



An integrative taxonomic study of the genus *Theromyzon* (Hirudinea: Glossiphoniidae), with description of a new North American species

Maddy Foote^{A,B,*} ID, Rafael Iwama^{A,B}, Danielle de Carle^{A,B} and Sebastian Kvist^{A,B}

For full list of author affiliations and declarations see end of paper

***Correspondence to:**

Maddy Foote
Department of Natural History,
Royal Ontario Museum, 100 Queens Park,
Toronto, ON, M5S 2C6, Canada
Email: madeleine.foote@alum.utoronto.ca

Handling Editor:
Greg Rouse

ABSTRACT

Theromyzon Philippi, 1867 is a genus of sanguivorous, freshwater leeches in the family Glossiphoniidae. The genus is broadly distributed across the globe, possibly due to the frequent feeding in the nasopharyngeal cavities of migratory waterfowl that may allow for long distance dispersal. The genus has a history of taxonomic confusion resulting from mischaracterisations of key morphological features of type specimens that have produced several re-descriptions and synonyms. Here, we bring partial order to this confusion through robust morphological investigations of newly collected North American (and a single South American) specimens, representing most of the known species diversity from this continent. We also produce the first species-level phylogeny for *Theromyzon* and attempt to understand species boundaries regarding both morphology and genetics. Our results demonstrate that there are at least five species of *Theromyzon* present in North America (*T. bifarium*, *T. tessulatum*, *T. rude*, *T. trizonare*, and a clade that needs further investigation), and a hitherto undescribed taxon that does not conform to any previously published description, and represents a unique lineage in the phylogeny; we describe this new species under the name *Theromyzon tigris* sp. nov. This study sheds light on the discriminatory power of select morphological characters and the distribution of phenotypes within the genus. We also provide a comprehensive classification framework for the known species within the genus designed to facilitate identification and minimise future taxonomic confusion.

Keywords: Glossiphoniidae, Hirudinea, Hirudinida, leech, phylogenetics, species diversity, taxonomy, *Theromyzon*, *Theromyzon tigris*.

Introduction

The obligate sanguivorous freshwater genus *Theromyzon* was proposed by Philippi (1867) to accommodate the South American leech *Theromyzon pallens* Philippi, 1867, and members of the genus have since been recorded throughout all of the major biogeographical regions with the exception of Australasia (e.g. Christoffersen 2007; Sket and Trontelj 2008; Bielecki *et al.* 2016). Members of *Theromyzon* are traditionally characterised by the presence of four pairs of eyespots arranged in two parallel rows on somites II, III, IV and V, making genus-level identifications relatively straightforward. Most, if not all species of *Theromyzon* are considered temporary ectoparasites of birds, including migratory waterfowl, leading researchers to propose host-mediated dispersal as an efficient way for the leeches to be transported across long distances (Elliott and Tullett 1982). This host affinity possibly underlies the exceedingly broad distribution of *Theromyzon* and may, in turn, negate the utility of geography as a character for specimen identifications and species circumscriptions.

Identification of *Theromyzon* specimens to the species level is also challenging due to the difficulty in observing key morphological characters and the absence of a robust and comprehensive phylogenetic hypothesis. One of the morphological features historically used to separate species is the number of annuli between gonopores. However, the

Received: 16 September 2021

Accepted: 20 February 2022

Published: 27 July 2022

Cite this:

Foote M *et al.* (2022)
Invertebrate Systematics
36(7), 631–646. doi:[10.1071/IS21062](https://doi.org/10.1071/IS21062)

© 2022 The Author(s) (or their employer(s)). Published by
CSIRO Publishing.

gelatinous tissues of several species within the genus, coupled with non-standardised specimen preparations, and the minute size of the gonopores has impaired identifications in the past. With no geographical component to fall back on, species names seem to have been applied to specimens in a more or less arbitrary manner that has led to extensive taxonomic confusion and erroneous identifications (Oosthuizen and Davies 1993). For example, in North America, the first *Theromyzon* species on record was collected by Baird (1869) from Great Bear Lake in the Northwest Territories, Canada. At the time, the specimens were identified as *Glossiphonia rufa* Baird, 1869, but syntypes were later determined to alternately be *Placobdella ornata* (Verrill, 1872) and *Theromyzon rude* (Baird, 1863) (Oosthuizen and Davies 1992). At present, five valid *Theromyzon* species have been recorded from North America: *Theromyzon rude* (Baird, 1869), *Theromyzon bifarium* Oosthuizen & Davies, 1993, *Theromyzon maculosum* (Rathke, 1862), *Theromyzon tessulatum* (Müller, 1774) and *Theromyzon trizonare* Davies & Oosthuizen, 1993 (Davies and Oosthuizen 1993).

Further complicating specimen identification is the apparent global distribution of some species. For example, *Theromyzon tessulatum* and *Theromyzon maculosum* both have type localities in Europe, but Oosthuizen and Davies (1993) collected specimens in North America that were considered morphologically identical to European type specimens, suggesting conspecificity. If these leeches do indeed disperse via their avian hosts, this could easily be a real pattern. However, given the paucity of reliable taxonomic information and the difficulty in observing key diagnostic characters, *T. tessulatum* and *T. maculosum* could equally likely contain undiscovered cryptic diversity.

To alleviate the confusions regarding the taxonomy and species separation for members of the genus, we rigorously analyse the morphology of select species, with emphasis on North American diversity. We corroborate the morphological results with a molecular multilocus phylogeny and provide authoritative DNA barcodes for the included species. We also calculate inter- and intraspecific genetic distances for cytochrome c oxidase subunit I as corroboration of the morphological analyses. Finally, we describe a new North American species that was evinced through our integrative taxonomic approach.

Materials and methods

Specimen collection and identification

In total, 65 specimens were collected during Royal Ontario Museum (ROM) field expeditions to the Canadian provinces of Quebec, Ontario, Manitoba, Saskatchewan, Alberta, and British Columbia in 2016 and 2018. In addition, a single specimen of *Theromyzon propinquum* (Ringuelet, 1947) was collected in western Chile during a trip led by the American

Museum of Natural History in 2016. The specimens were typically found attached to underwater vegetation, rocks or other debris, and attached to the exposed skin of the collectors. Most often, the specimens were spotted with the anterior ends of the bodies waving in the water, a behaviour that potentially maximises the chances of attaching to swimming waterfowl. Leeches were first relaxed by gradually increasing the concentration of ethanol to pond water (up to ~15% ethanol) and subsequently fixed in 95% ethanol. All specimens were stored at 4°C until further analysis and voucher specimens are permanently lodged in the Invertebrate Zoology collections at the ROM; Table 1 presents a list of the specimens and associated metadata, including GenBank accession numbers.

Specimens were initially identified based on diagnostic morphological traits for the genus (i.e. four pairs of eyes arranged in two parallel rows; Oosthuizen and Davies 1993), and later corroborated using specialised literature (e.g. Klemm 1985; Sawyer 1986; Oosthuizen and Davies 1993) following in-depth morphological analyses. External colouration patterns, number of annuli between gonopores, and the shape and size of reproductive structures were important characters for identifying specimens to the species level.

Imaging and morphological examination

External and internal morphology for all specimens was examined under a Leica Wild M10 dissecting microscope (Leica Microsystems, Richmond Hill, ON, Canada). Specimens were dissected ventrally for the examination of the internal morphology. Photographs of the specimens were taken using a Keyence VHX-6000 digital microscope (Keyence Corporation, Mississauga, ON, Canada). Key morphological differences were used to identify and separate specimens at the species-level using the original descriptions for each *Theromyzon* species.

DNA extraction, amplification and sequencing

To avoid potential contamination by ingesta of the leeches, tissue samples were cut from one side of the caudal sucker of each specimen. DNA was extracted from these samples using the Qiagen DNeasy Blood and Tissue extraction kit (Qiagen, Valencia, CA) following the manufacturer's protocol, except that DNA was eluted in 50 µL of Buffer AE. Thereafter, two mitochondrial loci (cytochrome c oxidase subunit I [COI] and 12S rRNA) and one nuclear locus (28S rRNA) were amplified using the following primers, for COI: LCO1490 (5' GGTCAACAAATCATAAAGATATTGG 3') and HCO2198 (5' TAAACTTCAGGGTGACCAAAAATCA 3') (Folmer et al. 1994); for 12S rRNA: 12S a (5' AACIIGGATTAGATACCC 3') and 12S b (5' GAGAGTGACGGGCGATGTGT 3') (Simon et al. 1994); for 28S rRNA: 28SA (5' GACCCGTCTGAAGCACG 3') (Whiting 2002) and 28SBout (5' CCCACAGCGCCAGTTTC TGCTTACC 3') (Schulmeister 2003). Polymerase Chain

Table I. List of specimens included in this study, respective collection numbers, locality and GenBank accession numbers.

