



HAL
open science

Complex mitogenomic rearrangements within the Pectinidae (Mollusca: Bivalvia)

Tamás Malkócs, Amélia Viricel, Vanessa Becquet, Louise Evin, Emmanuel
Dubillot, Eric Pante

► **To cite this version:**

Tamás Malkócs, Amélia Viricel, Vanessa Becquet, Louise Evin, Emmanuel Dubillot, et al.. Complex mitogenomic rearrangements within the Pectinidae (Mollusca: Bivalvia). *BMC Evolutionary Biology*, 2022, 22 (29), 10.1186/s12862-022-01976-0 . hal-03434007

HAL Id: hal-03434007

<https://hal.science/hal-03434007>

Submitted on 18 Nov 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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Complex mitogenomic rearrangements**
2 **within the Pectinidae (Mollusca:**
3 **Bivalvia)**

4 Tamás Malkócs^{a,b,c,d}, Amélia Viricel^a, Vanessa Becquet^a, Louise Evin^a, Emmanuel
5 Dubillot^a, Eric Pante^a

6 ^a Littoral Environnement et Sociétés (LIENSs), UMR 7266, CNRS-Université de La
7 Rochelle, 2 rue Olympe de Gouges, F-17042 La Rochelle Cedex 01, France

8 ^b Pál Juhász-Nagy Doctoral School of Biology and Environmental Sciences,
9 University of Debrecen, Egyetem tér 1, 4032 Debrecen, Hungary

10 ^c Institute of Biology and Ecology, University of Debrecen, Egyetem tér 1, 4032
11 Debrecen, Hungary

12 ^d Institute of Aquatic Ecology, Centre for Ecological Research, 4026 Debrecen,
13 Hungary

14 Corresponding author: Tamás Malkócs <tamas.malkocs@gmail.com>; Littoral
15 Environnement et Sociétés (LIENSs), UMR 7266, CNRS-Université de La Rochelle,
16 2 rue Olympe de Gouges, F-17042 La Rochelle Cedex 01, France

17 **Abstract**

18 **Background**

19 Scallops (Bivalvia: Pectinidae) present extraordinary variance in both mitochondrial
20 genome size, structure and content, even when compared to the extreme diversity
21 documented within Mollusca and Bivalvia. In pectinids, mitogenome rearrangements
22 involve protein coding and rRNA genes along with tRNAs, and different genome
23 organization patterns can be observed even at the level of Tribes. Existing pectinid
24 phylogenies fail to resolve some relationships in the family, Chlamydinae being an
25 especially problematic group.

26 **Results**

27 In our study, we sequenced, annotated and characterized the mitochondrial genome
28 of a member of Chlamydinae, *Mimachlamys varia*—a species of commercial interest
29 and an effective bioindicator—revealing yet another novel gene arrangement in the
30 Pectinidae. The phylogeny based on all mitochondrial protein coding and rRNA
31 genes suggests the paraphyly of the *Mimachlamys* genus, further commending the
32 taxonomic revision of the classification within the Chlamydinae subfamily. At the
33 scale of the Pectinidae, we found that 15 sequence blocks are involved in
34 mitogenome rearrangements, which behave as separate units.

35 **Conclusions**

36 Our study reveals incongruities between phylogenies based on mitochondrial
37 protein-coding versus rRNA genes within the Pectinidae, suggesting that locus
38 sampling affects phylogenetic inference at the scale of the family. We also conclude

39 that the available taxon sampling does not allow for understanding of the
40 mechanisms responsible for the high variability of mitogenome architecture observed
41 in the Pectinidae, and that unraveling these processes will require denser taxon
42 sampling.

43 **Keywords**

44 mitogenome, gene rearrangement, Pectinidae, de novo assembly, Variegated
45 scallop, RNA-seq

46 **Background**

47 Historically, animal mitochondrial genomes were thought to have a conserved size of
48 approximately 16 kbp, as well as a conserved structure, containing 13 protein coding
49 genes (PCGs), two rRNA genes, and 22 tRNA genes and a single non-coding,
50 ‘control region’ that regulates replication and transcription [1, 2]. In contrast,
51 molluscan mitochondrial genomes provide examples of exceptional variation in size,
52 structure and even function. Mitogenome size in Mollusca ranges from ~14 kbp up to
53 56.2 kbp [3–7], gene rearrangements, duplications, losses and inversions are
54 frequent e.g. [8], and there is evidence of the integration of novel genes e.g. [9].
55 Additionally, a unique, complex mitochondrial inheritance system, called ‘doubly
56 uniparental inheritance’ evolved in gonochoric bivalve mollusks [10–12].

57 Regarding mitogenome structure, scallops (Bivalvia: Pectinidae) constitute one of
58 the most peculiar groups within Mollusca. They include one of the largest
59 mitogenomes known (32–42 kbp; [7], with other species having mitogenomes of

60 different lengths ranging from 16 kbp to 21 kbp [13, 14]. They also present the most
61 variation in mitochondrial gene order, even at relatively shallow phylogenetic depths,
62 i.e. within subfamilies [14]. In Pectinidae, as well as among metazoans in general,
63 gene rearrangements are most common for tRNAs [14–17]. Several pectinids
64 contain extra tRNAs beyond the essential 22 needed for mitochondrial translation,
65 either duplicated genes, or genes encoding for the same amino acid, but using a
66 different anticodon [18–21]. However, even more species contain fewer than 22
67 mitochondrial tRNAs [13, 14, 19, 22, 23], possibly coming from annotation errors,
68 which might be the result of modified tRNA secondary structure, common in other
69 mollusks [24–26]. Similarly, *atp8* genes were reported earlier as missing from most
70 bivalves, including pectinids [13, 14, 18, 20], although reannotation efforts revealed
71 its presence in several species [15, 21, 27]. The difficulty of annotating the *atp8* gene
72 is thought to stem from the poorly conserved sequence of this gene in bivalves [28–
73 30]. Unlike what was observed in other animal groups, mitochondrial genome
74 rearrangements in Pectinidae often involve PCGs [13–15]. This, however, does not
75 seem to be correlated to sequence similarities among PCGs [14]. Furthermore, like
76 most bivalves, most pectinid mitochondrial genomes contain several inflated non-
77 coding regions, some of which are hypothesized to function as control regions [13,
78 14, 18, 20], although it remains controversial whether several control regions within a
79 mitogenome could provide normal function [20].

80 Most existing pectinid phylogenies are based on one, or a handful of gene
81 sequences, of either just mitochondrial, or of both mitochondrial and nuclear origin
82 [31–41]. These phylogenies are often incongruent with morphological classifications
83 of the Pectinidae [40]. Also, while they conclude that the Pectinidae is monophyletic,

84 many lower taxonomic levels are reported as being paraphyletic [31, 40].
85 Phylogenies based on full mitogenomes have shown their utility in disentangling
86 evolutionary relationships that were controversial when investigated using only a few
87 genes [16, 42–44]. There are a few examples of pectinid phylogenies based on
88 complete mitogenomes [14, 15, 18, 22, 27], however, most of these studies rely
89 solely on genetic distance-based methods (except: [15, 22]), without considering the
90 differences in substitution rates among different mitochondrial genes. Despite the
91 diversity of approaches, some phylogenetic relationships remain to be resolved, for
92 example, those within the Chlamydinae subfamily.

93 Numerous scallop species are harvested for human consumption and are used as
94 environmental indicators. One example is the variegated scallop (*Mimachlamys varia*
95 L., 1758), a species that inhabits the subtidal zone of the European Atlantic and the
96 Mediterranean Sea [45]. It was demonstrated that *Mim. varia* is a potent bioindicator,
97 as it presents significant physiological responses to chemical contaminants,
98 including changes in biomarkers connected to oxidative stress, immune system
99 function and mitochondrial respiration [46–48]. Most existing genetic research
100 involving *Mim. varia* focused on the use of a few genes for phylogenetic or
101 population genetic inference [31, 40, 49]. Viricel et al. [50] published the
102 transcriptome of *Mim. varia*, providing a valuable resource for further
103 ecophysiological and ecotoxicological research, including analyses of differential
104 gene expression in response to marine pollution and comparative genomic studies
105 within the Pectinidae.

106 In this study, our goals were to: I) assemble and characterize the mitochondrial
107 genome of the variegated scallop (*Mimachlamys varia*); II) reconstruct the most up to
108 date pectinid phylogeny using all available mitochondrial genomes, including *Mim.*
109 *varia*; III) perform comparative analyses of pectinid mitochondrial gene orders, to
110 investigate the evolution of mitogenome arrangements within this family.

111 **Results**

112 **Sequencing of the mitogenome of reference for the variegated scallop**

113 The NOVOPlasty assembly, based on a single *cox1* seed, yielded four contigs,
114 5,289, 530, 3,458 and 13,528 bp long, recruiting about 5.9 M reads (9.67% of the
115 61.1 M RNAseq reads submitted to the assembler), with an average depth of
116 coverage of 84,143. NOVOPlasty never retrieved a full, circular mitogenome based
117 on these data; the final assembly is based on the manual assembly of the four,
118 overlapping contigs in Geneious, resulting in a complete, 20,400 bp long circular
119 assembly. In total, 6.9 M reads were remapped onto the assembled genome. After
120 remapping, the breadth of coverage is 99.98%, only 5 bases are not covered around
121 the site where the circular genome was split into a linear sequence. Non-zero
122 coverage depth values ranged from 11–4,160,626 (Fig. 1). The mean and median
123 coverage depths were 100,437 and 14,687, respectively.

124 Genome composition

125 The full mitogenome contains 13 Protein Coding Genes (PCGs; including an *atp8*,
126 reported as missing for many bivalves; [13, 14, 18, 20]), 23 tRNAs, and 2 rRNAs

127 (Fig. 1, Table 1). All genes are coded on the “+” strand. Two 2 bp overlaps were
128 detected by MITOS2, between *nad1-trnR* and *cob-trnD*. Global A+T content is
129 56.3%. No heteroplasmy was detected with the two MAF settings used in
130 NOVOPlasty.

131 Protein-coding genes

132 All 13 PCGs commonly occurring in metazoan mitochondrial genomes were
133 identified in *Mim. varia*. The combined length of PCGs was 10,896 bp (45.4% of the
134 complete genome). The total amino acid (AA) length was 3619, excluding stop
135 codons. The most frequently used AAs were Leucine (12.86%), Valine (10.39%),
136 Phenylalanine (9.55%), Glycine (9.45%) and Serine (8.31%). Six genes (*cox1*, *atp8*,
137 *cob*, *atp6*, *cox2*, *nad4*) had the most common start codon ATG, three had ATT
138 (*nad6*, *nad3*, *nad5*), three TTG (*nad1*, *nad4l*, *cox3*) and *nad2* had the ATA start
139 codon (Table 1). Eight genes had the TAG and five had the TAA stop codon.

140 rRNA and tRNA genes

141 The *rrnS* and *rrnL* genes were 964 bp and 1,422 bp long, respectively. Out of the
142 total 23 tRNAs identified, most were identified by MITOS2, except *trnQ2*, which was
143 identified by ARWEN. The length of tRNAs ranges from 63 to 73 nt. Three tRNA
144 genes are present in two copies, *trnL* with different anticodons (*trnL1*: UAG, *trnL2*:
145 UAA), *trnM* and *trnQ* with the same anticodon in both copies, CAU and UUG,
146 respectively. All tRNAs were predicted to have the typical cloverleaf structure, except
147 *trnQ2* (Additional File 1). While present in two copies in most animal mitogenomes,
148 *trnS* was not identified in the mitogenome of *Mim. varia*.

