positions of putative paralogues and their counterparts in the COI topology are highlighted with dotted lines (see Table 4).

Table 5
Molecular variability and neutrality tests

Clades		Sample size	No. of haplotypes	P_i	Tajima's D	Fu and Li's D	Fu and Li's F
A	COI	66	22	0.0035	−2.428[a]	−2.6425[b]	−3.055[a]
	COB	51	26	0.0069	−0.326	−1.185	−1.589
B	COI	10	7	0.0178	−0.241	0.900	0.727
	COB	9	5	0.0102	−0.326	0.020	−0.144

P_i, nucleotide diversity; levels of significance: [a] $P < 0.01$, [b] $P < 0.05$.

4. Discussion

The data presented here contain several strong components of population structure. The most conspicuous is an apparent correspondence between genetics and geography on a broad geographic scale (Fig. 1). This finding is not entirely unexpected since genetic variability due to long-distance separation is a common phenomenon. However, despite this general correspondence, the genealogical/phylogenetic arrangement does not imply any simple succession

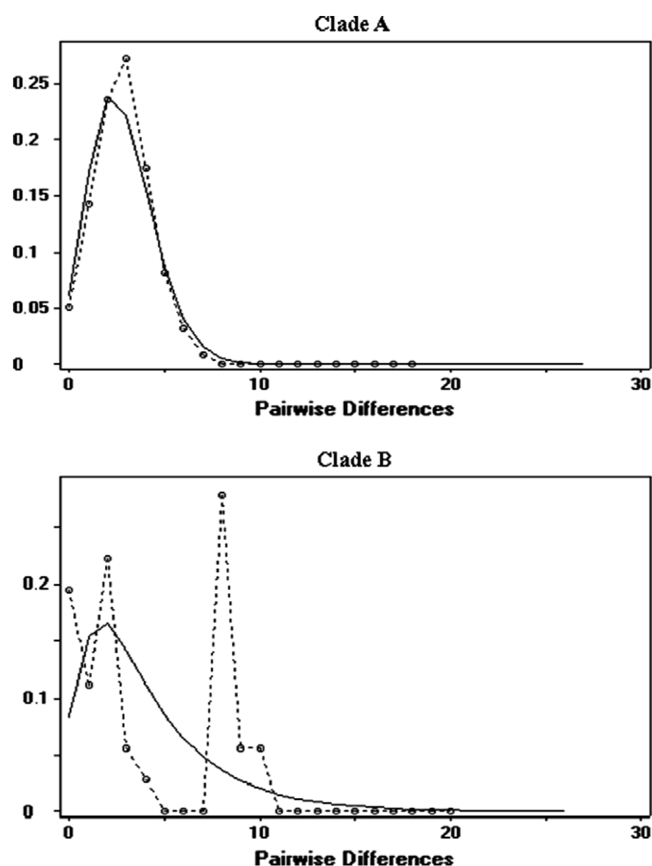

Fig. 3. Mismatch distribution of cytochrome oxidase subunit I haplotypes for two European *Ligula* clades. The observed frequency of numbers of pairwise mismatches is represented by a dotted line. The expected frequency (solid line) was estimated under the population-growth model in the program DNASP.

Table 6
Demographic parameters of European samples of *Ligula* given by coalescence analysis

Clade	*Theta*	*g*
A	0.118 (0.075–0.194)	953.1 (680.9–1208.6)
B	0.010 (0.005–0.023)	195.2 (163.3–214.1)

Values in parentheses represent 95% confidence intervals.

of vicariance events. The mitochondrial clades, although geographically restricted, do not cluster in any obvious pattern corresponding to the evolutionary history or distances between localities. It is reasonable to assume that at least in some cases the phylogeographic pattern is affected by unique long-distance transmissions mediated by introduced fish or migrating birds and followed by local diversification. For example, our set contains, to our knowledge, the first reported molecular data on *Ligula* specimens from the southern hemisphere, namely haplotypes H4 and H5 from Australia. Fish infections with *Ligula* have previously been reported from various sites in Australia and New Zealand, and both salmonid fish and fish-eating birds have been suggested as factors responsible for their introduction (Pollard, 1974; Weekes and Penlington, 1986; McDowall, 1990; Morgan, 2003). This view corresponds with our DNA-based results: despite considerable geographic distance and isolation, the Australian samples were placed firmly within clade B. However, since Australian fauna was only represented by samples from two adjacent localities, its overall diversity could be seriously underestimated. Similarly, Canadian, Mexican and Ethiopian haplotypes formed well-supported distinct clades, although based on only a few samples. Thus, with knowledge as it stands and the data currently available, it would be difficult to infer any global phylogeographic pattern.

Compared with the geography-based structure, the host-specificity component is expressed in a weaker but intriguing way. Within the most abundant clades A and B, no obvious correspondence was found between tapeworm phylogeny and specificity to fish-hosts. This indicates that within the geographically delimited clades, tapeworms are able to utilize several intermediate hosts without suffering from any apparent genetic isolation. Such mixing is consistent with the current view of the life cycle of *L. intestinalis*, since the tapeworms from different fish are supposed to cross-breed in their definitive bird host. In contrast to this intraclade homogeneity, striking differences in composition of intermediate-host spectra can be found among the clades, even if they occur in sympatry. For example, out of 66 samples collected from the European species and clustering within clade A, 65 were collected from derived cyprinids. The single exception is sample Ua Cc from Ukraine, isolated from a basal cyprinid fish *Carassius carassius*. No samples from derived cyprinids were placed into the related clade B, composed of tapeworms from basal cyprinids and other fish orders (Table 1).