Genus	Species	Voucher	Country	Province or state	Locality	Coordinates	COI	12S	28S
Ingroup									
<i>Theromyzon</i>	<i>bifarium</i>	ROMIZII0094	Canada	Ontario	Mijinemungshing Lake, Lake Superior Provincial Park	47.70277, -84.73026	OK586887	—	—
<i>Theromyzon</i>	<i>bifarium</i>	ROMIZII1276	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586886	OK032080	OK032085
<i>Theromyzon</i>	<i>bifarium</i>	ROMIZII2539	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586890	OK032079	OK032086
<i>Theromyzon</i>	<i>bifarium</i>	—					AY047330	—	—
<i>Theromyzon</i>	<i>pallens</i>	—	France	Brittany	Étang de La Musse, Paimpont,		AF003279	—	—
<i>Theromyzon</i>	<i>propinquum</i>	ROMIZII1688	Chile	Llanquihue	Lake by Gracias a la Vida Lodge, Pichilaguna	-41.2636, -73.055	OK586891	OK032084	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII0072	Canada	Ontario	Unnamed pond, Terra Cotta Conservation Area	43.7194778, -79.9598528	OK586870	—	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII0497	USA	Minnesota	Eagle Lake	44.2068833, -93.88975	OK586871	OK032075	OK032089
<i>Theromyzon</i>	<i>rude</i>	ROMIZII0501	USA	Minnesota	Eagle Lake	44.2068833, -93.88975	OK586868	—	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1440	Canada	Saskatchewan	Anglin Lake beach, Great Blue Heron Provincial Park	53.74263, -105.89503	OK586869	—	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1461	Canada	Saskatchewan	Henry's Pond, Great Blue Heron Provincial Park	53.7546, -105.87061	OK586867	—	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1463	Canada	Saskatchewan	Henry's Pond, Great Blue Heron Provincial Park	53.7546, -105.87061	OK586866	OK032076	OK032090
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1536	Canada	Manitoba	Unnamed pond, east of Lake Winnipeg	50.72285, -96.53915	OK586872	OK032074	OK032104
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1537	Canada	Manitoba	Unnamed pond, east of Lake Winnipeg	50.72285, -96.53915	OK586875	—	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1538i	Canada	Manitoba	Unnamed pond, east of Lake Winnipeg	50.72285, -96.53915	OK586874	OK032073	OK032102
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1538ii	Canada	Manitoba	Unnamed pond, east of Lake Winnipeg	50.72285, -96.53915	OK586873	OK032072	OK032103
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1601	Canada	Manitoba	Gull Lake, Duck Mountain Provincial Park	51.38466, -101.35353	OK586865	—	—
<i>Theromyzon</i>	<i>rude</i>	ROMIZII1606	Canada	Manitoba	Ashern	51.182222, -98.345556	OK586864	—	—

(Continued on next page)

Table I. (Continued)

Genus	Species	Voucher	Country	Province or state	Locality	Coordinates	COI	I2S	28S
<i>Theromyzon</i>	<i>rude</i>	—	Canada	Ontario	Lake of Two Rivers, Algonquin Provincial Park		AF003262	—	—
<i>Theromyzon</i>	sp.	ROMIZII1339	Canada	Alberta	Chump Lake	54.65085, -112.56135	OK586889	—	—
<i>Theromyzon</i>	<i>tessulatum</i>	ROMIZII1337	Canada	Alberta	Chump Lake	54.65085, -112.56135	OK586884	—	—
<i>Theromyzon</i>	<i>tessulatum</i>	ROMIZII1338	Canada	Alberta	Chump Lake	54.65085, -112.56135	OK586883	OK032078	OK032100
<i>Theromyzon</i>	<i>tessulatum</i>	ROMIZII12561	Canada	Alberta	Beaver Lake	54.7566667, -111.92145	OK586885	—	—
<i>Theromyzon</i>	<i>tessulatum</i>	ROMIZII12562	Canada	Alberta	Beaver Lake	54.7566667, -111.92145	OK586888	OK032077	OK032087
<i>Theromyzon</i>	<i>tessulatum</i>	ROMIZII12566	Canada	Alberta	Beaver Lake	54.7566667, -111.92145	OK586882	—	—
<i>Theromyzon</i>	<i>tessulatum</i>	—	France				AY047318	AF099957	AY425404
<i>Theromyzon</i>	<i>tigris</i> sp. nov.	ROMIZII1375	Canada	Alberta	Fish Lake	52.45356, -116.14535	OK586880	—	—
<i>Theromyzon</i>	<i>tigris</i> sp. nov.	ROMIZII1462	Canada	Saskatchewan	Henry's Pond, Great Blue Heron Provincial Park	53.7546, -105.87061	OK586878	—	—
<i>Theromyzon</i>	<i>tigris</i> sp. nov.	ROMIZII12569	Canada	Saskatchewan	Lake Namekus	53.8386667, -106.0471	OK586877	OK032082	OK032088
<i>Theromyzon</i>	<i>tigris</i> sp. nov.	ROMIZII12570	Canada	Saskatchewan	Lake Namekus	53.8386667, -106.0471	OK586881	OK032081	OK032099
<i>Theromyzon</i>	<i>tigris</i> sp. nov. — HOLOTYPE	ROMIZII12571	Canada	Saskatchewan	Lake Namekus	53.8386667, -106.0471	OK586879	OK032083	OK032101
<i>Theromyzon</i>	<i>tigris</i> sp. nov.	ROMIZII12948	Canada	Québec	Unnamed pond, Les Collines-de-l'Outaouais	45.6402, -75.83299	OK586876	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1243	Canada	Alberta	Elkwater Lake	49.66146, -110.2896	OK586846	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1244	Canada	Alberta	Elkwater Lake	49.66146, -110.2896	OK586851	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1250	Canada	Alberta	Big Island Lake	53.49371, -113.19718	OK586860	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1252	Canada	Alberta	Big Island Lake	53.49371, -113.19718	OK586845	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1254	Canada	Alberta	Big Island Lake	53.49371, -113.19718	OK586844	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1265	Canada	Alberta	Big Island Lake	53.49371, -113.19718	OK586849	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1273	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586862	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1274	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586843	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1286	Canada	Alberta	Long Lake	54.44325, -112.76613	OK586842	—	—
<i>Theromyzon</i>	<i>trizonare</i>	ROMIZII1311	Canada	Alberta	Unnamed pond outside Lakelands Provincial Park	54.77775, -111.711	OK586861	OK032071	OK032098

(Continued on next page)

Table I. (Continued)

Genus	Species	Voucher	Country	Province or state	Locality	Coordinates	COI	I2S	28S
<i>Theromyzon</i>	trizonare	ROMIZII1312	Canada	Alberta	Unnamed pond outside Lakelands Provincial Park	54.77775, -111.711	OK586841	OK032070	OK032091
<i>Theromyzon</i>	trizonare	ROMIZII1352	Canada	Alberta	Muir Lake	53.628, -113.95655	OK586855	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1353	Canada	Alberta	Muir Lake	53.628, -113.95655	OK586859	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1361	Canada	Alberta	Jackfish Lake	52.5018, -115.55861	OK586850	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1385	Canada	Saskatchewan	Cherry Lake	50.394, -103.65951	OK586848	OK032069	OK032094
<i>Theromyzon</i>	trizonare	ROMIZII1394	Canada	Saskatchewan	Unnamed pond on Beaver Lake Trail, Moose Mountain Provincial Park	49.83891, -102.31053	OK586847	OK032068	OK032092
<i>Theromyzon</i>	trizonare	ROMIZII1398	Canada	Saskatchewan	Unnamed pond, Moose Mountain Provincial Park	49.81298, -102.33606	OK586856	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1399	Canada	Saskatchewan	Unnamed pond, Moose Mountain Provincial Park	49.81298, -102.33606	OK586854	OK032067	OK032095
<i>Theromyzon</i>	trizonare	ROMIZII1400	Canada	Saskatchewan	Unnamed pond, Moose Mountain Provincial Park	49.81298, -102.33606	OK586853	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1409	Canada	Saskatchewan	Pipestone Creek, Moosomin Regional Park	50.07595, -101.70761	OK586858	OK032065	OK032097
<i>Theromyzon</i>	trizonare	ROMIZII1410	Canada	Saskatchewan	Pipestone Creek, Moosomin Regional Park	50.07595, -101.70761	OK586840	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1485	Canada	Saskatchewan	Cypress Lake West	49.47253, -109.56998	OK586852	—	—
<i>Theromyzon</i>	trizonare	ROMIZII1486	Canada	Saskatchewan	Cypress Lake West	49.47253, -109.56998	OK586857	OK032066	OK032096
<i>Theromyzon</i>	trizonare	ROMIZII2540	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586839	—	—
<i>Theromyzon</i>	trizonare	ROMIZII2541	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586838	—	—
<i>Theromyzon</i>	trizonare	ROMIZII2542	Canada	Alberta	Half Moon Lake	53.45585, -113.0835	OK586863	OK032064	OK032093
<i>Theromyzon</i>	trizonare	ROMIZII2565	Canada	Alberta	Beaver Lake	54.7566667, -111.92145	OK586837	—	—
Outgroup									
<i>Alboglossiphonia</i>	heteroclitia	—					AF116016	AF099955	—
<i>Calliobdella</i>	vivida	—					AF003260	AY425409	AY425360
<i>Glossiphonia</i>	complanata	—	United Kingdom				AF003277	AY425414	AY425371

(Continued on next page)

Table I. (Continued)

