

Type of the Paper (Article)

Mitochondrial genomes, phylogenetic associations, and SNP recovery for the key invasive Ponto-Caspian amphipods in Europe

Tomasz Mamos^{1,2,*}, Michał Grabowski¹, Tomasz Rewicz^{1,3}, Jamie Bojko⁴, Dominik Strapagiel⁵ and Artur Burzyński⁶

¹ University of Lodz, Faculty of Biology & Environmental Protection, Department of Invertebrate Zoology and Hydrobiology, Łódź, 90-237, Poland; Tomasz.mamos@biol.uni.lodz.pl (T.M.); michal.grabowski@biol.uni.lodz.pl (M.G.); tomasz.rewicz@biol.uni.lodz.pl (T.R.)

² University of Basel, Zoological Institute, Basel, 4051, Switzerland

³ University of Guelph, Centre for Biodiversity Genomics, Guelph, Ontario, N1G 2W1, Canada

⁴ National Horizons Centre, Teesside University, Darlington, DL1 1HG, United Kingdom; J.Bojko@tees.ac.uk (J. B.)

⁵ University of Lodz, Faculty of Biology & Environmental Protection, Biobank, Łódź, 90-237, Poland; dominik.strapagiel@biol.uni.lodz.pl (D. S.)

⁶ Polish Academy of Sciences, Institute of Oceanology, Department of Genetics and Marine Biotechnology, Powstańców Warszawy 55, 81-712 Sopot, Poland; aburzynski@iopan.pl (A. B.)

* Correspondence: tomasz.mamos@biol.uni.lodz.pl; Tel.: +48-42-635-44-46

Abstract: The Ponto-Caspian region is the main donor of invasive amphipods to freshwater ecosystems, with at least 13 species successfully established in European inland waters. *Dikerogammarus* spp. and *Pontogammarus robustoides* are among the most successful, due to their strong invasive impact on local biota. However, genomic knowledge on these invaders is scarce, while phylogeography and population genetics have been based on short fragments of mitochondrial markers or nuclear microsatellites. In this study, we provide: (i) reconstruction of four mitogenomes for four invasive gammarids; (ii) comparison between the structure of the newly obtained mitogenomes and to those from literature; (iii) SNP calling rates for individual *D. villosus* and *D. haemobaphes* from different invasion sites across Europe; and (iv) the first time-calibrated full mitogenome reconstruction of several Ponto-Caspian taxa.

We have sequenced, assembled and annotated four mitogenomes, each from *D. villosus*, *D. haemobaphes*, *D. bispinosus* and *Pontogammarus robustoides*. We found that, in comparison to other gammarids, the mitogenomes of Ponto-Caspian species show a translocation between the tRNA-E and tRNA-R positions. Phylogenetic reconstruction using the mitogenomes identified that Ponto-Caspian gammarids form a well-supported group that originated in the Miocene. Our study supports paraphyly in the family Gammaridae. These mitogenomes serve as vital genetic resources for the development of new markers for PCR-based identification methods and demographic studies.

Keywords: Amphipoda; invasive species; population genetics; mitogenome, Ponto-Caspian, SNP

1. Introduction

The Ponto-Caspian region (Azov, Black, and Caspian seas with surrounding areas) is a significant donor of invasive amphipods to European inland water bodies [1-3]. At least 13 morphospecies have successfully established populations to date [4,5]. Among them, *Dikerogammarus villosus* and *Dikerogammarus haemobaphes* have colonized most of the European main inland water bodies in less than 20 years, having a deteriorating effect upon local benthic communities [6-8]. *Dikerogammarus bispinosus* (a third invasive representative of the genus *Dikerogammarus*) reached the Rhine estuary via the southern

invasion corridor [1]; however, populations of this species have been declining in the lower section of the Danube in recent decades [9,10]. *Dikerogammarus villosus* and *D. haemobaphes* earned their nicknames “killer” and “demon” shrimp, respectively, due to their invasive impacts. They are the highly adaptable, physiologically tolerant, and efficient predators characterized by a high fecundity, which allows them to dominate local macroinvertebrate communities [8,11-13] as well as introduce pathogenic species [14-19]. In particular, *D. villosus* is regarded as one of the worst 100 invasive species in Europe [20,21] and has been deemed the worst non-native amphipod invader of English and Welsh waterways by the UK Environment Agency [22]. In recent years these species have accelerated their invasion, as new records, especially of *D. villosus*, were noted in England and Wales waterways by the Environment Agency [22], Baltic States [7] and Masurian Lake district in Poland [23]. *Dikerogammarus haemobaphes* has most recently invaded Boroughbridge in North Yorkshire, UK, carrying with it several invasive parasites [24]. The European invasion of *D. villosus* is followed by *D. haemobaphes* [6] and, most recently, by *P. robustoides* [25,26].

Phylogeographic and population genetics data have been attained for *D. villosus* throughout its invaded range [27-29]. Two invasive populations, one originating from the Danube and the other originating from the Dnieper were found in Central and Western Europe. They are genetically differentiated and allopatrically distributed, and neither of them show signs of a loss of genetic diversity compared to respective source areas. A recent phylogeographic study by [6] upon *D. haemobaphes* revealed the presence of cryptic lineages in the native region, as well as shallow differentiation in populations from the European invaded area. The evolutionary processes behind the invasion of these species are still under examination [6,27,28,30,31]. Little is known about the most recent colonisation of Europe by *P. robustoides*, including its molecular diversity [25]. Its dispersal routes and molecular divergence require exploration, since this species has the potential to be invasive [11,32].

The original descriptions of *D. villosus* (Sowinsky, 1894), *D. haemobaphes* (Eichwald, 1841) and *P. robustoides* placed them in the genus *Gammarus* (respectively *Gammarus marinus* var. *villosa* Sowinsky, 1894 and *Gammarus haemobaphes* Eichwald, 1841, *Gammarus robustoides* G.O. Sars, 1894) in the family Gammaridae Leach, 1814. Then, [33] established the genus *Dikerogammarus*, to which he moved *D. villosus* and *D. haemobaphes*, joined later by *Dikerogammarus bispinosus* Martynov, 1925. In 1904, [34] established the genus *Pontogammarus* into which *P. robustoides* was placed. [35], based on several morphological features shared by Ponto-Caspian species, challenged the former genus-level classification within Gammaridae, and placed *Dikerogammarus* together with several newly established genera into an informal taxonomic group, he called “*Dikerogammarus*-*Pontogammarus* complex” or “Ponto-Caspian complex of genera”. Following, [36], after re-evaluation of the morphological traits within the group, moved the whole “*Dikerogammarus*-*Pontogammarus* complex” to the family Pontogammaridae Bousfield 1977.

Several decades later, [37] conducted a general taxonomic revision of the Amphipoda, based on cladistic analysis of a number of morphological traits, which excluded *Dikerogammarus* from the Pontogammaridae and placed it back into the Gammaridae. More recently, molecular studies by [38] and [39], suggested high phylogenetic affinity of *Dikerogammarus* to other genera of the former “*Dikerogammarus*-*Pontogammarus* complex” (sensu [35]) that are classified within Pontogammaridae. Nevertheless, given the limited number of molecular markers used and the lack of resolution in the phylogenetic tree, the exact relationships within as well as between the pontogammarid clade and other gammarids were not fully identified. The above mentioned authors continued to refer to them as “Ponto-Caspian pontogammarids” or “Ponto-Caspian groups of genera”, without a conclusion into which family they should belong. Further, the within-genus taxonomy of *Dikerogammarus* continues to be a subject of discussion and is far from being fully resolved, mainly due to poor descriptions of some species, as well as missing type materials. For example, until recently, *D. bispinosus* was considered a subspecies or synonym of *D. villosus* Sowinsky 1894, while *D. villosus* was considered as a synonym of *D.*

haemobaphes Eichwald 1841 [40]. A molecular study by [31] resolved this question and confirmed the taxonomic status of *D. villosus*, *D. haemobaphes*, and *D. bispinosus* as three separate species. Nevertheless, the most recent studies show that while *D. villosus* seems to be a well-defined species over its geographic range [27], both *D. haemobaphes* and *D. bispinosus* contain divergent phylogenetic lineages, which may represent yet undescribed, cryptic or pseudocryptic species [6,41].

There is still relatively scarce knowledge on the genomics of the Amphipoda. To date, five amphipod (complete or partial) genomes are available for the species: *Orchestia grillus*, *Trinorchestia longiramus*, *Gammarus roeselii*, *Hyalella azteca*, and *Parhyale hawaiiensis* [42]. The main obstacle to progression is their large and repetitive genome. So far only two such genomes have been completed, first for *Parhyale hawaiiensis* with size reaching 3.6 Gb [43] and second for *Gammarus roeselii* with size estimated at 3.4Gb [44], both being among the largest reported arthropod genomes [43]. Another potential problem could be polyploidisation, which was documented for the Ohrid Lake *Gammarus* species flock [45]. Genetic data has also been gathered via a transcriptomic approach, providing nuclear and mitochondrial data for ~40 species of amphipod [46-49].

