Zoologica Scripta

KUNGL. VETENSKAPS-AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Molecular phylogeny of Nearctic species of *Rhynchelmis* (Annelida)

Hong Zhou, Steven V. Fend, Daniel L. Gustafson, Pierre De Wit & Christer Erséus

Submission: 26 February 2009
Accepted: 17 March 2010
doi:10.1111/j.1463-6409.2010.00429.x

Zhou, H., Fend, S. V., Gustafson, D. L., De Wit, P. & Erséus, C. (2010). Molecular phylogeny of Nearctic species of *Rhynchelmis* (Annelida). — *Zoologica Scripta*, *39*, 378–393.

The Nearctic species of *Rhynchelmis* (Clitellata, Lumbriculidae) are known primarily from cool-water habitats in western North America. Their taxonomy has so far been based on limited collections from isolated localities, using intuitive assessment of morphological characters. This approach has proved unsatisfactory when additional populations of closely related species were sampled and scrutinized for incorporation in the present classification. Therefore, in this study, mitochondrial (cytochrome c oxidase subunit I and 16S rDNA) and nuclear internal transcriber spacer (ITS rDNA) genes were analysed as phylogenetic markers of Nearctic *Rhynchelmis* species. A combined approach with all the three gene regions provided a better resolution than any of the individual genes by itself. The genes demonstrated monophyly of all major groupings proposed on the morphological basis. Within the *Rhynchelmis yakimorum* complex, however, the genetic data and distribution suggested that two clades initially referred to as a '*R. yakimorum* variant 1', one from the lower Snake River drainage in Idaho and one from southern coastal Oregon, might represent two separate species. On the other hand, the sympatric distribution and low genetic distance between *Rhynchelmis gustafsoni* and a form tentatively identified as '*R.* cf. *yakimorum*' (both collected in eastern Idaho) indicated conspecific status. This study also showed that the cytochrome c oxidase subunit I (COI) gene, which may be informative of recent and on-going speciation and useful for species discrimination (as a DNA barcode), is less suitable as a single molecular marker for phylogenetic inference. Regardless of whether one deals with very closely related species (such as those of the *yakimorum* complex), with taxa with a wide and disjunct distribution (such as *Rhynchelmis rostrata*), or with more distantly related species, COI data should be supplemented by other genetic markers as well as morphological and biogeographical information.

Corresponding author: *Christer Erséus, Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden. E-mail: christer.erseus@zool.gu.se*

Hong Zhou, Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden

Present address for Hong Zhou, College of Marine Life Sciences, Ocean University of China, 5 Yushan Road, 266003, Qingdao, China. E-mail: hzhou@ouc.edu.cn

Steven V. Fend, U.S. Geological Survey, 345 Middlefield Rd. M/S 465, Menlo Park CA 94025, USA. E-mail: svfend@usgs.gov

Daniel L. Gustafson, Department of Ecology, Montana State University, PO Box 173460, Bozeman MT 59717-3460, USA. E-mail: dlg@rapid.msu.montana.edu

Pierre De Wit, Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden. E-mail: pierre.de_wit @zool.gu.se

Introduction

Rhynchelmis Hoffmeister, 1843 is a genus in the freshwater clitellate family Lumbriculidae, with about 25 species, subspecies and variants, all with spermathecae in segment VIII and male pores in X. Monophyly of this taxon is supported by several apomorphic characters (Brinkhurst 1989): a filiform proboscis, a long, tubular atrium, a duct joining the spermathecal ampulla to the gut and branched lateral blood vessels. This assumes that the Baikalian group related to *Rhynchelmis olchonensis* Burow & Kozhow,

1932, which by convention used to be included in *Rhynchelmis*, is referred to *Pseudorhynchelmis* Hrabě, 1982 (Martin *et al.* 1998; Kaygorodova & Liventseva 2007; Kaygorodova *et al.* 2007).

Three distinctive groups within *Rhynchelmis* have sometimes been regarded as separate genera, i.e. *Rhynchelmis* s. str., *Sutroa* Eisen, 1888; and *Rhynchelmoides* Hrabě, 1936 (Eisen 1888; Hrabě 1936; footnote p. 11; Hrabě 1982; Holmquist 1976; Sokol'skaya 1983). The Palaearctic *Rhynchelmis* s. str., recently designated as a subgenus *R.* (*Rhynchelmis*) (Fend & Brinkhurst 2010) are distinguished from all Nearctic species of *Rhynchelmis* s. lat. by several apomorphies, including longitudinal muscle bands that curl inwards, very thick vasa deferentia, spermathecal ampullae with a single thick diverticulum and (often) a large, glandular structure sometimes referred to as a 'rudimentary atrium' (Timm 1979; Hrabě 1982; Fend & Brinkhurst 2000). The North American '*Sutroa* group' includes seven named species, plus two undescribed (S. Fend, unpublished data) and three morphological variants (Fend & Brinkhurst 2000). Some features of this group are unique within the genus and rare within the family: two or more elongate diverticula originating at the ectal duct of each spermatheca, large spermathecal and penial 'bulbs' (with muscular-glandular masses at the pores) and distinctive atrial musculature. *Rhynchelmis yakimorum*, *Rhynchelmis monsserratus*, *Rhynchelmis gustafsoni* and *Rhynchelmis utahensis*, all Fend & Brinkhurst, 2000; here form a distinctive group, the *yakimorum* species complex *sensu* Fend & Brinkhurst (2000), characterized by paired, multilobed spermathecal diverticula. The '*Rhynchelmoides* group' is known primarily from western North America, with four species described from the western Nearctic (*Rhynchelmis elrodi* Smith & Dickey, 1918, *Rhynchelmis glandula* Altman, 1936, *Rhynchelmis alaskana* Holmquist, 1976; *Rhynchelmis saxosa* Fend & Brinkhurst, 2000), two undescribed eastern Nearctic species (Fend & Lenat, in prep.) and *Rhynchelmis orientalis* Yamaguchi, 1936 in Japan. This group is more difficult to define; it basically lacks the above apomorphies, and five of the seven species also lack the connection between the spermatheca and the gut. The Beringian *Rhynchelmis brooksi* Holmquist, 1976 and *Rhynchelmis malevici* (Sokol'skaya 1983) are sometimes associated with *Rhynchelmoides*, but most likely represent an additional group.

Despite the differences between these groups, the lack of a well-supported phylogenetic hypothesis of the genus and the presence of 'intermediate' forms, such as *Rhynchelmis orientalis* and *Rhynchelmis aleutensis* Fend, 2005, have led recent authors to take the conservative approach of retaining *Rhynchelmis* for all of these taxa (Timm 1997; Kathman & Brinkhurst 1998; Fend & Brinkhurst 2000). A morphology-based, parsimony analysis by Kaygorodova & Liventseva (2007) supported the *R.* (*Rhynchelmis*) clade, but not the *Sutroa* and *Rhynchelmoides* groups.

The *yakimorum* complex (within the *Sutroa* group) is mostly associated with present or former Snake River drainages in the Columbia River, and the different species and variants first appeared to have limited, disjunct ranges (Fend & Brinkhurst 2000). Subsequent collections (Fig. 1; and unpublished) have indicated considerably wider ranges for some types, with spatial overlap, and some of the original taxonomic decisions now need to be questioned. Typical *R. yakimorum* initially appeared restricted to the Yakima River, but similar material has been collected in Idaho (Fig. 1). *Rhynchelmis yakimorum* variant 1 (Fend & Brinkhurst 2000) was originally described from the lower Klamath River, which has no current connection to the Columbia. More recent records include other Pacific drainages near the Klamath, as well as two widely separated lower Snake River tributaries in Idaho and Nevada. *Rhynchelmis gustafsoni*, previously known from Montana and eastern Idaho, now has an expanded range to southern-central Idaho, and *R. utahensis*, first known from northern Utah, also occurs in Snake River drainages in southern Idaho. Despite some large-scale geographical overlap, local sympatry of *yakimorum* complex species appears restricted to a single area near the Little Lost and Pahsimeroi Rivers in Idaho (Fig. 2), where *R. gustafsoni* occurs more or less sympatrically with worms similar to *R. yakimorum* ('*R.* cf. *yakimorum* 'in this paper). Typical *R. yakimorum*, *R. yakimorum* variant 1 and *R. gustafsoni* are distinguished by the lateral placement of male and spermathecal pores (Figs S1 and S2; see also Fend & Brinkhurst 2000).

In this study, we used mitochondrial cytochrome c oxidase subunit I (COI) and 16S (rDNA), and nuclear ITS (the internal transcriber spacers region = ITS1 + 5.8-S + ITS2) sequences of six Nearctic *Rhynchelmis* species and some variants morphologically similar to *R. yakimorum* and *R. rostrata*, as independent data, with the aim to establish species boundaries and a well-supported phylogeny of this group. An important objective was also to address speciation and biogeography in the *Sutroa* group in Western USA.

