

# The complete mitochondrial genome of the grooved carpet shell, *Ruditapes decussatus* (Bivalvia, Veneridae)

Fabrizio Ghiselli<sup>1,\*</sup>, Liliana Milani<sup>1,\*</sup>, Mariangela Iannello<sup>1</sup>, Emanuele Procopio<sup>1</sup>, Peter L. Chang<sup>2</sup>, Sergey V. Nuzhdin<sup>2</sup> and Marco Passamonti<sup>1</sup>

<sup>1</sup> Department of Biological, Geological and Environmental Sciences, University of Bologna, Italy, Bologna, Italy

<sup>2</sup> Department of Biological Sciences, Program in Molecular and Computational Biology, University of Southern California, Los Angeles, CA, USA

\* These authors contributed equally to this work.

## ABSTRACT

Despite the large number of animal complete mitochondrial genomes currently available in public databases, knowledge about mitochondrial genomics in invertebrates is uneven. This paper reports, for the first time, the complete mitochondrial genome of the grooved carpet shell, *Ruditapes decussatus*, also known as the European clam. *Ruditapes decussatus* is morphologically and ecologically similar to the Manila clam *Ruditapes philippinarum*, which has been recently introduced for aquaculture in the very same habitats of *Ruditapes decussatus*, and that is replacing the native species. Currently the production of the European clam is almost insignificant, nonetheless it is considered a high value product, and therefore it is an economically important species, especially in Portugal, Spain and Italy. In this work we: (i) assembled *Ruditapes decussatus* mitochondrial genome from RNA-Seq data, and validated it by Sanger sequencing; (ii) analyzed and characterized the *Ruditapes decussatus* mitochondrial genome, comparing its features with those of other venerid bivalves; (iii) assessed mitochondrial sequence polymorphism (SP) and copy number variation (CNV) of tandem repeats across 26 samples. Despite using high-throughput approaches we did not find evidence for the presence of two sex-linked mitochondrial genomes, typical of the doubly uniparental inheritance of mitochondria, a phenomenon known in ~100 bivalve species. According to our analyses, *Ruditapes decussatus* is more genetically similar to species of the Genus *Paphia* than to the congeneric *Ruditapes philippinarum*, a finding that bolsters the already-proposed need of a taxonomic revision. We also found a quite low genetic variability across the examined samples, with few SPs and little variability of the sequences flanking the control region (Largest Unassigned Regions (LURs)). Strikingly, although we found low nucleotide variability along the entire mitochondrial genome, we observed high levels of length polymorphism in the LUR due to CNV of tandem repeats, and even a LUR length heteroplasmy in two samples. It is not clear if the lack of genetic variability in the mitochondrial genome of *Ruditapes decussatus* is a cause or an effect of the ongoing replacement of *Ruditapes decussatus* with the invasive *Ruditapes philippinarum*, and more analyses, especially on nuclear sequences, are required to assess this point.

Submitted 19 May 2017  
Accepted 25 July 2017  
Published 22 August 2017

Corresponding author  
Fabrizio Ghiselli,  
fabrizio.ghiselli@unibo.it

Academic editor  
Tim Collins

Additional Information and  
Declarations can be found on  
page 23

DOI 10.7717/peerj.3692

© Copyright  
2017 Ghiselli et al.

Distributed under  
Creative Commons CC-BY 4.0

**OPEN ACCESS**

**Subjects** Zoology

**Keywords** Complete mitochondrial genome, Mitochondrial length polymorphism, Mitochondrial repeats, Codon usage, Bivalve molluscs, European clam, Comparative mitochondrial genomics, mtDNA *de novo* assembly, RNA-Seq, Doubly uniparental inheritance

## INTRODUCTION

Despite a large number of animal complete mitochondrial genomes (mtDNAs) being available in public databases (>55,000 in GenBank), up to now sequencing has been focused mostly on vertebrates (~50,000 in GenBank), and the current knowledge about mitochondrial genomics in invertebrates—with the notable exception of few model organisms (e.g., *Drosophila* and *Caenorhabditis elegans*)—is uneven. To better understand invertebrate mitochondrial biology—and, most importantly, mitochondrial biology and evolution in general—it is necessary to adopt a more widespread approach in gathering and analyzing data. Failing to do so would bias our knowledge toward a few taxonomic groups, with the risk of losing a big part of the molecular and functional diversity of mitochondria. Actually, despite maintaining its core features in terms of genetic content, mtDNA in Metazoa shows a wide range of variation in some other traits such as, for example, genome architecture, abundance of unassigned regions (URs)—namely regions with no assigned product (protein, RNA)—repeat content, gene duplications, introns, UTRs, and even additional coding genes (see [Breton et al., 2014](#) for a review) or genetic elements (e.g., small RNAs, see [Pozzi et al., 2017](#)). All this emerging diversity is in sharp contrast with the—at this point outdated—textbook notion about mtDNAs role being limited to the production of a few subunits of the protein complexes involved in oxidative phosphorylation.

This paper reports, for the first time, the complete mitochondrial genome of the grooved carpet shell, *Ruditapes decussatus* (Linnaeus, 1758). *Ruditapes decussatus*—also known as the European clam—is distributed all over the Mediterranean coasts, as well as on the Atlantic shores, from Lofoten Islands (Norway) to Mauritania, including the British Isles. *Ruditapes decussatus* lives in warm coastal waters, especially in lagoons, and it is morphologically and ecologically similar to the Manila clam *Ruditapes philippinarum*, which has been recently introduced for aquaculture in the very same habitats of *Ruditapes decussatus*. *Ruditapes philippinarum*, native from the Philippines, Korea, and Japan, was accidentally introduced into North America in the 1930s, and from there it was purposely introduced in France (1972), UK (1980), and Ireland (1982) for aquaculture purposes ([Gosling, 2003](#)). According to historical records, *Ruditapes decussatus* was one of the most important species for aquaculture in Europe, but overfishing, irregular yields, recruitment failure, and outbreaks of bacterial infection pushed the producers to introduce *Ruditapes philippinarum*; Italy imported large quantities of *Ruditapes philippinarum* seed from UK in 1983 and 1984. Compared to the European clam, the Manila clam turned out to be faster growing, more resistant to disease, to have a more extended breeding period and a greater number of spawning events, and to begin sexual maturation earlier (i.e., at a smaller size). Upon introduction of the more robust *Ruditapes philippinarum*, *Ruditapes decussatus* suffered a population decline in the Southwestern Europe ([Arias-Pérez et al., 2016](#)), and

currently the production of the European clam is almost insignificant. Nonetheless the grooved carpet shell is considered a high value product, and therefore it is an economically important species, especially in Portugal, Spain and Italy ([Gosling, 2003](#); [Leite et al., 2013](#); [de Sousa et al., 2014](#)).

Molluscs in general, and bivalves in particular, exhibit an extraordinary degree of mtDNA variability and unusual features, such as: large mitochondrial genomes (up to ~47Kb), high proportion of URs (i.e., number of base pairs annotated as URs over the total mtDNA length), novel protein coding genes with unknown function, frequent and extensive gene rearrangement, and differences in strand usage ([Gissi, Iannelli & Pesole, 2008](#); [Breton et al., 2011](#); [Ghiselli et al., 2013](#); [Milani et al., 2014b](#); [Plazzi, Puccio & Passamonti, 2016](#)). Moreover, mitochondrial genome size varies among bivalves because of gene duplications and losses ([Serb & Lydeard, 2003](#); [Passamonti et al., 2011](#); [Ghiselli et al., 2013](#)), and sometimes genes are fragmented as in the case of ribosomal genes in oysters ([Milbury et al., 2010](#)). The most notable feature of bivalve mtDNA is the doubly uniparental inheritance (DUI) system of transmission ([Skibinski, Gallagher & Beynon, 1994a, 1994b](#); [Zouros et al., 1994a, 1994b](#)). Under DUI, two different mitochondrial lineages (and their respective genomes) are transmitted to the progeny: one is inherited from the egg (female-transmitted or F-type mtDNA), the other is inherited from the spermatozoon (male-transmitted or M-type mtDNA). Following fertilization, the early embryo is heteroplasmic, but the type of mitochondria present in the adult is tightly linked to its sex. Females are commonly homoplasmic for F, while males are heteroplasmic with the following distribution of mtDNA types: the germ line is homoplasmic for the M-type (which will be transmitted via sperm to male progeny), the soma is heteroplasmic to various degrees, depending on tissue type and/or species ([Ghiselli, Milani & Passamonti, 2011](#); [Zouros, 2013](#)). To date, the only known animals exhibiting DUI are about 100 species of bivalve molluscs ([Gusman et al., 2016](#)). This natural and evolutionarily stable heteroplasmic system can be extremely useful to investigate several aspects of mitochondrial biology (see [Passamonti & Ghiselli, 2009](#); [Breton et al., 2014](#); [Milani & Ghiselli, 2015](#); [Milani, Ghiselli & Passamonti, 2016](#)). Indeed, despite the fact that many aspects of DUI are still unknown, there is evidence that DUI evolved from a strictly maternal inheritance (SMI) system ([Milani & Ghiselli, 2015](#); [Milani, Ghiselli & Passamonti, 2016](#)), by modifications of the molecular machinery involved in mitochondrial inheritance, through as-yet-unknown specific factors (see [Diz, Dudley & Skibinski, 2012](#); [Zouros, 2013](#) for proposed models). The detection of DUI is not a straightforward process, especially using PCR-based approaches: given that the divergence between F and M genomes is often comparable to the distance between mtDNAs of different classes of Vertebrates, primers may fail to amplify one of the two mtDNAs, yielding a false-negative result. Moreover, M-type mtDNA can be rare in somatic tissues, so it may be difficult to amplify from animals sampled outside of the reproductive season, when gonads are absent (thoroughly discussed in [Theologidis et al., 2008](#)). High-throughput sequencing (HTS) approaches can overcome such problems, because a prior knowledge of the mtDNA sequence is not needed, and low-copy variants can be easily unveiled (see [Ju et al., 2011](#); [King et al., 2014](#)). Until now, HTS has been scarcely

utilized to study mitochondrial transcriptomes and genomes ([Pesole et al., 2012](#); [Smith, 2013](#)), even if it showed very good potential ([Lubošny et al., 2017/2](#); [Yuan et al., 2016](#)). In this work we: (i) assembled *Ruditapes decussatus* mitochondrial genome from RNA-Seq data, and validated it by Sanger sequencing, (ii) analyzed and characterized *Ruditapes decussatus* mitochondrial genome, comparing its features with those of other venerid bivalves; (iii) assessed mitochondrial sequence polymorphism (SP) and structural variants—copy number variation (CNV) of tandem repeats—among the sampled animals.

## MATERIALS AND METHODS

### Sampling

The 26 *Ruditapes decussatus* specimens used in this study were collected from the Northern Adriatic Sea, in the river Po Delta region (Sacca di Goro, approximate GPS coordinates: 44°50'06"N, 12°17'55"E) at the end of July 2011, during the spawning season. Each individual was dissected, and gonadal liquid collected with a glass capillary tube. All the samples showed ripe gonads, consistently with the time of the year when the sampling occurred. The gonadal liquid was checked under a light microscope to assess the sex of the individual, and to make sure that the sample consisted of mature gametes. Both the gamete samples and the clam bodies were flash-frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ , until nucleic acid extraction. [Table S1](#) shows the sample list, and details about data availability.

### RNA-Seq

In total, 12 samples (six males and six females) were used for RNA-Seq. Total RNA extraction and library preparation were performed following the protocol described in [Mortazavi et al. \(2008\)](#), with the modifications specified in [Ghiselli et al. \(2012\)](#). The 12 samples were indexed, pooled and sequenced in two lanes (two technical replicates) of Illumina GA IIX, using 76 bp paired-end reads.

### De novo assembly

The mitochondrial genome of *Ruditapes decussatus* was not available in the databases, so we used the transcriptome data to generate a draft to be used as a guide for Sanger sequencing. Illumina reads from all 12 samples were pooled and compared to a set of 20 Bivalvia mitochondrial genomes to identify reads with mitochondrial origin. Alignment was done using BLASTN. All reads with similarity yielding  $E\text{-value} < 1\text{E-}5$  were then assembled into contigs using the A5 pipeline (version 2013032; [Tritt et al., 2012](#)) and joined into scaffolds using CAP3 ([Huang & Madan, 1999](#)). For the quality check step, we applied a PHRED Q-score cutoff threshold of 33; the other A5 parameters were set as default. CAP3 was run with default settings as well.