In the absence of genomic data, mitochondrial genomes are much easier to sequence and reconstruct and are often used in phylogenetic models [50-52]. The structure of the amphipod mitochondrial genome resembles the general structure of other animal mitogenomes; being small (14-18kb) circular molecules and having ~37 genes: 13 protein coding, 2 rRNA and 22 tRNA [53]. Due to their small size and abundance, the complete mitogenome sequences are relatively easy to obtain from next generation sequencing projects [43,54-57] as well as for the widely distributed freshwater Palearctic family, Gammaroidea [50-52,58]. It is common for an important resource, mitochondrial transcripts, to not be assembled and published, despite the fact that a large number of reads in each data set belong to the mitochondrial genome [59].

Pertinent to the *Dikerogammarus* spp., Bojko [55] published the first mitochondrial genome for a single *D. haemobaphes* from an invasive population in England, which became the first reported full mitogenome for the family Pontogammaridae. Bojko [55] reported a potential recombination and duplication of a tRNA, as well as a putative duplication of an ATP8-like gene within the control region (CR). The primary goal of our study is to provide a comparative reconstruction of mitogenomes of four invasive gammarids of Ponto-Caspian origin (*D. villosus*, *D. haemobaphes*, *D. bispinosus* and *P. robustoides*). Second, we will compare the structure of the obtained mitogenomes to those of the other gammarids available in public databases to search for putative rearrangements. Third, we will compare mitochondrial polymorphisms between individual *D. villosus* and *D. haemobaphes* that represent different invasive populations in Europe and, in the case of *D. villosus*, the native Ponto-Caspian population. Finally, we will provide the first time-calibrated reconstruction of the phylogenetic position of these Ponto-Caspian taxa versus other gammaroids using mitochondrial genes.

2. Results

2.1. Molecular species identification

The COI barcoding prior to sequencing supported morphological identification of studied Ponto-Caspian species. After using BLAST on all new assemblies done based on GenBank data, and precise identification through COI and 16S genes, there was a complete agreement most of the time. In two cases, the BLAST search of the assemblies showed misidentifications (Table S1). The first case was *Pandorites podoceroideus* (SRR3467097), which was identified as *Obesogammarus crassus* (the name used throughout the text). The second case was the transcriptome of *Gammarus pulex* (SRR8089725), which in fact showed presence of two mitochondrial genomes, one of *G. fossarum* and second belonging to *G. pulex*, only the latter was used in our analyses.

2.2. Structure

The three mitogenomes constructed for *D. villosus* consisted of 15,173 bp to 15,176 bp circular sequences in length. The mitogenome for *D. bispinosus* reached a length of 15,366 bp and the newly assembled mitogenome for *D. haemobaphes* (Germany) constituted 15,468 bp. For partial reconstruction of *P. robustoides* mitogenome we were able to obtain a single linear contig, 14,339 bp long. The 37 expected genes (Fig. 1) were annotated and resemble the canonical bilaterian gene set: 13 protein-coding genes, 2 rRNA genes, and 22 tRNA genes. Gene order and position for all Ponto-Caspian mitogenomes are the same (Fig. 2). The same pattern is shared by sister taxa: *Obesogammarus crassus* and *Homoeogammarus veneris* (Gammaridae). Relative positions of all protein and rRNA coding genes follow the Pancrustacea ground pattern, but the position and transcriptional polarity of some tRNA coding genes, in a few cases, is different (Fig. 2). A single rearrangement is observed for tRNA genes relative to the pattern observed for Gammarus spp. i.e. the tRNA-E and tRNA-R switched places in the Ponto-Caspian amphipods (*Dikerogammarus* spp., *O. crassus*) and in *H. veneris*. Since the sequences reconstructed from transcriptomic data have inherently poor coverage for tRNA expression (Fig S1), the data for some of the Ponto-Caspian sister taxa must be interpreted cautiously.

The mitogenomes have standard AT/GC ratio, having a GC content of 32.2% - *D. villosus*, 34% - *D. haemobaphes*, 30.6% - *D. bispinosus* and 32.2% in partial *P. robustoides* mtDNA. Rich AT composition is visible in the control region (Fig. 1), which is also characterised by the presence of a poly-T stretch and some tandemly repeated sequences.

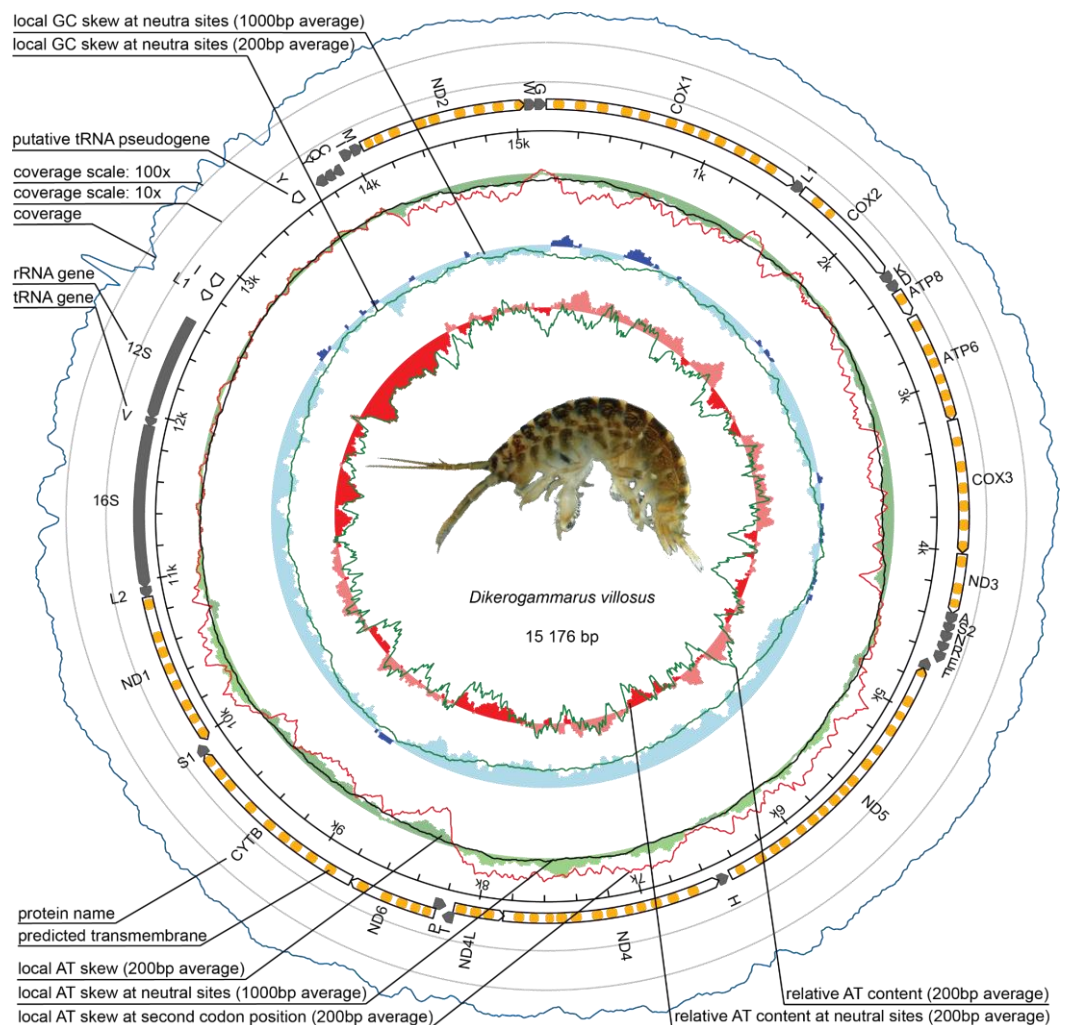


Figure 1. Genetic map of the mitochondrial genome for the invasive Ponto-Caspian amphipod, *Dikerogammarus villosus*. The mitogenome was generated using the mitoconstructor set of tools. Transfer RNA genes are labelled by their single-letter amino acid code (photo. of *D. villosus* by M. Grabowski).

2.3. SNP recovery

The comparison of reads from different ranges allowed us to identify several polymorphic sites across the mitogenome of both *D. villosus* and *D. haemobaphes* (Table S2). In total, 32 polymorphic sites differentiate *D. villosus* from Turkey and Poland, while 21 polymorphic sites differentiate between individuals from Poland and England. The *D. haemobaphes* from Germany differ from those collected in the UK by 52 polymorphic sites. All the polymorphic sites with the information on reading coverage, variant frequency and variant p-value are provided in Table S2. The polymorphic sites within genes, their position, and their impact on translation are provided in Table S2. For each species, a comparison of the highest number of polymorphic sites recorded are for NAD4 and NAD5 (4 to 7 SNPs). Equally, these are the longest genes (NAD5: 1659-1696 bp, NAD4 1242-1302 bp), except for COXI (1515-1535 bp); however, the COXI gene has only 2-3 SNP's recorded (Table S2).