The concept of cryptic species in which there are no discernible morphological characters despite genetic differences in mitochondrial loci is well documented. For example, sequence divergence between the chewing louse, *Columbicola passerinae*, parasitising the blue ground dove, *Claravis pretiosa*, differs from *C. passerinae* parasitising the common ground dove by 11.3% for a portion of the COI mitochondrial gene, indicating the two louse populations could be reclassified as two distinct species (Johnson et al., 2002). In our study, the genetic divergence seen in the mitochondrial COI gene tree between samples of haplotypes A and B was $8.1 \pm 0.02\%$. By comparison, the genetic divergence between two cestode species, *Paranoplocephala buryatiensis* and *Paranoplocephala longivaginata* using COI sequence data was $7.05 \pm 0.47\%$ and intraspecific genetic diversity was $0.5 \pm 0.5\%$ and $0.2 \pm 0.08\%$, respectively (Haukisalmi et al., 2007). Blouin et al. (1998) compared mtDNA sequence variation among individual nematodes of the same species and closely related species and showed that the typical intraspecific range was 2–6% and the interspecific range was 10–20%. Taking into account all these data, it is likely that genetic divergence alone cannot be straightforwardly used as proof of cryptic speciation. However, the analogy with closest known examples, i.e. the tapeworms of the genus *Paranoplocephala*,

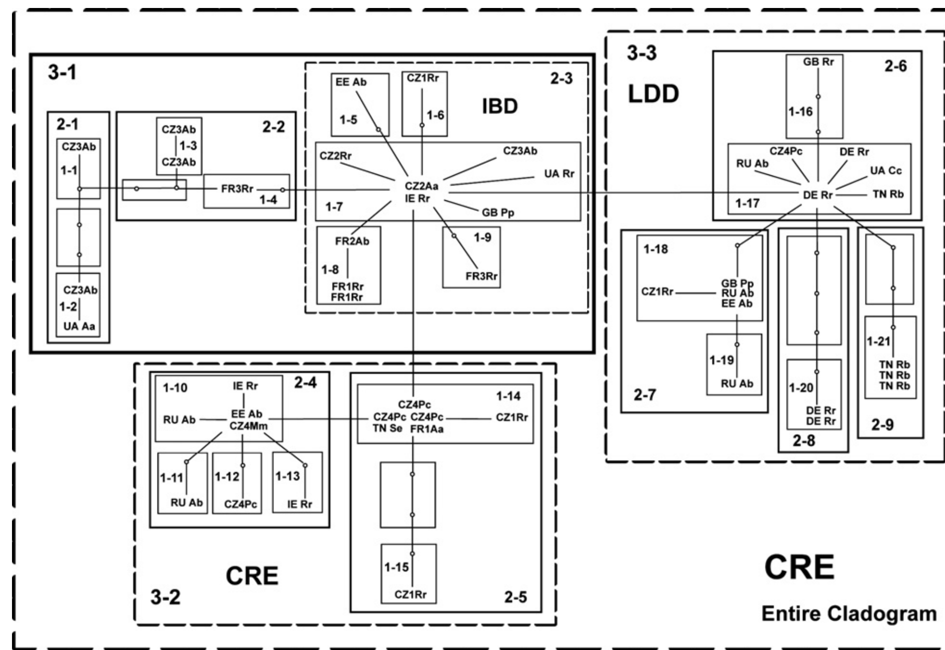

Fig. 4. Nested clade diagram obtained from concatenated matrix of the cytochrome oxidase subunit I and cytochrome *B* haplotypes of clade A. Neighbouring haplotypes are connected by a single mutation, open circles represent missing haplotypes along mutational pathways. Sample codes are given in Table 1. Nested clade levels are indicated by the numbers within a particular nested clade. Dashed lines highlight nesting categories where population events were recognized using the nested clade phylogeographic analysis. IBD, isolation by distance; CRE, contiguous range expansion; LDD, long distance dispersal. For levels of significance see Table 7.

indicates that the clades A and B could in principle represent two different biological species.

Regardless of their taxonomic status, the European haplotypes of clades A and B provide a typical instance of two genetically distinct lineages/species maintained in sympatry. Such situations constitute potential models for studying the speciation processes in parasites. It has been emphasized that to analyze the speciation scenario correctly, an appropriate usage of the sympatric/allopatric concept is an important prerequisite, which requires a priori knowledge of the parasite's biology and transmission mode (McCoy, 2003; Giraud, 2006). For example, parasite lineages inhabiting different host species living in sympatry should only be considered sympatric if the hosts do not pose an extrinsic barrier to their genetic mixture (Giraud, 2006). In the case of *L. intestinalis*, adult worms develop and mate in the intestine of a fish-eating bird serving as a definitive host. Despite their specificity to different intermediate hosts, the members of both clades A and B were found in the same definitive host, *P. cristatus*, sampled from the same locality. Therefore, the genetic isolation of clades A and B is maintained in sympatry and requires an explanation based on biological traits rather than physical isolation.

