REABIC

Aquatic Invasions (2017) Volume 12 Article in press

© 2017 The Author(s). Journal compilation © 2017 REABIC

Open Access

Research Article

CORRECTED PROOF

Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples—first record of *Gammarus fossarum* in the UK

Rosetta C. Blackman[1,*], Drew Constable[2], Christoph Hahn[1,3], Andrew M. Sheard[4], Jessica Durkota[5], Bernd Hänfling[1] and Lori Lawson Handley[1]

[1]*Evolutionary and Environmental Genomics Group, School of Environmental Sciences, University of Hull, Cottingham Road, Hull, HU6 7RX, UK*

[2]*Environment Agency, Bromholme Lane, Brampton, Huntingdon, Cambs, PE28 4NE, UK*

[3]*Institute of Zoology, University of Graz, Universitaetsplatz 2, A-8010 Graz, Austria*

[4]*Independent author, Beverley, UK*

[5]*Department of Geography, University College London, Gower Street, London, WC1E 6BT, UK*

Author e-mails: rosiecblackman@gmail.com *(RCB)*, drew.constable@environment-agency.gov.uk *(DC)*, C.Hahn@hull.ac.uk *(CH)*, andysheard@me.com *(AMS)*, j.durkota@ucl.ac.uk *(JD)*, B.Haenfling@hull.ac.uk *(BH)*, L.Lawson-Handley@hull.ac.uk *(LLH)*

**Corresponding author*

Received: 23 September 2016 / Accepted: 20 March 2017 / Published online: 24 April 2017

Handling editor: Michal Grabowski

Abstract

We report the discovery of a non-native gammarid, *Gammarus fossarum* (Koch, 1836) (Crustacea, Amphipoda), in UK rivers. *Gammarus fossarum* is a common freshwater gammarid in many parts of mainland Europe, but was previously considered absent from the UK. *Gammarus fossarum* was detected in a number of UK rivers following DNA metabarcoding of a mini-barcode region of the *COI* gene in macroinvertebrate kick samples, and environmental DNA (eDNA) from water and sediment samples. Subsequent morphological analysis and standard DNA barcoding showed that the species could be reliably identified and separated from *Gammarus pulex* (Linnaeus, 1758), the most dominant and widespread native freshwater gammarid in the UK. Our data demonstrate extensive geographical coverage of *G. fossarum* in the UK, spanning distant river catchments. At present there is no data to confirm the likely origin of *G. fossarum*'s introduction. Subsequent re-examination of historic archive material shows the species to have been present in the UK since at least 1964. This study is among the first to demonstrate the potential of eDNA metabarcoding for detection of new non-native species.

Key words: environmental DNA, metabarcoding, passive detection, early warning, cryptic species, Gammaridae, non-native

Introduction

Amphipods are successful invaders in freshwater ecosystems, with many invasive non-native species (INNS) having been observed to adversely impact indigenous species within Europe over the last century (Bij de Vaate et al. 2002; Grabowski et al. 2007). The introduction of non-native amphipods may not only lead to displacement of native congeners (e.g. Dick and Platvoet 2000; MacNeil and Platvoet 2005; Kinzler et al. 2009), but may also impact on ecosystem structure and functioning (MacNeil et al. 2011; Piscart et al. 2011; Constable and Birkby 2016) and introduce novel pathogens to newly colonised areas (Bacela-Spychalska et al. 2012).

Once non-native species are widely established, efforts to reduce their impacts are often problematic, hence management strategies are strongly focused on preventing introductions or spread (e.g. the "check, clean, dry" campaign in the UK). Early detection is key to such strategies, either to improve the success of eradication programs or to prevent further establishment and dispersal (Roy et al. 2014; Dejean et al.

2012). For freshwater macroinvertebrates, INNS detection methods typically rely on sampling programmes and morphological identification. However, the standard UK monitoring method for macroinvertebrates, a three minute kick sample, will typically recover 62% of families and 50% of species at a site (Furse et al. 1981). This can present considerable challenges when dealing with rare or elusive species. Morphological identification can also prove difficult when identifying taxonomically similar or cryptic species, or juvenile life stages, and is highly dependent on the taxonomical expertise of the investigator. Emerging molecular detection methods may provide significant benefit for detecting non-native species in aquatic environments (Darling and Mahon 2011; Lawson Handley 2015).

One new and rapidly developing method is the use of environmental DNA (eDNA) (Taberlet et al. 2012a, b; Rees et al. 2014; Lawson Handley 2015), which refers to cellular or extracellular DNA that can be extracted directly from environmental samples without prior separation of taxa (Taberlet et al. 2012a). Environmental DNA has been successfully used in numerous studies to detect specific taxa using a targeted approach based on standard or quantitative PCR (Dejean et al. 2012; Dougherty et al. 2016). In an alternative approach, called "metabarcoding", entire species assemblages are analysed by PCR with broadly conserved primers, followed by Next Generation Sequencing (NGS: see Lawson Handley 2015; Hänfling et al. 2016; Port et al. 2016; Valentini et al. 2016 for further detail). Environmental DNA metabarcoding has been successfully used in a small number of studies, for example, to describe entire communities of vertebrates (e.g. Lawson Handley 2015; Hänfling et al. 2016; Port et al. 2016; Valentini et al. 2016) and invertebrates (Deiner et al. 2016) from marine, lake and river samples. Metabarcoding has excellent potential as an early warning tool for detection of non-native species from samples collected from invasion pathways or natural/semi-natural habitats (Mahon and Jerde 2016; Lawson Handley 2015). For example, the technique was recently used as an early detection method for screening ship ballast, and detected non-indigenous zooplankton in Canadian ports (Brown et al. 2016). Environmental DNA metabarcoding has also identified non-native fish species present in samples from the live bait trade (white perch, *Morone americana* (Gmelin, 1789) Mahon et al. 2014) and in river samples (northern snakehead, *Channa argus* (Cantor, 1842) Simmons et al. 2015). However the number of applications of metabarcoding for detection of non-native species has so far been limited.

In this paper we describe the detection of *Gammarus fossarum* (Koch, 1836), a newly recognised freshwater amphipod to the UK, using macroinvertebrate community and eDNA metabarcoding. The species was found in several UK rivers following a preliminary non-targeted sampling programme for macroinvertebrate communities based on metabarcoding of a 313 bp mini-barcode region of the cytochrome c oxidase subunit I (*COI*) gene, and was subsequently confirmed using a combination of morphological analysis and standard full-length *COI* DNA barcoding (via Sanger sequencing). This study demonstrates the power of eDNA metabarcoding for detection of non-native species in natural habitats.

Methods

Metabarcoding surveys

Sampling

Field surveys were carried out in March 2015 within 8 UK river catchments (Figure 1, Maps A–H, excluding E). At each site (n = 65) environmental variables including water depth, width, substrate type and surrounding habitat were recorded. Three sample types were collected at each site: a three minute macroinvertebrate kick sample (Murray-Bligh 1999) for identification by microscopy analysis and high molecular weight DNA extraction from pools of individuals; and water and sediment samples were collected for eDNA extraction. Two litres of water was sampled from the surface by collecting 4 × 500 ml from points across the river width using a sterile bottle. Sediment samples were collected from points across the river width using a trowel, and the material was placed in a 42 fluid oz. sterile Whirl-pak® bag (Cole-Palmer, Hanwell, London). All sampling equipment was sterilized in 10% commercial bleach solution for 10 minutes then rinsed with 10% MicroSol detergent (Anachem, UK) and purified water between samples. Sample bottles filled with ddH_2O were taken into the field and later filtered as sample blanks.

