

# First annotated draft genomes of non-marine ostracods (Ostracoda, Crustacea) with different reproductive modes

Patrick Tran Van <sup>1,2</sup>, Yoann Anselmetti<sup>3,4</sup>, Jens Bast<sup>1,5</sup>, Zoé Dumas<sup>1</sup>, Nicolas Galtier<sup>3</sup>, Kamil S. Jaron<sup>1</sup>, Koen Martens<sup>6,7</sup>, Darren J. Parker<sup>1,2</sup>, Marc Robinson-Rechavi<sup>1,2</sup>, Tanja Schwander<sup>1</sup>, Paul Simion<sup>3,8</sup>, Isa Schön<sup>7,9</sup>

(1) Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

ORCID of PTV: 0000-0003-4236-8412

ORCID of KSJ: 0000-0003-1470-5450

ORCID of DJP: 0000-0003-4027-7779

ORCID of MRR: 0000-0002-3437-3329

ORCID of TS: 0000-0003-1945-5374

(2) Swiss Institute of Bioinformatics, Lausanne, Switzerland

(3) ISEM – Institut des Sciences de l’Evolution, Montpellier, France

ORCID of YA: 0000-0002-6689-1163

ORCID of NG: 0000-0002-0479-4878

ORCID of PS: 0000-0002-6667-3297

(4) Current address: CoBIUS lab, Department of Computer Science, University of Sherbrooke, 2500 Boulevard de l’Université, Sherbrooke, QC J1K 2R1, Canada

(5) Current address: Institute for Zoology, University of Cologne, Köln, Germany

ORCID of JB: 0000-0003-0017-3860

(6) Royal Belgian Institute of Natural Sciences, OD Nature, Freshwater Biology, Brussels, Belgium

(7) University of Ghent, Dept Biology, Ghent, Belgium

ORCID of KM: 0000-0001-8680-973X

(8) Université de Namur, LEGE, URBE, Namur, 5000, Belgium

(9) University of Hasselt, Research Group Zoology, Diepenbeek, Belgium

ORCID of IS: 0000-0001-9269-6487

Authors 2-11 in alphabetical order

## KEYWORDS:

ancient asexual, sexual, *Darwinula stevensoni*, *Cyprideis torosa*

## ABSTRACTS

Ostracods are one of the oldest crustacean groups with an excellent fossil record and high importance for phylogenetic analyses but genome resources for this class are still lacking. We have successfully assembled and annotated the first reference genomes for three species of non-marine ostracods; two with obligate sexual reproduction (*Cyprideis torosa* and *Notodromas monacha*) and the putative ancient asexual *Darwinula stevensoni*. This kind of genomic research has so far been impeded by the small size of most ostracods and the absence of genetic resources such as linkage maps or BAC libraries that were available for other crustaceans. For genome assembly, we used an Illumina-based sequencing technology, resulting in assemblies of similar sizes for the three species (335-382Mb) and with scaffold numbers and their N50 (19-56 kb) in the same orders of magnitude. Gene annotations were guided by transcriptome data from each species. The three assemblies are relatively complete with BUSCO scores of 92-96%, and thus exceed the quality of several other published crustacean genomes obtained with similar techniques. The number of predicted genes (13,771-17,776) is in the same range as Branchiopoda genomes but lower than in most malacostracan genomes. These three reference genomes from non-marine ostracods provide the urgently needed basis to further develop ostracods as models for evolutionary and ecological research.

## BACKGROUND

### Relevance of ostracods

Ostracoda are small, bivalved crustaceans, widely occurring in almost all aquatic habitats as part of the meiobenthos and periphyton. There are 2,330 formally described species of extant non-marine ostracods (Meisch *et al.* 2019) and at least another 7,000 described species of extant marine ostracod species (see Schön and Martens 2016 for an estimate by S. Brandao). Their calcified valves are preserved as microfossils, making them the extant arthropod group with the most extensive fossil record. The group has an estimated (Cambrian) age of c 500 myr (millions of years) according to a molecular clock (Oakley *et al.* 2013), and c. 450 myr (Ordovician; Maddocks 1982) to 509 myr (Wolfe *et al.* 2016) according to the fossil record. This makes them one of the oldest extant pancrustacean groups (Figure 1). Because of their excellent fossil data, evolutionary events can be dated with real time estimates making ostracods ideal models for evolutionary research (Butlin and Menozzi 2000; Oakley and Cunningham 2002; Oakley *et al.* 2013; Schön and Martens 2016;).

Contrary to the extensive focus on this group for palaeontological research, there is a total lack of published ostracod genomes, and even isolated genomic data from ostracods in open access databases are still rare. Thus, the only resources available beyond individual gene sequences are four mitogenomes (the marine ostracods *Vargula hilgendorfii* (Ogoh and Ohmiya 2004; GenBank accession number NC\_005306) and *Cypridina dentata* (Wang *et al.* 2019; NC\_042792); and two unpublished mitogenomes from *V. tsujii* (NC\_039175) and *Cyprideis torosa* (PRJNA302529)). Also, raw Illumina DNA sequencing reads of the podocopid ostracod *Eucypris virens* have been generated as part of a study testing DNA extraction methods for high throughput sequencing in zooplankton (SRX8021019; Beninde *et al.* 2020) but these have neither been assembled nor annotated. In studies on crustacean phylogenies and gene expression (see Table S1 for details), raw RNA-sequencing reads have been generated for a total of 12 species coming from the three major ostracod lineages (Mydocopida, Halocyprida and Podocopida), but the number of assembled and annotated ostracod genes in these studies remains very limited, ranging between 4 and 822 genes.

### Choice of model species

Extant non-marine ostracods show a high prevalence of asexual reproduction (Chaplin *et al.* 1994; Butlin *et al.* 1998; Martens *et al.* 1998), which has evolved several times independently in different ostracod lineages and is most frequent in the Cyprididae and the Darwinulidae. Ostracods are thus an ideal group to further study the paradox of sex, which remains one of the most puzzling questions in evolutionary biology (Bell 1982; Otto and Lenormand 2002; Schön *et al.* 2009a; Neiman *et al.* 2018). The most important sets of hypotheses explaining why sex is advantageous despite its direct costs are based on the fact that physical linkage among loci generates different forms of selective interference (recently reviewed in Otto 2020). Genome-wide data are very valuable to test if asexuals indeed are affected by these predictions (e.g., Glémin *et al.* 2019; Jaron *et al.* 2020) and to develop insights into mechanisms such as gene conversion (Omilian *et al.* 2006), DNA repair (Schön and Martens, 1998; Hecox-Lea and Mark Welch 2018) or horizontal gene transfer (Gladyshev *et al.* 2008; Danchin *et al.* 2010; Boschetti *et al.* 2012; Flot *et al.* 2013; Paganini *et al.* 2012). Such data are also needed to further test for general consequences of asexuality beyond lineage-specific

effects (Jaron *et al.* 2020). For many animal groups in which asexuality is frequent, genomic data are limited to a few representatives only (Tvedte *et al.* 2019) or are totally absent like in the Ostracoda.

Of all extant non-marine ostracods, the Cyprididae (cyprids) are most speciose, comprising 42% of all known species (Meisch *et al.* 2019). They would thus be an obvious choice for genomic studies, also because in this ostracod family, mixed reproduction with sexual and asexual females and geographic parthenogenesis is very common (Horne *et al.* 1998). Asexual cyprids, however, are often polyploid (Symonova *et al.* 2018; Adolfsson *et al.* 2010), probably because of hybridization between males and asexual females through accidental mating (Schmit *et al.* 2013). Consequently, genome sizes are relatively large (Jeffery *et al.* 2017; Gregory 2020) up to 3.13 pg which equals more than 3 Gb. These features are likely to seriously complicate genomic assemblies and annotations in the absence of any genomic resources for ostracods, which is why we did not choose any asexual cypridid ostracods for this genome project. Instead, we have selected three other species of non-marine ostracods, one putative ancient asexual ostracod and two species with obligate sexual reproduction.

The ostracod family Darwinulidae is one of the two last remaining animal groups which are most likely true ancient asexuals (Heethoff *et al.* 2009; Schön *et al.* 2009b; Schwander 2016) and comprises about 35 morphospecies (Meisch *et al.* 2019). All darwinulids are brooders with valve dimorphisms between males and females that are detectable in the fossil record. Martens *et al.* (2003) showed that males have been absent in this family for at least 200 myr. One study reported a few males in a single darwinulid species (Smith *et al.* 2006) but proof of the functionality of these males for successful mating and meaningful genetic exchange could not be provided. Such (potential) atavistic males have also been reported in other putative ancient asexuals (Heethoff *et al.* 2009). The type species of the Darwinulidae, *Darwinula stevensoni*, has been asexual since c 20 myr (Straub 1952), occurs on all continents except Antarctica (Schön *et al.* 2012) and in a wide range of habitats (Schön *et al.* 2009b). *Darwinula stevensoni* is the best investigated darwinulid ostracod so far and has been the subject of ecological (Van Doninck *et al.* 2002, 2003 a & b; Van den Broecke *et al.* 2013) and molecular research using DNA sequence data from single genes (Schön *et al.* 1998; Schön *et al.* 2003; Martens *et al.* 2005; Schön *et al.* 2012). These studies revealed that *D. stevensoni* is most likely apomictic or functionally mitotic (following the definition of apomixis in animals as in Schön *et al.* 2009a). The species also has low mutation rates. as there appears to be no (Schön *et al.* 1998) or low (Schön and Martens 2003; Schön *et al.* 2009b) allelic divergence within individuals, and genetic differences between populations from different continents can be attributed to ancient vicariant processes (Schön *et al.* 2012). It has also been suggested that gene conversion is common in this species (Schön and Martens 1998, 2003). These results, however, were based on a limited number of genes and require further confirmation with genome-wide data. *Darwinula stevensoni* has a life cycle of 1 year in Belgium (Van Doninck *et al.* 2003b) and up to 4 years in more northern regions (McGregor 1969 in Northern America; Ranta 1979 in Finland), which is exceptionally long for a non-marine ostracod. It can survive a wide range of temperatures, salinities (Van Doninck *et al.* 2002) and oxygen concentrations (Rossi *et al.* 2002). The total genome size of *D. stevensoni* has been estimated as 0.86-0.93 pg with flow cytometry (Pacziesniak,

unpublished), approximating 900 Mb. There is no information on the ploidy level of *D. stvensoni*, except for the study by Tétart (1979) showing 22 dot-like chromosomes. Because of its putative ancient asexuality, no close sexual relatives of *D. stvensoni* are available for comparative, genomic analyses. We have chosen two fully sexual non-marine ostracod species from the Cytherideidae and the Notodromadidae with high population densities in Belgium as comparisons to the putative ancient asexual: *Cyprideis torosa* and *Notodromas monacha* respectively. *Cyprideis torosa* inhabits brackish waters and is the only extant species of this genus in Europe (Meisch 2000). It has been the subject of various biological and especially palaeontological and geochemical studies (see for example: Heip, 1976 a, b; De Deckker *et al.* 1999; Keyser 2005). Frogley and Whittaker (2016) suggested that *C. torosa* is at least of Pleistocene origin (c 2.5 myr) but might be older. There are only two molecular studies of this species based on single genes (Schön and Martens 2003; Schön *et al.* 2017). No information on the genome size or the karyotype of *C. torosa* is currently available.

The second sexual ostracod species analysed here, *Notodromas monacha*, occurs throughout the Northern hemisphere and is a non-marine ostracod with a most peculiar behaviour: it is partially hyponeustonic, hanging upside down attached to the water surface (Meisch 2000). The fossil record of *N. monacha* goes back to the Miocene (max 23 myr - Janz 1997), and its genome size is at 0.87pg (Jeffery *et al.* 2017; Gregory 2020) very similar to that of *D. stvensoni*. This species has not yet been the subject of any molecular studies.

Our aim here is to provide the first reference genome data of non-marine ostracods from three different species with varying reproductive modes: the putative ancient asexual *D. stvensoni* and the two obligate sexuals, *C. torosa* and *N. monacha*. We also generate transcriptomes of these species to facilitate genome annotations.