Such sympatric occurrence of two phylogenetically related lineages can be explained either by sympatric speciation or by a secondary encounter of cryptic species originating from allopatry. In a context of parasitology, disruptive selection has regularly been invoked as a possible means of sympatric speciation dependent on host biology (Théron and Combes, 1995; McCoy et al., 2001). It has even been suggested that the fish feeding strategy can determine the composition of its parasitic fauna and thus affect the overall co-evolutionary scenario (Jousson et al., 2000; Hypša et al., 2005; Criscione et al., 2005). We may hypothesize that the feeding strategy of pelagic (clade A) versus benthic fish (clade B) represents a potential diversifying force in parasites. In Lake Tana in Ethiopia, host specificity of *L. intestinalis* was shown to be likely attributable to host food type. Dejen et al. (2006) reports that *L. intestinalis* shows host specificity to *Barbus tanapelagius*, a specialized zooplanktivore, compared with *B. humilis* which has both zooplanktivore and benthic invertebrate food types.

A secondary encounter of two or more cryptic species with different intermediate host spectra seems a more plausible explanation. We hypothesize the scenario of host-parasite co-speciation (Hafner and Nadler, 1990; Hafner and Page, 1995; Hoberg et al., 1997; Hughes, 2007). In such a scenario, the two genetically distinct European lineages would have originated in allopatry and adapted to different groups of intermediate hosts. After a secondary encounter, these cryptic species retained genetic isolation and different host specificities. Indeed, the two groups of parasites reflected the divergence in host ranges with *Ligula* from clade A recovered exclusively from derived cyprinid fish and *Ligula* from clade B from basal cyprinid species. The phylogenetic analysis of European cyprinids showed the existence of two main fish clades corresponding to the sub-families Cyprininae and Leuciscinae (Zardoya and Doadrio, 1998; Gilles et al., 2001). The divergence of

Table 7
Results of the nested clade phylogeographic analysis

Nesting category	D_c $P \leqslant$	$P \geqslant$	D_n	$P \geqslant$	Chain of inference/inferred pattern
Clade 2-3					1-2-3-4 NO/
Subclade					
1-5	1.000	1.000	1.000	0.079	Restricted gene flow with isolation by distance (IBD)
1-6	1.000	1.000	0.241	0.839	
1-7	0.485	0.517	0.691	0.311	
1-8	0.034[a]	0.979	0.334	0.678	
1-9	1.000	1.000	0.170	1.000	
1-T	0.882	0.119	0.664	0.336	
Clade 3-2					1-19-20-2-11-12 NO/
Subclade					
2-4	0.032[a]	0.970	0.015[a]	0.987	Contiguous range expansion (CRE)
2-5	0.928	0.073	0.980	0.021[a]	
1-T	0.032[a]	0.970	0.018[a]	0.984	
Clade 3-3					1-2-3-5-6*-7-8 YES/
Subclade					
2-6					Restricted gene flow/dispersal but with some long-distance dispersal (LDD) over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations
2-7	0.282	0.727	0.257	0.751	
2-8	0.568	0.433	0.788	0.213	
2-9	0.114	1.000	0.078	0.965	
1-T	0.012[a]	1.000	0.956	0.049[a]	Too few clades: insufficient genetic resolution to discriminate between range expansion/colonization and restricted dispersal/gene flow
	0.855	0.115	0.234	0.766	
Entire cladogram					1-2-11-12 NO/
Subclade					
3-1					Contiguous range expansion (CRE)
3-2	0.130	0.870	0.099	0.901	
3-3	0.753	0.247	0.671	0.329	
1-T	0.924	0.076	0.937	0.063[b]	
	0.072[b]	0.928	0.062[b]	0.938	

Nesting categories are as in Fig. 4.
[a] Values of $P \leqslant 0.05$.
[b] Values of P close to 0.05 level, considered as significant when running the inference key.

Cyprinidae based on fossil evidence took place approximately 27.7 million years ago (Zardoya and Doadrio, 1999) and we hypothesize that the two *Ligula* clades are the product of a co-evolution event. Unfortunately, calibrated rates are lacking for platyhelminths because they do not leave a fossil record (Haukisalmi et al., 2007).

This latter view (i.e. the encounter of already separate cryptic species) is also supported by several previous findings. First, Olson et al. (2002) described two sympatric lineages of *Ligula* sp. from Northern Ireland with different types of ITS rDNA. These lineages, characterized by their affinity to two different intermediate hosts, *Rutilus* and *Gobio*, obviously correspond to our clades A and B. As noted by Olson et al. (2002), *Rutilus* is not native to Ireland. Its introduction in the 1970s was later followed by an increased occurrence of the great crested grebe, the final host of *Ligula.* It is likely that grebes then secondarily introduced clade A *Ligula* into Ireland where clade B was already well established. We found a similar situation in *Ligula* populations from northern Africa. Whereas the Algerian samples collected from *Barbus* clustered within clade B, the samples recovered from *Rutilus rubilio* in a neighboring locality in Tunisia were placed into the "European" clade A. Since *R. rubilio* was introduced into Tunisia from Southern Europe in the 1960s (Losse, G.F., Nau, W., Winter, M., 1991. Le développement de la pêche en eau douce dans le nord de la Tunisie Projet de la coopération technique. Coopération technique Tuniso-Allemande, CGP/GTZ), we assume that *Ligula* clade A was either imported with the infected fish stock or immigrated later with the final bird host. Although both native and introduced fish species are nowadays common in the North African region, their *Ligula* fauna do not seem to undergo any substantial mixing.