Macroinvertebrate community sample processing

All macroinvertebrates from each kick sample were sorted and identified to the lowest taxonomic level possible, before being stored in sterile 50 ml falcon tubes filled with 100% ethanol. For DNA extraction, samples were dried to remove the ethanol and the entire macroinvertebrate community was lysed in a Qiagen Tissue Lyser® with Digisol (50mM Tris, 20M EDTA, 120 mM NaCl and 1% SDS) (3 × 30 sec). Samples were then incubated overnight at 55 °C with SDS and Proteinase K. DNA from a 200 µl subsample

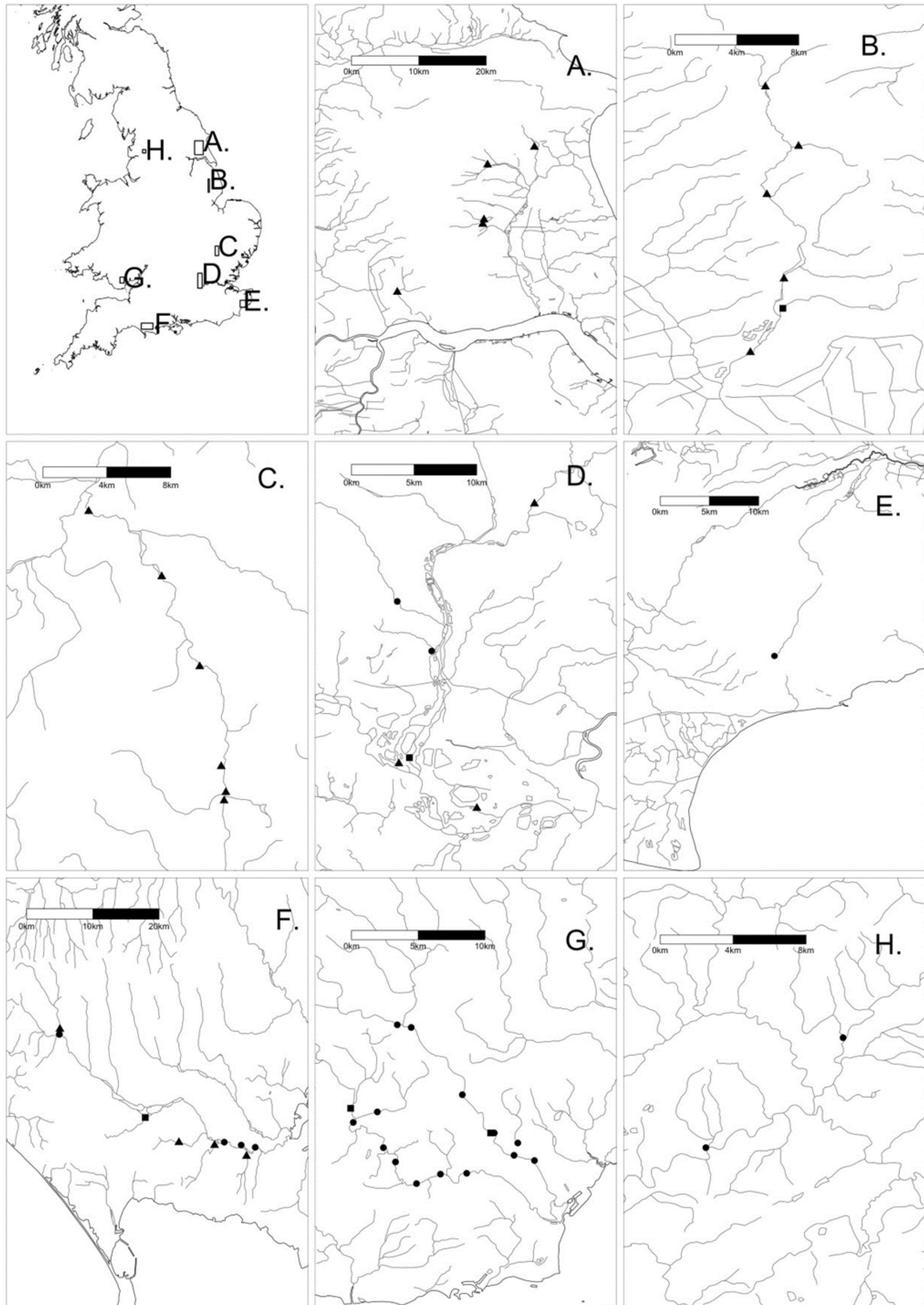

Figure 1. Distribution of Gammaridae species detected during this study. ● – *Gammarus fossarum*, ▲ – *Gammarus pulex* and ■ – both species present. A – River Hull, B – River Bain, C – River Cam, D – River Colne, E – Nailbourne, F – River Frome, G – Rivers Taff and Ely and H – River Ribble. See supplementary information Table S1 for further site information (Pebesma et al. 2005; Wickham. 2009; Bivand et al. 2013; Bivand et al. 2016; Gallic. 2016) Contains OS data © Crown copyright and database right (2016).

of the lysed tissue was extracted using the DNeasy Blood & Tissue Kit® (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

Environmental DNA sample processing

Water samples were filtered within 24 hours through sterile 47 mm diameter 0.45 µm cellulose nitrate membrane filters and pads (Whatman, GE Healthcare, UK), using Nalgene filtration units attached to a vacuum pump. Sediment samples were stored at −20 °C within 12 hours of sampling. The sample was defrosted, mixed and 200 ml of sediment placed in a sterile measuring cylinder with 500 ml of molecular grade water, then inverted 10 times and left to stand for 30 s, the supernatant was then poured off into a sterile container. This procedure was repeated twice. Two hundred and fifty millilitres of the supernatant was then prefiltered through sterile 20 µm filter paper (Whatman, GE Healthcare, UK), and the filtrate subsequently filtered through 0.45 µm cellulose nitrate filters, as for the water samples. Filter papers were stored in sterile petri dishes at −20 °C until extraction. Filtration blanks (2 L purified water) were run before the samples for each filtration run to test for contamination at the filtration stage (n = 5). Filtration equipment was sterilized in 10% commercial bleach solution for 10 minutes then rinsed with 10% MicroSol detergent and purified water after each filtration.

Environmental DNA from both water and sediment samples was extracted using PowerWater® DNA Isolation Kit (MoBio Laboratories, Inc. Carlsbad, USA) following the manufacturer's instructions.

PCR, library prep and sequencing

We chose to use *COI* for metabarcoding because this region has the broadest taxonomic coverage for macroinvertebrates in public sequence databases and is the most widely used DNA barcode for taxonomic discrimination in this group. A 313 bp fragment ("mini-barcode") was targeted using the primers described in Leray et al. (2013). For library preparation we used a nested tagging protocol, modified from the Illumina 16S two-step metabarcoding protocol (Illumina 2011) as outlined in Kitson et al. (2015).

In the first step, PCRs were performed with modified versions of the primers jgHCO2198 TAIA CYTCIGGRTGICCRAARAAYCA and mICOIintF GGWACWGGWTGAACWGTWTAYCCYCC (Leray et al. 2013). In addition to the standard primer sequence, primers included one of eight unique forward or 12 unique reverse 8-nucleotide Molecular Identification Tags (MID), plus a bridge site, which acts as a binding site for PCR 2 (see Kitson et al. 2015 for full details). PCRs were carried out in 25 µl volumes with MyFi High-Fidelity Taq (Bioline, UK) containing: 10 µM of each primer, and 2 µl of undiluted DNA template. PCRs were performed on an Applied Biosystems Veriti Thermal Cycler with the following profile: initial denaturation at 95 °C for 1 min, followed by 45 cycles of denaturation at 98 °C for 15 s, annealing at 51 °C for 15 s and extension at 72 °C for 30 s, with a final extension time of 10 min at 72 °C. This included PCR and filtering blanks (n = 3 and n = 5, respectively) and single species positives: *Triops cancriformis* (Bosc, 1801) (n = 2) and *Harmonia axyridis* (Pallas, 1773) (n = 2). PCR products were confirmed by gel electrophoresis on a 2% agarose gel stained with ethidium bromide. PCRs were carried out three times and then pooled. Pooled PCR products were then purified using the E.Z.N.A Cycle Pure Kit® (VWR International, Leicestershire).