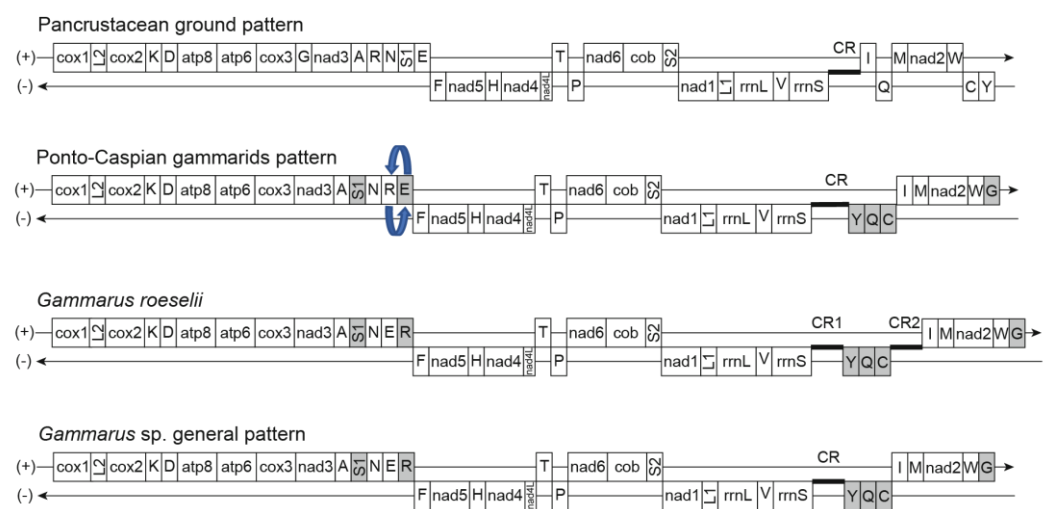


Figure 2. Organization of mitochondrial genomes in the Ponto-Caspian gammarids in comparison with the pancrustacean ground pattern and representative *Gammarus* genera. Gene features with altered location relative to the pancrustacean ground pattern are shown in a grey colour. Rearrangement of a tRNA is shown using blue arrows. A '+' indicates the forward DNA strand and '-' indicates the reverse DNA strand. Transfer RNA genes are labelled by their single-letter amino acid code. Figure modified after [50].

2.4. Phylogeny reconstruction

The reconstructed phylogenetic trees share topology, showing the same well supported clades (Fig. 3, Fig. S2, Fig. S3). For the Maximum Likelihood reconstruction, the protein-based tree gave a slightly higher support than the reconstruction based on mtDNA sequences. The Bayesian time-calibrated phylogeny reconstruction provided high posterior probabilities for most of the nodes (Fig. 3, Fig. S3).

According to this phylogeny, the superfamily Gammaroidea includes a few well-supported clades that diverged in the Early/Middle Eocene (Fig. 3, Fig. S3). First, the Baikalian families, Crypturopodidae and Micruropodidae, form a well-supported clade that forms a sister relationship to a clade containing all of the other representatives of the superfamily Gammaroidea included in our analysis. The latter is composed of two clades, whose phylogenetic relationships point to paraphyly for the family Gammaridae. One includes *E. berilloni*, *E. marinus*, *H. veneris* (Gammaridae) and the Ponto-Caspian genera i.e. *Dikerogammarus* (Gammaridae), *Pontogammarus* (Pontogammaridae) and *Obesogammarus* (Pontogammaridae). This clade started to diverge already in the middle Eocene, when the

E. berilloni lineage branched off, followed by the *E. marinus* lineage in the late Eocene. *Homoeogammarus veneris* is a sister lineage to the Ponto-Caspian genera, from which it diverged in the mid-Oligocene. The latter have diverged in the middle Miocene. The three *Dikerogammarus* species diverge at the mid-late Miocene and form a monophyletic clade relative to *P. robustoides* and *O. crassus*.

The second clade includes all of the Gammarus species (Gammaridae), plus the predominantly Baikalian Pallaseidae and the endemic Baikalian Acanthogammaridae and Eulimnogammaridae, as well as the Baikalian *Echiuropus macronychus*. However, this species is classified in the Crypturopodidae and instead is affiliated with the Eulimnogammaridae in our analysis. Interestingly, *Gammarus* sp. seems to be a polyphyletic genus provided that a set of highly divergent brackish water/freshwater species (*G. chevreuxi*, *G. duebeni*, *G. lacustris*) form a sister clade beside another one containing freshwater *Gammarus* sp. (*G. roeselii*, *G. fossarum*, *G. wautieri*, *G. pulex*), which branches as a sister clade to the Baikalian taxa. All of the above divergences may be dated to the middle and late Eocene. Finally, the *Gammarus* lineages seem to start their diversification in the late Eocene, while the diversification of the Baikalian taxa dates back further, to the late Oligocene/early Miocene.

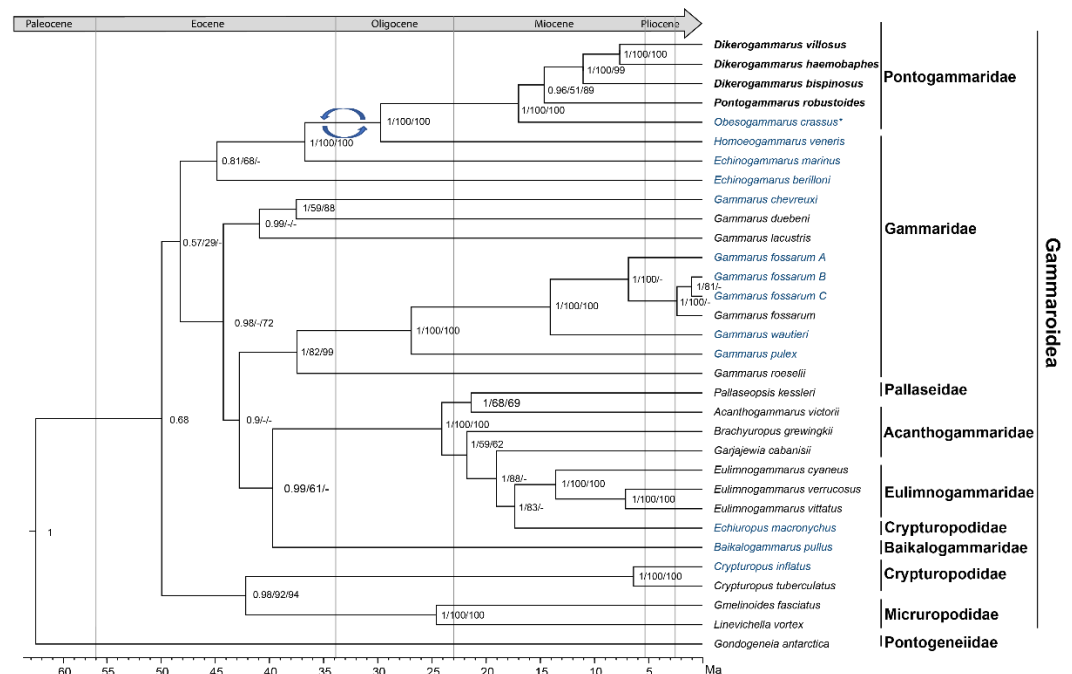


Figure 3. A Bayesian phylogenetic maximum clade credibility chronogram of the Gammaroidea, based on available mitogenomic and mitochondrial transcriptome data. The sources of sequences used are indicated in Table S1. Numbers next to branches indicate support: Bayesian posterior probabilities from chronogram/bootstraps values from Maximum Likelihood (ML) phylogeny on full data / bootstrap value for ML based on proteins (see Material and Methods). Arrows indicate a putative tRNA rearrangement. On the right, the vertical lines indicate the families and the superfamily. Species names in bold sequences indicated assembled “de novo” from own data, names in blue indicate newly reconstructed from third party transcriptomic data. *indicate *Pandorites podoceroideis* (SRR3467097), which was identified as *Obesogammarus crassus*

2.5. Substitution rates

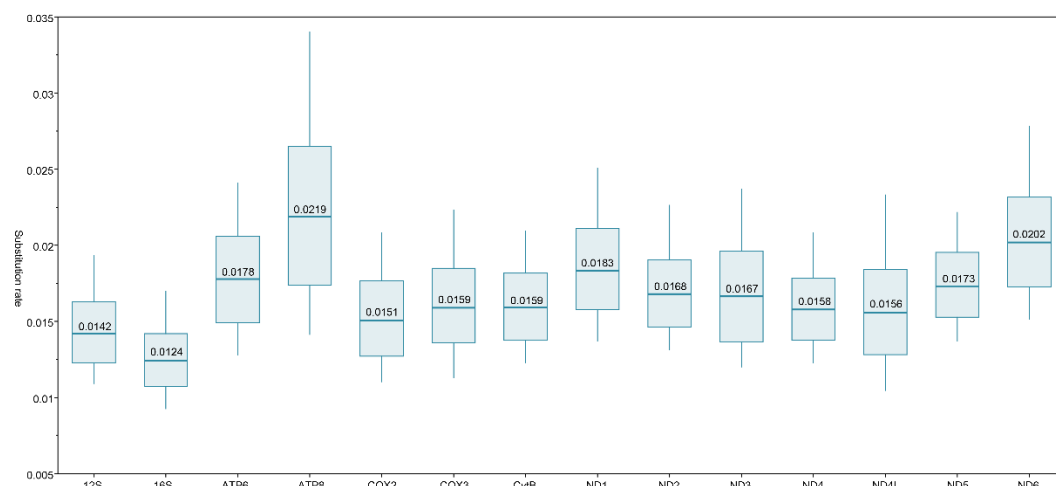


Figure 4. Substitution rates for 12 protein coding mitochondrial genes. The mean value is provided within the boxplot (substitution per site x My⁻¹), whiskers stand for value range, box stand for SD.

