First annotated draft genomes of nonmarine ostracods (Ostracoda, Crustacea) with different reproductive modes

Patrick Tran Van, Yoann Anselmetti, Jens Bast, Zoé Dumas, Nicolas Galtier, Kamil Jaron, Koen Martens, Darren Parker, Marc Robinson-Rechavi, Tanja Schwander, et al.

To cite this version:

Patrick Tran Van, Yoann Anselmetti, Jens Bast, Zoé Dumas, Nicolas Galtier, et al.. First annotated draft genomes of nonmarine ostracods (Ostracoda, Crustacea) with different reproductive modes. G3, Genetics Society of America, 2021, 11 (4), pp.jkab043. 10.1093/g3journal/jkab043 . hal-03328000

HAL Id: hal-03328000
https://hal.archives-ouvertes.fr/hal-03328000
Submitted on 27 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
First annotated draft genomes of nonmarine ostracods (Ostracoda, Crustacea) with different reproductive modes

Patrick Tran Van,1,2,* Yoann Anselmetti,3,† Jens Bast,1,4 Zoé Dumas,1 Nicolas Galtier,3 Kamil S. Jaron,1 Koen Martens,4,5 Darren J. Parker1,2,4 Marc Robinson-Rechavi1,2, Tanja Schwander,1 Paul Simion1,2, Marc Robinson-Rechavi1,2

1Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
2Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland
3ISEM—Institut des Sciences de l’Evolution, Montpellier 34090, France
4Royal Belgian Institute of Natural Sciences, OD Nature, Freshwater Biology, Brussels 1000, Belgium
5Department of Biology, University of Ghent, Ghent 9000, Belgium
6Université de Namur, LEGE, URBE, Namur 5000, Belgium
7University of Hasselt, Research Group Zoology, Diepenbeek 3590, Belgium
8Present address: Institute for Zoology, University of Cologne, Köln, Germany.
9†Present address: CoBIUS Lab, Department of Computer Science, University of Sherbrooke, 2500 Boulevard de l’Université, Sherbrooke, QC J1K 2R1, Canada.
10‡Present address: Institute for Zoology, University of Cologne, Köln, Germany.
11University of Lausanne, Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland.
12Present address: Institute for Zoology, University of Cologne, Köln, Germany.
13University of Hasselt, Research Group Zoology, Diepenbeek 3590, Belgium.
14University of Lausanne, Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland.
15Present address: Institute for Zoology, University of Cologne, Köln, Germany.
16Royal Belgian Institute of Natural Sciences, OD Nature, Freshwater Biology, Brussels 1000, Belgium.
17University of Hasselt, Research Group Zoology, Diepenbeek 3590, Belgium.
18University of Lausanne, Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland.
19Department of Biology, University of Ghent, Ghent 9000, Belgium.
20Université de Namur, LEGE, URBE, Namur 5000, Belgium.
21University of Hasselt, Research Group Zoology, Diepenbeek 3590, Belgium.
22University of Lausanne, Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland.
23Royal Belgian Institute of Natural Sciences, OD Nature, Freshwater Biology, Brussels 1000, Belgium.
24University of Hasselt, Research Group Zoology, Diepenbeek 3590, Belgium.
25University of Lausanne, Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland.
26Royal Belgian Institute of Natural Sciences, OD Nature, Freshwater Biology, Brussels 1000, Belgium.

Abstract

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 nonmarine 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 (335–382 Mb) 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. 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 nonmarine ostracods provide the urgently needed basis to further develop ostracods as models for evolutionary and ecological research.

Keywords: ancient asexual; sexual; Darwinula stevensoni; Cyprideis torosa

Introduction

Relevance of ostracods

Ostracoda are small, bivalved crustaceans, widely occurring in almost all aquatic habitats as part of the meiobenthos and periphyton. There are 2330 formally described species of nonmarine 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 (335–382 Mb) 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. 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 nonmarine ostracods provide the urgently needed basis to further develop ostracods as models for evolutionary and ecological research.
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 nonmarine 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; Schón 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., Glemín et al. 2019; Jaron et al. 2020) and to develop insights into mechanisms such as gene conversion (Omilian et al. 2006), DNA repair (Schon 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; Paganini et al. 2012; Flot et al. 2013). 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.