### Sanger validation

In total, 14 *Ruditapes decussatus* samples from the same collection campaign—sexed, and stored at  $-80^{\circ}\text{C}$ —were used for DNA extraction. DNA from the gonadic tissue was

extracted using the Qiagen DNeasy kit. Primers for mtDNA amplification were designed based on contigs obtained from RNA-Seq matching venerid mtDNA sequences, then the “primer walking” method was used to Sanger-sequence the complete mitochondrial genome of *Ruditapes decussatus*. The primers were designed with the software Primer3 (Rozen & Skaletsky, 2000) and tested on several samples, then a female was chosen as reference sample for Sanger validation of mtDNA *de novo* assembly. In addition, we amplified the largest unassigned region (LUR) of 13 females to assess its variability (see Results and Discussion). The list of the primers and their sequences are reported in Table S2. PCR reactions were performed in a final volume of 50  $\mu$ l using the GoTaq Flexi DNA Polymerase Kit (Promega, Madison, WI, USA), on a 2720 Thermal Cycler (Applied Biosystem, Foster City, CA, USA). The PCR reactions were set as follows: initial denaturation 95 °C for 1 min, then 30 cycles of amplification (denaturation 95 °C for 1 min, annealing 48–60 °C for 1 min, extension 72 °C for 1 min/kb), then the final extension at 72 °C for 5 min. PCR products were checked by electrophoretic run on 1% agarose gel, and then purified using the DNA Clean & Concentrator-25 kit (Zymo Research, Irvine, CA, USA).

Sanger sequencing was performed by Macrogen Inc. (<http://www.macrogen.com>).

Sequences were aligned with the software MEGA 6.0 (Tamura et al., 2013), using the contigs obtained by RNA-seq as a reference.

## Annotation

Open reading frames (ORFs) were identified with ORF finder (Wheeler et al., 2005). Alternative start codons were considered functional because they are common in Bivalvia. ORFs were annotated starting from the first available start codon (ATG, ATA, or ATC) downstream of the preceding gene, and ending with the first stop codon in frame (TAA or TAG). tRNA genes and their structure were identified with MITOS (Bernt et al., 2013) and ARWEN (Laslett & Canback, 2008). Secondary structures were predicted using the RNAFold Server, included in the ViennaRNA Web Services (<http://rna.tbi.univie.ac.at/>; Gruber et al., 2008); the folding temperature was set at 16 °C which is the average annual temperature of the water from which the *Ruditapes decussatus* specimens used in this work were fished (download RNAFold results from figshare: <https://ndownloader.figshare.com/files/8387672>). tRNAs and other secondary structures were drawn with the software Varna GUI (Darty, Denise & Ponty, 2009). Ribosomal small subunit (*rrnS*) and large subunit (*rrnL*) were identified with BLASTN, and annotated considering the start and the end of the adjacent genes as the boundaries of the rRNA genes. Non-genic regions were annotated as URs. In order to identify the putative D-loop/control region (CR), we analyzed the LUR with the MEME suite (Bailey et al., 2009) to find DNA motifs using the following bivalve species as comparison: *Acanthocardia tuberculata*, *Arctica islandica*, *Coelomactra antiquata*, *Fulvia mutica*, *Hiatella arctica*, *Loripes lacteus*, *Lucinella divaricata*, *Lutraria rhynchaena*, *Meretrix lamarckii* (F-type), *Meretrix lamarckii* (M-type), *Meretrix lusoria*, *Meretrix lyrata*, *Meretrix meretrix*, *Meretrix petechialis*, *Moerella iridescens*, *Nuttallia olivacea*, *Paphia amabilis*, *Paphia euglypta*, *Paphia textile*, *Paphia undulata*, *Ruditapes philippinarum* (F-type), *Ruditapes philippinarum* (M-type),



*Semele scabra*, *Sinonovacula constricta*, *Solecurtus divaricatus*, *Solen grandis*, *Solen strictus*, *Soletellina diphos* and the sea urchin *Strongylocentrotus purpuratus* (Echinoidea, Strongylocentrotidae). The list of the species used in the phylogenetic analysis and in the comparative analyses of DNA motifs, sequence similarity, and gene order are available in [Table S3](#). The GOMo (Gene Ontology for Motifs; [Buske et al., 2010](#)) tool of the MEME suite was used to assign GO terms to the motifs discovered.

The number of repeats in the LUR of the reference sample (F4) was calculated with tandem repeat finder (<http://tandem.bu.edu/trf/trf.html>), since the complete LUR sequence was available (download tandem repeat finder results from Figshare: <https://ndownloader.figshare.com/files/8387666>). In the other cases, in which the LUR could not be sequenced without gaps, the number of repeats was inferred from agarose gel electrophoresis.

### Other analyses

Comparisons among venerid complete mtDNAs were performed with BLAST Ring Image Generator (BRIG) ([Alikhan et al., 2011](#)) and Easyfig ([Sullivan, Petty & Beatson, 2011](#)). Descriptive statistics were obtained with MEGA v6.0 ([Tamura et al., 2013](#)), except for the codon usage table, which was obtained with the Sequence Manipulation Suite ([Stothard, 2000](#)). SP assessment from RNA-Seq reads was performed with the Genome Analysis Toolkit (GATK, [McKenna et al., 2010](#)), with the Sanger-sequenced mtDNA as reference. For SP discovery and genotyping we used standard hard filtering parameters or variant quality score recalibration ([DePristo et al., 2011](#)). The MitoPhast pipeline ([Tan et al., 2015](#)) was used to obtain the Maximum Likelihood (ML) tree, which was visualized with Evolview v2 ([He et al., 2016](#)). Briefly, MitoPhast takes as input GenBank files (.gb), extracts the coding sequences, profiles the sequences with Pfam ([Finn et al., 2016](#)) and PRINTS ([Attwood et al., 2003](#)), performs a multiple sequence alignment with Clustal Omega ([Sievers et al., 2011](#)), removes poorly aligned regions with trimAl ([Capella-Gutiérrez, Silla-Martinez & Gabaldón, 2009](#)), concatenates the coding sequences, performs data partitioning and model selection, and then carries out a ML analysis using RAxML ([Stamatakis, 2014](#)). The species used in the ML analysis, and their GenBank Accession Numbers are listed in [Table S3](#). Amino acid sequences of three different *cox3* ORFs inferred from Sanger sequencing and GATK polymorphism data were analyzed with InterProScan ([Jones et al., 2014](#)).

## RESULTS

### De novo assembly and Sanger validation

Despite using HTS on extracts of ripe gonads (i.e., mature gametes), and multiple assembly strategies (see Discussion for details) we could not find evidence for DUI. The de novo assembly process produced 9 contigs, of which 8 included multiple genes, and one included a single gene (see [Table 1](#)). The sequences of the contigs in FASTA format are available on figshare (<https://ndownloader.figshare.com/files/8906839>). In four cases (Contigs 1, 3, 6, and 7) a clear polyadenylation signal was present, in other four cases (Contigs 2, 5, 8, and 9) it was not. Contig 4, the only one including a single gene (*cox3*), ends with just 8 As, so it is not clear if a polyadenylation signal is present in this case.

**Table 1** Features of the contigs obtained by *de novo* assembly of mtDNA.

Contig	Length	Gene content	Poly-A	Notes
1	6,794	<i>atp6_nd3_nd5_cox1_tRNA-Leu1_nd1_nd2_nd4L</i>	Yes	Chimeric assembly. The contiguity between <i>nd5</i> and <i>cox1</i> is an artifact
2	1,884	<i>rrnS_cox3</i>	No	–
3	1,288	<i>atp6_nd3</i>	Yes	–
4	1,663	<i>cox3</i>	?	The contig ends with just 8 As
5	1,934	<i>atp8_nd4_tRNA-His_tRNA-Glu_tRNA-Ser2_tRNA-Tyr</i>	No	–
6	1,831	<i>atp8_nd4_tRNA-His</i>	Yes	–
7	5,478	<i>cox2_tRNA-Ile_nd4L_nd2_nd1_tRNA-Leu1_cox1</i>	Yes	There is a polyadenylation signal (56 As) after the <i>cox2</i> gene
8	2,879	<i>cytb_rrnL</i>	No	–
9	952	<i>nd6_tRNA-Lys_tRNA-Val_tRNA-Phe_tRNA-Trp_tRNA-Arg_tRNA-Leu2</i>	No	–

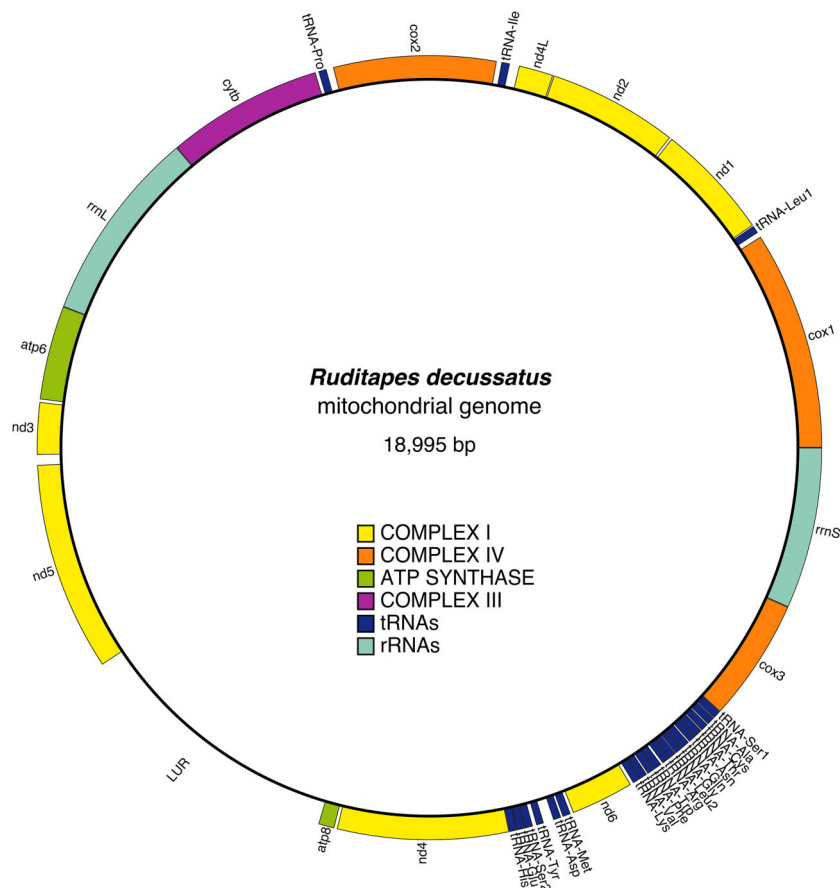
In Contig 7 (that includes *cox2*, *tRNA-Ile*, *nd4L*, *nd2*, *nd1*, *tRNA-Leu1*, and *cox1* genes) there is a polyadenylation signal (56 As) after the *cox2* gene.

The nine contigs were used as a scaffold for the primer walking procedure used for Sanger validation of the *de novo* assembly. We first tried to connect the contigs designing primers close to the 5' and 3' ends of each contig and pairing them following the gene order of *Paphia*, because the sequence of genes in the contigs suggested that *Ruditapes decussatus* gene order might have been similar. During such process, Contig 1 turned out to be a chimeric assembly between two non-contiguous portions of the mtDNA, one including *atp6*, *nd3*, and *nd5*, the other including *cox1*, *tRNA-Leu1*, *nd1*, *nd2*, and *nd4L*. Once we amplified and sequenced the portions of mtDNA between the contigs, we proceeded with the Sanger resequencing of the remaining parts.

### Annotation and mtDNA features

The mitochondrial genome contains 13 protein-coding genes, and in the reference female is 18,995 bp long (Fig. 1); the gene arrangement and other details are shown in Table 2. All genes are located on the heavy strand, and in addition to the classic start codon ATG (Met), the alternative start codons ATA (Met) and ATC (Ile) are present. The most frequently used start codons are: ATA (*cox1*, *nd1*, *nd4L*, *cox2*, *cob*, *atp8*, *nd4*), and ATG (*nd2*, *atp6*, *nd3*, *nd5*, *nd6*, *cox3*). The stop codons found are TAG (*cox1*, *nd2*, *nd4L*, *cox2*, *cytb*, *nd4*) and TAA (*nd1*, *atp6*, *nd3*, *atp8*, *nd6*). The *nd4* gene has an incomplete stop codon (TA-). 22 tRNA genes were identified, including two tRNAs for leucine, tRNA-Leu1 (TAG) and tRNA-Leu2(TAA), and two for serine, tRNA-Ser1(TCT) and tRNA-Ser2 (TGA), both showing degenerate D-arm branches. tRNA structures are shown in Fig. S1. The two rRNAs, *rrnS* and *rrnL*, were both identified: the *rrnS* is located between *cox3* and *cox1*, while *rrnL* is between *cytb* and *atp6*. URs were identified on the basis of unannotated spaces between different genes; we found 24 URs (Table 3).