The mean relative substitution rate for the superfamily Gammaroidea falls within a range of 0.0150 to 0.0213 substitutions x My⁻¹, being similar to the pre-established COXI rate of 0.01773. The substitution rate of the relatively short ATP-8 gene has a value of 0.0258 (Fig. 4). All of the statistics for the protein coding gene substitution rates are provided in Table S3.

3. Discussion

Gammarids of Ponto-Caspian origin are the most prominent colonisers of European inland water bodies [4]. In most cases they are considered to be invaders, and a significant threat to native biodiversity [8,60]. The mitogenomic diversity data collected in this study provides information towards the resolution of their taxonomic affinities [39] and presents a foundation for further genetic tool development to benefit their identification, colonization route(s) and population dynamics. The accumulation of mitogenomic data for the group is progressing [55]; however, it is important to stress that all genomic resources, prior to incorporating in analyses should be properly validated. For superfamily Gammaroidea we have detected two cases of species misidentification.

After inspecting the assemblies based on transcriptome sequences for *G. pulex* provided by [48] it appears that the reads belong to two species, *G. pulex* and *G. fossarum*. COXI barcoding fragments allow us to identify *G. fossarum*, as belonging to BIN: BOLD:ACG7784 - a widely distributed species in France and Germany [61]. Both species are morphologically similar and may be hard to distinguish. The mitogenome assembled from the *G. fossarum* reads, mislabelled as *G. pulex*, was used in studies by [51] and [50]. The second misidentification detected in our study relates to *Pandorites podoceroideis*, whose mtDNA fits with *O. crassus* from the Ponto-Caspian region; both species belong to Pontogammaridae. This mislabelled data has been used by [49]. Both mentioned misidentification could be avoided through simple barcoding, which could verify morphological determination. These findings support the need for a proper barcode reference library for all organisms. Such misidentifications may result in avoidable consequences, as more countries and stakeholders include eDNA and metabarcoding as standard biomonitoring protocols [62-64].

In this case, *G. pulex* and *G. fossarum* have wide European ranges and are both often viewed as bioindicators [65] and model organisms for ecotoxicological studies [66]. Contrary to this, proper morphological discrimination of both species is not trivial and we can find cryptic (possibly hyper-cryptic) diversity when using molecular studies [61,67,68]. Moreover, correct species identification is crucial when studying gene expression,

especially when discrepancies of certain expression patterns may be the result of species-specific factors and not the experimental factor or stressor.

The mitochondrial genomes of Ponto-Caspian gammarids reveal a conservative order of genes and transcriptional polarity similar to the pancrustacean ground pattern [69]. Changes in the position and location of tRNA polarity have already been shown in mitogenome reconstructions for representatives of other gammarids [50,51,58]. The outstanding feature, for the Ponto-Caspian amphipods and their sister species *Homoeogammarus veneris* in comparison to other species from the superfamily Gammaroidea, is switch of position of two tRNA (E and R). This rearrangement was first observed for *D. haemobaphes* [55]. However, analysis and gene annotation conducted via our pipeline, cannot unequivocally confirm a non-similar duplication of the ATP8, or duplication of tRNA Q genes in the structure of *D. haemobaphes* or other Ponto-Caspian gammarids. It is possible that these duplications are a result of software errors, likely at the assembly or annotation step(s), which may be avoided with longer read sequencing or follow-up PCR validation in future studies. However, duplications are not rare in the amphipod mitogenome. A duplication of the control region in some Gammaridae was putatively identified in the Baikalian *Garjajewia cabanisii* [52], and also reported for the mitogenome of *G. roeselii* [50]. The mitogenomes of the Ponto-Caspian species that we have sequenced do not show such structures. However, in our case, the coverage of the control region is lower than in the study of [50] and is rather problematic for assembly due to high AT content and putative repetitions.

To date, studies using amphipod mitochondrial DNA focus primarily on phylogeographic and population dynamics, including only fragments of COI and 16S genes [6,70]. In our study we detected multiple SNPs across most of the mitochondrial genes for different populations, both for *D. villosus* and *D. haemobaphes*. Detection of SNPs at the species level can be valuable to the study of invasive organisms, helping to identify their source population(s) and determine their invasion corridor(s) [71-73]. Growing accessibility, together with falling prices of high throughput sequencing, will undoubtedly make multigene SNP data a primary resource to study biological invasions [74]. Especially using mitochondrial transcript that are particularly rich when sequencing only RNA.

Our phylogenetic analysis reveals well-supported relationships between four members of the Pontogammaridae using mitogenomics. The results show that the Pontogammaridae includes a supported monophyletic clade with close affinities to *Homeogammarus veneris*, which supports the results of [38]. Our observation that *Echinogammarus* is a sister genus to the aforementioned taxa is also supported by nuclear data from other studies [39]. [38] suggest moving *E. marinus* to the genus *Marinogammarus*, created for that species by [75]. Given that *E. marinus* is the type species for *Marinogammarus*, our results support such a claim. The type species of *Echinogammarus* is *E. berilloni*, present in our analysis and clearly belonging to another lineage. The monophyletic clade composed of Pontogammaridae and *Echinogammarus* (Gammaridae) was also shown by [50] using a concatenated dataset of mitochondrial coding genes; however, in their study, this group was a sister lineage to all other families of the superfamily Gammaroidea. Our study, utilising all mitochondrial genes with separate partitions, suggests a more complex pattern of Gammaridae evolution, including clades formed by Baikalian Micruropodidae and Crypturopodidae as sister groups to all other families. This deeper structure remains unresolved, even when using whole transcriptome data [49]. The paraphyly of the Gammaridae is confirmed in our study; however, unlike the case of the study by [50], with Baikalian families as well as the Pontogammaridae. These results support complex and ancient origins of the family Gammaridae and are further supported by studies using nuclear data [38,39]. Our study, among others, highlights that the whole superfamily requires taxonomical revision using integrative methods.

Our results suggest that the Eocene could be a backdrop to the burst of diversification observed for the Gammaroidea, resulting in the diversity of families existing today. This early time frame of their speciation is supported by previous studies based on nuclear and mitochondrial sequence data [38]. Interestingly, the diversification of the Pontogammaridae according to our study takes place between ca. 17 and 8 Ma, and coincides with the

time scale of their earliest recorded potential representative, whose fossil record resides in the Upper Sarmatian (ca. 9Ma: [76,77]).

4. Materials and Methods

4.1. Material collection and DNA isolation

Dikerogammarus villosus were collected from its invasive range in Poland (50.412N, 18.108E, April 2009, n=2), England (N 52.3024, W 0.3208, September 2016, n=2) and native range from Turkey (41.316N, 28.620E, September 2007, n=1). *Dikerogammarus haemobaphes* was collected in Germany (47.973N, 11.352E, May 2011, n=1). *Dikerogammarus bispinosus* was collected in Hungary (47.518N, 19.042E, April 2012, n=1). Material was collected using standard benthic hand-net and identified using available keys (i.e.: [78]). The voucher specimens and isolated DNA are stored at the Department of Invertebrate Zoology and Hydrobiology, University of Lodz. Material from England is stored at the Centre for Environment, Fisheries and Aquaculture Science (Cefas) repository. DNA isolation for all specimens follows a standard phenol-chloroform protocol according to a procedure from [79]. To confirm the species identification, the COXI (cytochrome c oxidase subunit 1) fragment was amplified and sequenced, following a procedure from [27]. The COX1 sequences were used in BLAST searches to confirm morphological identification.