![Figure 1](image.png)
Of all extant nonmarine 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 (Adolfsson et al. 2010; Symonová et al. 2018), 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 nonmarine ostracods, one putative ancient asexual darwinulid ostracod and two species with obligate sexual reproduction.

The ostracod family Darwinulidae is one of the two last remaining animal groups which are still supported as being genuine 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 been 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, 2003a, 2003b; Van den Broecke et al. 2013) and molecular research using DNA sequence data from single genes (Schön et al. 1998; 2003; Martens et al. 2005; Schön et al. 2012). These studies revealed that D. stevensoni is most likely apomorphic 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), which could be an explanation for the low observed mutation rates. These results, however, were based on a limited number of genes and require further confirmation with genomewide 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 nonmarine 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 to be 0.86–0.93 pg with flow cytometry (Paczesniak, unpublished), approximating 900 Mb. There is no information on the ploidy level of D. stevensoni, except for the study by Tétart (1979) showing 22 dot-like chromosomes.

Because of its putative ancient asexuality, no close sexual relatives of D. stevensoni are available for comparative, genomic analyses. We have chosen two fully sexual nonmarine ostracod species from the Cytherideidae and the Notodromadidae with high population densities in Belgium as comparisons to the putative ancient asexual: C. torosa and Notodromas monacha, respectively. Cyprides 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, 1976a, 1976b; De Deckker et al. 1999; Keyser 2005). Fregley and Whittaker (2017) 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 (Schon 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 analyzed here, N. monacha, occurs throughout the Northern hemisphere and is a nonmarine ostracod with a most peculiar behavior: 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 0.87 pg (Jeffery et al. 2017; Gregory 2020) very similar to that of D. stevensoni. This species has not yet been the subject of any molecular studies.

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

**Materials and methods**

**Sample collection for genome and transcriptome sequencing**

All three nonmarine 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 C. torosa and N. monacha. C. torosa was sampled by moving the net on the water surface. Nonmarine 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 were extracted. More details on the origin of biological samples are provided in Supplementary Table S2.

For generating reference genomes, DNA was extracted from a single female of each species using the QIAamp DNA Micro kit according to the 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-extraction methods were applied to the homogenized...
tissue and the aqueous layer transferred to RNAasy MinElute Columns (Qiagen). Subsequent steps of RNA extraction were done following the RNAasy 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 2 × 125 bp paired-end libraries with average insert sizes of 250–300, 550 and 700 bp, and two mate-pair libraries with average insert sizes of 3000 and 5000 bp; see Supplementary 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 351× and 386× (Supplementary Table S3).

Reads were filtered with Trimmomatic v0.36 (Bolger et al. 2014) and NxTrim v0.4.1 (O’Connell et al. 2015). Because of uneven coverage produced by PCR-based whole-genome amplification (Chen et al. 2013; Oyola et al. 2014), we first normalized reads using BBMap 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. 2011). Scaffolds identified as contaminants were filtered out using Blobtools v1.0 (Laetsch and Blaxter 2017). The completeness of genomes assembled as contaminants were filtered out using Blobtools v1.0 (Laetsch and Blaxter 2017). The completeness of genomes assembled as contaminants were filtered out using Blobtools v1.0 (Laetsch and Blaxter 2017). The completeness of genomes assembled as contaminants were filtered out using Blobtools v1.0 (Laetsch and Blaxter 2017). The completeness of genomes assembled as contaminants were filtered out using Blobtools v1.0 (Laetsch and Blaxter 2017).

**Protein coding gene annotation**

Libraries were prepared using the Illumina TruSeq Stranded RNA kit, following the manufacturer’s instructions. RNA reads were generated with the Illumina HiSeq 2500 system (Supplementary Table S4). Reads were filtered with Trimomatic 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 arthropod protein database (v 2018-10).

More details of the annotation pipelines and the applied parameters can be found in Supplementary Material 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. stevensoni and N. monacha without whole genome amplification, preparing libraries with the NEBNext Ultra™ II DNA Library Prep Kit for Illumina. Reads were filtered with Trimomatic 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.

