Species delimitation in trematodes using DNA sequences: Middle-American *Clinostomum* as a case study

GERARDO PÉREZ-PONCE DE LEÓN[1]*, MARTÍN GARCÍA-VARELA[1], CARLOS D. PINACHO-PINACHO[1,2], ANA L. SERENO-URIBE[1] *and* ROBERT POULIN[3]

[1] *Departamento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Ap. Postal 70-153, México d.f., C.P. 04510, Mexico*
[2] *Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, México City, Mexico*
[3] *Department of Zoology, University of Otago, PO Box 56, Dunedin, New Zealand*

(*Received 20 April 2016; revised 19 July 2016; accepted 21 July 2016; first published online 30 August 2016*)

SUMMARY

The recent development of genetic methods allows the delineation of species boundaries, especially in organisms where morphological characters are not reliable to differentiate species. However, few empirical studies have used these tools to delineate species among parasitic metazoans. Here we investigate the species boundaries of *Clinostomum*, a cosmopolitan trematode genus with complex life cycle. We sequenced a mitochondrial [cytochrome *c* oxidase subunit I (COI)] gene for multiple individuals (adults and metacercariae) from Middle-America. Bayesian phylogenetic analysis of the COI uncovered five reciprocally monophyletic clades. COI sequences were then explored using the Automatic Barcode Gap Discovery to identify putative species; this species delimitation method recognized six species. A subsample was sequenced for a nuclear gene (ITS1, 5·8S, ITS2), and a concatenated phylogenetic analysis was performed through Bayesian inference. The species delimitation of Middle-American *Clinostomum* was finally validated using a multispecies coalescent analysis (species tree). In total, five putative species are recognized among our samples. Mapping the second intermediate hosts (fish) onto the species tree suggests that metacercariae of these five species exhibit some level of host specificity towards their fish intermediate host (at the family level), irrespective of geographical distribution.

Key words: Trematoda, host specificity, genetic diversity, species delimitation, species tree, biodiversity.

INTRODUCTION

Species are fundamental units of biological studies, and still no uniform guidelines exist for determining species boundaries in an objective manner. Although morphology has been commonly used to delineate species, the development of genetic tools has allowed researchers to use such data to infer species limits, where other lines of evidence (morphology in particular) may underestimate or overestimate species diversity. The advent of molecular tools has offered an unprecedented opportunity within parasitology to add new components to our discovery and description of parasite biodiversity, including the potential recognition of cryptic species (Pérez-Ponce de León and Nadler, 2010; Nadler and Pérez-Ponce de León, 2011; Poulin, 2011). This possibility challenges our capacity to establish reliable estimates of parasite diversity (Poulin, 2014). DNA sequences of trematodes, where more cryptic species are found than in any other helminth taxa for a given sampling effort (Poulin, 2011), have accumulated rapidly in the last decade, and even though a large proportion of taxonomic papers on trematodes do not use genetic data, some authors have recently advocated the need to generate sequence data from as many host species/parasite species/geographic location combinations as possible (Blasco-Costa *et al.* 2016*a*). In particular, a large genomic library of mitochondrial and nuclear genes has been developed for some trematodes whose metacercariae are commonly found in fish; that includes members of the Clinostomidae and Diplostomatidae, where molecular data have proven very useful in evaluating the taxonomic status of parasites, detecting cryptic species and linking larval forms in intermediate hosts (fish) with adults in their definitive hosts (birds) (e.g. Caffara *et al.* 2011; Locke *et al.* 2011, 2015*a*, *b*; Chibwana *et al.* 2013, 2015; Georgieva *et al.* 2013; Sereno-Uribe *et al.* 2013; Blasco-Costa *et al.* 2014, 2016*b*; García-Varela *et al.* 2016*a*, *b*).

The cosmopolitan trematode, genus *Clinostomum* comprises species whose adult forms are usually found in the mouth cavity and oesophagus of 12 bird families distributed worldwide (Matthews and Cribb, 1998; Caffara *et al.* 2011; Locke *et al.* 2015*a*). Matthews and Cribb (1998) pointed out that assessing the validity of species described within *Clinostomum* has been made difficult by the fact that many species were described inadequately, either solely from metacercariae, or even one from a cercaria, and not from gravid adults. In recent years, molecular methods have been used to obtain

* Corresponding author: Departamento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de México, México. E-mail: ppdleon@ib.unam.mx

Parasitology (2016), **143**, 1773–1789.
doi:10.1017/S0031182016001517

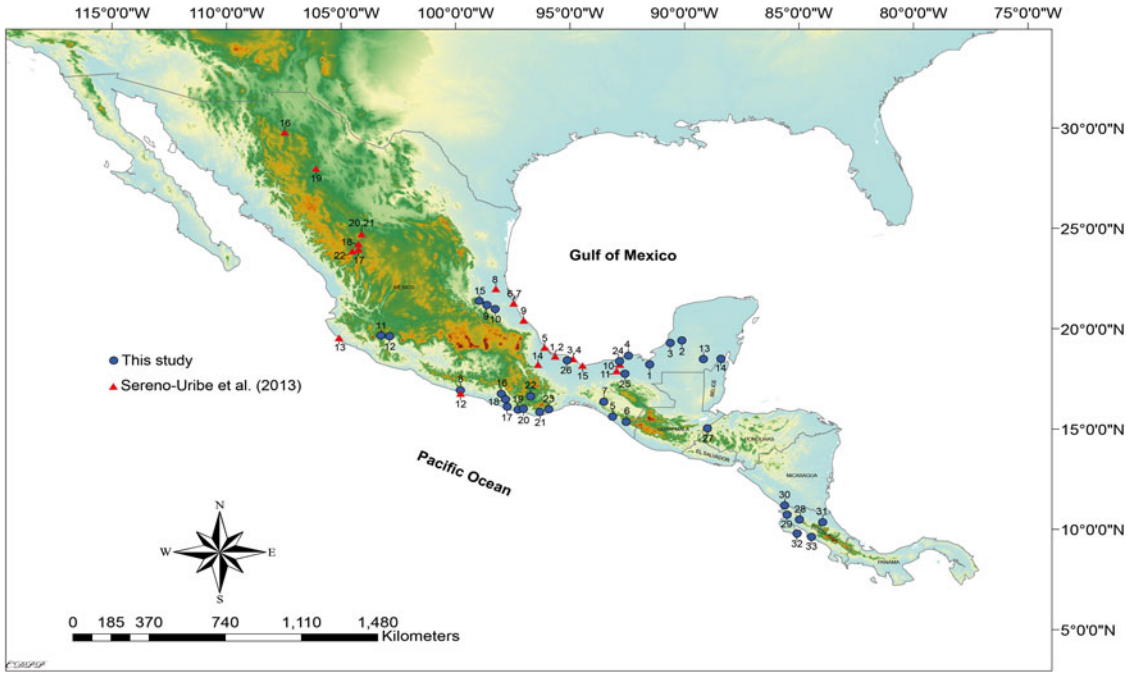

Fig. 1. Map of Middle-America showing the 33 sampling localities of *Clinostomum* spp. across Central and Southern Mexico, Honduras and Costa Rica. Also shown are the 22 localities sampled by Sereno-Uribe *et al*. (2013).

information that, in combination with morphology, has been very useful to distinguish among species of *Clinostomum* (e.g. Dzikowski *et al*. 2004; Gustinelli *et al*. 2010; Caffara *et al*. 2011, 2013, 2014; Sereno-Uribe *et al*. 2013; Athokpam *et al*. 2014; Senapin *et al*. 2014; Locke *et al*. 2015*a*; Pinto *et al*. 2015). A large database of mitochondrial and nuclear DNA sequences of *Clinostomum* individuals is now available, improving our ability to produce a more reliable estimate of species diversity in the genus. Locke *et al*. (2015*a*) conducted a large-scale molecular survey of *Clinostomum* metacercariae with sequences obtained from the DNA barcode region of cytochrome *c* oxidase subunit I (COI), and the internal transcribed spacers (ITS1 and ITS2), from specimens collected in different parts of the world. Eight candidate species that remain to be described were recognized with varying degrees of confidence as groups delineated by the Automatic Barcode Gap Discovery (ABGD) and the algorithm of Ratnasingham and Hebert (2013) which is an online resource only implemented for COI sequences, with ITS and other data considered as supporting evidence. Seven of the eight species occur across the Americas. Specimens from Middle-America were not sampled extensively by Locke *et al*. (2015*a*), however, and this is an area where we have conducted intensive survey work recent years, allowing us to complement their large-scale survey with numerous sequences from a relatively restricted geographic range.

We are engaged in a two-part study that aims to provide empirical evidence to address Matthews and Cribb's (1998) suggestion that putative species of *Clinostomum* should be subjected to a molecular analysis to determine species boundaries, and the extent to which morphological characters are reliable. Our first step, reported here, is to thoroughly characterize the genetic diversity of Middle-American *Clinostomum* using a dataset of two unlinked loci (mitochondrial and nuclear). Our goals are to: (1) establish a primary species delimitation hypothesis through DNA sequence analysis following a unified species concept (de Queiroz, 2007); (2) explore the primary species hypothesis through a species delimitation method such as ABGD, and validate it through a coalescent species tree analysis; and (3) show potential patterns of host specificity of the metacercariae of Middle-American *Clinostomum*, and discuss the importance of accurate species delimitation for assessments of host specificity. Our study of *Clinostomum* provides a case study with broad applications to other taxa.