4.2. Sequencing/Assembly

Isolates with confirmed species identification were tested for DNA quality and quantity using Qubit (Qubit dsDNA BR Assay) and Nanodrop machines. Libraries were prepared from 1ng of whole genomic DNA using the Nextera XT preparation kit (Illumina) for 2 times 150bp paired-end sequencing on an Illumina NextSeq500 sequencer (Illumina, San Diego, CA) in Biobank Lab, Department of Molecular Biophysics, University of Lodz.

The quality of the reads before and after the pre-processing steps were assessed using FastQC (v0.11.5) [80]. Quality trimming and removal of remaining sequencing adapters was performed with Trimmomatic [81]. Mitobim [82] and NOVOplasty [83] were used to pull out and assemble any mitochondrial genomes from raw DNA reads. The assemblies were then verified through mapping of mitochondrial reads using Bowtie2 [84]. Annotation and visualisation were performed using a mitoconstrictor set of tools [85].

In order to provide an extensive set of outgroups within the Gammaroidea superfamily, available databases were also searched, and data downloaded. The mitochondrial genes were obtained from annotated mitochondrial genomes available through NCBI (GenBank). Additionally, unannotated raw transcriptome data (NCBI SRA) were downloaded and assembled anew. A list of species and accession codes of new assemblies and read data are provided in Table S1. The same approach was used to obtain mitogenomic sequences from downloaded SRA data as described above, but the de novo assembly was conducted using Trinity [86] and an identification of the mitochondrial sequences was completed with wise2 [87] and infernal [88], using scripts from mitoconstrictor pipeline. The mitochondrial sequences were then validated through a realignment of the filtered reads. Mitochondrial contigs were further assembled in CLC genomic workbench (Quiagen) and verified by mapping the raw reads at the final, single mitochondrial contig.

Final annotation was performed using a mitoconstrictor set of tools. The reconstructed mitochondrial COX1 genes (for both: downloaded and own data) and, when available, rRNA genes were tested with blast vs. GenBank and our own unpublished database to ensure proper species identification of the data.

4.3. SNP recovery

To identify variation between specimens from each of the studied ranges, we performed SNP recovery. All reads per individual were mapped to the reference mitogenome using Bowtie2 [89]. In the case of *D. villosus*, the newly assembled mitogenome from Poland was used as the reference and the reads from native range (Turkey) and Invasive range (England) were used for mapping. For *D. haemobaphes*, the already published mitogenome from England (MK644228) was used as a reference and the reads from the newly

assembled mitogenome of *D. haemobaphes* from Germany were mapped. For the identification of the SNPs, we used the 'Find Variations/SNPs' tool through Geneious 11.1 software [90]. The initial minimum variant frequency was set to 0.35, the p-value was calculated using an approximate method, the maximum variant p-value was set to 10^{-6} and the minimum strand p-value to 10^{-5} . The variants were called from both coding and noncoding regions except for the AT rich control region.

4.4. Phylogenetic analysis with molecular clock calibration

To analyse the diversification of the Ponto-Caspian gammarids across both a phylogenetic and a temporal context, we reconstructed the amphipod phylogeny using a molecular clock approach. A set of 28 species with available mitogenome data from the superfamily Gammaroidea were used. *Gondogeneia antarctica* (Pontogeneiidae) was added as an outgroup (Table S1). The dataset included 13 protein-coding genes and 2 rRNA genes. The coding genes were aligned via MAFFT [91] with automatic determination of algorithm and gap open penalty set to 2.5, while rRNA were aligned using structural aligner software LocARNA [92]. The alignments were inspected by eye to identify possible misalignments of triplets, and in case of coding genes trimmed to codon position and, in all cases, trimmed when individual sequences were longer relative to the others, or due to putative miss-annotations (12S gene of *G. fossarum* - KY197961). Missing positions at the ends of the sequences were coded with unknown nucleotides to exclude them from the phylogenetic analyses. The whole set of genes reached a length of 12,862 bp.

The phylogeny was reconstructed using Bayesian inference in BEAST 2.6.2 [93]. The evolutionary rates for gammarids, especially for COX1, have been vastly studied and cross validated using both fossil and geological data (e.g. [94-96]), allowing us to apply the general rate of 0.01773 [76] for the COX1 gene. Four partitions were used: ND6 having a different substitution model was treated as a separate partition, 11 protein coding genes were concatenated in one partition and 12S and 16S rDNA genes with the same substitution model were also treated as one partition as well as COX1 as calibrated partition. The substitution model was selected via bModelTest [97]. Birth-Death tree model and relaxed clock were set as priors. Four runs of the MCMC, each 50 million generations long and sampled every 5,000 generations were performed and examined for convergence in TRACER 1.7 [98]. All parameters in each run reached the effective sample size (ESS) above 200 and were combined using LogCombiner2.6.2 [93]. The final tree was summarized with TreeAnnotator 2.6.2 [93].

To provide relative rates of evolution for each of the protein coding genes, the above analysis was also run using all genes as separate partitions. The difference in settings for the clock, in this case simple "strict clock", was used for each partition and the length of MCMC was set to 100 million generations.

To provide additional support for the BI topology, we have also reconstructed a phylogeny using the Maximum Likelihood approach through RAxML 8.2.8 [99]. For this analysis, all 15 genes were concatenated. The best-scoring ML tree was produced using the GTR+I+G substitution model. Bipartition information was drawn from phylogeny obtained with the rapid hill-climbing tree search algorithm. Statistical support was estimated with a thorough bootstrap test set to 1000 repetitions. To limit the impact on the phylogeny of highly divergent 3rd codon position, the phylogeny was reconstructed using amino acid sequences. All protein coding sequences were translated, realigned in MAFFT using BLOSUM65 [91] scoring matrix automatic selection of algorithm, and then concatenated. The tree was reconstructed in RAxML 8.2.8 using the same settings as above.

5. Conclusions

We present the first de novo full mitochondrial genomes of *D. villosus* and *D. bispinosus* as well as a new assembly for *D. haemobaphes* and a partial mitochondrial genome for *P. robustoides*. These mitogenomes serve as a vital resource for the development of new genetic markers for PCR-based identification methods, as well as SNP-based demographic studies. The reconstruction of these mitogenomes show that Ponto-Caspian

gammarids, and their sister lineage *Homoeogammarus veneris*, show a stable structural feature that involves the tRNA-E and tRNA-R being switched in place relative to other gammarids. Aside from this, the pancrustacean structural format is adhered to.

The phylogenetic reconstruction we provide for *Dikerogammarus* spp. and *P. robustoides*, based on the mitogenomes of Ponto-Caspian gammarids and other Gammaroidae, reveals a well-supported group that appears to have originated in the Miocene. Our study supports paraphyly in the Gammaridae family, advocating the need for detailed taxonomic revision using an integrative approach of the Gammaroidae and Ponotgammaridae families.

Supplementary Materials: Fig S1-S3, Table S1-S3

Author Contributions: Conceptualization, T.M. and A.B. with contribution of M.G. and T.R.; methodology, T.M. and A.B.; formal analysis, A.G. and T.M.; resources, D.S.; data curation, T.M. and A.G.; writing—original draft preparation, T.M.; writing—review and editing, M.G., T.R., J.B., A.B., D.S.; visualization, T.M. and A.G.; funding acquisition, D.S., T.R., T.M., J.B. All authors have read and agreed to the published version of the manuscript.

Funding: The study was partially funded by National Science Centre projects: 2018/31/D/NZ8/03061; 2017/01/X/NZ8/01086 and partially by the statutory funds of the Department of Invertebrate Zoology and Hydrobiology, University of Lodz. Tomasz Rewicz and Tomasz Mamos were supported by the Scholarship of the Polish National Agency for Academic Exchange (NAWA) at Bekker Programme (T.M. project nb. PPN/BEK/2018/1/00225, T.R. project nb. PPN/BEK/2018/1/00162). J.B. would like to thank NERC funding (#1368300) involved in the collection of *D. haemobaphes* data and personal funds at the University of Florida (USA) and Teesside University (UK) for further sequencing and analysis.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Nucleotide sequence data reported in this study are available in the Third Party Annotation Section of the DDBJ/ENA/GenBank databases under the accession numbers TPA: Provided after acceptance.