We found that *D. interrupta* clustered within the *L. intestinalis* tree. This has already been reported by Logan et al. (2004) and Kuchta et al. (2007) who synonymized *Digramma* with *Ligula*. In this study, we encountered a conflict between the position of our *D. interrupta* samples isolated from the fish genus *Hemiculter* in Russia and China (haplotypes 9 and 43) and the Chinese samples collected and characterized previously (haplotypes 12 and 13) by Li and Liao (2003). This conflict cannot be unequivocally explained; however several facts indicate that the molecular data from the study of Li and Liao (2003) should be treated with caution. Analyzing COI sequences from several Chinese samples of *Digramma* and *Ligula*, these authors found extremely low variability, represented by

only two haplotypes shared by both tapeworm species. Such genetic homogeneity contradicts not only the high genetic variation found in our COI set but also their own results obtained with a different mitochondrial gene; when using one sequence of NADH deshydrogenase subunit I, they detected clear differences between *Ligula* and *Digramma* samples. Since the authors worked with formalin-preserved tapeworms the contamination originated during the process of DNA isolation might be an explanation. Still less clear is the taxonomic status of another species, *L. colymbi*. In our sample set, we found the suggested morphological characteristics unreliable and difficult to assess. Since the great crested grebe, *P. cristatus*, known to harbor both *L. intestinalis* and *L. colymbi*, is often considered a typical final host of the latter species (Dubinina, 1980), we paid special attention to the phylogenetic relationships of the adults isolated from *P. cristatus*. The tapeworms obtained from a single locality clustered among typical *L. intestinalis* samples within the clades A (five samples; haplotypes 14, 23, 24) and B (two samples; haplotypes 25, 26). This finding shows that species identification based on the final host only may be unreliable.

ITS has been the most frequently used marker in previous phylogenetic studies in the genus *Ligula* and in diphyllobothriid tapeworms in general (Olson et al., 2002; Li and Liao, 2003; Luo et al., 2003; Logan et al., 2004), thus knowledge of the occurrence of paralogues and the degree of intragenomic variability should be an important basis for evolutionary interpretations in such studies. In our study, we confirmed the presence of numerous ITS2 copies within a single genome and demonstrated that in some cases the copies may differ considerably in their phylogenetic position. In spite of this potentially serious disruption of the phylogenetic signal, the COI and ITS2 topologies displayed an unexpectedly high degree of congruence. Almost all of the clades revealed by COI were mirrored by corresponding monophyletic ITS2 clades (Fig. 2). The only exception was clade A, where the difference between COI and ITS2 was due to different levels of resolution rather than contradicting signals. We assume that the occurrence of divergent copies within a single genome is relatively rare and does not affect the overall topology in population studies where the sample is sufficiently large. It may, however, cause serious distortions in phylogenetic studies where few samples are collected and each is represented by a single ITS2 copy.

As many hundreds or even thousands of ITS2 copies are present in a single genome (e.g. Buckler-Iv et al., 1997; Lewis and Doyle, 2002; Pfeil et al., 2004; Bayly and Ladiges, 2007; Ganley and Kobayashi, 2007), it is unfeasible to thoroughly analyse the complete intragenomic variability. It is therefore impossible to distinguish the portion of the heterogeneity that is due to standard long-term variability within a single haploid genome, and how much of the variability should be attributed to occasional crossing among the clades. Considering the overall fit between the COI and ITS2 phylogenies, we assume that occasional crossings rather than a high degree of ancestral polymorphism are responsible for the presence of highly diverged forms within a single tapeworm. In this respect, it is interesting to note that no such group of highly divergent copies was observed in *D. interrupta*. This seems to support the status of *D. interrupta* as an isolated cryptic species which does not cross-breed with other lineages/species of the complex. Similarly, the pattern shown in Fig. 2 (i.e. the overall correspondence between COI and ITS2 with occasional "switches") indicates several well-formed and mostly isolated lineages/species of *Ligula* with rare crossings.

Besides their different host ranges, the two Euro-Mediterranean clades A and B differ in the parameters of their population genetics. Although statistics may be affected by the low number of samples in clade B, the tests of neutrality (Table 5), mismatch distribution (Fig. 3) and coalescence analysis (Table 6) provided identical patterns. Population sizes of the two main lineages were shown to differ: clade A was characterized as fast-expanding and having a wider distribution than clade B. This is in concordance with the host distribution of the clades: while clade A infects widespread fish taxa with high population densities, the European lineage of clade B was sampled exclusively from less common host taxa (information obtained from www.fishbase.org). As discussed above, the distribution of tapeworms from different fish species within the inner topology of clade A did not suggest any obvious co-speciation between the parasite and its intermediate host. Rather, the different tapeworm frequencies found in different fish species indicate that a certain degree of specificity might be expressed in the differential prevalence of various genotypes. Loot et al. (2001) described preferential infection of *Rutilus* by *L. intestinalis* at several localities in France. Even though both roach and bream are abundant species in France, among the 1385 specimens of bream examined only two were infected (0.2%), whereas the prevalence of *Ligula* in roach at the same localities was 40% (G. Loot, unpublished data). Similarly, Griffiths and Bigsby (unpublished data ex Olson et al., 2002) did not find any *Ligula* infections in *Abramis* in Lough Neagh. By contrast, *L. intestinalis* infections of *Abramis* are very common in middle and eastern European areas (e.g. Dubinina, 1980; Barus and Prokes, 1994). In the future, additional studies using fine-scale molecular markers (microsatellite sequence) and experimental cross-infection are needed to explore patterns of local parasite adaptation (see Greischar and Koskella, 2007) in clade A.