Genus	Species	Voucher	Country	Province or state	Locality	Coordinates	COI	12S	28S
<i>Haementeria</i>	<i>acuecueyetzin</i>	—	Mexico	Veracruz			JN850901	JN850863	JN850887
<i>Haementeria</i>	<i>acuecueyetzin</i>	—	Mexico	Veracruz	Catemaco		JN850899	JN850862	JN850885
<i>Haementeria</i>	<i>depressa</i>	—	Uruguay	Maldonado			JN850902	JN850864	JN850888
<i>Haementeria</i>	<i>depressa</i>	—	Uruguay	Jaureguiberry			JN850903	JN850865	JN850889
<i>Haementeria</i>	<i>ghiliani</i>	—	French Guiana		BioPharm Leeches		AF329035	AY425417	AY425374
<i>Haementeria</i>	<i>gracilis</i>	—	Uruguay	Arroyo Espinas			AF329034	AY425418	AY425375
<i>Haementeria</i>	<i>lopezi</i>	—	Mexico	Colima			JN850898	JN850861	JN850884
<i>Haementeria</i>	<i>officinalis</i>	—	Mexico	Michoacan			JN850906	JN850868	JN850891
<i>Haementeria</i>	<i>paraguayensis</i>	—	Argentina	Chaco			JN850908	JN850870	JN850893
<i>Helobdella</i>	<i>paranensis</i>	—					AF329037	AY425412	AY425363
<i>Helobdella</i>	<i>stagnalis</i>	—	France				AF329041	AY425424	AY425382
<i>Marsupiobdella</i>	<i>africana</i>	—					AF116015	AY425433	AY425392
<i>Oligobdella</i>	<i>brasiliensis</i>	—	Brazil				JN850911	JN850874	JN850895
<i>Placobdella</i>	<i>biannulata</i>	—	USA	North Carolina			AF116021	AY425435	AY425397
<i>Placobdella</i>	<i>parasitica</i>	—	Canada	Ontario			AF003261	AY425438	AY425401
<i>Placobdella</i>	<i>picta</i>	—					AF116020	AY425413	AY425364
<i>Pontobdella</i>	<i>muricata</i>	—					KY659072	AF099958	KY659073

Reactions (PCR) were conducted in 25- μ L volumes, each sample containing 16.39 μ L of ddH₂O; 2.5 μ L of buffer; 2.5 μ L of MgCl₂; 1 μ L of each primer at 10- μ M concentration; 0.56 μ L of dNTPs at 10-mM concentration; 0.05 μ L of Platinum Taq polymerase (Invitrogen, Carlsbad, CA); and 1 μ L of total genomic DNA. The thermocycler protocol was as follows: for *COI*, denaturation at 94°C (1 min), five rounds of 94°C (30 s) 40°C (40 s) 72°C (1 min), 35 rounds of 94°C (30 s), 46°C (40 s), 72°C (1 min), followed by extension at 72°C (5 min); for *12S* rRNA, 94°C (2 min), followed by 32 rounds of 94°C (30 s), 48°C (30 s), 72°C (45 s), with subsequent extension at 72°C for 7 min; for *28S* rRNA, 94°C for 5 min, subsequently 39 rounds of 95°C (1 min), 52°C (1 min), 72°C (1 min), with a final extension at 72°C for 7 min. PCR products were run on a 1% agarose gel and successful amplifications were cleaned with Exo-SapIT (Affymetrix, Santa Clara, CA, USA) according to manufacturer's protocol. Cycle sequencing was performed using 3 μ L of PCR product to 0.5 μ L each of ABI Big Dye Terminator (ver. 3.1, Applied Biosystems) and Big Dye 5 \times Sequencing Buffer (Applied Biosystems, Carlsbad, CA, USA) and 2 μ L of primer at 10- μ M concentration. The samples were denatured at 96°C for 1 min followed by 30 cycles of 96°C (10 s), 50°C (5 s) and 60°C (4 min). Sequences were obtained using an ABI PRISM 3730 DNA Analyser at the ROM, and were assembled and edited using Geneious (ver. 11.1.5, Biomatters Ltd, Auckland, New Zealand, see www.geneious.com; Kearse *et al.* 2012).

Alignment and phylogenetic analysis

To guide subsequent sequencing efforts, we initially generated a *COI* gene tree for all specimens of *Theromyzon*. Guided by the resulting tree, the multilocus matrix (concatenated with Mesquite, ver. 3.11, W. P. Maddison and D. R. Maddison, see <https://www.mesquiteproject.org>) contained 23 of the 66 newly collected *Theromyzon* specimens with multiple representatives of each major clade in the *COI* gene tree; these specimens were further sequenced for *12S* rRNA and *28S* rRNA (Table 1). The multilocus dataset was also augmented by the inclusion of 4 *COI*, 1 *12S* and 1 *28S* sequence for *Theromyzon* specimens, and 27 outgroup sequences, from GenBank (Table 1). All sequences were aligned using the MAFFT (ver. 7, see <https://mafft.cbrc.jp/alignment/software>; Katoh and Standley 2013) plugin for Geneious, applying default settings. All trees were rooted at *Calliobdella vivida* (Verrill, 1872) and *Pontobdella muricata* (Linnaeus, 1758) (Piscicolidae) following the results of Trontelj *et al.* (1999) and Tessler *et al.* (2018).

We employed three optimality criteria for the phylogenetic analyses: parsimony, maximum likelihood (ML) and Bayesian inference (BI). For parsimony, TNT (ver. 1.5, see www.lillo.org.ar/phylogeny/tnt; Goloboff and Catalano 2016) used a New Technology search for 1000 replications with five rounds of ratcheting and five rounds of tree fusing after the initial

Wagner tree builds. Top scoring trees were subsequently subjected to branch swapping using the 'bbreak' command and support values were estimated by 1000 rounds of standard bootstrapping applying default settings.

For the ML and BI trees, the optimal models and partitioning scheme were determined using the 'TESTMERGE' option in IQ-TREE (ver. 1.6, see <http://www.iqtree.org>; Nguyen *et al.* 2015), with each locus treated as a potential partition. Results suggested a general time reversible model, with a discrete gamma model with four rate categories, and allowing for invariable sites (GTR + F + I + G4) for *COI*; discrete gamma model with four rate categories, not allowing for invariable sites (GTR + F + G4) for *12S* rRNA; and a Kimura two parameter (K2P) model for *28S* rRNA. For both the concatenated dataset and the *COI* dataset, the ML tree was produced in IQ-TREE using a tree search with 1000 replicates with default settings and 1000 pseudoreplicates were used to calculate bootstrap support values.

The BI tree for the concatenated dataset was generated using MrBayes (ver. 3.2.6, see <http://nbisweden.github.io/MrBayes>; Huelsenbeck and Ronquist 2001) on the CIPRES Science Gateway platform (see <https://www.phylo.org>; Miller *et al.* 2010) using 1 060 000 generations with two simultaneous runs, each with four chains, and the same models and partitions as indicated by IQ-TREE. Relative burn-in was set at 25% and the analysis was terminated once the average standard deviation of split frequencies went below 0.01. Tracer (ver. 1.6, see <https://beast.community/tracer>; Bouckaert *et al.* 2014; Rambaut *et al.* 2018) was used to ensure that the simulation had reached stationarity and that the runs had converged.

MEGAX (ver. 10.1.8, see <https://www.megasoftware.net>; Stecher *et al.* 2020) was used to compute uncorrected p-distances. The analyses included all *COI* sequences for *Theromyzon*, with pairwise deletion of gaps. Three sets of values were calculated: average intraspecific pairwise distance; average interspecific pairwise distance; and pairwise distance between each species and sister clade.

All specimens in the present study were collected under permits issued by Manitoba Conservation and Water Stewardship (permit number PP-PHQ-16-009), Saskatchewan Ministry of Parks, Culture and Sport, and Alberta Fish and Wildlife Policy Branch (collection licence 55242, research permit 55243).

Results

Species description

Some of the collected specimens did not conform to any known species description, yet showed distinct morphological features and formed a unique clade in both the multilocus and *COI* trees (see Phylogenetic analyses section below). Here we present the description of this new species, coupled with figures for both internal and external morphology.

Systematics

Order HIRUDINIDA sensu Siddall, Apakupakul, Burreson, Coates, Erséus, Gelder, Källersjö & Trapido-Rosenthal, 2001

Suborder GLOSSIPHONIIFORMES Tessler & de Carle, 2018

Family GLOSSIPHONIIDAE Vaillant, 1890

Genus *Theromyzon* Philippi, 1867

***Theromyzon tigris*, sp. nov.**

(Fig. 1, 2)

Diagnosis

Total body length ranging between 10.9 and 12.5 mm; chromatophores and unpigmented spots scattered on venter and dorsum; gonopores separated by two primary annuli; single female gonopore; 12 pairs of crop caeca; male atrium spherical, reduced; atrial cornuae present but reduced; 6 pairs of testisacs; common oviduct present.

Type locality. Namekus Lake, Saskatchewan, Canada (53.8386667, -106.0471).

Material examined

HOLOTYPE (ROMIZI12571). Dissected, full body specimen, fixed in 95% ethanol, collected on August 5, 2018 from type locality. Paratypes ($n = 4$) all dissected, full body specimens, fixed in 95% ethanol. ROMIZI12569 and ROMIZI12570 collected on 5 August 2018 from type locality; ROMIZI11462 collected on 19 June 2016 from Henry's

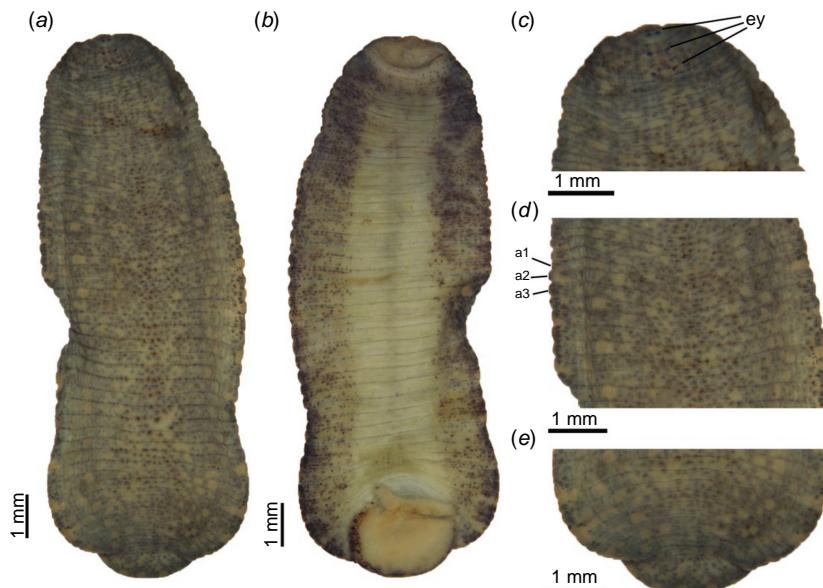


Fig. 1. External morphology of *Theromyzon tigris* sp. nov. (ROMIZI12571). (a) Dorsal view. (b) Ventral view. (c) Dorsal view of the nuchal region. (d) Dorsal view of mid-body somites. (e) Ventral view of the posterior region of the body. Abbreviations: ey, eyespots; a1–a3, primary annuli a1–a3.