MATERIALS AND METHODS

Samples collected

Individuals of *Clinostomum* were collected from February 2013 through February 2015 in their definitive (fish-eating birds), and second intermediate hosts (fish), in 26 localities across Southern Mexico, one locality in Honduras, and six localities in Costa Rica (Fig. 1). Information regarding host species, localities, geographical coordinates and GenBank accession numbers are provided in Table 1. Birds were captured with a shotgun and immediately kept in ice. Fish were captured with seine nets and electrofishing, transported to the

http://dx.doi.org/10.1017/S0031182016001517

laboratory, sacrificed with an overdose of anaesthetic (sodium pentobarbital) and immediately examined. Hosts were necropsied and all internal organs were examined for parasites under a dissecting microscope a few hours after their capture. Collected trematodes were preserved in 100% ethanol for DNA extraction and sequencing (Table 1).

Extraction, amplification and sequencing of DNA

Specimens were placed individually in tubes and digested overnight at 56 °C in a solution containing 10 mM Tris–HCl (pH 7·6), 20 mM NaCl, 100 mM Na_2 EDTA (pH 8·0), 1% Sarkosyl and 0·1 mg mL^{-1} proteinase K. Following digestion, DNA was extracted from the supernatant using the DNAzol reagent (Molecular Research Center, Cincinnati, Ohio) according to the manufacturer's instructions. The COI (~474 bp) and the ITS1–5·8S–ITS2 (ITS) (~1200 bp) were amplified using the polymerase chain reaction (PCR). A fragment of COI was amplified using primers modified from Moszczynska *et al*. (2009), with the forward primer 527F 5′–ATTCG(R)TTAAAT(Y)TKTGTGA–3′ and the reverse primer 528R 5′–CCAAAC(Y)AACAC(M)GACAT–3′ (Sereno-Uribe *et al*. 2013). The ITS was amplified using the forward primer BD1 5′–GTCGTAACAAGGTTTCCGTA–3′ and the reverse primer BD2 5′–ATCTAGACC GGACTAGGCTGTG–3′ (Luton *et al*. 1992). PCR reactions and cycling conditions followed those described in Sereno-Uribe *et al*. (2013). Sequencing reactions were performed using ABI Big Dye (Applied Biosystems, Boston, Massachusetts) terminator sequencing chemistry, and reaction products were separated and detected using an ABI 3730 capillary DNA sequencer. Contigs were assembled and base-calling differences resolved using Codoncode Aligner version 3.5·4 (Codoncode Corporation, Dedham, Massachusetts). Sequences of both molecular markers were deposited in the GenBank database (Table 1).

COI phylogenetic analyses

Sequences obtained in the current research were aligned with other sequences available in GenBank, i.e. *C. marginatum* (COI: JF718596–JF718619, HQ439564–HQ439577, HQ439579–HQ439586, JX630991–JX630997), *C. complanatum* (COI: JF718584, JF718588–JF718595), *C. cutaneum* (KP110515–KP110516), *C. phalacrocoracis* (COI: KJ786967–KJ786974), *C. attenuatum* (KP150305–06), *C. philippinense* (KP110523), *C. tataxumui* (COI: JX630998–JX631044) and *Clinostomum* sp. (COI: KJ818259–KJ818264) from experimental infections in *Poecilia reticulata* in Minas Gerais, Brazil (Pinto *et al*. 2015). Sequences of the eight putative species of *Clinostomum* from the study of Locke *et al*. (2015*a*) were also included: COI: KP110525–KP110543. Additionally, we added the sequence of *Euclinostomum* sp. as the only other member of the family Clinostomatidae for which sequences are available (COI: KC894795–KC894797) from three osphronemid fish, *Trichopsis vittata, T. schalleri* and *Betta imbellis*, in Thailand (Senapin *et al*. 2014) (Supplementary Table S1). This species, and the strigeids *Alaria mustelae* (JF904529) and *Diplostomum baeri* (GQ292501) were used as outgroups for rooting the trees. Alignment was constructed using the software Clustal W (Thompson *et al*. 1994) with default parameters and adjusted manually with the MacClade program (Maddison and Maddison, 2002). The COI dataset was analysed through Bayesian inference (BI). The best-fitting model (TPM1uf + I + G) was identified with the Akaike Information Criterion (AIC) implemented in jModelTest v0.1.1 (Posada, 2008). For the BI analyses, the implemented model was GTR + I + G, because the less complex TPM1uf + I + G models are not implemented in MrBayes. BI analysis was performed using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001), with two runs and four chains (one cold, three heated) per run. The Metropolis-coupled Markov chain Monte Carlo (MC^3) were run for 10 million generations, sampled every 1000 generations, and the first 2500 samples were discarded as burn-in (25%). The outputs of MrBayes were examined with Tracer v1·4 (Rambaut and Drummond, 2007) to check for convergence of different parameters, to determine the approximate number of generations at which log-likelihood values stabilized, to identify the effective sample size (EES > 200) for each parameter, and the estimated magnitude of model parameters in individual and combined runs. Topological convergence in the two independent MCMC runs was checked with the 'compare plot' in AWTY (Wilgenbusch *et al*. 2004). The initial 25% of MC^3s was verified to include all the generations before stationarity was achieved. Posterior probabilities (PPs) of clades were obtained from the 50% majority rule consensus of sampled trees after excluding the initial 25% as burn-in. Trees were visualized using FigTree program version 1.3.1 (Rambaut, 2006). Unique haplotypes of COI from all localities as well as unique haplotypes for sequences available from the GenBank dataset were identified using DnaSP v.5 (Librado and Rozas, 2009) and NETWORK v. 4.2.0.1 (Bandelt *et al*. 1999; www.fluxusengineering.com). The intra and interspecific genetic variations were determined using the Tamura-Nei distance (TrN) with the program MEGA v. 5 (Tamura *et al*. 2011).

Automatic Barcode Gap Discovery

We analysed the COI sequences using the ABGD to identify putative species, a method originally proposed by Puillandre *et al*. (2012). This method automatically finds the distance at which a barcode gap

http://dx.doi.org/10.1017/S0031182016001517

Table 1. Specimens of *Clinostomum* spp. collected in this study including collection sites (CS), locality (by state of the Mexican Republic), host species, geographical coordinates and GenBank accession number (with number of isolates sequenced per gene). The collection site number for each locality corresponds with the number in Figure 1

CS	Locality	Host	Family		Coordinates	CO1 (GenBank)	ITS (GenBank)	Taxa[b]
				N	W			
	MEXICO							
	Campeche State							
1	Laguna Silvictuc	*Ardea alba*[a]	Ardeidae	18°37′338″	90°16′57·1″	KJ477506-10 (5)	KJ477619-22 (4)	Lineage 5
		Cichlasoma urophthalmus	Cichlidae			KJ477511-15 (5)	KJ477623-27 (5)	Lineage 5
2	Santa Cruz	*Thorichthys meeki*	Cichlidae	19°24′24″	90°23′59″	KJ477531- 35 (5)	KJ477640-44 (5)	Lineage 5
		Cichlasoma urophthalmus	Cichlidae			KJ477526-30 (5)	KJ477635-39 (5) KJ477659 (1)	Lineage 5
		Thorichthys sp.	Cichlidae			KJ477557-58 (2)	KJ477660 (1)	Lineage 5
3	Río Champotón	*Tigrisoma mexicanum*[a]	Ardeidae	19°16′50″	90°36′51″	KJ477536-40 (5)	KJ477645-48 (4)	Lineage 5
		Cochlearius cochlearis[a]	Ardeidae			KJ477541-45 (5)	KJ477649-53 (5)	Lineage 5
4	Laguna El Milagro	*Cichlasoma urophthalmus*	Cichlidae	18°38′46″	92°26′16″	KJ477556 (1)		Lineage 5
		Rhamdia guatemalensis	Heptapteridae			KJ477546-47 (2)	KJ477654-55 (2)	Lineage 3
		Dormitator maculatus	Eleotridae			KJ477548-50 (3)	KJ477656-58 (3)	*C. tataxumui*
	Chiapas State							
5	Laguna Pijijiapan	*Tigrisoma mexicanum*[a]	Ardeidae	15°33′02″	93°15′53″	KU156776-778 (3)	KM676407-08 (2) KU156742 (1)	*C. tataxumui*
6	Río El Triunfo	*Profundulus punctatus*	Profundulidae	15°20′44″	92°32′30″	KU156779-781 (3)	KM676409-11 (3)	Lineage 4
7	Río San Juan	*Rhamdia guatemalensis*	Heptapteridae	16°21′00″	93°30′54″	KU156782-83 (2)	KU156743- 45 (3)	Lineage 1
	Guerrero State							
8	Laguna de Tres Palos	*Eleotris picta*	Eleotridae	16°48′00″	99°47′00″	KJ477551–55 (5)	KJ477661-62 (2)	*C. tataxumui*
	Hidalgo State							
9	San Felipe	*Herichthys cyanoguttatus*	Cichlidae	21°10′10″	98°35′48″	KJ477468-73 (6)	KJ477581-586 (6)	Lineage 5
10	Río Atlapexco	*Herichthys cyanoguttatus*	Cichlidae	21°0′57·6″	98°20′22″	KU156784-785 (2)	KU156746 (1)	Lineage 5
	Jalisco State							
11	Río La Rosa, S Vicente	*Astyanax aeneus*	Characidae	19°38′54″	103°15′23″	KU156786-88 (3)	KU156747-49 (3)	Lineage 2
12	Santa María del Oro	*Astyanax aeneus*	Characidae	29°35′49″	102°54′15″	KU156789-91 (3)	KU156750- 52 (3)	Lineage 2
	Quintana Roo State							
13	Laguna Milagros	*Vieja* sp.	Cichlidae	18°28′35″	89°09′52″	KJ477516-20 (5)	KJ477628-30 (3)	Lineage 5
14	Laguna Azul	*Cichlasoma urophthalmus*	Cichlidae	18°29′31″	88°23′48″	KJ477521-25 (5)	KJ477631-34 (4)	Lineage 5
	San Luís Potosí State							
15	Matlapa	*Herichthys cyanoguttatus*	Cichlidae	21°19′35″	98°49′16″	KJ477465-67 (3)	KJ477578-80 (3)	Lineage 5
	Oaxaca State							
16	El Platanar	*Profundulus balsanus*	Profundulidae	16°44′55″	97°59′32″	KJ504185-87 (3)	KJ477663-64 (2)	Lineage 4
17	Río Verde	*Cichlasoma trimaculatum*	Cichlidae	16°06′35″	97°43′58″	KJ504194-96 (3)	KJ477665-66 (2)	Lineage 5
		Tigrisoma mexicanum[a]	Ardeidae			KJ504197-201 (5)	KJ477667-68 (2)	Lineage 5
		Ardea alba[a]	Ardeidae			KJ504202 (1)	KJ477669 (1)	*C. tataxumui*
		Cochlearius cochlearis[a]	Ardeidae			KJ504203 (1)	KJ477670 (1)	Lineage 5
18	Flores Magón	*Cichlasoma trimaculatum*	Cichlidae	16°21′6·1″	97°45′38″	KJ504204-205 (2)	KJ477671 (1)	Lineage 5