The analysis of the nucleotide composition points out that the mitochondrial genome of this bivalve species exhibits high A+T content, totaling 63% vs 37% G+C. The



**Figure 1** *Ruditapes decussatus* mtDNA gene arrangement.

minimum values of A+T are found in *cytb* (60.1%) and *nd4* (61%). The nucleotide composition of every gene is shown in Table 4. According to the analysis above, both A and T occur very frequently at the third position of codons (64.6% on average of A+T), while the less frequent base in third position is C (12%). The most used codons are UUU (Phe), counted 269 times, and UUA (Leu) counted 210 times (6.78% and 5.29% of the total, respectively), while the less used codons are CGC (Arg) counted 6 times (0.15%), ACC (Thr) and CCG (Pro) each counted 16 times (0.4%) (Table 5). Only in four cases over 20 (Lys, Leu, Gln, Val), the most frequently used codon matches the correspondent mitochondrial tRNA anticodon.

The UR11 is the LUR and is located between *atp8* and *nd5* (Figs. 1 and 2A). The LUR of the female used for whole mtDNA Sanger sequencing (i.e., the reference female, F4) is 2,110 bp long, and includes 6.5 repeated sequences—each repeat having a length of 54 bp—localized in the 3' region of the LUR, just upstream the *atp8* gene (Fig. 2A). DNA secondary structure analysis predicted three stem-loop structures in such region (Fig. 2B and Supplemental Information files on figshare: <https://ndownloader.figshare.com/files/8387672>), with a change in Gibbs free energy ( $\Delta G$ ) of  $-71.38$  Kcal/mol. We also amplified and sequenced the LUR of 13 more females. We were not able to completely sequence LURs longer than 2,110 bp, because of the known difficulties in Sanger sequencing of



**Table 2** MtDNA gene arrangement of *Ruditapes decussatus*.

Name	Type	Start	Stop	Length (bp)	Start	Stop	Anticodon
<i>cox1</i>	Coding	1	1,716	1,716	ATA	TAG	
<i>tRNA-Leu1</i>	tRNA	1,754	1,815	62			TAG
<i>nd1</i>	Coding	1,822	2,739	918	ATA	TAA	
<i>nd2</i>	Coding	2,755	3,774	1,020	ATG	TAG	
<i>nd4l</i>	Coding	3,780	4,052	273	ATA	TAG	
<i>tRNA-Ile</i>	tRNA	4,125	4,190	66			GAT
<i>cox2</i>	Coding	4,228	5,499	1,272	ATA	TAG	
<i>tRNA-Pro</i>	tRNA	5,553	5,616	64			TGG
<i>cytb</i>	Coding	5,641	6,864	1,224	ATA	TAG	
<i>rrnL</i>	rRNA	6,865	8,385	1,521			
<i>atp6</i>	Coding	8,386	9,123	738	ATG	TAA	
<i>nd3</i>	Coding	9,145	9,552	408	ATG	TAA	
<i>nd5</i>	Coding	9,631	11,268	1,638	ATG	TAG	
<i>atp8</i>	Coding	13,379	13,504	126	ATA	TAA	
<i>nd4</i>	Coding	13,526	14,865	1,340	ATA	TA-	
<i>tRNA-His</i>	tRNA	14,866	14,928	63			GTG
<i>tRNA-Glu</i>	tRNA	14,929	14,990	62			TTC
<i>tRNA-Ser2</i>	tRNA	14,991	15,052	62			TGA
<i>tRNA-Tyr</i>	tRNA	15,081	15,140	60			GTA
<i>tRNA-Asp</i>	tRNA	15,218	15,280	63			GTC
<i>tRNA-Met</i>	tRNA	15,294	15,358	65			CAT
<i>nd6</i>	Coding	15,380	15,874	495	ATG	TAA	
<i>tRNA-Lys</i>	tRNA	15,897	15,959	63			TTT
<i>tRNA-Val</i>	tRNA	15,960	16,021	62			TAC
<i>tRNA-Phe</i>	tRNA	16,030	16,092	63			GAA
<i>tRNA-Trp</i>	tRNA	16,093	16,155	63			TCA
<i>tRNA-Arg</i>	tRNA	16,171	16,232	62			TCG
<i>tRNA-Leu2</i>	tRNA	16,233	16,295	63			TAA
<i>tRNA-Gly</i>	tRNA	16,297	16,358	62			TCC
<i>tRNA-Gln</i>	tRNA	16,359	16,427	69			TTG
<i>tRNA-Asn</i>	tRNA	16,435	16,497	63			GTT
<i>tRNA-Thr</i>	tRNA	16,498	16,560	63			TGT
<i>tRNA-Cys</i>	tRNA	16,565	16,626	62			GCA
<i>tRNA-Ala</i>	tRNA	16,632	16,696	65			TGC
<i>tRNA-Ser1</i>	tRNA	16,698	16,764	67			TCT
<i>cox3</i>	Coding	16,765	17,730	966	ATG	TAA	
<i>rrnS</i>	rRNA	17,731	18,995	1,265			

**Note:**

The anticodon of tRNAs are reported in the 5'-3' direction.

**Table 3** Unassigned regions (URs).

UR name	Start	Stop	Length (bp)
UR1	1,717	1,753	37
UR2	1,816	1,821	6
UR3	2,740	2,754	15
UR4	3,775	3,779	5
UR5	4,053	4,124	72
UR6	4,191	4,227	37
UR7	5,500	5,552	53
UR8	5,617	5,640	24
UR9	9,124	9,144	21
UR10	9,553	9,630	78
UR11 (LUR)	11,269	13,378	2,110
UR12	13,505	13,525	21
UR13	15,053	15,080	28
UR14	15,141	15,217	77
UR15	15,281	15,293	13
UR16	15,359	15,379	21
UR17	15,875	15,896	22
UR18	16,022	16,029	8
UR19	16,156	16,170	15
UR20	16,296	16,296	1
UR21	16,428	16,434	7
UR22	16,561	16,564	4
UR23	16,627	16,631	5
UR24	16,697	16,697	1

regions including multiple repeats. The sequence alignment of the 13 LURs is available for download from figshare (<https://ndownloader.figshare.com/files/8360789>). LUR lengths, inferred from gel electrophoresis, are reported in Table 6, and they range from 2,000 to 5,000 bp. Two females (F3 and F17) showed length heteroplasmy of the LUR. The portion of the genome occupied by URs varies between 14.11% and 29.38%, depending on LUR length. The analysis with MEME (output shown in Figs. S2 and S3) unveiled two motifs (Fig. 2C) that show a strong conservation within the Veneridae family, and with *S. purpuratus*. The sea urchin was included in the analysis because Cao et al. (2004) reported a match between some motifs found in the CR of the marine mussels *Mytilus edulis* and *Mytilus galloprovincialis* with regulatory elements of the sea urchin CR. Accordingly, the search with GOMo assigned a series of GO terms related to transcription to the two motifs (Table S4).

### Polymorphism

Table 7 (top) shows the statistics associated with the SP analysis performed with GATK on the 12 samples used for RNA-Seq, with the Sanger-sequenced mtDNA as reference. Overall, 257 SPs were called, of which 145 (56.4%) were located in coding sequences

**Table 4** Nucleotide composition.

Name	Length (bp)	T (%)	C (%)	A (%)	G (%)	A+T (%)	T3 (%)	C3 (%)	A3 (%)	G3 (%)	A3+T3 (%)
<i>cox1</i>	1,716	35.8	15.5	25.8	22.9	61.6	39	12.1	28.0	21.3	67.0
<i>nd1</i>	918	38.7	12.5	24.0	24.8	62.7	38	10.1	30.7	21.2	68.7
<i>nd2</i>	1,020	38.3	11.0	24.8	25.9	63.1	35	11.5	29.4	24.4	64.4
<i>nd4l</i>	273	39.9	12.8	25.3	22.0	65.2	34	14.3	30.8	20.9	64.8
<i>cox2</i>	1,272	29.7	14.8	29.1	26.4	58.8	30	15.3	27.4	27.6	57.4
<i>cob</i>	1,224	37.4	17.2	22.7	22.6	60.1	41	14.7	21.8	22.1	62.8
<i>rrnL</i>	1,749	33.2	11.5	32.6	22.6	65.8	33	10.6	33.4	23.0	66.4
<i>atp6</i>	510	42.0	15.7	20.8	21.6	62.8	45	13.5	21.8	20.0	66.8
<i>nd3</i>	408	39.5	11.0	24.8	24.8	64.3	33	11.0	30.1	25.7	63.1
<i>nd5</i>	1,638	37.6	11.7	27.7	23.0	65.3	35	11.0	34.2	19.8	69.2
<i>atp8</i>	126	44.4	11.9	19.0	24.6	63.4	45	4.8	23.8	26.2	68.8
<i>nd4</i>	1,340	38.9	12.9	22.1	26.1	61.0	41	10.8	24.9	23.5	65.9
<i>nd6</i>	495	39.2	12.1	23.0	25.7	62.2	38	13.9	27.9	20.0	65.9
<i>cox3</i>	966	36.9	12.7	24.8	25.6	61.7	39	9.6	28.6	23.0	67.6
<i>rrnS</i>	1,265	32.7	12.3	32.9	22.1	65.6	35	13.5	31.6	19.5	66.6
All coding	14,920	36.3	13.2	26.5	24.0	63.0	37	12.0	28.9	22.4	65.7
All <i>rRNAs</i>	3,014	32.9	23.8	32.7	22.3	65.7					
All <i>tRNAs</i>	1,394	35.4	12.8	30.2	21.7	65.6					
All URs	2,681	28.2	14.1	34.1	23.6	62.3					
All genic DNA	16,314	36.2	13.2	26.8	23.8	63.0					
All DNA	18,995	35.1	13.3	27.9	23.7	63.0					

**Note:**

URs, unassigned regions.

(CDS). Interestingly, most of the SPs were called because of private alleles of one single male specimen (mRDI01). More in detail, 151 SPs out of 257 (58.7%) along the whole mtDNA sequence, and 103 SPs out of 145 (71%) in CDS, were private of mRDI01. In CDS, if we exclude the SPs associated with this male, the number of polymorphisms drops to 42 over 14,920 bp of coding mtDNA (GATK output in VCF format and a detailed list of SPs in tabular format is available on figshare: <https://ndownloader.figshare.com/files/8902537>), of which 18 are represented by indels, 6 of which are located in 4 different coding genes: one each in *cox1*, *cytb*, and *nd5*, plus 3 in *cox3* (see Table 8). A file showing the ORF generated by the different variants of *cox3*, and alignments between them is available on figshare (<https://ndownloader.figshare.com/files/8402471>). Table 7 (bottom) shows the number of SPs in males, in males except mRDI01, and in females both along the whole mtDNA, and in CDS. The number in brackets represent the number of private SPs for each category.