## MATERIAL AND METHODS

### Sample collection for genome and transcriptome sequencing

All three non-marine ostracod species were sampled in Belgian lakes where previous research had shown that these species occurred (Schön and Martens 2003; Merckx *et al.* 2018). Living ostracods were sampled using a hand net with a mesh size of 150 µm. The hand net was swept in between the vegetation and forcefully right above the surface of the sediment for collecting *Darwinula stvensoni* and *Cyprideis torosa*. *Notodromas monacha* was sampled by moving the net on the water surface. Non-marine ostracods were kept in habitat water. Their taxonomic identity was confirmed, and they were sorted alive under a binocular microscope as described by Martens and Horne (2016). Individual ostracods were picked with a pipette and transferred into sterilized EPA water in which they were maintained until DNA and RNA was extracted. More details on the origin of biological samples are provided in Table S2.

For generating reference genomes, DNA was extracted from a single female of each species using the QIAamp DNA Micro kit according to manufacturer's instructions. The extracted DNA from single females was amplified in two independent reactions using the SYNGIS TruePrime WGA kit and then pooled, to generate sufficient DNA for preparing different libraries. To generate transcriptomes for annotation of reference genomes, RNA was

extracted from 40 pooled individuals per species from the same collection batch. For this, individuals were frozen in liquid nitrogen and, after addition of Trizol (Life Technologies), mechanically crushed with beads (Sigmund Lindner). Next, chloroform and ethanol-treatment was applied to the homogenized tissue and the aqueous layer transferred to RNeasy MinElute Columns (Qiagen). Subsequent steps of RNA extraction were done following the RNeasy Mini Kit protocol, including DNase digestion. Finally, RNA was eluted into water and stored at -80°C. RNA quantity and quality were estimated with the NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent).

### Genome assembly

We prepared five genomic DNA libraries for each reference genome (three 2x125bp paired-end libraries with average insert sizes of 250-300, 550 and 700bp, and two mate-pair libraries with average insert sizes of 3000 and 5000bp; see Table S3 for more details) with the Illumina TruSeq DNA Library Prep Kit. Reads were generated with the Illumina HiSeq 3000 system for a total coverage between 351X and 386X (Table S3).

Reads were filtered with Trimmomatic v0.36 (Bolger *et al.* 2014) and NxTrim v0.4.1 (O'Connell *et al.* 2015). We employed non-standard methods for *de novo* genome assemblies owing to uneven coverage produced by PCR-based whole-genome amplification (Chen *et al.* 2013; Oyola *et al.* 2014). Filtered reads were normalized using BBDuk v36.59 (Bushnell 2014) and then assembled into contigs with SPAdes v3.10.1 (Bankevich *et al.* 2012). Scaffolding was performed using SSPACE v3.0 (Boetzer *et al.* 2012). Scaffolds identified as contaminants were filtered out using Blobtools v1.0 (Laetsch and Blaxter 2017). The completeness of genomes assemblies was assessed with BUSCO v3.0.2 (Seppey *et al.* 2019) against the *arthropoda\_odb9* lineage. More details of the assembly pipelines and the applied parameters can be found in supplementary materials (SM1).

### Protein coding gene annotation

Libraries were prepared using the Illumina TruSeq Stranded RNA, following the manufacturer's instructions. RNA reads were generated with the Illumina HiSeq 2500 system (Table S4). Reads were filtered with Trimmomatic v0.36. All trimmed reads were mapped against the genomes with STAR v2.5.3a (Dobin *et al.* 2013) and further assembled with Trinity v2.5.1 (Haas *et al.* 2013) under the "genome guided" mode to produce transcriptome assemblies.

The obtained transcriptomes and protein evidence were used to train and predict protein coding genes using MAKER v2.31.8 (Holt and Yandell 2011). Predicted protein coding genes were functionally annotated with Blast2GO v5.5.1 (Conesa *et al.* 2005; Götz *et al.* 2008) against the NCBI *non-redundant arthropods* protein database (v 2018-10).

More details of the annotation pipelines and the applied parameters can be found in supplementary materials (SM2).

## GenomeScope analyses

The whole genome amplification approach, which we used in the present study because of the small body size of individual ostracods, generated unequal read coverage of ostracod genomes and prevented us from directly estimating genome sizes and levels of heterozygosity from the assemblies. To overcome this problem, we re-sequenced two individual ostracods each of *D. stvensoni* and *N. monacha* without whole genome amplification, preparing libraries with the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina. Reads were filtered with Trimmomatic v0.36 and analyzed using GenomeScope v2.0 (Ranallo-Benavidez *et al.* 2020) to correctly estimate genome size and heterozygosity. More details on the analyses are provided in the supplementary material (SM3).

## RESULTS AND DISCUSSION

### First ostracod reference genomes and their features

We successfully produced the first *de novo* reference genomes of non-marine ostracods, namely of the three species *Darwinula stvensoni*, *Cyprideis torosa* and *Notodromas monacha* with different reproductive modes (see SM1 and Tables S3-S4 for more details on the assemblies). Given the small size of individual non-marine ostracods and the limited amount of soft tissue suitable for DNA and RNA extractions (Schön and Martens 2016), this was not a trivial task. We used a whole genome amplification approach (WGA), because the TruSeq DNA Nano library prep kit for Illumina sequencing or low input protocols for PacBio (Duncan *et al.* 2019) were not yet available when these assemblies were generated. We would not recommend WGA for future studies because this PCR-based method generated uneven coverage, and consequently, problems for applying routine genome assembly methods and estimates of genome size and heterozygosity. Despite these limitations, our approach produced genome assemblies that are useful for future research as will be outlined below.

When assessing the quality of the obtained ostracod *de novo* genome assemblies, the assembly of the putative ancient asexual, *D. stvensoni*, had the best contiguity, with the largest N50 although the total number of scaffolds was similar to *N. monacha* (Table 1, Table S6). The genome of the putative ancient asexual is furthermore the most complete as shown by its total BUSCO score of 96% and of 94% for complete single copy genes (Table 1). The quality of the genome from the obligate sexual ostracod *Cyprideis torosa* is the lowest of the three ostracod species as it has the highest number of scaffolds, and the lowest N50; it is also less complete with a total BUSCO score of 92% (Table S7) and of 87% for complete single copy genes (Table 1). All three species have similar numbers of predicted genes and transcripts (Table S7).

Ostracod genome sizes estimated with flow cytometry are somewhat larger than the estimates that we obtained here from GenomeScope analyses of re-sequenced individual ostracods. The haploid genome size of *D. stvensoni* was estimated at 420-455 Mb with flow cytometry (Pacziesniak, unpublished) while we estimated 362 Mb from sequence reads (Figure S1 A-B). Similarly, the size of the haploid genome of *N. monacha* is estimated at 425 Mb with flow cytometry (Jeffery *et al.* 2017; Gregory 2020), which is larger than the 385 Mb (Figure S1 C-D) that we obtained from sequence reads. It thus seems that some parts of each genome are

missing from our sequencing reads. Transposons and repeat-rich genomic regions can contribute to gaps in genomic assemblies (Peona *et al.* 2020). Some of these missing regions could also be GC-rich, a feature which is known to cause a sequencing-bias with Illumina technology (see for example Chen *et al.* 2013, Botero-Castro *et al.* 2017). Acquiring more complete genome assemblies will require the additional application of long read technologies to ostracods.

Genome-wide estimates of heterozygosity are especially interesting for asexual taxa because the absence of recombination is expected to cause accumulation of mutations, resulting in increasing allelic divergences within individuals (Birky 1996). Jaron *et al.* (2020) identified three factors driving intragenomic heterozygosity in asexuals: how the transition to parthenogenesis occurred, which cytological mechanism underlies parthenogenesis and how long asexual reproduction has been ongoing. Based on sequencing reads from individual ostracods, we estimate heterozygosity of the putative ancient asexual ostracod *D. stevensoni* to be 0.92-0.99% (Figure S1 A-B) and 1.32-1.43% for the sexual *N. monacha* (Figure S1 C-D). The genome-wide heterozygosity of *D. stevensoni* matches to some extent an earlier study on intra-individual divergence in three nuclear genes of *D. stevensoni* (Schön and Martens 2003). The finding of almost 1% heterozygosity in *D. stevensoni* is remarkable, given that all previous genome-wide estimates for asexual arthropods that did not evolve via hybridization revealed extremely low levels of heterozygosity (Jaron *et al.* 2020). Yet heterozygosity is clearly less than the estimates for parthenogenetic species with known hybrid origin (1.73-8.5%) or polyploidy (1.84%-33.21%) (Jaron *et al.* (2020), supporting the view that *D. stevensoni* is neither a hybrid nor a polyploid. Asexual reproduction in ostracods is thought to be apomictic (Chaplin *et al.* 1994), implying that observed heterozygosity levels are largely depend on the relative impact of heterozygosity losses from gene conversion and heterozygosity gains from new mutations. Given the apparent absence of sex and recombination for millions of years (Straub 1952), it is perhaps surprising that heterozygosity in this putative ancient asexual ostracod is not larger. This may suggest that genome-wide rates of gene conversion and mutation are comparable in this species.

### **Genome contiguity of ostracod assemblies as compared to other crustaceans**

We here compare the qualities of our ostracod genome assemblies to those of 19 other crustacean species (Table 1) published in the last four years. We only include studies with complete assemblies and sufficient information to assess assembly qualities. We assessed the contiguity of the three *de novo* ostracod genome assemblies by the number of scaffolds and their N50. On the one hand, both features are comparable to those of the copepod *Apocylops royi* (Jørgensen *et al.* 2019) and the amphipod *Parhyale hawaiiensis* (Kao *et al.* 2016) (Table 1) and better than for crustaceans with larger genomes such as the decapods *Cherax quadricarinatus* (Tan *et al.* 2019), *Palaeomon carinicauda* (Li *et al.* 2019), *Penaeus monodon* (Van Quyen *et al.* 2020), *Marsupenaeus japonicus* (Yuan *et al.* 2018) and *Procambarus virginalis* (Gutekunst *et al.* 2018; Table 1). On the other hand, genome assemblies of several other crustaceans have smaller scaffold numbers and higher N50 and thus better contiguities than the assemblies obtained here for non-marine ostracods. For the two notostracan *Lepidurus* species (Savojardo *et al.* 2018), this can probably be explained by their smaller genome sizes. For other crustaceans, genome assemblies or linkage maps have been available

beforehand which have considerably improved assembly qualities (Table 1) as in the examples of the cladocerans *Daphnia pulex* (Ye *et al.* 2017), *D. magna* (Lee *et al.* 2019) and the copepod *Tigriopus japonicus* (Jeong *et al.* 2020). No such genomic resources are currently available for ostracods. Finally, other studies of crustacean genomes with better assembly contiguities (the branchiopod *Eulimnadia texana* - Baldwin-Brown *et al.* 2018, and the decapod *Erichthya japonica sinensis* - Tang *et al.* 2020), the copepod *Tigriopus californicus* - Jeong *et al.* 2020 and the isopod *Armadillidium vulgare* - Chebbi *et al.* 2019) have used a combination of Illumina and long read technologies (Table 1). Long-read technologies such as PacBio used to require a relatively large amount of high-molecular weight DNA (Solares *et al.* 2018), which could not be obtained for ostracods with their very low yields of high molecular weight DNA from individual specimens and their small body sizes as compared to many other crustaceans (Schön and Martens 2016). We hope that low input protocols for PacBio (Duncan *et al.* 2019) and other long read technologies can be successfully applied to ostracods in the future, in which case the genome assemblies obtained here could form the basis for subsequent hybrid assemblies. Optimizing Oxford Nanopore Technology for non-marine ostracods has already commenced (Schön *et al.* in prep.).

### Genome annotations of ostracods and other crustaceans

Because our *de novo* ostracod genome assemblies are relatively complete (see BUSCO scores in Table 1), we will here also briefly compare some features of predicted protein coding genes with those of other crustaceans (Table S8). We have predicted 13,771 to 17,776 protein coding genes in the three non-marine ostracod genomes, (Tables S7 and S8) with the highest number for the sexual *C. torosa* and an intermediate estimate for the putative ancient asexual *D. stenosoni*. The number of annotated protein coding genes in non-marine ostracods is similar to estimates for various branchiopods and the copepods *Oithona nana*, *Tigriopus californicus* and *T. kingsejongensis* but lower than in most malacostracans (Table S8). Not all genome studies of crustaceans cited here contain information on other features of coding genes, such as the average size of genes, introns and exons (Table S7). Comparisons of these features are therefore limited and will not be further discussed here but we provide available data of these features for ostracods and other crustacean genomes for reference.