19	Puente Manialtepec	*Dormitator latifrons*	Eleotridae	15°57′23″	97°14′90″	KJ504206-210 (5)	KJ477672-73 (2)	*C. tataxumui*
20	Laguna Manialtepec	*Tigrisoma mexicanum*[a]	Ardeidae	15°56′25″	97°11′40″	KJ504192-193 (2) KJ504211 (1)	KJ477674-75 (2)	*C. tataxumui*
21	Río Santa María Huatulco	*Profundulus punctatus*	Profundulidae	15°50′44″	96°18′57″	KJ504184 (1) KJ504188-91 (4)	KJ477676-77 (2)	Lineage 4
22	Presa Los Ocotes	*Ardea alba*[a]	Ardeidae	16°36′57″	96°43′13″	KJ504182-83 (2)	KJ477678-80 (3)	*C.marginatum*
		Pseudoxiphophorus jonesii	Poeciliidae			KJ477559–69 (10) KJ504181 (1)	KJ477587-96 (10) KJ477681 (1)	Lineage 4
23	Río Chacalapa	*Profundulus punctatus*	Profundulidae	15°55′46″	95°55′56″	KU156792-793 (2)	KU156753 (1)	Lineage 4
	Tabasco State							
24	El Espino	*Tigrisoma mexicanum*[a]	Ardeidae	18°14′47″	92°49′57″	KU156794-796 (3)	KU156754-757 (4)	Lineage 5
25	Emiliano Zapata	*Tigrisoma mexicanum*[a]	Ardeidae	17°44′44″	91°95′41″	KJ477498-500 (3)	KJ477612-613 (2)	Lineage 3
		Ardea herodias[a]	Ardeidae			KJ477501 (1)	KJ477614 (1)	Lineage 3
		Rhamdia guatemalensis	Heptapteridae			KJ477502-505 (4)	KJ477615-618 (4)	Lineage 3
	Veracruz State							
26	Lago de Catemaco	*Rhamdia guatemalensis*	Heptapteridae	18°25′	95°07′	KJ477443-450 (8)	KJ477571-577 (7)	Lineage 3
		Egretta thula[a]	Ardeidae			KU156797- 802 (6)	KU156758-759 (2)	Lineage 2 Lineage 4
		Ardea alba[a]	Ardeidae			KU156803 (1)	KU156760-761(2)	Lineage 2
	HONDURAS							
27	El Paraiso	*Rhamdia* sp.	Heptapteridae	15°01′26″	88°59′32″	KU156804-807 (4)	KU156763-767 (4)	Lineage 1 Lineage 3
	COSTA RICA							
28	Tilarán	*Rhamdia rogersi*	Heptapteridae	10°28′47″	84°58′15″	KU156808 (1)	KU156768-769 (2)	Lineage 3
29	Río Irigaray	*Parachromis managuensis*	Cichlidae	10°43′22″	85°30′37″	KU156809-811 (3)	KU156770 (1)	Lineage 5
		Amatitlania nigrofasciata	Cichlidae			KU156812-814 (3)	KU156771 (1)	Lineage 5
30	Río Las Vueltas	*Archocentrus siquia*	Cichlidae	11°10′56″	85°36′58″	KU156815-818 (4)	KU156772-773 (2)	Lineage 5
		Rhamdia rogersi	Heptapteridae			KU156819-820 (2)	KU156774-775 (2)	Lineage 3
31	Horquetas de Sarapiqui	*Amphilopus* sp.	Cichlidae	10°20′36″	83°57′42″	KJ477490-94 (5)	KJ477608-609 (2)	Lineage 5
		Amatitlania sp.	Cichlidae			KJ477495-97 (3)	KJ477610-611 (2)	Lineage 5
32	Río Grande	*Gobiomorus maculatus*	Eleotridae	9°51′45·4″	84°56′20″	KJ477474-80 (7)	KJ477597-601 (6)	Lineage 3
33	Quebrada Ganado	*Gobiomorus maculatus*	Eleotridae	09°31′55″	84°28′10″	KJ477481-89 (9)	KJ477602-607 (6)	Lineage 3

[a] Species of definitive hosts where adults were collected.
[b] Lineages and species correspond with the results of the phylogenetic tree based on COI obtained in this study.

occurs and sorts the sequences into putative species based on this distance, i.e. ABGD clusters sequences into candidate species based on the differences obtained among pairwise distances; in that way one of the aims of the method is to statistically infer the barcode gap from the data, and to partition the dataset into the maximum number of species (Puillandre *et al*. 2012). The COI alignment was then uploaded at the website: http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html. The analysis was run with the default settings [P_{min} = 0·001, P_{max} = 0·1, Steps = 10, X (relative gap width) = 1·5, Nb bins = 20], and pairwise differences were estimated with the TrN distance to contrast the result of ABGD with respect to the primary species delimitation hypothesis obtained with the COI Bayesian phylogenetic analysis. The relative gap width was subsequently modified first to 1 and then 0·5, to explore the potential effect on the number of recursive partitions and the number of species for partition.

ITS sequences and concatenated analysis

A subset of ITS sequences was obtained comprising individuals from the five COI lineages (on average, 30 specimens per lineage), and were aligned with other sequences available in GenBank, i.e. *C. attenuatum* (KP150307), *C. detruncatum* (KP110517–KP110519), *C. marginatum* (ITS: JN108032, JF718620, JF718622, JF718630–JF718643, JX631045–JX631049, JX631074–JX631101), *C. complanatum* (ITS: AY245701, FJ609420, JF718621, JF718623–JF718629), *C. phalacrocoracis* (ITS: FJ609422–FJ609423), *C. cutaneum* (ITS: FJ609421, GQ339114), *C. philippinense* (KP110570), *C. tataxumui* (ITS: JX631050–JX631073, X631102–JX631140) and *Clinostomum* sp. (ITS: KJ789384–KJ789387). Also, ITS sequences from the study of Locke *et al*. (2015*a*) were included: ITS: KP110571–KP110587. Additionally, the sequence of *Euclinostomum* sp. (ITS: KC894798–KC894801) from Senapin *et al*. (2014), and those of the strigeids *Alaria mustelae* (JF769478) and *D. baeri* (AY123042) were included as outgroups for rooting the trees (Supplementary Table S1). The ITS dataset was analysed through BI. The best-fitting model (TVM + G) was identified with the AIC implemented in jModelTest v0.1.1 (Posada, 2008). For the BI analyses, the implemented model was GTR + G, because the less complex TVM + G models are not implemented in MrBayes. BI analysis was performed using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) with the same parameters as for the COI tree. As we did for COI sequences, unique genotypes of ITS from all localities as well as unique genotypes for sequences available from the GenBank dataset were identified using DnaSP v.5. (Librado and Rozas, 2009) and NETWORK v. 4.2.0.1 (Bandelt *et al*. 1999; www.fluxusengineering.com). Finally, the concatenated (COI + ITS) datasets were analysed through BI with the same parameters, and with substitution models for each partition corresponding to each dataset.

Species tree estimation

To account for stochastic differences in the coalescent history of the mitochondrial and nuclear genes, the evolutionary history of clinostomids, including our samples from Middle-America, was reconstructed through the Species Tree Ancestral Reconstruction analysis using the program *BEAST (Heled and Drummond, 2010) as implemented in BEAST 2 (Bouckaert *et al*. 2014). *BEAST operates under a Bayesian framework and uses sequence information from different loci, and multiple individuals per taxon, to estimate the species tree, coestimating the posterior distribution of species and gene trees using a coalescent model. Based on the premise that multiple samples per species are necessary to complete the estimation and this may have a detrimental effect on inferring species topology (Heled and Drummond, 2010), we excluded the following species from the analysis because only one sequence for both molecular markers was available: *Clinostomum philippinense, C. attenuatum, C. detruncatum, Clinostomum* sp. 6 Locke *et al*. (2015*a*) and *Clinostomum* sp. 4 Locke *et al*. (2015*a*). Following the methodology described in Prévot *et al*. (2013), we ran the *BEAST analysis for 100 000 000 generations with a sample frequency of 10 000 using a lognormal clock (without fossil calibrations) and a mean rate fixed to one. Based on results from jModelTest v0.1.1 (Posada, 2008), models of DNA sequence evolution were assigned to each partition. The Yule tree prior was used for species-level analyses, while default values were used for remaining priors. The final *BEAST species tree was reconstructed using TreeAnnotator v. 2.1.2 (Bouckaert *et al*. 2014), and it was a maximum clade credibility tree with median node heights after burn-in of 15% trees. Nodal support was determined using PPs.