Acknowledgments: Authors would like to thank Jan-Niklas Macher (Naturalis Biodiversity Center, Leiden, Netherlands) for consultation and sharing of datasets at an early stage of work and Paulina Borówka for initiating contact between Dept. of Invert. Zoology and Hydrobiology UniLodz with Biobank UniLodz. Computer intensive tasks were run on PL-GRID infrastructure and sciCORE (<https://scicore.unibas.ch>) scientific computing center of the University of Basel.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bij de Vaate, A.; Jazdzewski, K.; Ketelaars, H.A.M.; Gollasch, S.; Van der Velde, G. Geographical patterns in range extension of Ponto-Caspian macroinvertebrate species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences* **2002**, *59*, 1159-1174, doi:Doi 10.1139/F02-098.
2. Copilaş-Ciocianu, D.; Sidorov, D. Taxonomic, ecological and morphological diversity of Ponto-Caspian gammaridean amphipods: a review. *bioRxiv* **2021**, 2021.2001.2021.427559, doi:10.1101/2021.01.21.427559.
3. Leuven, R.S.E.W.; van der Velde, G.; Baijens, I.; Snijders, J.; van der Zwart, C.; Lenders, H.J.R.; de Vaate, A.B. The river Rhine: a global highway for dispersal of aquatic invasive species. *Biological Invasions* **2009**, *11*, 1989-2008, doi:DOI 10.1007/s10530-009-9491-7.
4. Copilaş-Ciocianu, D.; Sidorov, D.; Šidagytė-Copilas, E. Global distribution and diversity of alien Ponto-Caspian amphipods. *bioRxiv* **2021**, 2021.2007.2019.452907, doi:10.1101/2021.07.19.452907.
5. Holdich, D.M.; Pöckl, M. Invasive crustaceans in European inland waters. In *Biological invaders in inland waters: profiles, distribution and threats.*, Gherardi, F., Ed.; Springer: Berlin, 2007; pp. 29-75.

6. Jazdzewska, A.M.; Rewicz, T.; Mamos, T.; Wattier, R.; Baćela-Spychalska, K.; Grabowski, M. Cryptic diversity and mtDNA phylogeography of the invasive demon shrimp, *Dikerogammarus haemobaphes* (Eichwald, 1841), in Europe. *NeoBiota* **2020**, *57*, 53.
7. Minchin, J.D.; Arbačiauskas, K.; Daunys, D.; Ezhova, E.; Grudule, N.; Kotta, J.; Molchanova, N.; Olenin, S.; Višinskienė, G.; Strake, S. Rapid expansion and facilitating factors of the Ponto-Caspian invader *Dikerogammarus villosus* within the eastern Baltic Sea. *Aquatic Invasions* **2019**, *14*, 165-181.
8. Rewicz, T.; Grabowski, M.; MacNeil, C.; Bacela-Spychalska, K. The profile of a 'perfect' invader – the case of killer shrimp, *Dikerogammarus villosus*. *Aquatic Invasions* **2014**, *9*, in press.
9. Borza, P.; Csányi, B.; Huber, T.; Leitner, P.; Paunović, M.; Remund, N.; Szekeres, J.; Wolfram, G. Longitudinal distributional patterns of Peracarida (Crustacea, Malacostraca) in the River Danube. *Fundamental and Applied Limnology / Archiv für Hydrobiologie* **2015**, *187*, 113-126, doi:10.1127/fal/2015/0769.
10. Borza, P.; Huber, T.; Leitner, P.; Remund, N.; Graf, W. Current velocity shapes co - existence patterns among invasive *Dikerogammarus* species. *Freshwater Biology* **2017**, *62*, 317-328.
11. Bacela - Spychalska, K.; van der Velde, G. There is more than one 'killer shrimp' : trophic positions and predatory abilities of invasive amphipods of Ponto - Caspian origin. *Freshwater Biology* **2013**, *58*, 730-741.
12. Constable, D.; Birkby, N.J. The impact of the invasive amphipod *Dikerogammarus haemobaphes* on leaf litter processing in UK rivers. *Aquatic ecology* **2016**, *50*, 273-281.
13. Grabowski, M.; Baćela, K.; Konopacka, A. How to be an invasive gammarid (Amphipoda : Gammaroidea)- comparison of life history traits. *Hydrobiologia* **2007**, *590*, 75-84, doi:DOI 10.1007/s10750-007-0759-6.
14. Bojko, J.; Stebbing, P.; Bateman, K.; Meatyard, J.; Bacela-Spychalska, K.; Dunn, A.; Stentiford, G. Baseline histopathological survey of a recently invading island population of 'killer shrimp', *Dikerogammarus villosus*. *Diseases of Aquatic Organisms* **2013**, *106*, 241-253.
15. Bojko, J.; Dunn, A.M.; Stebbing, P.D.; Ross, S.H.; Kerr, R.C.; Stentiford, G.D. *Cucumispora ornata* n. sp.(Fungi: Microsporidia) infecting invasive 'demon shrimp'(*Dikerogammarus haemobaphes*) in the United Kingdom. *Journal of Invertebrate Pathology* **2015**, *128*, 22-30.
16. Bacela-Spychalska, K.; Wattier, R.A.; Genton, C.; Rigaud, T. Microsporidian disease of the invasive amphipod *Dikerogammarus villosus* and the potential for its transfer to local invertebrate fauna. *Biological Invasions* **2012**, *14*, 1831-1842.
17. Subramaniam, K.; Behringer, D.C.; Bojko, J.; Yutin, N.; Clark, A.S.; Bateman, K.S.; van Aerle, R.; Bass, D.; Kerr, R.C.; Koonin, E.V. A new family of DNA viruses causing disease in crustaceans from diverse aquatic biomes. *MBio* **2020**, *11*, e02938-02919.
18. Allain, T.W.; Stentiford, G.D.; Bass, D.; Behringer, D.C.; Bojko, J. A novel nudivirus infecting the invasive demon shrimp *Dikerogammarus haemobaphes* (Amphipoda). *Scientific reports* **2020**, *10*, 1-13.
19. Bojko, J.; Stentiford, G.D.; Stebbing, P.D.; Hassall, C.; Deacon, A.; Cargill, B.; Pile, B.; Dunn, A.M. Pathogens of *Dikerogammarus haemobaphes* regulate host activity and survival, but also threaten native amphipod populations in the UK. *Diseases of aquatic organisms* **2019**, *136*, 63-78.
20. DAISIE. *Handbook of Alien Species in Europe*; Springer: Dordrecht, The Netherlands, 2009; Volume 3.
21. Devin, S.; Beisel, J.-N. *Dikerogammarus villosus* (Sowinsky), killer shrimp (Gammaridae, Crustacea). Chapter 13: Species accounts of 100 of the most invasive alien species in Europe. **2009**.
22. News, B. 'Killer' shrimp is the worst alien invader of Britain's waterways which officials say are costing billions of pounds to tackle. Available online: <http://www.bbc.co.uk/news/uk-14428585> (accessed on 01 September 2021).
23. Bacela-Spychalska, K.; Grabowski, M.; Mamos, T.; Rewicz, T. Data about distribution of *Dikerogammarus villosus* in Masurian Lake District in Poland. *Unpublished work* **2021**.

24. Burgess, A.; Bojko, J. Microsporidia are coming: *Cucumispora ornata* and *Dictyocoela berillonum* invade Northern Britain. *BioInvasions Records* **2021**, Submitted.
25. Csabai, Z.; Borza, P.; Rewicz, T.; Pernecker, B.; Berta, B.J.; Móra, A. Mass appearance of the Ponto-Caspian invader *Pontogammarus robustoides* in the River Tisza catchment: bypass in the southern invasion corridor? *Knowledge & Management of Aquatic Ecosystems* **2020**, *9*.
26. Moedt, S.; van Haaren, T. *Pontogammarus robustoides* (Sars, 1894), a new non-indigenous amphipod in the Netherlands (Crustacea: Amphipoda). *Lauterbornia* **2018**, *85*, 123-126.
27. Rewicz, T.; Wattier, R.; Grabowski, M.; Rigaud, T.; Bacela-Spychalska, K. Out of the Black Sea: phylogeography of the invasive killer shrimp *Dikerogammarus villosus* across Europe. *PLoS One* **2015**, *10*, e0118121, doi:10.1371/journal.pone.0118121.
28. Rewicz, T.; Wattier, R.; Rigaud, T.; Grabowski, M.; Mamos, T.; Bacela-Spychalska, K. 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* **2017**, *62*, 1036-1051, doi:<https://doi.org/10.1111/fwb.12923>.
29. Wattier, R.A.; Haine, E.R.; Beguet, J.; Martin, G.; Bollache, L.; Musko, I.B.; Platvoet, D.; Rigaud, T. No genetic bottleneck or associated microparasite loss in invasive populations of a freshwater amphipod. *Oikos* **2007**, *116*, 1941-1953, doi:DOI 10.1111/j.2007.0030-1299.15921.x.
30. Bacela-Spychalska, K.; Grabowski, M.; Rewicz, T.; Konopacka, A.; Wattier, R. The 'killer shrimp' *Dikerogammarus villosus* (Crustacea, Amphipoda) invading Alpine lakes: overland transport by recreational boats and scuba-diving gear as potential entry vectors? *Aquatic Conservation: Marine and Freshwater Ecosystems* **2013**, *23*, 606-618, doi:<https://doi.org/10.1002/aqc.2329>.
31. Müller, J.C.; Schramm, S.; Seitz, A. Genetic and morphological differentiation of *Dikerogammarus* invaders and their invasion history in Central Europe. *Freshwater Biology* **2002**, *47*, 2039-2048, doi:DOI 10.1046/j.1365-2427.2002.00944.x.
32. Kobak, J.; Rachalewski, M.; Bacela-Spychalska, K. Conquerors or exiles? Impact of interference competition among invasive Ponto-Caspian gammarideans on their dispersal rates. *Biological Invasions* **2016**, *18*, 1953-1965, doi:10.1007/s10530-016-1140-3.
33. Stebbing, T.R.R. Amphipoda from the Copenhagen Museum and other sources, part II. *The Transactions of the Linnean Society of London: Zoology* **1899**, *2*, 395-432.
34. Sowinsky, V.K. [Introduction a l'etude de la faune du bassin marin Ponto-Aralo-Kaspien sous le point de vue d'une province zoo-geographique independante (russ.)]. *Zapiski Kievskago Obshchestva Estestvoispytatelei (Memoires de la societe de naturalistes de Kiew)* **1904**, *18*, 1-487.
35. Stock, J.H. The systematics of certain Ponto-Caspian Gammaridae (Crustacea, Amphipoda). *Mitteilungen aus dem Hamburgischen Zoologischen Museum und Institut* **1974**, *70*, 75-95.
36. Bousfield, E.L. A new look at the systematics of Gammaridean amphipods of the world. *Crustaceana Supplement* **1977**, *4*, 282-316.
37. Lowry, J.K.; Myers, A.A. A Phylogeny and Classification of the Senticaudata subord. nov. Crustacea: Amphipoda). *Zootaxa* **2013**, *3610*, 1-80, doi:10.11646/zootaxa.3610.1.1.
38. Hou, Z.; Sket, B. A review of Gammaridae (Crustacea: Amphipoda): the family extent, its evolutionary history, and taxonomic redefinition of genera. *Zoological Journal of the Linnean Society* **2016**, *176*, 323-348, doi:<https://doi.org/10.1111/zoj.12318>.
39. Sket, B.; Hou, Z. Family Gammaridae (Crustacea: Amphipoda), mainly its *Echinogammarus* clade in SW Europe. Further elucidation of its phylogeny and taxonomy. *Acta biologica Slovenica* **2018**, *61*, 93-102.