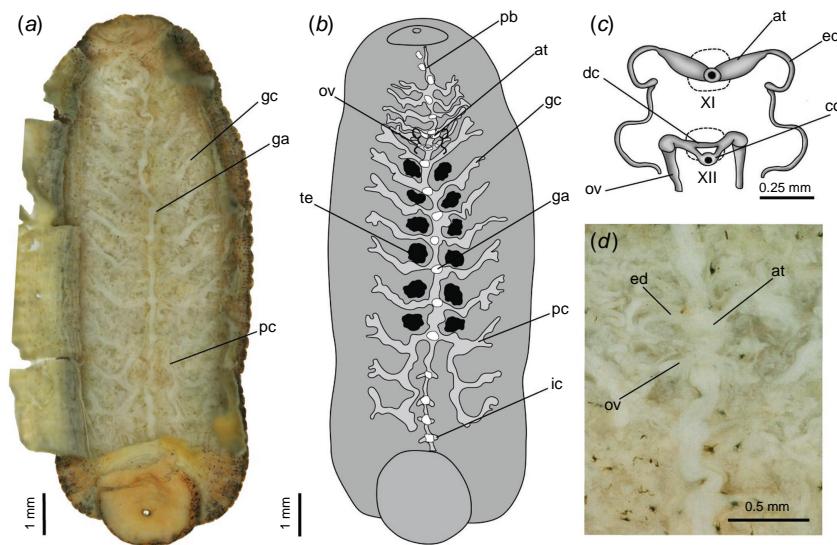


Fig. 2. Internal morphology of *Theromyzon tigris* sp. nov. (ROMIZI12571). (a) Ventral dissection showing the digestive system. (b) Ventral view of the internal morphology. (c) Ventral view of the male and female reproductive systems. (d) Ventral dissection showing the male and female reproductive systems. Abbreviations: gc, gastric caeca; ga, ganglia; pc, post-caeca; ov, ovisacs; te, testisacs; pb, proposic; at, male atrium; ic, intestinal caeca; dc, connecting ductile; ed, ejaculatory ducts; co, common oviduct.

pond, Great Blue Heron Provincial Park, Saskatchewan, Canada (53.7546, -105.87061); ROMIZ11375 collected on 3 July 2016 from Fish Lake, Alberta, Canada (52.45356, -116.14535).

Description

Based on entire type series. Body length 10.9–12.5 mm; maximum body width 3.3–4.2 mm; body sublanceolate, dorsoventrally flattened (Fig. 1a, b); anterior sucker well developed, ovoid, mouth pore located at the anterior border of the sucker (Fig. 1b); head circular (Fig. 1c); posterior sucker circular, ventrally directed, not pedunculated, maximum width 2.14 mm (Fig. 1b). Somites I and II uniannulate (Fig. 1c); somite III biannulate (Fig. 1c); somites IV–XXVII triannulate on both dorsum and venter (Fig. 1d: a1, a2, a3); one post anal annulus (Fig. 1e). Dorsum dark, from greyish-brown to dark olive green (Fig. 1a); pigmentation formed by scattered chromatophores and scattered lighter (cream to yellow) spots (4–8 spots on each annulus) sometimes fused to form larger blotches (Fig. 1a, d); lateral edge of a2 unpigmented on both dorsum and venter (Fig. 1a, d). Eyespots four pairs (Fig. 1c: ey), somewhat triangular, inner paramedial position, first pair close to each other on somite II, second pair on IIIa2 + a3 separated by half of the diameter of an eyespot, third pair on IVa2 separated by the diameter of an eyespot, fourth pair on Va2/a3 separated by the diameter of four eyespots; first pair of eyespots one-quarter the size of the fourth pair. Male gonopore inconspicuous, in furrow at XIa2/a3; single female gonopore in furrow at XIIa1/a2; gonopores separated by 2 annuli.

Pharynx protrusible (proboscis), short, linear, with base at VII (Fig. 2a, b: pb); salivary glands diffuse, visible from VI to IX; common salivary ducts absent; oesophagus short (Fig. 2a, b), limited to VI/VII. Twelve pairs of crop caeca, branched at distal end, first pair at VIII, last pair forming a post caeca from XIX to XXVII (Fig. 2a, b: gc, pc); 4 pairs of intestinal caecae (Fig. 2a, b: ic); atrium posterior to ganglion in XI, reduced, spherical with only reduced atrial cornuae (Fig. 2c, d: at). Ejaculatory ducts dorsolaterally directed in proximal region (Fig. 2c, d: ed), proximal region thicker, followed by thinner descendant portion from XI/XII to XII/XIII, diameter of ejaculatory ducts changes abruptly; descendant portion of the ejaculatory ducts forming loops towards the medial line of the body. Six pairs of testisacs, first pair at XIII/XIV, last pair at XVIII/XIX (Fig. 2b: te); ovisacs at XII (Fig. 2d: ov), short, simple, connection between paired ovisacs absent, common oviduct present.

Etymology

Theromyzon tigris sp. nov. is named for the dorsal pigmentation pattern. The chromatophores and unpigmented blotches form a pattern reminiscent of a tiger's stripes. The specific epithet is, fittingly, the name the ancients used for this barred feline.

Remarks

Theromyzon tigris sp. nov. can be distinguished from other species within the genus by the unique combination of five characters. *T. maculosum* and *T. trizonare* can be distinguished from the new species by the presence of a common oviduct in *T. tigris* sp. nov. and the number of female gonopores; whereas both *T. maculosum* and *T. trizonare* possess two female gonopores, *T. tigris* sp. nov. only possesses a single pore. Female and male gonopores are separated by two annuli in *Theromyzon tigris* sp. nov., and this contrasts with *T. trizonare* and *T. tessulatum* in which gonopores are separated by three and four annuli respectively. Fourthly, in *T. tessulatum* and *T. bifarium*, the ejaculatory ducts are attached to the atrium in a dorso-postero-lateral position, whereas the same structures are attached in a dorsomedial position in *Theromyzon tigris* sp. nov. The new species can be distinguished from *T. rude* by virtue of possessing atrial cornuae, even if somewhat reduced; these structures are lacking in *T. rude*.

Using similar characters, the new species is also distinguishable from other congeners. According to Ringuelet (1985), the gonopores of the South American species *T. propinquum* are separated by three annuli and the crop is formed by 11 pairs of gastric caecae. In *Theromyzon tigris* sp. nov., the gonopores are separated by two annuli and the crop is composed of 12 gastric caecae. The original description of *Theromyzon pallens* mentions only general body shape and colour pattern, and lacks any diagnostic characters (Philippi 1867). Sawyer (1986) highlights that this species is not well differentiated from *T. tessulatum* and *T. propinquum*, and that previous records of *T. tessulatum* in South America by Oka (1930), Moore (1911) and Blanchard (1893) probably constitute records of *T. pallens*. Ringuelet (1985) also recognises the lack of a proper morphological circumscription for each of *T. pallens*, *T. tessulatum* and *T. propinquum*. The number of annuli between male and female gonopores distinguishes *Theromyzon tigris* sp. nov. from the three Palearctic species *T. tessellatoides* (Livanow, 1902), *T. mollissimum* (Grube, 1871) and *T. mathaii* Bhatia, 1939. Whereas the gonopores are separated by five annuli in *T. tessellatoides* and *T. mollissimum*, these are separated by three annuli in *T. mathaii*. However, in both *T. cooperi* (Harding, 1932) and *T. garjaewi* (Livanow, 1902), as in *Theromyzon tigris* sp. nov., the gonopores are separated by two annuli. Despite this, *Theromyzon garjaewi* can be differentiated from *Theromyzon tigris* sp. nov. by the absence of a common oviduct (Livanow 1902). *Theromyzon tigris* sp. nov. can be differentiated from *T. cooperi* by a larger atrium in *T. cooperi* (Oosthuizen 1993).

Known distribution

Theromyzon tigris sp. nov. is known only from lakes and ponds in the Canadian provinces of Alberta and Saskatchewan (see Table 1). Specimens were generally

found in water bodies with mixed sandy and rocky bottoms, along with some vegetation. Specimens were collected from submerged rocks and vegetation along the edge of the water.