RESULTS

Thirty-three localities, including 26 across southern Mexico, one in Honduras and six in Costa Rica, were sampled to collect metacercariae and/or adults of *Clinostomum* spp. in their second intermediate and/or definitive hosts (Fig. 1); in total, we analysed specimens from 23 fish species belonging to six families (Cichlidae, Eleotridae, Heptapteridae, Poeciliidae, Characidae and Profundulidae) in freshwater and estuarine habitats, and five bird species (all belonging to the family Ardeidae) (Table 1). In some cases both fish and birds were sampled at the same locality, but in most of them only fish were collected.

http://dx.doi.org/10.1017/S0031182016001517

Sequences of COI ($n = 213$) were obtained from 168 metacercariae and 45 adults of Middle-American *Clinostomum*. The final alignment (including outgroups) consisted of 370 sequences of COI, including data from Genbank.

Phylogenetic analysis of COI

The alignment of the COI dataset contained no gaps and was 474 bp in length with 210 variable sites. Seventy-eight unique haplotypes were found among the 213 sequences of Middle-American *Clinostomum* generated in this study (Table 2). Base frequencies were $A = 0{\cdot}18$, $C = 0{\cdot}12$, $G = 0{\cdot}25$ and $T = 44$. Overall genetic divergence ranged from 0 to 26·9% among the 370 COI sequences of *Clinostomum* spp. Among the lineages and species of *Clinostomum* occurring in Middle-America, genetic divergence varied between 0·9 and 23% (Supplementary Table S2). The lower divergence values correspond to those of *Clinostomum* sp. 1 and *Clinostomum* sp. 2 of Locke *et al*. (2015*a*) with respect to the other isolates of Lineage 3, a result that shows conspecificity. However, only considering Lineage 3 as a whole, genetic divergence among Middle-American *Clinostomum* varies from 9 to 23% (Supplementary Table S2). Further phylogenetic analysis considered only unique haplotypes. The 78 haplotypes of our samples are clustered in five lineages that seem to represent independent evolutionary units since they form reciprocally monophyletic groups (Fig. 2), and this constitutes the primary species hypothesis for the specimens of Middle-American *Clinostomum*. Two of these lineages contain few haplotypes; Lineage 1 includes three haplotypes (1–3) and Lineage 2 includes six (14–19). However, the other three lineages contain ten or more haplotypes, with ten for Lineage 4 (4–13), 23 for Lineage 3 (20–42) and 36 (43–78) for Lineage 5. Interestingly, three of the ten candidate species of *Clinostomum* in Locke *et al*. (2015*a*) are nested within lineages uncovered in our study (Fig. 2). *Clinostomum* sp. 1 (1 COI sequence of a metacercaria from *Rhamdia guatemalensis* from Yucatan, Mexico) and *Clinostomum* sp. 2 (5 COI sequences of metacercariae from *Scycidium salvini* from Oaxaca, Mexico) cluster within Lineage 3 (39 COI sequences), while *Clinostomum* sp. 3 (1 COI sequence of a metacercaria from *Poecilia mexicana* from Veracruz, Mexico) clusters within Lineage 4 (28 COI sequences).

Species delimitation through ABGD

The ABGD was used as a species delimitation tool, to explore the primary species hypothesis obtained through the Bayesian COI phylogenetic analysis (Fig. 2). The 20 recursive steps in the ABGD analysis resulted in ten different sequence partitions (Supplementary Fig. S1), ranging from two to 74 groups (=species). However, the best correspondence between number of species obtained by ABGD and those from the phylogenetic analyses was found in the sequence partitions 7–8, with 21 groups for the outgroups (*Alaria, Diplostomum* and *Euclinostomum*) and *Clinostomum* spp. The gap width and the prior of maximum intraspecific divergence (P_{max}) were modified, and the result did not varied. All the other partitions were not considered due to the excessive splitting or lumping of identified COI lineages, respectively (Supplementary Fig. S1). Of the 21 groups differentiated by genetic divergence values between 2·1 and 3·5%, six (instead of five from the COI Bayesian analysis) were recovered for the Middle-American samples of *Clinostomum* herein studied (see Supplementary Fig. S1). Lineage 4 from the phylogenetic circumscription is split into two species in ABGD, with haplotypes 12 and 13, representing samples of adult *Clinostomum* from the fish-eating bird *Egretta thula* from Catemaco Lake in Mexico, as a separate species.

Concatenated analyses (COI + ITS)

To further corroborate the primary species hypothesis from a single-locus analysis (COI), a nuclear gene was incorporated into the analysis. A subsample ($n = 150$) that included specimens from the five lineages recovered by COI was sequenced for the ITS (ITS1, 5·8S and ITS2), and a phylogenetic tree was constructed (Supplementary Fig. S2). The final alignment (including outgroups) consisted of 307 sequences of ITS, including the data from GenBank. In total, 26 unique genotypes were obtained for the ITS dataset (Table 2). Further phylogenetic analysis considered only those sequences. The alignment of the ITS sequences was 1098 bp in length, with 297 variable sites. Gaps were treated as missing data in phylogenetic analysis. Base frequencies were $A = 0{\cdot}23$, $C = 0{\cdot}21$, $G = 0{\cdot}25$ and $T = 0{\cdot}30$. Overall ITS sequence divergence among species and lineages of *Clinostomum* ranged from 0 to 9%. BI recovered relationships similar to those of the COI tree, although only three lineages instead of five were found within Middle-American *Clinostomum* (Supplementary Fig. S2). Due to the different number of lineages of Middle-American *Clinostomum* uncovered by the mitochondrial and the nuclear genes, a concatenated analysis of both datasets was performed. The alignment consisted of 144 sequences with 1572 bp length. Figure 3 depicts the BI tree where five lineages are recovered for samples of *Clinostomum* from Middle-America. Relationships are supported by relatively high PP values.

Species tree analysis

A Species Tree Ancestral Reconstruction (*BEAST) was run to look for congruence between the BI tree

http://dx.doi.org/10.1017/S0031182016001517

Table 2. Locality, host, sample size (*n*), number of unique haplotypes and genotypes corresponding to COI lineages for the Middle-American specimens of *Clinostomum* sp. detected in this study

Locality	Host	*n*		Haplotypes	Genotypes	
		COI	ITS	COI	ITS	Lineage (COI)
Río San Juan	*Rhamdia guatemalensis*	2	3	1	1	1
El Paraiso	*Rhamdia* sp.	4	4	2,3,34	1, 21	1,3
Presa Los Ocotes	*Pseudoxiphophorus jonesii*	11	11	4,5	7,8	4
Río El Triunfo	*Profundulus punctatus*	3	3	6,7	4,5,6	4
El Platanar	*Profundulus balsanus*	3	2	8,9	9	4
Río Santa María Huatulco	*Profundulus punctatus*	5	2	10	9	4
Río Chacalapa	*Profundulus punctatus*	2	1	11	10	4
Lago de Catemaco	*Egretta thula*[a]	3	3	12,13	3	4
	Egretta thula[a]	3	3	14	3	2
	Ardea alba[a]	1	2	18	11	2
	Rhamdia guatemalensis	8	7	39–42	14,15,16,18,19,20	3
Río La Rosa	*Astyanax aeneus*	3	3	15,19	2, 12	2
Santa María del Oro	*Astyanax aeneus*	3	3	16,17	13	2
Quebrada Ganado	*Gobiomorus maculatus*	9	6	20, 23,24,27–29	23,26	3
Río Las Vueltas	*Rhamdia rogersi*	2	2	21	22	3,5
Río Grande	*Gobiomorus maculatus*	7	6	22,25, 26,35	25	3
Laguna El Milagro	*Rhamdia guatemalensis*	2	2	30	14	3,5
Emiliano Zapata	*Tigrisoma mexicanum*[a]	3	3	31,32	15	3
	Ardea herodias[a]	1	1	33	23	3
	Rhamdia guatemalensis	4	4	36–38	14	3
El Espino	*Tigrisoma mexicanum*[a]	3	4	43,46,47	17	5
Teapa	*Petenia splendida*	9	5	44,48,49	–	5
Matlapa	*Herichthys cyanoguttatus*	3	3	45,66	–	5
San Felipe	*Herichthys cyanoguttatus*	6	6	50,51	–	5
Laguna Milagros	*Vieja* sp.	5	3	52,72	–	5
Santa Cruz	*Thorichthys meeki* *Cichlasoma urophthalmus* *Thorichthys* sp.	12	12	53–56,73,74	–	5
Río Champotón	*Tigrisoma mexicanum*[a] *Cochlearius cochlearis*[a]	10	9	57,75–78	–	5
Río Verde	*Cichlasoma trimaculatum* *Tigrisoma mexicanum*[a] *Cochlearius cochlearis*[a]	9	5	58–60	–	5
Flores Magón	*Cichlasoma trimaculatum*	2	1	61	–	5
Río Atlapexco	*Herichthys cyanoguttatus*	2	1	62, 67	–	5
Río Irigaray	*Parachromis managuensis* *Amatitlania nigrofasciata*	6	2	63–65	–	5
Horquetas de Sarapiqui	*Amphilopus* sp. *Amatitlania* sp.	8	4	68,69	–	5

Table 2. (*Cont.*)

Locality	Host	*n*		Haplotypes	Genotypes	Lineage (COI)
		COI	ITS	COI	ITS	
Laguna Silvictuc	*Ardea alba*	10	9	70,71	–	5
	Cichlasoma urophthalmus					
Tilaran	*Rhamdia rogersi*	1	2	2	24	3
Laguna Azul	*Cichlasoma urophthalmus*	5	4	52		5
Total				78	26	

For ITS, a subsample of the five COI lineages was sequenced (not all localities represented).
[a] Adult specimens in fish-eating birds.

topologies of the COI and ITS datasets. Figure 4 depicts the final *BEAST species tree as a maximum clade credibility tree. This tree clearly distinguishes the five putative species of Middle-American *Clinostomum* from the morphologically described species and candidate species recognized by Locke *et al*. (2015*a*). The species tree also shows that the five genetic lineages uncovered here do not form a monophyletic group, but are nested with other species occurring in the Nearctic and Neotropical biogeographical regions. Lineage 1 from heptapterids from Mexico and Honduras is the sister taxa of *Clinostomum* sp. 7 of Locke *et al*. (2015*a*), a species found in Brazil and Bolivia. A second group consists of Lineage 4 from profundulids and poeciliids in Mexico as the sister taxon of Lineage 3 from eleotrids and heptapterids in Mexico and Lineage 2 from characids in Mexico. A third group consists of *Clinostomum* sp. 5 of Locke *et al*. (2015*a*), from a cichlid in Bolivia, as the sister taxon of *C. tataxumui* from Mexican eleotrids plus Lineage 5 found in cichlids across Mexico and Costa Rica. The latter two groups are the sister group of *C. marginatum* that includes samples that distribute from Canada to Central Mexico (Fig. 4).