Current knowledge of *Rhynchelmis*, and access to material for molecular study led us to test the following hypotheses: (1) *Rhynchelmis* s. lat. (excluding *Pseudorhynchelmis*) is monophyletic and separated from other lumbriculids with long atria and/or a prostomial proboscis. (2) The Nearctic taxa, roughly corresponding to the *Sutroa* plus *Rhynchelmoides* groups, are a monophyletic group separated from the Palaearctic *Rhynchelmis* s. str. (3) The

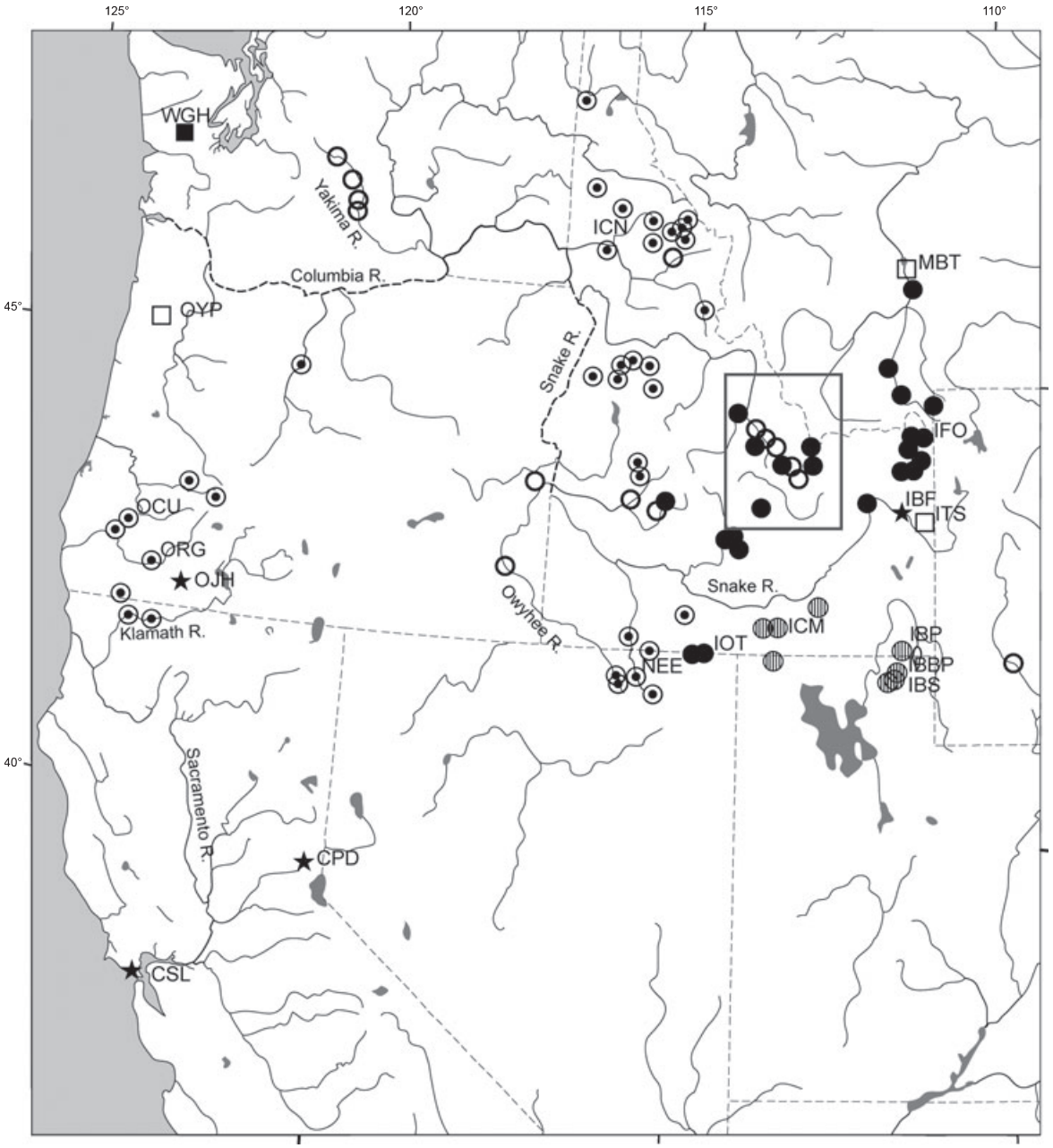

Fig. 1 Map of western USA, showing distributions of *Sutroa* group species used in the analyses. Refer to Table 1 for site abbreviations (three or four letter acronyms on map). The rectangular area indicates the area shown in greater detail in Fig. 2. Morphological taxa are: *Rhynchelmis rostrata* black stars (only the localities used in this analysis are shown), *Rhynchelmis yakimorum* (typical morphotype) open circle, *R. yakimorum* variant 1 circle with central dot, *Rhynchelmis gustafsoni* black circle, *Rhynchelmis utahensis* hatched circle. *Rhynchelmoides* group species (only the localities used in this analysis are shown): *Rhynchelmis saxosa* black square, *Rhynchelmis elrodi (MBT, ITS) and Rhynchelmis glandula* (OYP) open square.

Sutroa and *Rhynchelmoides* groups are each monophyletic. (4) Various forms attributed to *Rhynchelmis rostrata* are a monophyletic group. (5) The *yakimorum* species complex is monophyletic within the *Sutroa* group, and possibly the sister to *R. rostrata.* (6) Within this complex, *R. yakimorum* (with variants), *R. gustafsoni* and *R. utahensis* represent three clades. (7) The various morphotypes of *R. yakimorum* represent distinct lineages, despite their overlapping distributions [no type locality material, however, was available for the study.]

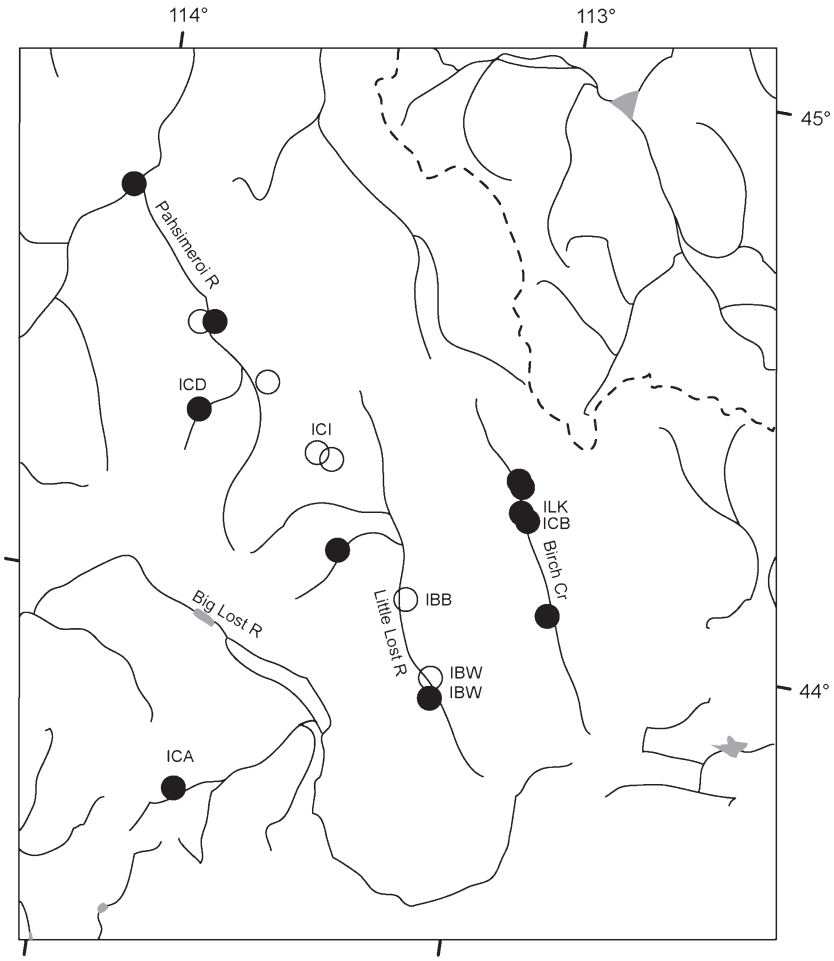

Fig. 2 Map of 'Lost Rivers' area of Idaho, showing distributions of *Rhynchelmis yakimorum* complex morphotypes. Refer to Table 1 for site abbreviations. *Rhynchelmis yakimorum* (typical morphotype) open circle, *Rhynchelmis gustafsoni* black circle.

Material and methods

Taxon and specimen sampling

Thirty-five specimens representing six Nearctic species and three variants of *Rhynchelmis* were collected for DNA, mostly from springs and rivers in north-western USA by D. L. Gustafson or S. Fend, and identified by S. Fend. Voucher specimens from the same collections were morphologically examined, and *R. yakimorum* complex specimens were either sectioned or transversely dissected for drawings of the reproductive organs (Figs S1 and S2).

DNA sampling was intended to encompass the known range of the *yakimorum* complex (Fig. 1), with particular emphasis on the region of the Little Lost and Pahsimeroi Rivers in Idaho (Fig. 2), where local sympatry of different morphotypes has been documented. Our material also covers much of the range of *R. rostrata*, and includes type locality specimens of both *R. rostrata* (CSL) and *Sutroa alpestris* Eisen, 1892 (Table 1: CPD), later designated a variant of *R. rostrata* (Brinkhurst & Cook 1966; Fend & Brinkhurst 2000). To represent the Palaearctic *Rhynchelmis* s. str. group, specimens of *R. tetratheca* Michaelsen, 1920 were collected from sites in Sweden, by S. Lundberg or C. Erséus. The choice of outgroup taxa was based on two distinctive characters of *Rhynchelmis* also shared with other lumbriculids: the unusually elongate atria and the filiform proboscis. The group with very elongate atria included four species of *Eclipidrilus* Eisen, 1881 (*Eclipidrilus lacustris* [Verrill, 1871], *Eclipidrilus frigidus* Eisen, 1881, *Eclipidrilus palustris* [Smith, 1900] and *E. pacificus* Fend, 2005); the group with a filiform proboscis included *E. palustris*, *Kincaidiana hexatheca* Altman, 1936, *Eremidrilus coyote* Fend & Rodriguez, 2003, and an undescribed Nearctic species of *Guestphalinus* Michaelsen, 1933 (S. Fend unpublished data). We used three sequences already available in GenBank; all other sequences are new. Vouchers of some of the sequenced worms, or topotypic specimens of a target population, were deposited in the Swedish Museum of Natural History, Stockholm. Specimens included in the study, with sampling sites (three-letter abbreviation denoting each local population), GenBank accession nos. and vouchers are listed in Table 1.