149 Gene order

150 The *Mim. varia* mitogenome represents a novel gene order for the Pectinidae. The
151 conserved pectinid gene clusters "*nad1-trnR-rnrL-cox1*" and "*nad4-trnH-trnW*" are
152 present, however, the "*nad6-trnL2-cob*" cluster is split by the insertion of *trnF*, the
153 MNR and *atp8*, and the "*cob-cox2*" cluster is also missing. Notable is the relocation
154 of the "*rnrS-nad5*" pair, and the location of the MNR between the *nad6* and *cob*
155 genes, a pattern observed in only one other species, *Pl. magellanicus*.

156 Structure of the MNR

157 The major non-coding region (MNR) spans 4,206 nt and has an elevated A+T
158 content of 60.1%. There were four repeated regions, with repeat sizes between 10
159 and 683 bp (Table 2). The largest repeat region contained 4.7 copies of a 683 bp
160 long sequence, with a total length of 3,205 bp, making up 76.2% of the MNR. We
161 PCR-amplified the MNR but failed to obtain its complete sequence using Sanger
162 sequencing, possibly because of numerous stem-and loop structures and high %AT
163 in this region (Additional file 2).

164 **Annotation of pectinid genomes**

165 Out of the 29 Pectinidae and 1 outgroup (*O. lurida*) mitogenomes retrieved from
166 Genbank, only two were incomplete. All protein coding and rRNA genes identified in
167 previous descriptions of these genomes were found and reannotated in this study.
168 The *atp8* gene was annotated in all 25 genomes, in which it was not previously
169 annotated, including two paralogous copies in *Mizuhopecten yessoensis* and *Pl.*
170 *magellanicus*, and unusually long versions of the gene in two *Ar. irradians* GenBank

171 accessions—with alternative start codons—showing only 78.5% sequence similarity
172 with conspecific *atp8* genes (Table 3, Additional file 3). Between 20 and 40 tRNA
173 genes were annotated per mitogenome in the 29 pectinid species. Most
174 mitogenomes contained 23 tRNA genes, compared to the 22 commonly found in
175 animals. Gene content was mainly identical within species, but there are slight
176 variations, some of which bears a phylogenetic pattern. For example, the *trnS1*
177 genes in *Amusium pleuronectes*, *Pecten albicans* and *Pecten maximus* all have
178 GCU as anticodon, and all species in the Chlamydinae subfamily contain two copies
179 of the *trnM* gene. Some variations are present in only one species, *Pe. albicans*
180 missing *trnL1* and *trnM*, *Mimachlamys nobilis* containing an extra *trnT* gene, *Miz.*
181 *yessoensis* lacking five tRNA genes, while containing two supernumerary ones, and
182 *Pl. magellanicus* containing 12 *trnM*, 4 *trnF* and 3 *trnS2* genes (Fig. 2), although
183 most of these were identified as pseudogenes by Smith et al. [20].

184 **Mito-phylogenomics**

185 All four Subfamilies and 7 out of 9 Tribes of the Pectinidae family are represented in
186 this study (Table 4). Two different topologies can be observed among the ML trees
187 reconstructed from the eight datasets (Fig. 3). The two rRNA datasets (G and H)
188 show a distinct topology compared to the other 6 datasets, but the topology based
189 on the rRNA datasets is not well-supported. All PCG ML trees had very similar
190 topology. *Mim. varia* is placed within the same clade with the other two congeneric
191 species only in the tree based on the E dataset (PCG amino acid sequences). In
192 every other tree *Mim. varia* is placed in a clade with species in the Crassadomini and
193 Chlamydini tribes. The monophyly of neither of these clades is supported. Similarly,

194 although the topology suggests that *Pl. magellanicus* is sister to all other pectinids,
195 its placement is never supported as such, rather the topology of the Pectinidae
196 should be considered as a trichotomy represented by the subfamilies Palliolinae,
197 Chlamydinae and a Pectininae+Aequipectini group. The B dataset resulted in the
198 most well-supported tree (Fig. 4A). All eight ML trees presented with proportional
199 branch lengths are shown in Fig. 3.

200 Every Bayesian phylogenetic reconstruction run on every dataset reached
201 convergence (ESS>200). The results of these analyses slightly differ from the results
202 of the ML method (Fig. 4B). Support values are generally higher than in the ML
203 approach. *Pl. magellanicus* is placed in the same clade with Chlamydinae in the
204 trees B, E, F, G and H with moderate to high posterior probabilities (0.81-0.98). In
205 the rRNA trees (G and H), there is a trichotomy at the root of the Pectinidae, where
206 the three clades are formed by Palliolinae+Chlamydinae, Pectininae and
207 Aequipectini groups, respectively. The placement of *Mim. varia* is identical to its
208 placement in the ML trees. All eight Bayesian trees presented with proportional
209 branch lengths are shown in Fig 5.

210 **Genome rearrangements**

211 Gene collinearity

212 Fifteen Locally Collinear Blocks were identified among 13 (1 per species) pectinid
213 mitogenomes (Fig. 6). Most of these 15 blocks were present in all mitogenomes,
214 except for LCB2 (missing from *Mim. nobilis*, *Mimachlamys senatoria*, *Pe. albicans*
215 and *Pl. magellanicus*), LCB9 and LCB14 (both missing from *Pl. magellanicus*). Ten

216 blocks contained one major gene (PCG or rRNA) and none or some tRNA genes.
217 LCB6 and LCB9 contained one or two PCGs and none or some tRNA genes, LCB7
218 two to four PCGs and some tRNAs, while LCB2 and LCB4 both contained a set of
219 tRNA genes only. LCBs 4, 7 and 12 were the most variable in terms of length (Fig.
220 7), due to the fact that these three blocks contained the Major Non-coding Region in
221 8 out of the 13 species (Fig. 6). The order of LCBs was generally more similar in
222 more closely related species, and identical within *Argopecten*, between *Mim. nobilis*
223 and *Mim. senatoria* and between *Chlamys farreri* and *Crassadoma gigantea*. *Mim.*
224 *varia*, *Miz. yessoensis* and *Pl. magellanicus* presented substantially different genome
225 organization from what could be expected by looking at their phylogenetic position
226 alone. When we look at only a subset of the whole species list, several larger blocks
227 can be observed. Larger blocks present in more than two species are the following:
228 LCBs 5, 1 and 13 are present in the same order in 11 species (different in *Mim.*
229 *yessoensis* and *Pl. magellanicus*); LCBs 5, 1, 13, 7, 8, 9 and 10 form a single block
230 in the Pectininae and the three *Argopecten* species; LCBs 10, 14, 8, 11, 15, 12 and 4
231 form a block in the Chlamydinae.

232 Common interval analysis

233 Tandem-duplication/random loss and transposition events occurred seven times
234 along the pectinid phylogeny (Fig. 8). Out of the 7 internal nodes, only one (the
235 common ancestor of chlamydins) was consistent, with a high level of certainty, two
236 were k-consistent, i.e., less reliable, with an intermediate level of certainty (the
237 common ancestor of Chlamydinae+*Placopecten* and the common ancestor of

238 Pectininae and *Argopecten*) and the rest were fallback nodes (representing the
239 highest level of uncertainty), including the common ancestor of Pectinidae.

240 Three genes – *nad1*, *rnl* and *cox1* – were involved in the fewest rearrangement
241 events (Additional file 4).

242 Maximum likelihood analysis of gene orders

243 In the maximum likelihood phylogeny reconstructed from gene orders, *Argopecten* is
244 sister to pectinins, as in the eight ML and eight Bayesian trees, however, within
245 Pectininae, *Am. pleuronectes* and *Pe. maximus* form the crown (Fig. 9). The other
246 major differences are the positions of *Mim. varia* and *Miz. yessoensis*, the latter
247 becoming the basal species in the Chlamydinae and *Mim. varia* becoming sister to
248 the crown of the Chlamydinae, formed by the pairs of *Mim. nobilis* – *Mim. senatoria*
249 and *Cr. gigantea* – *Ch. farreri* (Fig. 9).

250 **Discussion**

251 To the best of our knowledge, our study presents the most complete mitochondrial
252 comparative- and phylogenomic analysis to date in the Pectinidae family, including
253 the newly-assembled mitogenome of *Mim. varia*. The composition of the latter
254 conforms to most other pectinid mitogenomes published so far. However, it
255 represents a completely novel gene order, previously not described in any other
256 species of Pectinidae [13, 14, 18, 20–23, 27].

257 **Gene content**

258 In addition to the 12 PCGs present in every bivalve mitogenomes, the *atp8* gene was
259 annotated in *Mim. varia*. This is not surprising, given that *atp8* is being annotated in
260 more and more newly published bivalve mitogenomes, as well as identified in
261 mitogenomes where it was originally thought to be missing [15]. In our study, the
262 *atp8* gene was annotated in 25 mitogenomes in which it was not previously found, in
263 two species (*Miz. yessoensis*, *Pl. magellanicus*) it was even detected in two
264 paralogous copies, and in two accessions of *Ar. irradians*, only as an unusual
265 version of the gene (alternative start codon and a total length of 222 bp),(Table 3).
266 The *atp8* gene is very short (135 – 222 bp in the Pectinidae), and is among the most
267 variable mitochondrial PCGs, making it very difficult to detect. The discovery of a
268 functional *atp8* gene in most pectinids hints at it not being accessory, as previously
269 proposed [51–53]. Duplicated genes in mitochondrial genomes are not unheard of
270 [54–56], although whether all copies are functional remains unknown.

271 While it is common for metazoan mitochondrial genomes to contain two trnL and
272 trnS genes, pectinid mitogenomes also commonly contain two trnM genes. In
273 accordance, the mitogenome of *Mim. varia* contains two trnL and trnM genes,
274 however, it contains no trnS gene, and two trnF and trnQ genes. The two trnF genes
275 show low sequence similarity and trnF (AAA) shows low sequence identity to trnF
276 genes in other species as well as every tRNA gene in *Mim. varia*. Although with our
277 methodology, trnQ2 was annotated in *Mim. varia*, this might be an artifact. We base
278 this assumption on the fact that the two trnQ genes in *Mim. varia* are found next to
279 each other, in the gene block (trnK, trnF, trnQ2, trnQ1, trnE), but the same gene
280 block, containing trnS2 instead of trnQ2, is identified in *Mim. senatoria* [21] and *Cr.*
281 *gigantea* [18], and revealed during reannotation in the current study in *Mim. nobilis*,

282 *Ch. farreri* and *Miz. yessoensis* (Fig. 2, Additional file 3). Also, this trnQ2 gene shows
283 higher sequence similarity to the trnS2 gene than to the trnQ gene in the
284 aforementioned species. While we assume that this difference is the result of an
285 artifact, it is not impossible that in *Mim. varia*, the trnS gene evolved into trnQ. This is
286 supported by the fact that tRNA genes are often lost in mitochondrial genomes, and
287 that remolding of tRNA anticodons is known to happen in mollusks [9, 21, 57].

288 **Structure of the MNR**

289 High A+T content and stem-loop secondary structures are the common diagnostic
290 traits in identifying the mitochondrial control region [58–60]. Both can be observed in
291 the MNR of *Mim. varia* (Fig. 1, Additional file 2), pointing towards it serving as the
292 control region.

293 **Annotation of pectinid genomes**

294 The fact that we were able to annotate both the atp8 gene and some new tRNA
295 genes in most pectinid accessions downloaded for this study show that the lack of
296 these genes were most likely artifacts emerging from the difficulty of annotating
297 these features in bivalves. With the expansion of annotation databases and the
298 advancement of annotation tools, these difficulties are getting easier and easier to
299 overcome, leading to progressively more precise annotations.