### Comparison with other veneridae

Figure 3 shows the *Ruditapes decussatus* mtDNA map (external gray circle), and the BLASTN identity (colored inner circles) with complete mtDNAs of other 10 venerid species (see list in Table S3). Figure 4 shows the ML tree obtained with the MitoPhast

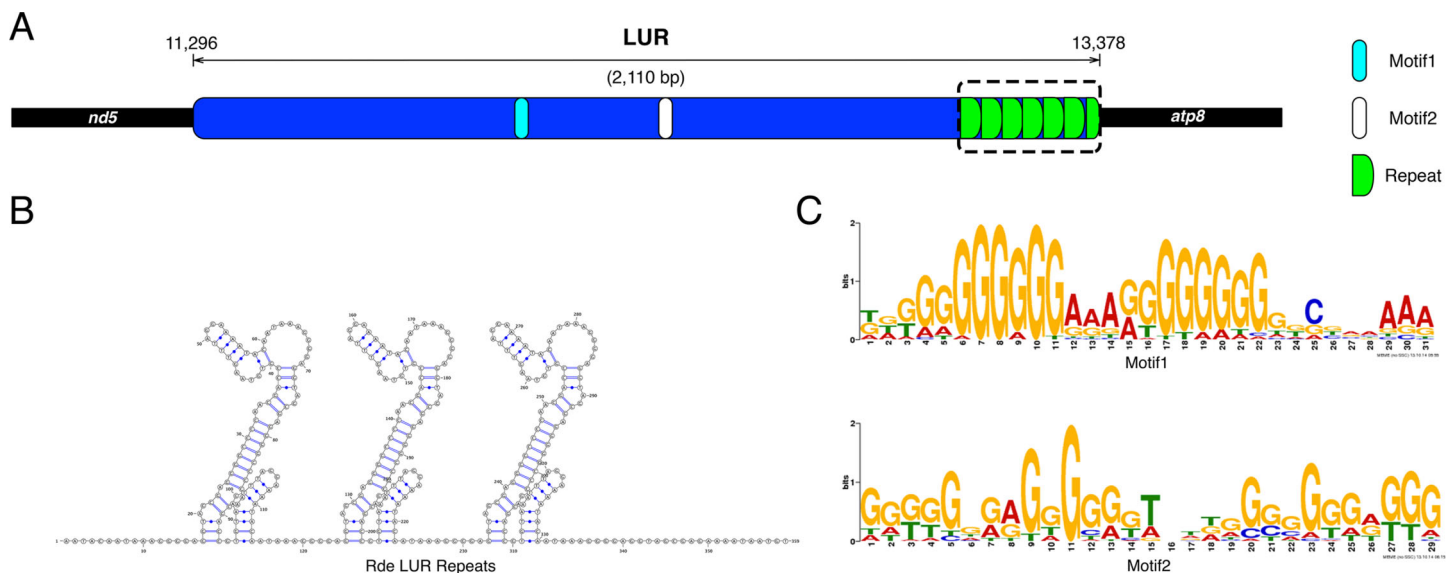
Table 5 Codon usage.

Amino acid	Codon	#	Frequency	%TOT	Amino acid	Codon	#	Frequency	%TOT
Ala	GCG	29	0.15	0.73	Pro	CCG	16	0.12	0.40
	<u>GCA</u>	44	0.23	1.11		<u>CCA</u>	36	0.27	0.91
	GCT	85	<u>0.45</u>	2.14		CCT	58	<u>0.43</u>	1.46
	GCC	30	0.16	0.76		CCC	24	0.18	0.61
Cys	TGT	94	<u>0.76</u>	2.37	Gln	CAG	25	0.44	0.63
	<u>TGC</u>	30	0.24	0.76		<u>CAA</u>	32	<u>0.56</u>	0.81
Asp	GAT	54	<u>0.66</u>	1.36	Arg	CGG	23	0.31	0.58
	<u>GAC</u>	28	0.34	0.71		<u>CGA</u>	21	0.28	0.53
Glu	GAG	87	<u>0.6</u>	2.19		CGT	25	<u>0.33</u>	0.63
	<u>GAA</u>	58	0.4	1.46		CGC	6	0.08	0.15
Phe	TTT	269	<u>0.78</u>	6.78	Ser	AGG	69	0.19	1.74
	<u>TTC</u>	78	0.22	1.97		<u>AGA</u>	69	0.19	1.74
Gly	GGG	131	<u>0.4</u>	3.30		AGT	55	0.15	1.39
	<u>GGA</u>	61	0.19	1.54		AGC	23	0.06	0.58
	GGT	98	0.3	2.47		TCG	18	0.05	0.45
	GGC	36	0.11	0.91		<u>TCA</u>	33	0.09	0.83
His	CAT	37	<u>0.62</u>	0.93		TCT	76	<u>0.21</u>	1.92
	<u>CAC</u>	23	0.38	0.58		TCC	22	0.06	0.55
Ile	ATT	165	<u>0.8</u>	4.16	Thr	ACG	21	0.17	0.53
	<u>ATC</u>	40	0.2	1.01		<u>ACA</u>	30	0.24	0.76
Lys	AAG	61	0.41	1.54		ACT	57	<u>0.46</u>	1.44
	<u>AAA</u>	87	<u>0.59</u>	2.19		ACC	16	0.13	0.40
Leu	TTG	122	0.23	3.08	Val	GTG	113	0.3	2.85
	<u>TTA</u>	210	<u>0.39</u>	5.29		<u>GTA</u>	121	<u>0.32</u>	3.05
	CTG	43	0.08	1.08		GTT	119	0.32	3.00
	<u>CTA</u>	70	0.13	1.76		GTC	23	0.06	0.58
	CTT	75	0.14	1.89		Trp	TGG	58	<u>0.54</u>
CTC	20	0.04	0.50	<u>TGA</u>	49		0.46	1.24	
Met	<u>ATG</u>	86	0.36	2.17	Tyr	TAT	103	<u>0.69</u>	2.60
	ATA	155	<u>0.64</u>	3.91		<u>TAC</u>	47	0.31	1.18
Asn	AAT	76	<u>0.66</u>	1.92	STOP	TAG	34	0.58	0.86
	<u>AAC</u>	39	0.34	0.98		TAA	25	0.42	0.63

**Note:**

The codons corresponding to a tRNA present in the mitochondrial genome are underlined and in bold. The highest frequency among synonymous codons is also underlined and in bold. #, number of codons; Frequency, frequency of each codon among synonymous codons; %TOT, frequency of each codon among all the codons.

pipeline; the complete input and output of this analysis is available on figshare (<https://ndownloader.figshare.com/files/8360792>). Figure 5 shows the variation in gene order between *Ruditapes decussatus* and *P. euglypta* (Fig. 5A), *M. lamarckii* F-type (Fig. 5B), *Ruditapes philippinarum* F-type (Fig. 5C), and among all the four species (Fig. 5D).



**Figure 2** Principal features of the Largest Unassigned Region (LUR). (A): map of the LUR; (B): DNA secondary structure predicted in the repeat region (boxed in A); (C): Logos of the two DNA motifs found in the LUR.

**Table 6** LUR length and number of repeats in the 13 female samples analyzed.

Specimen	Length (bp)	Number of repeats	GenBank Acc. No.
F3	2,100–3,500	6.5–25	MF055702
F5	5,000	45	MF055703
F7	3,500	25	MF055704
F9	3,500	25	MF055705
F10	3,000	20	MF055706
F11	3,000	20	MF055707
F13	3,500	25	MF055708
F15	3,000	20	MF055709
F16	3,500	25	MF055710
F17	2,500–3,500	8–25	MF055711
F19	3,500	25	MF055712
F20	2,500	8	MF055713
F21	2,100	6.5	MF055714

**Note:**

F3 and F17 are heteroplasmic with LURs of different length.

## DISCUSSION

### RNA-Seq-guided sequencing of mtDNA

The *de novo* assembly of the mtDNA from RNA-Seq data turned out to be informative, simplifying the primer walking procedure used for Sanger sequencing. Only one contig (Contig 1) resulted to be a chimeric sequence obtained by the misassembly of two smaller contigs. Most of the contigs (eight out of nine) contained more than one gene, and most of the tRNA genes were included in the *de novo* assembly. Except for *tRNA-Pro*,

**Table 7** Sequence Polymorphism (SP): SPs and small indels called by GATK.

Feature	Value	Min	Median	Mean	Max
Depth (all SPs)	–	6	1,357	1,521	3,880
Phred score (all SPs)	–	3.30E+01	5.76E+03	4.18E+07	2.15E+09
Depth (SPs in CDS)	–	222	2,038	2,150	3,880
Phred score (SPs in CDS)	–	1.18E+02	1.01E+04	4.45E+07	2.15E+09
Total number of SPs	257	–	–	–	–
Number of mRDI01 private SPs	151 (58.7% of the total)	–	–	–	–
Number of SPs in CDS	145 (56.4% of the total)	–	–	–	–
Number of mRDI01 private SPs in CDS	103 (71% of the SP in CDS)	–	–	–	–
Number of SPs in CDS (excluding mRDI01)	42	–	–	–	–
Frequency of SPs in CDS	0.0097 (~1 every 103 bp)	–	–	–	–
Frequency of SPs in CDS (excluding mRDI01)	0.0028 (~1 every 355 bp)	–	–	–	–
Total number of indels	18	–	–	–	–
Number of indels in CDS	6	–	–	–	–
Number of indels causing frameshift	4	–	–	–	–
<b># Of SPs</b>	<b>Whole mtDNA</b>	<b>CDS</b>			
Males	234 (160)	136 (107)			
Males (no mRDI01)	84 (15)	32 (6)			
Females	97 (23)	38 (9)			

**Note:**

CDS, coding sequences; Whole mtDNA, polymorphism in the whole mitochondrial genome; the number in brackets the bottom of the table represent private SPs (e.g., there are 23 female specific SPs in the whole mtDNA and 9 female specific SPs in CDS); *p*-value, significance of the Fisher's exact test on number of SPs between sexes (i.e., all males vs females, males except mRDI01 vs females).

**Table 8** Indels located in coding sequences.

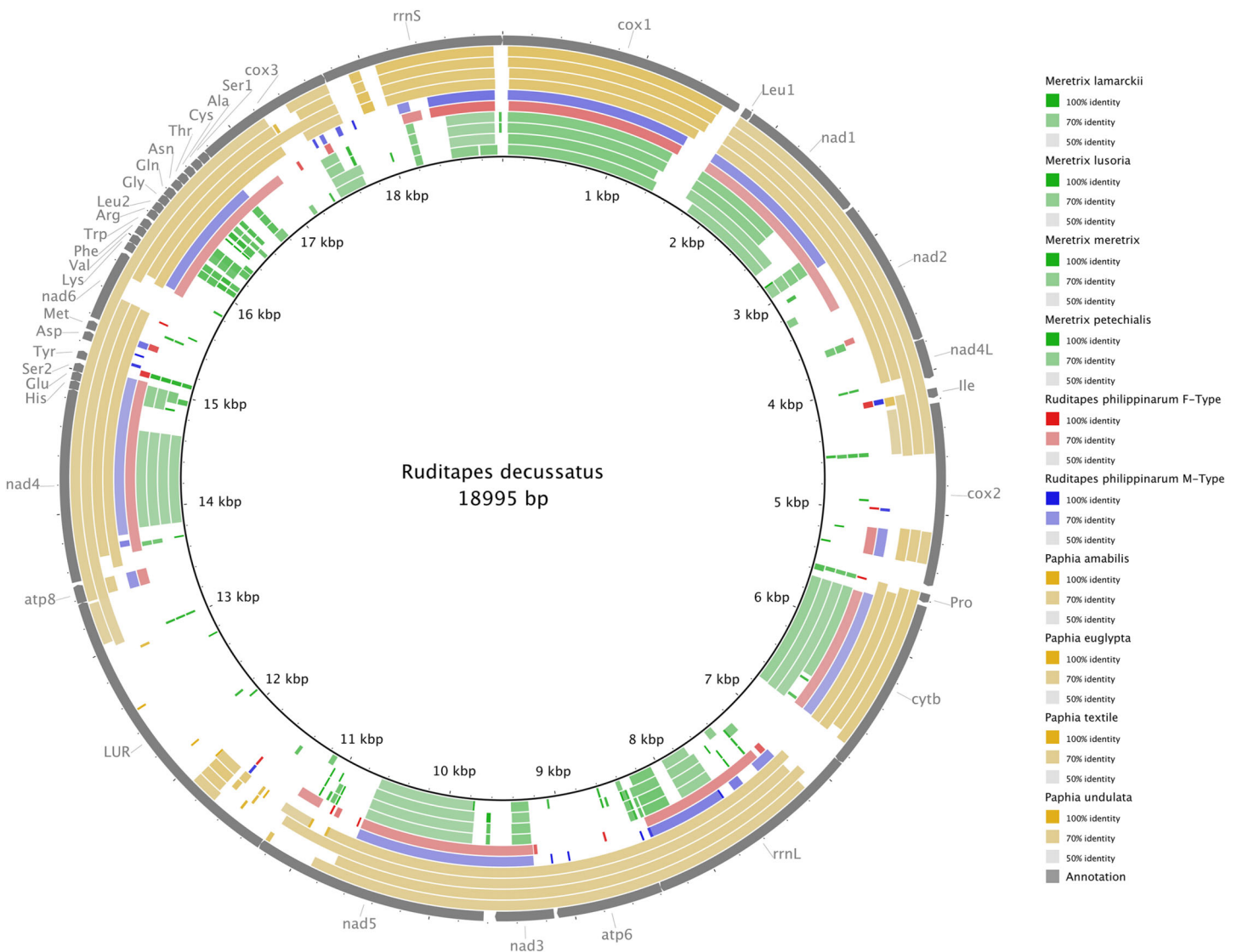
Position	Depth	Qual	Gene	SP	Frameshift	Sample	Allele frequency	Notes
1,698	3,732	1.38E+04	<i>cox1</i>	C/CAAA	No	mRDI02, mRDI03	0.089, 0.85	Insertion of 1 Lysine
6,364	1,929	2.15E+09	<i>cytb</i>	CT/C	Yes	fRDI04, mRDI05	0.80, 0.81	Yields a shorter <i>Cytb</i> . Possible sequencing error due to the homopolymer CTTTTTT
10,449	1,780	2.15E+09	<i>nd5</i>	C/CT	Yes	fRDI01, fRDI04, fRDI05	0.11, 0.10, 0.11	Yields a <i>nd5</i> gene divided in 2 ORFs. Possible sequencing error due to the homopolymer CTTTTTT
17,619	2,272	5.98E+03	<i>cox3</i>	AGCG/A	No	mRDI01	0.97	Deletion of one Alanine
17,621	2,188	9.99E+04	<i>cox3</i>	CG/C	Yes	mRDI01	0.99	Always combined with SP_17624. Together change the last 35 amino acids
17,624	2,287	5.98E+03	<i>cox3</i>	C/CAT	Yes	mRDI01	0.99	Always combined with SP_17621. Together change the last 35 amino acids

**Note:**

Depth, sequencing depth; Qual, quality of the called SP expressed in Phred score; Allele frequency, frequency of the alternative allele in each sample indicated in the "Sample" column.