Table 1 Localities, GenBank accession numbers, and voucher specimen registration numbers of the six *Rhynchelmis* species and three variants included in the study and of the outgroup taxa (the four *Eclipidrilus* species, *Kincaidiana hexatheca*, *Guestphalinus* new species, and *Eremidrilus coyote*).

Taxon	Locality	Population code	CE Ref. no. or GenBank	GenBank accession no.			Combined analysis	Specimen voucher
				COI	16S	ITS		
Rhynchelmis gustafsoni	ID, Butte Co., Warm Spring Head	IBW	507	–	GU592339	–	–	SMNH105587
R. gustafsoni	ID, Lemhi Co., Kaufman Cabin Springs	ILK	508	GU592308	GU592340	GU592369	+	SMNH105588
R. gustafsoni	ID, Custer Co., Antelope Big Spring	ICA	512	GU592309	GU592341	GU592370	+	SMNH105589
R. gustafsoni	ID, Custer Co., Antelope Big Spring	ICA	GB	–	AY764153	–	–	–
R. gustafsoni	ID, Custer Co., Doublespring	ICD	582–1	–	–	FJ639329	+	SMNH103930
R. gustafsoni	ID, Custer Co., Doublespring	ICD	582–2	GU592310	GU592342	–	+	SMNH105590
R. gustafsoni	ID, Freemont Co., Osborne Spring	IFO	583–1	–	–	GU592371	+	SMNH105591
R. gustafsoni	ID, Freemont Co., Osborne Spring	IFO	583–2	GU592311	GU592343	–	+	–
R. gustafsoni	ID, Clark Co., Birch Cr. at Kaufman	ICB	584–1	–	–	GU592372	+	SMNH105592
R. gustafsoni	ID, Clark Co., Birch Cr. at Kaufman	ICB	584–2	GU592312	GU592344	–	+	–
R. gustafsoni	ID, Owyhee Co., Twin Springs at Shoshone	IOT	1083	–	–	GU592368	–	SMNH105593
Rhynchelmis cf. *yakimorum*	ID, Custer Co., Iron Springs	ICI	498	GU592326	GU592357	GU592388	+	SMNH105594
R. cf. *yakimorum*	ID, Butte Co., Big Spring Cr.	IBB	501	–	–	GU592389	+	SMNH105595
R. cf. *yakimorum*	ID, Butte Co., Big Spring Cr.	IBB	502	GU592327	GU592358	–	+	–
R. cf. *yakimorum*	ID, Butte Co., Warm Spring Head	IBW	504	–	GU592359	GU592390	+	SMNH105596
R. yakimorum variant 1	NV, Elko Co., East Fork Owyhee R. at Beaver Cr.	NEE	1086	GU592323	GU592353	GU592384	+	SMNH105597
R. yakimorum variant 1	ID, Clearwater Co., North Fork Clearwater R. below dam	ICN	1087	GU592324	GU592354	GU592385	+	SMNH105598
R. yakimorum variant 1	OR, Cow Cr., tributary to Umpqua R.	OCU	860	GU592325	GU592355	GU592386	+	SMNH105599
R. yakimorum variant 1	OR, Jackson Co., Rogue R. at Gold Hill	ORG	867–1	–	–	GU592387	+	–
R. yakimorum variant 1	OR, Jackson Co., Rogue R. at Gold Hill	ORG	867–2	–	GU592356	–	+	–
Rhynchelmis utahensis	ID, Bear Lake Co., St. Charles Spring	IBS	491	GU592319	GU592349	GU592380	+	SMNH105600
R. utahensis	ID, Bear Lake Co., Porcupine Camp Spring	IBP	492	GU592320	GU592350	GU592381	+	SMNH105601
R. utahensis	ID, Cassia Co., McClenden Spring	ICM	496	GU592321	GU592351	GU592382	+	SMNH105602
R. utahensis	ID, Bear Lake Co., Blue Pond Spring	IBBP	585	GU592322	GU592352	GU592383	+	SMNH105603
Rhynchelmis rostrata	ID, Bonneville Co., Fisher Bottom, Lower Spring	IBF	581	GU592313	–	GU592374	+	SMNH105604
R. rostrata	CA, San Francisco, Lobos Cr. below Mountain Lake	CSL	556	–	–	GU592373	–	SMNH105615
R. rostrata	OR, Jackson Co., Ice House Cr.	OJH	1070	GU592305	GU592336	GU592366	+	SMNH105616
R. rostrata variant	CA, Placer Co., spring on North shore of Donner Lake	CPD	555	GU592314	GU592345	GU592375	+	–
Rhynchelmis elrodi	ID, Teton Co., Sherman Spring West	ITS	GB	–	AY764152	–	–	SMNH105617
R. elrodi	MT, Broadwater Co., Toston Big Spring	MBT	580	GU592306	GU592337	FJ639328	+	SMNH105618*
Rhynchelmis glandula	OR, Yamhill Co., spring at Peavine Ridge, W of McMinnville	OYP	864–1	–	–	GU592367	+	SMNH105619
R. glandula	OR, Yamhill Co., spring at Peavine Ridge, W of McMinnville	OYP	864–2	GU592307	GU592338	–	+	–
Rhynchelmis saxosa	WA, Grays Harbor Co., Humptulips R. at Campbell Tree Grove	WGH	706–1	–	GU592346	GU592376	+	SMNH105620
R. saxosa	WA, Grays Harbor Co., Humptulips R. at Campbell Tree Grove	WGH	706–2	GU592315	–	–	+	–
Rhynchelmis tetratheca	SE, Södertälje, Moraån R. near Järna	SSM	322	GU592316	AY340477	GU592377	+	–
R. tetratheca	SE, Örebro, Karlslund, Svartån R.	SKS	586	GU592317	GU592347	GU592378	+	–
R. tetratheca	SE, Gotland, Lummelunda, southern branch of Lummelundaån R.	SGL	641	GU592318	GU592348	GU592379	+	–
Eclipidrilus palustris	FL, Orange Co., near Christmas Tootoosahatchee Cr.	FOT	868	–	GU592333	–	–	SMNH105621
Ec. palustris	FL, Orange Co., small tributary to Rock Springs Run	FOR	1076	GU592302	GU592332	–	+	SMNH105622

Table 1 (Continued).

Taxon	Locality	Population code	CE Ref. no. or GenBank	GenBank accession no.			Combined analysis	Specimen voucher
				COI	16S	ITS		
Eclipidrilus lacustris	OR, Jackson Co., Rogue R. at Gold Hill	ORG	862–1	–	–	GU592361	+	SMNH105623
Ec. lacustris	OR, Jackson Co., Rogue R. at Gold Hill	ORG	862–2	–	GU592330	–	+	–
Eclipidrilus frigidus	CA, Sierra Co., Big Spring at Bassets	CSB	557	GU592300	GU592329	GU592360	+	SMNH105624
Ec. frigidus	CA, Yuba Co., Jackass Cr. at Deadwood Cr.	CYJ	GB	AY040706	–	–	–	SMNH105625
Eclipidrilus pacificus variant	OR, Eugene, spring at Fox Hollow Road	OEF	863–1	–	–	GU592362	+	SMNH105626
Ec. pacificus variant	OR, Eugene, spring at Fox Hollow Road	OEF	863–2	GU592301	GU592331	–	+	–
Eremidrilus coyote	CA, Santa Clara Co., Coyote Cr. below Madrone Soda Springs	CSC	702	–	GU592328	GU592363	+	SMNH105627
Kincaidiana hexatheca	OR, Cow Cr, tributary to Umpqua R.	OCU	861	GU592304	GU592335	GU592365	+	–
Guestphalinus n.sp.	WA, Jefferson Co., Clearwater R. at Copper Mine Bottom Camp	WJC	865–1	GU592303	–	–	+	SMNH105628
Guestphalinus n.sp.	WA, Jefferson Co., Clearwater R. at Copper Mine Bottom Camp	WJC	865–2	–	GU592334	GU592364	+	–

CA, California; Co., County; Cr., Creek; FL, Florida; ID, Idaho; MT, Montana; NV, Nevada; OR, Oregon; R., River; SE, Sweden; WA, Washington. For populations ICA, ICD, IFO, ICB, IBB, ORG, OYP, ORG, OEF and WJC, sequences of two specimens were combined for the combined analysis. Depository of vouchers, Swedish Museum of Natural History (SMNH), Stockholm.

DNA extraction, amplification and sequencing

DNA was extracted from 95% ethanol-preserved animals using the E.Z.N.A. Tissue DNA Kit (Omega Bio-Tek, Inc., Norcross, GA, USA) according to the protocol supplied by the manufacturer. About 1100 bp of the ITS region were amplified using the primers ITS-5 and ITS-4 (White *et al.* 1990). About 650 bp of the COI gene were amplified using the universal primers LCO1490 and HCO2198 (Folmer *et al.* 1994). Amplification of 16S was performed either by the primer pair 16SAR-L/16SBR-H (Palumbi 1996; ca. 480 bp) or by Ann 16SF/Ann 16SR (Sjölin *et al.* 2005, ca. 300 bp). PCR was carried out with standard products and by protocols as recommended by the manufacturers.