**Table 1** Quality features of published crustacean genomic assemblies of the last 4 years and of the current study

| Class          | Order       | Species                      | Size     | No. of scaffolds | N50        | BUSCO       | Reference |
|---------------|-------------|------------------------------|----------|-----------------|------------|-------------|-----------|
| Branchiopoda  | Diplostraca | Daphnia pulexa               | 156      | 1,822           | 1,661      | 96          | Ye et al. (2017) |
| Branchiopoda  | Diplostraca | D. magna                    | 130      | 4,193           | 10,124     | 96.7 (C)   | Lee et al. (2019) |
| Branchiopoda  | Notostraca  | Lepidurus arcticus          | 73       | 7,167           | 116        | 98.4 (C)   | Savojardo et al. (2019) |
| Branchiopoda  | Notostraca  | L. apus lubbiicki           | 90       | 20,738          | 402        | 97.8 (C)   | Savojardo et al. (2019) |
| Branchiopoda  | Spinicaudata| Eulimnadia texanac          | 120      | 112             | 18,000     | n.i.        | Baldwin-Brown et al. (2018) |
| Copepoda      | Cyclopoida  | Apocylops royi               | 258      | 97,072          | n.i.       | 50 (C)     | Jørgensen et al. (2019) |
| Copepoda      | Cyclopoida  | Oithona nama                 | 85       | 4,626           | 401        | n.i.        | Madoui et al. (2017) |
| Copepoda      | Harpacticoa | Tigriopus Californicus      | 190      | 459             | 298        | 94.5 (C)   | Barreto et al. (2018) |
| Copepoda      | Harpacticoa | T. japonicus                | 197      | 339             | 10,650     | 96 (C)     | Jeong et al. (2020) |
| Copepoda      | Harpacticoa | T. kingsejongensis          | 295      | 270,823         | 159        | 61.1 (C)   | Kang et al. (2017) |
| Ostracoda     | Podocopida  | Cyprideis torosa            | 335      | 132,611         | 19         | 86.6 (C)   | Current study |
| Ostracoda     | Podocopida  | Daruinula stevensoni        | 382      | 62,118          | 56         | 93.7 (C)   | Current study |
| Ostracoda     | Podocopida  | Notodromas monacha          | 377      | 62,251          | 42         | 95.8 (C + F)| Current study |
| Malacostraca  | Amphipoda   | Parhyale hawaiensiseb       | 4,024    | 100,000         | 69         | n.i.       | Kao et al. (2016) |
| Malacostraca  | Isopoda     | Armadillidium vulgarec      | 1,725    | 43,451          | 51         | 87.9 (C)   | Chebbi et al. (2019) |
| Malacostraca  | Decapoda    | Cherax quadricarinatus      | 3,237    | 508,682         | 33         | 81.3 (C)   | Tan et al. (2020) |
| Malacostraca  | Decapoda    | Eriocheir japonica sinensis | 1,270    | 1,368           | 3,185      | 92.7 (C)   | Tang et al. (2020) |
| Malacostraca  | Decapoda    | Palaemon carinicaudab       | 2,185    | 28,089,718      | 586        | n.i.       | Li et al. (2019) |
| Malacostraca  | Decapoda    | Pemaonos monodon            | 1,600    | 1,211,364       | 2          | 96.8 (C + F)| Van Quyen et al. (2020) |
| Malacostraca  | Decapoda    | Litopenaeus vannameibc      | 1,664    | 4,682           | 606        | 95         | Zhang et al. (2019) |
| Malacostraca  | Decapoda    | Marsupenaeus japonicus      | 924      | 37,192,281      | 1          | 97         | Yuan et al. (2018) |
| Malacostraca  | Decapoda    | Procambarus virginalis      | 3,300    | 3,752,011       | 39         | n.i.       | Gutekunst et al. (2018) |
Data availability

Raw sequence reads have been deposited in NCBI’s sequence read archive under the following bioprojects: PRJNA515625 (reference genomes, Supplementary Table S3) and PRJNA631617 (RNA-seq for annotations and resequenced individuals, Supplementary Tables S4 and S5).

Genome assemblies and annotations have been deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB38362 (Supplementary Tables S6 and S7). Codes for the analyses are available at: https://github.com/Asex GenomeEvol/Ostracoda_genomes. Additional supplementary material is available at the figshare collection page, including detailed descriptions of the methods (Supplementary Methods SM1–SM3), Tables S1 and S2, and S6 and S8, and the results of the GenomeScope analyses of D. stevensoni and N. monacha (Supplementary Figures S1, A–D). Supplemental Material available at figshare: https://doi.org/10.25387/g3.13858817.