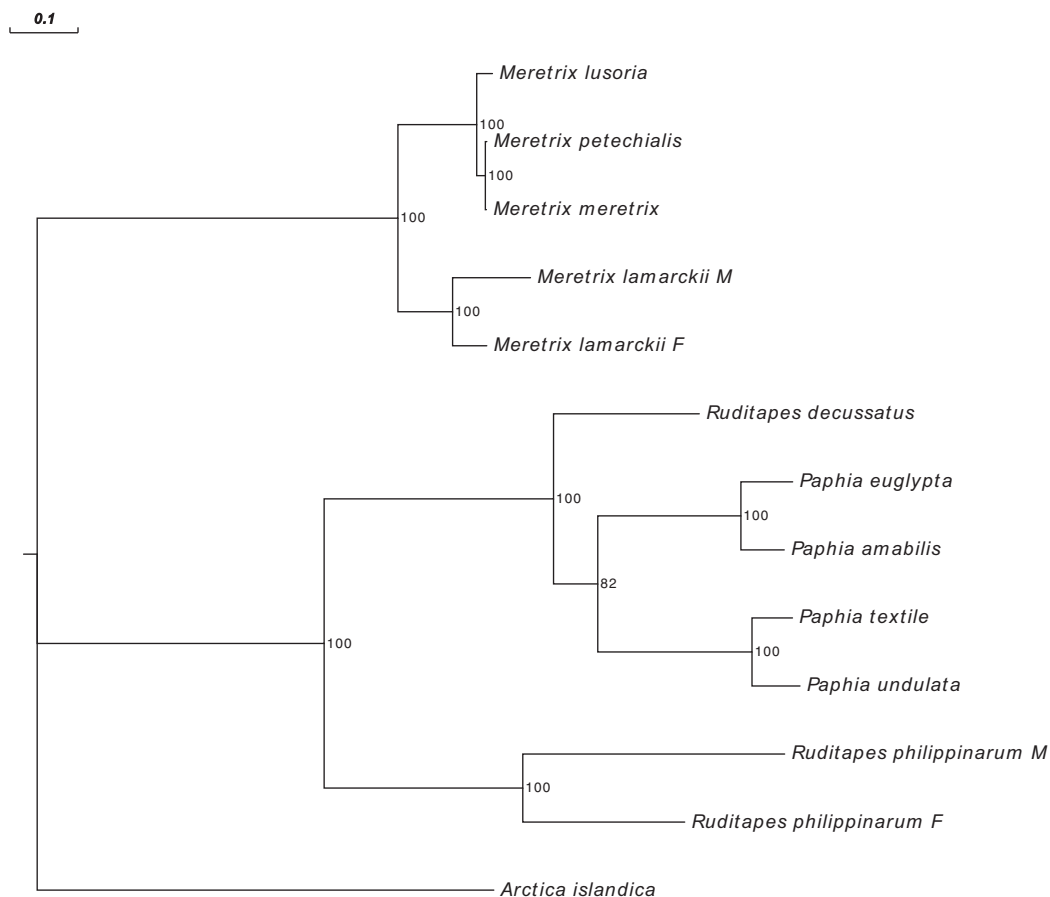
*tRNA-Ile*, and *tRNA-Leu1*, all the other tRNA genes are organized in two big clusters: a 13-gene cluster positioned between *cox3* and *nd6*, and a 6-gene cluster between *nd6* and *nd4*. The assembly retrieved 6 out of 13 tRNAs from the first cluster (missing *tRNA-Gly*, *tRNA-*





**Figure 3** BLASTN comparison of *Ruditapes decussatus* and other Veneridae. *Ruditapes decussatus* mtDNA map (external gray circle), and BLASTN identity (colored inner circles) with complete mtDNAs of other 10 venerid species (see list in Table S3).

*Glu*, *tRNA-Asn*, *tRNA-Thr*, *tRNA-Cys*, *tRNA-Ala*, and *tRNA-Ser1*), and 4 out of 6 tRNAs from the second cluster (missing *tRNA-Met* and *tRNA-Asp*). All the tRNA genes not located in these two clusters (*tRNA-Pro*, *tRNA-Ile*, and *tRNA-Leu1*) were included in the contigs. The presence of a clear polyadenylation signal in four of the assembled contigs (see Table 1) seems to indicate the existence of multiple polycistronic transcripts. It is also noteworthy that poly-A sequences seem to be absent in contigs having tRNA or rRNA genes at one end (Contigs 2, 5, 8 and 9). This could be either an evidence supporting the “tRNA punctuation model” of RNA processing proposed by *Ojala, Montoya & Attardi (1981)* for human mitochondria, or a result of difficulties in sequencing/assembly of such regions. More analyses are required to address this point.

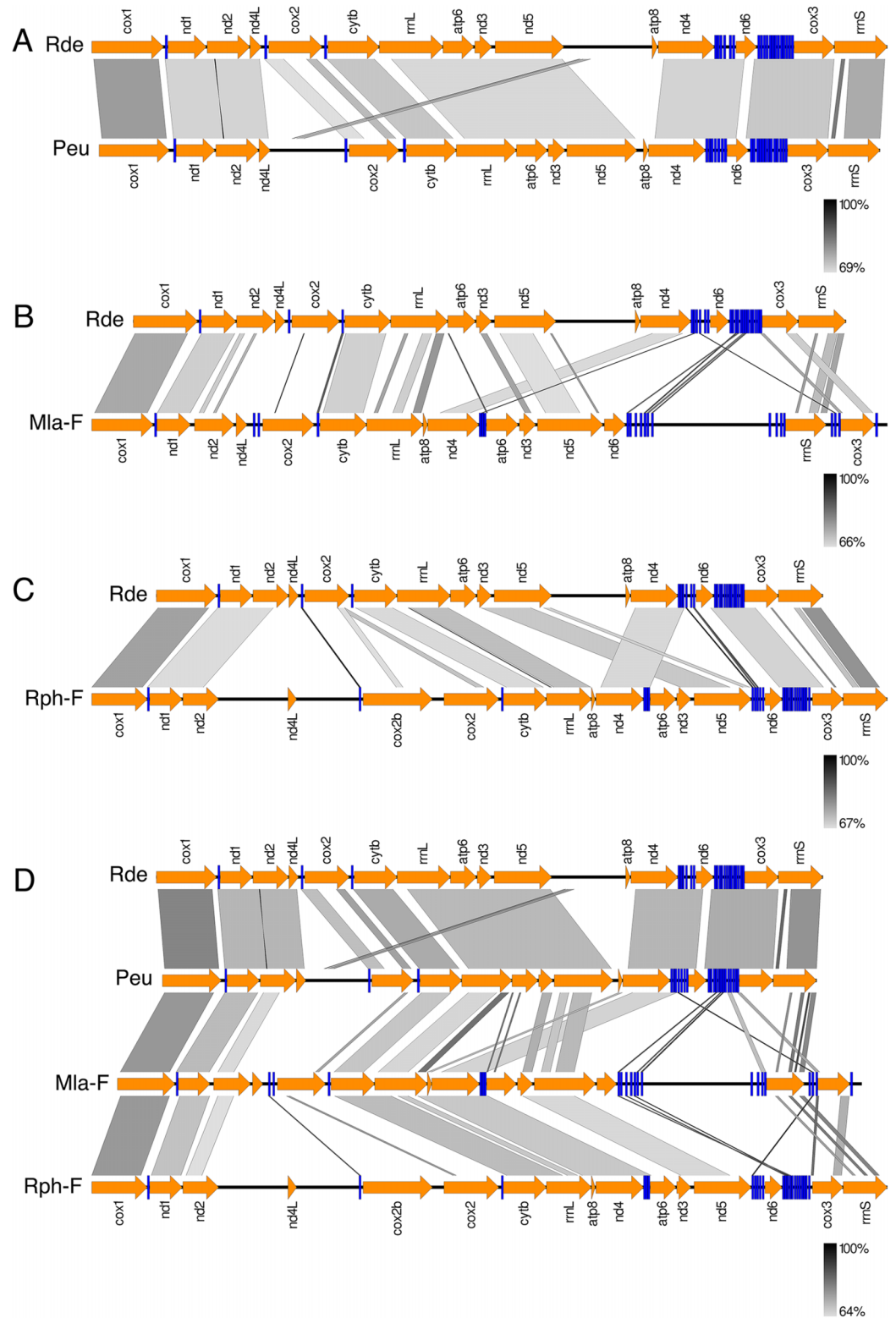


**Figure 4** Maximum Likelihood (ML) tree of Veneridae obtained with all mitochondrial coding genes. ML tree obtained with the MitoPhast pipeline; the complete input and output of this analysis is available on figshare (<https://doi.org/10.6084/m9.figshare.4970762.v1>).

### General features

The size of the fully Sanger-sequenced mitochondrial genome of *Ruditapes decussatus* (reference female F4) is of 18,995 bp, and it includes 13 protein-coding genes, 22 tRNAs and 2 rRNAs. Our data support the presence of the *atp8* gene in the mtDNA of *Ruditapes decussatus*; *atp8* has been reported as missing in several bivalve species, however, more accurate searches often led to the identification of the gene, so, in most cases, the alleged lack of *atp8* is likely ascribable to annotation inaccuracies due to the extreme variability and the small size of the gene (Breton, Stewart & Hoeh, 2010; Breton et al., 2014; Plazzi, Puccio & Passamonti, 2016).

The mitochondrial genome of *Ruditapes decussatus* shows a high content of A-T (63%), a common feature in bivalve mtDNAs; moreover, T is the most common nucleotide at the third codon base (64.6%). The most common codon is UUU (Phe), which is also the most commonly used in bivalves, as well as in other invertebrates (Passamonti et al., 2011).



**Figure 5** Comparison of gene order in venerid mtDNAs. Variation in gene order between *Ruditapes decussatus* and *P. euglypta* (A) *M. lamarckii* F-type (B) *Ruditapes philippinarum* F-type (C) and among all the four species (D).

## Codon usage

As shown in Table 5, in 16 cases out of 20, the most frequently used codon does not correspond to the anticodon of the inferred tRNA. In other words, there is not a correspondence between the most abundant codons and the anticodons of the 22 mitochondrial tRNAs. According to the “wobble hypothesis”—first proposed by Crick (1966)—the conformation of the tRNA anticodon loop enables some flexibility at the first base of the anticodon, so a Watson–Crick type of base pairing in the third position of the codon is not strictly necessary. This allows an amino acid to be correctly incorporated by ribosomes even if the tRNA is not fully complementary to the codon; according to Crick, this explains the degeneracy of the genetic code. This feature is particularly interesting in the light of the debate about natural selection acting at synonymous sites: since the early 1980s, evidence of a correlation between synonymous codon usage and tRNA abundances started accumulating. According to these authors, synonymous codon usage is biased to match skews in tRNA abundance, as a result of selective pressure maximizing protein synthesis rates (reviewed in Chamary, Parmley & Hurst, 2006). Following this rationale, the results here reported and data from other marine bivalves and metazoans (Yu & Li, 2011; Passamonti et al., 2011) would suggest that in some mitochondrial genomes translation efficiency is not maximized, and this observation deserves further investigation.