Host association and host specificity of the metacercariae

The host families from which the metacercariae of Middle-American *Clinostomum* were collected were mapped onto the species tree (Fig. 4). Based on this result, and irrespective of host sample size and geographical distribution, it seems that the metacercariae of Middle-American *Clinostomum* uncovered in our study exhibit some level of specificity for particular host groups. In some cases metacercariae are restricted to a particular locality, and in other cases they are found across a wide geographic range, with host specificity patterns independent of geographical distribution (see Figs 2 and 4). Lineage 1 was found in heptapterids (*R. guatemalensis* and *Rhamdia* sp., from Mexico and Honduras); lineage 2 was only found in *Astyanax aeneus* (Characidae) in western Mexico; lineage 3 seems to be the least host-specific since it was found in eleotrids from Costa Rica, and heptapterids and gobiids from Mexico; lineage 4 was only found in cyprinodontiform fish (poeciliids and profundulids) in Mexico; and finally, lineage 5 was only found in cichlids (at least 12 species) across a wide geographic range comprising central Mexico southwards to Costa Rica.

DISCUSSION

Clinostomum is widely distributed across Middle-America; the metacercariae allocated to this genus have been recorded (either as *Clinostomum complanatum* or *Clinostomum* sp.) in at least 76 fish species

http://dx.doi.org/10.1017/S0031182016001517

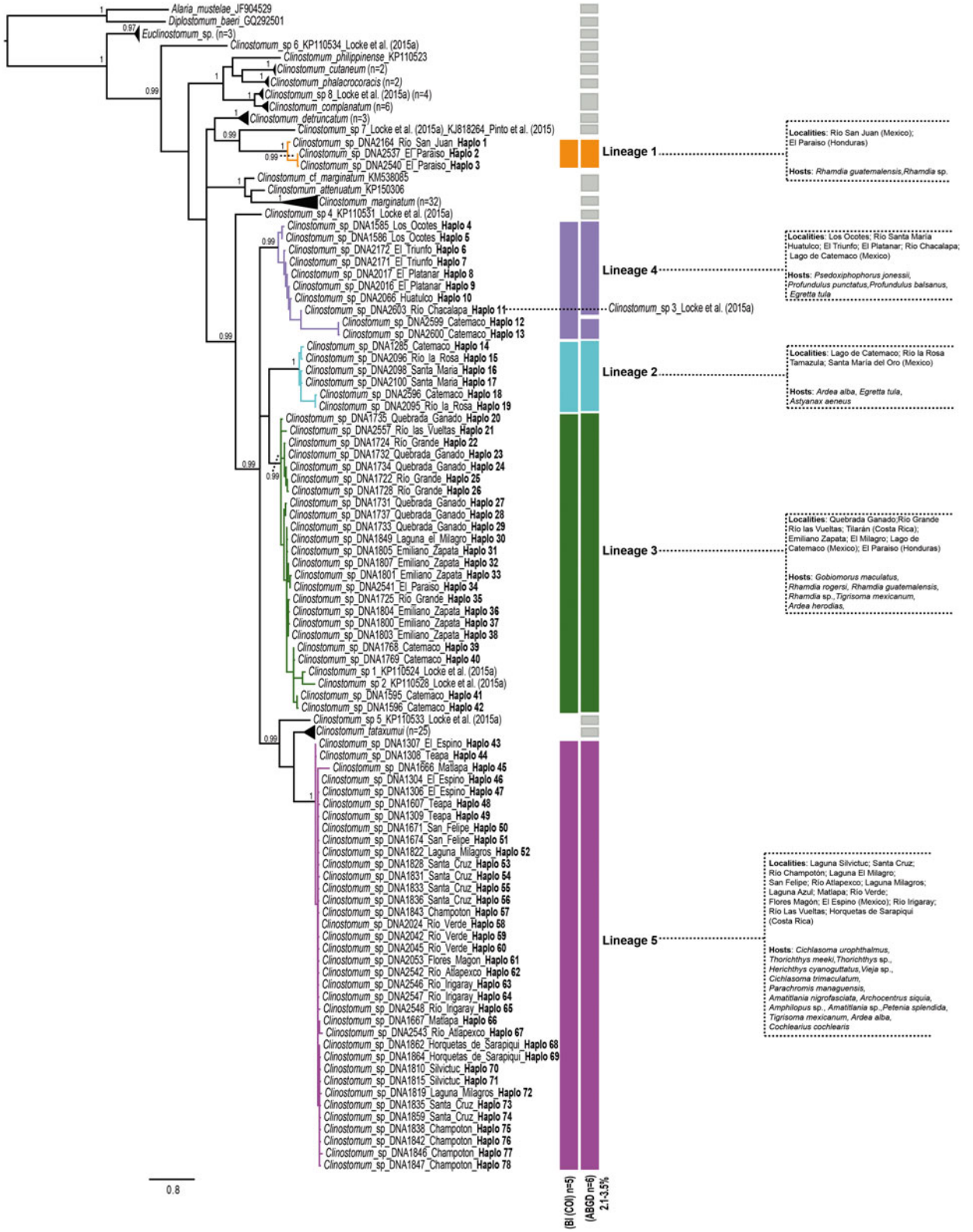

Fig. 2. Phylogram for *Clinostomum* spp. derived from partial sequences of the mitochondrial COI gene (474 bp) using BI analysis (50% majority rule tree). Sequences generated in this study were aligned with available sequences for *Clinostomum* in GenBank. Nodal support values next to the branches correspond to the PPs (⩾0·95). The scale bar represents the number of nucleotide substitutions per site. Column in colour correspond with the five genetic lineages of 78 unique COI haplotypes. The second column shows the estimated entities through the ABGD method for partitions with pairwise distance of TrN between 1·3 and 3·5%. The intermediate and definitive hosts as well as the localities for each lineage are included.

Fig. 3. Tree (50% majority rule) inferred from the concatenated (COI + ITS) datasets for *Clinostomum* using BI analysis. The alignment consisted of sequences generated in this study for samples of *Clinostomum* from Middle-America, and those available in GenBank. The scale bar represents the number of nucleotide substitutions per site. Nodal support values next to the branches correspond to the Bayesian PPs (⩾0·95). Column in colour correspond with the five genetic lineages.