Results and discussion

First ostracod reference genomes and their attributes

We produced the first de novo reference genomes of nonmarine ostracods, namely of the three species D. stevensoni, C. torosa, and N. monacha with different reproductive modes (see Supplementary Material SM1 and Tables S3 and S4 for more details on the assemblies). 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 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. stevensoni, had the best contiguity, with the largest N50 although the total number of scaffolds was similar to N. monacha (Table 1). 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 C. 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% (Supplementary Table S7) and of 87% for complete single copy genes (Table 1). All three species have similar numbers of predicted genes and transcripts (Supplementary 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. stevensoni was estimated at 420–455 Mb with flow cytometry (Paczesniak, unpublished) while we estimated 362 Mb from sequence reads (Supplementary Figure S1, A and 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 (Supplementary Figure S1, C and D) that we obtained from sequence reads. It thus seems that either the genome size estimates by flow cytometry are incorrect or 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% (Supplementary Figure S1, A and B) and 1.32%–1.43% for the sexual N. monacha (Supplementary Figure S1, C and 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 apomorphic (Chaplin et al. 1994), implying that observed heterozygosity levels are largely dependent 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 4 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. Both features are comparable to those of the copepod Apoeclops royi (Jorgensen et al. 2019) and the amphipod Parhyale hawaiaiensis (Kao et al. 2016) (Table 1) and better than for crustaceans with larger genomes such as the decapods Cherax quadrinarinatus (Tan et al. 2020), Palaeomon carinicauda (Li et al. 2019), Penaeus monodon (Van Quyen et al. 2020), Marsupenaeus japonicus (Yuan et al. 2018), and Procambus virginalis (Gutekunst et al. 2018; Table 1). Genome assemblies of several other crustaceans, however, have smaller scaffold numbers and higher N50 and thus better contiguities than the assemblies obtained here for nonmarine ostracods. For the two notostracan Lepidurus species (Savojardo et al. 2019), 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 Erichthion japonica sinensis—Tang et al. 2020, the copepod Tigriopus californicus—Jeong et al. 2020, and the isopod Armadillium 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 nonmarine ostracods has already commenced (Shon 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 (Supplementary Table S8). We have predicted 13,771–17,776 protein coding genes in the three nonmarine ostracod genomes (Supplementary Tables S7 and S8), with the highest number for the sexual C. torosa and an intermediate estimate for the putative ancient asexual D. stevensoni. The number of annotated protein coding genes in nonmarine ostracods is similar to estimates for various branchiopods and the copepods Oithona nana, Tigriopus californicus, and T. kingsejongensis but lower than in most malacostracans (Supplementary 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 (Supplementary 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 (Rotllant et al. 2018). The novel data on predicted genes and transcripts from nonmarine 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 nonmarine and marine ostracods.

Conclusions

We have successfully obtained de novo genome assemblies for three species of nonmarine 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 genomes presented here can also provide the first step toward a genomic assessment of the putative ancient asexual status of nonmarine darwinulid ostracod species.

Acknowledgments

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.

Funding

This research was funded by Belgian Federal Science Policy (BR/314/PI/LATTECO) and a grant from the Swiss National Science Foundation (CRSII3_160723).

Conflicts of interest: None declared.

Literature cited

Adolfsson S, Michalakis Y, Paczesniak D, Bode SNS, Butlin RK, 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 JG, Weeks SC, Long AD. 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, Nurk S, Antipov D, Gurevich AA, Dvorkin M, et al. 2012. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19:455–477. doi: 10.1089/cmb.2012.0021.

Barroto FS, Watson ET, Lima TG, Willett CS, Edmands S, et al. 2018. Genomic signatures of mitonuclear coevolution across populations of Tigriopus californicus. Nat Ecol Evol. 2:1250–1257. doi: 10.1038/s41559-018-0588-1.

Bell G. 1982. The Masterpiece of Nature: The Evolution and Genetics of Sexuality. Berkeley, CA: University of California Press.

Benede J, Möst M, Meyer A. 2020. Optimized and affordable high-throughput sequencing workflow for preserved and nonpreserved small zooplankton specimens. Mol Ecol Res. 20:1632–1646. doi: 10.1111/1755-0998.13228.