In the second PCR step, Illumina adapters and additional forward and reverse MID tags were added in a second PCR with 10 µM of each tagging primer and 2 µl of purified PCR product. PCR settings were: initial denaturation at 95 °C for 3 min, followed by 12 cycles of denaturation at 98 °C for 20 s, annealing at 72 °C for 1 min and extension at 72 °C for 5 mins, with a final extension time of 10 mins at 4 °C (Kitson et al. 2015).

Samples were then classified into five categories based on the strength of band produced on ethidium bromide-stained agarose gels. Negative controls (including filtration blanks) produced no bands on the agarose gel so were categorised with samples with the lowest band strengths when being added to the library. All positive control (i.e. extracted tissue) samples were categorised as high band strength. Volumes of the samples were then pooled according 5 band strength categories: 10 µl for the lowest band strength, then decreasing volumes of 8 µl, 6 µl, 4 µl, and 2 µl for increasing band strength. The library was then pooled and cleaned using AMPure XP beads following the recommended manufacturer's protocol (Agencourt AMPure XP, Beckman Coulter Inc. US). The library was run at a 12 pM concentration on an Illumina MiSeq, at the in-house facility at the University of Hull, using the 2 × 300 bp V3 chemistry.

Specimen confirmation – microscopy and standard DNA barcode sequencing:

Verification of the results from DNA metabarcoding was carried out using a combination of morphological identification and standard DNA barcoding (by Sanger sequencing).

Gammarus fossarum is a well-studied diverse species complex, which has three well established cryptic species

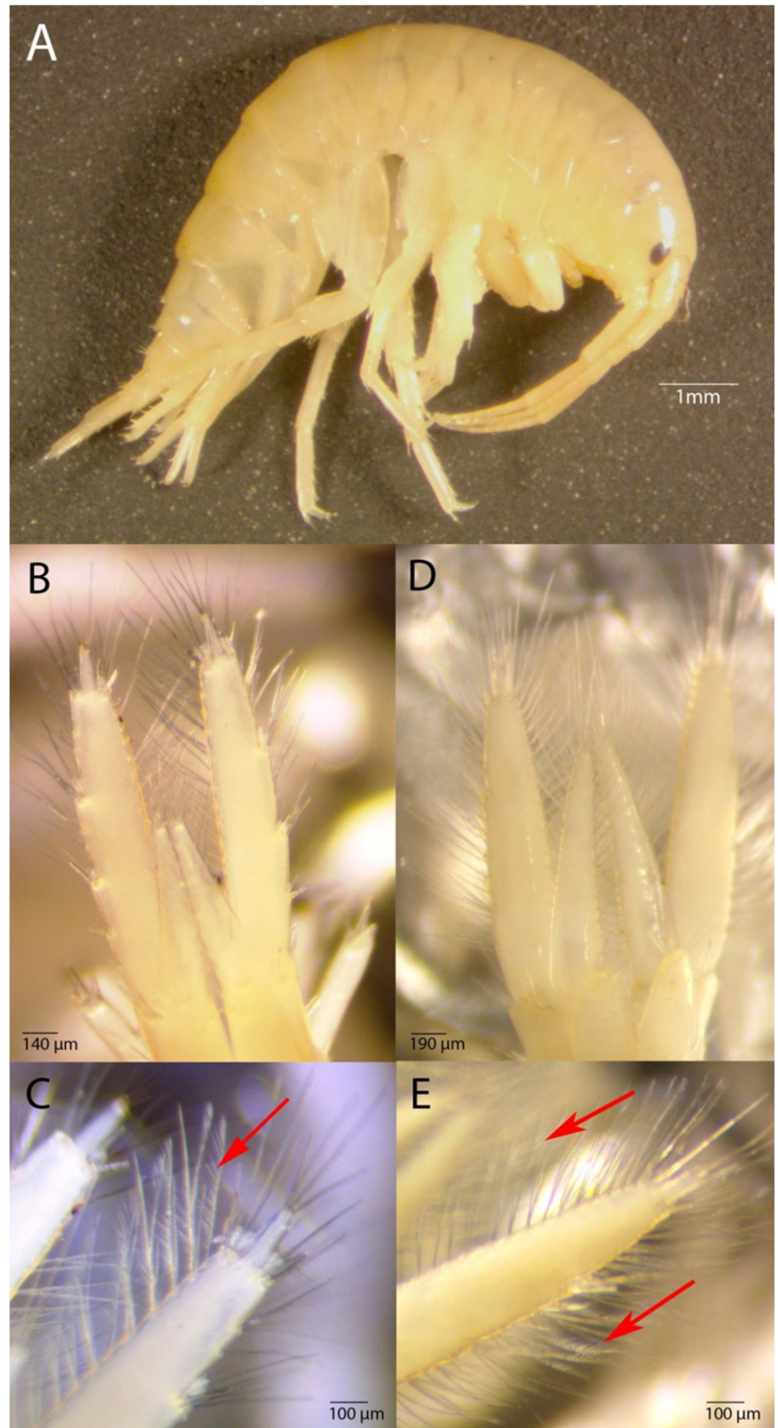

Figure 2. Picture of *Gammarus fossarum* found in the River Taff, UK, 7/6/2016, A) male adult specimen, B) male uropod III and C) male plumose hairs on inside of exopod of uropod III (↗); and picture of male *Gammarus pulex* features for comparison D) uropod III and E) plumose hairs on inner and outer edge of exopod of uropod III (↗) (Photographs by D. Constable).

(types A, B and C) with a further 36–53 different cryptic lineages being identified through phylogenetic studies (Weiss et al. 2014; Copilaş-Ciocianu and Petrusek 2015). Species within this complex are known to differ in their ecology both in terms of their environmental requirements and geographic distributions (Copilaş-Ciocianu and Petrusek 2015; Eisenring et al. 2016). The *G. fossarum* complex belongs to the *G. pulex*-group, which means it has small oval or kidney shaped eyes (less than twice as long as wide)

and the pereopods 5–7 are armed with spines and few setae (Pinkster 1972). Within the UK, these features alone would help to separate it from *G. duebeni*, *G. tigrinus* and *G. zaddachi*. It can be distinguished from all five known UK freshwater *Gammarus* residents by examining uropod III. In *G. fossarum* the ratio length of the endopod versus the exopod is about 0.5, whilst in the other five it is >0.5, typically 0.75 (see Figure 2B and 2D respectively). Another feature of *G. fossarum* is that only the inside margin of the exopod has plumose setae, whilst the other five have plumose setae on both inner and outer margins (see Figure 2C and 2E respectively). The latter feature should however be used with caution, as plumose setae on the outer margin of the exopod can show up in very old males of *G. fossarum* (Meijering 1972).

A *post hoc* morphological examination of UK *Gammarus* specimens was carried out to confirm the presence of *G. fossarum.* Since the entire macroinvertebrate samples from the original sampling program had been lysed for metabarcoding, new specimens were collected by hand net from two catchments where *G. fossarum* was detected by metabarcoding in close proximity to previously sampled sites; River Taff, Wales (n = 38) on 7/6/2016 and River Frome, England (n = 39) on 27/6/2016. Additional, archived specimens obtained from the Nailbourne (Little Stour catchment), England (n = 2) on 20/4/2013, were also analysed; (see Table 1 and Figure 1, Maps: E, F and G). Collected individuals were then subject to morphological examination and identified using Karaman and Pinkster (1977), Eggers and Martens (2001) and Piscart and Bollache (2012).

Microscopic identification was carried out on all specimens collected for morphological confirmation. Both *G. fossarum* (n = 37) and *G. pulex* (n = 1) were identified from individuals collected from the River Taff and only *G. fossarum* (n = 39) was found in a sample from the River Frome. Standard DNA barcoding was performed on some of the individuals identified morphologically as *G. fossarum* (n = 3) and *G. pulex* (n = 1) from the River Taff, and *G. fossarum* from the Nailbourne (Little Stour catchment) (n = 2). DNA was extracted using the DNeasy Blood & Tissue Kit® (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The full length *COI* DNA barcoding fragment was amplified (Folmer et al. 1994) using the following protocol: PCRs were performed in 25 µl volumes with MyTaq (Bioline, UK), 10 µM of each primer and 2 µl of DNA template. The PCR profile consisted of: initial denaturation at 95 °C for 1 min, followed by 35 cycles of denaturation at 95 °C for 15 s, annealing at 50 °C for 15 s and extension at 72 °C for 10 s, with a final extension time of 10 min at 72 °C. PCR products were checked on agarose gels and commercially sequenced using HCO2198 (Macrogen Europe, Amsterdam, Netherlands).