Sequencing was performed on a Beckman Coulter CEQ 8000 using the Genome Lab Dye Terminator Cycle Sequencing Quick Start Kit (Beckman Coulter, Inc., Brea, CA, USA), and the cycling profile for the sequencing reactions was as follows: (96 °C/20 s–50 °C/20 s–60 °C/4 min) × 29. Both strands were sequenced using the same primers as for the PCR, but for ITS, four overlapping fragments of ca. 600 bp each were sequenced using the primers ITS-4 and ITS-5, and two internal primers, (Källersjö *et al.* 2005) 5.8SF and 5.8SR, which both anneal to 5.8S rDNA. Contigs were assembled using the DNASTAR package (Lasergene DNASTAR, Inc., Madison, WI, USA) and examined for sequencing errors.

Alignment and phylogenetic inference

Sequences were edited with Geneious Pro (Biomatters Ltd., Auckland, New Zealand). COI and 16S data were aligned using the MUSCLE web server (available at http://www.ebi.ac.uk/Tools/muscle/index.html) (Edgar 2004), while the ITS sequences, which showed considerable length variation, were aligned using ClustalX 2.0.12 (Thompson *et al.* 1997); both programs with default settings.

We used two methods for phylogenetic estimation: parsimony and Bayesian analyses of both separate and combined DNA sequence datasets. For about half of our terminal taxa, the three genes were successfully sequenced from the same specimen. In some cases, however, the three gene sequences were obtained from two different individuals of the same collecting site; i.e. we assumed that there was no or only trivial within-site variation. Thus, the combined analyses were designed to encompass those site-populations with at least two gene sequences, resulting in 25 ingroup and seven outgroup terminals (Table 1).

PAUP* 4.0b10 (Swofford 2003) was used to find the most parsimonious trees with heuristic search, the TBR branch-swapping algorithm, and branches collapsed if minimum length is zero ('amb-'). Gaps in the alignments

of 16S and ITS sequences were either treated as missing data or as a fifth character state, and correspondingly two sets of clade support were assessed by parsimony bootstrapping with 1000 replicates, TBR branch swapping and 10 random addition sequences per replicate.

Bayesian inference analyses using Metropolis-coupled Markov chain Monte Carlo (MCMCMC) estimation of posterior probability distributions (Nylander *et al.* 2004) were conducted with MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003; Huelsenbeck *et al.* 2005), using the default values of one cold and three heated Markov chains for two simultaneous and independent runs. Three separate analyses were performed starting from random trees to assure convergence (Huelsenbeck *et al.* 2002). In each analysis, the MCMC sampling procedure was run 50 000 000 generations and sampled every 1000 generations; the first 10 000 000 generations were discarded as burn-in, leaving 40 001 trees for estimating posterior probabilities. Models were selected for the three gene datasets and codon positions in the amino acid triplets of COI based on the Akaike Information Criterion (Akaike 1973), and calculated in MrModeltest 2.2 (Nylander 2004). The combined dataset was partitioned by genes and by codon positions in COI, and allowed for varied DNA substitution models and different evolutionary rates across partitions. The models selected for the first, second and third codon positions of COI were GTR + I, F81 and GTR + G respectively. The model of best-fit to both of 16S and ITS gene data was GTR + I + G.

Results

Alignments

The aligned partial COI sequences consist of 569 nucleotide sites, of which 233 (41%) are variable and 201 (35%) are parsimony informative. The aligned 16S sequence data are composed of 493 positions (including 26 with gaps), of which 187/197 (38%/40%) are variable and 128/138 (26%/28%) parsimony informative, when gaps are treated as missing data or as a fifth character state respectively ('missing data'/'fifth character'). The ITS gene sequences have 1380 positions after alignment, including a 159-bp-long, highly conserved 5.8S region. The ITS-1 and ITS-2 spacer regions are about 550 (range 545–573) and 350 (range 350–381) bp long, respectively, for the investigated *Rhynchelmis* species. There are 650/795 (47%/58%) variable sites, 386/546 (28%/40%) of which are parsimony informative if gaps are treated as missing data or as fifth characters respectively.

COI divergence

As reflected in Fig. 3 (see also below), the average COI haplotype divergence varies considerably within species/variants, from 0.1% in *R. tetratheca* to 9.0% in *R. rostrata*; all three populations of the latter (including the variant) were widely separated, both genetically and geographically. For *R. gustafsoni*, divergence among populations range from 0% to 4.0%. Among the four populations of *R. utahensis*, ICM is 9.8% different from the other three, which vary by <1%. Among species/variants, the average variation is only 2.5% (range 1.4–4.0%) between *R.* cf. *yakimorum* and *R. gustafsoni*, 4.0% (range 3.9–4.2%) between *R.* cf. *yakimorum* and the NEE-ICN variant 1, but 7.6% between *R.* cf. *yakimorum* and the OCU variant 1 (range 7.2–8.1%). Pairwise COI *p*-distances among populations of *R. gustafsoni* and various forms of *R. yakimorum* all exceed 1% (1.1–9.3%) except for the zero distance between the ILK-ICB populations of *R. gustafsoni*. Other interspecific variation in *Rhynchelmis* is between 10.9% (*elrodi* vs. *glandula*) and 20.1% (*saxosa* vs. *rostrata*). The *p*-distance between the ingroup (all species of *Rhynchelmis*) and the outgroup taxa is on average 19.0%.

Phylogenetic analyses

The Bayesian analysis of the COI gene (Fig. 3) provided good resolution for most closely related lineages, and even subdivision of populations, but left some deeper relationships unresolved, including the possible monophyly of *Rhynchelmis* s.lat.. Fourteen nodes are supported by posterior probabilities (pp) ≥0.95, two of which are the *Sutroa* group and the *yakimorum* complex (both with pp 1.00). The *Rhynchelmoides* group, however, is not supported; the tree suggests (with pp 0.99) that two of its species (*R. elrodi*, *R. glandula*) are more closely related to the *Sutroa* group than they are to the third (*R. saxosa*). Monophyly of the three populations of *R. rostrata* (pp 1.00), with the closest relationship between populations IBF and OJH (pp 0.98), is implied. The *yakimorum* complex is the sister group of *R. rostrata*. Within this complex, the *R. utahensis* clade (pp 0.95) is the sister group of a clade consisting of *R. gustafsoni* and various forms of '*yakimorum*' (including variants); within this latter clade, however, resolution is poor. The parsimony bootstrap analysis (of COI) resulted in a tree (see bootstrap support values in Fig. 3) less resolved than, but not incongruent with the Bayesian tree.

In the Bayesian tree based on the 16S gene (Fig. 4), 10 nodes gained pp values ≥0.95. These nodes include the *Sutroa* group (pp 0.96) and a large terminal group in the *yakimorum* complex, i.e. *R. gustafsoni*, *R.* cf. *yakimorum* and the ICN and NEE populations of *yakimorum* variant 1 (pp 0.98). Monophyly of the *yakimorum* complex as a whole is suggested, but with rather low support (pp 0.92), and so is monophyly of *Rhynchelmoides*, but with poor support (pp only 0.65). The parsimony bootstrap trees (of 16S) are