300 While variations in tRNA gene content presented by singular species (especially
301 those with only one mitogenome annotated: *Pe. albicans* and *Mim. varia*) might well
302 be resulting from sequencing errors, incomplete genomes, or simply misannotations.
303 While the possibility of these cannot be completely excluded, even in the case of

304 species that have multiple genomes published (*Mim. nobilis*, *Miz. yessoensis*, *Pl.*
305 *magellanicus*), those presenting a phylogenetic pattern are most probably real. A
306 peculiar case is that of *trnM*, as it is present in more than half of all studied species
307 (Fig. 2). This gene is present in two copies in several invertebrate mitogenomes [61,
308 62], including some bivalves [25]. The invertebrate mitochondrial genetic code
309 contains five start codons, most of which code for amino acids other than methionine
310 in vertebrate mitochondrial genomes. This means that *trnM* must do 'double duty' as
311 the tRNA for both methionine and formyl-methionine, matching five instead of two
312 start codons [58], which could explain why some invertebrates, including some
313 pectinids evolved to have more copies of it in their mitochondrial genomes.

314 Numerous pectinid mitogenomes are longer than the usual metazoan mitochondrial
315 genome (~16 kbp), for example *Aequipecten opercularis* 21–28.2 kbp, *Chlamys*
316 *hastata* 23.9–27.2 kbp, *Chlamys islandica* 22.5–25 kbp, *Cr. gigantea* 22.8–24.8 kbp
317 [63], *Mim. varia* 20.4 kbp (the current study), *Miz. yessoensis* 20.5–21 kbp [14], *Pl.*
318 *magellanicus* 30.7–40.7 kbp [20], and *Pe. maximus* [22, 63, 64], although the
319 available mitogenome for the latter species is only 17.2 kbp long. This is mostly due
320 to inflated non-coding regions, as demonstrated by Ghiselli et al. [58] in the whole
321 Mollusca phylum. We identified tandem repeats of total length larger than 500 bp in
322 the non-coding regions of *Pl. magellanicus*, *Ch. farreri* (compare with [13]), *Mim.*
323 *varia* and *Am. pleuronectes* (Additional file 5) and repeats of around 500 bp length
324 scattered throughout the mitogenome of *Miz. yessoensis* (Additional file 3). Gjetvaj et
325 al. [63] also found that most of the pectinid mitogenomes they studied, contain
326 tandem repeats of various sizes and repeat numbers in their non-coding region, and
327 they found no significant sequence similarity among these repeats. This supports the

328 assumption that the repeats arose independently in every lineage. There is also
329 apparent intraspecific variation in mitochondrial genome size within the Pectinidae
330 [63–65].

331 While it is likely that correlation exist between the number and length of tandem
332 repeats and genome size, the available number of pectinid mitogenomes does not
333 allow to establish correlations between these traits.

334 The presence of multiple repeats make it seem very likely that tandem duplication is
335 frequent in pectinid mitogenomes, perhaps enabled by the fact that all genes are
336 coded on the “+” strand. However, this does not seem to influence the mitogenome
337 architecture of other marine bivalves, most of which also code all genes on the “+”
338 strand [13]. In plants, which usually have significantly larger mitogenomes than
339 animals and contain numerous repeated sequences, it is suggested that these
340 repeats promote recombination [66–68]. This might play a role in the evolution of the
341 mitochondrial genome in the Pectinidae as well.

342 The prevalence of non-coding regions and, in some instances, extra tRNAs in
343 Pectinidae make this group an interesting candidate to study the taxonomic
344 distribution of small mitochondrial highly-transcribed RNAs (smithRNAs, [69]). This
345 novel class of sncRNAs (short non-coding RNAs), initially described from the Manila
346 clam *Ruditapes philippinarum* (class Venerida) and detected in zebrafish, fruitflies
347 and mice [70], could influence the expression of nuclear-encoded genes and be
348 involved in gonad formation. They were quasi-exclusively found in non-coding
349 regions and tRNAs [69].

350 Mito-phylogenomics

351 The phylogenies presented in our study are in accordance with previous results, both
352 with those using a few genes, and those using the complete set of mitochondrial
353 PCGs. One novel result is that in this group, the two rRNA genes seem to have little
354 power to resolve deep divergences, as seen in the different placement of the
355 Pectininae in both the ML and Bayesian approaches. In accordance with Puslednik
356 et al. [40] and Alejandrino et al. [31], we found that *Mim. varia* is not placed into one
357 monophyletic clade with the other two *Mimachlamys* species. Generally, although
358 Chlamydinae itself is monophyletic, its lower taxonomic levels are not well-resolved,
359 and several genera are paraphyletic, as it was presented in previous studies [22, 27,
360 31, 40, 71]. In contrast to Puslednik et al. [40], who concluded Aequipectini to be the
361 basal clade in Pectinidae (similarly as shown in our rRNA datasets, G and H), we
362 have found that Aequipectini always form a monophyletic group with Pectininae in
363 both of our ML and Bayesian approaches, when we included PCGs in the analysis. It
364 appears that while mitochondrial rRNA genes and the nuclear H3 gene are in
365 accordance [40], mitochondrial PCGs paint a slightly different picture, as can also be
366 seen in, for example, Marín et al. [22] and Yao et al. [27]. However, this could at
367 least be partly caused by the small number of taxa involved in our study, and the
368 different choice of outgroups, as both taxon sampling and outgroup selection are
369 known to influence topologies [40]. Despite the Pectinidae being one of the largest
370 families in Bivalvia, containing around 350 species [72], existing phylogenies usually
371 contain only a handful of species, with the most making up only 31% of all species
372 [31]. Chlamydinae is an especially problematic group within the Pectinidae, perhaps
373 partly because of the low sampling relative to the high number of species in the

374 group. While Puslednik et al. [40] conclude that Chlamydinae—in contrast to Waller’s
375 hypothesis (i.e., that Chlamydinae is paraphyletic, and have provided the ancestral
376 stock for Palliolinae and Pectininae)—is in fact the crown group of the Pectinidae.
377 However, their more recent study [31], doubled taxon sampling and used one more
378 nuclear gene (28S), and subsequently came to the opposite conclusion, confirming
379 Waller’s original hypothesis. Another, recent study by Smedley et al. [71] included
380 yet another nuclear gene (18S), and—as they were studying the Pectinoidea—
381 numerous outgroup species to Pectinidae. They revealed a similar topology to
382 Alejandrino et al. [31], but with substantially lower support. This shows the
383 importance of marker selection and appropriate taxon sampling on pectinid
384 phylogenetics.

385 The phylogeny reconstructed from gene order data shows a similar, although not
386 identical topology to the one reconstructed from sequence data. The difference
387 between the two trees within the Pectininae can be explained with the likely
388 incompleteness of the *Pe. albicans* mitogenome, missing a few tRNA genes. In
389 accordance with our phylogenomic results, the gene order phylogeny fails to properly
390 resolve Chlamydinae, which, again, can be attributed to low taxon sampling within
391 this group.

392 The lack of sampling is even more prominent if we look at phylogenomic studies, as
393 the number of published pectinid mitogenomes is currently low and only slowly
394 growing. Although our study does not resolve this problem, it extends the list of
395 published pectinid mitogenomes, shows the utility of mito-phylogenomics within the

396 Pectinidae and expands our knowledge of the evolutionary history of mitochondria
397 within this remarkable group of bivalves.

398 **Genome rearrangements**

399 Gene collinearity

400 According to the 'punctuation model' [73], mitochondrial genomes are generally
401 transcribed as a single polycistronic RNA from each strand, followed by the
402 enzymatic removal of tRNAs, leading to gene-specific mRNAs [74]. In pectinids,
403 major genes are not always separated by tRNAs (Fig. 6). It is possible that these
404 cistrons are punctuated by secondary structures instead of tRNAs, or that they
405 remain bicistrons. However, given that LCBs are the units of genome rearrangement,
406 and most of them contain one major gene and some tRNA genes, it is probable that
407 these are all separated during mRNA maturation. The one exception from this rule is
408 LCB7, that contains the *nad6* and *cob* genes in every studied pectinid species but
409 contains the *atp8* gene in addition to these in *Mim. varia* and
410 Pectininae+Aequipectini, although in different position, and *atp8* together with *atp6* in
411 *Pl. magellanicus* (Fig. 6, Additional file 3). Nevertheless, the variance in the
412 composition of this block points towards it not being a single unit in rearrangements.
413 The order of LCBs also shows a clear phylogenetic pattern. The divergence from this
414 pattern in the case of the three outlier species, *Mim. varia*, *Miz. yessoensis* and *Pl.*
415 *magellanicus* is most likely stemming from limited taxon sampling, i.e. they are
416 relatively divergent from their closest relatives in our sampling, and their true close
417 relatives would most likely show a similar gene order to theirs, as evidenced in other,
418 better sampled groups in our study (e.g. *Argopecten*).

419 Common interval analysis

420 To the best of our knowledge, our study is the first attempt to reconstruct, at least
421 partially, the ancient pectinid mitogenome. The CREx method is widely used in other
422 animal groups, for example insects, e.g. [75, 76]. While the ancestral gene order
423 could be inferred for Insecta, shared by most groups, and some lineages showing
424 some rearrangement, this is not the case for Mollusca. *Katharina tunicata* [17, 77]
425 and *Solemya velum* [30] are hypothesized to carry a mitochondrial genome similar in
426 organization to the ancestral Molluscan and Conchiferan mitogenomes, respectively.
427 However, pectinid mitogenomes are very divergent from these two species, and from
428 closely related groups with available mitogenomes, which severely limits the
429 effectiveness of using these as outgroups, given the large chance for homoplasy.
430 Our method tries to overcome this difficulty, with first reconstructing the putative
431 gene order of internal nodes, including the common ancestor of the Pectinidae,
432 followed by the inference of rearrangement events. Although some internal nodes
433 were not consistent, we recovered the putative gene order of the common ancestor
434 of the Chlamydiae, a monophyletic group in most phylogenetic analyses, including
435 ours. While most previous studies focusing on mitochondrial genome
436 rearrangements in the Pectinidae involved only a few species, we compare genome
437 rearrangements among all available pectinid mitogenomes, with methods not
438 previously used in this group. Confirming the assumption of Marín et al. [22], we
439 have found that TDRL (Tandem Duplication/Random Loss) events are equally
440 important as transpositions in pectinid mitogenome rearrangements.

441

442 **Conclusions**

443 Scallop fisheries are commercially important and steadily growing worldwide [78, 79].
444 Considering that mitochondrial function is implicated in growth and environmental
445 resilience [80, 81], it is important to understand the evolution of the unusual
446 mitochondrial genomes of pectinids, as it would aid fisheries management in our
447 changing climate.

448 In our study, we annotate and characterize the mitochondrial genome of
449 *Mimachlamys varia*, we demonstrate that complete mitochondrial genomes are
450 powerful resources in reconstructing pectinid phylogenies, and we also show the
451 utility of several tools, not previously used in this group, to investigate mitochondrial
452 genome rearrangements within the Pectinidae. We present annotations of the *atp8*
453 gene in several pectinid species, where it was previously thought to be missing.