## Length and sequence polymorphism

The mtDNA of *Ruditapes decussatus* has a high proportion of URs mostly depending on the length of the LUR (Table 6); on average, bivalve mtDNAs have 1.7× the amount of URs in respect to other analyzed Metazoa (Ghiselli et al., 2013), and it is still unclear whether there is an accumulation of non-functional sequences in bivalve mtDNA due to genetic drift, or if such URs are maintained by natural selection because they contain—so far unknown—functional elements (see Milani et al., 2013, 2014b; Breton et al., 2014; Pozzi et al., 2017). The LUR of *Ruditapes decussatus* most likely includes the mitochondrial CR, as indicated by the presence of two motifs (Fig. 2C; Figs S2 and S3) similar to two regulatory elements identified in the sea urchin CR. These two motifs are the same identified in previous analyses on the clam *Ruditapes philippinarum* and the mussel *Musculista senhousia* (Ghiselli et al., 2013; Guerra, Ghiselli & Passamonti, 2014) so they are conserved across distant bivalve taxa, and the GO terms associated with such motifs are related to transcriptional control (Table S4). An interesting feature of *Ruditapes decussatus* LUR is its variable length (Table 6), most likely due to different repeat content. As a matter of fact, the very same repeat sequence was present in every LUR, and our data strongly suggest that LUR length variation is actually due to repeat CNV (see Supplemental Information files on figshare: <https://ndownloader.figshare.com/files/8387666> and <https://ndownloader.figshare.com/files/8360789>), as observed in other bivalve species (see Ghiselli et al., 2013; Guerra, Ghiselli & Passamonti, 2014). Tandem repeats have been also reported in the mitochondrial genomes of the bivalves *Acanthocardia tuberculata* (Dreyer & Steiner, 2006), *Placopecten magellanicus* (La Roche et al., 1990), *Moerella iridescens*, *Sanguinolaria olivacea*, *Semele scaba*, *Sinonovacula constricta*, *Solecurtus divaricatus*

([Yuan et al., 2012](#)), *Ruditapes philippinarum* ([Ghiselli et al., 2013](#)), and *Musculista senhousia* ([Guerra, Ghiselli & Passamonti, 2014](#)). These repeats are believed to arise from duplications caused by replication slippage ([Buroker et al., 1990](#); [Hayasaka, Ishida & Horai, 1991](#); [Broughton & Dowling, 1994](#)). The tandem repeats found at the 3' end of *Ruditapes decussatus* LUR are predicted to form a secondary structure (see [Fig. 2B](#) and [Supplemental Information](#) files on figshare) composed by multiple stem-loops, which obviously increase in number with the increment of the number of tandem repeats. The effect, if any, of tandem repeats in mtDNA is unknown: since the repeats are almost always localized in proximity of the CR, they might interact with regulatory elements—or even contain some—influencing replication and/or transcription initiation, and such interactions might also be altered by the formation of secondary structures ([Passamonti et al., 2011](#); [Ghiselli et al., 2013](#); [Guerra, Ghiselli & Passamonti, 2014](#)).

We assessed the genetic variability of *Ruditapes decussatus* mtDNA using two different approaches: by SP calling in CDS (RNA-Seq data on 12 individuals), and by analysis of the LUR (Sanger sequencing of 14 individuals). The CR and its flanking regions are known to be hypervariable, so they are commonly used to assess polymorphism at low taxonomic levels. Our data strongly support a very low genetic variability: the number of SPs in CDS is 145, of which 103 are private of a single individual (mRDI01)—thus reducing the number to 42—while the number of variable sites in the analyzed LURs is 98 over 3,095 aligned positions. Considering the known variability of mtDNA in bivalves ([Gissi, Iannelli & Pesole, 2008](#); [Ghiselli et al., 2013](#); [Breton et al., 2014](#); [Plazzi, Puccio & Passamonti, 2016](#)), this is a surprising result. Even more if we compare the results of the present work to a methodologically identical analysis performed on 12 *Ruditapes philippinarum* samples from the Pacific coast of USA, performed by [Ghiselli et al. \(2013\)](#): in that work, GATK yielded 194 SPs in the M-type mtDNA and 293 in the F-type. Strikingly, the 12 *Ruditapes philippinarum* samples analyzed were actually two families (6 siblings + 6 siblings). This means that randomly sampled individuals of *Ruditapes decussatus* used in this work showed a much lower mtDNA variability than *Ruditapes philippinarum* siblings. A previous analysis on the *cox1* gene of *Ruditapes decussatus* reported a nucleotide diversity ( $\pi$ ) of 0.15 for a population from the Northern Adriatic Sea ([Cordero, Peña & Saavedra, 2014](#)). Another analysis on the same gene of *Ruditapes philippinarum* from the same range resulted in a  $\pi = 0.25$  ([Cordero et al., 2017](#)), so *Ruditapes decussatus* has a lower nucleotide diversity at the *cox1* locus. The difference between the variability in mtDNA of *Ruditapes decussatus* that we are reporting here and that of *Ruditapes philippinarum* reported in [Ghiselli et al. \(2013\)](#) appears to be more marked. It is known that the genetic variability of *Ruditapes philippinarum* in the Adriatic Sea is lower than in populations from its native range in Asia ([Cordero et al., 2017](#)), probably because of the bottlenecks that this species had to go through during the multiple colonization events. The introduction in North America from Asia happened first (in the 1930s), and from there the Manila clam was introduced in Atlantic Europe (in the 1970s and 1980s), and lastly into the Adriatic Sea (1983 and 1984), and it is plausible that the genetic diversity decreased at each introduction event. Accordingly, [Cordero et al. \(2017\)](#) observed that *Ruditapes philippinarum* genetic variability in Europe is lower compared to that of the Pacific coast



of the USA, so the samples analyzed in *Ghiselli et al. (2013)* could have been more polymorphic than those analyzed in *Cordero, Peña & Saavedra (2014)*, thus explaining the more pronounced differences in genetic variability between the Manila clam and the European clam discussed above. In any case, all the available data point to a lower genetic diversity of *Ruditapes decussatus* mtDNA, and it would be interesting to know whether it is a cause or an effect of the ongoing replacement of *Ruditapes decussatus* with the invasive *Ruditapes philippinarum*. It will also be important to investigate genetic variability of the nuclear genes, especially after *Cordero, Peña & Saavedra (2014)* reported contrasting levels of differentiation between mitochondrial and nuclear markers.

With respect to SP effects, we found six indels in CDS, 2 of which do not cause frameshift, but a simple insertion/deletion of one amino acid (SP\_1698, and SP\_17619, see [Table 8](#)). Of the remaining four, SP\_6364 and SP\_10449 consist of a deletion and an insertion of a single T in two homopolymeric sequences (CTTTTTTTT and CTTTTTTT, respectively), raising the possibility of a sequencing error. In any case, the two SPs yield a shorter CDS (*cytb* and *nd5*, respectively), and are present at relatively low frequencies in the specimens carrying them, except for SP\_6364 which has a frequency of 80% in fRDI04. The *cox3* gene shows three SPs: the first one, SP\_17619, does not cause a frameshift, and results in the deletion of one alanine residue, and its frequency in mRDI01 is 97%. The second one, SP\_17621, consists of a deletion of a G with respect to the reference sequence, which is the Sanger-sequenced mtDNA of sample F4; all the individuals analyzed with RNA-Seq carry this deletion except for mRDI01 which, at that position, has the same sequence of the reference mtDNA (reference-like allele frequency in mRDI01 = 99%). The third indel, SP\_17624, consists of an insertion of two nucleotides, and its frequency in mRDI01 is 99%. So, basically, for *cox3* we have three types of sequences: (i) the Sanger-sequenced reference, which yields a 966 bp (321 aa) ORF; (ii) a sequence found in 11/12 of samples analyzed with RNA-Seq (except mRDI01) that carries a single-nucleotide deletion (SP\_17621), and yields a 963 bp (320 aa) ORF; (iii) a sequence, private of mRDI01, which is obtained by combining SP\_17624 and SP\_17621 (both 99% of frequency, so most likely co-occurring), which produces a 963 bp (320 aa) ORF. Interestingly, the ORFs obtained from the sequences described in (ii) and (iii), are almost identical, namely the sequence obtained by RNA-seq in 11/12 samples and the sequence obtained by RNA-Seq in mRDI01 are basically the same, and differ from the Sanger-sequenced reference, yielding an amino acid sequence that differs in the last 35 residues (all data available in Supplemental Information files on figshare: <https://doi.org/10.6084/m9.figshare.4970762.v3>). Given this consistent difference between the sequence obtained by Sanger-sequencing of DNA, and those obtained by RNA-Seq, it is tempting to speculate that this difference might be caused by RNA editing, a mechanism observed in mtDNA of some animals (*Lavrov & Pett, 2016*), and recently reported to be common in cephalopods (*Liscovitch-Brauer et al., 2017*). Actually, *Liscovitch-Brauer et al. (2017)* reported only A-to-I editing, which is not the kind of change we are observing here, but other types of editing are known across eukaryotes (see *Gott & Emeson, 2000* for a review), and some others, still unknown, might exist as well. Post-transcriptional modifications (thus including RNA-editing) are still poorly understood mechanisms, but they appear to be responsible for



most of the mitochondrial gene expression regulation ([Scheibye-Alsing et al., 2007](#); [Scheffler, 2008](#); [Milani et al., 2014a](#)). What we propose here is a pure conjecture, but we think in the future it might be worthy to investigate mitochondrial transcriptomes looking for such kind of “unexpected” biological features.

Interestingly, in contrast with a low nucleotide variability along the entire mitochondrial genome, we observed a pretty high polymorphism in LUR length due to CNV of tandem repeats, and even a LUR length heteroplasmy: two females yielded two electrophoretic bands each (~2,100 and ~3,500 bp in F3; ~2,500 and ~3,500 bp in F17; see [Table 6](#)). A possible explanation is that the diversity (CNV) detected in the LURs could be recent: the accumulation of nucleotide variation at different sites along the mitochondrial genome needs time, while the kind structural variability we observed can be achieved in few generations (or even one) considering that replication slippage is common in repeat-rich regions.

### Phylogenetic relationship with *Ruditapes philippinarum*

Despite *Ruditapes decussatus* and *Ruditapes philippinarum* being morphologically similar and being ascribed to the same genus, the results here reported clearly show that they are quite different both for mtDNA sequence ([Figs. 3 and 4](#)) and mtDNA gene arrangement ([Fig. 5](#)). This is an unusual finding, even among bivalves, which are known to be fast-evolving for these characters. This may point to the fact that these two species are less related than previously thought. Actually, this is not the first clue that *Ruditapes decussatus* and *Ruditapes philippinarum* are quite different genetically, as allozyme electrophoresis ([Passamonti, Mantovani & Scali, 1997, 1999](#)) and satellite DNA content ([Passamonti, Mantovani & Scali, 1998](#)) pointed out. More in-depth analyses are therefore needed to correctly trace the phylogenetic relationships of these two *Ruditapes* species, which may eventually end up in two different Genera. As shown in [Figs. 3–5](#), the Genus *Paphia* is the most similar to *Ruditapes decussatus*.

### Presence/absence of DUI

We could not find evidence for sex-specific mtDNAs, typical of DUI. As stated in the Introduction, the search for DUI is not a straightforward process. HTS can help thanks to a much deeper sequencing coverage (in respect to the cloning-and-Sanger-sequencing approach), and because it overcomes the problem of primer specificity, a limitation of the classical approach. One possible concern about using HTS approaches based on short reads in presence of DUI is about the ability of softwares to detect divergent reads and assemble them correctly. More specifically, one could ask what is the divergence threshold under which the assemblers are not able to partition the contigs into two sex-linked groups. We do not know such a threshold, but we used different assembly strategies trying to retrieve sex-specific mtDNA sequences from our data. Other than the approach reported in Materials and Methods (which is the one that produced the data reported here), we tried other techniques. After identifying reads that blasted to bivalve mitochondrial sequences present in GenBank and discarding all the other reads, we generated A5+CAP3 assemblies: (i) for each of the samples (obtaining 12 separate assemblies), and (ii) pooling the six males together and the six females together, and

assembling the two sex-specific pools. Both these approaches did not show evidence of sex-specific mtDNAs. Then we took the assembly obtained from the females and removed the reads from each of the samples that mapped (<8 mismatches) to these sequences. We then used the remaining reads as A5 input. The program could not assemble anything. Lastly, we tried the software MetaVelvet ([Namiki et al., 2012](#))—that assembles metagenomes—on all the reads matching bivalve mtDNAs, and only one genome was produced. After all these alternative approaches failed to find two sex-linked mtDNAs, we decided to proceed with the assembly as indicated in Materials and Methods, because it was the technique that yielded the best quality contigs, most likely because using the reads from all 12 the individuals granted a higher coverage of the mtDNA. Given these results, we can propose three different explanations.