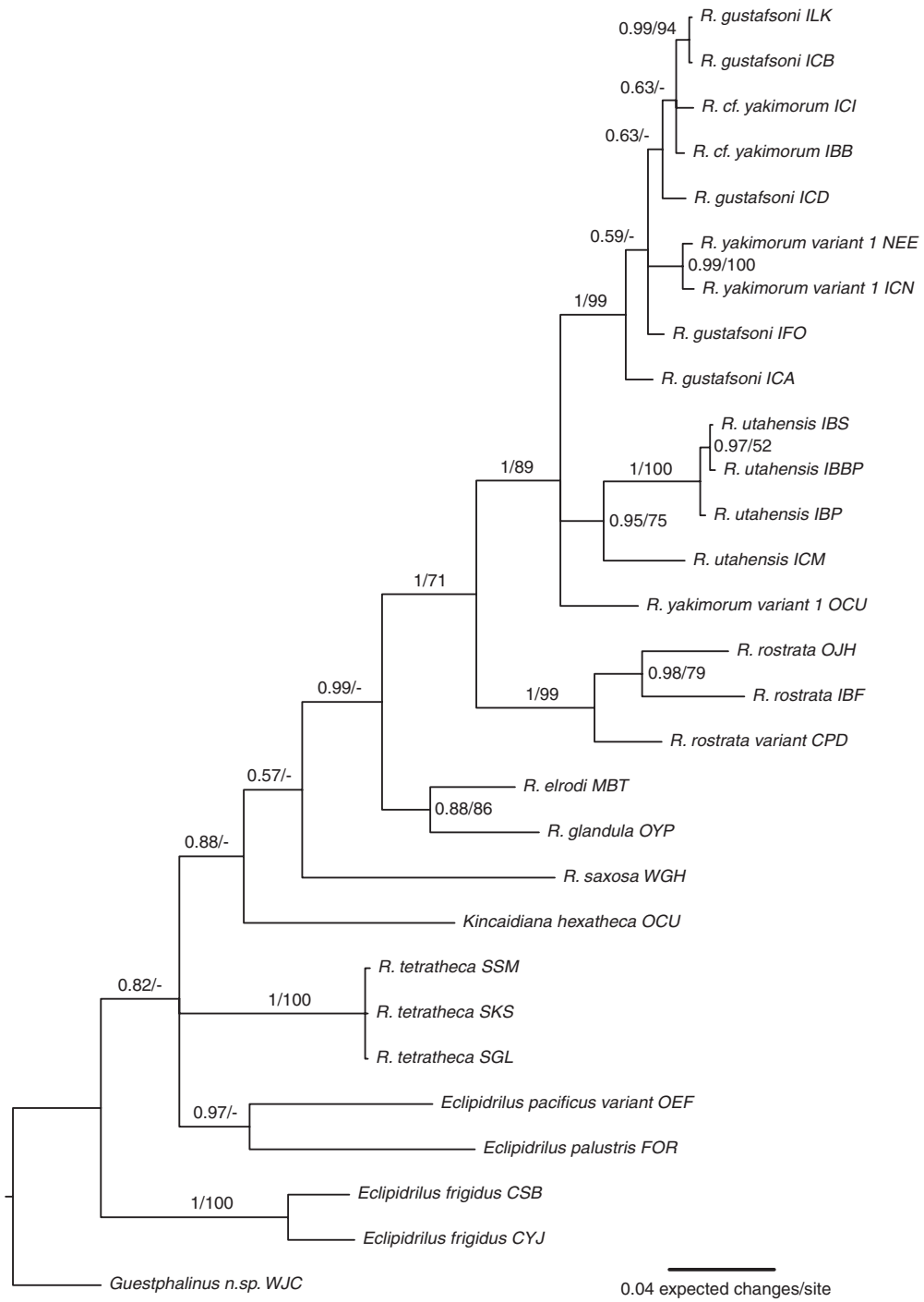

Fig. 3 Bayesian consensus tree of mitochondrial cytochrome c oxidase subunit I (COI) sequences. Two node support values (≥50%) are given using the formula: posterior probabilities/bootstrap.

congruent with the Bayesian tree (see Fig. 4) regarding all *Rhynchelmis* taxa; however, there are minor differences in the suggested relationships among the outgroup taxa (not shown here).

The ITS Bayesian tree (Fig. 5), with 17 nodes supported by pp ≥0.95, corroborates both *Rhynchelmis* s.lat. (pp 1.00), with *R. tetratheca* sister to the remaining taxa (pp 1.00), and the *Sutroa* group (pp 1.00), including the

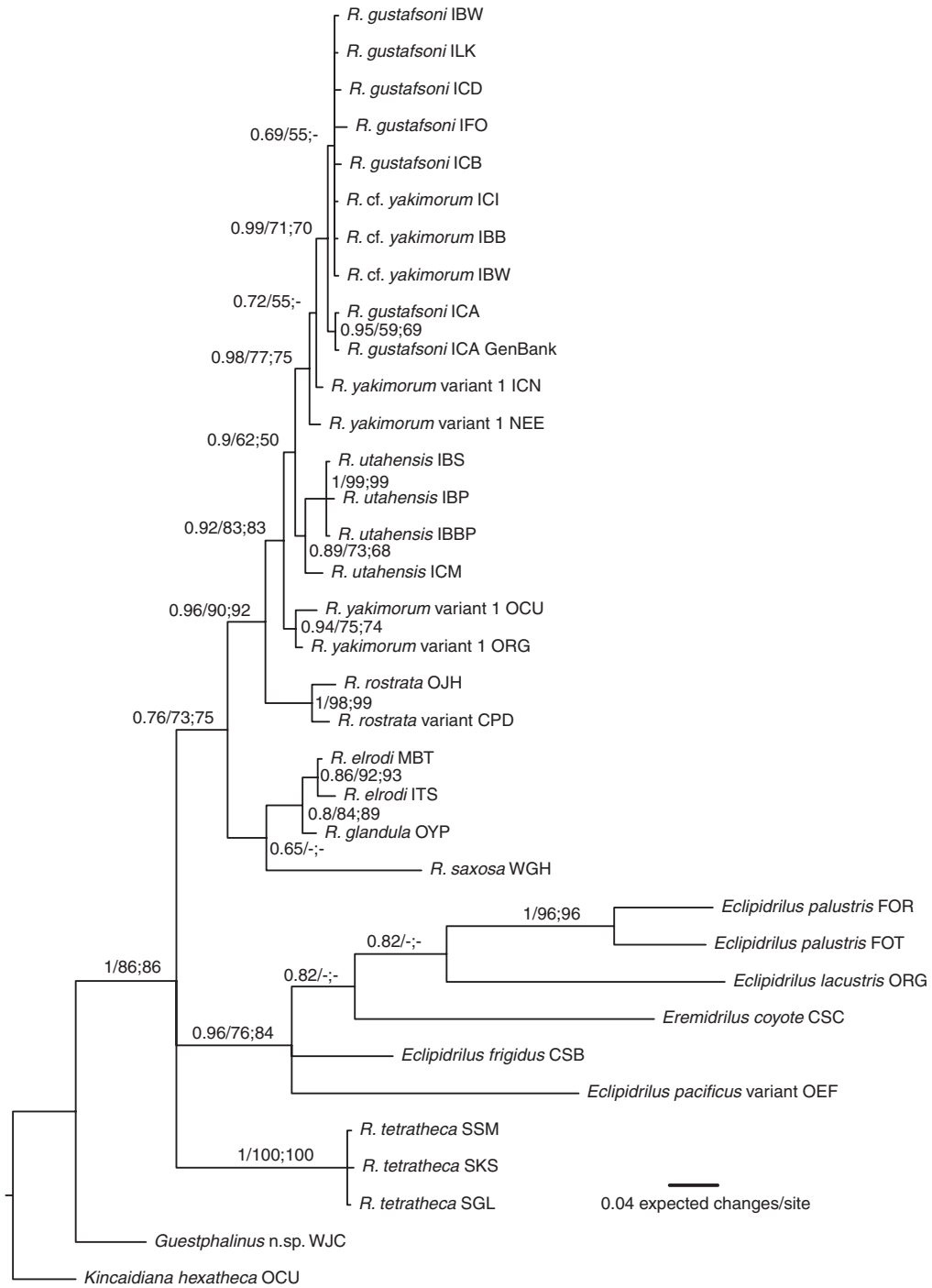

Fig. 4 Bayesian consensus tree of mitochondrial 16S sequences Three node support values (≥50%) are given using the formula: posterior probabilities/bootstrap with gaps = missing data; bootstrap with gaps = 5th character state.

yakimorum complex (supported by pp 1.00) and *R. rostrata* (pp 1.00) as sister groups. Furthermore, with pp 1.00, the *Rhynchelmoides* group is sister to *Sutroa* (although the support for *Rhynchelmoides* is rather low, pp 0.93). Within *R. rostrata*, the variant (CPD) is suggested as sister group to a group of the three remaining populations (IBF, CSL, OJH; pp 0.87); among the latter, a close relationship between IBF and CSL is supported (pp 0.95). Within the

yakimorum complex, *R. utahensis* is not, as a whole, resolved from the rest, but populations IBS, IBP and IBBP form a strongly supported group (pp 1.00). Among the populations of *R. yakimorum* variant 1, OCU and ORG form one group (pp. 1.00), NEE and ICN another group (pp 0.99); and with pp 0.97, the latter is most closely related to a group containing all populations of *R.* cf. *yakimorum* and *R. gustafsoni*. However, the latter two are not supported as separate species. The parsimony bootstrap trees (of ITS) are largely congruent with the Bayesian tree (see bootstrap values in Fig. 5).

The Bayesian analysis of the three genes combined provided a tree (Fig. 6) with 24 well-supported nodes (pp ≥0.95); the tree is thus close to fully resolved. *Rhynchelmis* s. lat., the *Rhynchelmoides* and *Sutroa* groups, and the nominal taxa *R. rostrata* and *R. utahensis* all have maximum support. The combined dataset also gives a better resolution of the *yakimorum* complex (Fig. 6) than any of the individual gene trees (Figs 3–5). Within the *Rhynchelmoides* group, *R. elrodi* and *R. glandula* form the sister group (pp 0.97) to *R. saxosa*. Moreover, *Eclipidrilus* (a genus with elongate atria) stands out as a well-supported taxon (pp 1.00). The bootstrap values of the parsimony analyses (of all data combined) show no important incongruence with the Bayesian tree, except for the lack of support for *E. lacustris* being sister to *E. pacificus* variant; the bootstrap analyses instead suggest that *E. lacustris* is sister to *E. palustris*.

Discussion

COI, 16S and ITS as phylogenetic markers

The three genes analysed in our study gave similar results. Although differing in detailed branching and node support, the mitochondrial (COI and 16S) and nuclear trees (ITS) were largely congruent. Thus, in our material, the maternal genealogy (only eggs carry mitochondria between generations) runs in parallel with the genealogy of the continuously recombined DNA of the nuclear genome. This strengthens the case for ceased gene flow (and speciation) between lineages with significant COI and 16S differences.

Our analyses have not refuted the five higher-level phylogenetic hypotheses proposed based on morphology (see Introduction). *Rhynchelmis* s. lat. (but excluding *Pseudorhynchelmis*) is monophyletic and separated from other Nearctic lumbriculids having *Rhynchelmis*-like characters such as a prostomial proboscis or elongate atria (shown by ITS as well as combined data; Figs 5 and 6). The Nearctic groups corresponding to *Sutroa* and *Rhynchelmoides* together form a monophyly with a sister group relationship to the Palaearctic *Rhynchelmis* s. str. (also Figs 5 and 6). These results support traditional groupings (Hrabe 1982), as well as the morphology-based analyses by Kaygorodova & Liventseva (2007), which used an outgroup of Palaearctic taxa. Furthermore, both the *Sutroa* and the morphologically more generalized *Rhynchelmoides* groups are monophyletic and sister groups (Figs 5 and 6). The *Sutroa* group, characterized by distinctive characters such as spermathecal diverticula and penial bulbs, is supported by all analyses (Figs 3–6). The various widely distributed populations attributed to *Rhynchelmis rostrata* form a monophyletic group, separate from all other *Rhynchelmis* species (Figs 3–6). The *yakimorum* species complex is monophyletic within the *Sutroa* group and the sister group of *R. rostrata* (COI and combined data; Figs 3 and 6).