454 We show incongruities between phylogenies based on mitochondrial PCGs and
455 rRNA genes, suggesting that locus sampling affects phylogenetic inference at the
456 scale of the Pectinidae family. Both the lack of suitable outgroups—more closely
457 related to the Pectinidae (e.g. Limidae, Entoliidae, Spondylidae and
458 Propeamussiidae) than the currently available species—and the lack of sufficient
459 taxon sampling, especially within Chlamydinae and Palliolinae, are limiting factors for
460 current research. We therefore believe that sequencing and assembly of more
461 mitogenomes, particularly from the aforementioned groups, would greatly improve
462 the power of further research of both mito-phylogenomics and the evolution of
463 mitogenome rearrangements in the Pectinidae.

464 **Methods**

465 **Sequencing of the mitogenome of reference for the variegated scallop**

466 One individual of *Mim. varia* (adult male, shell length: 46 mm, shell height: 40mm)
467 was collected in Angoulins, France, in order to sequence and annotate a reference
468 transcriptome (nuclear and mitochondrial) for the species. RNA was purified from
469 five tissues (digestive gland, mantle, gills, adductor muscle and gonads) using the
470 Nucleospin RNA Set for Nucleozol kit (Macherey-Nagel). After quality control,
471 extractions were pooled in equal amounts (4 µg of RNA per tissue type). The library
472 was sequenced on Illumina HiSeq 2500 (2 × 300 bp). Full sample preparation and
473 sequencing information is provided in Viricel et al. [50]; raw reads can be accessed
474 on the NCBI Sequence Read Archive (Acc. SRP127478), and transcripts can be
475 accessed at the Transcriptome Shotgun Assembly database (GGGO01000000).

476 Paired reads, minimally cleaned with Trimmomatic v. 0.36 ([82]; only adapters were
477 removed as recommended in the NOVOPlasty manual), were used as input data for
478 assembly with NOVOPlasty v. 4.2 [83] to assemble the mitochondrial genome.
479 Quality-control before and after read filtering was done using FastQC v0.11.5 [84].
480 The size of the mitogenome being highly variable among pectinids (from 16,079 bp
481 in *Argopecten ventricosus*, KT161261.1, to 32,115 bp in *Placopecten magellanicus*,
482 NC_007234.1, [85]), we set a wide range for possible mitogenome sizes (15–35 kbp)
483 as a parameter in NOVOPlasty. K-mer size was set to 39. We used a portion of the
484 mitochondrial gene *cox1* as an initial seed for mitogenome assembly (we used the
485 most common haplotype found along the coast of France, KU680872.1; [49]). When

486 NOVOPlasty runs resulted in multiple contigs (mitogenome fragments), they were
487 assembled in Geneious v2021.0.3 (<https://www.geneious.com>). Reads were mapped
488 onto the assembly using bowtie2 v2.3.5.1 [86] and coverage statistics were
489 generated with samtools v1.10 [87]. Once the reference mitogenome was
490 assembled, we used NOVOPlasty to detect intra-individual heteroplasmy [88] using
491 a minimum Minor Allele Frequency (MAF) of 0.1 and 0.01 in two separate runs.
492 Finally, the MITOS2 web server (<http://mitos2.bioinf.uni-leipzig.de/index.py>) [89] was
493 used to annotate the reference with default parameters, using the RefSeq 63
494 Metazoa database as reference and the protein-coding gene prediction method of Al
495 Arab et al. [90]. The mitogenome was loaded into Geneious v2021.0.3. and checked
496 manually for errors in annotations. ARWEN v1.2.3 [91], was used to look for
497 additional tRNA genes and secondary structures of tRNAs were predicted using
498 MITOS2 and the RNAfold web server, with default options. The Major Non-coding
499 Region (MNR) was scanned for repeated regions using Tandem Repeats Finder v
500 4.09 (<https://tandem.bu.edu/trf/trf.html>) [92] and secondary structures were predicted
501 using the mfold Web Server (<http://www.unafold.org/>) [93]. The circular mitogenome
502 was visualised using the CGView Web Server (<http://cgview.ca/>) [94], with
503 annotations imported from Geneious.

504 **Assembly and annotation of pectinid mitogenomes from GenBank.**

505 All currently available pectinid mitogenomes (as of January 2021) were downloaded
506 from GenBank (Table 4) along with *Ostrea lurida* (Ostreidae) as an outgroup. We
507 chose an oyster species as outgroup as there are no published mitogenomes of taxa
508 more closely related to Pectinidae. Using only one outgroup made interpreting

509 comparisons of mitogenome rearrangements substantially less difficult. The largest
510 mitogenome from each species was reannotated using the MITOS2 web server with
511 the same settings as for the *Mim. varia* mitogenome. New annotations of pectinid
512 mitogenomes were reviewed manually, in the same way as it was described for *Mim.*
513 *varia*. If there were multiple annotations for a gene, they were filtered manually
514 based on length and sequence similarity to other species. Additionally, tRNA genes
515 were compared with previously published annotations. Next, this reference sequence
516 of each species was used to annotate other sequences of the same species using
517 the "Live Annotate & Predict" feature in Geneious.

518 **Mito-phylogenomics**

519 Phylogenetic analysis was conducted on the 13 protein coding genes (PCGs) and
520 two Ribosomal RNA (rRNA) genes of all available pectinid species in GenBank, *Mim.*
521 *varia* and *O. lurida* as an outgroup. This approach provides the most comprehensive
522 dataset without the ambiguities arising from tRNA annotations, non-coding regions
523 and genome rearrangements. Every gene was extracted and aligned separately in
524 Geneious v2021.0.3. In the case of the *atp8* gene, where more copies were
525 identified in some species, a Maximum Likelihood (ML) tree was built with the
526 RAxML [95] plugin of Geneious using default parameters, and the least divergent
527 copy was selected for further analysis. The *atp8* gene of two accessions of
528 *Argopecten irradians* (EU023915; NC_009687) were not included in the phylogenetic
529 analyses given their unusual structure, and divergent sequences. The 13 PCG
530 sequences were aligned separately based on translated amino-acids with MAFFT
531 [96] with the FFT-NS-2 algorithm as implemented in TranslatorX [97]. As the initial

532 alignments contained indels and poorly aligned regions, these regions were masked
533 with Gblocks [98] as implemented in TranslatorX, with default settings (i.e., allowing
534 smaller final blocks, gap positions within the final blocks and less strict flanking
535 positions). The two rRNA genes were aligned with MAFFT in Geneious using the
536 FFT-NS-2 algorithm. Regions of uncertainty in the alignment were then masked with
537 Gblocks using the same settings as for the PCGs. Individual gene alignments were
538 then concatenated in Geneious resulting in eight datasets used for phylogenetic
539 analyses: A) all PCG and rRNA genes treated with Gblocks; B) all PCG amino acid
540 sequences and rRNA genes treated with Gblocks; C) all PCGs; D) all PCGs treated
541 with Gblocks; E) all PCG amino acid sequences; F) all PCG amino acid sequences
542 treated with Gblocks; G) rRNA genes; H) rRNA genes treated with Gblocks. These
543 datasets were prepared in order to investigate the effect of rRNA and protein-coding
544 genes as well as the presence/absence of ambiguously aligned regions on the
545 topology and branch support of the phylogeny.

546 Maximum Likelihood phylogenetic reconstruction was achieved using IQtree v2.0.6
547 [99] (except for the B dataset, where IQtree v1.6.12 was used), with automatic model
548 selection and using the best partition scheme in ModelFinder (option MFP+MERGE)
549 [100], 10,000 Ultrafast Bootstrap replicates and 10,000 replicates of the SH-like
550 approximate likelihood ratio test (SH-aLRT) [101] on all eight datasets. Partitioning
551 and substitution model information for every run can be found in Additional file 6.

552 Bayesian analyses were done using MrBayes v3.2.7 [102]. For each dataset, two
553 runs were performed over 1.5 million generations, with four chains each (three
554 heated and one cold). Trees were sampled every 500 generations, and the first 25%

555 were discarded as burn-in. Finally, a majority consensus tree was constructed to
556 estimate posterior probabilities of branches. Convergence of each run was evaluated
557 through Effective Sample Size (ESS) values and trace plots in the software Tracer
558 v.1.6 [103]. Partitioning and prior model parameter settings can be found in
559 Additional file 7.

560 **Genome rearrangements**

561 Three approaches were used to infer mitogenome rearrangements in the Pectinidae:
562 1) gene collinearity analysis, which detects mosaic patterns of homology among a
563 set of genomes, based on whole-genome sequence alignment; 2) common interval
564 analysis, which determines pairwise rearrangement events between genomes,
565 based on gene order data, considering several rearrangement types, but requiring
566 identical gene content; and 3) maximum likelihood analysis of gene orders, which
567 reconstructs phylogenies based solely on gene order data, without requiring an
568 identical set of genes between genomes.

569 Gene collinearity

570 Gene collinearity among pectinid mitogenomes (one individual/species, 13
571 individuals in total) was explored using the progressiveMauve algorithm and default
572 parameters (automatically calculating seed weight and the minimum LCB score,
573 computing Locally Collinear Blocks (LCBs) and using full alignments) in the MAUVE
574 Geneious plugin [104].

575 Common interval analysis

576 Two joint softwares – both relying on the same algorithm – CREx [105] and TreeREx
577 [106] were implemented for common interval analyses. Both software consider the
578 following types of rearrangement events: transpositions, inversions, inverse
579 transpositions and tandem-duplication/random loss (TDRL) involving the duplication
580 of a gene block followed by the random loss of one copy of a gene in either of the
581 two blocks. Both tools are only able to handle identical gene sets among species. To
582 conform to this condition, only major genes (PCGs and rRNA genes) and the MNR
583 were included in this analysis, as tRNA gene content is variable among pectinid
584 mitogenomes.

585 The MrBayes tree based on the most supported dataset (B), with only 8 pectinids
586 and one outgroup species representing unique gene orders, along with the gene
587 order dataset described above, was used in the first step, using the TreeREx [106]
588 software. TreeREx uses the phylogenetic information in the tree to ascertain
589 rearrangement events among nodes of the tree, and to infer putative gene orders of
590 common ancestors at internal nodes. TreeREx was used with default settings: strong
591 consistency method applied (-s); weak consistency method applied (-w);
592 parsimonious weak consistency method applied (-W); and the maximum number of
593 inversions (-m) set to 2. Finally, the input dataset in the previous analysis,
594 complemented with the putative gene orders inferred at internal nodes was put into
595 CREx again, to visualize each rearrangement event between each pair of
596 neighboring nodes.

597 Maximum likelihood analysis of gene orders

598 The MLGO (Maximum Likelihood analysis of Gene Order) [107] web server can be
599 used to directly infer a phylogeny based on gene order data. The algorithm handles
600 insertions, deletions and duplications along with transpositions. Unlike
601 CREx/TreeREx, it is not limited to a common set of genes across taxa. Complete
602 gene order data from 13 pectinids (one individual per species) and one outgroup (*O.*
603 *lurida*) were used to infer a phylogeny of the Pectinidae. Branch support was
604 computed with a bootstrap analysis of 1000 replicates.

605 **Declarations**

606 **Ethics approval and consent to participate**

607 No vertebrates or regulated invertebrates were involved in this study.

608 **Consent for publication**

609 Not applicable.

610 **Availability of data and materials**

611 Raw *Mimachlamys varia* RNAseq reads are deposited in the NCBI Sequence Read
612 Archive (Acc. SRP127478), and transcripts can be accessed at the Transcriptome
613 Shotgun Assembly database (GGGO01000000). The assembled mitochondrial
614 genome is uploaded to GenBank (MZ520326), and the GenBank accession numbers
615 of all mitogenomes used in this study can be found in Table A. The sequence
616 alignments used this article will be uploaded upon acceptance to a public repository.