40. Pjatakova, G.M.; Tarasov, A.G. Caspian Sea amphipods: biodiversity, systematic position and ecological peculiarities of some species. *Journal of Salt Lake Research* **1996**, *5*, 63-79.
41. Morhun, H.; Copilaş-Ciocianu, D.; Rewicz, T.; Son, M.O.; Khomenko, A.; Huseynov, M.; Utevsky, S.; Grabowski, M. Molecular markers and SEM imaging reveal pseudocryptic diversity within the Ponto-Caspian low-profile amphipod invader *Dikerogammarus bispinosus*. *The European Zoological Journal* **2021**, submitted.
42. Coordinators, N.R. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **2016**, *44*, D7-D19, doi:10.1093/nar/gkv1290.
43. Kao, D.; Lai, A.G.; Stamataki, E.; Rosic, S.; Konstantinides, N.; Jarvis, E.; Di Donfrancesco, A.; Pouchkina-Stancheva, N.; Sémon, M.; Grillo, M.; et al. The genome of the crustacean *Parhyale hawaiiensis*, a model for animal development, regeneration, immunity and lignocellulose digestion. *eLife* **2016**, *5*, doi:10.7554/elife.20062.
44. Cormier, A.; Chebbi, M.A.; Giraud, I.; Wattier, R.; Teixeira, M.; Gilbert, C.; Rigaud, T.; Cordaux, R. Comparative Genomics of Strictly Vertically Transmitted, Feminizing Microsporidia Endosymbionts of Amphipod Crustaceans. *Genome Biology and Evolution* **2021**, *13*, doi:10.1093/gbe/evaa245.
45. Salemaa, H.; Kamal'tynov, R.M. The chromosome numbers of endemic Amphipoda and Isopoda – an evolutionary paradox in the ancient lakes Ohrid and Baikal. *Archiv fuer Hydrobiologie Ergebnisse der Limnologie, Beiheft* **1994**, *44*, 247-256.
46. Caputo, D.R.; Robson, S.C.; Werner, I.; Ford, A.T. Complete transcriptome assembly and annotation of a critically important amphipod species in freshwater ecotoxicological risk assessment: *Gammarus fossarum*. *Environment International* **2020**, *137*, 105319, doi:<https://doi.org/10.1016/j.envint.2019.105319>.
47. Cogne, Y.; Degli-Esposti, D.; Pible, O.; Gouveia, D.; François, A.; Bouchez, O.; Eché, C.; Ford, A.; Geffard, O.; Armengaud, J.; et al. De novo transcriptomes of 14 gammarid individuals for proteogenomic analysis of seven taxonomic groups. *Scientific Data* **2019**, *6*, 184, doi:10.1038/s41597-019-0192-5.
48. Gismondi, E.; Thomé, J.P. Transcriptome of the freshwater amphipod *Gammarus pulex hepatopancreas*. *Genomics Data* **2016**, *8*, 91-92, doi:<https://doi.org/10.1016/j.gdata.2016.04.002>.
49. Naumenko, S.A.; Logacheva, M.D.; Popova, N.V.; Klepikova, A.V.; Penin, A.A.; Bazykin, G.A.; Etingova, A.E.; Mugue, N.S.; Kondrashov, A.S.; Yampolsky, L.Y. Transcriptome - based phylogeny of endemic Lake Baikal amphipod species flock: fast speciation accompanied by frequent episodes of positive selection. *Molecular ecology* **2017**, *26*, 536-553.
50. Cormier, A.; Wattier, R.; Teixeira, M.; Rigaud, T.; Cordaux, R. The complete mitochondrial genome of *Gammarus roeselii* (Crustacea, Amphipoda): insights into mitogenome plasticity and evolution. *Hydrobiologia* **2018**, *825*, 197-210, doi:10.1007/s10750-018-3578-z.
51. Macher, J.N.; Leese, F.; Weigand, A.M.; Rozenberg, A. The complete mitochondrial genome of a cryptic amphipod species from the *Gammarus fossarum* complex. *Mitochondrial DNA Part B* **2017**, *2*, 17-18, doi:10.1080/23802359.2016.1275844.
52. Romanova, E.V.; Aleoshin, V.V.; Kamal'tynov, R.M.; Mikhailov, K.V.; Logacheva, M.D.; Sirotinina, E.A.; Gornov, A.Y.; Anikin, A.S.; Sherbakov, D.Y. Evolution of mitochondrial genomes in Baikalian amphipods. *BMC Genomics* **2016**, *17*, 1016, doi:10.1186/s12864-016-3357-z.
53. Boore, J.L. Animal mitochondrial genomes. *Nucleic Acids Res* **1999**, *27*, 1767-1780, doi:10.1093/nar/27.8.1767.
54. Aunins, A.W.; Nelms, D.L.; Hobson, C.S.; King, T.L. Comparative mitogenomic analyses of three North American stygobiont amphipods of the genus *Stygobromus* (Crustacea: Amphipoda). *Mitochondrial DNA Part B* **2016**, *1*, 560-563, doi:10.1080/23802359.2016.1174086.