Birky CW, Jr. 1996. Heterozygosity, heteromorphomy, and phylogenetic trees in asexual eukaryotes. Genetics 144:427–437.

Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SPAspace. Bioinformatics 27:578–579. doi: 10.1093/bioinformatics/btq683.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. doi: 10.1093/bioinformatics/btu170.
Boschetti C, Carr A, Crisp A, Eyres I, Wang-Koh Y, et al. 2012. Biochemical diversification through foreign gene expression in bdelloid rotifers. PLoS Genet. 8:e1003035. doi: 10.1371/journal.pgen.1003035.

Botelho-Castro F, Figuet E, Tlak MK, Nabholz B, Galtier N. 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/mzx236.

Bushnell B. 2014. BBMap: A fast, accurate, splice-aware aligner. 9th Annual Genomics of Energy & Environment Meeting, Walnut Creek, CA, USA.

Butlin RK, Menozzi P. 2000. Open questions in evolutionary ecology: do ostracods have the answers? In: Horne DJ, Martens K (eds.), Evolutionary Biology and Ecology of Ostracoda. Dordrecht: Springer. p. 1–14. doi: 10.1023/A:1003925705735.

Butlin RK, Schön I, Martens K. 1998. Asexual reproduction in non-marine ostracods. Heredity 81:473–480. doi: 10.1046/j.1365-2540.1998.00454.x.

Chaplin JA, Havel JE, Hebert PDN. 1994. Sex and ostracods. Trends Ecol Evol. 9:435–439. doi: 10.1016/0169-5347(94)90127-9.

Chebbi MA, Becking T, Moumen B, Giraud I, Gilbert C, Chaplin JA, Havel JE, Hebert PDN. 1994. Sex and ostracods. Trends Ecol Evol. 9:435–439. doi: 10.1016/0169-5347(94)90127-9.

Chen Y-C, Liu T, Yu C-H, Chiang T-Y, Hwang C-C. 2013. Effects of non-marine ostracods. Heredity 81:473–480. doi: 10.1046/j.1365-2540.1998.00454.x.

De Deckker P, Chivas AR, Shelley JMG. 1999. Uptake of Mg and Sr in non-marine ostracods. Heredity 81:473–480. doi: 10.1046/j.1365-2540.1998.00454.x.

Conesa A, Götz S, García-Gómez JM, Terol J, Talon M, et al. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. doi: 10.1093/bioinformatics/bti610.

Danchin EG, Rosso MN, Vieira P, de Almeida-Engler J, Coutinho PM, et al. 2010. Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. Proc Natl Acad Sci USA. 107:17651–17656. doi: 10.1073/pnas.1008486107.

De Deckker P, Chivas AR, Shelley JMG. 1999. Uptake of Mg and Sr in non-marine ostracods. Heredity 81:473–480. doi: 10.1046/j.1365-2540.1998.00454.x.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, et al. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. doi: 10.1093/bioinformatics/bts635.

Duncan T, Kingan SB, Lambert CC, Baybayan P, Korlach J. 2019. A low DNA input protocol for high-quality PacBio de novo genome assemblies. J Biomol Techn. 30:S1–S2.

Flot J-F, Hespeels B, Li X, Noel B, Arkhipova I, Duncan T, Kingan SB, Lambert CC, Baybayan P, Korlach J. 2019. A low DNA input protocol for high-quality PacBio de novo genome assemblies. J Biomol Techn. 30:S1–S2.

Frogley MR, Whittaker JE. 2017. On the origins of Cyprideis torosa and Armadillidium vulgare as models for animal development, regeneration, immunity and lignocellulose digestion. elife 5:e20062. doi: 10.7554/eLife.20062.

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, et al. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36:3420–3435. doi: 10.1093/nar/gkn176.

Gregory TR. 2020. Animal genome size database. Available online: http://genomesize.com (Accessed: 2020 October 7).

Gutekunst J, Andriantsao R, Falckenhayn C, Hanna K, Stein W, et al. 2018. Clonal genome evolution and rapid invasive spread of the marbled crayfish. Nat Ecol Evol. 2:567–573. doi: 10.1038/s41559-018-0467-9.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8:1494–1512. doi: 10.1038/nprot.2013.084.