Bioinformatics

Processing of Illumina read data and taxonomic assignment were performed using a custom bioinformatics pipeline (metaBEAT, v.0.97.7-global; see Github reference 1) as described previously (Hänfling et al. 2016), with minor modifications. For each sample, raw Illumina sequences were filtered to retain only read pairs containing the expected forward/reverse in-line barcode combination (perfect matches only) using the program *process_shortreads* from the Stacks v1.20 program suite (Catchen et al. 2013) and subsequently quality trimmed using the program Trimmomatic v0.32 (Bolger et al. 2014). Specifically, read quality was assessed across 5 bp sliding windows starting from the 3'-end, and reads were clipped until the per window average read quality reached a minimum of phred 30. Any reads shorter than 100 bp after the quality clipping were discarded. To remove PCR primers and spacer sequences the first 30 bp of the reads was clipped off. Remaining sequence pairs were merged into single high quality reads using the program FLASH v1.2.11 (Magoč and Salzberg 2011). For any read pairs not merged successfully, only the forward read was retained for downstream analyses. Sequences were clustered at 97% identity using vsearch v1.1 (see Github reference 2). Any clusters represented by less than three sequences were excluded from further analyses, as these likely represent sequencing error. Each of the remaining distinct sequence clusters was collapsed to a single representative sequence (aka centroid). Only centroid sequences of the expected length as determined by the primers (313 bp ± 5%) were retained for downstream analyses. To obtain a final set of non-redundant (nr) queries for taxonomic assignment, centroid sequences across all samples were clustered globally at 97% identity using vsearch v1.1. The global set of nr queries was subjected to a BLAST (Zhang et al. 2000) search (blastn) against a custom reference database consisting of gammarid sequences from Weiss et al. (2014) and two *CO1* sequences from *T. cancriformis* (GenBank accession numbers EF189678.1 and JX110644.1) and *H. axyridis* (accession numbers KU188381.1 and KU188380.1), respectively. Taxonomic assignment was performed using a lowest common ancestor (LCA) approach. In brief, after the BLAST search the algorithm identifies the most significant matches

to the reference database (top 10% bit-scores) for each of the query sequences. If only a single taxon is present in this list of matches then the query is assigned directly to this taxon. If more than one taxon is present, the query is assigned to the lowest taxonomic level that is shared by all taxa in the list. Queries yielding best BLAST matches below a bit-score of 80 or with less than 85% identity were binned as "unassigned". To assure full reproducibility of our analyses we have deposited the entire workflow in an additional dedicated Github repository (see Github reference 3). To reduce the possibility of false positives based on our single species positive samples and in order to obtain a conservative estimate of the distribution of *G. fossarum* in the UK, we only report *G. fossarum* as present at a given site if it was supported by at least 1% of the total quality trimmed reads per sample.

Phylogeny

Phylogenetic analysis was performed to further confirm the identity of the putative *Gammarus* sp. sequences obtained as part of the current study. We downloaded a previously published *CO1* dataset (Weiss et al. 2014; Copilaş-Ciocianu and Petrusek 2015) from Genbank, comprising 89 sequences of *G. fossarum*, six *G. pulex* (Linnaeus, 1758) sequences and a single sequence each from four further outgroup species (*G. balcanicus* (Schaferna, 1922), *G. glabratus* (Hou and Li, 2003), *G. roeselii* (Gervais, 1835) and *G. tigrinus* (Sexton, 1939) (Radulovici et al. 2009; Hou et al. 2011; Feckler et al. 2012; Weiss et al. 2014). This set of previously published sequences was extended by the sequences obtained via standard full-length DNA barcoding and mini-barcode meta-barcoding. Prior to phylogenetic analysis we extracted the most abundant sequence, i.e. haplotype, from each sample from the initially obtained 97% sequence clusters assigned to *G. fossarum* and *G. pulex*, respectively. Nucleotide sequences of *G. fossarum* and *G. pulex* used in the phylogenetic analysis were deposited in Genbank (GenBank accession KY464959–KY464977). Phylogenetic analysis was performed in the Reprophylo environment (Szitenberg et al. 2015). In brief, sequences were aligned using the program MAFFT v7.123b (Katoh and Standley 2013) and the alignment was trimmed using the program trimAl v1.2rev59 (Capella-Gutiérrez et al. 2009). Maximum-likelihood tree inference was performed using RAxML v8.0.12 (Stamatakis 2014). The full, detailed analysis is provided as Jupyter notebook in the dedicated Github repository (Github reference 3), which also contains the alignment underlying the phylogenetic tree and further supplementary information.

Comparison of data from eDNA/DNA and microscopy analysis

A correlation was performed to compare the Gammaridae abundance data generated from the kick sample microscopy analysis and the DNA/eDNA metabarcoding. Specifically, the relationship between DNA/eDNA data (read count) and data from micro-scopy analysis (biomass calculated from average Gammaridae specimen weight) was investigated by calculating Pearson's Correlation Coefficient in R v3.1.3 (R Core team 2013). Note that *G. fossarum* and *G. pulex* sequencing data have been combined here as the species were not distinguished during the initial morphological determination.

Results

Metabarcoding survey

The total sequence read count passing quality control, before removal of chimeric sequences, was 4,290,271. We quantified the level of possible contamination using sequence information from single species positive samples, which enabled us to choose a suitable threshold level (1% of total sample reads) for filtering and removal of low level contamination. This conservative threshold is comparable to recent, similar studies (e.g. Hänfling et al. 2016; Port et al. 2016). After applying this threshold, over the 195 samples the total read count was 933,457.

Gammarus fossarum was detected in 28 sites in total: 25 via metabarcoding, 1 site by morphological identification, 1 site by standard DNA barcoding and 1 site by morphological identification and DNA bar-coding (See Table 1 and Supplementary material Table S1). Of the 25 metabarcoding samples, *G. fossarum* was found in: 25 DNA macroinvertebrate samples, 8 water eDNA samples and 9 sediment eDNA samples. *G. pulex* was detected in 27 of the sites in the metabarcoding DNA macroinvertebrate samples only and a single site using Sanger sequencing.

A full breakdown of gammarid sequences per sample and proportion of gammarid biomass per sample are included in Supplementary material Table S1. A further 36 freshwater macroinvertebrate families were detected by metabarcoding: data from these non-gammarid species form part of a wider macroinvertebrate data set which is being analysed separately and will be published elsewhere.

The average read count of the samples with gammarid species present was 3512. At those sites the proportion of *G. fossarum* reads per sample ranged from 1.68 – 100% in the macroinvertebrate DNA,

Table 1. Specimen identification and identification method for morphologically identified and DNA barcoded specimens. (*Specimens collected from the River Frome were subject to morphological identification only. **Specimens collected from Nailbourne were DNA sequenced only due to damaged specimens).