Gene annotation in general but especially in the crustaceans is challenging; this is for example illustrated by the much lower numbers of protein coding genes (18,440) which are predicted in the novel reference genome of the cladoceran *Daphnia pulex* by Ye *et al.* (2017) as compared to the first assembly of *D. pulex* with more than 30,000 predicted genes (Colbourne *et al.* 2011). Even more difficult is assigning gene functions to annotated crustacean genomes (Rotlantt *et al.* 2018). The novel data on predicted genes and transcripts from non-marine ostracods in the current study will significantly contribute to future genome annotations in crustaceans and other arthropods. The genes and transcripts predicted here can also provide the baseline for future gene expression studies of non-marine and marine ostracods.

## CONCLUSIONS

We have successfully obtained *de novo* genome assemblies for three species of non-marine ostracods with different reproductive modes. These represent the first quality reference genomes for ostracods. Given the paucity of genome assemblies from crustaceans as compared to insects or other arthropods, these assemblies are important tools to further develop ostracods as models for evolutionary and ecological research, also including marine species. Even if the *de novo* genome assemblies are somewhat fragmented and not yet at the chromosome-level, they have a high level of completeness and will thus facilitate future studies of ostracods. The reference genomes of this paper can also provide the first step towards a genomic assessment of the putative ancient asexual status of non-marine darwinulid ostracod species.

### DATA AVAILABILITY

Raw sequence reads have been deposited in NCBI's sequence read archive under the following bioprojects: PRJNA515625 (reference genomes, Table S3) and PRJNA631617 (RNA-seq for annotations and resequenced individuals, Table S4 and Table S5). Genome assemblies and annotations have been deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB38362 (Table S6 and Table S7). Codes for the analyses are available at: [https://github.com/AsexGenomeEvol/Ostracoda\\_genomes](https://github.com/AsexGenomeEvol/Ostracoda_genomes).

### ACKNOWLEDGEMENTS, FUNDING, CONFLICT OF INTEREST

This research was funded by Belgian Federal Science Policy (BR/314/PI/LATTECO) and a grant from the Swiss National Science Foundation (CRSII3\_160723). Marie Cours, Tijs Van Den Bergen and Jeroen Venderickx are acknowledged for technical support in sampling and sorting ostracod samples. We also thank Kristiaan Hoedemakers and Jeroen Venderickx for their assistance in finalizing the figure.

### LITERATURE CITED

- Adolfsson S., Y. Michalakis, D. Paczesniak, S.N.S. Bode, R.K. Butlin, et al., 2010 Evaluation of elevated ploidy and asexual reproduction as alternative explanations for geographic parthenogenesis in *Eucypris virens* ostracods. *Evolution* 64: 986–997. doi: 10.1111/j.1558-5646.2009.00872.x
- Baldwin-Brown, J.G., S.C. Weeks, and A.D. Long, 2018 A new standard for crustacean genomes: the highly contiguous, annotated genome assembly of the clam shrimp *Eulimnadia texana* reveals HOX gene order and identifies the sex chromosome. *Genome Biol. Evol.* 10: 143-156. doi: 10.1093/gbe/evx280
- Bankevich, A., S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, et al., 2012 Spades : A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19: 455-477. <https://doi.org/10.1089/cmb.2012.0021>
- Barreto, F. S., E.T. Watson, T.G. Lima, C.S. Willett, S. Edmands, et al., 2018 Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nat. Ecol. Evol.* 2: 1250-1257. <https://doi.org/10.1038/s41559-018-0588-1>
- Bell, G. 1982 *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. University of California Press: Berkeley, CA, USA.

- Beninde, J., M. Möst, and A. Meyer, 2020 Optimized and affordable high-throughput sequencing workflow for preserved and nonpreserved small zooplankton specimens. *Mol. Ecol. Res.* <https://doi.org/10.1111/1755-0998.13228>
- Birky C.W. Jr., 1996 Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* 144: 427-437.
- Boetzer, M., C. V. Henkel, H. J. Jansen, D. Butler, and W. Pirovano, 2011 Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578-579. <https://doi.org/10.1093/bioinformatics/btq683>
- Bolger, A.M., M. Lohse, and B. Usadel, 2014 Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Boschetti C., A. Carr, A. Crisp, I. Eyres, Y. Wang-Koh, E. et al., 2012 Biochemical diversification through foreign gene expression in bdelloid rotifers. *PLoS Genet.* 8: e1003035. <https://doi.org/10.1371/journal.pgen.1003035>
- Botero-Castro, F., E. Figuet, M.K. Tilak, B. Nabholz, and N. Galtier, 2017 Avian genomes revisited: Hidden genes uncovered and the rates versus traits paradox in birds. *Mol. Biol. Evol.* 34:3123-3131. doi: 10.1093/molbev/msx236
- Bushnell, B., 2014 BBMap: A fast, accurate, splice-aware aligner. United States.
- Butlin, R. K., and P. Menozzi, 2000. Open questions in evolutionary ecology: do ostracods have the answers?. In *Evolutionary Biology and Ecology of Ostracoda*, pp. 1-14. Springer, Dordrecht. <https://doi.org/10.1023/A:1003925705735>
- Butlin R. K., I. Schön, and K. Martens, 1998 Asexual reproduction in non-marine ostracods. *Heredity* 81: 473-480. <https://doi.org/10.1046/j.1365-2540.1998.00454.x>
- Chaplin, J.A., J.E. Havel, and P.D.N. Hebert, 1994 Sex and ostracods. *Trends Ecol. Evol.* 9: 435–439. [https://doi.org/10.1016/0169-5347\(94\)90127-9](https://doi.org/10.1016/0169-5347(94)90127-9)
- Chebbi, M.A., T. Becking, B. Moumen, I. Giraud, C. Gilbert, et al., 2019 The genome of *Armadillidium vulgare* (Crustacea, Isopoda) provides insights into sex chromosome evolution in the context of cytoplasmic sex determination. *Mol. Biol. Evol.* 36: 727-741. <https://doi.org/10.1093/molbev/msz010>
- Chen, Y.-C., T. Liu, C.-H. Yu, T.-Y. Chiang, and C.-C. Hwang, 2013 Effects of gc bias in next-generation-sequencing data on de novo genome assembly. *PLoS ONE* 8: e62856. <https://doi.org/10.1371/journal.pone.0062856>
- Colbourne, J. K., M.E. Pfrender, D. Gilbert, W.K. Thomas, A. Tucker, et al., 2011 The ecoresponsive genome of *Daphnia pulex*. *Science* 331: 555-561. 10.1126/science.1197761
- Conesa, A., S. Götz, J.M. García-Gómez, J. Terol, M. Talon, et al., 2005 Blast2GO : A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674-3676. <https://doi.org/10.1093/bioinformatics/bti610>
- Danchin, E.G., M.N. Rosso, P. Vieira, J. de Almeida-Engler, P.M. Coutinho, et al., 2010 Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. *Proc. Natl. Acad. Sci. USA* 107: 17651–17656. <https://doi.org/10.1073/pnas.1008486107>
- De Deckker, P., A.R. Chivas, and J.M.G. Shelley, 1999 Uptake of Mg and Sr in the euryhaline ostracod *Cyprideis* determined from in vitro experiments. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 148: 105–116. [https://doi.org/10.1016/S0031-0182\(98\)00178-3](https://doi.org/10.1016/S0031-0182(98)00178-3)

- Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, et al., 2013 STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15-21. <https://doi.org/10.1093/bioinformatics/bts635>
- Duncan, T., S.B. Kingan, C.C. Lambert, P. Baybayan, and J. Korlach, 2019 A low DNA input protocol for high-quality PacBio de novo genome assemblies. *J. Biomol. Techn. JBT* 30 (Suppl): S1-S2.
- Flot, J.-F., B. Hespeels, X. Li, B. Noel, I. Arkhipova et al., 2013 Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* 500: 453–457. <https://doi.org/10.1038/nature12326>
- Frogley, M.R., and J.E. Whittaker, 2016 On the origins of *Cyprideis torosa* (Jones, 1850) and a short biography of Professor T. R. Jones. *J. Micropalaeontol.* 36: 8-12. <https://doi.org/10.1144/jmpaleo2015-023>
- Gladyshev, E.A., M. Meselson, and I.R. Arkhipova, 2008 Massive horizontal gene transfer in bdelloid rotifers. *Science* 320: 1210–1213. [10.1126/science.1156407](https://doi.org/10.1126/science.1156407)
- Glémin, S., C.M. François, and N. Galtier, 2019 Genome Evolution in Outcrossing vs. Selfing vs. Asexual Species. In *Evolutionary Genomics* (pp. 331-369). Humana, New York, NY. [https://doi.org/10.1007/978-1-4939-9074-0\\_11](https://doi.org/10.1007/978-1-4939-9074-0_11)
- Götz, S., J.M. García-Gómez, J. Terol, T.D. Williams, et al., 2008 High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acid. Res.* 36: 3420-3435. <https://doi.org/10.1093/nar/gkn176>
- Gregory, T.R. Animal Genome Size Database. Available online: <http://genomesize.com>
- Gutekunst, J., R. Andriantsoa, C. Falckenhayn, K. Hanna, W. Stein, et al., 2018 Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nat. Ecol. Evol.* 2: 567. <https://doi.org/10.1038/s41559-018-0467-9>
- Haas, B.J., A. Papanicolaou, M. Yassour, M. Grabherr, M., P.D. Blood, et al., 2013 *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Prot.* 8: 1494-1512. <https://doi.org/10.1038/nprot.2013.084>
- Hecox-Lea, B.J., and D.B. Mark Welch, 2018 Evolutionary diversity and novelty of DNA repair genes in asexual Bdelloid rotifers. *BMC Evol. Biol.* 18: 177. doi:10.1186/s12862-018-1288-9.
- Heethoff, M., R.A. Norton, S. Scheu, and M. Maraun, 2009 Parthenogenesis in oribatid mites (Acari, Oribatida): evolution without sex. In Schön, I., K. Martens, and P. Van Dijk (eds) *Lost sex. The evolutionary biology of parthenogenesis*. Springer Academic Publishers, Dordrecht, The Netherlands, pp. 241-258. [https://doi.org/10.1007/978-90-481-2770-2\\_12](https://doi.org/10.1007/978-90-481-2770-2_12)
- Heip, C., 1976a The life-cycle of *Cyprideis torosa* (Crustacea, Ostracoda). *Oecologia* 24: 229–245. <https://doi.org/10.1007/BF00345475>
- Heip, C., 1976b The spatial pattern of *Cyprideis torosa* (Jones, 1850) (Crustacea: Ostracoda). *J. Mar. Biol. Assoc. U. K.*, 56: 179-189. <https://doi.org/10.1017/S002531540002052X>
- Holt, C., and M. Yandell, 2011 MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12: 491. <https://doi.org/10.1186/1471-2105-12-491>
- Horne, D. J., A. Baltanas, and G. Paris, 1998 Geographical distribution of reproductive modes in living non-marine ostracods. In: Martens, K. (ed) *Sex and Parthenogenesis*.