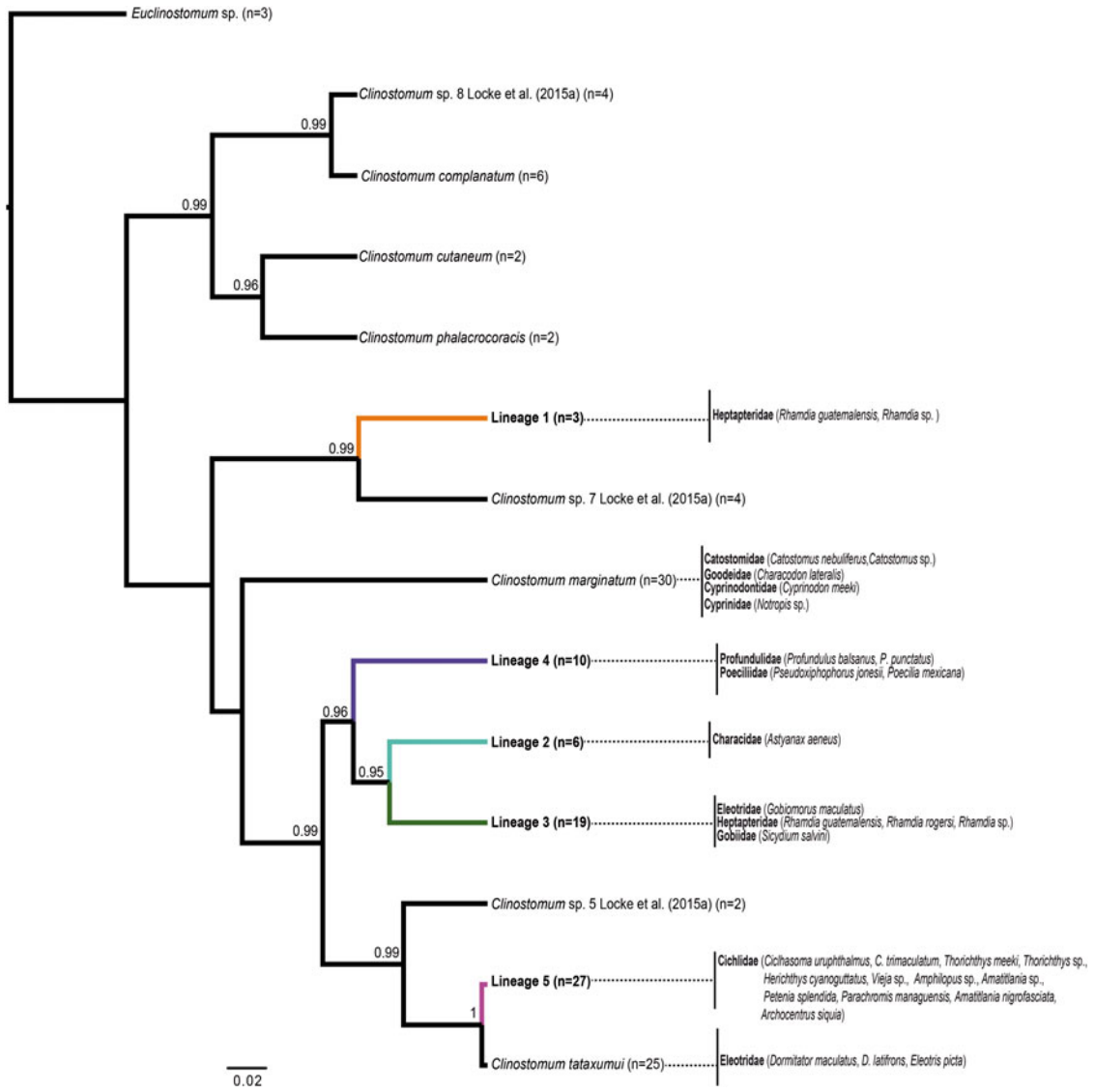

Fig. 4. Maximum clade credibility tree of *Clinostomum* spp. based on COI and ITS1–5·8S–ITS2 using Bayesian coalescence. Main lineages are denoted as 1–5, and colours follow those in Figs 2 and 3. Note that the following species were excluded to avoid a detrimental effect on inferring species topology because only one sequence was available: *Clinostomum philippinense, C. attenuatum, C. detruncatum, Clinostomum* sp. 6 Locke *et al*. (2015*a*) and *Clinostomum* sp. 4 Locke *et al*. (2015*a*). The number of concatenated ITS and COI sequences per species (*n*) follows the species name. Putative species of *Clinostomum* sp. of Locke *et al*. (2015*a*) are indicated. Bayesian PPs ⩾0·95 are shown above the node.

belonging to 13 families, in two localities of Nicaragua, six in Costa Rica and 101 across Mexico (Aguirre-Macedo *et al*. 2001; Pérez-Ponce de León *et al*. 2007; Sandlund *et al*. 2010); in contrast, adults have only been found in ten species of fish-eating birds in 11 localities of Mexico (see Sereno-Uribe *et al*. 2013 and references therein). In a recent study aiming to identify the species of *Clinostomum* across Mexico, Sereno-Uribe *et al*. (2013) showed that the species *C. marginatum* (and not *C. complanatum*) is found in northern areas of Mexico, and a second species genetically and morphologically distinct, *C. tataxumui*, was described from southern areas corresponding to Middle-America. In the present study, extensive samples of *Clinostomum* spp. were obtained in that geographic region, and numerous individuals were sequenced for both mitochondrial and nuclear genes, and analysed in combination with those available in the GenBank dataset from the rest of the world. The level of genetic divergence uncovered was unexpected since we had recently established the presence of just the two aforementioned species of *Clinostomum* across Mexico. To our surprise, both nuclear and mitochondrial molecular markers revealed divergence levels indicating greater diversity than previously thought, with the potential existence of independent evolutionary units, or species. We used COI as a proxy to obtain a primary species hypothesis.

High levels of genetic divergence were found among COI lineages, reaching up to 26·9%; typical

pairwise divergence at the COI gene among trematode congeners varies between 3·9 and 25% (Moszczynska *et al*. 2009; Herrmann *et al*. 2014; Vilas *et al*. 2015). For instance, in diplostomids, for which a large genetic library has been built (Locke *et al*. 2010, 2015*b*; Georgieva *et al*. 2013; Blasco-Costa *et al*. 2014), molecular prospecting studies show similar levels of genetic divergence for this molecular marker. Georgieva *et al*. (2013) found a range for average divergence between 4·6 and 14·9% among Palaearctic *Diplostomum* spp.; Chibwana *et al*. (2013) found interspecific divergences between 11·7 and 14·8% among African diplostomids; Otachi *et al*. (2014) recorded divergence levels between 7·2 and 15·8% among species of *Tylodelphys* in Kenya; and finally, Pinto *et al*. (2015) found variation levels between 21·3 and 28·1% between samples of *Clinostomum* recovered from a freshwater poeciliid fish in Brazil, and sequences available in GenBank for other congeners. Clearly, using COI pairwise distance thresholds (i.e. a certain percentage of base pair differences) for initial assessment of species diversity based on sequence data appears reasonable because high levels of divergence are unlikely among individuals of single species. However, this approach is problematic for species delimitation because critical assumptions implicit in these benchmark comparisons may be violated. Modern taxonomic approaches show that the use of a subjective genetic threshold alone in supporting the delimitation of putative species is not satisfactory, and at least evidence of reciprocal monophyly of lineages through phylogenetic analyses of more than one locus must be found as a starting point to delineate species by hypothesis-testing and operational procedures, including explicitly evolutionary methods (see Nadler and Pérez-Ponce de León, 2011; Blasco-Costa *et al*. 2016*a*). The aforementioned studies used a hypothesis-testing procedure through proper phylogenetic analyses, in addition to pairwise divergence values.

A single gene-tree (COI) was estimated using BI; the primary species hypothesis was drawn from the haplotype groups recovered from that analysis. As a result, five putative species were recognized; following the 'unified species concept' of de Queiroz (2007), we consider species as 'lineages evolving separately from others'. However, we acknowledge that the best practice to infer more robust species limits in parasitic organisms would be to use different lines of evidence, i.e. molecular, morphological, ecological (host associations), and biogeographical (geographical distribution) through an integrative taxonomic approach. Regarding the analysis of molecular data, it might be necessary first to explore the resulting hypotheses of species delimitation through phylogenetic methods and pairwise divergence, with other methods available in the taxonomic literature (see Puillandre *et al*. 2012; Camargo and Sites, 2013; Carstens *et al*. 2013; Flot, 2015), but this should be determined on a case by case basis because it is possible that in some trematode groups researchers may fail to recognize candidate species through these methods (see Carstens *et al*. 2013). Certainly, the extent to which these methods provide compelling evidence to discriminate among trematode species has not been explored in great detail, as very few empirical studies are available. Martínez-Aquino *et al*. (2013) used the General Mixed Yule Coalescent model (GMYC) to establish species boundaries among populations of the allocreadiid *Margotrema* spp. in central Mexico. Blasco-Costa *et al*. (2014) provided eight lines of evidence to delineate species of *Diplostomum* in Iceland. In addition to phylogenetic inference, these authors followed a character-based DNA barcoding approach to identify diagnostic characters discriminating the novel Icelandic lineages through the CAOS (Characteristic Attribute Organization System) framework and further estimated a species tree topology from mitochondrial and nuclear gene trees under a Bayesian multispecies coalescent model. Herrmann *et al*. (2014) used multi-locus data and implemented the Bayesian species delimitation method, with distinct COI lineages as the maximum putative number of species, to delineate two cryptic species of *Stegodexamene anguillae* in New Zealand. Locke *et al*. (2015*a*, *b*) applied the ABGD and the Barcode Index Numbers (BINs) to establish species limits among larval clinostomids and diplostomids, respectively. Pérez-Ponce de León *et al*. (2015) used a multispecies coalescent analysis (species tree) approach from a dataset combining nuclear and mitochondrial genes to determine the species limits of two new species of *Phyllodistomum*.

In our study, a distance-based algorithm (ABGD) (Puillandre *et al*. 2012) was used as a method to explore the primary species hypothesis obtained through the COI phylogenetic circumscription of Middle-American species of *Clinostomum*. In principle, a set of a priori threshold values was considered for delineating genetically distinct species (Puillandre *et al*. 2012) and this was used to partition the data into groups of sequences, each representing a hypothetical species. Even though the gap width and the prior of maximum intraspecific divergence were modified, only one additional species was recognized using ABGD among the Middle-American samples. Six to seven of the ten partitions were not even considered due to excessive splitting (up to 81 species) or lumping (two species) of identified COI lineages (see Supplementary Fig. S2). In this way, ABGD species delineation tool recognized six instead of five species. Irrespective of the gap width used ($X = 1{\cdot}5$, 1, 0·5 and variation of priors of maximum intraspecific divergence), two particular haplotypes (12 and 13 from adult specimens collected in fish-eating birds in Catemaco, Veracruz), are separated as putative

http://dx.doi.org/10.1017/S0031182016001517

species from Lineage 4 considering the pairwise distance (see Fig. 2). The genetic divergence value obtained through the TrN model between these two haplotypes, and those included in Lineage 4 varies between 10·4 and 12·6%, while the divergence among all the other haplotypes (4–11) only varies from 0·4 to 4·2%. However, the COI Bayesian analysis, and the concatenated analysis of the COI and ITS datasets (Fig. 3), unequivocally recognized five putative species among the Middle-American samples of *Clinostomum*. These two haplotypes recognized by ABGD are not recovered as a reciprocally monophyletic group in the COI and the COI + ITS concatenated analysis, and the nodal support of the Bayesian analysis shows that altogether form a well-supported monophyletic group. For that reason we reject their validity as a separate species. In our opinion, ABGD should not be used as the only method to establish species boundaries among trematodes because species diversity might be overestimated. This method should be used in combination with a phylogenetic analysis preferably of two unlinked loci. We further validated the species delineation by undertaking a concatenated analysis and a final assessment was achieved through a species tree analysis. Our study was fundamentally aimed at documenting current species diversity among these trematodes. Still, the five putative species uncovered here require proper description; this will follow in a separate contribution with formal description, and name, of the species based on adult characters, complemented with a characterization of their metacercariae. Fortunately, we obtained adult individuals from fish-eating birds for four of the five lineages herein validated as putative species, and we were able to establish a link between metacercariae and adults. The detailed morphological study that will follow as the second part of the study will determine if particular traits are reliable to separate the putative species. Characters such as the organization of the genital complex, body width, distance between suckers, and position of the genital pore seem reliable, but still species of *Clinostomum* are very similar to each other and differ only in relatively minor characters of the adults (Matthews and Cribb, 1998) and metacercariae (Caffara *et al.* 2011). Interestingly, our relatively small-scale study of *Clinostomum* in Middle-America yielded five putative species that include three of the eight congeneric species recognized by Locke *et al.* (2015*a*) in their large-scale study. As a result, at least ten additional species have to be added to the genus *Clinostomum*, currently composed by 15 valid species, raising the number to 25.