Unique ID	Catchment	Site Name	Coordinates		*G. fossarum*		*G. pulex*	
			Lat	Long	Micro-scopy	DNA sequen-cing	Micro-scopy	DNA sequen-cing
DC003	Taff	Forest Farm Country Park	51.516	−3.242	✓	✓		
DC004	Taff	Forest Farm Country Park	51.516	−3.242			✓	✓
DC005	Taff	Forest Farm Country Park	51.516	−3.242	✓	✓		
DC006	Taff	Forest Farm Country Park	51.516	−3.242	✓	✓		
DC007-045	Frome	East Stoke	50.681	−2.185	✓*			
JD001	Nailbourne	Adj Saint Ethelburga well	51.126	1.087		✓**		
JD002	Nailbourne	Adj Saint Ethelburga well	51.126	1.087		✓**		

1.67 – 55.35% in the water eDNA and 1.59 – 18.05% in sediment eDNA samples (Table S1). Similarly, *G. pulex* reads ranged from 1.65 – 97.41% in the DNA macroinvertebrate samples. There was a significant positive correlation between the percentage of *Gammarus* biomass in the sample, and the percentage of *Gammarus* sequence reads (Pearson's $r = 0.747$, df = 46, $P = 1.098 \times 10^{-9}$, Supplementary material Figure S1). Importantly, *Gammarus* sequences were detected when gammarids constituted as little as 2.6% of the total biomass (Table S1).

Verification of Gammarus fossarum by microscopy

Gammarus fossarum was not identified morphologically in any samples surveyed in March 2015 prior to metabarcoding. Of the 38 gammarid specimens recovered from the River Taff on 7/6/2016, 37 *G. fossarum* morphological identifications were made. Adult males ranged between 8–12 mm (n = 21) and adult females 7–10 mm (n = 15). Four females were ovigerous. The other gammarid specimen encountered was a male *G. pulex* (13 mm). Of the 39 gammarid specimens collected from the River Frome on 27/6/2016, all were identified as *G. fossarum* morphologically. Adult males of this population ranged from 8–11.5 mm (n = 24) and adult females 7–9 mm (n = 15). Again, four ovigerous females were recorded. The relative abundance of size distribution in the two sampled populations can be seen in the Supplementary information (Figure S2). The two individuals collected from the Nailbourne on 20/4/2013 were not verified using microscopy as the specimens were too heavily damaged for morphological identification.

The size ranges encountered for *G. fossarum* fall within the expected range for the species, with Goedmakers (1972), Pinkster (1972), Karaman and Pinkster (1977) and Piscart and Bollache (2012) reporting that the largest males typically reach 14–15 mm.

Verification of Gammarus fossarum by DNA barcoding

Morphological identifications were confirmed by DNA sequencing for specimens collected from the River Taff (n = 4): 3 specimens of *G. fossarum* and a single *G. pulex*. The individuals collected from the Nailbourne (n = 2) were also both identified as *G. fossarum* using subsequent DNA barcoding (see Table 1).

Phylogeny

The phylogeny (Figure 3) is congruent with the findings of the morphological identification. The *G.* cf. *fossarum* and *G.* cf. *pulex* sequences cluster with their respective lineages (identified in Weiss et al. 2014; Copilaş-Ciocianu and Petrusek 2015). *Gammarus fossarum* sequences obtained by both metabarcoding and standard DNA barcoding show little divergence and cluster together in the phylogeny, indicating closely related sequences. The *G. fossarum* sequences obtained in the current study group with high statistical support within Clade 11, as defined using the distance based Automatic Barcode Gap Discovery (ABGD) approach in Weiss et al. (2014). Sequences further group in a subclade with samples from south-western Germany, Southern Black Forest and Eastern Sauerland in Germany, i.e. clade 14, as delineated using the tree-based GMYC in Weiss et al. (2014). Aligning the UK *G. fossarum* specimens within Clade 11 confirms previous studies which show this clade to be the most widely distributed across Europe within the species complex (Copilaş-Ciocianu and Petrusek 2015; Weiss and Leese 2016).

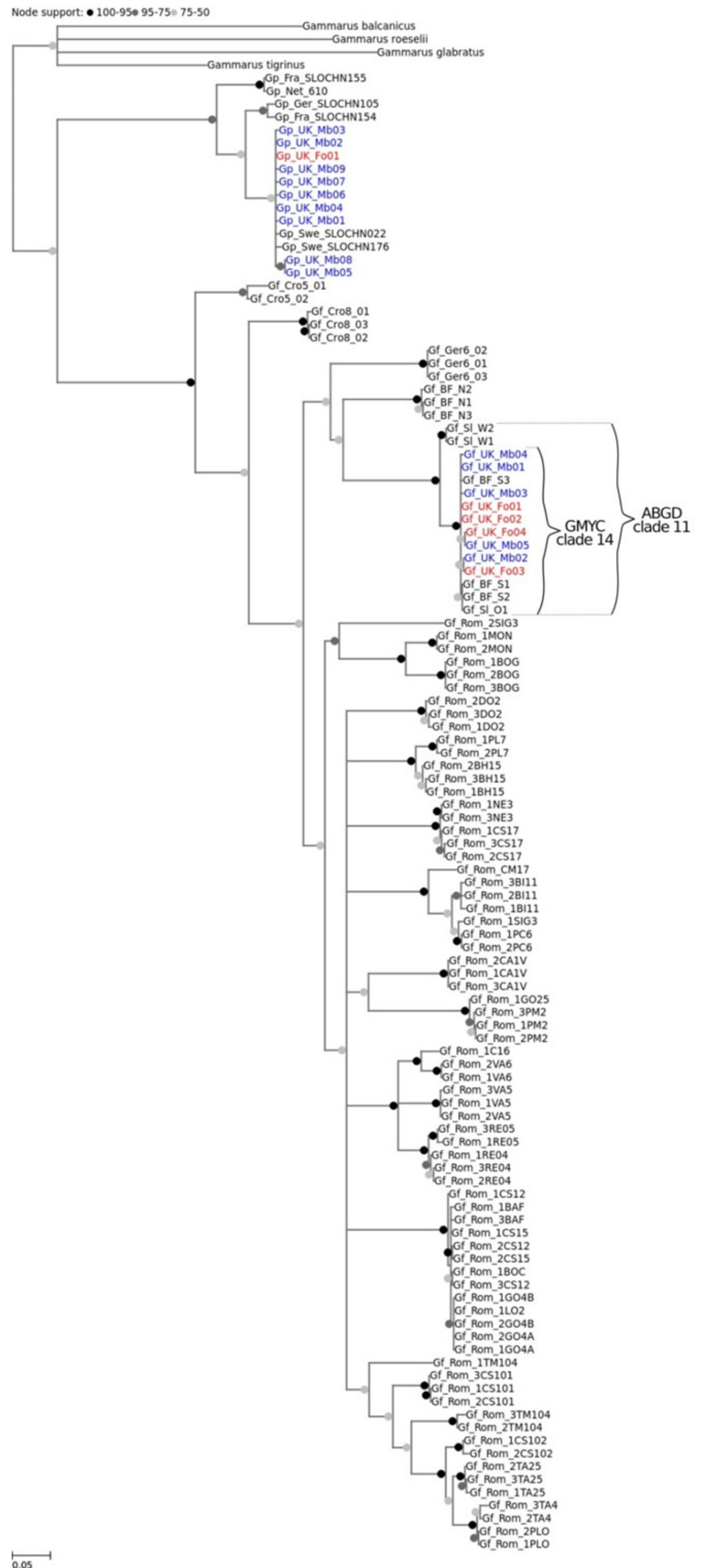

Figure 3. Maximum likelihood phylogenetic tree for the *COI* gene – based on sequences obtained from previously published and newly obtained Gammaridae sequences. The mini-barcode (metabarcoding) and standard *COI* barcode sequences from this study are represented in blue and red, respectively. (See supplementary material Table S2, for accession numbers and origin of individual sequences). GMYC – General Mixed Yule Coalescent, ABGD – Automatic Barcode Gap Discovery (Puillandre et al. 2011) indicate the approaches used by Weiss et al (2014) to detect the different clades in their study.

Discussion

Non-targeted detection by direct and environmental DNA metabarcoding has the potential to revolutionise early warning systems for non-native species, but this utility of the new technology has so far been demonstrated only a limited number of times (Mahon et al. 2014; Brown et al. 2016). In this study, *G. fossarum,* a newly recognised non-native species for the UK, was detected during the course of a wider metabarcoding survey of macroinvertebrate communities. The identification of *G. fossarum* was subsequently confirmed by microscopy and standard DNA barcoding. The sequences generated from this study indicate that the UK populations of *G. fossarum* sampled here fall within the previously identified Clade 11, *sensu* Weiss et al. (2014), of this highly diverse species complex (Figure 3). Importantly this is the most widely distributed clade within the *G. fossarum* complex (Weiss et al. 2014; Copilaş-Ciocianu and Petrusek 2015; Weiss and Leese 2016).