617 **Competing interests**

618 The authors declare no competing interests.

619 **Funding**

620 The first author thanks the Erasmus+ programme of the European Union for funding
621 his mobility and stay at the LIENSs laboratory in La Rochelle. This work was
622 supported by the Contrat de Plan Etat-Région (CPER/FEDER) ECONAT (RPC
623 DYPOMAR).

624 **Authors' contributions**

625 ED and VB collected the samples, AV and VB performed laboratory analyses, TM,
626 EP and LE performed data analyses, TM wrote the manuscript, AV and EP reviewed
627 and edited the manuscript. All authors read and approved the final manuscript.

628 **Acknowledgements**

629 We thank Nathalie Imbert and Denis Fichet for coordinating axe 2 of the DYPOMAR
630 project and Benoit Simon-Bouhet for additional funding. RNA extractions were
631 prepared at the Molecular Core Facility at the University of La Rochelle. We thank
632 the bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo GenoToul).

633 **References**

634 1. Boore JL. Animal mitochondrial genomes. *Nucleic Acids Res.* 1999;27:1767–80.
635 doi:10.1093/NAR/27.8.1767.

636 2. Burger G, Forget L, Zhu Y, Gray MW, Franz Lang B. Unique mitochondrial
637 genome architecture in unicellular relatives of animals.
638 www.ncbi.nlm.nih.gov/PMGifs/Genomes/euko.html,. Accessed 9 Nov 2021.

639 3. Dreyer H, Steiner G. The complete sequence and gene organization of the
640 mitochondrial genome of the gadilid scaphopod *Siphonodontalium lobatum*
641 (Mollusca). *Mol Phylogenet Evol.* 2004;31:605–17.

642 4. Feldmeyer B, Hoffmeier K, Pfenninger M. The complete mitochondrial genome of
643 *Radix balthica* (Pulmonata, Basommatophora), obtained by low coverage shot gun
644 next generation sequencing. *Mol Phylogenet Evol.* 2010;57:1329–33.

645 5. Hatzoglou E, Rodakis GC, Lecanidou R. Complete sequence and gene
646 organization of the mitochondrial genome of the land snail *Albinaria coerulea*.
647 *Genetics.* 1995;140:1353–66. doi:10.1093/genetics/140.4.1353.

648 6. Liu YG, Kurokawa T, Sekino M, Tanabe T, Watanabe K. Complete mitochondrial
649 DNA sequence of the ark shell *Scapharca broughtonii*: An ultra-large metazoan
650 mitochondrial genome. *Comp Biochem Physiol - Part D Genomics Proteomics.*
651 2013;8:72–81.

652 7. Snyder M, Fraser AR, Laroche J, Gartner-Kepkay KE, Zouros E, Kafatos FC.
653 Atypical mitochondrial DNA from the deep-sea scallop *Placopecten magellanicus*.
654 *Proc Natl Acad Sci U S A.* 1987;84:7595–9. doi:10.1073/pnas.84.21.7595.

655 8. Xin Y, Ren J, Liu X. Mitogenome of the small abalone *Haliotis diversicolor* Reeve
656 and phylogenetic analysis within Gastropoda. *Mar Genomics.* 2011;4:253–62.
657 doi:10.1016/j.margen.2011.06.005.

- 658 9. Wu X, Li X, Li L, Xu X, Xia J, Yu Z. New features of Asian *Crassostrea* oyster
659 mitochondrial genomes: A novel alloacceptor tRNA gene recruitment and two novel
660 ORFs. *Gene*. 2012;507:112–8.
- 661 10. Skibinski DOF, Gallagher C, Beynon CM. Mitochondrial DNA inheritance [11].
662 *Nature*. 1994;368:817–8. doi:10.1038/368817b0.
- 663 11. Zouros E, Ball AO, Saavedra C, Freeman KR. Mitochondrial DNA inheritance
664 [12]. *Nature*. 1994;368:818. doi:10.1038/368818a0.
- 665 12. Zouros E. Doubly uniparental inheritance of mitochondrial DNA: Might it be
666 simpler than we thought? *J Zool Syst Evol Res*. 2020;58:624–31.
- 667 13. Ren J, Shen X, Jiang F, Liu B. The mitochondrial genomes of two scallops,
668 *Argopecten irradians* and *Chlamys farreri* (Mollusca: Bivalvia): the most highly
669 rearranged gene order in the family Pectinidae. *J Mol Evol*. 2010;70:57–68.
- 670 14. Wu X, Xu X, Yu Z, Kong X. Comparative mitogenomic analyses of three scallops
671 (Bivalvia: Pectinidae) reveal high level variation of genomic organization and a
672 diversity of transfer RNA gene sets. *BMC Res Notes*. 2009;2:1–7.
- 673 15. Plazzi F, Puccio G, Passamonti M. Comparative large-scale mitogenomics
674 evidences clade-specific evolutionary trends in mitochondrial DNAs of Bivalvia.
675 *Genome Biol Evol*. 2016;8:2544–64.
- 676 16. Liu H, Yang Y, Sun S, Kong L, Li Q. Mitogenomic phylogeny of the Naticidae
677 (Gastropoda: Littorinimorpha) reveals monophyly of the Polinicinae. *Zool Scr*.
678 2020;49:295–306. doi:10.1111/zsc.12412.

- 679 17. Vallès Y, Boore JL. Lophotrochozoan mitochondrial genomes. *Integr Comp Biol.*
680 2006;46:544–57. doi:10.1093/icb/icj056.
- 681 18. Liao D, Zhou Y, Tong J, Cao S, Yu X, Fu B, et al. Characterization and
682 phylogenetic analysis of the complete mitochondrial genome from Rock Scallop
683 (*Crassadoma gigantea*) using next-generation sequencing. *Mitochondrial DNA Part*
684 *B Resour.* 2018;3:827–8. doi:10.1080/23802359.2018.1483752.
- 685 19. Sato M, Nagashima K. Molecular characterization of a mitochondrial DNA
686 segment from the Japanese scallop (*Patinopecten yessoensis*): Demonstration of a
687 region showing sequence polymorphism in the population. *Mar Biotechnol.*
688 2001;3:370–9. doi:10.1007/s10126001-0015-4.
- 689 20. Smith DR, Snyder M. Complete mitochondrial DNA sequence of the scallop
690 *Placopecten magellanicus*: Evidence of transposition leading to an
691 uncharacteristically large mitochondrial genome. *J Mol Evol.* 2007;65:380–91.
- 692 21. Wu X, Li X, Yu Z. The mitochondrial genome of the scallop *Mimachlamys*
693 *senatoria* (Bivalvia, Pectinidae). *Mitochondrial DNA.* 2015;26:242–4.
- 694 22. Marín A, Fujimoto T, Arai K. The mitochondrial genomes of *Pecten albicans* and
695 *Pecten maximus* (Bivalvia: Pectinidae) reveal a novel gene arrangement with low
696 genetic differentiation. *Biochem Syst Ecol.* 2015;61:208–17.
697 doi:10.1016/j.bse.2015.06.015.
- 698 23. Xu K, Kanno M, Yu H, Li Q, Kijima A. Complete mitochondrial DNA sequence
699 and phylogenetic analysis of Zhikong scallop *Chlamys farreri* (Bivalvia: Pectinidae).
700 *Mol Biol Rep.* 2011;38:3067–74. doi:10.1007/s11033-010-9974-8.

701 24. Boore JL, Medina M, Rosenberg LA. Complete sequences of the highly
702 rearranged molluscan mitochondrial genomes of the scaphopod *Graptacme eborea*
703 and the bivalve *Mytilus edulis*. *Mol Biol Evol.* 2004;21:1492–503.
704 doi:10.1093/molbev/msh090.

705 25. Wu X, Xiao S, Li X, Li L, Shi W, Yu Z. Evolution of the tRNA gene family in
706 mitochondrial genomes of five *Meretrix* clams (*Bivalvia*, *Veneridae*). *Gene.*
707 2014;533:439–46.

708 26. Yamazaki N, Ueshima R, Terrett JA, Yokobori S, Kaifu M, Segawa R, et al.
709 Evolution of Pulmonate Gastropod Mitochondrial Genomes: Comparisons of Gene
710 Organizations of *Euhadra*, *Cepaea* and *Albinaria* and Implications of Unusual tRNA
711 Secondary Structures. *Genetics.* 1997;145.

712 27. Yao L, Yu H, Liu Y, Zhang Y, Li Y. The complete mitochondrial genome and
713 phylogenetic analysis of *Amusium pleuronectes*. *Mitochondrial DNA Part B.*
714 2020;5:2318–9. doi:10.1080/23802359.2020.1772691.

715 28. Bettinazzi S, Plazzi F, Passamonti M. The complete female- and male-
716 transmitted mitochondrial genome of *Meretrix lamarckii*. *PLoS One.*
717 2016;11:e0153631. doi:10.1371/journal.pone.0153631.

718 29. Gaitán-Espitia JD, Quintero-Galvis JF, Mesas A, D'Elía G. Mitogenomics of
719 southern hemisphere blue mussels (*Bivalvia*: *Pteriomorpha*): Insights into the
720 evolutionary characteristics of the *Mytilus edulis* complex. *Sci Rep.* 2016;6:1–10.
721 doi:10.1038/srep26853.

- 722 30. Plazzi F, Ribani A, Passamonti M. The complete mitochondrial genome of
723 *Solemya velum* (Mollusca: Bivalvia) and its relationships with Conchifera. BMC
724 Genomics. 2013;14:1–15. doi:10.1186/1471-2164-14-409.
- 725 31. Alejandrino A, Puslednik L, Serb JM. Convergent and parallel evolution in life
726 habit of the scallops (Bivalvia: Pectinidae). BMC Evol Biol. 2011;11:1–9.
727 doi:10.1186/1471-2148-11-164.
- 728 32. Canapa A, Barucca M, Marinelli A, Olmo E. Molecular data from the 16s rRNA
729 gene for the phylogeny of pectinidae (Mollusca: Bivalvia). J Mol Evol. 2000;50:93–7.
730 doi:10.1007/s002399910010.
- 731 33. Saavedra C, Peña JB. Phylogenetics of American scallops (Bivalvia: Pectinidae)
732 based on partial 16S and 12S ribosomal RNA gene sequences. Mar Biol.
733 2006;150:111–9. doi:10.1007/s00227-006-0335-z.
- 734 34. Feng Y, Li Q, Kong L, Zheng X. DNA barcoding and phylogenetic analysis of
735 Pectinidae (Mollusca: Bivalvia) based on mitochondrial COI and 16S rRNA genes.
736 Mol Biol Rep. 2011;38:291–9. doi:10.1007/s11033-010-0107-1.
- 737 35. López-Piñón MJ, Freire R, Insua A, Méndez J. Sequence characterization and
738 phylogenetic analysis of the 5S ribosomal DNA in some scallops (Bivalvia:
739 Pectinidae). Hereditas. 2008;145:9–19. doi:10.1111/j.0018-0661.2008.2034.x.
- 740 36. Mahidol C, Na-Nakorn U, Sukmanomon S, Yoosuk W, Taniguchi N, Nguyen
741 TTT. Phylogenetic relationships among nine scallop species (Bivalvia: Pectinidae)
742 inferred from nucleotide sequences of one mitochondrial and three nuclear gene
743 regions. J Shellfish Res. 2007;26:25–32. doi:10.2983/0730-
744 8000(2007)26[25:PRANSS]2.0.CO;2.