In the same way, no effect of the intermediate fish host on the tapeworm population structure was found using the haplotype network as a basis for NCPA. In this analysis, the flat population structure with respect to geography was mainly attributed to contiguous range expansion. Of the two equally probable hypotheses suggested for clade 3-3, we consider the range expansion with long distance dispersal more plausible since subclade 2-9 is formed by Tunisian samples. Indeed, these tapeworms are thought to have been recently imported into Tunisia and to have

become genetically isolated from continental Europe by the Mediterranean Sea (W. Bouzid et al., unpublished data).

Within clade A, genealogical discordance between parasite and host phylogenies suggests extensive gene flow among parasites across the host species spectra. Among the factors reducing genetic variation within populations of parasites, an important role is usually attributed to the dissemination of eggs by highly mobile definitive hosts (Kennedy, 1998; Gittenberger et al., 2006). It has recently been demonstrated in a system of cladoceran crustaceans that waterfowl migrations can affect the genetic structure of aquatic invertebrates (Figuerola et al., 2005). The birds that act as potential definitive hosts of *Ligula* in Europe are mostly migrating and wintering species (Curry-Lindahl, 1980). For example, the Black-headed gull, *Larus ridibundus*, and the Herring gull, *Larus argentatus*, have a widespread dispersion and display wintering erraticism over central and southern Europe (Curry-Lindahl, 1980). Some populations even migrate to North African coasts and can reach as far as the African tropics during the winter (Curry-Lindahl, 1980). The Great Crested Grebe *P. cristatus*, from which several mature tapeworms were analyzed in this study, winters in several European countries and spends summer on Atlantic and North Sea coasts. Some of them migrate to North Africa and join the sedentary populations of the same species. Similar migration behaviour can be observed in many other definitive host species. Even though the duration of *Ligula* infection in a bird host is very short, usually lasting only 2–5 days (Dubinina, 1980), birds are probably able to disseminate the eggs over long distances within a short period of time during seasonal migration (Wyatt and Kennedy, 1989).

The origin and evolution of host specificity in parasites is one of the perpetual questions in parasitology. In host-parasite associations with strong co-evolutionary signals, co-speciation events are usually considered as the mechanisms behind host specificity. However, in parasites with free-living stages, mobile hosts and thus the capacity for population mixing, local adaptation rather than co-speciation processes might be a more important source of specificity (Gandon and Michalakis, 2002; Greischar and Koskella, 2007). Since most of the previous analyses on tapeworm phylogeny were performed at higher phylogenetic levels, it is difficult to make any strong generalization by comparing our results with other published analyses. It is, however, interesting to note that various phylogenetic studies on tapeworms indicate little sign of co-evolution or even completely uncoupled phylogenies of the parasites and their hosts (Caira and Jensen, 2001; Škeriková et al., 2001; Wickström et al., 2003). The phylogenetic patterns of the *Ligula* samples discussed above might provide an interesting insight into the origin of host specificity in this type of parasite. At low phylogenetic levels (i.e. within the clades) the homogenization guaranteed by mobile definitive hosts results in remarkably shallow population structures and prevents co-speciation between host and parasite. In genetically distinct lineages or cryptic species (clades A and B), we suggest a scenario of co-evolution between the tapeworm *Ligula* and its host range. In sufficiently divergent cryptic species, the different host specificities naturally persist even when the parasite populations meet again and coexist in sympatry. Finally, at high phylogenetic levels, geographically isolated lineages diverge due to physical isolation and are thus able to adapt to local host fauna.

Acknowledgements

The following persons kindly provided samples of *Ligula* tapeworms or helped considerably in their sampling: J. Sitko, Czech Republic; C. Arme, United Kingdom; A. Aydogdu, Turkey; T. Boutorina, Russia; J.M. Caspeta-Mandujano, Mexico; A. de Chambrier, Switzerland; M. Davidová, Czech Republic; M. Borga Ergonul, Turkey; A. Kostadinova, Bulgaria; R. Kuchta, Czech Republic; S. Frank, Germany; A. Abdessalem, Algeria; A. Lymbery, Australia; P. Nie, China; A. Kangur and K. Kangur, Estonia. Samples of *Ligula* tapeworms from Lake Tana could not be obtained without the help of Abebe Getahun, Seyoum Mengistou, Eshete Dejen Dresilign, Moges Beleteu, and the staff of the ANRS Agricultural Research Institute, Bahir Dar, Ethiopia. The study was supported by Grants LC06073 and MSM 60076605801 (Ministry of Education, Czech Republic), Barrande project (Egide, France), Grant Agency of the Czech Republic (projects Nos. 206/08/1019 and 524/04/0342) and Institute of Parasitology (research projects Z60220518 and LC 522).

References

Anderson, T.J.C., 2001. The dangers of using single locus markers in parasite epidemiology: ascaris as a case study. Trends Parasitol. 17, 183–188.