Gammarus fossarum was found in seven distant river catchments within the UK, indicating a widespread distribution (Figure 1). Initial detection of *G. fossarum* was made using non-targeted metabarcoding of macroinvertebrate DNA, water eDNA and sediment eDNA samples. Of the sites where *G. fossarum* was detected using this method (n = 25), *G. fossarum* was detected in all 25 DNA macroinvertebrate samples (100%), in 8 of water (32%) and 9 sediment (36%) samples. The lower detection of *G. fossarum* in eDNA samples compared to macroinvertebrate samples is not surprising due to the dilution of eDNA and effects of flow on DNA availability in lotic systems.

At 23 of the 28 sites (including *post hoc* samples) where *G. fossarum* was present it was the only Gammaridae species detected. This suggests it is not only widespread in the UK but could also be the dominant gammarid in some locations, possibly even having displaced the native *G. pulex* locally. With the new species discovery, recent re-examination of historical archived gammarid samples was undertaken from available Environment Agency and Natural History Museum (NHM), London, collections. Material from the Environment Agency had overlooked records of *G. fossarum* dating back to 2005 from the River Len, Maidstone, Kent (51.2619ºN; 0.56451ºE) whilst re-examination of material from the NHM revealed the earliest record to date, 1964 from the River Darent, Kent. This shows that *G. fossarum* has remained undetected and overlooked by conventional means for a substantial length of time.

Gammarus fossarum is indigenous and widespread in mainland Europe, and typically inhabits springs and upper reaches of mountainous streams, with *G. pulex* being more dominant in lower river sections (Nijssen 1963; Goedmakers 1972; Karaman and Pinkster 1977; Chen et al. 2012). This distribution pattern is linked to *G. fossarum*'s comparative preference for shallower streams and higher current velocities, and its reduced tolerance of low dissolved oxygen conditions (Meijering 1971; Peeters and Gardeniers 1998). It may also be found in middle sections of rivers and is able to coexist with *G. pulex* (Janetzky 1994; Piscart and Bollache 2012; Copilaş-Ciocianu et al. 2014). In such areas of coexistence, *G. fossarum* will often occupy faster flowing areas where vegetation is sparse or absent, and *G. pulex* will be found near marginal shore zones, with reduced currents and rich vegetation growth (Karaman and Pinkster 1977). The distributions of *G. fossarum* in this study covered a range of habitats, mainly lowland rivers (altitude <90 m) with the exception of the Nailbourne spring, adjacent to Saint Ethelburga Well and Maiden Newton on the Upper Frome, with altitudes of 106 m and 109 m, respectively (see Supplementary information). The river depths at *G. fossarum* locations were shallow, seldom reaching more than 20 cm. It is important that further exploration of UK upland systems is undertaken as the sites surveyed for this study were mostly lowland, and at this stage are an indication of habitat suitability rather than preference for *G. fossarum* in the UK. Of our five study sites where *G. fossarum* and *G. pulex* coexisted, all had a mean depth >20 cm and featured both fast and slow currents as well as vegetative marginal areas, however there appears to be no other pattern in the distribution of sites where both species were found to co-exist. Four of the five sites were from the metabarcoding samples, the percentage read count for both species varied substantially, hence no species dominance can be inferred from this data (see Supplementary material Table S1).

Gammarus fossarum is the third non-native freshwater gammarid to be found in the UK within the last six years, following the discoveries of *Dikerogammarus villosus* in 2010 (MacNeil et al. 2010) and *Dikerogammarus haemobaphes* in 2012 (Aldridge 2013). The record is rather unforeseen and the species has not been included on the UK's non-native species watch list with more focus being placed on Ponto-Caspian species that have invaded western Europe (Gallardo and Aldridge 2015). A detailed risk assessment of the threat that *G. fossarum* poses to native Gammaridae within the UK does not currently exist; further research into how *G.*

pulex and *G. fossarum* co-exist within UK habitats should be carried out to decide if this action is warranted. However, the importance of this discovery as a new non-native species to the UK should not be overlooked as it has important implications for future ecological assessments.

In conclusion, we detected a newly recognised non-native species to UK fauna using non-targeted DNA metabarcoding, and confirmed its presence using microscopy and standard DNA barcoding. It is well known that the effectiveness of INNS control or management relies heavily upon early detection (Lodge et al. 2006; Vander Zanden et al. 2010). In future, for other species, non-targeted monitoring of high risk invasion pathways using eDNA may ensure that early eradication or containment are possible management options (Davis 2009; Hulme 2009; Jerde et al. 2011; Thomsen et al. 2012; Lawson Handley 2015). It is important that future research should now focus on establishing the true distribution, ecology and potential implications of *G. fossarum* within the UK, as well as exploring how the non-targeted eDNA metabarcoding approach can be used to detect non-native species.

Acknowledgements

This work was funded by the UK Environment Agency. We are particularly grateful to Dr Kerry Walsh and Dr Alice Hiley for initiating the metabarcoding study and for support throughout. We gratefully acknowledge Amanda Arnold, Frances Attwood, Charlotte Davey, Carole Fitzpatrick, Tim Jones, Mel Lacan, Ellen Moss and Paul Nichols for providing invaluable help during field work, Julia Day, Robert Donnelly, Helen Kimbell, James Kitson and Graham Sellers for contributing to laboratory work and Miranda Lowe and Ben Price from the Natural History Museum, London and Susan Skipp from the Environment Agency for allowing access to archive material. We would like to thank F. Leese and anonymous reviewers for their constructive criticism on the initial submission which greatly helped strengthen the manuscript. The views expressed in this paper are our own, and not necessarily of the institutions we represent.

References

Aldridge DC (2013) GB non-native organism rapid risk assessment for *Dikerogammarus haemobaphes* (Eichwald, 1841). http://www.nonnativespecies.org/downloadDocument.cfm?id=870 (accessed 12 August 2016)

Bacela-Spychalska K, Wattier RA, Genton C, Rigaud T (2012) Microsporidian disease of the invasive amphipod *Dikerogammarus villosus* and the potential for its transfer to local invertebrate fauna. *Biological Invasions* 14: 1831–1842, https://doi.org/10.1007/s10530-012-0193-1

Bij de Vaate A, Jazdzewski K, Ketelaars HAM, Gollasch S, Van der Velde G (2002) Geographical patterns in range extension of Ponto-Caspian macroinvertebrate species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences* 59: 1159–1174, https://doi.org/10.1139/f02-098

Bivand RS, Pebesma EJ, Gomez-Rubio V (2013) Applied spatial data analysis with R, Second edition. Springer, NY. http://www.asdar-book.org

Bivand RS, Keitt T, Rowlingson B (2016) rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 1.1-10. https://CRAN.R-project.org/package=rgdal

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30: 2114–2120, https://doi.org/10.1093/bioinformatics/btu170

Brown EA, Chain FJJ, Zhan A, MacIsaac HJ, Cristescu ME (2016) Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. *Diversity and Distributions* 22: 1045–1059, https://doi.org/10.1111/ddi.12465

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* 25: 1972–1973, https://doi.org/10.1093/bioinformatics/btp348

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: An Analysis Tool Set for Population Genomics. *Molecular Ecology* 22: 3124–3140, https://doi.org/10.1111/mec.12354

Chen W, Bierbach D, Plath M, Streit B, Klaus S (2012) Distribution of amphipod communities in the Middle to Upper Rhine and five of its tributaries. *BioInvasions Records* 1: 263–271, https://doi.org/10.3391/bir.2012.1.4.04