Hecox-Lea BJ, Mark Welch DB. 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, Horton RA, Scheu S, Maraun M. 2009. Parthenogenesis in oribatid mites (Acari, Oribatida): evolution without sex. In: I Schön, K Martens, P Van Dijk, editor. Lost Sex. The Evolutionary Biology of Parthenogenesis. Dordrecht, The Netherlands: Springer Academic Publishers. p. 241–258. doi: 10.1007/978-90-481-2770-2_12.

Heip C. 1976a. The life-cycle of Cyprideis torosa (Crustacea, Ostracoda). Oecologia 24:229–245. doi: 10.1007/BF00345475.

Heip C. 1976b. The spatial pattern of Cyprideis torosa (Jones, 1850) (Crustacea: Ostracoda). J Mar Biol Assoc. 56:179–189. doi: 10.1017/S002531540002052X.

Holt C, Yandell M. 2011. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics. 12:491.doi: 10.1186/1471-2105-12-491.

Horne DJ, Baltanas A, Paris, G. 1998. Geographical distribution of reproductive modes in living non-marine ostracods. In: K Martens, editor. Sex and Parthenogenesis. Evolutionary Biology of Reproductive Modes in Non-Marine Ostracods. Leiden: Backhuys Publishers. p. 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 KS, Bast J, Novell RW, Ranallo-Benavidez TR, Robinson-Rechavi M, et al. 2020. Genomic features of asexual animals. J Hered. esa031.doi: 10.1093/jhered/esa031.

Jeffery NW, Ellis EA, Oakley TH, Gregory TR. 2017. The genome sizes of ostracod crustaceans correlate with body size and evolutionary history, but not environment. J Hered. 108:701–706. doi: 10.1093/jhered/esx055.

Jeong CB, Lee BY, Choi BS, Kim MS, Park JC, et al. 2020. The genome of Tigriopus japonicus, a model for animal development, regeneration, immunity and lignocellulose digestion. elife 5:e20062. doi: 10.7554/eLife.20062.

Kang S, Ahn D-H, Lee J H, Lee S G, Shin S C, et al. 2017. The genome of the Antarctic-endemic copepod, Tigriopus kingsejongensis. Gigascience. 6:1–9.

Kao D, Lai AG, Stamataki E, Rosic S, Konstantinides N, et al. 2016. The genome of the crustacean Parhyale hawaiensis, a model for animal development, regeneration, immunity and lignocellulose digestion. elife 5:e20062. doi: 10.7554/eLife.20062.
Keyser D. 2005. Histological peculiarities of the noding process in Cyprides toros (Jones) (Crustacea, Ostracoda). Hydrobiologia 538: 95–106. doi: 10.1007/s10750-004-4940-x.

Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. F1000Research. 6:1287. doi: 10.12688/f1000research.12232.1.

Lee BY, Choi BS, Kim MS, Park JC, Jeong CB, Madoui M-A, Poulain J, Sugier K, Wessner M, Noel B, Martens K, Rossetti G, Butlin RK, Schön I. 2005. Molecular and morphological phylogeny of the ancient asexual Darwinulidae (Crustacea, Ostracoda). Hydrobiologia 538:153–165. doi: 10.1007/Pl00001867.

Martens K, Horne DJ. 2016. Collecting and processing living non-marine ostracods. J Crust Biol. 36:849–854. doi: 10.1163/1937240X-000002488.

Martens K, Horne DJ, Griffiths HI. 1998. Age and diversity of non-marine ostracods. In: K Martens, editor. Sex and Parthenogenesis. Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods. Leiden: Backhuys Publishers. p. 37–55.

Martens K, Rossetti G, Butlin RK, Schön I. 2005. Molecular and morphological phylogeny of the ancient asexual Darwinulidae (Crustacea, Ostracoda). Hydrobiologia 538:153–165. doi: 10.1007/Pl00001867.

Martens K, Rossetti G, Horne D. 2003. How ancient are ancient asexuals? Proc R Soc Lond B 270:723–729. doi: 10.1098/rspb.2002.2270.

McGregor DL. 1969. The reproductive potential, life history and parasitism of the freshwater Ostracods Darwinula stevensoni (Brady and Robertson). In: JW Neale, editor. The Taxonomy, Morphology and Ecology of Recent Ostracoda. Edinburgh: Oliver and Boyd. p. 194–221.