1. *Ruditapes decussatus* is characterized by SMI of mitochondria, so a male-transmitted mtDNA is not present in this species.
2. The divergence between the two sex-specific mtDNAs is too low to be detected. This could be the outcome of two different situations.
  - a) DUI is very young in this species, so the two sex-linked mtDNAs did not have the time to diverge.
  - b) A role-reversal event occurred recently. Role reversal (a.k.a. “route reversal” or “masculinization”) is a process—observed so far only in species of the *Mytilus* complex—by which F-type genomes invades the male germ line becoming sperm-transmitted, thus turning into M-type mtDNAs ([Hoeh et al., 1997](#)). This event actually resets to zero the divergence between F- and M-type, although substantial differences in the control regions were reported between the original F-type and the “masculinized” one (see [Zouros, 2013](#) for a thorough review). The hypothesis that role reversal could have occurred multiple times in the evolutionary history of bivalves and could have led to the complete replacement of M or F mtDNAs in several species was proposed by [Hoeh et al. \(1997\)](#) to explain the scattered phylogenetic distribution of DUI across Bivalvia. Indeed, according to the hypothesis of a single origin, DUI arose >400 Mya, approximately at the origin of Autolamellibranchia, but, as said, such hypothesis requires the assumption of multiple role-reversal and/or DUI loss events in several branches of the bivalve tree (see [Zouros, 2013](#) for a detailed discussion). Recently, a multiple origin of DUI was proposed ([Milani et al., 2013, 2014b](#); [Milani, Ghiselli & Passamonti, 2016](#); [Mitchell et al., 2016](#)), and in such case there would be no need of multiple role-reversal events to explain its phylogeny. In our opinion, until further evidence will be provided, role-reversal should not be considered a rule, but rather an exception. Of course, we cannot rule out that a masculinization event might have occurred in *Ruditapes decussatus*, so this hypothesis must be taken into consideration.
3. In our data, even if there is no clear evidence of a male-specific mtDNA, a male sample (mRDI01) clearly stood out from the others, both males and females (see [Table 7](#)). Overall, the divergence between mRDI01 and the other 11 samples calculated

considering its private SPs is of 151 sites over 18,995 bp (considering the whole mtDNA), and of 103 sites over 14,920 bp (considering only CDS). In both cases the divergence is very low (0.8% and 0.7%, respectively), which explains why the mtDNA of mRDI01, although different, was not assembled as a separate genome. We have no sufficient data to evaluate if such divergence is normal within *Ruditapes decussatus* populations, but considered the variability usually observed in bivalves, we find the difference unsurprising. On the contrary, the lack of variability among the other 11 samples is remarkable. For these reasons, we are inclined to believe that mRDI01 divergence is compatible with hypotheses (1) and (2). That said, there still could be a third, quite conjectural, hypothesis by which these data might indicate an incipient DUI, not yet fixed in the population.

All in all, we have a preference for the first explanation, but the present data are not sufficient to exclude the others, and a more thorough investigation is necessary to assess this point.

Up to now DUI was identified in only three Veneridae species: *Cyclina sinensis*, *Ruditapes philippinarum*, and *Meretrix lamarckii* (Gusman *et al.*, 2016). The status of *Paphia* is still unknown, and in future works it would be interesting to investigate more Heterodonta species to understand better the distribution of DUI in this derived group of bivalves.

## ACKNOWLEDGEMENTS

We would like to thank Edoardo Turolla (Istituto Delta Ecologia Applicata, Ferrara, Italy) for providing the specimens, and Massimo Milan for bibliographic suggestions. We also gratefully thank the Editor Tim Collins, and the reviewers Carlos Saavedra, Shallee Page, and one anonymous colleague for their comments and suggestions.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This study was supported by the Italian Ministry of Education, University and Research (MIUR) FIR Programme no. RBFR13T97A funded to FG, MIUR SIR Programme no. RBSI14G0P5 funded to LM, Zumberge Foundation to SVN, and by the Canziani bequest funded to MP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Italian Ministry of Education, University and Research (MIUR) FIR: RBFR13T97A.

MIUR SIR: RBSI14G0P5.

Zumberge Foundation.

Canziani bequest.

## Competing Interests

The authors declare that they have no competing interests.

## Author Contributions

- Fabrizio Ghiselli conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Liliana Milani conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper.
- Mariangela Iannello performed the experiments, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper.
- Emanuele Procopio performed the experiments, analyzed the data.
- Peter L. Chang analyzed the data.
- Sergey V. Nuzhdin conceived and designed the experiments, contributed reagents/materials/analysis tools, reviewed drafts of the paper.
- Marco Passamonti conceived and designed the experiments, contributed reagents/materials/analysis tools, reviewed drafts of the paper.

## DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

GenBank accession numbers [MF055702](#) to [MF055714](#), and [KP089983](#).

GenBank BioProject [PRJNA170478](#).

## Data Availability

The following information was supplied regarding data availability:

Ghiselli, Fabrizio; Milani, Liliana; Iannello, Mariangela; Procopio, Emanuele; L. Chang, Peter; Nuzhdin, Sergey; Passamonti, Marco (2017): The Complete Mitochondrial Genome of the Grooved Carpet Shell, *Ruditapes decussatus* (Bivalvia, Veneridae). figshare. <https://doi.org/10.6084/m9.figshare.4970762.v3>.

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.3692#supplemental-information>.

## REFERENCES

- Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**(1):402 DOI [10.1186/1471-2164-12-402](#).
- Arias-Pérez A, Cordero D, Borrell Y, Sánchez JA, Blanco G, Freire R, Insua A, Saavedra C. 2016. Assessing the geographic scale of genetic population management with microsatellites and introns in the clam *Ruditapes decussatus*. *Ecology and Evolution* **6**(10):3380–3404 DOI [10.1002/ece3.2052](#).
- Attwood TK, Bradley P, Flower DR, Gaulton A, Maudling N, Mitchell AL, Moulton G, Nordle A, Paine K, Taylor P, Uddin A, Zygouri C. 2003. PRINTS and its automatic supplement, prePRINTS. *Nucleic Acids Research* **31**(1):400–402 DOI [10.1093/nar/gkg030](#).

- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37:W202–W208 DOI 10.1093/nar/gkp335.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsche G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* 69(2):313–319 DOI 10.1016/j.ympev.2012.08.023.
- Breton S, Milani L, Ghiselli F, Guerra D, Stewart DT, Passamonti M. 2014. A resourceful genome: updating the functional repertoire and evolutionary role of animal mitochondrial DNAs. *Trends in Genetics* 30(12):555–564 DOI 10.1016/j.tig.2014.09.002.
- Breton S, Stewart DT, Hoeh WR. 2010. Characterization of a mitochondrial ORF from the gender-associated mtDNAs of *Mytilus* spp. (Bivalvia: Mytilidae): identification of the “missing” ATPase 8 gene. *Marine Genomics* 3(1):11–18 DOI 10.1016/j.margen.2010.01.001.
- Breton S, Stewart DT, Shepardson S, Trdan RJ, Bogan AE, Chapman EG, Ruminas AJ, Piontkivska H, Hoeh WR. 2011. Novel protein genes in animal mtDNA: a new sex determination system in freshwater mussels (Bivalvia: Unionoida)? *Molecular Biology and Evolution* 28(5):1645–1659 DOI 10.1093/molbev/msq345.
- Broughton RE, Dowling TE. 1994. Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. *Genetics* 138:179–190.
- Buroker NE, Brown JR, Gilbert TA, O’Hara PJ, Beckenbach AT, Thomas WK, Smith MJ. 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. *Genetics* 124:157–163.
- Buske FA, Bodén M, Bauer DC, Bailey TL. 2010. Assigning roles to DNA regulatory motifs using comparative genomics. *Bioinformatics* 26(7):860–866 DOI 10.1093/bioinformatics/btq049.
- Cao L, Kenchington E, Zouros E, Rodakis GC. 2004. Evidence that the large noncoding sequence is the main control region of maternally and paternally transmitted mitochondrial genomes of the marine mussel (*Mytilus* spp.). *Genetics* 167(2):835–850 DOI 10.1534/genetics.103.026187.
- Capella-Gutiérrez S, Silla-Martinez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973 DOI 10.1093/bioinformatics/btp348.
- Chamary JV, Parmley JL, Hurst LD. 2006. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nature Reviews Genetics* 7(2):98–108 DOI 10.1038/nrg1770.
- Cordero D, Delgado M, Liu B, Ruesink J, Saavedra C. 2017. Population genetics of the Manila clam (*Ruditapes philippinarum*) introduced in North America and Europe. *Scientific Reports* 7:39745 DOI 10.1038/srep39745.
- Cordero D, Peña JB, Saavedra C. 2014. Phylogeographic analysis of introns and mitochondrial DNA in the clam *Ruditapes decussatus* uncovers the effects of Pleistocene glaciations and endogenous barriers to gene flow. *Molecular Phylogenetics and Evolution* 71:274–287 DOI 10.1016/j.ympev.2013.11.003.
- Crick FHC. 1966. Codon—anticodon pairing: the wobble hypothesis. *Journal of Molecular Biology* 19(2):548–555 DOI 10.1016/s0022-2836(66)80022-0.
- Darty K, Denise A, Ponty Y. 2009. VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25(15):1974–1975 DOI 10.1093/bioinformatics/btp250.
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* 43(5):491–498 DOI 10.1038/ng.806.

- de Sousa JT, Milan M, Bargelloni L, Pauletto M, Matias D, Joaquim S, Matias AM, Quillien V, Leitão A, Huvet A. 2014. A microarray-based analysis of gametogenesis in two Portuguese populations of the European clam *Ruditapes decussatus*. *PLOS ONE* 9:e92202 DOI 10.1371/journal.pone.0092202.
- Diz AP, Dudley E, Skibinski DOF. 2012. Identification and characterisation of highly expressed proteins in sperm cells of the marine mussel *Mytilus edulis*. *Proteomics* 12(12):1949–1956 DOI 10.1002/pmic.201100500.
- Dreyer H, Steiner G. 2006. The complete sequences and gene organisation of the mitochondrial genomes of the heterodont bivalves *Acanthocardia tuberculata* and *Hiatella arctica*—and the first record for a putative Atpase subunit 8 gene in marine bivalves. *Frontiers in Zoology* 3:13.
- Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* 44(D1):D279–D285 DOI 10.1093/nar/gkv1344.
- Ghiselli F, Milani L, Chang PL, Hedgecock D, Davis JP, Nuzhdin SV, Passamonti M. 2012. *De Novo* assembly of the Manila clam *Ruditapes philippinarum* transcriptome provides new insights into expression bias, mitochondrial doubly uniparental inheritance and sex determination. *Molecular Biology and Evolution* 29(2):771–786 DOI 10.1093/molbev/msr248.
- Ghiselli F, Milani L, Guerra D, Chang PL, Breton S, Nuzhdin SV, Passamonti M. 2013. Structure, transcription, and variability of metazoan mitochondrial genome: perspectives from an unusual mitochondrial inheritance system. *Genome Biology and Evolution* 5(8):1535–1554 DOI 10.1093/gbe/evt112.
- Ghiselli F, Milani L, Passamonti M. 2011. Strict sex-specific mtDNA segregation in the germ line of the DUI species *Venerupis philippinarum* (Bivalvia: Veneridae). *Molecular Biology and Evolution* 28(2):949–961 DOI 10.1093/molbev/msq271.
- Gissi C, Iannelli F, Pesole G. 2008. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity* 101(4):301–320 DOI 10.1038/hdy.2008.62.
- Gosling EM. 2003. *Bivalve Molluscs: Biology, Ecology and Culture*. Hoboken: Wiley Online Library.
- Gott JM, Emeson RB. 2000. Functions and mechanisms of RNA editing. *Annual Review of Genetics* 34(1):499–531 DOI 10.1146/annurev.genet.34.1.499.
- Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. 2008. The Vienna RNA websuite. *Nucleic Acids Research* 36:W70–W74 DOI 10.1093/nar/gkn188.
- Guerra D, Ghiselli F, Passamonti M. 2014. The Largest Unassigned Regions of the male- and female-transmitted mitochondrial DNAs in *Musculista senhousia* (Bivalvia Mytilidae). *Gene* 536(2):316–325 DOI 10.1016/j.gene.2013.12.005.
- Gusman A, Lecomte S, Stewart DT, Passamonti M, Breton S. 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ* 4:e2760 DOI 10.7717/peerj.2760.
- Hayasaka K, Ishida T, Horai S. 1991. Heteroplasmy and polymorphism in the major noncoding region of mitochondrial DNA in Japanese monkeys: association with tandemly repeated sequences. *Molecular Biology and Evolution* 8:399–415 DOI 10.1093/oxfordjournals.molbev.a040660.
- He Z, Zhang H, Gao S, Lercher MJ, Chen W-H, Hu S. 2016. Evolvview v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Research* 44(W1):W236–W241 DOI 10.1093/nar/gkw370.