Arme, C., 1997. Ligulosis in two cyprinid hosts: *Rutilus rutilus* and *Gobio gobio*. Helminthologia 34, 191–196.

Arme, C., Owen, R.W., 1968. Occurrence and pathology of *Ligula intestinalis* infections in British fishes. J. Parasitol. 54, 272–280.

Barus, V., Prokes, M., 1994. Parasite load of *Ligula intestinalis* plerocercoids in adult silver bream, *Blicca bjoerkna*. Helminthologia 31, 91–94.

Barus, V., Prokes, M., 2002. Length and weight of *Ligula intestinalis* plerocercoids (Cestoda) parasitizing adult cyprinid fishes (Cyprinidae): a comparative analysis. Helminthologia 39, 29–34.

Bayly, M.J., Ladiges, P.Y., 2007. Divergent paralogues of ribosomal DNA in eucalypts (Myrtaceae). Mol. Phylogenet. Evol. 44, 346–356.

Bean, C.C., Winfield, I.J., 1992. Influences of the tapeworm *Ligula intestinalis* (L.) on the spatial distributions of juvenile roach *Rutilus rutilus* (L.) and gudgeon *Gobio gobio* (L.) in Lough Neagh, Northern Ireland. Neth. J. Zool. 42, 416–429.

Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1998. Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. Mol. Biol. Evol. 15, 1719–1727.

Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1995. Host movement and the genetic structure of populations of parasitic nematodes. Genetics 141, 1007–1014.

Brant, S.V., Ortí, G., 2003. Evidence for gene flow in parasitic nematodes between two host species of shrews. Mol. Ecol. 12, 2853–2859.

Briolay, J., Galtier, N., Brito, R.M., Bouvet, Y., 1998. Molecular phylogeny of cyprinidae inferred from cytochrome b DNA sequences. Mol. Phylogenet. Evol. 9, 100–108.

Buckler-Iv, E.S., Ippolito, A., Holtsford, T.P., 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145, 821–832.

Caira, J.N., Jensen, K., 2001. An investigation of the co-evolutionary relationships between onchobothriid tapeworms and their elasmobranch hosts. Int. J. Parasitol. 31, 960–975.

Carter, V., Pierce, R., Dufour, S., Arme, C., Hoole, D., 2005. The tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea) inhibits LH expression and puberty in its teleost host, *Rutilus rutilus*. Reproduction 130, 939–945.

Chapman, A., Hobbs, R.P., Morgan, D.L., Gill, H.S., 2006. Helminth parasitism of *Galaxias maculatus* (Jenyns 1842) in southwestern Australia. Ecol. Freshw. Fish 15, 559–564.

Criscione, C.D., Blouin, M.S., 2004. Life cycles shape parasite evolution: comparative population genetics of salmon trematodes. Evolution 58, 198–202.

Criscione, C.D., Poulin, R., Blouin, M.S., 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Mol. Ecol. 14, 2247–2257.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1660.

Cunha, C., Mesquita, N., Dowling, T.E., Gilles, A., Coelho, M.M., 2002. Phylogenetic relationships of Eurasian and American cyprinids using cytochrome b sequences. J. Fish. Biol. 61, 929–944.

Curry-Lindahl, K., 1980. Les oiseaux migrateurs à travers mer et terre. Delachaux et Niestle, Neuchatel, Paris.

Dejen, E., Vijverberg, J., Sibbing, F.A., 2006. Spatial and temporal variation of cestode infection and its effects on two small barbs (*Barbus humilis* and *B. tanapelagius*) in Lake Tana, Ethiopia. Hydrobiologia 556, 109–117.

Dubinina, M.N., 1980. Tapeworms (Cestoda, Ligulidae) of the Fauna of the USSR. Amerind Publishing, New Delhi, India.

Durand, J.D., Tsigenopoulos, C.S., Unlu, E., Berrebi, P., 2002. Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA-evolutionary significance of this region. Mol. Phylogenet. Evol. 22, 91–100.

Figuerola, J., Green, A.J., Michot, T.C., 2005. Invertebrate eggs can fly: evidence of waterfowl-mediated gene flow in aquatic invertebrates. Am. Nat. 165, 274–280.

Gandon, S., Michalakis, Y., 2002. Local adaptation, evolutionary potential and host-parasite coevolution: interactions between migration, mutation, population size and generation time. J. Evol. Biol. 15, 451–462.

Ganley, ARD., Kobayashi, T., 2007. Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res. 17, 184–191.

Gilles, A., Lecointre, G., Miquelis, A., Loerstcher, M., Chappaz, R.M., Brun, G., 2001. Partial combination applied to phylogeny of European cyprinids using the mitochondrial control region. Mol. Phylogenet. Evol. 19, 22–33.

Giraud, T., 2006. Speciation in parasites: host switching does not automatically lead to allopatry. Trends Parasitol. 22, 151–152.

Giribet, G., 2005. TNT: tree analysis using new technology. Syst. Biol. 54, 176–178.

Gittenberger, E., Groenenberg, D.S.J., Kokshoorn, B., Preece, R.C., 2006. Molecular trails from hitch-hiking snails. Nature 439, 409.

Greischar, M.A., Koskella, B., 2007. A synthesis of experimental work on parasite local adaptation. Ecol. Lett. 10, 418–434.