- 745 37. Malkowsky Y, Klussmann-Kolb A. Phylogeny and spatio-temporal distribution of
746 European Pectinidae (Mollusca: Bivalvia). *Syst Biodivers.* 2012;10:233–42.
747 doi:10.1080/14772000.2012.676572.
- 748 38. Matsumoto M. Phylogenetic analysis of the subclass Pteriomorpha (Bivalvia)
749 from mtDNA COI sequences. *Mol Phylogenet Evol.* 2003;27:429–40.
- 750 39. Pujolar JM, Marčeta T, Saavedra C, Bressan M, Zane L. Inferring the
751 demographic history of the Adriatic *Flexopecten* complex. *Mol Phylogenet Evol.*
752 2010;57:942–7.
- 753 40. Puslednik L, Serb JM. Molecular phylogenetics of the Pectinidae (Mollusca:
754 Bivalvia) and effect of increased taxon sampling and outgroup selection on tree
755 topology. *Mol Phylogenet Evol.* 2008;48:1178–88.
- 756 41. Saavedra C, Peña JB. Phylogenetic relationships of commercial European and
757 Australasian king scallops (*Pecten* spp.) based on partial 16S ribosomal RNA gene
758 sequences. *Aquaculture.* 2004;235:153–66.
- 759 42. Brugler MR, France SC. The complete mitochondrial genome of the black coral
760 *Chrysopathes formosa* (Cnidaria:Anthozoa:Antipatharia) supports classification of
761 antipatharians within the subclass Hexacorallia. *Mol Phylogenet Evol.* 2007;42:776–
762 88.
- 763 43. Lavrov D V., Wang X, Kelly M. Reconstructing ordinal relationships in the
764 Demospongiae using mitochondrial genomic data. *Mol Phylogenet Evol.*
765 2008;49:111–24.

- 766 44. Kayal E, Roure B, Philippe H, Collins AG, Lavrov D V. Cnidarian phylogenetic
767 relationships as revealed by mitogenomics. *BMC Evol Biol.* 2013;13:1–18.
768 doi:10.1186/1471-2148-13-5.
- 769 45. Brand AR. Chapter 12 Scallop ecology: Distributions and behaviour. *Dev Aquac*
770 *Fish Sci.* 2006;35 C:651–744.
- 771 46. Breitwieser M, Viricel A, Graber M, Murillo L, Becquet V, Churlaud C, et al. Short-
772 term and long-term biological effects of chronic chemical contamination on natural
773 populations of a marine bivalve. *PLoS One.* 2016;11:150184.
774 doi:10.1371/journal.pone.0150184.
- 775 47. Breitwieser M, Viricel A, Churlaud C, Guillot B, Martin E, Stenger PL, et al. First
776 data on three bivalve species exposed to an intra-harbour polymetallic contamination
777 (La Rochelle, France). *Comp Biochem Physiol Part - C Toxicol Pharmacol.*
778 2017;199:28–37.
- 779 48. Milinkovitch T, Bustamante P, Huet V, Reigner A, Churlaud C, Thomas-Guyon H.
780 In situ evaluation of oxidative stress and immunological parameters as
781 ecotoxicological biomarkers in a novel sentinel species (*Mimachlamys varia*). *Aquat*
782 *Toxicol.* 2015;161:170–5.
- 783 49. Breitwieser M, Becquet V, Thomas-Guyon H, Pillet V, Sauriau PG, Graber M, et
784 al. Population structure and genetic diversity in the variegated scallop, *Mimachlamys*
785 *varia* (Linnaeus, 1758), a novel bioindicator of chemical pollution on the French
786 coastline. *J Molluscan Stud.* 2018;84:417–25.

787 50. Viricel A, Becquet V, Dubillot E, Pante E. De novo assembly and functional
788 annotation of the transcriptome of *Mimachlamys varia*, a bioindicator marine bivalve.
789 *Mar Genomics*. 2018;41 April:42–5. doi:10.1016/j.margen.2018.04.002.

790 51. Gissi C, Iannelli F, Pesole G. Evolution of the mitochondrial genome of Metazoa
791 as exemplified by comparison of congeneric species. *Heredity*. 2008;101:301–20.
792 doi:10.1038/hdy.2008.62.

793 52. Papakonstantinou T, Law RHP, Nagley P, Devenish RJ. Non-functional variants
794 of yeast mitochondrial ATP synthase subunit 8 that assemble into the complex.
795 *Biochem Mol Biol Int*. 1996;39:253–60.

796 53. Stephens AN, Khan MA, Roucou X, Nagley P, Devenish RJ. The molecular
797 neighborhood of subunit 8 of yeast mitochondrial F1F0-ATP synthase probed by
798 cysteine scanning mutagenesis and chemical modification. *J Biol Chem*.
799 2003;278:17867–75. doi:10.1074/jbc.M300967200.

800 54. Bojko J. The mitochondrial genome of UK (non-native) *Dikerogammarus*
801 *haemobaphes* (Amphipoda: Gammaridae) informs upon *Dikerogammarus* evolution,
802 invasions and associated microparasites. *Hydrobiol* 2019 8471. 2019;847:229–42.
803 doi:10.1007/S10750-019-04084-1.

804 55. Yokobori S, Fukuda N, Nakamura M, Aoyama T, Oshima T. Long-Term
805 Conservation of Six Duplicated Structural Genes in Cephalopod Mitochondrial
806 Genomes. *Mol Biol Evol*. 2004;21:2034–46. doi:10.1093/MOLBEV/MSH227.

807 56. Kawashima Y, Nishihara H, Akasaki T, Nikaido M, Tsuchiya K, Segawa S, et al.
808 The complete mitochondrial genomes of deep-sea squid (*Bathyteuthis abyssicola*),
809 bob-tail squid (*Semirossia patagonica*) and four giant cuttlefish (*Sepia apama*, S.

810 latimanus, *S. lycidas* and *S. pharaonis*), and their application to the phylogenetic
811 analysis of Decapodiformes. *Mol Phylogenet Evol.* 2013;69:980–93.

812 57. Guerra D, Bouvet K, Breton S. Mitochondrial gene order evolution in Mollusca:
813 Inference of the ancestral state from the mtDNA of *Chaetopleura apiculata*
814 (*Polyplacophora*, *Chaetopleuridae*). *Mol Phylogenet Evol.* 2018;120:233–9.

815 58. Ghiselli F, Gomes-Dos-Santos A, Adema CM, Lopes-Lima M, Sharbrough J,
816 Boore JL. Molluscan mitochondrial genomes break the rules. *Philos Trans R Soc B*
817 *Biol Sci.* 2021;376.

818 59. Grande C, Templado J, Zardoya R. Evolution of gastropod mitochondrial
819 genome arrangements. *BMC Evol Biol.* 2008;8:1–15. doi:10.1186/1471-2148-8-61.

820 60. Shadel GS, Clayton DA. Mitochondrial DNA maintenance in vertebrates. *Annual*
821 *Review of Biochemistry.* 1997;66:409–36. doi:10.1146/annurev.biochem.66.1.409.

822 61. Voigt O, Erpenbeck D, Wörheide G. A fragmented metazoan organellar genome:
823 The two mitochondrial chromosomes of *Hydra magnipapillata*. *BMC Genomics.*
824 2008;9:1–10. doi:10.1186/1471-2164-9-350.

825 62. Zhang J, Zhou C, Gai Y, Song D, Zhou K. The complete mitochondrial genome
826 of *Parafronurus youi* (Insecta: Ephemeroptera) and phylogenetic position of the
827 Ephemeroptera. *Gene.* 2008;424:18–24.

828 63. Gjetvaj B, Cook DI, Zouros E. Repeated sequences and large-scale size
829 variation of mitochondrial DNA: A common feature among scallops (*Bivalvia*:
830 *Pectinidae*). *Mol Biol Evol.* 1992;9:106–24.

- 831 64. Rigaa A, Monnerot M, Sellos D. Molecular cloning and complete nucleotide
832 sequence of the repeated unit and flanking gene of the scallop *Pecten maximus*
833 mitochondrial DNA: Putative replication origin features. *J Mol Evol.* 1995;41:189–95.
- 834 65. Fuller KM. Polymorphism and heteroplasmy of mitochondrial DNA in the scallop
835 *Placopecten magellanicus* (Gmelin). 1991.
- 836 66. Lonsdale, D. M., Brears, T., Hodge, T. P., Melville, S. E., & Rottmann W, H. The
837 plant mitochondrial genome: homologous recombination as a mechanism for
838 generating heterogeneity. *Philos Trans R Soc London B, Biol Sci.* 1988;319:149–63.
839 doi:10.1098/rstb.1988.0039.
- 840 67. Palmer JD, Herbon LA. Plant mitochondrial DNA evolved rapidly in structure, but
841 slowly in sequence. *J Mol Evol.* 1988;28:87–97. doi:10.1007/BF02143500.
- 842 68. Van de Paer C, Bouchez O, Besnard G. Prospects on the evolutionary
843 mitogenomics of plants: A case study on the olive family (Oleaceae). *Mol Ecol*
844 *Resour.* 2018;18:407–23.
- 845 69. Pozzi A, Plazzi F, Milani L, Ghiselli F, Passamonti M. SmithRNAs: Could
846 Mitochondria “Bend” Nuclear Regulation? *Mol Biol Evol.* 2017;34:1960–73.
847 doi:10.1093/MOLBEV/MSX140.
- 848 70. Passamonti M, Calderone M, Delpero M, Plazzi F. Clues of in vivo nuclear gene
849 regulation by mitochondrial short non-coding RNAs. *Sci Reports* 2020 101.
850 2020;10:1–12. doi:10.1038/s41598-020-65084-z.
- 851 71. Smedley GD, Audino JA, Grula C, Porath-Krause A, Pairett AN, Alejandrino A, et
852 al. Molecular phylogeny of the Pectinoidea (Bivalvia) indicates Propeamussiidae to

853 be a non-monophyletic family with one clade sister to the scallops (Pectinidae). *Mol*
854 *Phylogenet Evol.* 2019;137:293–9.

855 72. Biscotti MA, Canapa A, Olmo E, Barucca M, Teo CH, Schwarzacher T, et al.
856 Repetitive DNA, molecular cytogenetics and genome organization in the King scallop
857 (*Pecten maximus*). *Gene.* 2007;406:91–8.

858 73. Ojala D, Montoya J, Attardi G. TRNA punctuation model of RNA processing in
859 human mitochondria. *Nature.* 1981;290:470–4. doi:10.1038/290470a0.

860 74. D’Souza AR, Minczuk M. Mitochondrial transcription and translation: Overview.
861 *Essays in Biochemistry.* 2018;62:309–20. doi:10.1042/EBC20170102.

862 75. Tyagi K, Chakraborty R, Cameron SL, Sweet AD, Chandra K, Kumar V.
863 Rearrangement and evolution of mitochondrial genomes in Thysanoptera (Insecta).
864 *Sci Rep.* 2020;10:1–16.

865 76. Yoshizawa K, Johnson KP, Sweet AD, Yao I, Ferreira RL, Cameron SL.
866 Mitochondrial phylogenomics and genome rearrangements in the barklice (Insecta:
867 Psocodea). *Mol Phylogenet Evol.* 2018;119 October 2017:118–27.
868 doi:10.1016/j.ympev.2017.10.014.

869 77. Xu W, Jameson D, Tang B, Higgs PG. The relationship between the rate of
870 molecular evolution and the rate of genome rearrangement in animal mitochondrial
871 genomes. *J Mol Evol.* 2006;63:375–92.