Patterns of host specificity of the metacercariae

Hosts belonging to six fish families (Cichlidae, Characidae, Eleotridae, Heptapteridae, Poecilidae and Profundulidae) were studied in each sampled locality (data not shown), but irrespective of host sample size which greatly varied among localities, the metacercariae of the five lineages herein uncovered seem to show some fidelity for a particular family of freshwater fish. Mapping the second intermediate host onto the species tree suggests that metacercariae of the five putative species of Middle-American *Clinostomum* exhibit some level of host specificity to fish taxa with distinct ecological and phylogenetic affinities (Fig. 4). Overall, host specificity in clinostomid metacercariae in Middle-American fish seems to be more strongly related to the physiological compatibility of host and parasite than to ecological factors. Locke *et al.* (2010) observed this pattern among diplostomid metacercariae in sympatric fish along the St. Lawrence River, Canada. In our study, in one particular locality, Laguna el Milagro, Campeche (collection site 4, see Table 1, Fig. 1), three fish species occurring in sympatry were studied, i.e. the Mayan cichlid *Cichlasoma urophthalmus*, the Guatemalan chulin *R. guatemalensis*, and the fat sleeper goby *Dormitator maculatus*. These hosts were infected with Lineage 5, Lineage 3 and *Clinostomum tataxumui*, respectively. Without molecular or experimental evidence regarding the life cycle, these three lineages would probably have been considered as a single species, and because they represent larval forms, they would be designated as *Clinostomum* sp. Taxonomic surveys provide the data for interpretation of host specificity. In this context, since host specificity is a fundamental property of parasitic organisms (Poulin *et al.* 2011), the use of objective tools to delimitate parasite species is crucial to its accurate measurement. This corresponds with the scenario described by Poulin and Keeney (2008) in that morphologically identical individuals included under one species name may consist of several different isolated gene pools. This idea lead these authors to conclude that even though morphological species descriptions remain essential, the specificity of most parasite taxa will need to be reassessed by using both morphological and genetic data.

Concluding remarks

Our data suggest high levels of genetic diversity in *Clinostomum* across a range spanning from Mexico southwards to Costa Rica. Five putative species are recognized. Based on our results we predict that an even larger hidden diversity remains to be discovered within *Clinostomum* at least in the Americas, particularly if a large number of specimens are sequenced from a wide geographic range as suggested by Blasco-Costa *et al.* (2016*a*). This study originally aimed to obtain genetic data to show that the species of *Clinostomum* occurring in Middle-America was *C. tataxumui*, the species we recently

http://dx.doi.org/10.1017/S0031182016001517

described from Southern Mexico. Unexpectedly, we discovered several genetic lineages. Based on the fact that *Clinostomum* individuals are commonly found in freshwater fishes across the Americas, we contend that if a more extensive sequencing effort is made, the result will be an increase in the number of species recognized. In trematodes such as *Clinostomum*, where high levels of genetic variation are found among individuals, we recommend the use of a species delimitation method to explore species circumscription based on single-locus or even multilocus phylogenetic analyses, and the use of several lines of evidence to further corroborate the species hypothesis (de Queiroz, 2007) such as interruption of gene flow, phenotypic variation and ecological niche differentiation. For parasitic organisms, where a test of reproductive isolation is methodologically complicated, patterns of host association and biogeography are informative. Such approach will certainly contribute towards reliable parasite biodiversity estimates. Our study also illustrates one of the key uses of molecular markers in parasite species identification (Criscione *et al.* 2005), i.e. the partial elucidation of life cycles by establishing the species that may serve as intermediate or paratenic hosts for larval stages. Adult specimens were obtained from the mouth cavity and oesophagus of five bird species from at least ten localities across Mexico, and molecular data now allow us to link adults and metacercariae of at least four of the five putative species herein validated. The next step in our research, as mentioned before, is to properly describe (and name) these four lineages and collect more data from fish-eating birds in the search for adults of the lineage known solely from metacercariae. This will provide also an opportunity for a closer look at the diagnostic morphological characters used to delimit species in sexually mature specimens, and find those that are reliable for species identification (see Matthews and Cribb, 1998). Still, much remains to be learned about host specificity patterns of clinostomids on a world-wide scale. The study of a wide array of definitive hosts, as well as the first intermediate hosts (gastropods), will be crucial to our understanding of the evolutionary biology and biogeography of this parasitic group.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at http://dx.doi.org/10.1017/S0031182016001517.

ACKNOWLEDGEMENTS

We are grateful to Arturo Angulo and Carlos Garita-Alvarado for their help during fieldwork in Costa Rica through the collecting permit issued by the Costa Rican government, and Leopoldo Andrade and Eduardo Hernández for their help during field work in Mexico. We also thank Luis García Prieto for providing specimens deposited at the CNHE. Specimens in Mexico were collected under the Cartilla Nacional de Colector Científico (FAUT 0202 and 0057) issued by the Secretaría del Medio Ambiente y Recursos Naturales (SEMARNAT), to MGV and GPPL, respectively. CDPP thanks the support of the Programa de Posgrado en Ciencias Biológicas, UNAM and CONACYT for granting a scholarship to complete his PhD program.

FINANCIAL SUPPORT

This paper was written during the sabbatical leave of G. P. P. L. to the Department of Zoology, University of Otago, New Zealand. Thanks are due to Dirección General de Asuntos del Personal Académico (DGAPA-UNAM) and CONACyT (Consejo Nacional de Ciencia y Tecnología), for their financial support to travel to New Zealand. This research was supported by grants from the Programa de Apoyo a Proyectos de Investigación e Inovación Tecnológica (PAPIIT-UNAM) IN204514 and IN206716 to G. P. P. L. and M. G. V., respectively, and CONACYT 179048 to M. G. V.). We also thank CONACYT for supporting the development of the Mexican Barcode of Life (MEXBOL) network through grant 251085 to Virginia León.

REFERENCES

Aguirre-Macedo, L., Scholz, T., González-Solís, D., Vidal-Martínez, V., Posel, P., Arjona-Torres, G., Siu-Estrada, E. and Dumailo, S. (2001). Larval helminths parasitizing freshwater fishes from the Atlantic coast of Nicaragua. *Comparative Parasitology* **68**, 42–51.

Athokpam, V. D., Jyrwa, D. B. and Tandon, V. (2014). Utilizing ribosomal DNA gene marker regions to characterize the metacercariae (Trematoda: Digenea) parasitizing piscine intermediate hosts in Manipur, Northeast India. *Journal Parasitic Diseases* **40**, 330–338.

Bandelt, H. J., Forster, P. and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37–48.

Blasco-Costa, I., Faltynková, A., Georgieva, S., Skírnisson, K., Scholz, T. and Kostadinova, A. (2014). Fish pathogens near the Arctic Circle: molecular, morphological and ecological evidence for unexpected diversity of *Diplostomum* (Digenea: Diplostomidae) in Iceland. *International Journal for Parasitology* **44**, 703–715.

Blasco-Costa, I., Cutmore, S. C., Miller, T. L. and Nolan, M. J. (2016*a*). Molecular approaches to trematode systematics: 'best practice' and implications for future study. *Systematic Parasitology* **93**, 295–306.

Blasco-Costa, I., Poulin, R. and Presswell, B. (2016*b*). Species of *Apatemon* Szidat, 1928 and *Australapatemon* Sudarikov, 1959 (Trematoda: Strigeidae) from New Zealand: linking and characterising life cycle stages with morphology and molecules. *Parasitology Research* **115**, 271–289.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A. and Drummond, A. J. (2014). BEAST 2: a Software platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology* **10**, e1003537.

Caffara, M., Locke, S. A., Gustinelli, A., Marcogliese, D. J. and Fiovaranti, M. L. (2011). Morphological and molecular differentiation of *Clinostomum complanatum* and *Clinostomum marginatum* (Digenea: Clinostomidae) metacercariae and adults. *Journal of Parasitology* **97**, 884–891.

Caffara, M., Bruni, G., Paoletti, C., Gustinelli, A. and Fioravanti, M. L. (2013). Metacercariae of *Clinostomum complanatum* (Trematoda: Digenea) in European newts *Triturus carnifex* and *Lissotriton vulgaris* (Caudata: Salamandridae). *Journal of Helminthology* **88**, 278–285.

Caffara, M., Davidovich, N., Falk, R., Smirnov, M., Ofek, T., Cummings, D., Gustinelli, A. and Fioravanti, M. L. (2014). Redescription of *Clinostomum phalacrocoracis* metacercariae (Digenea: Clinostomidae) in cichlids from Lake Kinneret, Israel. *Parasite* **21**, 32.

Camargo, A. and Sites, J. W. (2013). Species delimitation: a decade after the Renaissance. In *The Species Problem-Ongoing Issues* (ed. Pavlinov, I.), pp. 225–247. InTech., Croatia.

Carstens, B. C., Pelletier, T. A., Reid, N. M. and Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology* **22**, 4369–4383.

Chibwana, F. D., Blasco-Costa, I., Georgieva, S., Hosea, K. M., Nkwengulila, G., Scholz, T. and Kostadinova, A. (2013). A first insight into the barcodes for African diplostomids (Digenea: Diplostomidae): brain parasites in *Clarias gariepinus* (Siluriformes: Clariidae). *Infectious Genetics and Evolution* **17**, 62–70.