Groves, K.L., Shields, B.A., 2001. Observations on the plerocercoid stage of the tapeworm *Ligula* in three species of fish from the lower crooked river of central Oregon. J. Aquat. Anim. Health 13, 285–289.

Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online – a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 33, W557–W559.

Hafner, M.S., Nadler, S.A., 1990. Cospeciation in host-parasite assemblages: comparative analysis of rates of evolution and timing of cospeciation events. Syst. Zool. 39, 192–204.

Hafner, M.S., Page, R.D.M., 1995. Molecular phylogenies and host-parasite cospeciation: Gophers and Lice as a model system. Philos. T. Roy. Soc. B 349, 77–83.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Haukisalmi, V., Hardman, L.M., Hardman, M., Laakkonen, J., Niemimaa, J., Henttonen, H., 2007. Morphological and molecular characterisation of *Paranoplocephala buryatiensis n.* sp. and *P. longivaginata* Chechulin & Gulyaev, 1998 (Cestoda: Anoplocephalidae) in voles of the genus *Clethrionomys*. Systematic Parasitol. 66, 55–71.

Hoberg, E.P., Brooks, D.R., Siegel-Causey, D., 1997. Host-parasite cospeciation: history, principles, and prospects. In: Clayton, D., Moore, J. (Eds.), Host-Parasite Evolution: General Principles and Avian Models. Oxford University Press, Oxford, pp. 212–235.

Hughes, J., 2007. Multiple cophylogenetic analyses reveal frequent cospeciation between pelecaniform birds and *Pectinopygus* Lice. Syst. Biol. 56, 232–251.

Hypša, V., Škeříková, A., Scholz, T., 2005. Phylogeny, evolution and host-parasite relationships of the order Proteocephalidea (Eucestoda) as revealed by combined analysis and secondary structure characters. Parasitology 130, 359–371.

Jarne, P., Théron, A., 2003. Genetic structure in natural populations of flukes and snails: a practical approach and review. Parasitology 123, 27–40.

Johnson, K.P., Williams, B.L., Drown, D.M., Adams, R.J., Clayton, D.H., 2002. The population genetics of host specificity: genetic differentiation in dove lice (Insecta: Phthiraptera). Mol. Ecol. 11, 25–38.

Jousson, O., Bartoli, P., Pawlowski, J., 2000. Cryptic speciation among intestinal parasites (Trematoda: Digenea) infecting sympatric host fishes (Sparidae). J. Evol. Biol. 13, 778–785.

Kennedy, C.R., 1998. Aquatic birds as agents of parasite dispersal: a field test of the effectiveness of helminth colonisation strategies. Bull. Scand. Soc. Parasitol. 8, 23–28.

Kennedy, C.R., Burrough, R.J., 1981. The establishment and subsequent history of a population of *Ligula intestinalis* in roach *Rutilus rutilus* (L.). J. Fish Biol. 19, 105–126.

Kuchta, R., Scholz, T., Brabec, J., Bray, R.A., 2007. Suppression of the tapeworm order pseudophyillidea (Platyhelminthes: Eucestoda) and the proposal of two new orders, Bothricephalidea and Diphyllobothriidea. Int. J. Parasitol. 38, 49–55.

Kuhner, M.K., 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22, 768–770.

Lewis, C.E., Doyle, J.J., 2002. A phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes. Plant Syst. Evol. 236, 1–17.

Li, J., Liao, X., 2003. The taxonomic status of *Digramma* (Pseudophyllidae: Ligulidae) inferred from DNA sequences. J. Parasitol. 89, 792–799.

Li, J., Liao, X., Yang, H., 2000. Molecular characterization of a parasitic tapeworm (*Ligula*) based on DNA sequences from formalin-fixed specimens. Biochem. Genet. 38, 309–322.

Liu, H.Z., Chen, Y.J., 2003. Phylogeny of the East Asian cyprinids inferred from sequences of the mitochondrial DNA control region. Can. J. Zool. 81, 1938–1946.

Logan, F.J., Horák, A., Štefka, J., Aydogdu, A., Scholz, T., 2004. The phylogeny of diphyllobothriid tapeworms (Cestoda: Pseudophyllidea) based on ITS-2 rDNA sequences. Parasitol. Res. 94, 10–15.

Loot, G., Aulagnier, S., Lek, S., Thomas, F., Guegan, J.F., 2002. Experimental demonstration of a behavioural modification in a cyprinid fish, *Rutilus rutilus* (L.), induced by a parasite, *Ligula intestinalis* (L.). Can. J. Zool. 80, 738–744.

Loot, G., Francisco, P., Santoul, F., Lek, S., Guegan, J.F., 2001. The three hosts of the *Ligula intestinalis* (Cestoda) life cycle in Lavernose-Lacasse gravel pit, France. Arch. Hydrobiol. 152, 511–525.

Loot, G., Park, Y.S., Lek, S., Brosse, S., 2006. Encounter rate between local populations shapes host selection in complex parasite life cycle. Biol. J. Linn. Soc. 89, 99–106.

Luo, H.Y., Nie, P., Yao, W.J., Wang, G.T., Gao, Q., 2003. Is the genus *Digramma* synonymous to the genus *Ligula* (Cestoda: Pseudophyllidea)? Parasitol. Res. 89, 419–421.