- Hoeh WR, Stewart DT, Saavedra C, Sutherland BW, Zouros E. 1997. Phylogenetic evidence for role-reversals of gender-associated mitochondrial DNA in *Mytilus* (Bivalvia: Mytilidae). *Molecular Biology and Evolution* 14(9):959–967 DOI 10.1093/oxfordjournals.molbev.a025839.
- Huang X, Madan A. 1999. CAP3: a DNA sequence assembly program. *Genome Research* 9(9):868–877 DOI 10.1101/gr.9.9.868.
- Jones P, Binns D, Chang H-YY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240 DOI 10.1093/bioinformatics/btu031.
- Ju YS, Kim J-I, Kim S, Hong D, Park H, Shin J-Y, Lee S, Lee W-C, Kim S, Yu S-B, Park SS, Seo S-H, Yun J-Y, Kim H-J, Lee D-S, Yavartanoo M, Kang HP, Gokcumen O, Govindaraju DR, Jung JH, Chong H, Yang K-S, Kim H, Lee C, Seo J-S. 2011. Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals. *Nature Genetics* 43(8):745–752 DOI 10.1038/ng.872.
- King JL, LaRue BL, Novroski NM, Stoljarova M, Seo SB, Zeng X, Warshauer DH, Davis CP, Parson W, Sajantila A, Budowle B. 2014. High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. *Forensic Science International: Genetics* 12:128–135 DOI 10.1016/j.fsigen.2014.06.001.
- La Roche J, Snyder M, Cook DI, Fuller K, Zouros E. 1990. Molecular characterization of a repeat element causing large-scale size variation in the mitochondrial DNA of the sea scallop *Placopecten magellanicus*. *Molecular Biology and Evolution* 7(1):45–64 DOI 10.1093/oxfordjournals.molbev.a040586.
- Laslett D, Canback B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* 24(2):172–175 DOI 10.1093/bioinformatics/btm573.
- Lavrov DV, Pett W. 2016. Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. *Genome Biology and Evolution* 8(9):2896–2913 DOI 10.1093/gbe/evw195.
- Leite RB, Milan M, Coppe A, Bortoluzzi S, dos Anjos A, Reinhardt R, Saavedra C, Patarnello T, Cancela ML, Bargelloni L. 2013. mRNA-Seq and microarray development for the Grooved Carpet shell clam, *Ruditapes decussatus*: a functional approach to unravel host-parasite interaction. *BMC Genomics* 14(1):741 DOI 10.1186/1471-2164-14-741.
- Liscovitch-Brauer N, Alon S, Porath HT, Elstein B, Unger R, Ziv T, Admon A, Levanon EY, Rosenthal JJC, Eisenberg E. 2017. Trade-off between transcriptome plasticity and genome evolution in cephalopods. *Cell* 169(2):191–202.e11 DOI 10.1016/j.cell.2017.03.025.
- Lubośny M, Przyłucka A, Sańko TJ, Śmietanka B, Rosenfeld S, Burzyński A. 2017/2. Next generation sequencing of gonadal transcriptome suggests standard maternal inheritance of mitochondrial DNA in *Eurhomalea rufa* (Veneridae). *Marine Genomics* 31:21–23 DOI 10.1016/j.margen.2016.11.002.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20(9):1297–1303 DOI 10.1101/gr.107524.110.
- Milani L, Ghiselli F. 2015. Mitochondrial activity in gametes and transmission of viable mtDNA. *Biology Direct* 10:22 DOI 10.1186/s13062-015-0057-6.
- Milani L, Ghiselli F, Guerra D, Breton S, Passamonti M. 2013. A comparative analysis of mitochondrial ORFans: new clues on their origin and role in species with doubly uniparental

inheritance of mitochondria. *Genome Biology and Evolution* 5(7):1408–1434  
DOI 10.1093/gbe/evt101.

- Milani L, Ghiselli F, Iannello M, Passamonti M. 2014a.** Evidence for somatic transcription of male-transmitted mitochondrial genome in the DUI species *Ruditapes philippinarum* (Bivalvia: Veneridae). *Current Genetics* 60(3):163–173 DOI 10.1007/s00294-014-0420-7.
- Milani L, Ghiselli F, Maurizii MG, Nuzhdin SV, Passamonti M. 2014b.** Paternally transmitted mitochondria express a new gene of potential viral origin. *Genome Biology and Evolution* 6(2):391–405 DOI 10.1093/gbe/evu021.
- Milani L, Ghiselli F, Passamonti M. 2016.** Mitochondrial selfish elements and the evolution of biological novelties. *Current Zoology* 62(6):687–697 DOI 10.1093/cz/zow044.
- Milbury CA, Lee JC, Cannone JJ, Gaffney PM, Gutell RR. 2010.** Fragmentation of the large subunit ribosomal RNA gene in oyster mitochondrial genomes. *BMC Genomics* 11(1):485 DOI 10.1186/1471-2164-11-485.
- Mitchell A, Guerra D, Stewart D, Breton S. 2016.** *In silico* analyses of mitochondrial ORFans in freshwater mussels (Bivalvia: Unionoida) provide a framework for future studies of their origin and function. *BMC Genomics* 17(1):597 DOI 10.1186/s12864-016-2986-6.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008.** Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5(7):621–628 DOI 10.1038/nmeth.1226.
- Namiki T, Hachiya T, Tanaka H, Sakakibara Y. 2012.** MetaVelvet: an extension of Velvet assembler to *de novo* metagenome assembly from short sequence reads. *Nucleic Acids Research* 40(20):e155 DOI 10.1093/nar/gks678.
- Ojala D, Montoya J, Attardi G. 1981.** tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290(5806):470–474 DOI 10.1038/290470a0.
- Passamonti M, Ghiselli F. 2009.** Doubly uniparental inheritance: two mitochondrial genomes, one precious model for organelle DNA inheritance and evolution. *DNA and Cell Biology* 28(2):79–89 DOI 10.1089/dna.2008.0807.
- Passamonti M, Mantovani B, Scali V. 1997.** Allozymic characterization and genetic relationships among four species of Tapetinae (Bivalvia, Veneridae). *Italian Journal of Zoology* 64:117–124 DOI 10.1080/11250009709356183.
- Passamonti M, Mantovani B, Scali V. 1998.** Characterization of a highly repeated DNA family in tapetinae species (mollusca bivalvia: veneridae). *Zoological Science* 15(4):599–605 DOI 10.2108/zsj.15.599.
- Passamonti M, Mantovani B, Scali V. 1999.** Allozymic analysis of some Mediterranean Veneridae (Mollusca: Bivalvia): preliminary notes on taxonomy and systematics of the family. *Journal of the Marine Biological Association of the UK* 79(5):899–906 DOI 10.1017/s0025315498001064.
- Passamonti M, Ricci A, Milani L, Ghiselli F. 2011.** Mitochondrial genomes and Doubly Uniparental Inheritance: new insights from *Musculista senhousia* sex-linked mitochondrial DNAs (Bivalvia Mytilidae). *BMC Genomics* 12(1):442 DOI 10.1186/1471-2164-12-442.
- Pesole G, Allen JF, Lane N, Martin W, Rand DM, Schatz G, Saccone C. 2012.** The neglected genome. *EMBO Reports* 13:473–474.
- Plazzi F, Puccio G, Passamonti M. 2016.** Comparative large-scale mitogenomics evidences clade-specific evolutionary trends in mitochondrial DNAs of Bivalvia. *Genome Biology and Evolution* 8(8):2544–2564 DOI 10.1093/gbe/evw187.
- Pozzi A, Plazzi F, Milani L, Ghiselli F, Passamonti M. 2017.** SmithRNAs: could mitochondria “bend” nuclear regulation? *Molecular Biology and Evolution* 34(8):1960–1973 DOI 10.1093/molbev/msx140.

- Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* 132:365–386 DOI 10.1385/1-59259-192-2:365.
- Scheffler IE. 2008. *Mitochondria*. Hoboken, NJ: Wiley.
- Scheibye-Alsing K, Cirera S, Gilchrist MJ, Fredholm M, Gorodkin J. 2007. EST analysis on pig mitochondria reveal novel expression differences between developmental and adult tissues. *BMC Genomics* 8(1):367 DOI 10.1186/1471-2164-8-367.
- Serb JM, Lydeard C. 2003. Complete mtDNA sequence of the North American freshwater mussel, *Lampsilis ornata* (Unionidae): an examination of the evolution and phylogenetic utility of mitochondrial genome organization in Bivalvia (Mollusca). *Molecular Biology and Evolution* 20(11):1854–1866 DOI 10.1093/molbev/msg218.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7(1):539 DOI 10.1038/msb.2011.75.
- Skibinski DO, Gallagher C, Beynon CM. 1994a. Mitochondrial DNA inheritance. *Nature* 368:817–818.
- Skibinski DO, Gallagher C, Beynon CM. 1994b. Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. *Genetics* 138:801–809.
- Smith DR. 2013. RNA-Seq data: a goldmine for organelle research. *Briefings in Functional Genomics* 12(5):454–456 DOI 10.1093/bfgp/els066.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313 DOI 10.1093/bioinformatics/btu033.
- Stothard P. 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* 28:1102–1104.
- Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27(7):1009–1010 DOI 10.1093/bioinformatics/btr039.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30(12):2725–2729 DOI 10.1093/molbev/mst197.
- Tan MH, Gan HM, Schultz MB, Austin CM. 2015. MitoPhAST, a new automated mitogenomic phylogeny tool in the post-genomic era with a case study of 89 decapod mitogenomes including eight new freshwater crayfish mitogenomes. *Molecular Phylogenetics and Evolution* 85:180–188 DOI 10.1016/j.ympev.2015.02.009.
- Theologidis I, Fodelianakis S, Gaspar MB, Zouros E. 2008. Doubly uniparental inheritance (DUI) of mitochondrial DNA in *Donax trunculus* (Bivalvia: Donacidae) and the problem of its sporadic detection in Bivalvia. *Evolution; International Journal of Organic Evolution* 62(4):959–970 DOI 10.1111/j.1558-5646.2008.00329.x.
- Tritt A, Eisen JA, Facciotti MT, Darling AE. 2012. An integrated pipeline for *de novo* assembly of microbial genomes. *PLOS ONE* 7(9):e42304 DOI 10.1371/journal.pone.0042304.
- Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, Wagner L. 2005. Database resources of the National Center for Biotechnology. *Nucleic Acid Research* 33:D39–D45 DOI 10.1093/nar/gki062.
- Yu H, Li Q. 2011. Mutation and selection on the wobble nucleotide in tRNA anticodons in marine bivalve mitochondrial genomes. *PLOS ONE* 6(1):e16147 DOI 10.1371/journal.pone.0016147.

- Yuan Y, Li Q, Yu H, Kong L. 2012.** The complete mitochondrial genomes of six heterodont bivalves (Tellinoidea and Solenoidea): variable gene arrangements and phylogenetic implications. *PLOS ONE* 7(2):e32353 DOI [10.1371/journal.pone.0032353](https://doi.org/10.1371/journal.pone.0032353).
- Yuan S, Xia Y, Zheng Y, Zeng X. 2016.** Next-generation sequencing of mixed genomic DNA allows efficient assembly of rearranged mitochondrial genomes in *Amolops chunganensis* and *Quasipaa boulengeri*. *PeerJ* 4:e2786 DOI [10.7717/peerj.2786](https://doi.org/10.7717/peerj.2786).
- Zouros E. 2013.** Biparental inheritance through uniparental transmission: the doubly uniparental inheritance (DUI) of mitochondrial DNA. *Evolutionary Biology* 40(1):1–31 DOI [10.1007/s11692-012-9195-2](https://doi.org/10.1007/s11692-012-9195-2).
- Zouros E, Ball AO, Saavedra C, Freeman KR. 1994a.** Mitochondrial DNA inheritance. *Nature* 368:818.
- Zouros E, Oberhauser Ball A, Saavedra C, Freeman KR. 1994b.** An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. *Proceedings of the National Academy of Sciences of USA* 91(16):7463–7467 DOI [10.1073/pnas.91.16.7463](https://doi.org/10.1073/pnas.91.16.7463).