Constable D, Birkby NJ (2016) The impact of the invasive amphipod *Dikerogammarus haemobaphes* on leaf litter processing in UK rivers. *Aquatic Ecology* 50: 273–281, https://doi.org/10.1007/s10452-016-9574-3

Copilaş-Ciocianu D, Grabowski M, Pârvulescu L, Petrusek A (2014) Zoogeography of epigean freshwater Amphipoda (Crustacea) in Romania: fragmented distributions and wide altitudinal variability. *Zootaxa* 3893: 243–260, https://doi.org/10.11646/zootaxa.3893.2.5

Copilaş-Ciocianu D, Petrusek A (2015) The Southwestern Carpathians as an Ancient Centre of Diversity of Freshwater Gammarid Amphipods: Insights from the *Gammarus fossarum* Species Complex. *Molecular Ecology* 24: 3980–3992, https://doi.org/10.1111/mec.13286

Davis MA (2009) Invasion Biology. Oxford University Press, Oxford, UK, 288 pp

Darling JA, Mahon AR (2011) From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Research* 111: 978–988, https://doi.org/10.1016/j.envres.2011.02.001

Deiner K, Fronhofer EA, Mächler E, Walse JC, Altermatt F (2016) Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications* 7: 12544, https://doi.org/10.1038/ncomms12544

Dejean T, Valentini A, Miquel C, Taberlet P, Bellemain E, Miaud C (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. *Journal of Applied Ecology* 49: 953–959, https://doi.org/10.1111/j.1365-2664.2012.02171.x

Dick JTA, Platvoet D (2000) Invading predatory crustacean *Dikerogammarus villosus* eliminates both native and exotic species. *Proceedings of the Royal Society B: Biological Sciences* 267: 977–983, https://doi.org/10.1098/rspb.2000.1099

Dougherty MM, Larson ER, Renshaw MA, Gantz CA, Egan SP, Erickson DM, Lodge DM (2016) Environmental DNA (eDNA) detects the invasive rusty crayfish *Orconectes rusticus* at low abundances. *Journal of Applied Ecology* 53: 722–732, https://doi.org/10.1111/1365-2664.12621

Eggers TO, Martens A (2001) A key to the freshwater Amphipoda (Crustacea) of Germany. *Lauterbornia* 55: 1–68

Eisenring M, Altermatt F, Westram AM, Jokela J (2016) Habitat Requirements and Ecological Niche of Two Cryptic Amphipod Species at Landscape and Local Scales. *Ecosphere* 7: e01319, https://doi.org/10.1002/ecs2.1319

Feckler A, Thielsch A, Schwenk K, Schulz R, Bundschuh M (2012) Differences in the Sensitivity among Cryptic Lineages of the *Gammarus fossarum* Complex. *The Science of the Total Environment* 439: 158–164, https://doi.org/10.1016/j.scitotenv.2012.09.003

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

Furse MT, Wright PD, Armitage I, Moss D (1981) An appraisal of pond net samples for biological monitoring of lotic macro-invertebrates. *Water Research* 15: 679–689, https://doi.org/10.1016/0043-1354(81)90160-3

Gallardo B, Aldridge DC (2015) Is Great Britain heading for a Ponto-Caspian invasional meltdown? *Journal of Applied Ecology* 52: 41–49, https://doi.org/10.1111/1365-2664.12348

Gallic E (2016) LegendMap. https://github.com/3wen/legendMap

Goedmakers A (1972) *Gammarus fossarum* Koch, 1835: Redescription based on neotype material and notes on its local variation (Crustacea, Amphipoda). *Bijdragen Tot de Dierkunde* 42: 124–138

Grabowski M, Jazdzewski K, Konopacka A (2007) Alien crustacea in Polish waters – Amphipoda. *Aquatic Invasions* 2: 25–38, https://doi.org/10.3391/ai.2007.2.1.3

Hänfling B, Handley LL, Read DS, Hahn C, Li J, Nichols P, Blackman RC, Oliver A, Winfield IJ (2016) Environmental DNA Metabarcoding of Lake Fish Communities Reflects Long-Term Data from Established Survey Methods. *Molecular Ecology* 25: 3101–3119, https://doi.org/10.1111/mec.13660

Hou Z, Sket B, Fiser C, Li S (2011) Eocene Habitat Shift from Saline to Freshwater Promoted Tethyan Amphipod Diversification. *Proceedings of the National Academy of Sciences of the United States of America* 108: 14533–14538, https://doi.org/10.1073/pnas.1104636108

Hulme PE (2009) Trade, Transport and Trouble: Managing Invasive Species Pathways in an Era of Globalization. *The Journal of Applied Ecology* 46: 10–18, https://doi.org/10.1111/j.1365-2664.2008.01600.x

Illumina (2011) Preapring 16S Ribosomal RNA Gene Amplicons for Illumina MiSeq system. Illumina technical note

Janetzky W (1994) Distribution of the genus *Gammarus* (Amphipoda: Gammaridae) in the River Hunte and its tributaries (Lower Saxony, northern Germany). *Hydrobiologia* 294: 23–34, https://doi.org/10.1007/BF00017622

Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) "Sight-Unseen" Detection of Rare Aquatic Species Using Environmental DNA. *Conservation Letters* 4: 150–157, https://doi.org/10.1111/j.1755-263X.2010.00158.x

Karaman GS, Pinkster S (1977) Freshwater *Gammarus* species from Europe, North Africa and adjacent regions of Asia (Crustacea, Amphipoda). Part I. *Gammarus pulex*-group and related species. *Bijdragen Tot de Dierkunde* 47: 1–97

Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30: 772–780, https://doi.org/10.1093/molbev/mst010

Kinzler W, Kley A, Mayer G, Waloszek D, Maier G (2009) Mutual predation between and cannibalism within several freshwater gammarids: *Dikerogammarus villosus* versus one native and three invasives. *Aquatic Ecology* 43: 457–464, https://doi.org/10.1007/s10452-008-9206-7

Kitson JJN, Hahn C, Sands RJ, Straw NA, Evans DM, Lunt DH (2015) Nested metabarcode tagging: a robust tool for studying species interactions in ecology and evolution. *bioRxiv* 035071, https://doi.org/10.1101/035071

Lawson Handley L (2015) How will the "molecular revolution" contribute to biological recording? *Biological Journal of the Linnean Society* 115: 750–766, https://doi.org/10.1111/bij.12516

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A New Versatile Primer Set Targeting a Short Fragment of the Mitochondrial COI Region for Metabarcoding Metazoan Diversity: Application for Characterizing Coral Reef Fish Gut Contents. *Frontiers in Zoology* 10: 34, https://doi.org/10.1186/1742-9994-10-34

Lodge DM, Williams S, MacIsaac HJ, Hayes KR, Leung B, Reichard S, Mack RN, Moyle PB, Smith M, Andow DA, Carlton JT, McMichael A (2006) Biological Invasions: Recommendations for U.S. Policy and Management. *Ecological Applications* 16: 2035–2054, https://doi.org/10.1890/1051-0761(2006)016[2035:BIRFUP]2.0.CO;2

MacNeil C, Platvoet D (2005) The predatory impact of the freshwater invader *Dikerogammarus villosus* on native *Gammarus pulex* (Crustacea: Amphipoda); influences of differential microdistribution and food resources. *Journal of Zoology* 267: 31–38, https://doi.org/10.1017/S0952836905007351

MacNeil C, Platvoet D, Dick JTA, Fielding N, Constable A, Hall N, Aldridge D, Renals T, Diamond M (2010) The Ponto-Caspian "killer shrimp", *Dikerogammarus villosus* (Sowinsky, 1894), invades the British Isles. *Aquatic Invasions* 5: 441–445, https://doi.org/10.3391/ai.2010.5.4.15

MacNeil C, Dick JTA, Platvoet D, Briffa M (2011) Direct and indirect effects of species displacements: an invading freshwater amphipod can disrupt leaf-litter processing and shredder efficiency. *Journal of the North American Benthological Society* 30: 38–48, https://doi.org/10.1899/10-056.1