872 78. Delargy A. Quantitative Methods for Producing Evidence to Support Sustainable
873 King Scallop Management. Bangor University; 2020.

- 874 79. Stewart BD, Howarth LM. Quantifying and Managing the Ecosystem Effects of
875 Scallop Dredge Fisheries. *Dev Aquac Fish Sci.* 2016;40:585–609.
- 876 80. Rustin P. Mitochondria, from cell death to proliferation. *Nat Genet.* 2002;30:352–
877 3. doi:10.1038/NG0402-352.
- 878 81. Quirós PM, Mottis A, Auwerx J. Mitonuclear communication in homeostasis and
879 stress. *Nat Rev Mol Cell Biol.* 2016;17:213–26. doi:10.1038/nrm.2016.23.
- 880 82. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina
881 sequence data. *Bioinformatics.* 2014;30:2114–20. doi:10.1093/bioinformatics/btu170.
- 882 83. Dierckxsens N, Mardulyn P, Smits G. NOVOPlasty: De novo assembly of
883 organelle genomes from whole genome data. *Nucleic Acids Res.* 2017;45.
- 884 84. Andrews S. FastQC: a quality control tool for high throughput sequence data.
885 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 886 85. La Roche J, Snyder M, Cook DI, Fuller K, Zouros E. Molecular characterization
887 of a repeat element causing large-scale size variation in the mitochondrial DNA of
888 the sea scallop *Placopecten magellanicus*. *Mol Biol Evol.* 1990;7:45–64.
- 889 86. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat*
890 *Methods.* 2012;9.
- 891 87. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The
892 Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009.
- 893 88. Dierckxsens N, Mardulyn P, Smits G. Unraveling heteroplasmy patterns with
894 NOVOPlasty. *NAR Genomics Bioinforma.* 2020;2:1–10.

895 89. Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, et al.
896 Improved annotation of protein-coding genes boundaries in metazoan mitochondrial
897 genomes. *Nucleic Acids Res.* 2019;47:10543–52. doi:10.1093/nar/gkz833.

898 90. Al Arab M, Höner zu Siederdisen C, Tout K, Sahyoun AH, Stadler PF, Bernt M.
899 Accurate annotation of protein-coding genes in mitochondrial genomes. *Mol*
900 *Phylogenet Evol.* 2017;106:209–16.

901 91. Laslett D, Canbäck B. ARWEN: A program to detect tRNA genes in metazoan
902 mitochondrial nucleotide sequences. *Bioinformatics.* 2008;24:172–5.
903 doi:10.1093/bioinformatics/btm573.

904 92. Benson G. Tandem repeats finder: A program to analyze DNA sequences.
905 *Nucleic Acids Res.* 1999;27.

906 93. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction.
907 *Nucleic Acids Res.* 2003;31:3406–15. doi:10.1093/nar/gkg595.

908 94. Grant JR, Stothard P. The CGView Server: a comparative genomics tool for
909 circular genomes. *Nucleic Acids Res.* 2008;36 Web Server issue.

910 95. Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic
911 analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006.

912 96. Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with
913 MAFFT. *Methods Mol Biol.* 2009;537:39–64. doi:10.1007/978-1-59745-251-9_3.

914 97. Abascal F, Zardoya R, Telford MJ. TranslatorX: Multiple alignment of nucleotide
915 sequences guided by amino acid translations. *Nucleic Acids Res.* 2010;38 SUPPL.
916 2:W7–13. doi:10.1093/nar/gkq291.

917 98. Castresana J. Selection of conserved blocks from multiple alignments for their
918 use in phylogenetic analysis. *Mol Biol Evol.* 2000;17:540–52.
919 doi:10.1093/oxfordjournals.molbev.a026334.

920 99. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: A fast and
921 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol*
922 *Biol Evol.* 2015.

923 100. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS.
924 ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat*
925 *Methods.* 2017.

926 101. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New
927 algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the
928 performance of PhyML 3.0. *Syst Biol.* 2010;59:307–21. doi:10.1093/sysbio/syq010.

929 102. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al.
930 MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a
931 Large Model Space. *Syst Biol.* 2012;61:539–42. doi:10.1093/sysbio/sys029.

932 103. Rambaut A, Drummond AJ. Tracer v1.6. Available from
933 <http://tree.bio.ed.ac.uk/software/tracer/>. 2013.

934 104. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: Multiple alignment of
935 conserved genomic sequence with rearrangements. *Genome Res.* 2004;14:1394–
936 403. doi:10.1101/gr.2289704.

937 105. Bernt M, Merkle D, Ramsch K, Fritzsich G, Perseke M, Bernhard D, et al. CREx:
938 Inferring genomic rearrangements based on common intervals. *Bioinformatics*.
939 2007;23:2957–8. doi:10.1093/bioinformatics/btm468.

940 106. Bernt M, Merkle D, Middendorf M. An algorithm for inferring mitogenome
941 rearrangements in a phylogenetic tree. In: *Lecture Notes in Computer Science*
942 (including subseries *Lecture Notes in Artificial Intelligence* and *Lecture Notes in*
943 *Bioinformatics*). Springer, Berlin, Heidelberg; 2008. p. 143–57. doi:10.1007/978-3-
944 540-87989-3_11.

945 107. Hu F, Lin Y, Tang J. MLGO: Phylogeny reconstruction and ancestral inference
946 from gene-order data. *BMC Bioinformatics*. 2014;15:354. doi:10.1186/s12859-014-
947 0354-6.

948 **Figure legends**

949 **Fig. 1.** Map of the *Mimachlamys varia* mitogenome. Protein coding (CDS), rRNA and
950 tRNA genes and the Major Non-coding Region (MNR) are marked with differently
951 colored arrows pointing in the direction of transcription. Innermost scale represents
952 size in kbp. GC content and GC skew are represented by black and purple-orange
953 histograms, respectively. The outermost circle in light gray shows the natural
954 logarithm of the depth of coverage for each base. Photo credit: Thierry Guyot.

955 **Fig. 2.** Heatmap representing mitochondrial tRNA gene sets observed in the 13
956 pectinid species included in this study. tRNAs are represented with their one-letter
957 amino-acid code, with the anticodon in parentheses. The deeper blue color
958 corresponds to higher copy number of a tRNA gene.

959 **Fig. 3.** Maximum likelihood trees based on all eight datasets (A-H). Branch lengths
960 are scaled by setting the length of the outgroup branch to 1.5 on each tree. Numbers
961 on top of the outgroup branches represent the original length of the branches. Every
962 genus is marked with a different color, which is identical to the color coding on
963 Figures 4 and 5. *Mimachlamys varia* is marked with a black diamond shape, the gray
964 line connects the corresponding internal node with the SH-like approximate
965 likelihood ratio test (SH-aLRT) value on the left and ultrafast bootstrap support on
966 the right. Dataset descriptions: A) all PCG and rRNA genes treated with Gblocks; B)
967 all PCG amino acid sequences and rRNA genes treated with Gblocks; C) all PCGs;
968 D) all PCGs treated with Gblocks; E) all PCG amino acid sequences; F) all PCG
969 amino acid sequences treated with Gblocks; G) rRNA genes; H) rRNA genes treated
970 with Gblocks.

971 **Fig. 4.** Phylogenetic reconstruction of 30 Pectinidae mitogenomes with *Ostrea lurida*
972 (*Ostreidae*) as the outgroup, based on the “B” dataset (PCG amino acid and rRNA
973 DNA sequences). Every genus is marked with a different color, identical to the color
974 coding on Figs. 3 and 5. *Mimachlamys varia* is marked with a black diamond shape.
975 Fully supported branches are marked with a blue circle on the corresponding node.
976 A: Maximum likelihood tree. SH-like approximate likelihood ratio test (SH-aLRT) is
977 presented on the left, ultrafast bootstrap support projected from the extended
978 majority-rule consensus tree is presented on the right. B: Bayesian tree. Numbers
979 represent posterior probabilities; the staggered and gray lines connect species
980 names with the corresponding tips. For original branch lengths, see Fig. 3B and Fig.
981 5B.

982 **Fig. 5.** Bayesian consensus trees based on all eight datasets (A-H). Branch lengths
983 are scaled by setting the length of the outgroup branch to 1.5 on each tree. Numbers
984 on top of the outgroup branches represent the original length of the branches. Every
985 genus is marked with a different color, that is identical to the color coding of Figs. 3
986 and 4. *Mimachlamys varia* is marked with a black diamond shape. Posterior
987 probabilities of the corresponding internal node are presented. Dataset descriptions:
988 A) all PCG and rRNA genes masked with Gblocks; B) all PCG amino acid
989 sequences and rRNA genes masked with Gblocks; C) all PCGs; D) all PCGs
990 masked with Gblocks; E) all PCG amino acid sequences; F) all PCG amino acid
991 sequences masked with Gblocks; G) rRNA genes; H) rRNA genes masked with
992 Gblocks.

993 **Fig. 6.** Gene collinearity analysis of 13 Pectinid species performed in Mauve v2.3.1.
994 Color blocks represent homologous regions among different mitogenomes. The level
995 of similarity at each position is shown in the blocks. Below these blocks, the white,
996 red, green and pink boxes represent protein coding, rRNA, tRNA genes, and tandem
997 repeats, respectively.

998 **Fig. 7.** Boxplot showing the length distribution of each LCB (locally colinear block)
999 from the gene collinearity analysis.

1000 **Fig. 8.** Common interval analysis presented on the most supported Bayesian tree,
1001 obtained from dataset "B", where each tip represents a unique gene order. Internal
1002 nodes are named as "An", where "n" is a number between 0 and 7, and represents
1003 ancestral gene orders. C and P indicate consistency and parsimony values,
1004 respectively. The colors code consistent (green; highest level of certainty), k-

1005 consistent (yellow; intermediate level of certainty) and inconsistent (red; low level of
1006 certainty) nodes. Mutation types are indicated on corresponding branches, T:
1007 transposition, TDRL: tandem duplication and random loss.

1008 **Fig. 9.** Maximum likelihood tree based on gene orders, 42 genes in total, including
1009 all 13 species included in this study plus the outgroup, and the complete set of genes
1010 in each mitogenome. Numbers represent bootstrap support. *Mimachlamys varia* is
1011 marked with a black diamond shape.