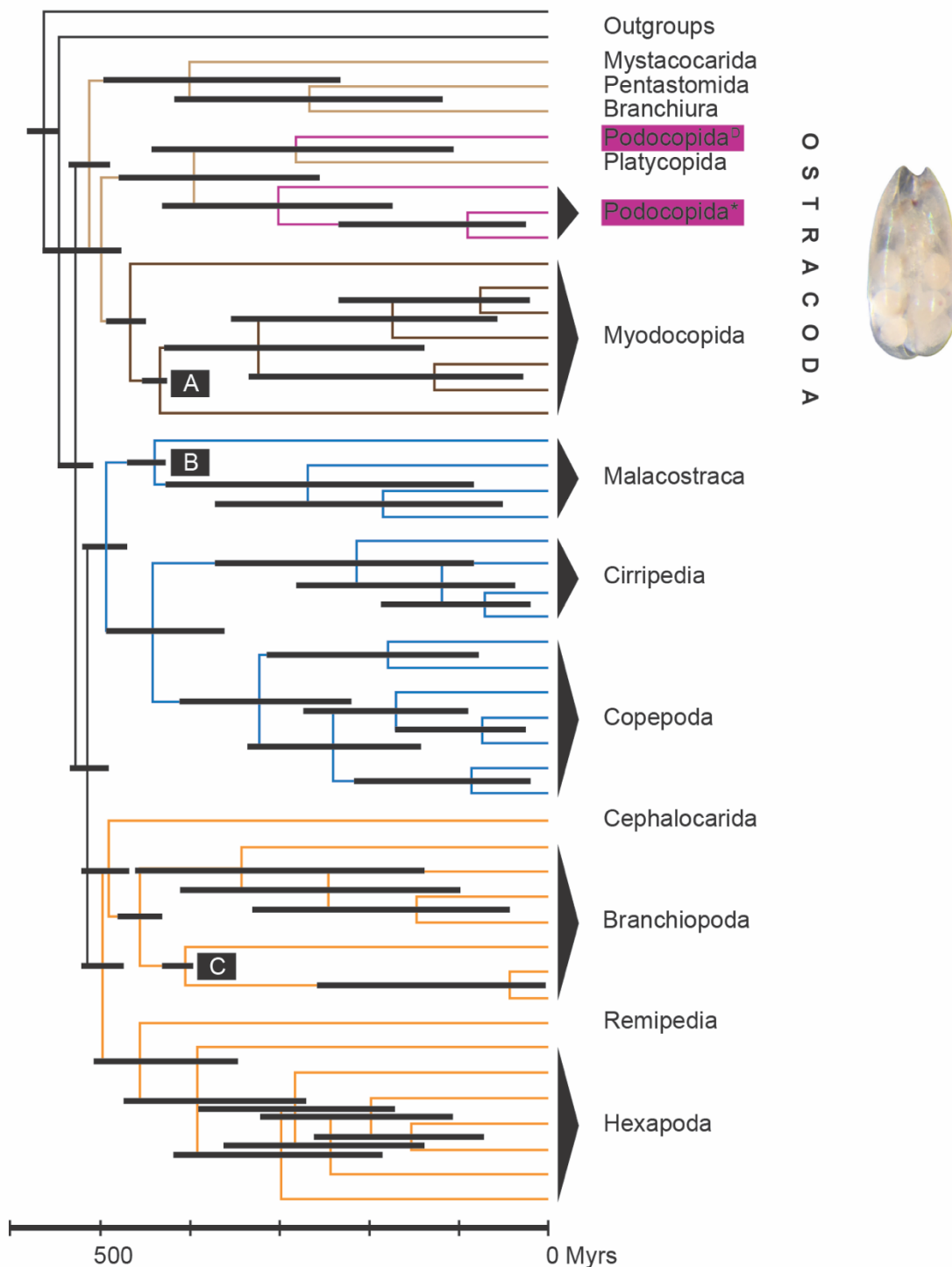
- Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods*. Backhuys Publishers, Leiden, pp. 77–99.
- Janz, H., 1997 Die Ostrakoden der kleini-Schichten des miozänen Kratersees von Steinheim am Albuch (Süddeutschland). *Stutt. Beitr. Naturkd. B* 251: 1–101.
- Jaron, K.S., J. Bast, R.W. Novell, T.R. Ranallo-Benavidez, M. Robinson-Rechavi, et al., 2020 Genomic features of asexual animals. *J. Heredity*: esaa031. <https://doi.org/10.1093/jhered/esaa031>
- Jeffery, N.W., E.A. Ellis, T.H. Oakley, and T.R. Gregory, 2017 The genome sizes of ostracod crustaceans correlate with body size and evolutionary history, but not environment. *J. Hered.* 108: 701–706. [10.1093/jhered/esx055](https://doi.org/10.1093/jhered/esx055)
- Jeong, C.B., B.Y. Lee, B.S. Choi, M.S. Kim, J.C. Park, et al., 2020 The genome of the harpacticoid copepod *Tigriopus japonicus*: Potential for its use in marine molecular ecotoxicology. *Aquat. Toxicol.* 222: 105462. <https://doi.org/10.1016/j.aquatox.2020.105462>
- Jørgensen, T.S., B.L.H. Nielsen, B. Petersen, P.D. Browne, B.W. Hansen, et al., 2019 The whole genome sequence and mRNA transcriptome of the tropical cyclopoid copepod *Apocyclops royi*. *G3: Genes Genom. Genet.* 9: 1295–1302. <https://doi.org/10.1534/g3.119.400085>
- Kao, D., A.G. Lai, E. Stamatakis, S. Rosic, et al., 2016 The genome of the crustacean *Parhyale hawaiiensis*, a model for animal development, regeneration, immunity and lignocellulose digestion. *elife* 5: e20062. [10.7554/eLife.20062](https://doi.org/10.7554/eLife.20062)
- Keyser, D., 2005 Histological peculiarities of the nodding process in *Cyprideis torosa* (Jones) (Crustacea, Ostracoda). *Hydrobiologia* 538: 95–106. <https://doi.org/10.1007/s10750-004-4940-x>
- Laetsch, D. R., and M.L. Blaxter, 2017 BlobTools: Interrogation of genome assemblies. *F1000Research* 6: 1287. <https://doi.org/10.12688/f1000research.12232.1>
- Lee, B.Y., B.S. Choi, M.S. Kim, J.C. Park, C.B. Jeong, et al., 2019 The genome of the freshwater water flea *Daphnia magna*: A potential use for freshwater molecular ecotoxicology. *Aquat. Toxicol.* 210: 69–84. <https://doi.org/10.1016/j.aquatox.2019.02.009>
- Li, J., J. Lv, P. Liu, P. Chen, J. Wang, et al., 2019 Genome survey and high-resolution backcross genetic linkage map construction of the ridgetail white prawn *Exopalaemon carinicauda* applications to QTL mapping of growth traits. *BMC Genomics* 20: 598. <https://doi.org/10.1186/s12864-019-5981-x>
- Maddocks, R., 1982 Part 4. Ostracoda. In Abele, L.G. (ed) *The Biology of Crustacea. Vol 1: Systematics, the Fossil Record and Biogeography*. Academ. Press, pp. 221–239.
- Martens, K., and D.J. Horne, 2016 Collecting and processing living non-marine ostracods. *J. Crust. Biol.* 36: 849–854. <https://doi.org/10.1163/1937240X-00002488>
- Martens, K., D.J. Horne, and H.I. Griffiths, 1998 Age and diversity of non-marine ostracods. In: Martens, K. (ed) *Sex and Parthenogenesis. Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods*. Backhuys Publishers, Leiden, pp. 37–55.
- Martens, K., G. Rossetti, R.K. Butlin, and I. Schön, 2005 Molecular and morphological phylogeny of the ancient asexual Darwinulidae (Crustacea, Ostracoda). *Hydrobiologia* 538: 153–165. <https://doi.org/10.1007/PL00021867>
- Martens, K., G. Rossetti, and D. Horne, 2003 How ancient are ancient asexuals? *Proc. R. Soc. Lond. Ser. B* 270: 723–729. <https://doi.org/10.1098/rspb.2002.2270>

- McGregor, D. L., 1969 The reproductive potential, life history and parasitism of the freshwater Ostracods *Darwinula stevensoni* (Brady and Robertson). In: Neale, J.W. (ed), *The Taxonomy, Morphology and Ecology of Recent Ostracoda*. Oliver and Boyd, Edinburgh, pp. 194–221.
- Meisch, C., 2000 *Freshwater Ostracoda of Western and Central Europe*. Springer: Heidelberg, Germany.
- Meisch, C., R.J. Smith, and K. Martens, 2019 A subjective global checklist of the extant non-marine Ostracoda (Crustacea). *Eur. J. Tax.* 25: 492. <https://doi.org/10.5852/ejt.2019.492>
- Merckx T., Souffreau C., Kaiser A., Baardsen L.F., et al., 2018 Body size shifts in aquatic and terrestrial urban communities. *Nature* 558: 113–116. <https://doi.org/10.1038/s41586-018-0140-0>
- Neiman, M., P. G. Meermans, T. Schwander, and S. Meermans, 2018 Sex in the wild: How and why field-based studies contribute to solving the problem of sex. *Evolution* 72: 1194–1203. <https://doi.org/10.1111/evo.13485>
- O’Connell, J., O. Schulz-Trieglaff, E. Carlson, M.M. Hims, N.A. Gormley, et al., 2015 NxTrim: Optimized trimming of Illumina mate pair reads. *Bioinformatics* 31: 2035–2037. <https://doi.org/10.1093/bioinformatics/btv057>
- Oakley, T.H., and C.W. Cunningham, 2002 Molecular phylogenetic evidence for the independent evolutionary origin of an arthropod compound eye. *Proc. Natl. Acad. Sci. U. S. A.* 99: 1426–1430. <https://doi.org/10.1073/pnas.032483599>
- Oakley, T.H., J.M. Wolfe, A.R. Lindgren, and A.K. Zaharoff, 2013 Phylotranscriptomics to bring the understudied into the fold: monophyletic Ostracoda, fossil placement, and pancrustacean phylogeny. *Mol. Biol. Evol.* 30: 215–233. <https://doi.org/10.1093/molbev/mss216>
- Ogoh, K., and Y. Ohmiya, 2004 Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorfii* (Crustacea, Ostracoda) with duplicate control regions. *Gene* 327: 131–139. [10.1016/j.gene.2003.11.011](https://doi.org/10.1016/j.gene.2003.11.011)
- Omilian, A.R., M.E. Cristescu, J.L. Dudycha, and M. Lynch, 2006 Ameiotic recombination in asexual lineages of *Daphnia*. *Proc. Natl. Acad. Sci.* 103: 18638–18643. <https://doi.org/10.1073/pnas.0606435103>
- Otto, S.P., 2020 Selective interference and the evolution of sex. *J. Hered.*: esaa026. <https://doi.org/10.1093/jhered/esaa026>
- Otto, S.P., and T. Lenormand, 2002 Resolving the paradox of sex and recombination. *Nature Rev. Genet.* 3: 252–261. <https://doi.org/10.1038/nrg761>
- Oyola, S. O., M. Manske, S. Campino, A. Claessens, et al., 2014 Optimized whole-genome amplification strategy for extremely AT-biased template. *DNA Res.* 21: 661–671. [10.1093/dnares/dsu028](https://doi.org/10.1093/dnares/dsu028)
- Paganini, J., A. Campan-Fournier, M. Da Rocha, P. Gouret, P. Pontarotti et al., 2012 Contribution of lateral gene transfers to the genome composition and parasitic ability of root-knot nematodes. *PLoS One* 7: e50875. <https://doi.org/10.1371/journal.pone.0050875>
- Peona, V., M.P. Blom, L. Xu, R. Burri, S. Sullivan, et al., 2020 Identifying the causes and consequences of assembly gaps using a multiplatform genome assembly of a bird-of-paradise. *Mol. Ecol. Res.* online published. doi:10.1111/1755-0998.13252