McCoy, K.D., Boulinier, T., Tirard, C., Michalakis, Y., 2001. Host specificity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick *Ixodes uriae*. J. Evol. Biol. 14, 395–405.

McCoy, K.D., 2003. Sympatric speciation in parasites – what is sympatry? Trends Parasitol. 19, 400–404.

McDowall, R.M., 1990. New Zealand Freshwater Fishes: A Natural History and Guide. Heinemann Reed, Auckland, New Zealand.

McManus, D.P., 1985. Enzyme analyses of natural populations of *Schistocephalus solidus* and *Ligula intestinalis*. J. Helminthol. 59, 323–332.

Morgan, D.L., 2003. Distribution and biology of *Galaxias truttaceus* (Galaxiidae) in south-western Australia, including first evidence of parasitism of fishes in Western Australia by *Ligula intestinalis* (Cestoda). Environ. Biol. Fish 66, 155–167.

Museth, J., 2001. Effects of *Ligula intestinalis* on habitat use, predation risk and catchability in European minnows. J. Fish Biol. 59, 1070–1080.

Olson, P.D., Littlewood, D.T.J., Griffiths, D., Kennedy, C.R., Arme, C., 2002. Evidence for the co-existence of separate strains or species of *Ligula* in Lough Neagh, Northern Ireland. J. Helminthol. 76, 171–174.

Page, R.D.M., Lee, P.L.M., Becher, S.A., Griffiths, R., Clayton, D.H., 1998. A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. Mol. Phylogenet. Evol. 9, 276–293.

Panchal, M., 2007. The automation of nested clade phylogeographic analysis. Bioinformatics 23, 509–510.

Pfeil, B.E., Brubaker, C.L., Craven, L.A., Crisp, M.D., 2004. Paralogy and orthology in the Malvaceae rpb2 gene family: investigation of gene duplication in *Hibiscus*. Mol. Biol. Evol. 21, 1428–1437.

Pollard, D.A., 1974. The biology of a landlocked form of the normally catadromous salmoniform fish *Galaxias maculatus* (Jenyns) VI. Effects of cestode and nematide parasites. Aust. J. Mar. Freshwat. Res. 25, 105–120.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Posada, D., Crandall, K.A., Templeton, A.R., 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Mol. Ecol. 9, 487–488.

Poulin, R., 1998. Evolutionary Ecology of Parasites: From Individuals to Communities. Chapman and Hall, London.

Prugnolle, F., Liu, H., de Meeûs, T., Balloux, F., 2005. Population genetics of complex life-cycle parasites: an illustration with trematodes. Int. J. Parasitol. 35, 255–263.

Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Saitoh, K., Sado, T., Mayden, R.L., Hanzawa, N., Nakamura, K., Nishida, M., Miya, M., 2006. Mitogenomic evolution and interrelationships of the cypriniformes (Actinopterygii : Ostariophysi): the first evidence toward resolution of higher-level relationships of the world's largest freshwater fish clade based on 59 whole mitogenome sequences. J. Mol. Evol. 63, 826–841.

Škeríková, A., Hypša, V., Scholz, T., 2001. Phylogenetic analysis of European species of *Proteocephalus* (Cestoda: Proteocephalidea): compatibility of molecular and morphological data, and parasite-host coevolution. Int. J. Parasitol. 31, 1121–1128.

Sweeting, R.A., 1977. Studies on *Ligula intestinalis*. Some aspects of the pathology in the second intermediate host. J. Fish Biol. 10, 43–50.

Taylor, M., Hoole, D., 1989. *Ligula intestinalis* (L.) (Cestoda: Pseudophyllidea): an ultrastructural study of the cellular response of roach fry, *Rutilus rutilus* L., to an unusual intramuscular infection. J. Fish Dis. 12, 523–528.

Templeton, A.R., 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. Mol. Ecol. 13, 789–809.

Théron, A., Combes, C., 1995. Asynchrony of infection timing, habitat preference, and sympatric speciation of schistosome parasites. Evolution 49, 372–375.

von Nickisch-Rosenegk, M., Brown, W.M., Boore, J.L., 2001. Complete sequence of the mitochondrial genome of the tapeworm hymenolepis diminuta: gene arrangements indicate that platyhelminths are eutrochozoans. Mol. Biol. Evol. 18, 721–730.

Weekes, P.J., Penlington, B., 1986. First records of *Ligula intestinalis* (Cestoda) in rainbow trout, *Salmo gairdneri*, and common bully *Gobiomorphus cotidianus* in New Zealand. J. Fish Biol. 28, 183–190.

Wickström, L.M., Haukisalmi, V., Varis, S., Hantula, J., Fedorov, V.B., Henttonen, H., 2003. Phylogeography of the circumpolar *Paranoplocephala arctica* species complex (Cestoda: Anoplocephalidae) parasitizing collared lemmings (*Dicrostonyx* spp.). Mol. Ecol. 12, 3359–3371.

Wyatt, R.J., Kennedy, C.R., 1989. Host-constrained epidemiology of the fish tapeworm *Ligula intestinalis* (L.). J. Fish Biol. 35, 215–227.

Zardoya, R., Doadrio, I., 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. Proc. Biol. Sci. 265, 1365–1372.

Zardoya, R., Doadrio, I., 1999. Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. J. Mol. Evol. 49, 227–237.