Diagnoses for other species

Imagery for external and internal morphology for representatives of each clade in the multilocus phylogeny is presented in Fig. 3–6. Key diagnostic characters are summarised in Table 2. The morphological analyses focused on the external pigmentation patterns and the number of annuli between gonopores, and the shape and size of the internal reproductive organs. The gonopores in *Theromyzon* are markedly small when compared to leeches of other genera and, for

this reason, counting the number of annuli between gonopores from an external viewpoint was occasionally difficult; in such cases, the number of annuli was calculated based on the internal placements of the male and female pores. The diminutive size of the single *Theromyzon propinquum* specimen prohibited inspection of the internal morphology. In addition, the resulting phylogeny evinced a lineage that showed enough genetic variation to warrant investigations into the identity of the specimen (ROMIZI11339). Unfortunately, due to the minute size of the specimen, we could not provide a reliable identity based on morphology. More specimens are needed to accurately identify this lineage and we refrain from making any inferences to this effect herein.

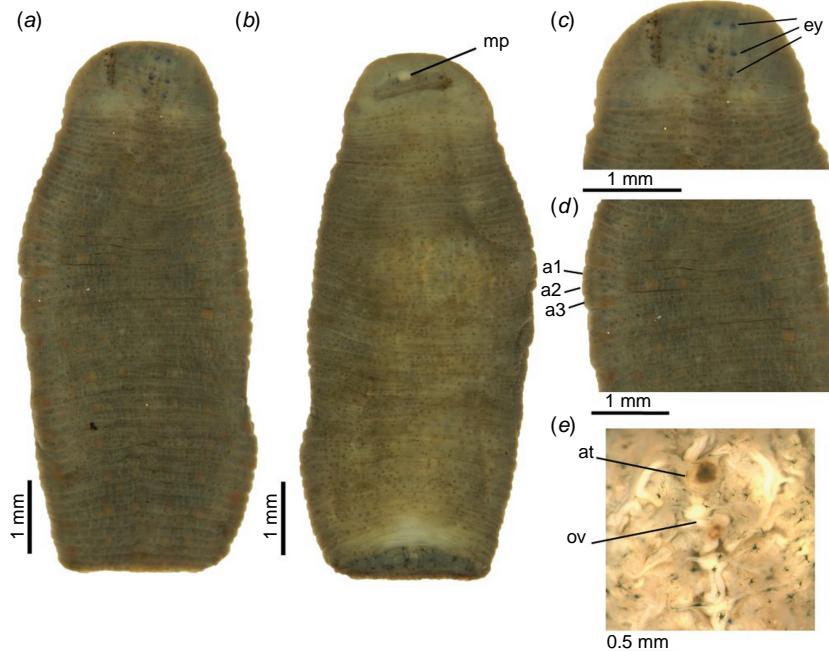


Fig. 3. External and internal morphology of *Theromyzon bifarium*. (a) Dorsal view (ROMIZII1276). (b) Ventral view (ROMIZII1276). (c) Dorsal view of the nuchal region (ROMIZII1276). (d) Dorsal view of mid-body somites (ROMIZII1276). (e) Ventral dissection showing the male and female reproductive systems (ROMIZII12539). Abbreviations: mp, mouth pore; ey, eyespots; a1–a3, primary annuli a1–a3; at, male atrium; ov, ovisacs.

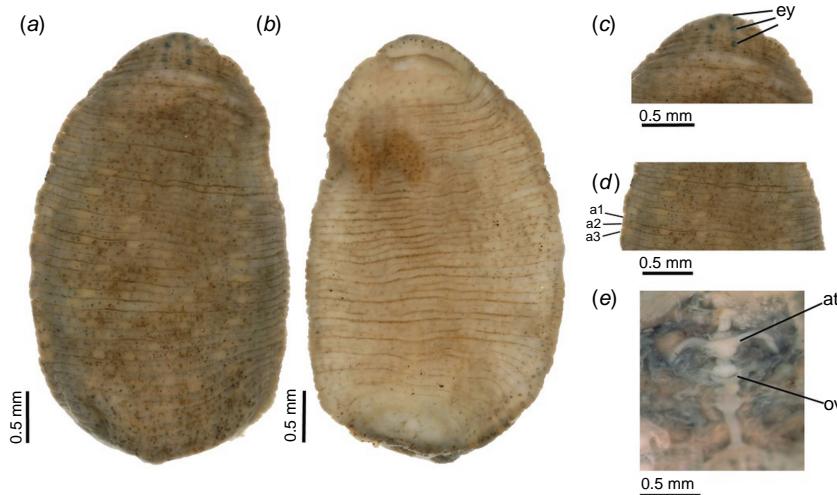


Fig. 4. External and internal morphology of *Theromyzon tessulatum*. (a) Dorsal view (ROMIZII1338). (b) Ventral view (ROMIZII1338). (c) Dorsal view of the nuchal region (ROMIZII1338). (d) Dorsal view of mid-body somites (ROMIZII1338). (e) Ventral dissection showing the male and female reproductive systems (ROMIZII12562). Abbreviations: ey, eyespots; a1–a3, primary annuli a1–a3; at, male atrium; ov, ovisacs.

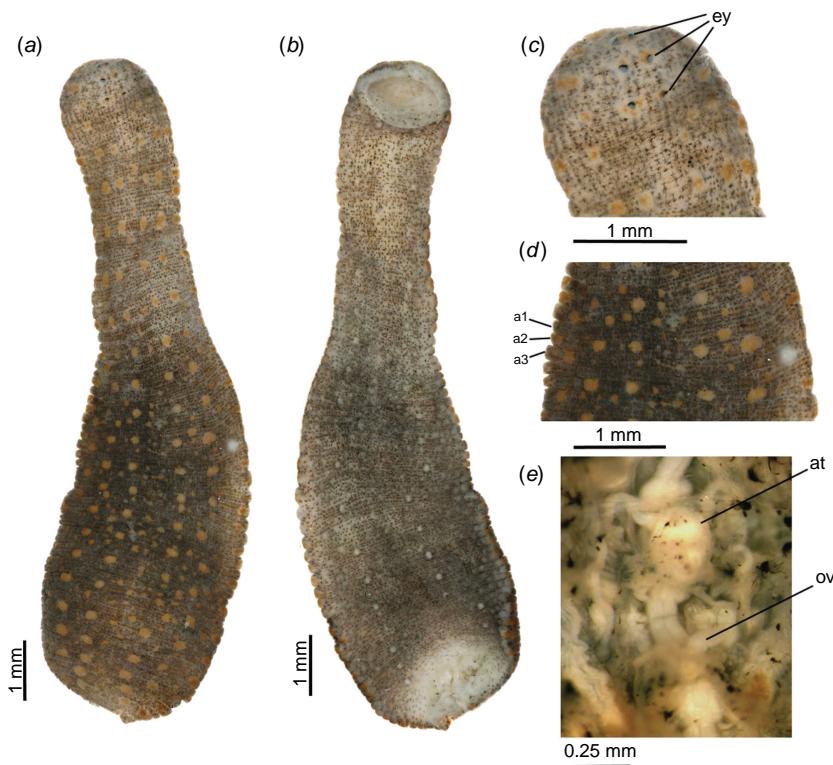


Fig. 5. External and internal morphology of *Theromyzon rude* (ROMIZII 1463). (a) Dorsal view. (b) Ventral view. (c) Dorsal view of the nuchal region. (d) Dorsal view of the mid-body somite. (e) Ventral dissection showing the male and female reproductive systems. Abbreviations: ey, eyesposts; a1–a3, primary annuli a1–a3; at, male atrium; ov, ovisacs.

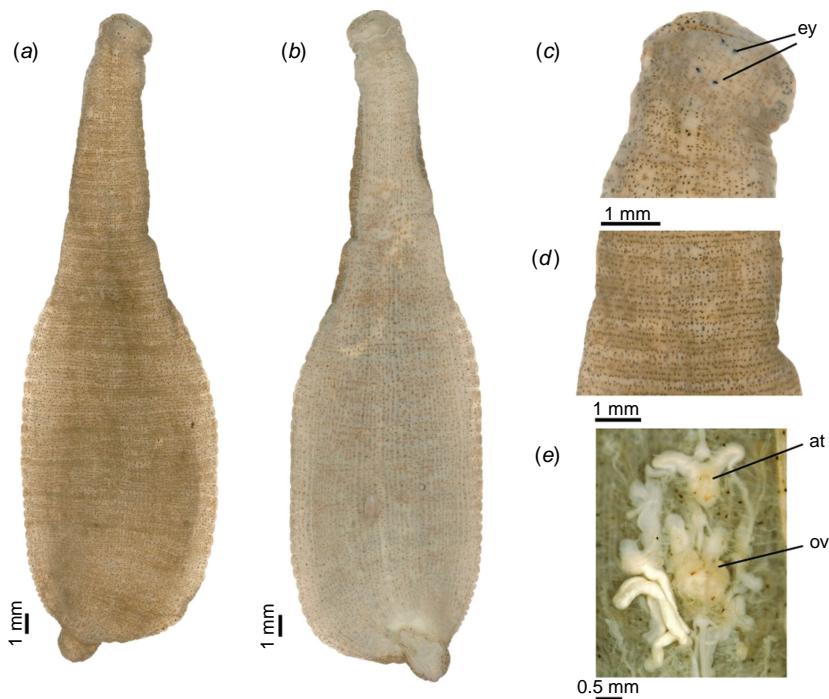


Fig. 6. External and internal morphology of *Theromyzon trizonare* (ROMIZII 1486). (a) Dorsal view. (b) Ventral view. (c) Dorsal view of the nuchal region. (d) Dorsal view of mid-body somites. (e) Ventral dissection showing male and female reproductive systems. Abbreviations: ey, eyesposts; at, male atrium; ov, ovisacs.