- Ranallo-Benavidez, T.R., K.S. Jaron, and M.C. Schatz, 2020 GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Commun.* 11: 1-10. <https://doi.org/10.1038/s41467-020-14998-3>
- Ranta, E., 1979 Population biology of *Darwinula stevensoni* (Crustacea, Ostracoda) in an oligotrophic lake. *Ann. Zool. Fenn.* 16: 28–35. <https://www.jstor.org/stable/2373370>
- Rossi, V., E.B.A. Todeschi, A. Gandolfi, M. Invidia, and P. Menozzi, 2002 Hypoxia and starvation tolerance in individuals from a riverine and a lacustrine population of *Darwinula stevensoni* (Crustacea: Ostracoda). *Arch. Hydrobiol.* 154: 151-171. [10.1127/archiv-hydrobiol/154/2002/151](https://doi.org/10.1127/archiv-hydrobiol/154/2002/151)
- Rotllant, G., F. Palero, P.B. Mather, H.D. Bracken-Grissom, and M.B. Santos, 2018 Preface: Recent advances in crustacean genomics. *Hydrobiologia* 825: 1-4. <https://doi.org/10.1007/s10750-018-3773-y>
- Roughgarden, J., 1991 The evolution of sex. *Am. Nat.* 138:934–953. <https://www.jstor.org/stable/2462442>
- Savojardo, C., A. Luchetti, P.L. Martelli, R. Casadio, and B. Mantovani, 2019 Draft genomes and genomic divergence of two *Lepidurus* tadpole shrimp species (Crustacea, Branchiopoda, Notostraca). *Mol. Ecol. Resour.* 19: 235-244. <https://doi.org/10.1111/1755-0998.12952>
- Schmit, O., L. Fukova, J. Vandekerckhove, Y. Michalakis, R. Matzke-Karasz et al., 2013 Mate recognition as a reproductive barrier in sexual and parthenogenetic *Eucypris virens* (Crustacea, Ostracoda). *Anim. Beh.* 85: 977-985. <https://doi.org/10.1016/j.anbehav.2013.02.021>
- Schön, I., R. K. Butlin, H. I. Griffiths, and K. Martens, 1998 Slow molecular evolution in an ancient asexual ostracod. *Proc. R. Soc. Lond. B* 265: 235-242. <https://doi.org/10.1098/rspb.1998.0287>
- Schön, I., S. Halse, and K. Martens, 2017 *Cyprideis* (Crustacea, Ostracoda) in Australia. *J. Micropalaeontol.* 36: 31-37. <https://doi.org/10.1144/jmpaleo2016-032>
- Schön, I., and K. Martens, 1998 Opinion: DNA repair in an ancient asexual - a new solution for an old problem? *J. Nat. Hist.* 32: 943-948. <https://doi.org/10.1080/00222939800770481>
- Schön, I., and K. Martens, 2003 No slave to sex. *Proc. R. Soc. Lond. B* 270: 827-833. <https://doi.org/10.1098/rspb.2002.2314>
- Schön, I., and K. Martens, 2016 Ostracod (Ostracoda, Crustacea) genomics - Promises and challenges. *Mar. Genomics.* 29: 19-25. doi: 10.1016/j.margen.2016.03.008
- Schön, I., K. Martens, and P. Van Dijk, 2009a (eds) *Lost sex. The evolutionary biology of parthenogenesis* Springer Academic Publishers, Dordrecht, The Netherlands.
- Schön, I., K. Martens, K. Van Doninck, and R.K. Butlin, 2003. Evolution in the slow lane: molecular rates of evolution in sexual and asexual ostracods (Crustacea: Ostracoda). *Biol. J. Linn. Soc.* 79: 93-100. <https://doi.org/10.1046/j.1095-8312.2003.00186.x>
- Schön, I., R. L. Pinto, S. Halse, A. J. Smith, K. Martens, and C.W. Birky, 2012 Cryptic species in putative ancient asexual Darwinulids (Crustacea, Ostracoda). *PLoS One* 7: e39844. doi: 10.1371/journal.pone.0039844.
- Schön, I., G. Rossetti, and K. Martens, 2009b Ancient asexual darwinulids: ancient scandals or scandalous gossip? In: Schön, I., K. Martens, and P. Van Dijk (eds) *Lost sex. The evolutionary biology of parthenogenesis*. Springer Academic Publishers, Dordrecht, The Netherlands, pp. 217-240. [https://doi.org/10.1007/978-90-481-2770-2\\_11](https://doi.org/10.1007/978-90-481-2770-2_11)

- Schwander, T., 2016 The end of an ancient asexual scandal. *Curr. Biol.* 26: R233-235. <https://doi.org/10.1016/j.cub.2016.01.034>
- Seppey, M., Manni, M., and E. M. Zdobnov, 2019 Busco: Assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962: 227-245. [https://doi.org/10.1007/978-1-4939-9173-0\\_14](https://doi.org/10.1007/978-1-4939-9173-0_14)
- Smith, R., T. Kamiya, and D.J. Horne, 2006 Living males of the 'ancient asexual' Darwinulidae (Ostracoda: Crustacea). *Proc. R. Soc. Lond. B* 273: 1569-1578. <https://doi.org/10.1098/rspb.2005.3452>
- Solares, E.A., M. Chakraborty, D.E. Miller, S. Kalsow, et al., 2018 Rapid low-cost assembly of the *Drosophila melanogaster* reference genome using low-coverage, long-read sequencing. *G3: Genes Genom. Genet.* 8: 3143-3154. 10.1534/g3.118.200162
- Straub, E.B., 1952 Mikropaläontologische Untersuchungen im Tertiär zwischen Ehingen und Ulm a.d. Donau. *Geol. Jb.* 66: 433-523.
- Symonová, R., I. Vrbová, D.K. Lamatsch, J. Paar, R. Matzke-Karasz, et al., 2018 Karyotype variability and inter-population genomic differences in freshwater ostracods (Crustacea) showing geographical parthenogenesis. *Genes* 9: 150. doi: 10.3390/genes9030150.
- Tan, M.H., H.M. Gan, Y.P. Lee, F. Grandjean, L.J. Croft, et al., 2020 A giant genome for a giant crayfish (*Cherax quadricarinatus*) with insights into cox1 pseudogenes in decapod genomes. *Front. Genet.* 11: 201. <https://doi.org/10.3389/fgene.2020.00201>
- Tang, B., Z. Wang, Q. Liu, H. Zhang, S. Jiang, et al., 2020 High-quality genome assembly of *Eriocheir japonica sinensis* reveals its unique genome evolution. *Front. Genet.* 10: 1340. 10.3389/fgene.2019.01340
- Tetart, J., 1979 Les garnitures chromosomiques des ostracodes d'eau douce. *Trav. Lab. Hydrobiol.* 69-70:113-140. (In French)
- Tvedte, E.S., J.M. Jr Logsdon, and A.A. Forbes, 2019 Sex loss in insects: causes of asexuality and consequences for genomes. *Curr. Opin. Insect Sci.* 31: 77-83. 10.1016/j.cois.2018.11.007.
- Van den Broecke, L., J. Vanfleteren, K. Martens, and I. Schön, 2013 Hurdles in investigating UVB damage in the putative ancient asexual *Darwinula stevensoni* (Ostracoda, Crustacea). *B. J. Zool.* 143: 106-118.
- Van Doninck, K., I. Schön, L. De Bruyn, and K. Martens, 2002 A general purpose genotype in an ancient asexual. *Oecologia* 132: 205-212. <https://doi.org/10.1007/s00442-002-0939-z>
- Van Doninck, K., I. Schön, F. Maes, L. De Bruyn, and K. Martens, 2003a Ecological strategies in the ancient asexual animal group Darwinulidae. *Freshwater Biol.* 48: 1285-1294. <https://doi.org/10.1046/j.1365-2427.2003.01078.x>
- Van Doninck, K., I. Schön, K. Martens, and B. Godderis, 2003b The life cycle of the ancient asexual ostracod *Darwinula stevensoni* (Brady & Robertson, 1870) (Crustacea, Ostracoda) in a temperate pond. *Hydrobiologia* 500: 331-340. <https://doi.org/10.1023/A:1024656920904>
- Van Quyen, D., H. M. Gan, Y.P. Lee, D.D. Nguyen, et al., 2020 Improved genomic resources for the black tiger prawn (*Penaeus monodon*). *Marine Genom.* 52: 100751. <https://doi.org/10.1016/j.margen.2020.100751>

- Wang, X., Q. Xu, J. Xiao, X. Miao, P. Liu, and Z. Wang, 2019 First record of the complete mitochondrial genome of *Cypridina dentata* (Myodocopida: Cypridinidae). Mitochondrial DNA B Resour 4: 1607-1608. <https://doi.org/10.1080/23802359.2019.1604097>
- Wolfe, J.M., A.C. Daley, D.A. Legg, and G.D. Edgecombe, 2016 Fossil calibrations for the arthropod Tree of Life. Earth-Sci. Rev. 160: 43-110. <https://doi.org/10.1016/j.earscirev.2016.06.008>
- Wolfe, J.M., and T.A. Hegna, 2014 Testing the phylogenetic position of Cambrian pancrustacean larval fossils by coding ontogenetic stages. Cladistics 30: 366-399. <https://doi.org/10.1111/cla.12051>
- Ye, Z., S. Xu, K. Spitze, J. Asselman, X. Jiang, et al., 2017 A new reference genome assembly for the microcrustacean *Daphnia pulex*. G3: Genes Genom. Genet. 7: 1405-1416. <https://doi.org/10.1534/g3.116.038638>
- Yuan, J., X. Zhang, C. Liu, Y. Yu, J. Wei, et al., 2018 Genomic resources and comparative analyses of two economical penaeid shrimp species, *Marsupenaeus japonicus* and *Penaeus monodon*. Mar. Genomics 39: 22-25. <https://doi.org/10.1016/j.margen.2017.12.006>



**Figure 1: The phylogenetic position of the Ostracoda among the pancrustaceans and their age estimated from fossil and molecular data.**

Modified from Oakley et al. (2013). Different pancrustaceans are indicated by branches in different colours. The Ostracoda include the Podocopida, Platycopida and Myodocopida. Here, three representatives of the Platycopida (indicated in purple) have been sequenced. The

phylogenetic clade to which *Darwinula stevensoni* belongs, is indicated by <sup>D</sup>, the clade to which *Cyprideis torosa* and *Notodromas monacha* belong, is indicated by \*. Black horizontal bars represent the range of age estimates in myr from Bayesian analyses by Oakley et al. (2013). The letters A-C in the black boxes indicated fossils that were used for calibrations of age estimates.

**Table 1: Quality features of published crustacean genomic assemblies of the last four years and of the current study.**

Assembly size is provided in million base pairs, scaffold N50 in kilo base pairs and BUSCO scores in %. Letters behind BUSCO scores indicate the % of complete single copy genes (C) or % of single and fragmented single copy genes (C+F), respectively. Where BUSCO scores lack brackets, no further information on completeness of single copy genes was provided. \$ anchoring of scaffolds in existing genome assembly; L = linkage map available; \* long read technology; # BAC library available. n. i. = no information available.

Class	Order	Species	Size	# scaffolds	N50	BUSCO	Reference
Branchiopoda	Diplostraca	<i>Daphnia pulex</i> <sup>SL</sup>	156	1,822	1,661	96	Ye <i>et al.</i> 2017
Branchiopoda	Diplostraca	<i>D. magna</i> <sup>SL</sup>	130	4,193	10,124	96.7 (C)	Lee <i>et al.</i> 2019
Branchiopoda	Notostraca	<i>Lepidurus arcticus</i>	73	7,167	116	98.4 (C)	Savojardo <i>et al.</i> 2018
Branchiopoda	Notostraca	<i>L. apus lubbocki</i>	90	20,738	402	97.8 (C)	Savojardo <i>et al.</i> 2018
Branchiopoda	Spinicaudata	<i>Eulimnadia texana</i> *	120	112	18,000	n. i.	Baldwin-Brown <i>et al.</i> 2017
Copepoda	Cyclopoida	<i>Apocyclops royi</i>	258	97,072	n. i.	50 (C)	Jørgensen <i>et al.</i> 2019
Copepoda	Cyclopoida	<i>Oithona nana</i>	85	4,626	401	n. i.	Madoui <i>et al.</i> 2017
Copepoda	Harpaticoida	<i>Tigriopus californicus</i> *	190	459	298	94.5 (C)	Barreto <i>et al.</i> 2018
Copepoda	Harpaticoida	<i>T. japonicus</i> <sup>\$</sup>	197	339	10,650	96 (C)	Jeong <i>et al.</i> 2020
Copepoda	Harpaticoida	<i>T. kingsejongensis</i>	295	270,823	159	61.1 (C)	Kang <i>et al.</i> 2017
<b>Ostracoda</b>	<b>Podocopida</b>	<b><i>Cyprideis torosa</i></b>	<b>335</b>	<b>132,611</b>	<b>19</b>	<b>86.6 (C)</b> <b>91.9 (C+F)</b>	<b>Current study</b>
<b>Ostracoda</b>	<b>Podocopida</b>	<b><i>Darwinula stevensoni</i></b>	<b>382</b>	<b>62,118</b>	<b>56</b>	<b>93.7 (C)</b> <b>95.8 (C+F)</b>	<b>Current study</b>
<b>Ostracoda</b>	<b>Podocopida</b>	<b><i>Notodromas monacha</i></b>	<b>377</b>	<b>62,251</b>	<b>42</b>	<b>92.7 (C)</b> <b>94.4 (C +F)</b>	<b>Current study</b>
Malacostraca	Amphipoda	<i>Parhyale hawaiiensis</i> <sup>#L</sup>	4,024	100,000	69	n. i.	Kao <i>et al.</i> 2016
Malacostraca	Isopoda	<i>Armadillidium vulgare</i> *	1,725	43,451	51	87.9 (C)	Chebbi <i>et al.</i> 2019
Malacostraca	Decapoda	<i>Cherax quadricarinatus</i> *	3,237	508,682	33	81.3 (C)	Tan <i>et al.</i> 2020
Malacostraca	Decapoda	<i>Eriocheir japonica sinensis</i> *	1,270	1,368	3,185	92.7 (C)	Tang <i>et al.</i> 2020
Malacostraca	Decapoda	<i>Palaeomon carinicauda</i> <sup>L</sup>	9,185	28,089,718	586	n. i.	Li <i>et al.</i> 2019
Malacostraca	Decapoda	<i>Penaeus monodon</i> *	1,600	1,211,364	2	96.8 (C+F)	Van Quyen <i>et al.</i> 2020
Malacostraca	Decapoda	<i>Litopenaeus vannamei</i> <sup>*L</sup>	1,664	4,682	606	95	Zhang <i>et al.</i> 2019
Malacostraca	Decapoda	<i>Marsupenaeus japonicus</i>	924	37,192,281	1	97	Yuan <i>et al.</i> 2018
Malacostraca	Decapoda	<i>Procambarus virginalis</i>	3,300	3,752,011	39	n. i.	Gutekunst <i>et al.</i> 2018

## **Supplementary material:**

### **SM1: Details of genome assembly pipelines and parameters**

The paired-end reads were filtered using Trimmomatic v0.36 (Bolger *et al.* 2014). Adapters were removed using adapters provided by Trimmomatic. Leading and trailing bases below quality 3 were removed. Reads were scanned using a 4-base sliding window and trimmed when the average quality dropped below 20. Reads were discarded if the size dropped below 60 bp. (parameters: PE ILLUMINACLIP:TruSeq-PE.fa:2:30:10:4 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 AVGQUAL:20 MINLEN:60).