Magoč T, Salzberg SL (2011) FLASH: Fast Length Adjustment of Short Reads to Improve Genome Assemblies. *Bioinformatics* 27: 2957–2963, https://doi.org/10.1093/bioinformatics/btr507

Mahon AR, Nathan LR, Jerde CL (2014) Meta-Genomic Surveillance of Invasive Species in the Bait Trade. *Conservation Genetics Resources* 6: 563–567, https://doi.org/10.1007/s12686-014-0213-9

Mahon AR, Jerde CL (2016) Using Environmental DNA for Invasive Species Surveillance and Monitoring. *Marine Genomics* 1452: 131–142

Meijering MPD (1971) Die *Gammarus*-Fauna der Schlitzerländer Fließgewässer. *Archiv für Hydrobiologie* 68: 575–608

Meijering MPD (1972) Physiologische Beiträge zur Frage der systematischen Stellung von *Gammarus pulex* (L.) und *Gammarus fossarum* Koch (Amphipoda). *Crustaceana* Suppl., 3: 313–325

Murray-Bligh JAD (1999) Procedure for Collecting and Analysing Macroinvertebrate Samples, Quality Management Systems for Environmental Monitoring: Biological Techniques BT001 Version 2.0. Environment Agency, Bristol, UK, pp 40–45

Nijssen H (1963) Some notes on the distribution and the ecology of the amphipod *Gammarus fossarum* Koch, 1835, in the Netherlands (Crustacea, Malacostraca). *Beaufortia* 10: 40–43

Peeters ETHM, Gardeniers JJP (1998) Logistic regression as a tool for defining habitat requirements of two common gammarids. *Freshwater Biology* 39: 605–615, https://doi.org/10.1046/j.1365-2427.1998.00304.x

Pebesma EJ, Bivand RS (2005) Classes and methods for spatial data in R. R News 5(2). http://cran.r-project.org/doc/Rnews/

Pinkster S (1972) On members of the *Gammarus pulex*-group (Crustacea-Amphipoda) from western Europe. *Bijdragen Tot de Dierkunde* 42: 164–191

Piscart C, Bollache L (2012) Crustacés amphipodes de surface (Gammare d'eau douce). Association Française de Limnologie, Thonon les Bains, 113 pp

Piscart C, Mermillod-Blondin F, Maazouzi C, Merigoux S, Marmonier P (2011) Potential impact of invasive amphipods on leaf litter recycling in aquatic ecosystems. *Biological Invasions* 13: 2861–2868, https://doi.org/10.1007/s10530-011-9969-y

Port JA, O'Donnell JL, Romero-Maraccini OC, Leary PR, Litvin SY, Nickols KJ, Yamahara KM, Kelly RP (2016) Assessing Vertebrate Biodiversity in a Kelp Forest Ecosystem Using Environmental DNA. *Molecular Ecology* 25: 527–541, https://doi.org/10.1111/mec.13481

Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology* 21: 1864–1877, https://doi.org/10.1111/j.1365-294X.2011.05239.x

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org

Radulovici AE, Sainte-Marie B, Dufresne F (2009) DNA Barcoding of Marine Crustaceans from the Estuary and Gulf of St Lawrence: A Regional-Scale Approach. *Molecular Ecology Resources* 9: 181–187, https://doi.org/10.1111/j.1755-0998.2009.02643.x

Rees HC, Maddison BC, Middleditch DJ, Patmore JR, Gough KC (2014) The detection of aquatic animal species using environmental DNA—a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* 51: 1450–1459, https://doi.org/10.1111/1365-2664.12306

Roy HE, Peyton J, Aldridge DC, Bantock T, Blackburn TM, Britton R, Clark P, Cook E, Dehnen-Schmutz K, Dines T, Dobson M, Edwards F, Harrower C, Harvey MC, Minchin D, Noble DG, Parrott D, Pocock MJO, Preston CD, Roy S, Salisbury A, Schönrogge K, Sewell J, Shaw RH, Stebbing P, Stewart AJA, Walker KJ (2014) Horizon Scanning for Invasive Alien Species with the Potential to Threaten Biodiversity in Great Britain. *Global Change Biology* 20: 3859–3871, https://doi.org/10.1111/gcb.12603

Simmons M, Tucker AJ, Chadderton WJ, Mahon AR (2015) Active and passive environmental DNA surveillance of aquatic invasive species *Canadian Journal of Fisheries and Aquatic Sciences* 73: 76–83, https://doi.org/10.1139/cjfas-2015-0262

Stamatakis A (2014) RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 30: 1312–1313, https://doi.org/10.1093/bioinformatics/btu033

Szitenberg AMJ, Blaxter ML, Lunt DH (2015) ReproPhylo: An Environment for Reproducible Phylogenomics. *PLoS Computational Biology* 11: e1004447, https://doi.org/10.1371/journal.pcbi.1004447

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012a) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* 21: 2045–2050, https://doi.org/10.1111/j.1365-294X.2012.05470.x

Taberlet P, Coissac E, Hajiabaei M, Rieseberg LH (2012b) Environmental DNA. *Molecular Ecology* 21: 1789–1793, https://doi.org/10.1111/j.1365-294X.2012.05542.x

Thomsen PF, Kielgast J, Iversen LL, Wiuf C, Rasmussen M, Gilbert MTP, Orlando L, Willerslev E (2012) Monitoring Endangered Freshwater Biodiversity Using Environmental DNA. *Molecular Ecology* 21: 2565–2573, https://doi.org/10.1111/j.1365-294X.2011.05418.x

Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Bellemain E, Besnard A, Coissac E, Boyer F, Gaboriaud C, Jean P, Poulet N, Roset N, Copp GH, Geniez P, Pont D, Argillier C, Baudoin JM, Peroux T, Crivelli AJ, Olivier A, Acqueberge M, Le Brun M, Møller PR, Willerslev E, Dejean T (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology* 25: 929–942, https://doi.org/10.1111/mec.13428

Vander Zanden MJ, Hansen GJ, Higgins SN, Kornis MS (2010) A pound of prevention, plus a pound of cure: early detection and eradication of invasive species in the Laurentian Great Lakes. *Journal of Great Lakes Research* 36: 199–205, https://doi.org/10.1016/j.jglr.2009.11.002

Weiss M, Leese F (2016) Widely Distributed and Regionally Isolated! Drivers of Genetic Structure in *Gammarus fossarum* in a Human-Impacted Landscape. *BMC Evolutionary Biology* 16 (July): 153, https://doi.org/10.1186/s12862-016-0723-z

Weiss M, Macher JN, Seefeldt MA, Leese F (2014) Molecular Evidence for Further Overlooked Species within the *Gammarus fossarum*. *Hydrobiologia* 721: 165–184, https://doi.org/10.1007/s10750-013-1658-7

Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Zhang Z, Schwartz S, Wagner L, Miller W (2000) A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 7: 203–214, https://doi.org/10.1089/10665270050081478

Github References:

Github reference 1: https://github.com/HullUni-bioinformatics/metaBEAT; v.0.97.7-global

Github reference 2: https://github.com/torognes/vsearch

Github reference 3: https://github.com/HullUni-bioinformatics/Blackman_et_al_Gfossarum_UK. DOI: https://doi.org/10.5281/zenodo.495075

Supplementary material

The following supplementary material is available for this article:

Table S1. Specimen identification, identification method and site information for metabarcoding samples.

Table S2. Information on specimens from own and published studies that were used in the phylogenetic tree.

This material is available as part of online article from:

http://www.aquaticinvasions.net/2017/Supplements/AI_2017_Blackman_etal_SupplementaryTables.xls

Figure S1. Correlation of the % *Gammarus* biomass in the sample, and the percentage *Gammarus* sequence reads.

Figure S2. Frequency distribution of body length of male and female *Gammarus fossarum* individuals collected from the River Taff and the River Frome.

This material is available as part of online article from:

http://www.aquaticinvasions.net/2017/Supplements/AI_2017_Blackman_etal_SupplementaryFigures.pdf