Theromyzon bifarium Oosthuizen & Davies, 1993

Diagnosis based on 2 specimens; total body length ranging between 7.5 and 9.7 mm. Eyes in two rows with one pair on each of somites II, III, IV and V (Fig. 3a, c, ey). Dorsal base colouration greenish-brown with four rows of beige papillae

throughout length of dorsum (Fig. 3a, d). Ventrum light brown to cream-coloured (Fig. 3b). Specimens with two annuli between gonopores. Male atrium slightly larger than female atrium, extending anteriorly to form weakly developed atrial cornuae (Fig. 3e, at). Ovaries with common, looped oviduct (Fig. 3e; ov) and with single female gonopore.

Table 2. Morphological comparison of diagnostic characters of North American species. Diagnostic characters for *T. maculosum* was obtained from (Oosthuizen and Davies 1993).

	Annuli between male and female gonopores	Number of female gonopores	Entrance of male ducts to atrium	Atrial cornua	Common oviduct
<i>T. maculosum</i>	2	2	Dorsolateral	Present	Absent
<i>T. bifarium</i>	2	1	Dorsolateral	Present	Present
<i>T. rude</i>	2	1	Dorsomedial	Absent	Present
<i>T. trizonare</i>	3	2	Dorsolateral	Present	Absent
<i>T. tessulatum</i>	4	1	Dorso-postero-lateral	Present	Present
<i>T. tigris</i> sp. nov.	2	1	Dorsolateral	Reduced	Present

Theromyzon tessulatum (Müller, 1774)

Diagnosis based on 2 specimens; total body length ranging between 4.1 and 7.5 mm. Eyes in two rows with one pair on each of somites II, III, IV and V (Fig. 4a, c, ey). One specimen (ROMIZI12562) lacking (or with inconspicuous) pigmentation patterns on dorsum, possibly due to preservation. Other specimen (ROMIZI11338) with light brown speckled base colouration on dorsum and with faint, light orange spots randomly interspersed (Fig. 4a, d). Venter uniformly light brown (Fig. 4b). Gonopores separated by four annuli (only visible by removal of the epidermis in one specimen). Male atrium slightly larger than female atrium, extending anteriorly to form weakly developed atrial cornuae (Fig. 4e, at). Ovaries with common oviduct and single female gonopore (Fig. 4e). Ovaries extend laterally and anteriorly, wrapping around ventral nerve cord (Fig. 4e).

Theromyzon rude (Baird, 1863)

Diagnosis based on 5 specimens; total body length ranging between 6.3 and 12 mm. Eyes in two rows with one pair on each of somites II, III, IV and V (Fig. 5a, c, ey). Specimens heavily dorsoventrally flattened with oblong, soft and gelatinous appearance. Anterior body narrower than posterior end (Fig. 5a, b). Dorsum with light to dark brown base colouration and orange specks forming four rows longitudinally (Fig. 5a, d). Venter grey with small black specks and white papillae forming two rows longitudinally (Fig. 5b). Gonopores only evident when epidermis removed; gonopores separated by two annuli. Male atrium much larger than female atrium and male atrium oval-shaped with gonopore at anterior end (Fig. 5e, at). Ovaries with common oviduct and single female gonopore (Fig. 5e, ov).

Theromyzon trizonare Davies & Oosthuizen, 1993

Diagnosis based on 8 specimens; total body length ranging between 14.9 and 35.2 mm. Eyes in two rows with one pair on each of somites II, III, IV and V (Fig. 6a, c, ey).

Specimens extremely dorsoventrally flattened, appearing almost translucent at lateral edges of body and with very gelatinous appearance. Epidermis very thin with crop caeca visible through skin on ventral surface. Body shape oblong with posterior end wider than anterior end (Fig. 6a, b). Dorsal base colouration light, appearing beige with brown specks in a random pattern (Fig. 6a, d). Gonopores readily observable (except in ROMIZI11394); with three annuli separating gonopores. Reproductive organs enlarged relative to other species and to body size. Male atrium extending anteriorly and laterally, forming distinct atrial cornuae (Fig. 6e, ac). Ovaries without common oviduct (Fig. 6e, av), each ending in own distinct gonopore (i.e. two female gonopores present).

Phylogenetic analyses

The ML COI gene tree had a $\ln L$ of $-11\ 285.547$ and suggested the presence of at least seven major clades (Supplementary Fig. S1). We were unable to identify the specimen forming the unidentified eighth clade in Supplementary Fig. S1 as this was a juvenile and as such the diagnostic morphological characters were not yet developed or identifiable. Insofar as this gene tree was chiefly used to guide further sequencing efforts, we will not extensively discuss the topology here, other than to mention that *Theromyzon* as a whole, and each of the putative species in the dataset, were recovered as monophyletic.

For the multilocus dataset, the three optimality criteria produced trees with almost identical topologies, although the parsimony tree was less resolved in some places. The best scoring ML tree for the multilocus analysis ($\ln L = -11\ 754.981$) is shown in Fig. 7, and the parsimony and BI trees are shown in Supplementary Files S1, S2. The analyses recovered *Theromyzon* as monophyletic with high support (likelihood bootstrap support (LBS) = 97%; parsimony bootstrap support (PBS) = 82%; and posterior probability (PP) = 0.99). *Theromyzon propinquum* placed as the sister taxon (LBS = 88%; PBS = 75%; PP = 0.99) to a clade formed by North American specimens of *T. bifarium* and *T. tessulatum* (LBS = 98%; PBS = 100%; PP = 1.00). In turn, this clade is the sister group to *T. tigris* sp. nov. (LBS = 93%; PBS = 74%;

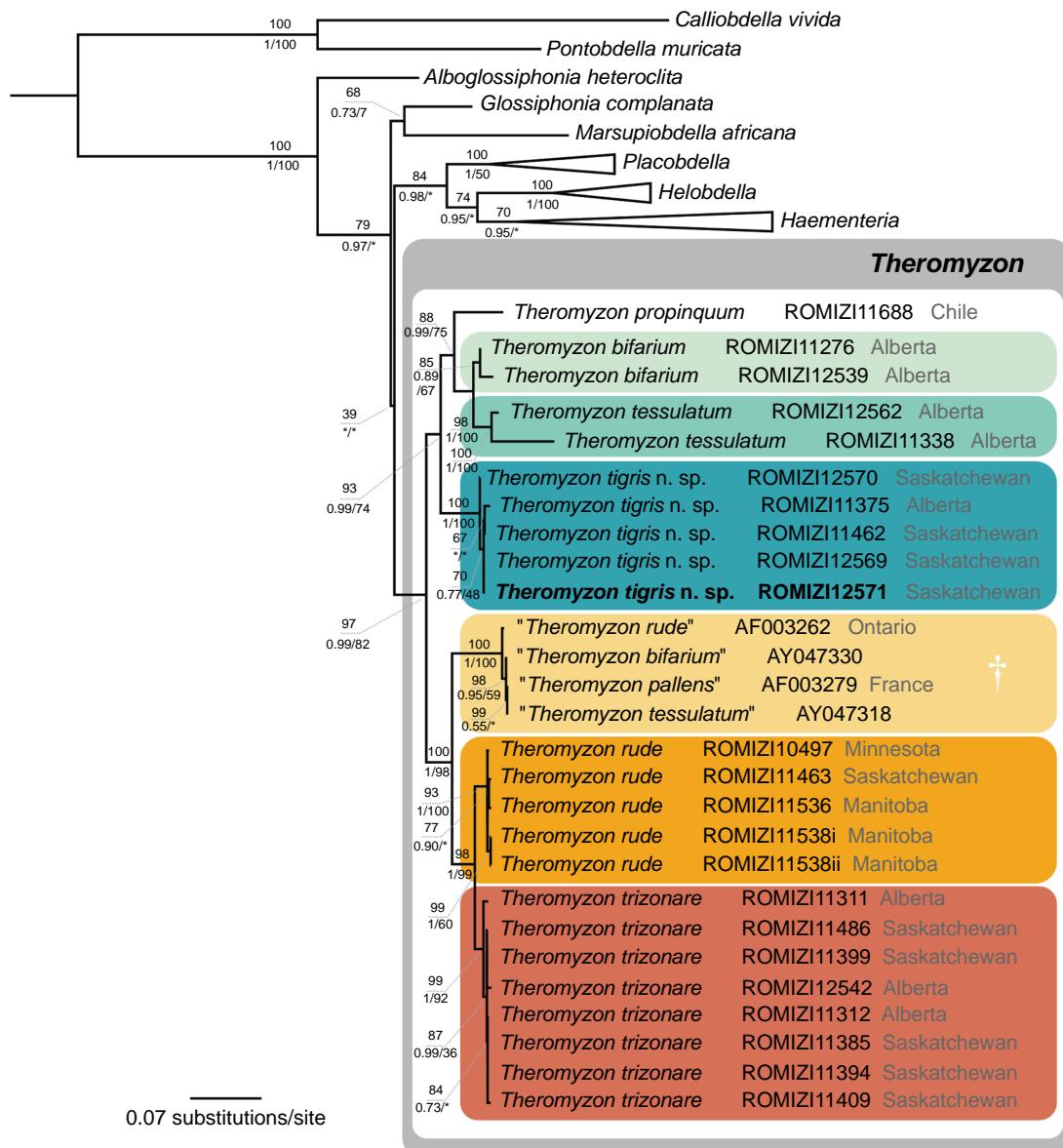


Fig. 7. Maximum likelihood phylogeny inferred from the multilocus dataset (COI, 12S rRNA, and 28S rRNA) ($\ln L = -11754.981$). Likelihood bootstrap support values are noted above each node; values below nodes are – from left to right – Bayesian posterior probability and parsimony bootstrap values. Asterisks denote nodes that were not present in the Bayesian or parsimony trees. Coloured boxes delimit species-level clades. ROM specimen numbers and collecting localities are listed next to tips where available and relevant. Holotype specimen for *Theromyzon tigris* sp. nov. (ROMIZI12571) is denoted by bold formatting. The dagger indicates the clade comprising sequence data downloaded from GenBank (accession numbers listed). Species identifications for these clades are further discussed in the text (see Discussion).