The mate-pair reads were filtered using NxTrim v0.4.1 (O'Connell *et al.* 2015) (parameters: --preserve-mp --separate). As a result, mate-pair reads flagged as MP and UNKNOWN were concatenated following the authors' suggestion. Finally, the concatenated mate-pair reads were trimmed using Trimmomatic with the same parameters described previously.

We employed non-standard methods for genome assemblies due to uneven coverage produced by PCR-based whole-genome amplification (Chen *et al.* 2013; Oyola *et al.* 2014). Only the overlapped libraries were used for contig assemblies. Filtered paired-end reads were merged into single reads by overlap detection using BBDiff v36.59 (Bushnell 2014) and the module BBMerge (parameters: minoverlap=15 mismatches=0 ecct strict). The merged reads were then concatenated with the rest of the single reads and normalization was performed using the module BBNorm with a target average depth of 65x.

Normalized data were *de novo* assembled using SPAdes v3.10.1 (Bankevich *et al.* 2012) (parameters: --careful -k 21,33,55,77,99,111,127). Contigs were then scaffolded with SSPACE v3.0 (Boetzer *et al.* 2012) using the remaining libraries and gap-filled with GapCloser v1.12-r6, a module of the SOAP package (Luo *et al.* 2012) with default parameters. Then, the genome assemblies were decontaminated using BlobTools v1.0 (Laetsch & Blaxter, 2017) under the taxrule "bestsumorder". Hit files were generated after a blastn v2.7.1+ against the NCBI nt database (v 2016-06), searching for hits with an e-value below 1e-25 (parameters: -max\_target\_seqs 10 -max\_hsps 1 -evalue 1e-25). Scaffolds without hits to metazoans were filtered out from the assemblies.

The completeness of genomes assemblies was assessed with BUSCO v3.0.2 (Seppey *et al.* 2019) against the *arthropoda\_odb9* lineage and the --long option.

### **SM2: Annotation of protein coding genes**

Raw paired-end RNA-sequence reads were trimmed according to quality with Trimmomatic v0.36 (Bolger *et al.* 2014) using the maximum information approach with tolerant parameters.

Any reads with a sequence length of < 80 bp after trimming were discarded (parameters: adapters.fa:2:30:12:1:true LEADING:3 TRAILING:3 MAXINFO:40:0.4 MINLEN:80). All trimmed RNA-sequence reads were then mapped against the genomes using STAR v2.5.3a (Dobin *et al.* 2013) under the "2-pass mapping" mode and default parameters. Then, the STAR outputs were used to produce transcriptome assemblies using Trinity v2.5.1 (Haas *et al.* 2013) "genome guided" mode (parameters: --genome\_guided\_max\_intron 100000 --SS\_lib\_type RF --jaccard\_clip). Finally, the transcriptome assemblies were filtered following Trinity developers recommendations

(<https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-FAQ>):

Briefly, filtered RNA-seq reads were mapped back against the transcriptomes using Kallisto v0.43.1 (Bray *et al.* 2016) with options --bias and --rf-stranded then transcripts with at least 1 TPM in any samples were retained.

Protein coding genes were predicted using MAKER v2.31.8 (Holt & Yandell 2011) in a 2-iterative way described in Campbell *et al.* (2014) with minor modifications following author recommendations. Only scaffolds above 500bp were annotated. Prior to gene prediction, MAKER used RepeatMasker v4.0.7 (Tarailo-Graovac & Chen 2009) for masking repetitive regions. For the first iteration, genes were predicted using Augustus v3.2.3 (Stanke *et al.* 2006) trained with the BUSCO v3.0.2 (Seppey *et al.* 2019) results (SM1). A combination of UniProtKB/Swiss-Prot (release 2018\_01) and the BUSCO *arthropoda\_odb9* proteome were used as protein evidence. The Trinity assembled mRNA-seq reads (described above) were used as transcript evidence. The resulting gene models were then used to retrain Augustus as well as SNAP v2013.11.29 (Korf 2004) and a second iteration was performed. Subsequently, predicted protein coding genes were functionally annotated using Blast2GO v5.5.1 (Conesa *et al.* 2005; Götz *et al.* 2008) with default parameters against the NCBI *non-redundant arthropods* protein database (v 2018-10).

### **SM3: Estimation of genome size and heterozygosity by genome profiling analysis**

Raw paired-end resequencing reads were trimmed using the same strategy as described in SM2. K-mer frequencies were computed using KMC v3.1.1 (Kokot *et al.* 2017). Genome sizes and heterozygosities were estimated with GenomeScope v2.0 (Ranallo-Benavidez *et al.* 2020) using parameters recommended by the authors.

**Table S1: Overview of RNA sequencing reads from ostracods in GenBank.**

No full assemblies or annotations are available from these studies. # of genes = total number of orthologous genes used for phylogenetic or gene expression studies, respectively.

Order	Species	GenBank accession number	Sequencing technique	# of genes	Study purpose	Reference
Mydocopida	<i>Vargula tsujii</i>	SRX532406	454	261	Phylogeny of pancrustaceans	Oakley <i>et al.</i> 2012
Mydocopida	<i>Euphilomedes morini</i>	SRX1097309	454	261	Phylogeny of pancrustaceans	Oakley <i>et al.</i> 2012
Mydocopida	<i>Alternochelata lizardensis</i>	PRJNA436031	Illumina		1KITE project	unpublished
Mydocopida	<i>Vargula hilgendorfii</i>	SRX884494	Illumina		1KITE project	unpublished
Mydocopida	<i>Vargula hilgendorfii</i>	DRX059606	Illumina			unpublished
Mydocopida	<i>Eusarsiella sp.</i>	SRX2085852	Illumina	244	Phylogeny of pancrustaceans	Lozano-Fernandez <i>et al.</i> 2019
Mydocopida	<i>Cylindroleberidinae</i>	SRX2085850	Illumina	244	Phylogeny of pancrustaceans	Lozano-Fernandez <i>et al.</i> 2019
Mydocopida	<i>Rutiderma sp.</i>	SRX2458829	Illumina	277	Phylogeny of pancrustaceans	Schwenter <i>et al.</i> 2018
Halocyprida	<i>Conchoecia obtusata</i>	SRX2458823	Illumina	277	Phylogeny of pancrustaceans	Schwenter <i>et al.</i> 2018
Podocopida	Undetermined; family Cypridinae	JL207200	454	822	Phylogeny of pancrustaceans	Von Reumont <i>et al.</i> 2012
Podocopida	<i>Heterocypris incongruens</i>	ICLE01000001	Illumina	4	Expression of antioxidant genes	Hiki <i>et al.</i> 2019
Podocopida	<i>Pontocypris mytiloides</i>	SRX4048873	Illumina	277		unpublished
Podocopida	<i>Paranesidea sp.</i>	SRX2085851	Illumina	277	Phylogeny of pancrustaceans	Schwenter <i>et al.</i> 2018
Podocopida	<i>Pterygocythereis sp.</i>	SRX2458816	Illumina	277	Phylogeny of pancrustaceans	Schwenter <i>et al.</i> 2018

**Table S2: Origin of biological material.**

All species were collected in Belgium in 2018. The species printed in blue is a putative ancient asexual, species printed in red are sexually reproducing.

Species	Population name	Habitat
<i>Cyprideis torosa</i>	Dievengat	Natural brackish lake
<i>Darwinula stevensoni</i>	Zaventem	Artificial lake
<i>Notodromas monacha</i>	Overijse	Artificial lake

**Table S3: Statistics of ostracod genome sequence data.**

bp = basepairs. G = giga. The coverage is estimated from final assembly sizes. The species printed in blue is a putative ancient asexual, species printed in red are sexually reproducing. Coverage is provided in %.

Species	Data type	Insert size (bp)	Yield (Gbp)	Coverage
<i>Cyprideis torosa</i>	Paired-end 2x150 bp	250-300	45.8	136.9
		350	41.3	123.3
		550	16.2	48.5
	Mate-pair 2x150 bp	3000	13.6	40.4
		5000	12.2	36.5
<i>Darwinula stevensoni</i>	Paired-end 2x150 bp	250-300	51.8	135.5
		350	39.1	102.2
		550	13.0	34.0
	Mate-pair 2x150 bp	3000	18.7	48.9
		5000	12.2	32.0
<i>Notodromas monacha</i>	Paired-end 2x150 bp	250-300	36.9	98.0
		350	51.3	136.2
		550	15.0	39.9
	Mate-pair 2x150 bp	3000	15.5	41.1
		5000	13.4	35.6

**Table S4: Statistics of ostracod transcriptome sequence data.**

Data type: paired-end 2x 101bp. bp = basepairs. G = giga. The species printed in blue is a putative ancient asexual, species printed in red reproduce sexually.

Species	Description	Yield (Gbp)
<i>Cyprideis torosa</i>	Pool 1 (40 males and females)	7.2
	Pool 2 (40 males and females)	8.2
	Pool 3 (40 males and females)	10.1
	Pool 4 (40 males and females)	9.8
<i>Darwinula stevensoni</i>	Pool 1 (40 females)	6.7
	Pool 2 (40 females)	7.1
	Pool 3 (40 females)	7.4
<i>Notodromas monacha</i>	Pool 1 (40 females)	3.8
	Pool 2 (40 males)	20.2

**Table S5: Statistics of ostracod resequencing data.**

Data type: paired-end 2x 101bp. bp = basepairs. G = giga. The coverage is estimated from final assembly sizes. The species in blue is a putative ancient asexual, species printed in red reproduce sexually.

Species	Description	Yield (Gbp)	Coverage
<i>Darwinula stevensoni</i>	Female DS_02	21.7	56.7
	Female DS_05	23.9	62.5
<i>Notodromas monacha</i>	Female Nm_F03	20.9	55.6
	Female Nm_F07	24.0	63.8

**Table S6: Statistics of genome assemblies of ostracod species.**

$\Sigma$  represents the sum of all scaffolds in million basepairs (Mbp). The average N50 was calculated per scaffold in kilo basepairs (Kbp). The BUSCO score is the proportion of conserved single copy ortholog genes among arthropods. N is the proportion of unknown nucleotides (gaps) in the assembly. The species in blue is a putative ancient asexual, species printed in red reproduce sexually.

Species	$\Sigma$ [Mbp]	N50 [kbp]	BUSCO [%]	Ns [%]
<i>Cyprideis torosa</i>	334.9	19.0	91.8	1.3
<i>Darwinula stevensoni</i>	382.1	56.4	95.8	0.9
<i>Notodromas monacha</i>	376.7	42.3	94.4	1.6

**Table S7: Statistics of protein coding gene annotations and other gene features in ostracod genomes.**

bp = basepairs. The genome coverage is the proportion of each genome covered by genes. Red species reproduce sexually, while the species indicated in blue is a putative ancient asexual.