The lower-level, morphology-based hypotheses (items 6 and 7; see Introduction) were generally not supported, particularly where they also appeared contradicted by biogeography (all sets of data). Of the three nominal species within the *yakimorum* complex, only *R. utahensis* represents a separate clade (Figs 3 and 6), whereas *R. gustafsoni* and various forms of '*yakimorum*' do not. The *yakimorum* variants also did not appear to represent phylogenetic clades corresponding to morphological categories. Two clades, both attributed to as *R. yakimorum* variant 1 on the basis of morphology (NEE-ICN and OCU-ORG clades, respectively) appear well separated from each other (Figs 3, 5 and 6), while all specimens of *R.* cf. *yakimorum* are nested within *R. gustafsoni* (all analysis). As our study did not include the population of *R. yakimorum* originally described from Washington (Fend & Brinkhurst 2000), we were unable to establish where *R. yakimorum*, in its most restricted sense, would be positioned in our tree (see further below).

In our study, the combined dataset was superior for overall resolution, but at the particular taxonomic levels studied, we found that 16S and ITS were useful more or less across the spectrum of levels, while the COI gene was best suited for describing the shallow phylogeny within species or among closely related species. This corroborates earlier conclusions that COI largely lacks the capability of uncovering deep phylogenetic relationships (Chen *et al.* 2004; Harrison 2004). Being a protein-coding gene, COI has most of its variation in third positions and reaches saturation in nucleotide substitutions rather soon after a divergence event.

Thus, the three genes combined gave a more resolved phylogeny of the Nearctic *Rhynchelmis* than each gene by itself. Moreover, this analysis was also more congruent with the morphologically defined groups. This emphasizes the importance of simultaneously evaluating several kinds of data in phylogenetic inference (see, e.g., Kluge 1989, 1998).

Fig. 5 Bayesian consensus tree of nuclear internal transcriber spacer (ITS) sequences. Three node support values (≥50%) are given using the formula: posterior probabilities/bootstrap with gaps = missing data; bootstrap with gaps = 5th character state.

COI distances and species delineation

DNA barcoding using a standard COI divergence threshold for species diagnosis may be an appealing prospect in situations where prior taxonomic work has been limited (Hebert *et al.* 2003). The latter authors established that more than 98% of 13 320 congeneric species pairs representing 11 phyla showed sequence divergence (using p-distances) >2%, with a mean value of 11.3 ± 5.3%. The mean

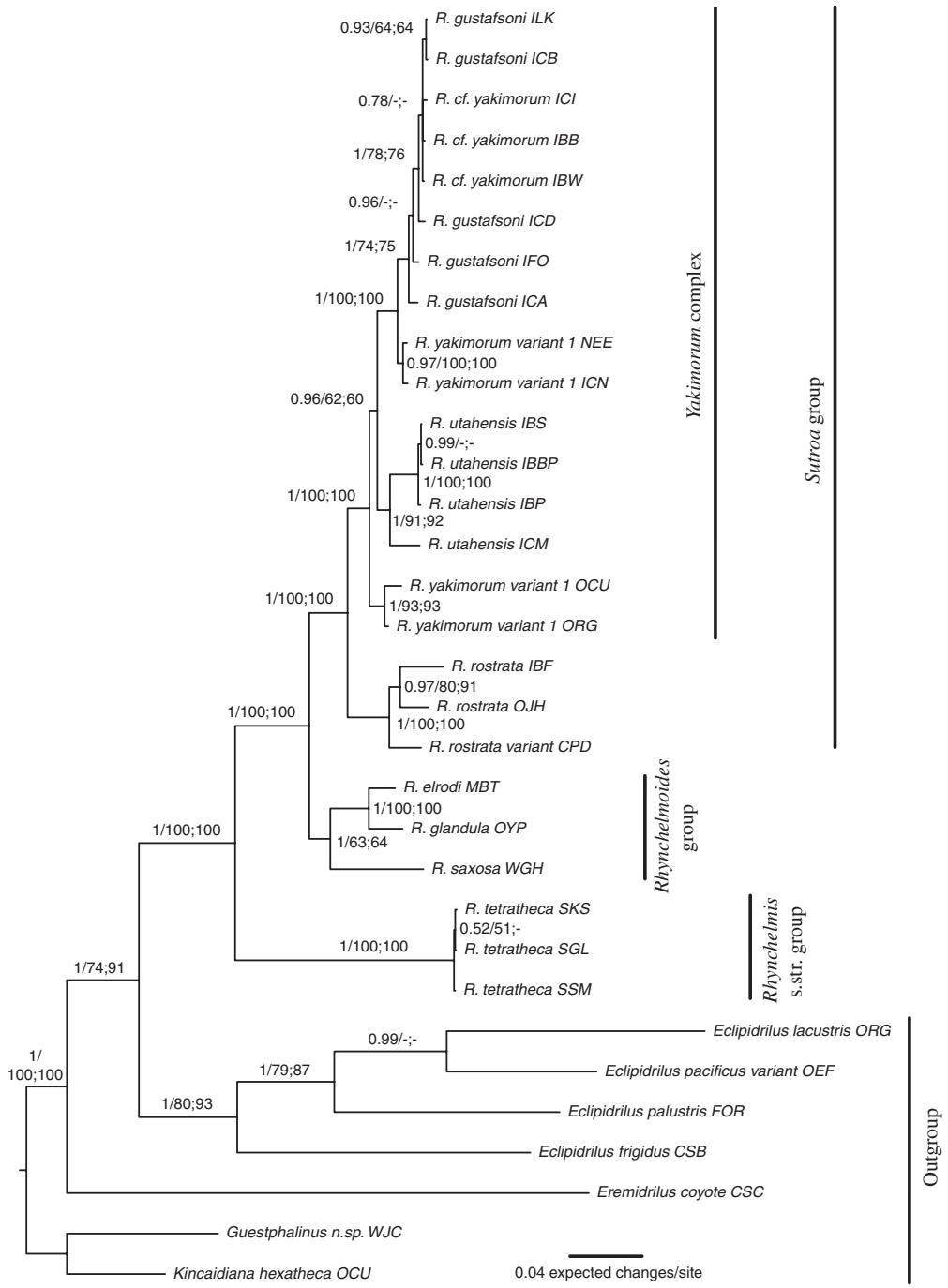

Fig. 6 Bayesian consensus tree of combined sequences of cytochrome c oxidase subunit I (COI), 16S and internal transcriber spacer (ITS). Three node support values (≥50%) are given using the formula: posterior probabilities/bootstrap with gaps = missing data; bootstrap with gaps = 5th character state.

percentage was a bit higher for the 128 species pairs analysed within Annelida (15.7 ± 6.2%). Erséus & Kvist (2007) showed that between-species distances are over 20 times higher than within-species p-distances, 19–25% as opposed to <1%, in four species of the marine clitellate genus *Tubificoides* (Annelida, Clitellata). However, this was based on material from a limited area, and the intraspecific variation is likely to increase when sampling is extended

to represent the full geographical range of a species; furthermore, the interspecific variation gets lower when only true sister species are compared (Moritz & Cicero 2004; Meyer & Paulay 2005). Using a geographically broader sample, Gustafsson *et al.* (2009) revealed cryptic speciation within the Holarctic freshwater lumbriculid, *Lumbriculus variegatus*, with two main clades being separated by 17.7 ± 0.3% COI p-distances, and a within-clade variation of, respectively, 0.6 ± 0.6% and 1.3 ± 0.8% Other examples of COI variation in clitellates are reviewed by Erséus & Gustafsson (2009), and as a preliminary conclusion, it seems that congeneric species of Clitellata indeed differ from each other by about 10% or more, while intraspecific variation normally is about one order of magnitude lower. Therefore, COI distances intermediate with regard to this (say, around 5% and up) are putative signs of ongoing speciation.

Applying this to our study (Fig. 3), it seems probable that *R. tetratheca*, *R. saxosa*, *R. glandula* and *R. elrodi* are good species, as they are clearly delineated from their congeners by COI p-distances of 11–20%. Although *R. rostrata* (including its variant) and *R. utahensis* were both resolved as well supported clades, COI p-distances suggest that each contains more than one species. Within the *Rhynchelmoides* group, *R. saxosa* differs from *R. glandula* and *R. elrodi* in lacking the connection between the spermatheca and the gut; genetically, it was also the most distinct taxon within the group, with a COI p-distance to the others of 18–19%. Moreover, Brinkhurst & Cook's (1966) suggestion that *R. glandula* and *R. elrodi* are conspecific, based on their similar morphology, is not supported by our data. The three sampled populations of *R. rostrata* (one of which being the 'variant') differed from each other by about 9%. Similarly, the ICM specimens of *R. utahensis* differed from the three other specimens (IBP, IBS, IBBP) by about 10%; separate evolution of a clade containing the three latter specimens is also supported by the ITS data (see Fig. 5).

The *yakimorum* complex appears to be a heterogeneous group, with *R. yakimorum* variant 1 even being polyphyletic; the 'variant 1' from western Oregon is well separated from all other forms in the complex by 7–9% in COI. The others (*R. gustafsoni*, *R.* cf. *yakimorum* and '*R. yakimorum* variant 1' from Nevada and Idaho) may be in the middle of a complicated speciation process, being separated from each others by up to 5%. The pronounced differences in orientation of copulatory organs (Fig. S1) certainly suggest a possible isolating mechanism, particularly with respect to the *R. gustafsoni* vs. *R. yakimorum* morphs occurring sympatrically at some 'Lost Rivers' localities. Intermediate forms were not seen at these sites. Nevertheless, the actual copulatory function of the male structures is not known, and it seems unlikely that the robust porophores function as typical penes. For that matter, it is possible that the appearance of the everted porophores is an artefact of fixation.

Biogeography

Despite significant range expansions for some species (Fig. 1), the *yakimorum* complex still appears largely associated with the modern Snake River drainage and adjacent Columbia River tributaries. Populations associated with other river systems, including the Bear River, upper Missouri drainages and a few Pacific coastal drainages (Klamath, Rogue, and Umpqua Rivers) can be associated with former Snake connections, based on studies of other organisms (Taylor 1985; Minckley *et al.* 1986).

The genetic relationships among *yakimorum* complex populations seem to correspond more closely to geography than to morphotype. This is particularly evident in the Little Lost/Pahsimeroi river drainages, where worms closely resembling *R. yakimorum* (ICI, IBB, IBW) in the Little Lost River drainage are genetically close to morphological '*gustafsoni*' (ICD) in the Pahsimeroi drainage (Fig. 6), and the two morphotypes occur sympatrically throughout this area (Fig. 2). Although a few specimens resembling typical *yakimorum* have been collected at other scattered sites in western Idaho (Fig. 1), the type locality in Washington is distant both geographically and in terms of river connections (Fig. 1). The ITS data are not indicative of reproductive isolation between the *gustafsoni* and cf. *yakimorum* morphotypes, but the significant variation in the mitochondrial genes (COI distances up to 3%) suggests the possibility of a mixing of Little Lost and Pahsimeroi stocks after a period of isolation. Therefore, the occurrence of a *yakimorum*-like morphotype is possibly the result of hybridization and/or introgression. Both the Pahsimeroi and Little Lost Rivers lie within a common trench, and local geology suggests drainage reversals, possibly between the late Pliocene and mid-Pleistocene (Ruppel 1967). *Rhynchelmis gustafsoni* seems associated with the Snake River to the south, but the *yakimorum* group morphotype associated with the Salmon River was described as *Rhynchelmis monsserratus* by Fend & Brinkhurst (2000); the latter was not available for this study, but is easily distinguished from *R. yakimorum* by its single median spermatheca. Clearly, resolution of this problem will require testing of more populations in the region, including examples of *R. monsserratus*. Two *R. gustafsoni* populations (ILK & ICB) from the neighbouring Birch Creek (also a 'lost' river) have a unique COI sequence, but the ITS tree does not suggest long-term reproductive isolation of the Birch Creek worms from those of Little Lost/Pahsimeroi rivers.

Rhynchelmis gustafsoni from a more distant population, Henrys Fork (IFO, Fig. 1), and the Big Lost River *gustafsoni* (ICA, Fig. 2) population also appear closely related to the terminal group of *R.* cf. *yakimorum* and other *R. gustafsoni* (see Fig. 6). The COI of the ICA worm is roughly 4% different from the corresponding haplotypes of the other *gustafsoni*/cf. *yakimorum* samples, and the two available 16S sequences from the ICA population are also separated from those of the other populations (see Fig. 4). The best solution at this point is to retain *gustafsoni* as the upper Snake River species, and to expand its morphological description to include the Little Lost River worms. The single *R. gustafsoni* population from the south side of the Snake River (IOT, Fig. 1) was only sequenced for ITS; although it grouped with the populations on the north side, its relationship to them was unresolved (Fig. 5).

Two widely separated populations of the *yakimorum* variant 1 morphotype (NEE, ICN; Fig. 1) represent major tributaries of the lower Snake River (the Owyhee and Clearwater rivers). The COI sequences of these two populations differ by only 1.1%, but they both differ from all analysed specimens of *R. gustafsoni* and *R.* cf. *yakimorum* by 3.5–4.7%. Moreover, the tree of the nuclear ITS gene (Fig. 5) suggests, with good support, that these populations have been isolated from the latter for a considerable amount of time.

The well-supported group including all of the *R. utahensis* populations (Fig. 6) is a morphotype that differs from the other *yakimorum* group species not only in positions of reproductive structures, but also in their morphology (Figs S1 and S2). Most collections are from tributaries to the Bear River, which currently drains southward to the Bonneville (present day Great Salt Lake) drainage. Although the worms are morphologically similar (compare Figs S1H and S2J–K with figs 36–40 in Fend & Brinkhurst 2000), the COI of the McClenden Spring (Snake River drainage) *utahensis* population (ICM, Fig. 1) differs by about 10% from that of the Bear River drainage populations (IBP, IBPP, IBS). As the coherent ITS of the latter populations (see Fig. 5) appears to indicate reproductive isolation, the *utahensis* morphotype is likely to contain at least two sibling species. The Snake and Bear Rivers have had repeated connections through the Pleistocene, and their fish and mollusc faunas show many similarities (Taylor 1985; Minckley *et al.* 1986). In this case, however, more localized factors may contribute to isolation of these populations. McClenden Spring (ICM) is isolated within a rather arid landscape, and surface flow normally does not extend to a Snake tributary.

Two of the *yakimorum* variant 1 morphotype populations analysed in the present study were collected in the Umpqua and Rogue River basins in southern coastal Oregon (OCU, ORG; Fig 1). If the other *yakimorum* complex species are retained as separate taxa, the coastal populations cannot be conspecific with the morphologically similar populations of the lower Snake River tributaries (NEE, ICN); thus OCU and ORG most likely represent a cryptic species, which is the sister group to all of the Snake River species.

The morphotype identified as *R. rostrata* is associated with both west coast and more inland drainages, the latter including tributaries to the Snake River and Great Salt Lake (Fend & Brinkhurst 2000), but it appears to be absent from the intervening Great Basin region. The great genetic differences, particularly in COI (Fig. 3) but also in ITS (Fig. 5), between the populations sampled here suggest that they have been isolated for a considerable time, and may represent cryptic species The most distinctive *R. rostrata* population analysed is CPD (Fig. 1), i.e. a worm from the type locality of *R. alpestris* (Eisen), later synonymized with *R. rostrata* by Brinkhurst & Cook (1966). This population was considered a variant by Fend & Brinkhurst (2000) because of a tendency towards more highly branched spermathecal diverticula. The CPD record is from a spring at Donner Lake, associated with the inland-draining Truckee River (Lahontan drainage), rather than the Sacramento-San Joaquin Rivers, where the typical *R. rostrata* morphotype is widespread (S. Fend, unpublished data**).**

Although *R. elrodi* and *R. glandula* are morphologically similar enough that they have been combined in the past (Brinkhurst & Cook 1966), they are genetically different and their distributions appear disjunct, with *elrodi* inhabiting the Rocky Mountains and *glandula* found only in the Coast Ranges.

Conclusions

Our study demonstrated the utility of a combined analysis of mtDNA (COI, 16S) and nuclear (ITS) sequences as molecular markers for phylogenetic reconstruction of Nearctic *Rhychelmis* species. Neither the morphology by itself nor a single mitochondrial gene is enough to delineate all species within the *yakimorum* species complex, but our combined genetic approach, including nuclear data, suggests that the *Sutroa* clade of *Rhynchelmis* contains several cryptic species, hidden among forms hitherto identified as nominal morphospecies or as 'variants' thereof (*R. rostrata*, *R. utahensis*, '*R. yakimorum* variant 1'). The status of *R. gustafsoni* vis-à-vis Idaho specimens identified as '*R.* cf. *yakimorum*' is more uncertain. Although there is genetic variation within this lot, it seems possible that they all constitute a single, but morphologically heterogeneous species. Some of this heterogeneity may be

due to hybridization with populations morphologically more similar to *yakimorum*.

Although COI barcoding as an identification system for animal species is generally supported by our study, caution should be made when using COI gene divergence to signal the boundary of those species which are closely allied or otherwise have a wide and disjunct distribution. In these cases, in particular, COI data should be supplemented by nuclear genetic markers (which may provide evidence of ceased gene flow) as well as by morphological and biogeographical information.

Acknowledgements

This study was supported by the Swedish Research Council (grant #621–2004–2397 to CE). We are grateful to Stefan Lundberg, Swedish Museum of Natural History, for providing specimens, to Anna Ansebo, Department of Zoology, University of Gothenburg, for technical assistance, and to two anonymous reviewers for excellent suggestions for improvement of the manuscript.

References

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds) *Proceedings of the Second International Symposium on Information Theory* (pp. 267–281). Budapest: Akademiai Kiado.

Brinkhurst, R. O. (1989). A phylogenetic analysis of the Lumbriculidae (Annelida, Oligochaeta). *Canadian Journal of Zoology*, *67*, 2731–2739.

Brinkhurst, R. O. & Cook, D. G. (1966). Studies on the North American aquatic Oligochaeta Part III. Lumbriculidae and additional notes and records of other species. *Proceedings of the Academy of Natural Sciences of Philadelphia*, *118*, 1–33.

Chen, Y., Xiao, H., Fu, J. & Huang, D.-W. (2004). A molecular phylogeny of eurytomid wasps inferred from DNA sequence data of 28S, 18S, 16S, and COI genes. *Molecular Phylogenetics and Evolution*, *31*, 300–307.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797.

Eisen, G. (1888). On the anatomy of *Sutroa rostrata*, a new annelid of the family Lumbriculina. *Memoirs of the Californian Academy of Sciences*, *2*, 3–8.

Erséus, C. & Gustafsson, D. (2009). Cryptic speciation in clitellate model organisms. In D. Shain (Ed) *Annelids in Modern Biology* (pp. 31–46). Hoboken, NJ: John Wiley & Sons.

Erséus, C. & Kvist, S. (2007). COI variation in Scandinavian marine species of *Tubificoides* (Annelida, Clitellata, Tubificidae). *Journal of the Marine Biological Association of the United Kingdom*, *87*, 1121–1126.

Fend, S. V. & Brinkhurst, R. O. (2010). New species of *Rhynchelmis* (Clitellata, Lumbriculidae), with observations on the Nearctic species. *Hydrobiologia*, *428*, 1–59.

Fend, S. V. & Brinkhurst, R. O. (2010). Contributions towards a review of the genus *Rhynchelmis* Hoffmeister (Clitellata: Lumbriculidae). *Zootaxa*, *2407*, 1–27.

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.

Gustafsson, D. R., Price, D. A. & Erséus, C. (2009). Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Molecular Phylogenetics and Evolution*, *51*, 182–189.

Harrison, J. S. (2004). Evolution, biogeography, and the utility of mitochondrial 16S and COI genes in phylogenetic analysis of the crab genus *Austinixa* (Decapoda: Pinnotheridae). *Molecular Phylogenetics and Evolution*, *30*, 743–754.

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. (2003). Barcoding animal life: cytochrome *c* oxidase subunit I divergences among closely related species. *Proceedings of the Royal Society, London, B (Supplement)*, *270*, 96–99.

Holmquist, C. (1976). Lumbriculids (Oligochaeta) of northern Alaska and northwestern Canada. *Zoologische Jahrbücher Abteilung für Systematik, Ökologie und Geographie der Tiere*, *103*, 337–431.

Hrabě, S. (1936). Über *Dorydrilus* (*Piguetia*) *mirabilis* n. subg. n. sp. aus einem Sodbrunnen in der Umbebung von Basel sowie über *Dorydrilus* (*Dorydrilus*) *michaelseni* Pig. und *Bichaeta sanguinea* Bret. *Spisy vydávané Přírodovédeckou Fakultou Masarykovy University*, *227*, 3–18.

Hrabě, S. (1982). Contribution to the knowledge of Oligochaeta from the Lake Baikal. *Věstni. Československé Společnosti Zoologické*, *46*, 129–146.

Huelsenbeck, J. P., Larget, B., Miller, R. E. & Ronquist, F. (2002). Potential applications and pitfalls of Bayesian inferences of phylogeny. *Systematic Biology*, *51*, 673–688.

Huelsenbeck, J. P., van der Larget, B., Mark, P. & Ronquist, F. (2005). *MrBayes: A Program for the Bayesian Inference of Phylogeny. Version 3.1.1*. Available via http://mrbayes.csit.fsu.edu.

Källersjö, M., von Proschwitz, T., Lundberg, S., Eldenäs, P. & Erséus, C. (2005). Evaluation of ITS rDNA as a complement to mitochondrial gene sequences for phylogenetic studies in freshwater mussels: an example using Unionidae from northwestern Europe. *Zoologica Scripta*, *34*, 415–424.

Kathman, R. D. & Brinkhurst, R. O. (1998). *Guide to the Freshwater Oligochaetes of North America* (264 pp). College Grove, TN: Aquatic Resources Center.

Kaygorodova, I. A. & Liventseva, V. G. (2007). A test of monophyly of the genus *Rhynchelmis* Hoffmeister, 1843 (Oligochaeta, Lumbriculidae): morphology based phylogeny. *Investigated in Russia*, *10*, 1517–1527.

Kaygorodova, I. A., Sherbakov, D. Yu. & Martin, P. (2007). Molecular phylogeny of Baikalian Lumbriculidae (Oligochaeta): evidence for recent explosive speciation. *Comparative Cytogenetics*, *1*, 71–84.

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology*, *38*, 7–25.

Kluge, A. G. (1998). Total evidence or taxonomic congruence: Cladistics or consensus classification. *Cladistics*, *14*, 151–158.

Martin, P., Ferraguti, M. & Kaygorodova, I. (1998). Description of two new species of *Rhynchelmis* (Oligochaeta: Lumbriculidae) from Lake Baikal (Russia), using classical morphology and ultrastructure of spermatozoa. *Annales de Limnologie*, *34*, 283–293.

Meyer, C. P. & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, *3*, 2229–2238.

Minckley, W. L., Hendrickson, D. A. & Bond, C. E. (1986). Geography of western North American fishes: description and relationships to intracontinental tectonism. In C. H. Hocutt & E. O. Wiley (Eds) *The Zoogeography of North American Freshwater Fishes* (pp. 519–613). New York: John Wiley and Sons, Inc.

Moritz, C. & Cicero, C. (2004). DNA barcoding: promise and pitfalls. *PLoS Biology*, *2*, 1529–1531.

Nylander, J. A. A. (2004). *Mr Modeltest. Version 2.2*. Available via http://www.ebc.uu.se/systzoo/staff/nylander.html.

Nylander, J. A. A., Ronquist, F., Huelsenbeck, J. P. & Nieves-Aldrey, J. L. (2004). Bayesian phylogenetic analysis of combined data. *Systematic Biology*, *53*, 1–21.

Palumbi, S. R. (1996). Nucleic acids II: the polymerase chain reaction. In D. M. Hillis, C. Moritz & B. K Mable (Eds) *Molecular Systematics*, 2nd ed. (pp. 205–247). Sunderland, MA: Sinauer Associates.

Ronquist, F. & Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574.

Ruppel, E. T. (1967). Late Cenozoic drainage reversal, east-central Idaho, and its relation to possible undiscovered placer deposits. *Economic Geology*, *62*, 648–663.

Sjölin, E., Erséus, C. & Källersjö, M. (2005). Phylogeny of Tubificidae (Annelida, Clitellata) based on mitochondrial and nuclear sequence data. *Molecular Phylogenetics and Evolution*, *35*, 431–441.

Sokol'skaya, N. L. (1983). Freshwater Oligochaeta of Kamchatka and the Koryak upland. *Sbornik Trudov Zoologicheskogo Muzeya Moskovskogo Gosudarstvennogo Universiteta*, *20*, 22–119.

Swofford, D. L. (2003). *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*. Sunderland, MA: Sinauer Associates.

Taylor, D. W. (1985). Evolution of freshwater drainages and molluscs in western North America. In C. J. Smiley & A. J. Leviton (Eds) *Late Cenozoic History of the Pacific Northwest, Interdisciplinary Studies of the Clarkia Fossil Beds of Northern Idaho*. (pp. 265–321). San Francisco: American Association for the Advancement of Science.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, *24*, 4876–4882.

Timm, T. (1979). The Estonian Lumbriculidae. *Eesti NSV Teadoste Akadeemia Toimetised (Biologia)*, *28*, 158–171.

Timm, T. (1997). Freshwater Oligochaeta of some urban watercourses in the Russian Far East. *Internationale Revue der gesamten Hydrobiologie*, *82*, 437–467.

White, T. J., Bruns, T. D., Lee, S. & Taylor, J. W. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. A. Innis, D. H. Gelfand, J. J. Sninsky & T. J. White (Eds) *PCR Protocols: A Guide to Methods and Applications*. (pp. 315–322). New York: Academic Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Genital segments of *Rhynchelmis yakimorum* complex species used in the analyses; left column with porophores retracted, right column with porophores everted or partially everted. All specimens are facing to the left, so the spermathecal pores (in segment VIII) are to the left of the male pores (in segment X). —A and B. *Rhynchelmis yakimorum* from Warm Spring, Idaho (A; IBW) and Iron Spring, Idaho (B; ICI). —C and D. *Rhynchelmis yakimorum* variant 1 from Cow Creek, Oregon (C; OCU) and Clearwater River, Idaho (D; ICN). —E and F. *Rhynchelmis gustafsoni* from Kaufman Cabin Spring, Idaho (ILK). – G and H. *Rhynchelmis utahensis* from McClenden Spring, Idaho (ICM).

Fig. S2 Transverse views of *Rhynchelmis yakimorum* complex species used in the analyses; all reconstructed from serial sections. Columns from left to right show spermathecal pores, male pores with penial structures (porophores) retracted, male pores with porophores everted. —A–C. *Rhynchelmis yakimorum* from Iron Springs, Idaho (ICI); typical morph with male and spermathecal pores on chaetal lines. —D–F. *Rhynchelmis yakimorum* variant 1 from Clearwater River, Idaho (ICN); spermathecal pores are closely appressed within a single muscular-glandular 'bulb'.—G–I. *Rhynchelmis gustafsoni* Warm Spring (IBW), Doublespring (ICD), and Kaufman Cabin Spring (ILK) (all eastern Idaho); with both sets of pores closely appressed; the everted porophores fused. —J–L. *Rhynchelmis utahensis* from Blue Pond Spring, Idaho (IBBP) (J–K) and Rick Spring, Utah (near type locality); male structures are large and acuminate when everted. Scale bars 500 μm.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.