Chibwana, F. D., Nkwengulila, G., Locke, S. A., McLughlin, J. D. and Marcogliese, D. J. (2015). Completion of the life cycle of *Tylodelphys mashonense* (Sudarikov, 1971) (Digenea: Diplostomidae) with DNA barcodes and rDNA sequences. *Parasitology Research* **114**, 3675–3682.

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

de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology* **56**, 879–886.

Dzikowski, R., Levy, M. G., Poore, M. F., Flowers, J. R. and Paperna, I. (2004). *Clinostomum complanatum* and *Clinostomum marginatum* (Rudolphi, 1819) (Digenea: Clinostomidae) are separate species based on differences in ribosomal DNA. *Journal of Parasitology* **90**, 413–414.

Flot, J. F. (2015). Species delimitation's coming of age. *Systematic Biology* **64**, 897–899.

García-Varela, M., Sereno-Uribe, A. L., Pinacho-Pinacho, C. D., Domínguez- Domínguez, O. and Pérez-Ponce de León, G. (2016*a*). Molecular and morphological characterization of *Austrodiplostomum ostrowskiae* Dronen (Digenea: Diplostomatidae), a parasite of cormorants in the Americas. *Journal of Helminthology* **90**, 174–185.

García-Varela, M., Sereno-Uribe, A. L., Pinacho-Pinacho, C. D., Hernández- Cruz, E. and Pérez-Ponce de León, G. (2016*b*). An integrative taxonomic study reveals a new species of Tylodelphys Diesing, 1950 (Digenea: Diplostomidae) in central and northern Mexico. *Journal of Helminthology*. doi: 10.1017/S0022149X15000917.

Georgieva, S., Soldánová, M., Pérez-del-Olmo, A., Dangel, D. R., Sitko, J., Sures, B. and Kostadinova, A. (2013). Molecular prospecting for European *Diplostomum* (Digenea: Diplostomidae) reveals cryptic diversity. *International Journal for Parasitology* **43**, 57–72.

Gustinelli, A., Caffara, M., Florio, D., Otachi, E. O., Wathuta, E. M. and Fiovaranti, M. L. (2010). First description of the adult stage of *Clinostomum cutaneum* Paperna, 1964 (Digenea: Clinostomidae) from grey herons *Ardea cinerea* L. and a redescription of the metacercaria from the Nile tilapia *Oreochromis niloticus* niloticus (L.) in Kenya. *Systematic Parasitology* **76**, 39–51.

Heled, J. and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology Evolution* **27**, 570–580.

Herrmann, K. K., Poulin, R., Keeney, D. B. and Blasco-Costa, I. (2014). Genetic structure in a progenetic trematode: signs of cryptic species with contrasting reproductive strategies. *International Journal for Parasitology* **44**, 811–818.

Huelsenbeck, J. P. and Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Librado, P. and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452.

Locke, S. A., McLaughlin, J. D., Dayanandan, S. and Marcogliese, D. J. (2010). DNA barcodes show cryptic diversity and a potential physiological basis for host specificity among Diplostomoidea (Platyhelminthes: Digenea) parasitizing freshwater fishes in the St. Lawrence River, Canada. *Molecular Ecology* **19**, 2813–2827.

Locke, S. A., McLaughlin, J. D., Lapierre, A. R., Johnson, P. T. J. and Marcogliese, D. J. (2011). Linking larvae and adults of *Apharyngostrigea cornu, Hysteromorpha triloba*, and *Alaria mustelae* (Diplostomoidea: Digenea) using molecular data. *Journal of Parasitology* **97**, 846–851.

Locke, S. A., Caffara, M., Marcogliese, D. J. and Fioravanti, M. L. (2015*a*). A large-scale molecular survey of *Clinostomum* (Digenea, Clinostomidae). *Zoologica Scripta* **44**, 203–217.

Locke, S. A., Al-Nasiri, F. S., Caffara, M., Drago, F., Kalbe, M., Lapierre, A. R., McLaughlin, J. D., Nie, P., Overstreet, R., Souza, G. T. R., Takemoto, R. M. and Marcogliese, D. J. (2015*b*). Diversity, specificity and speciation in larval Diplostomidae (Platyhelminthes: Digenea) in the eyes of freshwater fish, as revealed by DNA barcodes. *International Journal for Parasitology* **45**, 841–855.

Luton, K., Walker, D. and Blair, D. (1992). Comparison of ribosomal internal transcribed spacer from two congeneric species of flukes (Plathyhelminthes: Trematoda: Digenea). *Molecular Biochemical Parasitology* **56**, 323–328.

Maddison, D. R. and Maddison, W. P. (2002). *MacClade Version 4.0*. Sinauer Associates, Sunderland.

Martínez-Aquino, A., Ceccarelli, F. S. and Pérez-Ponce de León, G. (2013). Molecular phylogeny of the genus *Margotrema* (Digenea: Allocreadiidae), parasitic flatworms of goodeid freshwater fishes across central Mexico: species boundaries, host specificity and geographical congruence. *Zoological Journal of the Linnean Society of London* **168**, 1–16.

Matthews, D., Cribb, T. H. (1998). Digenetic trematodes of the genus *Clinostomum* Leidy, 1856 (Digenea: Clinostomidae) from birds of Queensland, Australia, including *C. wilsoni* n. sp. from *Egretta intermedia*. *Systematic Parasitology* **39**, 199–208.

Moszczynska, A., Locke, S. A., McLaughlin, D., Marcogliese, D. J., Crease, T. J. (2009). Development of primers for the mitochondrial cytochrome *c* oxidase I gene in digenetic trematodes (Platyhelminthes) illustrates the challenge of barcoding parasitic helminths. *Molecular Ecology Resorces* **9**, 75–82.

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

Otachi, E. O., Locke, S. A., Jirsa, F., Fellner-Franck, C. and Marcogliese, D. (2014). Morphometric and molecular analyses of *Tylodelphys* sp. metacercariae (Digenea: Diplostomidae) from the vitreous humour of four fish species from Lake Naivasha, Kenya. *Journal of Helminthology* **89**, 404–414.

Pérez-Ponce de León, G. and Nadler, S. A. (2010). What we don't recognize can hurt us: a plea for awareness about cryptic species. *Journal of Parasitology* **96**, 453–464.

Pérez Ponce de León, G., García Prieto, L. and Mendoza-Garfías, B. (2007). Trematode parasites (Platyhelminthes) of wildlife vertebrates in Mexico. *Zootaxa* **1534**, 1–247.

Pérez-Ponce de León, G., Martínez-Aquino, A. and Mendoza-Garfias, B. (2015). Two new species of *Phyllodistomum* Braun, 1899 (Digenea: Gorgoderidae), from freshwater fishes (Cyprinodontiformes: Goodeidae: Goodeinae) in central Mexico: an integrative taxonomy approach using morphology, ultrastructure and molecular phylogenetics. *Zootaxa* **4013**, 87–99.

Pinto, H. A., Caffara, M., Fioravanti, M. L. and Melo, A. L. (2015). Experimental and molecular study of cercariae of *Clinostomum* sp. (Trematoda: Clinostomidae) from *Biomphalaria* spp. (Mollusca: Planorbidae) in Brazil. *Journal of Parasitology* **101**, 108–113.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology Evolution* **25**, 1253–1256.

Poulin, R. (2011). Uneven distribution of cryptic diversity among higher taxa of parasitic worms. *Biology Letters* **7**, 241–244.

Poulin, R. (2014). Parasite biodiversity revisited: frontiers and constraints. *International Journal for Parasitology* **44**, 581–589.

Poulin, R. and Keeney, D. B. (2008). Host specificity under molecular and experimental scrutiny. *Trends in Parasitology* **24**, 24–28.

Poulin, R., Krasnov, B. R. and Mouillot, D. (2011). Host specificity in phylogenetic and geographic space. *Trends in Parasitology* **27**, 355–361.

Prévot, V., Jordaens, K., Sonet, G. and Backeljau, T. (2013). Exploring species level taxonomy and species delimitation methods in the facultatively self-fertilizing land snail genus *Rumina* (Gastropoda: Pulmonata). *PLoS ONE* **8**, e60736.

Puillandre, N., Lambert, A., Brouillet, S. and Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology* **21**, 1864–1877.

Rambaut, A. (2006). *FigTree v1.3.1*. Institute of Evolutionary Biology. University of Edinburgh, UK.

Rambaut, A. and Drummond, A. J. (2007). Tracer v1.4. http://beast.bio.ed.ac.uk/Tracer.

Ratnasingham, S. and Hebert, P. D. (2013). A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS ONE* **8**, e66213.

Sandlund, O. T., Daverdin, R. H., Choudhury, A., Brooks, D. R. and Diserud, O. H. (2010). A survey of freshwater fishes and their macroparasites in the Guanacaste Conservation Area (ACG), Costa Rica. Norwegian Institute for Nature Research Report 635 Trondheim, Norway.

Senapin, S., Phiwsaiya, K., Laosinchai, P., Kowasupat, C., Ruenwongsa, P. and Panijpan, B. (2014). Phylogenetic analysis of parasitic trematodes of the genus *Euclinostomum* found in Trichopsis and Betta fish. *Journal of Parasitology* **100**, 368–371.

Sereno-Uribe, A. L., Pinacho-Pinacho, C. D., García-Varela, M. and Pérez-Ponce de León, G. (2013). Using mitochondrial and ribosomal DNA sequences to test the taxonomic validity of *Clinostomum complanatum* Rudolphi, 1814 in fish-eating birds and freshwater fishes in Mexico, with the description of a new species. *Parasitology Research* **112**, 2855–2870.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MEGA5: molecular evolutionary genetic analysis

using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology Evolution* **28**, 2731–2739.

Thompson, J. D., Higgins, H. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

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

Wilgenbusch, J. C., Warren, D. L. and Swofford, D. L. (2004). AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.