Species	Gene count	Transcript count	Average gene length [bp]	Genome coverage [%]
<i>Cyprideis torosa</i>	17776	18069	4734	25.1
<i>Darwinula stevensoni</i>	15453	15922	7558	30.6
<i>Notodromas monacha</i>	13771	14294	7040	25.7

**Table S8: Annotations and gene features of crustacean genomes of the last four years and of the current study.**

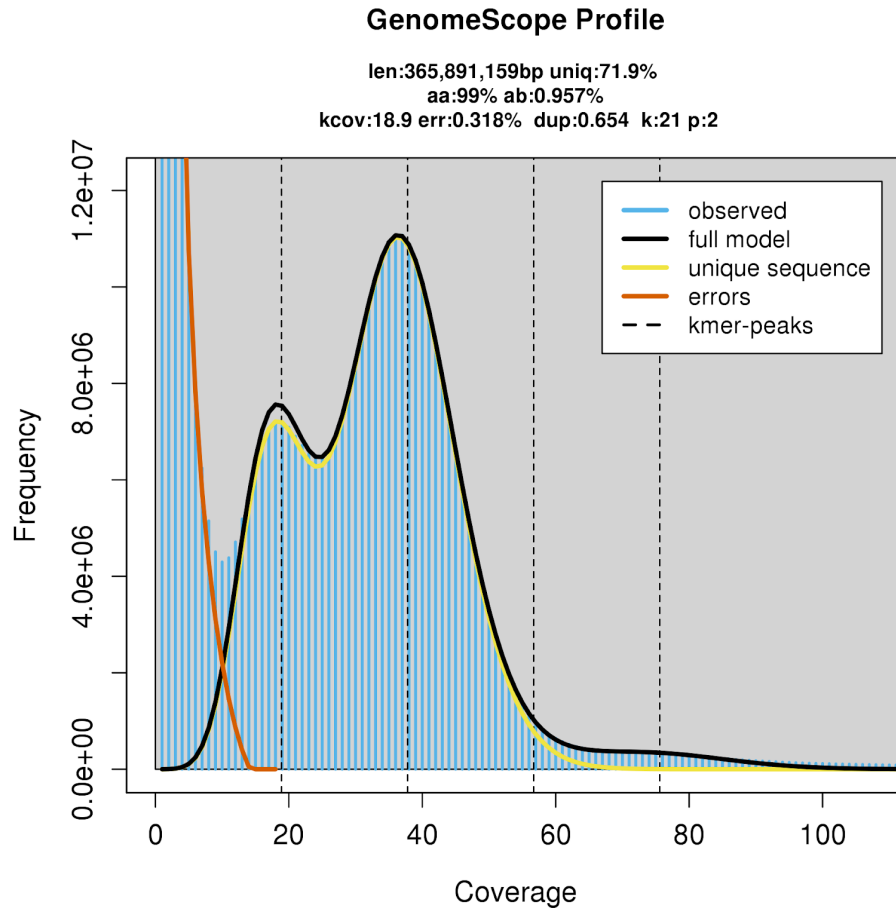
Average gene, intron and exon sizes are provided in base pairs.\* including all transcripts from different experiments and life stages. n. i. = no information.

Class	Order	Species	Predicted genes	Gene size	Intron size	Exon size	Predicted transcripts	Reference
Branchiopoda	Diplostraca	<i>Daphnia pulex</i>	18,440	2,998	223	237	177,598*	Ye <i>et al.</i> 2017
Branchiopoda	Diplostraca	<i>D. magna</i>	15,721	2,407	392	n. i.	25,997	Lee <i>et al.</i> 2019
Branchiopoda	Notostraca	<i>Lepidurus arcticus</i>	10,718	n. i.	269	233	n. i.	Savojardo <i>et al.</i> 2018
Branchiopoda	Notostraca	<i>L. apus lubbocki</i>	16,383	n. i.	258	253	n. i.	Savojardo <i>et al.</i> 2018
Branchiopoda	Spinicaudata	<i>Eulimnadia texana</i>	17,667	n. i.	n. i.	n. i.	23,965	Baldwin-Brown <i>et al.</i> 2017
Copepoda	Cyclopoida	<i>Apocylops royi</i>	29,730	n. i.	n. i.	n. i.	45,756	Jørgensen <i>et al.</i> 2019
Copepoda	Cyclopoida	<i>Oithona nana</i>	15,359	2,477	779	391	n. i.	Madoui <i>et al.</i> 2017
Copepoda	Harpaticoida	<i>Tigriopus californicus</i>	14,233	4,760	n. i.	n. i.	14,233	Barreto <i>et al.</i> 2018
Copepoda	Harpaticoida	<i>T. japonicus</i>	25,143	4,174	n. i.	n. i.	n. i.	Jeong <i>et al.</i> 2020
Copepoda	Harpaticoida	<i>T. kingsejongensis</i>	12,772	6,443	n. i.	n. i.	20,392	Kang <i>et al.</i> 2017
<b>Ostracoda</b>	<b>Podocopida</b>	<b><i>Cyprideis torosa</i></b>	<b>17,776</b>	<b>4,734</b>	<b>749</b>	<b>272</b>	<b>18,069</b>	<b>Current study</b>
<b>Ostracoda</b>	<b>Podocopida</b>	<b><i>Darwinula stevensoni</i></b>	<b>15,453</b>	<b>7,558</b>	<b>961</b>	<b>235</b>	<b>15,922</b>	<b>Current study</b>
<b>Ostracoda</b>	<b>Podocopida</b>	<b><i>Notodromas monacha</i></b>	<b>13,771</b>	<b>7,040</b>	<b>805</b>	<b>264</b>	<b>14,294</b>	<b>Current study</b>
Malacostraca	Amphipoda	<i>Parhyale hawaiiensis</i>	28,165	20,000	5,400	n. i.	26,715	Kao <i>et al.</i> 2016
Malacostraca	Isopoda	<i>Armadillidium vulgare</i>	19,051	8,639	1,506	213	n. i.	Chebbi <i>et al.</i> 2019
Malacostraca	Decapoda	<i>Cherax quadricarinatus</i>	19,494	9,768	n. i.	n. i.	n. i.	Tan <i>et al.</i> 2020
Malacostraca	Decapoda	<i>Eriocheir japonica sinensis</i>	12,772	n. i.	1,693	226	n. i.	Tang <i>et al.</i> 2020
Malacostraca	Decapoda	<i>Palaeomon carinicauda</i>	65,772	1,386	n. i.	n. i.	n. i.	Li <i>et al.</i> 2019
Malacostraca	Decapoda	<i>Penaeus monodon</i>	36,685	n. i.	n. i.	n. i.	n. i.	Van Quyen <i>et al.</i> 2020
Malacostraca	Decapoda	<i>Litopenaeus vannamei</i>	25,596	n. i.	n. i.	259	n. i.	Zhang <i>et al.</i> 2019
Malacostraca	Decapoda	<i>Marsupenaeus japonicus</i>	16,716	n. i.	n. i.	n. i.	n. i.	Yuang <i>et al.</i> 2018
Malacostraca	Decapoda	<i>Procambarus virginalis</i>	> 21,000	6,700	2,000	300	22,338	Gutekunst <i>et al.</i> 2018

### Figure S1 A-D: Results of GenomeScope analyses.

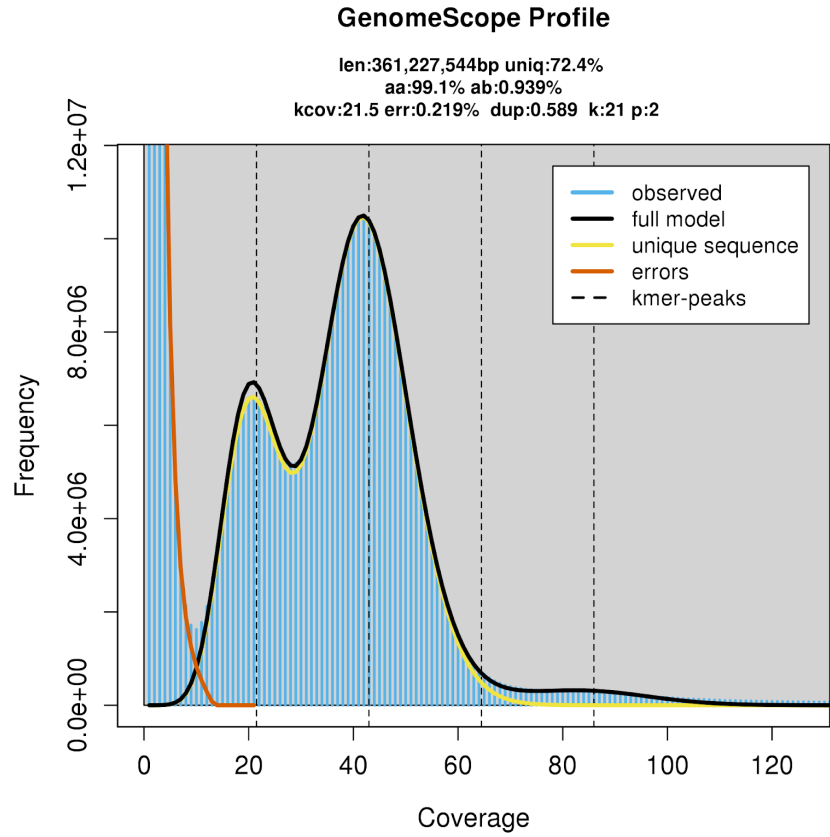
Two individuals each of *Darwinula stevensoni* (Figure S1 A-B) and *Notodromas monacha* (Figure S1 C-D) were resequenced and their reads analyzed to estimate genome sizes and heterozygosities.

Figure S1 A – DS\_02



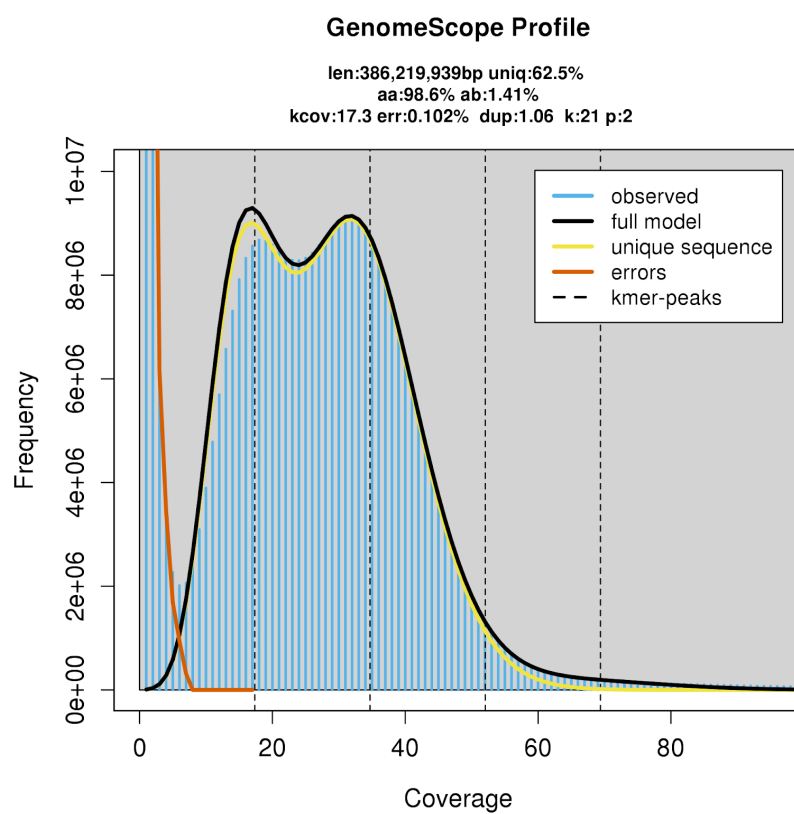
property	min	max
Homozygous (aa)	99.0133%	99.0724%
Heterozygous (ab)	0.9276%	0.9867%
Genome Haploid Length	365,167,484 bp	365,891,159 bp
Genome Repeat Length	102,728,873 bp	102,932,457 bp
Genome Unique Length	262,438,611 bp	262,958,702 bp
Model Fit	77.2855%	98.7670%
Read Error Rate	0.3181%	0.3181%