PP = 0.99). This larger clade forms the sister group to the remaining congeneric specimens (LBS = 100%; PBS = 98%; PP = 1.00) that are split into three species-level clades. The first is represented by four *Theromyzon* sequences from GenBank (each with a different taxonomic label) that form a clade with maximum support in all analyses, yet with astonishingly short branch lengths given that these supposedly represent four different species (but see Discussion).

The second and third clades are represented by *T. rude* and *T. trizonare* respectively. These place as sister taxa to each other (LBS = 98%; PBS = 99%; PP = 1.00) and in turn, as sister to the clade consisting of GenBank sequences.

The COI pairwise distance analyses showed an average interspecific distance of 9.8% between species of *Theromyzon*, with a minimum value of 4.3% and a maximum value of 12.4%. The average intraspecific distance was 0.9%, with a

minimum distance of 0.4% and a maximum distance of 2.1%. For reference, the average pairwise distance between each species of *Theromyzon* and the sister clade were 6.6%; the minimum distance was 4.3% and the maximum distance was 8.0%. Supplementary Table S1 summarises the results of the pairwise distance analyses.

Discussion

Through detailed morphological examinations, this paper aids in the circumscription of North American species of the leech genus *Theromyzon* and presents diagnoses for each of the studied species that might ease future taxonomic work. In addition, we present the first phylogenetic analysis based on molecular data for all but one of the previously recorded North American species. Both molecular and morphological data suggest that there are seven identifiable species in our dataset, six of which are predominantly North American. By virtue of forming unique clades and having distinct morphological features, the following North American species can be identified in our dataset: *T. trizonare*, *T. bifarium*, *T. rude*, *T. tessulatum* and *T. tigris* sp. nov.; the latter species is described by the present study. Aside from *T. propinquum*, the remaining species-level taxon is composed solely of GenBank sequences from both North America and Europe (notably, with four different taxonomic labels) such that determining which species these represent is difficult (see below). Whereas almost half of the currently recognised *Theromyzon* species worldwide (Oosthuizen and Davies 1993) are represented in our analyses, our study focuses on North American diversity and is almost completely devoid of data for taxa outside of this continent. There are three exceptions to this: our Chilean specimen of *T. propinquum* and GenBank sequences for *T. pallens* (AF003279) and *T. tessulatum* (AY047318), both specimens collected in France (Siddall and Burreson 1998; Borda and Siddall 2004).

The phylogenetic placement of the Chilean and French specimens deserves some further attention. These two sequences, along with two others downloaded from GenBank (AF003263: *T. rude* and AY047330: *T. bifarium*), form a clade. Given the very short branch lengths (in some cases zero) separating these terminals in the trees, these seem highly likely to represent different species from those reported from the identifications. Without studying the specimens from which the sequences were derived, drawing any conclusions as to the identities is impossible. However, for the reason that *T. rude*, *T. bifarium* and *T. tessulatum* are present in other, separate clades in our trees, with identities confirmed through morphological analyses, the likelihood that this clade represents any of those species is very small. If, indeed, the clade represents *T. pallens*, this is significant, as this species has never formally been recorded from North America; yet two of the four specimens in this clade are

North American in origin. Additionally, as our collection efforts failed to produce any specimens for *Theromyzon maculosum* and as this species has been recorded from North America, this ambiguous clade possibly represents *T. maculosum*. Owing to the odd placements of these specimens and the complete absence of morphological data, we refrain from drawing any firm conclusions regarding species-level identities.

Our phylogenetic analyses recovered a clade represented by *T. propinquum*, *T. bifarium* and *T. tessulatum*. Whereas *T. propinquum* was collected in Chile, the two remaining species were collected in North America (specifically in the Canadian province of Alberta). *T. bifarium* (a North American species) and *T. tessulatum* (a Holarctic species) are notably more closely related to *T. propinquum* than any of the other North American or Holarctic species represented in this dataset. This might be explained by the propensity of *Theromyzon* species to feed in the nasopharyngeal orifices of migratory waterfowl (Elliott and Tullett 1982; Oosthuizen and Fourie 1985) that may lead to inadvertent dispersal via the hosts. This could also suggest that the geographic affinities matter less for the species identities than for other genera of leeches (Sawyer 1986). Long distance leech dispersal by avian hosts has previously been posited for members of the families Praobdellidae and Haemadipsidae (Siddall et al. 2013; Nakano et al. 2020), and the broad geographic distributions and lack of strong phylogeographic signal for *Theromyzon* are possibly the result of a similar phenomenon. Supporting this hypothesis, the recovered clade consisting of GenBank sequences with unknown specific affinity includes specimens from both North America and Europe (Fig. 7). We argue that any future phylogenetic work on this genus would benefit from considering the ecology and behaviour of the species, and that the genus represents a possible framework for studies on population structure and host-mediated dispersal.

Worth mentioning, however, is that recent investigations into wide-spread leech taxa have frequently suggested that cryptic diversity is prevalent (de Carle et al. 2017; Saglam et al. 2018; Iwama et al. 2019; Anderson et al. 2020). Although some species of *Theromyzon* are recorded from localities throughout the Northern Hemisphere (e.g. *T. tessulatum* and *T. maculosum*), specimens from disparate locations have not been compared since the advent of widely accessible DNA sequencing. Given the current muddled state of taxonomic literature for this genus, suggesting that cryptic diversity within *Theromyzon* may yet be discovered is reasonable.

Taxonomic accounts of *Theromyzon* species have been plagued by inconsistent or inadequate species descriptions that have led to an unusually large body of synonyms for several taxa. To clarify the current state of *Theromyzon* taxonomy, we include a list of currently valid species names and the synonyms in Supplementary File S3. Much of the historical taxonomic confusion surrounds the use of

the number of annuli separating the gonopores as a distinguishing feature. The gonopores are often extremely small in species of *Theromyzon* and in many cases impossible to observe. For some species (e.g. *T. trizonare* and *T. tigris* sp. nov.), identifications based on external morphological characters may be possible, although more research is needed to understand intra-specific phenotypic variation. Further useful external characters include colour and pigmentation patterns on the ventral and dorsal surfaces, body texture (soft and gelatinous or hard), and level of translucency. However, dissection is the more reliable way to morphologically identify members of this genus to species level. The reproductive structures characterised by Oosthuizen and Davies (1993), in combination with external characters, are useful but there is no single morphological character that will definitively distinguish between species.

We resolve several outstanding taxonomic issues within *Theromyzon* that were established in previous studies (Sawyer 1986; Oosthuizen and Davies 1992, 1993; Davies and Oosthuizen 1993; Oosthuizen 1993). Additionally, we contribute *COI* genetic variation, a metric that has proven useful in species delimitation studies but that is, unfortunately, not widely available for leeches. These values contribute to knowledge of genetic variation within leeches more generally. Following our morphological analyses, the sequences generated herein make identifications of these species based on DNA barcoding possible. In particular, this study contributes authoritative DNA sequences and comprehensive morphological circumscriptions for five species that will facilitate future identifications of North American specimens. This will reduce misidentifications and lay the groundwork for further research into this fascinating but taxonomically confusing genus. We also recognise a total of 14 valid species of *Theromyzon* worldwide. In all, we hope this study represents a significant contribution that will prove useful in deciphering this elusive genus.

Supplementary material

Supplementary material is available [online](#).