Meisch C. 2000. Freshwater Ostracoda of Western and Central Europe. Heidelberg, Germany: Springer.

Meisch C, Smith RJ, Martens K. 2019. A subjective global checklist of the extant non-marine Ostracoda (Crustacea). EJT 25: doi: 10.5852/ejt.2019.492.

Merckx T, Souffreau C, Kaiser A, Baardsen LF, Backeljau T, et al. 2018. Body size shifts in aquatic and terrestrial urban communities. Nature 558:113–116. doi: 10.1038/s41586-018-0140-0.

Neiman M, Meirmans PG, Ranallo-Benavidez TR, Jaron KS, Schatz MC. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 11:1–10. doi: 10.1038/s41467-020-14998-3.

Otto SP. 2020. Selective interference and the evolution of sex. J Hered. essa026. doi: 10.1093/jhered/essa026.

Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination. Nat Rev Genet. 3:252–261. doi: 10.1038/nrg6761.

Oyola SQ, Manske M, Campino S, Claessens A, Hamilton WL, et al. 2014. Optimized whole-genome amplification strategy for extremely AT-biased template. DNA Res. 21:661–671. doi: 10.1093/dnares/dsu028.

Paganini, Campan-Fournier A, Da Rocha M, Gouret P, Pontarotti P, et al. 2012. Contribution of lateral gene transfers to the genome composition and parasitic ability of root-knot nematodes. PLoS One 7:e50875. doi: 10.1371/journal.pone.0050875.

Peona V, Blom MP, Xu L, Burri R, Sullivan S, et al. 2020. Identifying the causes and consequences of assembly gaps using a multiplat-form genome assembly of a bird-of-paradise. Mol Ecol Res. doi: 10.1111/1755-0998.13252.

Ranallo-Benavidez TR, Jaron KS, Schatz MC. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 11:1–10. doi: 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, Todeschi EBA, Gandolfi A, Invidia M, Menozzi F. 2002. Hypoxia and starvation tolerance in individuals from a riverine and a lacustrine population of Darwinula stevensoni (Crustacea: Ostracoda). Arch Hydrobiol. 154:151–171. doi: 10.1127/archiv-hydrobiol/154/2002/151.

Rolflant G, Palero F, Mather PB, Bracken-Grissom HD, Santos MB. 2018. Preface: recent advances in crustacean genomics. Hydrobiologia 825:1–4. doi: 10.1007/s10750-018-3773-y.

Savojardo C, Luchetti A, Martelli PL, Casadio R, Mantovani B. 2019. Draft genomes and genomic divergence of two Lepidurus tadalafil shrimp species (Crustacea, Branchiopoda, Notostraca). Mol Ecol Resour. 19:235–244. doi: 10.1111/1755-0998.12952.

Schmit O, Fukova L, Vandekerkhove J, Michalakis Y, Matzke-Karasz R, et al. 2012. Mate recognition as a reproductive barrier in sexual and parthenogenetic Eucypris virens (Crustacea, Ostracoda). Anim Behav. 85:977–985. doi: 10.1016/j.anbehav.2013.02.021.

Schön I, Butlin RK, Griffiths HI, Martens K. 1998. Slow molecular evolution in an ancient asexual ostracod. Proc R Soc Lond B 265: 235–242. doi: 10.1098/rspb.1998.0287.

Schön I, Halse S, Martens K. 2017. Cyprides (Crustacea, Ostracoda) in Australia. J Micropalaeontol. 36:31–37. doi: 10.1144/jmpaleo2016-032.

Schön I, Martens K. 1998. Opinion: DNA repair in an ancient asexual - a new solution for an old problem? J Nat Hist. 32:943–948. doi: 10.1080/00222939800770481.

Schön I, Martens K. 2003. No slave to sex. Proc R Soc Lond B 270: 827–833. doi: 10.1098/rspb.2002.2314.

Schön I, Martens K. 2016. Ostracod (Ostracoda, Crustacea) genomics - Promises and challenges. Mar Genomics 29:19–25. doi: 10.1016/j.margen.2016.03.008.

Schön I, Martens K, Van Dijk P, editors. 2009a. Lost Sex. The Evolutionary Biology of Parthenogenesis. Dordrecht, The Netherlands: Springer Academic Publishers.