Figure S1 B – DS\_05



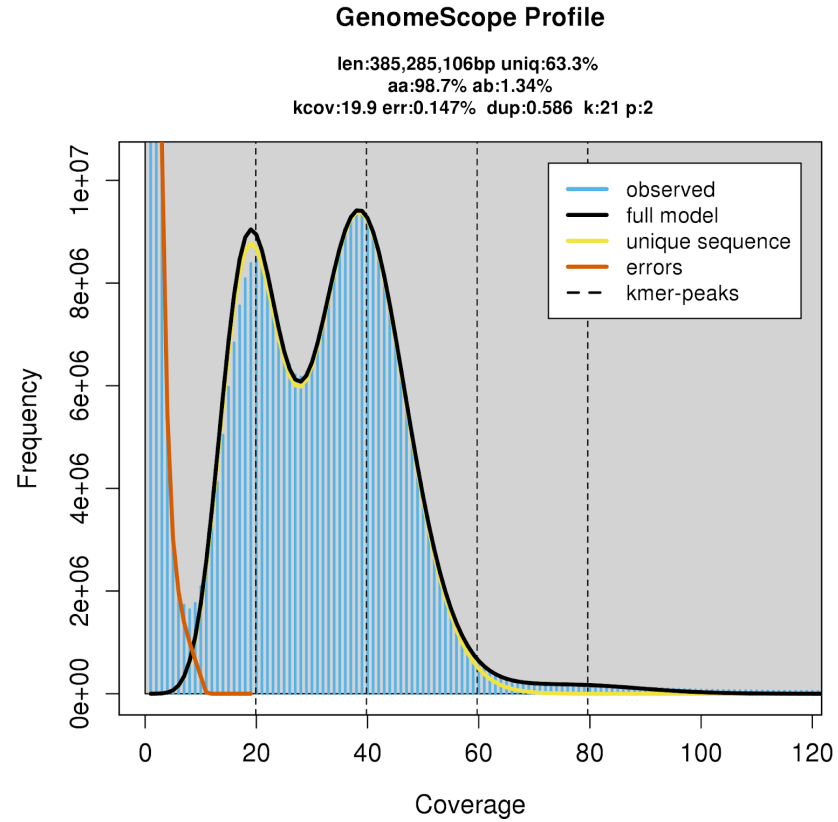
property	min	max
Homozygous (aa)	99.0465%	99.0755%
Heterozygous (ab)	0.9245%	0.9535%
Genome Haploid Length	360,688,136 bp	361,227,544 bp
Genome Repeat Length	99,559,402 bp	99,708,293 bp
Genome Unique Length	261,128,735 bp	261,519,252 bp
Model Fit	77.5377%	98.1080%
Read Error Rate	0.2192%	0.2192%

Figure S1 C – Nm\_F03



property	min	max
Homozygous (aa)	98.5730%	98.6159%
Heterozygous (ab)	1.3841%	1.4270%
Genome Haploid Length	385,081,521 bp	386,219,939 bp
Genome Repeat Length	144,297,289 bp	144,723,876 bp
Genome Unique Length	240,784,232 bp	241,496,063 bp
Model Fit	64.0406%	97.5643%
Read Error Rate	0.1021%	0.1021%

Figure S1 D – Nm\_F07



property	min	max
Homozygous (aa)	98.6515%	98.6778%
Heterozygous (ab)	1.3222%	1.3485%
Genome Haploid Length	384,661,719 bp	385,285,106 bp
Genome Repeat Length	141,131,428 bp	141,360,147 bp
Genome Unique Length	243,530,291 bp	243,924,959 bp
Model Fit	65.1164%	98.0966%
Read Error Rate	0.1470%	0.1470%

## References

- Baldwin-Brown, J. G., Weeks, S.C. & Long, A.D. (2018). A new standard for crustacean genomes: the highly contiguous, annotated genome assembly of the clam shrimp *Eulimnadia texana* reveals HOX gene order and identifies the sex chromosome. *Genome Biol. Evol.* 10, 143-156.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. & Pevzner, P.A. (2012). Spades : A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455-477. <https://doi.org/10.1089/cmb.2012.0021>
- Barreto, F.S., Watson, E.T., Lima, T.G., Willett, C.S., Edmands, S., Li, W. & Burton, R.S. (2018). Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nature Ecol. Evol.* 2, 1250-1257.
- Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D. & Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578-579. <https://doi.org/10.1093/bioinformatics/btq683>
- Bolger, A.M., Lohse, M. & Usadel, B. (2014). Trimmomatic : A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bray, N.L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnol.* 34, 525-527. <https://doi.org/10.1038/nbt.3519>
- Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. United States.
- Campbell, M.S., Holt, C., Moore, B. & Yandell, M. (2014). Genome annotation and curation using maker and maker-p. *Curr. Protoc. Bioinformatics* 48. <https://doi.org/10.1002/0471250953.bi0411s48>
- Chebbi, M.A., Becking, T., Moumen, B., Giraud, I., Gilbert, C., Peccoud, J. & Cordaux, R. (2019). The genome of *Armadillidium vulgare* (Crustacea, Isopoda) provides insights into sex chromosome evolution in the context of cytoplasmic sex determination. *Mol. Biol. Evol.* 36, 727-741.
- Chen, Y.-C., Liu, T., Yu, C.-H., Chiang, T.-Y. & Hwang, C.-C. (2013). Effects of gc bias in next-generation-sequencing data on de novo genome assembly. *PLoS ONE* 8, e62856. <https://doi.org/10.1371/journal.pone.0062856>
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M. & Robles, M. (2005). Blast2GO : A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674-3676. <https://doi.org/10.1093/bioinformatics/bti610>
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T.R. (2013). STAR : Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21. <https://doi.org/10.1093/bioinformatics/bts635>
- Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J. & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420-3435. <https://doi.org/10.1093/nar/gkn176>
- Gutekunst, J., Andriantsoa, R., Falckenhayn, C., Hanna, K., Stein, W., Rasamy, J. & Lyko, F. (2018). Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nature Ecol. Evol.* 2, 567.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protoc.* 8, 1494-1512. <https://doi.org/10.1038/nprot.2013.084>
- Hiki, K., Tsuchikawa, M., Nakajima, F. & Tobino, T. (2019). Identification of antioxidant genes in the ostracod *Heterocypris incongruens* through de novo transcriptome sequencing. *J. JSCE* 7, 133-142.
- Holt, C., & Yandell, M. (2011). MAKER2 : An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12, 491. <https://doi.org/10.1186/1471-2105-12-491>
- Jeong, C.B., Lee, B.Y., Choi, B.S., Kim, M.S., Park, J.C., Kim, D.H., et al. (2020). The genome of the harpacticoid copepod *Tigriopus japonicus*: Potential for its use in marine molecular ecotoxicology. *Aquatic Toxicol.* 222, 105462.

- Jørgensen, T. S., Nielsen, B.L.H., Petersen, B., Browne, P.D., Hansen, B.W. & Hansen, L. H. (2019). The whole genome sequence and mRNA transcriptome of the tropical cyclopoid copepod *Apocyclops royi*. *G3: Genes Genomes Genet.* 9, 1295-1302.
- Kang, S., Ahn, D.H., Lee, J.H., Lee, S.G., et al. (2017). The genome of the Antarctic-endemic copepod, *Tigriopus kingsejongensis*. *GigaScience* 6, 1 giw010.
- Kao, D., Lai, A.G., Stamatakis, E., Rosic, S., et al. (2016). The genome of the crustacean *Parhyale hawaiiensis*, a model for animal development, regeneration, immunity and lignocellulose digestion. *elife* 5, e20062.
- Kokot, M., Długosz, M. & Deorowicz, S. (2017). KMC 3 : Counting and manipulating k-mer statistics. *Bioinformatics* 33, 2759-2761. <https://doi.org/10.1093/bioinformatics/btx304>
- Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics* 5, 59. <https://doi.org/10.1186/1471-2105-5-59>
- Laetsch, D.R., & Blaxter, M.L. (2017). BlobTools : Interrogation of genome assemblies. *F1000Research* 6, 1287. <https://doi.org/10.12688/f1000research.12232.1>
- Lee, B.Y., Choi, B.S., Kim, M.S., Park, J.C., Jeong, C.B., Han, J. & Lee, J.S. (2019). The genome of the freshwater water flea *Daphnia magna*: A potential use for freshwater molecular ecotoxicology. *Aquatic Toxicol.* 210, 69-84.
- Li, J., Lv, J., Liu, P., Chen, P., Wang, J. & Li, J. (2019). Genome survey and high-resolution backcross genetic linkage map construction of the ridgetail white prawn *Exopalaemon carinicauda* applications to QTL mapping of growth traits. *BMC Genomics* 20, 598.
- Lozano-Fernandez, J., Giacomelli, M., Fleming, J.F., Chen, A., Vinther, J. et al. (2019). Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. *Genome Biol. Evol.* 11, 2055-2070. doi: 10.1093/gbe/evz097
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., et al. (2012). SOAPdenovo2 : An empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1, 18. <https://doi.org/10.1186/2047-217X-1-18>
- Madoui, M.A., Poulain, J., Sugier, K., Wessner, M., Noel, B., et al. (2017). New insights into global biogeography, population structure and natural selection from the genome of the epipelagic copepod *Oithona*. *Mol. Ecol.* 26, 4467-4482.
- Oakley, T.H., Wolfe, J.M., Lindgren, A.R. & Zaharoff, A.K. (2013). Phylotranscriptomics to bring the understudied into the fold: monophyletic Ostracoda, fossil placement, and pancrustacean phylogeny. *Mol. Biol. Evol.* 30, 215-233.
- O'Connell, J., Schulz-Trieglaff, O., Carlson, E., Hims, M.M., Gormley, N.A. & Cox, A.J. (2015). NxTrim : Optimized trimming of Illumina mate pair reads. *Bioinformatics* 31, 2035-2037. <https://doi.org/10.1093/bioinformatics/btv057>
- Oyola, S.O., Manske, M., Campino, S., Claessens, A., Hamilton, W.L., et al. (2014). Optimized whole-genome amplification strategy for extremely AT-biased template. *DNA Res.* 21, 661-671. <https://doi.org/10.1093/dnares/dsu028>
- Ranallo-Benavidez, T.R., Jaron, K.S. & Schatz, M.C. (2020). GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Commun.* 11, 1-10. <https://doi.org/10.1038/s41467-020-14998-3>
- Savojardo, C., Luchetti, A., Martelli, P.L., Casadio, R. & Mantovani, B. (2019). Draft genomes and genomic divergence of two *Lepidurus* tadpole shrimp species (Crustacea, Branchiopoda, Notostraca). *Mol. Ecol. Resour.* 19, 235-244.
- Schwentner, M., Richter, S., Rogers, D.C. & Giribet, G. (2018). Tetraconatan phylogeny with special focus on Malacostraca and Branchiopoda: highlighting the strength of taxon-specific matrices in phylogenomics. *Proc. R. Soc. B* 285, 20181524. <https://doi.org/10.1098/rspb.2018.1524>
- Seppy, M., Manni, M. & Zdobnov, E.M. (2019). Busco : Assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962, 227-245. [https://doi.org/10.1007/978-1-4939-9173-0\\_14](https://doi.org/10.1007/978-1-4939-9173-0_14)

- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S. & Morgenstern, B. (2006). AUGUSTUS : Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34, W435-439.  
<https://doi.org/10.1093/nar/gkl200>
- Tan, M.H., Gan, H.M., Lee, Y.P., Grandjean, F., Croft, L.J. & Austin, C. M. (2020). A giant genome for a giant crayfish (*Cherax quadricarinatus*) with insights into cox1 pseudogenes in decapod genomes. *Front. Genet.* 11, 201.
- Tang, B., Wang, Z., Liu, Q., Zhang, H., Jiang, S., et al. (2020). High-quality genome assembly of *Eriocheir japonica sinensis* reveals its unique genome evolution. *Front. Genet.* 10, 1340.
- Tarailo-Graovac, M. & Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics, Chapter 4*, Unit 4.10.  
<https://doi.org/10.1002/0471250953.bi0410s25>
- Van Quyen, D., Gan, H.M., Lee, Y.P., Nguyen, D.D., et al. (2020). Improved genomic resources for the black tiger prawn (*Penaeus monodon*). *Marine Genomics* 52,100751.
- von Reumont, B.M., Jenner, R.A., Wills, M. A., Dell'Ampio, E., Pass, G. et al. (2012). Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Mol. Biol. Evol.* 29, 1031–1045.
- Ye, Z., Xu, S., Spitze, K., Asselman, J. et al. (2017). A new reference genome assembly for the microcrustacean *Daphnia pulex*. *G3: Genes Genomes Genet.* 7, 1405-1416.
- Yuan, J., Zhang, X., Liu, C., Yu, Y., Wei, J., Li, F. & Xiang, J. (2018). Genomic resources and comparative analyses of two economical penaeid shrimp species, *Marsupenaeus japonicus* and *Penaeus monodon*. *Marine Genomics* 39, 22-25.
- Zhang, X., Yuan, J., Sun, Y., Li, S., Gao, Y. et al., (2019). Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. *Nature Commun.* 10, 1-14.