55. Bojko, J. The mitochondrial genome of UK (non-native) *Dikerogammarus haemobaphes* (Amphipoda: Gammaridae) informs upon *Dikerogammarus* evolution, invasions and associated microparasites. *Hydrobiologia* **2020**, *847*, 229-242, doi:10.1007/s10750-019-04084-1.
56. Lan, Y.; Sun, J.; Bartlett, D.H.; Rouse, G.W.; Tabata, H.G.; Qian, P.-Y. The deepest mitochondrial genome sequenced from Mariana Trench *Hirondellea gigas* (Amphipoda). *Mitochondrial DNA Part B* **2016**, *1*, 802-803, doi:10.1080/23802359.2016.1214549.
57. Pons, J.; Bauzà-Ribot, M.M.; Jaume, D.; Juan, C. Next-generation sequencing, phylogenetic signal and comparative mitogenomic analyses in Metacrangonyctidae (Amphipoda: Crustacea). *BMC genomics* **2014**, *15*, 1-16.
58. Krebs, L.; Bastrop, R. The mitogenome of *Gammarus duebeni* (Crustacea Amphipoda): A new gene order and non-neutral sequence evolution of tandem repeats in the control region. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **2012**, *7*, 201-211.
59. Smith, D.R. RNA-Seq data: a goldmine for organelle research. *Briefings in Functional Genomics* **2013**, *12*, 454-456, doi:10.1093/bfpg/els066.
60. Dick, J.T.A.; Platvoet, D.; Kelly, D.W. Predatory impact of the freshwater invader *Dikerogammarus villosus* (Crustacea : Amphipoda). *Canadian Journal of Fisheries and Aquatic Sciences* **2002**, *59*, 1078-1084, doi:10.1139/F02-074.
61. Wattier, R.; Mamos, T.; Copilaș-Ciocianu, D.; Jelić, M.; Ollivier, A.; Chaumot, A.; Danger, M.; Felten, V.; Piscart, C.; Žganec, K.; et al. Continental-scale patterns of hyper-cryptic diversity within the freshwater model taxon *Gammarus fossarum* (Crustacea, Amphipoda). *Scientific Reports* **2020**, *10*, 16536, doi:10.1038/s41598-020-73739-0.
62. Blackman, R.C.; Mächler, E.; Altermatt, F.; Arnold, A.; Beja, P.; Boets, P.; Egeter, B.; Elbrecht, V.; Filipe, A.F.; Jones, J.I.; et al. Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. *Metabarcoding and Metagenomics* **2019**, *3*, e34735.
63. Gold, Z.; Sprague, J.; Kushner, D.J.; Zerecero Marin, E.; Barber, P.H. eDNA metabarcoding as a biomonitoring tool for marine protected areas. *PloS one* **2021**, *16*, e0238557.
64. Pawlowski, J.; Bonin, A.; Boyer, F.; Cordier, T.; Taberlet, P. Environmental DNA for biomonitoring. *Molecular Ecology* **2021**, *30*, 2931-2936, doi:<https://doi.org/10.1111/mec.16023>.
65. Windisch, U.; Springer, F.; Stahl, T. Freshwater amphipods (*Gammarus pulex/fossarum*) and brown trout as bioindicators for PFC contamination with regard to the aquatic ecological status of a small stream. *Environmental Sciences Europe* **2020**, *32*, 108, doi:10.1186/s12302-020-00384-9.
66. Džeroski, S.; Demšar, D.; Grbović, J. Predicting Chemical Parameters of River Water Quality from Bioindicator Data. *Applied Intelligence* **2000**, *13*, 7-17, doi:10.1023/A:1008323212047.
67. Lagrue, C.; Wattier, R.; Galipaud, M.; Gauthey, Z.; Rullmann, J.-P.; Dubreuil, C.; Rigaud, T.; Bollache, L. Confrontation of cryptic diversity and mate discrimination within *Gammarus pulex* and *Gammarus fossarum* species complexes. *Freshwater Biology* **2014**, *59*, 2555-2570, doi:<https://doi.org/10.1111/fwb.12453>.
68. Weiss, M.; Macher, J.N.; Seefeldt, M.A.; Leese, F. Molecular evidence for further overlooked species within the *Gammarus fossarum* complex (Crustacea: Amphipoda). *Hydrobiologia* **2014**, *721*, 165-184.
69. Boore, J.L.; Lavrov, D.V.; Brown, W.M. Gene translocation links insects and crustaceans. *Nature* **1998**, *392*, 667-668, doi:10.1038/33577.
70. Csapó, H.; Krzywoźniak, P.; Grabowski, M.; Wattier, R.; Bącela-Spychalska, K.; Mamos, T.; Jelić, M.; Rewicz, T. Successful post-glacial colonization of Europe by single lineage of freshwater amphipod from its Pannonian Plio-Pleistocene diversification hotspot. *Scientific Reports* **2020**, *10*, 18695, doi:10.1038/s41598-020-75568-7.

71. Schmidt, T.L.; van Rooyen, A.R.; Chung, J.; Endersby-Harshman, N.M.; Griffin, P.C.; Sly, A.; Hoffmann, A.A.; Weeks, A.R. Tracking genetic invasions: Genome-wide single nucleotide polymorphisms reveal the source of pyrethroid-resistant *Aedes aegypti* (yellow fever mosquito) incursions at international ports. *Evolutionary Applications* **2019**, *12*, 1136-1146, doi:<https://doi.org/10.1111/eva.12787>.
72. Taylor, A.T.; Bangs, M.; Long, J.M. Sibship reconstruction with SNPs illuminates the scope of a cryptic invasion of Asian Swamp Eels (*Monopterus albus*) in Georgia, USA. *Biological Invasions* **2020**, *23*, 569-580.
73. Wu, Y.; Bogdanowicz, S.M.; Andres, J.A.; Vieira, K.A.; Wang, B.; Cossé, A.; Pfister, S.E. Tracking invasions of a destructive defoliator, the gypsy moth (*Erebidae: Lymantria dispar*): Population structure, origin of intercepted specimens, and Asian introgression into North America. *Evolutionary Applications* **2020**, *13*, 2056-2070, doi:<https://doi.org/10.1111/eva.12962>.
74. Rius, M.; Bourne, S.; Hornsby, H.G.; Chapman, M. Applications of next-generation sequencing to the study of biological invasions. *Curr Zool* **2015**, *61*, 488-504.
75. Schellenberg, A. Schlüssel und Diagnosen der dem Süßwasser-Gammarus nahestehenden Einheiten ausschliesslich der Arten des Baikalses und Australiens. *Zoologischer Anzeiger* **1937**, *117*, 267-280.
76. Copilaş-Ciocianu, D.; Sidorov, D.; Gontcharov, A. Adrift across tectonic plates: molecular phylogenetics supports the ancient Laurasian origin of old limnic crangonyctid amphipods. *Organisms Diversity & Evolution* **2019**, *19*, 191-207.
77. Derzhavin, A. Notes on the Upper Sarmatian amphipods of the Ponto-Caspian region. *Bulletin de la Societe des Naturalistes de Moscou* **1927**, *2*, 183-196.
78. Zettler, M.L.; Zettler, A. *Marine and freshwater Amphipoda from the Baltic Sea and adjacent territories*; ConchBooks: 2017.
79. Grabowski, M.; Rewicz, T.; Bacela-Spychalska, K.; Konopacka, A.; Mamos, T.; Jazdzewski, K. Cryptic invasion of Baltic lowlands by freshwater amphipod of Pontic origin. *Aquatic Invasions* **2012**, *7*.
80. Andrews, S. FastQC: a quality control tool for high throughput sequence data. **2010**.
81. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114-2120, doi:10.1093/bioinformatics/btu170.
82. Hahn, C.; Bachmann, L.; Chevreux, B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res* **2013**, *41*, e129-e129.
83. Dierckxsens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res* **2017**, *45*, e18-e18.
84. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **2012**, *9*, 357-359.
85. Lubośny, M.; Przyłucka, A.; Śmietanka, B.; Burzyński, A. Semimytilus alcosus: first known hermaphroditic mussel with doubly uniparental inheritance of mitochondrial DNA. *Scientific reports* **2020**, *10*, 1-12.
86. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **2011**, *29*, 644-652.
87. Birney, E.; Clamp, M.; Durbin, R. GeneWise and genomewise. *Genome research* **2004**, *14*, 988-995.
88. Nawrocki, E.P.; Eddy, S.R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **2013**, *29*, 2933-2935.
89. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754-1760.
90. Geneious 11.1. Available online: <https://www.geneious.com> (accessed on 01 September 2021).
91. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **2013**, *30*, 772-780.

92. Raden, M.; Ali, S.M.; Alkhnabashi, O.S.; Busch, A.; Costa, F.; Davis, J.A.; Eggenhofer, F.; Gelhausen, R.; Georg, J.; Heyne, S.; et al. Freiburg RNA tools: a central online resource for RNA-focused research and teaching. *Nucleic Acids Res* **2018**, *46*, W25-W29, doi:10.1093/nar/gky329.
93. Bouckaert, R.; Heled, J.; Kühnert, D.; Vaughan, T.; Wu, C.-H.; Xie, D.; Suchard, M.A.; Rambaut, A.; Drummond, A.J. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology* **2014**, *10*, e1003537, doi:10.1371/journal.pcbi.1003537.
94. Copilaș-Ciocianu, D.; Petrusek, A. Phylogeography of a freshwater crustacean species complex reflects a long-gone archipelago. *Journal of Biogeography* **2017**, *44*, 421-432, doi:<https://doi.org/10.1111/jbi.12853>.
95. Grabowski, M.; Mamos, T.; Bączela-Spychalska, K.; Rewicz, T.; Wattier, R.A. Neogene paleogeography provides context for understanding the origin and spatial distribution of cryptic diversity in a widespread Balkan freshwater amphipod. *PeerJ* **2017**, *5*, e3016, doi:10.7717/peerj.3016.
96. Mamos, T.; Wattier, R.; Burzyński, A.; Grabowski, M. The legacy of a vanished sea: a high level of diversification within a European freshwater amphipod species complex driven by 15 My of Paratethys regression. *Molecular Ecology* **2016**, *25*, 795-810, doi:<https://doi.org/10.1111/mec.13499>.
97. Bouckaert, R.R.; Drummond, A.J. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology* **2017**, *17*, 42, doi:10.1186/s12862-017-0890-6.
98. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic biology* **2018**, *67*, 901.
99. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* **2014**, *30*, 1312-1313, doi:10.1093/bioinformatics/btu033.