References

- Anderson K, Braoudakis G, Kvist S (2020) Genetic variation, pseudo-cryptic diversity, and phylogeny of *Erpobdella* (Annelida: Hirudinida: Erpobdelliformes), with emphasis on Canadian species. *Molecular Phylogenetics and Evolution* **143**, 106688. doi:[10.1016/j.ympev.2019.106688](https://doi.org/10.1016/j.ympev.2019.106688)
- Baird W (1869) Descriptions of some new suctorial annelids in the collection of the British Museum. *Proceedings of the Zoological Society of London* **37**, 310–318. doi:[10.1111/j.1469-7998.1869.tb07332.x](https://doi.org/10.1111/j.1469-7998.1869.tb07332.x)
- Bielecki A, Jelen I, Cichocka J, Walczak A (2016) *Theromyzon tessulatum* (O.F. Muller, 1774) (Clitellata: Hirudinida: Glossiphoniidae) – morphometry and structure of the digestive and reproductive systems. *Annals of Parasitology* **62**, 45.
- Blanchard R (1893) Révision des Hirudinées du Musée de Dresden. *Abhandlungen und Berichte des Konigl. Zoologischen und Anthropologisch-Etnographischen Museums zu Dresden* **4**, 1–8.
- Borda E, Siddall ME (2004) Arhynchobdellida (Annelida: Oligochaeta: Hirudinida): phylogenetic relationships and evolution. *Molecular Phylogenetics and Evolution* **30**, 213–225. doi:[10.1016/j.ympev.2003.09.002](https://doi.org/10.1016/j.ympev.2003.09.002)
- Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10**(4), e1003537. doi:[10.1371/journal.pcbi.1003537](https://doi.org/10.1371/journal.pcbi.1003537)
- Christoffersen ML (2007) Clitellate evolution and leech diversity: Glossiphoniidae excl. *Helobdella* (Annelida: Hirudinea: Rhynchobdellida) from South America. *Gaia. Scientia* **1**, 131–140.
- Davies RW, Oosthuizen JH (1993) A new species of duck leech from North America formerly confused with *Theromyzon rude* (Rhynchobdellida: Glossiphoniidae). *Canadian Journal of Zoology* **71**, 770–775. doi:[10.1139/z93-102](https://doi.org/10.1139/z93-102)
- de Carle D, Oceguera-Figueroa A, Tessler M, Siddall ME, Kvist S (2017) Phylogenetic analysis of *Placobdella* (Hirudinea: Rhynchobdellida: Glossiphoniidae) with consideration of *COI* variation. *Molecular Phylogenetics and Evolution* **114**, 234–248. doi:[10.1016/j.ympev.2017.06.017](https://doi.org/10.1016/j.ympev.2017.06.017)
- Elliott JM, Tullett PA (1982) Leech parasitism of waterfowl in the British Isles. *Wildfowl* **33**, 164–170.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.
- Goloboff PA, Catalano SA (2016) TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* **32**, 221–238. doi:[10.1111/cla.12160](https://doi.org/10.1111/cla.12160)
- Huelskenbeck JP, Ronquist F (2001) MrBAYES: Bayesian inference for phylogeny. *Bioinformatics* **17**, 754–755. doi:[10.1093/bioinformatics/17.8.754](https://doi.org/10.1093/bioinformatics/17.8.754)
- Iwama RE, Oceguera-Figueroa A, de Carle D, Manglicmot C, Erséus C, Miles NM-S, Siddall ME, Kvist S (2019) Broad geographic sampling and DNA barcoding do not support the presence of *Helobdella stagnalis* (Linnaeus, 1758) (Clitellata: Glossiphoniidae) in North America. *Zootaxa* **4671**, 1–25. doi:[10.11646/zootaxa.4671.1.1](https://doi.org/10.11646/zootaxa.4671.1.1)
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772–780. doi:[10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010)
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thaler T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649. doi:[10.1093/bioinformatics/bts199](https://doi.org/10.1093/bioinformatics/bts199)
- Klemm DJ (1985) Freshwater leeches (Annelida: Hirudinea). In ‘A guide to the freshwater Annelida (Polychaeta, naidid and tubificid Oligochaeta, and Hirudinea) of North America’. (Kendall/Hunt Publishing Co.: Dubuque, IA, USA)
- Livanow N (1902) Die Hirudineen-Gattung *Hemiclepsis* Vejd. *Zoologische Jahrbücher* **17**, 339–362.
- Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In ‘Proceedings of the Gateway Computing Environments Workshop (GCE)’, 14 November 2010, New Orleans, LA, USA. INSPEC Accession Number 11705685. (IEEE) doi:[10.1109/GCE.2010.5676129](https://doi.org/10.1109/GCE.2010.5676129)
- Moore PJ (1911) Hirudinea of Southern Patagonia. *Reports of the Princeton University Expeditions to Patagonia 1896–1899*(3), 669–687.
- Nakano T, Suzuki H, Suzuki N, Kimura Y, Sato T, Kamigaichi H, Tomita N, Yamasaki T (2020) Host-parasite relationships between seabirds and the haemadipsid leech *Chtonobdella palmyrae* (Annelida: Clitellata) inhabiting oceanic islands in the Pacific Ocean. *Parasitology* **147**(14), 1765–1773. doi:[10.1017/S0031182020001729](https://doi.org/10.1017/S0031182020001729)
- Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**, 268–274. doi:[10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300)
- Oka A (1930) Sur un nouveau genre d’Hirudinees provenant de l’Amérique du Sud. *Proceeding of the Imperial Academy (Tokyo)* **6**, 239–242. doi:[10.2183/pjab1912.6.239](https://doi.org/10.2183/pjab1912.6.239)

- Oosthuizen JH (1993) Redescription of the African duck leech *Theromyzon cooperi* (Harding, 1932) (Hirudinea:Glossiphoniidae). *South African Journal of Zoology* **28**, 67–70. doi:[10.1080/02541858.1993.11448294](https://doi.org/10.1080/02541858.1993.11448294).
- Oosthuizen JH, Davies RW (1992) Redescription of the duck leech, *Theromyzon rude* (Baird, 1869) (Rhynchobdellida: Glossiphoniidae). *Canadian Journal of Zoology* **70**, 2028–2033. doi:[10.1139/z92-274](https://doi.org/10.1139/z92-274)
- Oosthuizen JH, Davies RW (1993) A new species of *Theromyzon* (Rhynchobdellida: Glossiphoniidae), with a review of the genus in North America. *Canadian Journal of Zoology* **71**, 1311–1318. doi:[10.1139/z93-181](https://doi.org/10.1139/z93-181)
- Oosthuizen JH, Fourie FL (1985) Mortality amongst waterbirds caused by the African duck leech *Theromyzon cooperi*. *South African Journal of Wildlife Research* **15**, 98–106.
- Philippi RA (1867) Kurze Notiz über zwei Chilenische Blutegel. *Archiv für Naturgeschichte* **33**, 76–78.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**(5), 901–904. doi:[10.1093/sysbio/syy032](https://doi.org/10.1093/sysbio/syy032)
- Ringuelet R (1985) 'Fauna de agua dulce de la República Argentina.' (Fundación para la Educacion, la Ciencia y la Cultura: Buenos Aires, Argentina)
- Saglam N, Kutschera U, Saunders R, Saidel WM, Balombini KLW, Shain DH (2018) Phylogenetic and morphological resolution of the *Helobdella stagnalis* species-complex (Annelida: Clitellata: Hirudinea). *Zootaxa* **4403**, 61–86. doi:[10.11646/zootaxa.4403.1.3](https://doi.org/10.11646/zootaxa.4403.1.3)
- Sawyer, RT (1986). 'Leech Biology and Behaviour.' (Clarendon Press: Oxford, UK)
- Schulmeister S (2003) Simultaneous analysis of basal Hymenoptera (Insecta): introducing robust-choice sensitivity analysis. *Biological Journal of the Linnean Society. Linnean Society of London* **79**, 245–275. doi:[10.1046/j.1095-8312.2003.00233.x](https://doi.org/10.1046/j.1095-8312.2003.00233.x)
- Siddall ME, Burreson EM (1998) Phylogeny of leeches (Hirudinea) based on mitochondrial cytochrome oxidase subunit I. *Molecular Phylogenetics and Evolution* **9**, 156–162. doi:[10.1006/mpev.1997.0455](https://doi.org/10.1006/mpev.1997.0455)
- Siddall ME, Rood-Goldman R, Barrio A, Barboutis C (2013) The eyes have it: long-distance dispersal by an intraorbital leech parasite of birds. *The Journal of Parasitology* **99**(6), 1137–1139. doi:[10.1645/12-172.1](https://doi.org/10.1645/12-172.1)
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flok P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**, 651–701. doi:[10.1093/aesa/87.6.651](https://doi.org/10.1093/aesa/87.6.651)
- Sket B, Trontelj P (2008) Global diversity of leeches (Hirudinea) in freshwater. *Hydrobiologia* **595**, 129–137. doi:[10.1007/s10750-007-9010-8](https://doi.org/10.1007/s10750-007-9010-8)
- Stecher G, Tamura K, Kumar S (2020) Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular Biology and Evolution* **37**, 1237–1239. doi:[10.1093/molbev/msz312](https://doi.org/10.1093/molbev/msz312)
- Tessler M, de Carle D, Voiklis ML, Gresham OA, Neumann JS, Cios S, Siddall ME (2018) Worms that suck: phylogenetic analysis of Hirudinea solidifies the position of Acanthobdellida and necessitates the dissolution of Rhynchobdellida. *Molecular Phylogenetics and Evolution* **127**, 129–134. doi:[10.1016/j.ympev.2018.05.001](https://doi.org/10.1016/j.ympev.2018.05.001)
- Trontelj P, Sket B, Steinbrück G (1999) Molecular phylogeny of leeches: Congruence of nuclear and mitochondrial rDNA data sets and the origin of bloodsucking. *Journal of Zoological Systematics and Evolutionary Research* **37**, 141–147. doi:[10.1111/j.1439-0469.1999.tb00976.x](https://doi.org/10.1111/j.1439-0469.1999.tb00976.x)
- Whiting MF (2002) Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoologica Scripta* **31**, 93–104. doi:[10.1046/j.0300-3256.2001.00095.x](https://doi.org/10.1046/j.0300-3256.2001.00095.x)

Data availability. The data that support this study are available in the article and accompanying online supplementary material.

Conflicts of interest. The authors declare that they have no conflicts of interest.

Declaration of funding. This research was funded in part by NSERC Postgraduate Doctoral Scholarship (PGSD2-518435-2018, awarded to D. De Carle); NSERC Discovery Grant (RGPIN-2016-0612) and Helge Ax:son Johnson's Foundation Stipend (awarded to S. Kvist).

Acknowledgements. The authors thank Charlotte Calmerfalk Kvist and Michael Tessler for help with collecting specimens. We thank Maureen Zubowki and Don Stacey for invaluable work with the collections and Kristen Choffe for help with sequencing. We thank the Editor and two anonymous reviewers for insightful comments that improved the paper. We thank our generous funders: NSERC for the Postgraduate Doctoral Scholarship awarded to D. de Carle, and NSERC for the Discovery Grant and the Helge Ax:son Johnson's Foundation for the stipend awarded to S. Kvist.

Author affiliations

^ADepartment of Natural History, Royal Ontario Museum, 100 Queens Park, Toronto, ON, M5S 2C6, Canada.

^BDepartment of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada.