



The mitochondrial genome of UK (non-native) *Dikerogammarus haemobaphes* (Amphipoda: Gammaridae) informs upon *Dikerogammarus* evolution, invasions and associated microparasites

Jamie Bojko

Received: 17 May 2019 / Revised: 24 September 2019 / Accepted: 27 September 2019 / Published online: 14 October 2019
© Springer Nature Switzerland AG 2019

Abstract The amphipod *Dikerogammarus haemobaphes* is a high-risk carrier of parasites that impact wildlife in its non-native range. Studies using the mitochondrial genes, Cytochrome Oxidase Sub-Unit 1 (cox1) and small-subunit ribosomal RNA gene (16S), provide some nucleotide detail for understanding the evolution and phylogeography of this species. Despite this, the origins of the invasion remain unknown, as do the origins of its parasites. This study provides the full annotated mitochondrial genome (15,460 bp) of *D. haemobaphes*, consisting of 2 rRNAs, 24 tRNAs and 14 protein coding genes. Mitochondrial genes from the UK isolate are compared to existing data on NCBI and are used in a concatenated phylogenetic approach and identify *D. haemobaphes* as an early member of the Gammaridae (Amphipoda). Viral, bacterial, protistan and microsporidian parasites are present across the Gammaridae, including *D. haemobaphes*, suggesting

the ancestor of the Gammaridae harboured related diseases, and that further screening of amphipods is likely to reveal further microparasite diversity. This correlation suggests that other gammarid invaders have the potential to harbour a range of microparasites. The mitochondrial genome of this species will act a resource to facilitate our understanding of gene flow, disease epidemiology and evolutionary history in this invasion-disease model.

Keywords Parasitology · Invasive · Amphipoda · Invasion biology · Cytochrome oxidase

Introduction

The demon shrimp, *Dikerogammarus haemobaphes* (Eichwald 1841), is a non-native freshwater amphipod in the UK that exerts low levels of ecological damage and inter-species competition (Bovy et al., 2015). The species hosts multiple mortality-inducing and behaviour-altering pathogens that have been carried alongside the invasion into the UK (Bojko et al., 2018a). Infection with the microsporidian pathogen *Cucumispora ornata* Bojko, Dunn, Stebbing, Ross, Kerr, Stentford 2015 was noted to reduce activity in heavily infected hosts and was associated with mortality in both *D. haemobaphes* and non-target *Gammarus pulex* (L.), which also have the infection in wild populations

Handling editor: Diego Fontaneto

J. Bojko (✉)
Emerging Pathogens Institute, University of Florida,
Gainesville, FL 32611, USA
e-mail: Jamie.Bojko@ufl.edu

J. Bojko
Institute of Food and Agricultural Sciences, University of
Florida, Gainesville, FL 32611, USA

J. Bojko
School of Forest Resource and Conservation, University
of Florida, Gainesville, FL 32611, USA

(Bojko et al., 2018a). ‘Dikerogammarus haemobaphes bi-facies-like Virus’ (DhbflV), was also identified as a mortality-inducing virus at low prevalence within the *D. haemobaphes* population in the UK (Bojko et al., 2018a). Finally, a likely novel member of the *Nudiviridae*, ‘Dikerogammarus haemobaphes Bacilliform Virus’ (DhBV) was found to increase the activity of its host and potentially alter the rate of invasion spread (Bojko et al., 2018a).

This species, and specifically its parasites, are now considered a high-risk invasion system that requires the development of diagnostic methods to track the invasion, associated diseases and their effects. To date, mitochondrial data for this species are restricted to short ~ 600 bp sequence tags of the Cytochrome Oxidase Sub-Unit 1 gene (*cox1*) (Grabner et al., 2015) and partial 16S. Next generation sequencing platforms and bioinformatic tools provide the ability to rapidly provide data on the genomic composition of the demon shrimp and aid the development of diagnostic tools. Recent advances in the sequencing of mitochondrial genomes from amphipods has also allowed for increased phylogenetic information, with an excess of 50 mitochondrial genomes being available for Amphipoda (Romanova et al., 2016; Macher et al., 2017; Cormier et al., 2018).

Herein, the mitochondrial genome of the demon shrimp is presented. The mitochondrial genome of this UK-based individual will act as a resource to develop additional PCR diagnostics for population genetics studies to determine the genetic diversity and likely origins of invasive populations. Furthermore, this genome provides detailed information on the evolution of *Dikerogammarus* sp. and can be used in tandem with disease screening data to identify the potential origins of its parasites.

Materials and methods

Specimen collection and mitochondrial genome assembly

In 2016, a single animal was collected by hand from Carlton Brook, UK (British National Grid [BNG] ref: SK3870004400). The urosome of this individual underwent phenol:chloroform DNA extraction after an overnight digestion with Proteinase K. This extract was prepared into a DNA library using a NEXTERA-

XT library preparation method for MiSeq sequencing (Illumina; www.illumina.com) and Illumina TruSeq® DNA PCR-Free library preparation kit for HiSeq (Illumina; www.illumina.com). Raw data were trimmed (Illuminatrim-TRIMMOMATIC) and then assembled using SPAdes v.3.13.0 (default settings with km: 21, 33, 55, 77, 99, 127) (Bankevich et al., 2012; Bolger et al., 2014).

This resulted in a 15,460 bp circular contig with 243.97X coverage. Trimmed reads were re-aligned to the sequence to confirm even coverage across the circular sequence. This sequence was submitted to MITOS (invertebrate) to provide detailed annotation of protein coding (PCG) and non-coding RNA (ncRNA) genetic regions (Bernt et al., 2013), which were further edited and confirmed using data available on NCBI. Individual ncRNA and PCGs were compared to available sequence data from alternative *D. haemobaphes* and other Amphipoda using NCBI, BLASTp and BLASTn. Circa (www.omgenomics.com/circa) and CLC (www.qiagenbioinformatics.com) were used to develop diagrammatic representations of the genetic data.

Sequence data for the *D. haemobaphes* mitochondrial genome can be acquired from NCBI (accession number: MK644228).

Phylogenetic analyses

Three maximum likelihood phylogenetic trees were calculated using the mitochondrial genome of *D. haemobaphes*. The first two used the 16S (276 positions) or *cox1* (614 positions) gene to compare *Dikerogammarus* sp. from NCBI ($n = 16$ sequences and 39 sequences, respectively) (evolutionary model: HKY + F + I). The final tree used individually aligned and subsequently concatenated amino acid (AA) sequences (13 genes: *atp6*, *atp8-0*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*) ($n = 38$ Amphipoda and 1 Isopoda outgroup) (evolutionary model: mtInv + F + I + G4). In all cases the sequences were trimmed and aligned using MAFFT in Geneious v.10.0.2 (gap: 1.53, cost: 0.123) before phylogenetic analysis and model matching according to Bayesian Information Criterion (BIC). IQtree was used to calculate the phylogenetic trees (Nguyen et al., 2015) and included the use of ultrafast approximated bootstraps ($n = 1000$) (Minh et al. 2013). ‘+F’ refers to the empirical base frequencies

and counts base frequencies directly from the alignment; ‘+I’ refers to a fix the proportion of invariable sites; finally, ‘+G4’ refers to the addition of the discrete Gamma model.

Multiple sources of literature were used to compare known microparasites [*Nudiviridae*, ‘*Candidatus* *Aquirickettsiella*’, Microsporidia (*Cucumispora* and *Dictyocoela*) and gregarines (Apicomplexa)] of each amphipod with a known mitochondrial genome to the phylogenetic information determined by this study (Madyarova et al., 2015; Bojko, 2017; Bojko et al., 2017; Bacela-Spychalska et al., 2018; Dimova et al., 2018; Ironside & Wilkinson, 2018; Bojko & Ovcharenko, 2019).

Results

Mitochondrial genome composition and similarity

The mitochondrial genome of *D. haemobaphes* is 15,460 bp in length (coverage = 243.97%) and encodes 24 tRNA, 2 rRNA and 14 protein coding genes (including a duplication of *atp8*) (Table 1; Fig. 1). The closest associated genome is that of *Gammarus duebeni* (NC017760), which shares two closely related tRNAs and six protein coding genes, primarily linked with the cytochrome complex. The *cox1* and 16S (*rrnL*) genes of the *D. haemobaphes* mitochondrial genome showed closest similarity to *D. haemobaphes* haplotypes from Germany (Main River and North Rhine-Westphalia) (Table 1).

Structurally, the mitochondrial genome is A+T rich with 33.8% GC content across the circular genome. The closest relatives with full mitochondrial genome availability were *Gammarus duebeni* Lilljeborg, 1852 (NC017760) and *Eulimnogammarus cyaneus* (Dybowsky, 1874) (NC033360), which show high levels of relative gene organisation along the circular mitochondrial genome but with some small reorganisation of tRNAs. The *trnR* and *trnE* are present in that order instead of *trnE-trnR* as seen in the genomes of *G. duebeni* and *E. cyaneus* (Fig. 2). *Dikerogammarus haemobaphes* also has a duplication of the *trnQ*. A duplication of *atp8*, which is part of the Adenosine Tri-Phosphate synthesis pathway (genes: *atp8-0* and *atp8-1*), is present.

Phylogenetics of related species and conspecifics of the genus *Dikerogammarus*

The 16S phylogeny, incorporating multiple *Dikerogammarus* sp., identifies a *Dikerogammarus villosus* (Sowinsky, 1894) group (bootstrap = 100), a *D. haemobaphes* group (bootstrap = 98), a *Dikerogammarus caspius* (Pallas, 1771) group and a *Dikerogammarus bispinosus* Martynov, 1925 group. The *D. haemobaphes* 16S isolate from the UK branches alongside an isolate from Poland (Vistula River) (bootstrap = 51) (Fig. 3), but the UK isolate is genetically dissimilar by 3.27%, suggesting that although this is the closest isolate, they are not genetically identical. Greater genetic variation is visible between the *D. haemobaphes* 16S gene than the *D. villosus* 16S data (Fig. 3).

The tree based on the *cox1* gene also results in two distinct clades of *D. villosus* and *D. haemobaphes*, both with high (95–100) bootstrap support. The UK isolate of *D. haemobaphes* shows closest nucleotide similarity to a *D. haemobaphes* haplotype 1 (KY075268) sampled in Germany (sim. = 100%, cov. = 100%, *e*-value = 0.0). In Fig. 3 all the *D. haemobaphes* isolates branch together apart from one individual (AY529049), which originates from the North Caspian Sea, the species native range.

A concatenated phylogeny of all available mitochondrial protein sequences (*n* = 13 genes) from available amphipod mitochondrial genomes confirms that *D. haemobaphes* sits within the Gammaridae, but also identifies the species as an early branching member (bootstrap support = 100) relative to the *Gammarus* genus and other species from Europe and the Ponto-Caspian region (Fig. 4). Other Amphipoda show predicted branching throughout the tree, with all genera represented by multiple species (*Eulimnogammarus*, *Gammarus*, *Platorchestia*, *Hyallela*, *Epimeria*, *Pseudoniphargus*, *Stygobromus*, *Metacrangonyx* and *Caprella*) branching together (Fig. 4). The tree shows low bootstrap support close to the root (55 or 38), suggesting that further sequencing of highly derived amphipods may help to add detail to the tree and its topology, providing further detail to the tree and increase its accuracy at predicting topology.

Using available literature (Madyarova et al., 2015; Bojko, 2017; Bojko et al., 2017; Bacela-Spychalska et al., 2018; Dimova et al., 2018; Ironside & Wilkinson, 2018; Bojko & Ovcharenko, 2019), the known

Table 1 A table including each coding region on the mitochondrial genome of *Dikerogammarus haemobaphes* in addition the nucleotide and translated protein similarity of each coding region as determined via BLASTN and BLASTP comparison to existing NCBI data

Genome	Start (+)	End (+)	Gene	Strand	Gene hit	Gene similarity	Gene coverage
<i>D. haemobaphes</i> mitochondrial genome	1	1550	cox1	+	<i>Dikerogammarus haemobaphes</i> isolate Dh_H01 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	100.00	42
	1551	1611	trnL2(tta)	+	–	–	–
	1611	2266	cox2	+	<i>Gammarus lacustris</i> mitochondrion, complete genome	74.81	99
	2277	2339	trnK(aaa)	+	–	–	–
	2340	2401	trnD(gac)	+	<i>Gammarus duebeni</i> mitochondrion, complete genome	92.06	100
	2402	2560	atp8-0	+	–	–	–
	2554	3225	atp6	+	<i>Brachyuropus grewingkii</i> mitochondrion, complete genome	70.08	94
	3225	4010	cox3	+	<i>Pseudoniphargus mercadali</i> mitochondrion, partial genome	73.32	98
	4017	4368	nad3	+	<i>Eulimnogammarus verrucosus</i> mitochondrion, complete genome	72.59	76
	4369	4431	trnA(gca)	+	–	–	–
	4431	4483	trnS1(aga)	+	–	–	–
	4482	4542	trnN(aac)	+	<i>Coenobita perlatus</i> isolate CP140 mitochondrion, complete genome	95.65	75
	4547	4598	trnR(cga)	+	–	–	–
	4603	4664	trnE(gaa)	+	<i>Gammarus roeselii</i> mitochondrion, complete genome	88.52	96
	4666	4725	trnF(ttc)	–	–	–	–
	4730	6394	nad5	–	<i>Pontogammarus maeoticus</i> isolate N8 NADH dehydrogenase subunit 5 (ND5) gene, partial cds; mitochondrial	81.80	37
	6469	6529	trnH(cac)	–	–	–	–
	6529	7794	nad4	–	<i>Gammarus duebeni</i> mitochondrion, complete genome	71.11	86
	7836	8111	nad4L	–	<i>Solenocera crassicornis</i> mitochondrion, complete genome	94.44	13
	8131	8191	trnT(aca)	+	–	–	–
	8191	8256	trnP(cca)	–	–	–	–
	8266	8766	nad6	+	<i>Gammarus lacustris</i> mitochondrion, complete genome	84.95	18
	8778	9900	cob	+	<i>Gammarus duebeni</i> mitochondrion, complete genome	73.45	94
	9901	9959	trnS2(tca)	+	–	–	–
	9962	10,885	nad1	–	<i>Gammarus duebeni</i> mitochondrion, complete genome	69.79	92
	10,895	10,955	trnL1(cta)	–	–	–	–

Table 1 continued

Genome	Start (+)	End (+)	Gene	Strand	Gene hit	Gene similarity	Gene coverage
	10,956	11,982	rrnL	–	<i>Dikerogammarus haemobaphes</i> mitochondrial partial 16S rRNA gene, haplotype H4, from Germany:Main	99.37	56
	11,929	11,982	trnV(gta)	–	–	–	–
	11,982	12,723	rrnS	–	<i>Gammarus roeselii</i> mitochondrion, complete genome	74.34	53
	13,164	13,274	atp8-1	–	Zebrafish DNA sequence from clone DKEY-174N5 in linkage group 21, complete sequence	82.46	52
	13,787	13,846	trnY(tac)	–	<i>Gammarus duebeni</i> mitochondrion, complete genome	100	60
	13,847	13,903	trnQ(caa)	–	–	–	–
	14,037	14,097	trnQ(caa)	–	–	–	–
	14,097	14,153	trnC(tgc)	–	–	–	–
	14,228	14,291	trnI(ata)	+	–	–	–
	14,295	14,356	trnM(atg)	+	–	–	–
	14,387	15,307	nad2	+	<i>Gammarus duebeni</i> mitochondrion, complete genome	65.06	97
	15,339	15,399	trnW(tga)	+	–	–	–
	15,400	15,460	trnG(gga)	+	<i>Eulimnogammarus vittatus</i> mitochondrion, complete genome	95.65	75

Genome	Gene-e-value	Gene accession	Protein hit	Protein similarity	Protein coverage	Protein-e-value	Protein accession
<i>D. haemobaphes</i> mitochondrial genome	0	KT075268.1	COX1 gene product [Gammarus duebeni]	88.85	99	0	YP_006234443.1
	–	–	–	–	–	–	–
	3.00E–116	NC_044469.1	COX2 gene product [Gammarus duebeni]	82.95	100	2.00E – 136	YP_006234444.1
	–	–	–	–	–	–	–
	4.00E–12	JN704067.1	–	–	–	–	–
	–	–	–	–	–	–	–
	5.00E–62	KP161875.1	ATP synthase F0 subunit 6 [Eulimnogammarus cyaneus]	72.73	100	2.00E – 108	YP_009339283.1
	2.00E–118	MH592131.1	cytochrome c oxidase subunit 3 [Pseudoniphargus cupicola]	75.77	99	6.00E – 143	AXI98550.1
	1.00E–31	KF690638.1	NADH dehydrogenase subunit 3 [Eulimnogammarus vittatus]	59.83	100	1.00E – 40	YP_009107165.1
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–

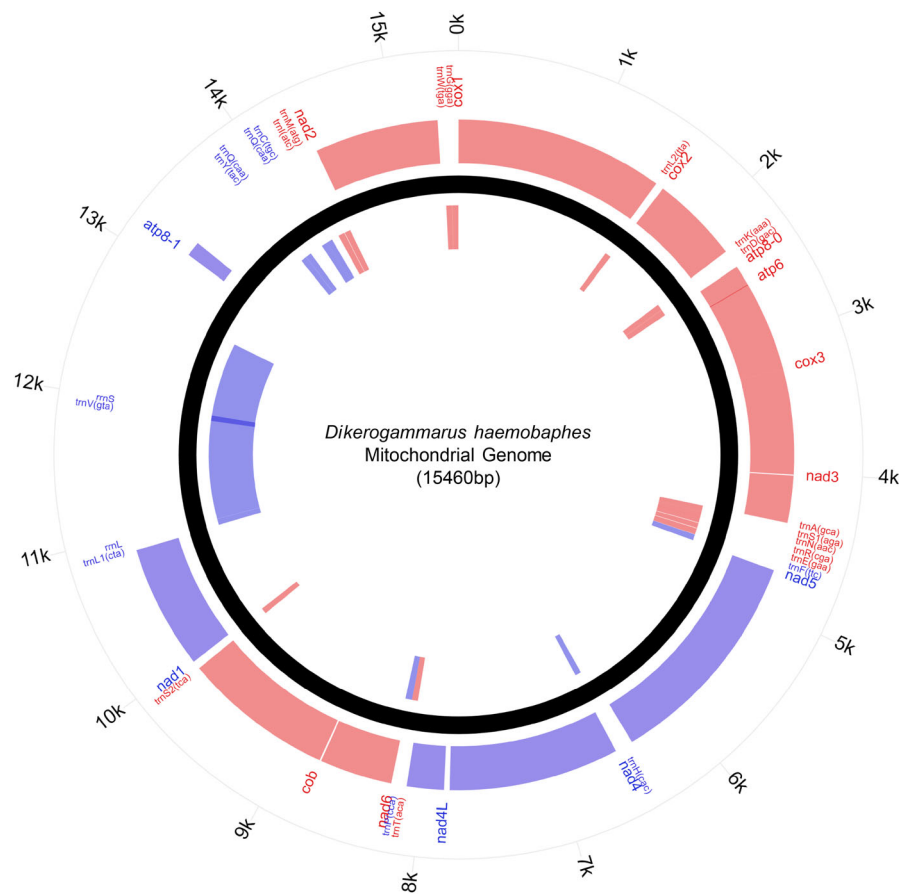
Table 1 continued

Genome	Gene-e-value	Gene accession	Protein hit	Protein similarity	Protein coverage	Protein-e-value	Protein accession
	8.00E–09	KY352234.1	–	–	–	–	–
	–	–	–	–	–	–	–
	3.00E–08	NC_037481.1	–	–	–	–	–
	–	–	–	–	–	–	–
	1.00E–177	KC797017.1	NADH dehydrogenase subunit 5 [Linevichella vortex]	57.56	92	0	APL97252.1
	–	–	–	–	–	–	–
	4.00E–140	JN704067.1	ND4 gene product [Gammarus duebeni]	66.83	99	0	YP_006234450.1
	0.002	MF379621.1	ND4L gene product [Gammarus duebeni]	68.29	90	5.00E – 25	YP_006234451.1
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	5.00E–18	NC_044469.1	NADH dehydrogenase subunit 6 [Limulus polyphemus]	47.93	96	1.00E – 23	APL97219.1
	3.00E–173	JN704067.1	CYTB gene product [Gammarus duebeni]	79.40	97	0	YP_006234453.1
	–	–	–	–	–	–	–
	2.00E–92	JN704067.1	ND1 gene product [Gammarus duebeni]	70.71	96	2.00E – 140	YP_006234454.1
	–	–	–	–	–	–	–
	1.00E–160	AJ440890.1	–	–	–	–	–
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	2.00E–34	NC_037481.1	–	–	–	–	–
	0.008	CR352267.11	–	–	–	–	–
	1.00E–06	JN704067.1	–	–	–	–	–
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	–	–	–	–	–	–	–
	3.00E–27	JN704067.1	NADH dehydrogenase subunit 2 [Brachyuropus grewingkii]	53.65	92	1.00E – 105	YP_009118052.1
	–	–	–	–	–	–	–
	8.00E–09	KM287572.1	–	–	–	–	–

microparasites of those amphipods from Europe and the Ponto-Caspian region are presented alongside the phylogenetics conducted by this study to identify possible points of parasite evolution and yet undetermined hosts that may harbour infection (Fig. 5). Bacilliform viruses (*Nudiviridae*), intracellular

bacteria (*Candidatus* *Aquirickettsiella*), species from two genera of Microsporidia (*Cucumispora* and *Dictyocoela*) and the presence of gregarines (Apicomplexa) are presented on the tree alongside known hosts (Fig. 5). Bacilliform viruses are present in two *Gammarus* sp. and *D. haemobaphes*, which sit at the

Fig. 1 A map of the circular mitochondrial genome of *Dikerogammarus haemobaphes*. The genome is represented as a single circular black line. Protein coding genes are present on the outside of the black circle, with positive strand sequences in red and negative strand sequences in blue. Non-coding RNA sequences are represented internal to the black circle, with positive strand coding regions in red and negative strand sequences in blue. The labels for each protein coding gene or ncRNA gene are listed around the outside of the diagram before the genome size markers. Please refer to the NCBI accession MK644228 for electronic annotation



base of the Gammaridae in Fig. 4. These same individuals also host intracellular bacteria (*Candidatus Aquirickettsiella*). Systematically identified *Cucumispora* sp. are restricted to two host species with mitochondrial genome data, *D. haemobaphes* and *Gammarus roeselii* (L.); however, multiple SSU sequences from Ponto-Caspian amphipod hosts place *Cucumispora* candidates across Fig. 5. Gregarines have been observed in *D. haemobaphes*, all *Gammarus* sp. and in members of the *Eulimnogammarus* genus (Fig. 5).

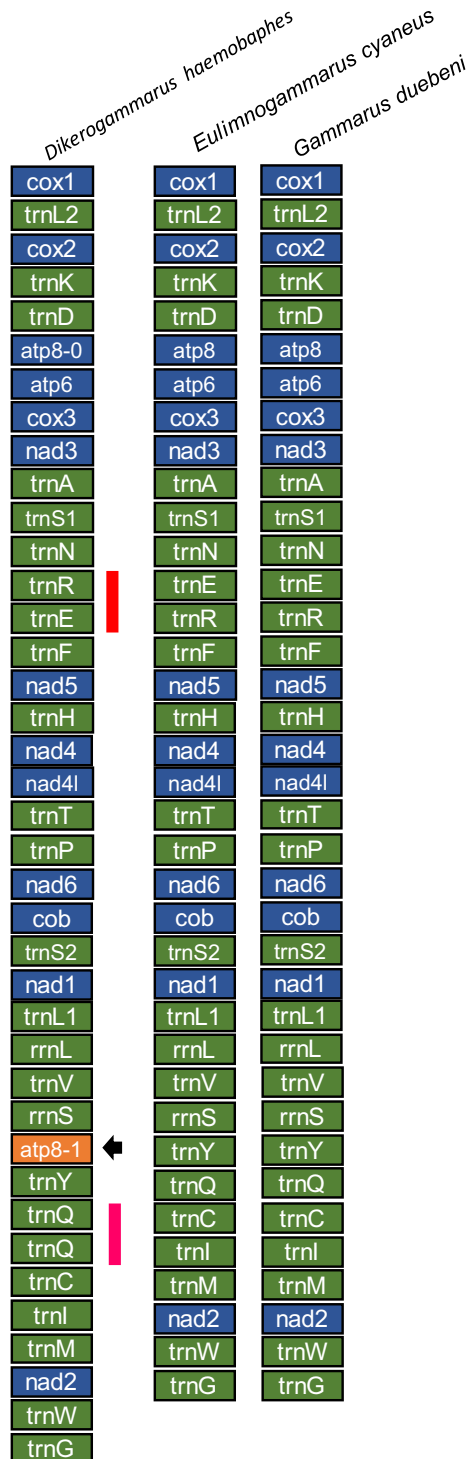
Two species, *Pallaseopsis kessleri* (Dyb.) and *Crypturopus tuberculatus* (Dyb.), do not yet have any identified microparasite groups explored herein.

Discussion

The mitochondrial genomes of eukaryotic organisms have been used to infer phylogenetic relationships

(Cormier et al., 2018), to understand energy and metabolism (Abele et al., 2007) and to better inform upon the genetic diversity of a population (Ma et al., 2015). Increased availability of mitochondrial data associated with biological invasions can provide a valuable resource to better understand invasion dynamics through population genetics. This information can be used to determine potential entry points and locate source populations of invasive species (Lallias et al., 2015), to determine the rates of evolution in invaders (Cormier et al., 2018) and cumulatively provide information on the potency of biosecurity and management efforts (Anderson et al., 2015).

This study provides the first complete mitochondrial genome for a *Dikerogammarus* sp., identifying a total 40 predicted coding regions for ncRNAs and PCGs that can be used to gain greater genetic-level data for understanding demon shrimp invasions, origins and evolution. These data are used to explore



the position of *Dikerogammarus* sp. within the Amphipoda and identifies the UK population to be similar to populations on mainland Europe. The

Fig. 2 Gene synteny comparison between *Dikerogammarus haemobaphes*, *Eulimnogammarus cyaneus* and *Gammarus duebeni*, using non-coding RNA (green) and protein coding regions (blue). Additional partial genes identified by MITOS are presented in orange (black arrow). The red line represents a region on the *D. haemobaphes* genome that has been rearranged. The pink line represents a tRNA duplication

largest mitochondrial phylogeny for the Amphipoda is presented herein and is correlated with known amphipod diseases (Bojko & Ovcharenko, 2019) to explore potential evolutionary origins in addition to host species that may harbour interesting infections.

Tracking biological invasions associated with disease introduction

Biological invasions that introduce disease tend to be understudied, with the majority focussing on the host introduction pathway and host-associated impact in novel environments (Roy et al., 2017). *Dikerogammarus haemobaphes* is a high-risk species for the introduction of disease; therefore it is important to note that the use of molecular resources in combination with disease screening efforts may be able to define the invasion pathway of the host and its parasites. This mitochondrial genome scaffold has already indicated that the UK population is tentatively related to non-native populations of demon shrimp collected from the Rhine and Main rivers in Germany (Grabner et al., 2015) and Vistula, Poland. In these locations, disease has also been observed from the lethal microsporidian parasite, *C. ornata* (aka: *Microsporidium* sp. G) (Bojko et al., 2015; Grabner et al., 2015; Bojko et al., 2017) and in the related *D. villosus* (Bacela-Spychalska et al., 2012).

There has been high success in the tracking of populations of invasive amphipods through Europe to their native range(s) using population genetics (Rewicz et al., 2015, 2017). Increased availability of molecular tools may allow a phylogeographic understanding for the origins of *D. haemobaphes* and potentially its parasites. This capability also extends to future invasions, such as the impending threat of invasion to the Great Lakes (USA), whereby multiple Ponto-Caspian species (e.g. *Dreissina polymorpha* Boettger, 1913) have already successfully invaded (Ricciardi & MacIsaac, 2000). Whether these species

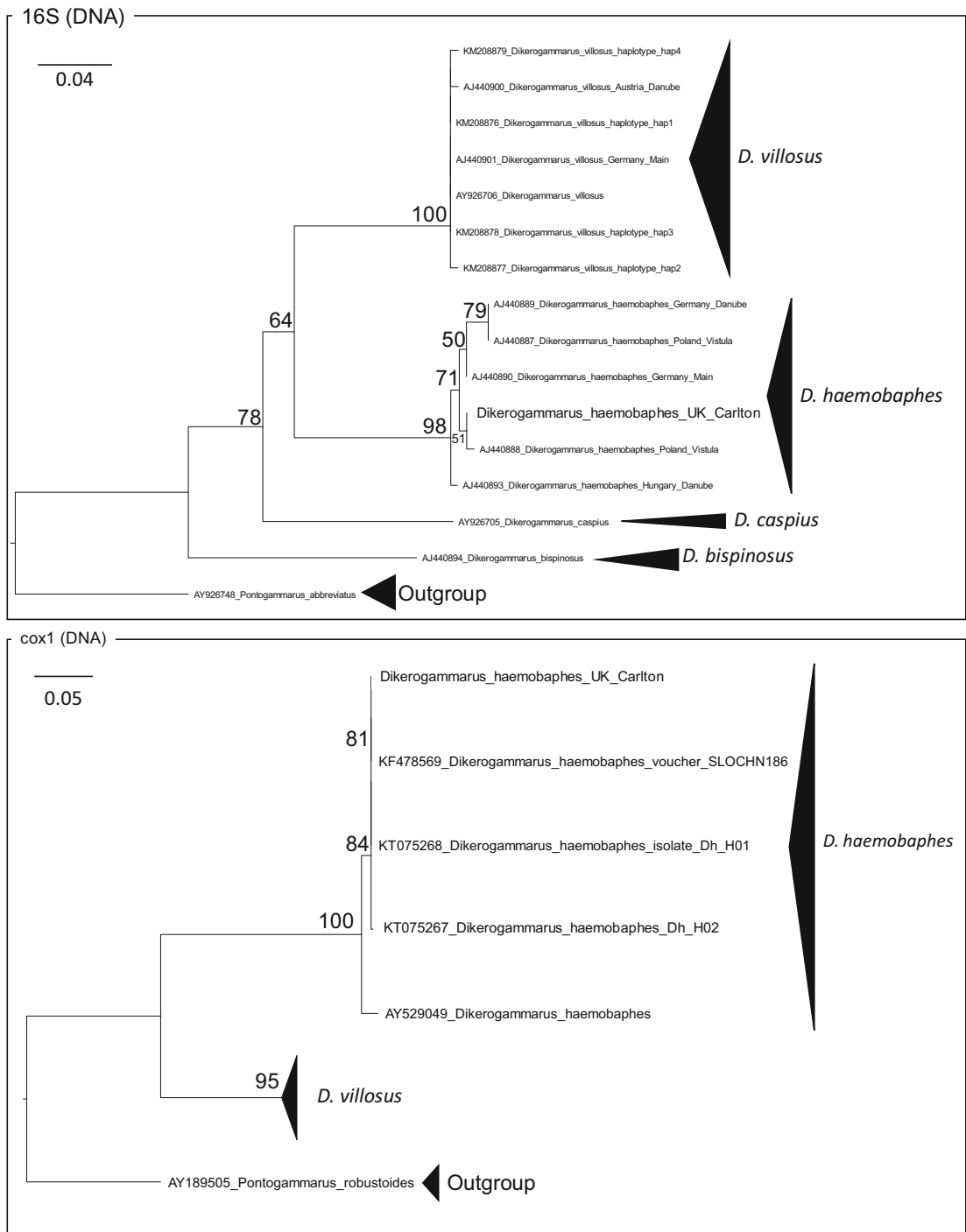


Fig. 3 Phylogenetic trees for the 16S and cox1 genes of *Dikerogammarus* sp., and outgroups. The 16S tree results in the presence of several clades of *Dikerogammarus* sp., including a ‘*D. haemobaphes*’ clade, a ‘*D. bispinosus*’ clade, a ‘*D. caspius*’

clade and a ‘*D. villosus*’ clade. The trees were calculated using MAFFT aligned and trimmed nucleotide sequence data with maximum likelihood

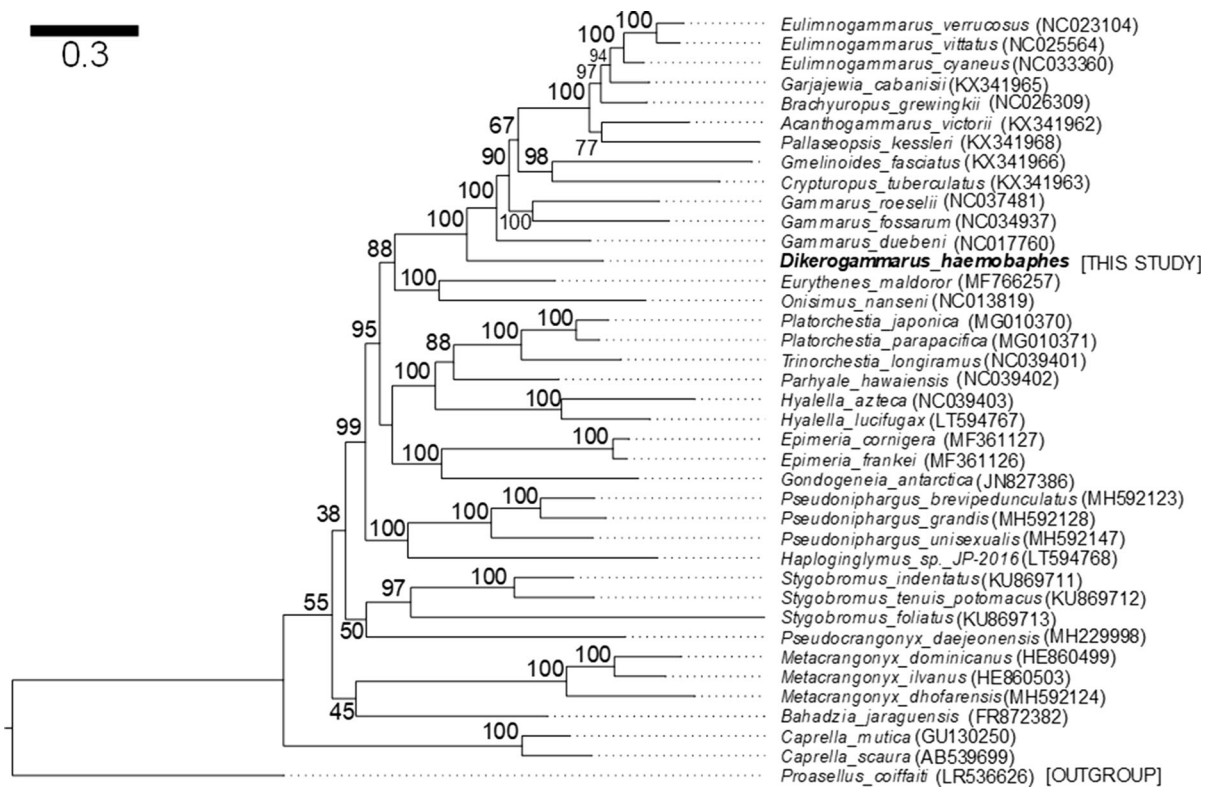


Fig. 4 A concatenated phylogenetic tree including all available amphipod mitochondrial genomes and representative protein sequences. The bootstrap values are indicated at the nodes and

the isopod, *Proasellus coiffaiti* is used to root the tree. The data obtained from the mitochondrial genome of *Dikerogammarus haemobaphes* are present in bold on the tree

have introduced disease to the Great Lakes remains unknown.

The mitochondrial data provided herein for *D. haemobaphes* represent only a single specimen, but based on the 16S data, it constituted a unique haplotype. The 16S gene showed closest similarity to *D. haemobaphes* (AJ440888) from Poland (97%). The *cox1* gene of the UK individual is 100% identical over a 658 bp region to *D. haemobaphes* ‘haplotype 1’ from Germany (North Rhine-Westphalia). In conclusion, it appears that the *D. haemobaphes* in the UK (Carlton Brook) likely arrived from invasive populations in central Europe and not from the native range.

Evolutionary history of *Dikerogammarus haemobaphes*

The genus *Dikerogammarus* (Gammaridea) contains freshwater and brackish amphipods and was first described by Stebbing (1899). The genus contains nine species to date: *D. aralychensis*, *D. bispinosus*, *D.*

caspius, *D. fluvitalis*, *D. gruberi*, *D. istanbulensis*, *D. oskari*, *D. villosus* and *D. haemobaphes* (Özbek & Özkan, 2011). These species are naturally distributed around the Ponto-Caspian region (Black Sea, Caspian Sea and Sea of Azov) and several have become invasive throughout Europe and on the island of the UK. *Dikerogammarus villosus* and *D. haemobaphes* have both invaded the UK and continue to impact freshwater systems, both directly and through the introduction of pathogens (Bojko et al., 2013, 2018b).

The mitochondrial genome of *D. haemobaphes* shares synteny and gene similarity with closely related amphipods, excluding the presence of some duplicate gene regions and a tRNA rearrangement (Figs. 1, 2). Specifically, the presence of a duplicated *atp8* gene (*atp8-1*) on the opposite coding strand (Fig. 1) is absent from other Gammaridae. This gene shows no genetic similarity to other Gammaridae and may be a motif specific to this species, or possibly the *Dikerogammarus* genus pending further research. If

this is the case it could be a clear molecular tag for use in future systematics of the *Dikerogammarus* group.

Phylogenetically and morphologically, *Dikerogammarus* sp. have been identified as members of the Gammaridae and this study supports their inclusion using mitochondrial data from a *D. haemobaphes* representative (Müller et al., 2002). The phylogenetic data in Fig. 4 suggest that *D. haemobaphes* is likely an early member of the Gammaridae [sensu Hou and Sket (2016)], branching with strong support before the other members. *Eurythenes maldoror* d'Udekem, d'Acoz, Havermans, 2015 and *Onisimus nansenii* (Sars, 1900) both branch at the node separating the Gammaridae and Eurythenidae/Uristidae and greater numbers of sequenced amphipods in both the *Dikerogammarus* and other related genera would greatly increase the evolutionary detail of the early formation of the Gammaridae.

Evolution of microparasites in the Gammaridae

Dikerogammarus sp. are thought to be high invasion risks to the UK, including their co-invasive diseases [summarised in Bojko et al., (2018b)], which are of importance to freshwater ecosystem health (Roy et al., 2017, 2019). The data presented herein have identified that the *D. haemobaphes* population seeding the UK invasion was likely originating in central Europe and not the native range, which means that the diseases identified from extensive screening in the UK are likely also present in the European range (Bojko et al., 2015, 2018b). Many of these diseases can impact the activity of their host, but also cause mortality in non-target amphipod hosts, such as *G. pulex*, potentially threatening biodiversity (Bojko et al., 2018b).

Disease screening in the Gammaridae is lacking in research effort; however, some studies have provided insight the parasite diversity of some that also have mitochondrial genomes available (Madyarova et al., 2015; Bojko, 2017; Bojko et al., 2017; Dimova et al., 2018; Bacela-Spychalska et al., 2018; Ironside & Wilkinson, 2018; Bojko & Ovcharenko, 2019). Combining data from any disease screening efforts with the phylogenetics conducted herein indicated that Microsporidia [*Cucumispora* sp. (including *candidatus* species) and *Dictyocoela* sp.] are present across the Gammaridae, with the exception of *P. kessleri* and *C. tuberculatus*, likely due to a lack of screening effort. Gregarine parasites (Apicomplexa) are present in *Eulimnogammarus* sp., *Gammarus* sp. and *D.*

haemobaphes, suggesting that this group is also likely present across the Gammaridae. The recent discovery of '*Candidatus* Aquirickettsiella gammari' Bojko, Dunn, Stebbing, van Aerle, Bacela-Spychalska, Bean, Urrutia, Stentiford 2018b and similar pathologies in *D. haemobaphes* (Bojko, 2017) and *G. roeselii* (Bojko et al., 2017) suggest that increased screening will also discover related pathogens across the Gammaridae. Finally, bacilliform viruses in the hepatopancreas of crustaceans (now thought to be part of the *Nudiviridae*) have been found in multiple Gammaridae, including: *Dikerogammarus* sp. (Bojko et al., 2013; 2018b) and *Gammarus* sp. (Bojko et al., 2017). The presence of this virus in this early branching gammarid host suggests the viral group are also likely present in the other Gammaridae (Fig. 5).

Concluding remarks

Dikerogammarus haemobaphes is the earliest member of the Gammaridae identified to date using concatenated mitochondrial phylogenetics. Knowledge of its diseases suggests that many of the other Gammaridae likely also co-evolved with microsporidian, protistan, bacterial and viral diseases; many yet to be discovered.

The mitochondrial genome of this host will provide further insight into the development of genetic identification tools and the ability to track this species and its diseases, perhaps in combination with eDNA tools to explore invasion presence (Mauvisseau et al., 2019). Knowledge of the mitochondrial genome will help to differentiate host haplotypes to explore disease susceptibility and identify regions of similarity and difference between *Dikerogammarus* populations.

Finally, this study has determined that the population in the UK seems to have been seeded by populations in Europe and not the native range, suggesting that the diseases in the UK are likely to be present in continental Europe and may pose risk to native Gammarids and the related freshwater ecology.

Acknowledgements I would like to thank Dr Donald Behringer for his contribution to the collection of HiSeq data for this study and for providing me some time to construct this manuscript. Additionally, I would like to acknowledge NERC funding (#:1368300) to JB, Dr Grant Stentiford, Dr Alison Dunn and Dr Paul Stebbing, which contributed to the initial collection of Illumina MiSeq data.

Data availability The mitochondrial genome is submitted under accession number: MK644228, in the NCBI database.

References

- Abele, D., E. Philipp, P. Gonzalez & S. Puntarulo, 2007. Marine invertebrate mitochondria and oxidative stress. *Frontiers in Bioscience* 12: 933–946.
- Anderson, L. G., A. M. Dunn, P. J. Rosewarne & P. D. Stebbing, 2015. Invaders in hot water: a simple decontamination method to prevent the accidental spread of aquatic invasive non-native species. *Biological Invasions* 17: 2287–2297.
- Bacela-Spychalska, K., R. A. Wattier, C. Genton & T. Rigaud, 2012. Microsporidian disease of the invasive amphipod *Dikerogammarus villosus* and the potential for its transfer to local invertebrate fauna. *Biological Invasions* 14: 1831–1842.
- Bacela-Spychalska, K., et al., 2018. Europe-wide reassessment of Dictyocoela (Microsporidia) infecting native and invasive amphipods (Crustacea): molecular versus ultrastructural traits. *Scientific Reports* 8: 8945.
- Bankevich, A., et al., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.
- Bernt, M., et al., 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* 69: 313–319.
- Bojko, J. 2017. Parasites of invasive Crustacea: risks and opportunities for control (Doctoral dissertation, University of Leeds).
- Bojko, J. & M. Ovcharenko, 2019. Pathogens and other symbionts of the Amphipoda: taxonomic diversity and pathological significance. *Diseases of Aquatic Organisms [AMPH]*. <https://doi.org/10.3354/dao03321>.
- Bojko, J., et al., 2013. Baseline histopathological survey of a recently invading island population of ‘killer shrimp’, *Dikerogammarus villosus*. *Diseases of Aquatic Organisms* 106: 241–253.
- Bojko, J., et al., 2015. *Cucumispora ornata* n. sp. (Fungi: Microsporidia) infecting invasive ‘demon shrimp’ (*Dikerogammarus haemobaphes*) in the United Kingdom. *Journal of Invertebrate Pathology* 128: 22–30.
- Bojko, J., et al., 2017. Parasites, pathogens and commensals in the “low-impact” non-native amphipod host *Gammarus roeselii*. *Parasites & Vectors* 10: 193.
- Bojko, J., et al., 2018a. Pathogens of *Dikerogammarus haemobaphes* regulate host activity and survival, but also threaten native amphipod populations in the UK. *Diseases of Aquatic Organisms [AMPH]*. <https://doi.org/10.3354/dao03195>.
- Bojko, J., et al., 2018b. ‘*Candidatus* Aquirickettsiella gammari’ (Gammaproteobacteria: Legionellales: Coxiellaceae): a bacterial pathogen of the freshwater crustacean *Gammarus fossarum* (Malacostraca: Amphipoda). *Journal of Invertebrate Pathology* 156: 41–53.
- Bolger, A. M., M. Lohse & B. Usadel, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.
- Bovy, H. C., et al., 2015. Predicting the predatory impacts of the “demon shrimp” *Dikerogammarus haemobaphes*, on native and previously introduced species. *Biological Invasions* 17(2): 597–607.
- Cormier, A., et al., 2018. The complete mitochondrial genome of *Gammarus roeselii* (Crustacea, Amphipoda): insights into mitogenome plasticity and evolution. *Hydrobiologia* 825: 197–210.
- Dimova, M., et al., 2018. Genetic diversity of Microsporidia in the circulatory system of endemic amphipods from different locations and depths of ancient Lake Baikal. *PeerJ* 6: e5329.
- Grabner, D. S., et al., 2015. Invaders, natives and their enemies: distribution patterns of amphipods and their microsporidian parasites in the Ruhr Metropolis, Germany. *Parasites & Vectors* 8: 419.
- Hou, Z., & B. Sket, 2016. A review of Gammaridae (Crustacea: Amphipoda): the family extent, its evolutionary history, and taxonomic redefinition of genera. *Zoological Journal of the Linnean Society* 176(2): 323–348.
- Ironside, J. E. & T. J. Wilkinson, 2018. Accumulation and exchange of parasites during adaptive radiation in an ancient lake. *International Journal for Parasitology* 48: 297–307.
- Lallias, D., et al., 2015. Invasion genetics of the Pacific oyster *Crassostrea gigas* in the British Isles inferred from microsatellite and mitochondrial markers. *Biological Invasions* 17: 2581–2595.
- Ma, H., et al., 2015. First mitochondrial genome for the red crab (*Charybdis feriata*) with implication of phylogenomics and population genetics. *Scientific Reports* 5: 11524.
- Macher, J. N., F. Leese, A. M. Weigand & A. Rozenberg, 2017. The complete mitochondrial genome of a cryptic amphipod species from the *Gammarus fossarum* complex. *Mitochondrial DNA Part B* 2: 17–18.
- Madyarova, E. V., et al., 2015. Microsporidian parasites found in the hemolymph of four baikalian endemic amphipods. *PLoS ONE* 10: e0130311.
- Mauvisseau, Q., C. Troth, E. Young, A. Burian & M. Sweet, 2019. The development of an eDNA based detection method for the invasive shrimp *Dikerogammarus haemobaphes*. *Management of Biological Invasions*. <https://doi.org/10.3391/mbi.2019.10.3.03>.
- Minh, B. Q., M. A. T. Nguyen & A. von Haeseler, 2013. Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* 30: 1188–1195.
- Müller, J. C., S. Schramm & A. Seitz, 2002. Genetic and morphological differentiation of *Dikerogammarus* invaders and their invasion history in Central Europe. *Freshwater Biology* 47: 2039–2048.
- Nguyen, L.-T., H. A. Schmidt, A. von Haeseler & B. Q. Minh, 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.
- Özbek, M. & N. Özkan, 2011. *Dikerogammarus istanbulensis* sp. n., A New Amphipod Species from Turkey (Crustacea: Amphipoda) with a Key for the Genus. *Zootaxa* 2813: 55–64.
- Rewicz, T., R. Wattier, M. Grabowski, T. Rigaud & K. Bacela-Spychalska, 2015. Out of the Black Sea: phylogeography

- of the invasive killer shrimp *Dikerogammarus villosus* across Europe. PLoS ONE 10: e0118121.
- Rewicz, T., R. Wattier, T. Rigaud, M. Grabowski, T. Mamos & K. Bączela-Spychalska, 2017. The killer shrimp, *Dikerogammarus villosus*, invading European Alpine Lakes: a single main source but independent founder events with an overall loss of genetic diversity. Freshwater Biology 62(6): 1036–1051.
- Ricciardi, A. & H. J. MacIsaac, 2000. Recent mass invasion of the North American Great Lakes by Ponto-Caspian species. Trends in Ecology & Evolution 15: 62–65.
- Romanova, E. V., et al., 2016. Evolution of mitochondrial genomes in Baikalian amphipods. BMC Genomics 17: 1016.
- Roy, H. E., et al., 2017. Alien pathogens on the horizon: Opportunities for predicting their threat to wildlife. Conservation Letters 10: 477–484.
- Roy, H. E., et al., 2019. Developing a list of invasive alien species likely to threaten biodiversity and ecosystems in the European Union. Global Change Biology 25: 1032–1048.
- Stebbing, T. R. R., 1899. Amphipoda from the Copenhagen Museum and other sources, part II. The Transactions of the Linnean Society of London: Zoology (ser 2) 8: 395–432.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.