1012 **Tables**

1013 **Table 1.** Main structural features of the *Mimachlamys varia* mitogenome.

| Gene | Location | Size (bp) | Start codon | Stop codon | Anticodon s | Intergenic nucleotide | A+T No. content |
|-------|-------------|--------------|----------------|---------------|----------------|--------------------------|--------------------|
| | | | | | | | AA (%) |
| trnN | 1–69 | 69 | | | GUU | – | 54 |
| nad1 | 94–993 | 900 | TTG | TAG | | 24 299 | 56.1 |
| trnR | 992–1055 | 64 | | | UCG | -2 | 56.2 |
| rrnL | 1111–2532 | 1422 | | | | 55 | 56.6 |
| cox1 | 2604–4127 | 1524 | ATG | TAG | | 71 507 | 55 |
| nad4l | 4384–4671 | 288 | TTG | TAG | | 256 95 | 50.3 |
| trnG | 4677–4739 | 63 | | | UCC | 5 | 60.3 |
| nad6 | 4774–5271 | 498 | ATT | TAA | | 34 165 | 54.6 |
| trnF | 5414–5483 | 70 | | | AAA | 142 | 58.6 |
| atp8 | 9690–9848 | 159 | ATG | TAG | | 4206 52 | 55.3 |
| trnL2 | 9859–9924 | 66 | | | UAA | 10 | 50 |
| cob | 10133–11257 | 1125 | ATG | TAG | | 208 374 | 56 |
| trnD | 11256–11321 | 66 | | | GUC | -2 | 47 |
| cox3 | 11401–12171 | 771 | TTG | TAA | | 79 256 | 55.5 |
| trnK | 12215–12283 | 69 | | | UUU | 43 | 46.4 |
| trnF | 12288–12352 | 65 | | | GAA | 4 | 49.2 |
| trnQ2 | 12359–12425 | 67 | | | UUG | 6 | 52.2 |

| | | | | | | |
|-------|-------------|----------|-----|-----|--------|------|
| trnQ1 | 12429–12497 | 69 | | UUG | 3 | 44.1 |
| trnE | 12503–12568 | 66 | | UUC | 5 | 54.5 |
| atp6 | 12642–13313 | 672 ATG | TAA | | 73 223 | 56.8 |
| cox2 | 13377–14078 | 702 ATG | TAG | | 63 233 | 54.7 |
| nad2 | 14084–15040 | 957 ATA | TAA | | 41 306 | 56 |
| trnV | 15051–15116 | 66 | | UAC | 10 | 50 |
| nad3 | 15125–15472 | 348 ATT | TAG | | 8 115 | 54.9 |
| nad4 | 15552–16835 | 1284 ATG | TAG | | 79 427 | 54.8 |
| trnH | 16837–16902 | 66 | | GUG | 1 | 56.9 |
| trnW | 16908–16977 | 70 | | CCA | 5 | 54.3 |
| trnY | 16981–17045 | 65 | | GUA | 3 | 49.2 |
| trnT | 17050–17112 | 63 | | UGU | 4 | 57.1 |
| trnP | 17116–17183 | 68 | | UGG | 3 | 48.5 |
| trnI | 17186–17256 | 71 | | GAU | 2 | 57.7 |
| trnL1 | 17265–17333 | 69 | | UAG | 8 | 53.6 |
| trnM2 | 17336–17402 | 67 | | CAU | 2 | 47.8 |
| trnM1 | 17411–17483 | 73 | | CAU | 8 | 41.1 |
| trnC | 17492–17559 | 68 | | GCA | 8 | 42.6 |
| trnA | 17571–17638 | 68 | | UGC | 11 | 59.1 |
| rrnS | 17675–18638 | 964 | | | 36 | 52.7 |
| nad5 | 18690–20393 | 1704 ATT | TAA | | 51 567 | 55.9 |

1014 **Table 2.** Tandem repeats in the *Mimachlamys varia* mitogenome.

| Position in | Perio | Copy | Consensus | Percent | Percent | Alignm | A | C | G | T | Entropy |
|-------------|-------|------|-----------|---------|---------|--------|---|---|---|---|---------|
|-------------|-------|------|-----------|---------|---------|--------|---|---|---|---|---------|

| genome | d size | number | size | matches | indels | ent | | | | | (0-2) | |
|-----------|--------|--------|------|---------|--------|-------|----|----|----|----|-------|--|
| | | | | | | score | | | | | | |
| 5601–5625 | 13 | 1.9 | 13 | 100 | 0 | 50 | 84 | 0 | 16 | 0 | 0.63 | |
| 5894–5935 | 10 | 4.3 | 10 | 87 | 6 | 68 | 21 | 28 | 40 | 9 | 1.84 | |
| 5896–5933 | 19 | 2 | 20 | 94 | 5 | 69 | 21 | 26 | 42 | 10 | 1.85 | |
| 6440–9644 | 683 | 4.7 | 683 | 99 | 0 | 6392 | 28 | 13 | 22 | 35 | 1.92 | |

1015 **Table 3.** Newly annotated *atp8* genes.

| Species | GenBank accession no. | Start | Stop | Length | AA | 5' | 3' |
|---------------------------------|--------------------------|-------|-------|--------|----|----------|----------|
| | | | | | | neighbor | neighbor |
| <i>Placopecten magellanicus</i> | DQ088274– <i>atp8-1</i> | 23484 | 23660 | 177 | 58 | trnM | atp6 |
| | DQ088274– <i>atp8-2</i> | 64 | 231 | 168 | 55 | trnG | trnM |
| | NC_007234– <i>atp8-1</i> | 23484 | 23660 | 177 | 58 | trnM | atp6 |
| | NC_007234– <i>atp8-2</i> | 64 | 231 | 168 | 55 | trnG | trnM |
| <i>Chlamys farreri</i> | EF473269 | 10041 | 10187 | 147 | 48 | nad5 | trnD |
| | EU715252 | 10041 | 10187 | 147 | 48 | nad5 | trnD |
| | NC_012138 | 10041 | 10187 | 147 | 48 | nad5 | trnD |
| <i>Mizuhopecten yessoensis</i> | AB271769– <i>atp8-1</i> | 9797 | 9943 | 147 | 48 | nad4l | trnD |
| | AB271769– <i>atp8-2</i> | 9239 | 9385 | 147 | 48 | nad5 | nad4l |
| | FJ595959– <i>atp8-1</i> | 9797 | 9943 | 147 | 48 | nad4l | trnD |
| | FJ595959– <i>atp8-2</i> | 9239 | 9385 | 147 | 48 | nad5 | nad4l |

| | | | | | | | |
|--|--------------------------|-------|-------|-----|----|-------|-------|
| | NC_009081- <i>atp8-1</i> | 9797 | 9943 | 147 | 48 | nad4l | trnD |
| | NC_009081- <i>atp8-2</i> | 9239 | 9385 | 147 | 48 | nad5 | nad4l |
| <i>Crassadoma</i> <i>gigantea</i> | MH016739 | 11145 | 11303 | 159 | 52 | nad5 | trnD |
| <i>Mimachlamys</i> <i>nobilis</i> | FJ415225 | 7938 | 8090 | 153 | 50 | nad4l | trnM |
| | FJ595958 | 7937 | 8089 | 153 | 50 | nad4l | trnM |
| | NC_011608 | 7938 | 8090 | 153 | 50 | nad4l | trnM |
| <i>Mimachlamys</i> <i>varia</i> | MZ520326 | 9690 | 9848 | 159 | 52 | trnF | trnL2 |
| <i>Pecten</i> <i>albicans</i> | KP900974 | 6000 | 6134 | 135 | 44 | trnE | nad6 |
| <i>Pecten</i> <i>maximus</i> | KP900975 | 5993 | 6127 | 135 | 44 | trnE | nad6 |
| <i>Argopecten</i> <i>irradians</i> | DQ665851 | 4272 | 4406 | 135 | 44 | trnE | nad6 |
| | EU023915 | 4185 | 4406 | 222 | 73 | trnF | nad6 |
| | KT161259 | 4272 | 4406 | 135 | 44 | trnE | nad6 |
| | KT161262 | 4272 | 4406 | 135 | 44 | trnE | nad6 |
| | KU589290 | 4272 | 4406 | 135 | 44 | trnE | nad6 |
| | NC_009687 | 4185 | 4406 | 222 | 73 | trnF | nad6 |
| | NC_012977 | 4272 | 4406 | 135 | 44 | trnE | nad6 |
| <i>Argopecten</i> <i>purpuratus</i> | KF601246 | 4276 | 4410 | 135 | 44 | trnE | nad6 |

| | | | | | | | |
|-------------------|-----------|------|------|-----|----|------|------|
| | KT161260 | 4276 | 4410 | 135 | 44 | trnE | nad6 |
| | NC_027943 | 4276 | 4410 | 135 | 44 | trnE | nad6 |
| <i>Argopecten</i> | KT161261 | 4238 | 4372 | 135 | 44 | trnE | nad6 |

ventricosus

1016

1017 **Table 4.** List of species included in the present study.

1018

| Subfamily | Tribe | Species | GenBank accession no. | Genome length (bp) | Source |
|-------------|------------|--|-----------------------------|--------------------------|--------------------------|
| Palliolinae | Palliolini | <i>Placopecten magellanicus</i> (Gmelin, 1791) | DQ088274 | 32115 | La Roche et al. [77]□ |
| | | | NC_007234 | 32115 | La Roche et al. [77]□ |
| Chlamydinae | Chlamydini | <i>Chlamys farreri</i> (Jones & Preston, 1904) | EF473269 | 20889 | Ren et al. [11]□ |
| | | | EU715252 | 21695 | Xu et al. [18]□ |
| | | | NC_012138 | 21695 | Xu et al. [18]□ |

| | | | | |
|----------------|--|-----------|-------|--------------------------------|
| | <i>Mizuhopecten yessoensis</i> (Jay, 1857) | AB271769 | 20414 | Sato et al. [14]□ |
| | | FJ595959 | 20964 | Wu et al. [12]□ |
| | | NC_009081 | 20414 | Sato et al. [14]□ |
| Crassadomini | <i>Crassadoma gigantea</i> (J.E. Gray, 1825) | MH016739 | 18495 | Liao et al. [13]□ |
| Mimachlamydini | <i>Mimachlamys nobilis</i> (Reeve, 1852) | FJ415225 | 17963 | Xu et al., unpublished |
| | | FJ595958 | 17935 | Wu et al. [12]□ |
| | | NC_011608 | 17963 | Xu et al., unpublished |
| | <i>Mimachlamys senatoria</i> (Gmelin, 1791) | KF214684 | 17383 | Wu et al. [16]□ |
| | | NC_022416 | 17383 | Wu et al. [16]□ |
| | <i>Mimachlamys varia</i> (Linnaeus, | MZ520326 | 20400 | Malkócs et al., unpublished |

| | | | | | |
|-----------------------|--------------|--|----------|-------|--------------------------------------|
| | | 1758) | | | |
| Pectininae | Amusiini | <i>Amusium</i> <i>pleuronectes</i> (Linnaeus, 1758) | MT419374 | 18044 | Yao et al. [23]□ |
| | Pectinini | <i>Pecten</i> <i>albicans</i> (Schröter, 1802) | KP900974 | 16653 | Marín et al. [17]□ |
| | | <i>Pecten</i> <i>maximus</i> (Linnaeus, 1758) | KP900975 | 17252 | Marín et al. [17]□ |
| Aequipectini group | Aequipectini | <i>Argopecten</i> <i>irradians</i> (Lamarck, 1819) | DQ665851 | 16211 | Ren et al. [11]□ |
| | | | EU023915 | 16221 | Petten and Snyder, unpublished |
| | | | KT161259 | 16286 | Li, unpublished |
| | | | KT161262 | 16211 | Li, unpublished |

| | | | | | |
|----------|-----------|---|-----------|-------|--------------------------------------|
| | | | KU589290 | 16212 | Liu, unpublished |
| | | | NC_009687 | 16221 | Petten and Snyder, unpublished |
| | | | NC_012977 | 16211 | Ren et al. [11]□ |
| | | <i>Argopecten purpuratus</i> (Lamarck, 1819) | KF601246 | 16266 | Marín et al. [17]□ |
| | | | KT161260 | 16270 | Li, unpublished |
| | | | KY321561 | 16224 | Romero et al., unpublished |
| | | | NC_027943 | 16270 | Li, unpublished |
| | | <i>Argopecten ventricosus</i> (G. B. Sowerby II, 1842) | KT161261 | 16079 | Li, unpublished |
| Outgroup | | | | | |
| Family: | Ostreidae | <i>Ostrea lurida</i> Carpenter, 1864 | NC_022688 | 16344 | Xiao et al. [100]□ |

1019 Taxonomic authorities according to the World Register of Marine Species (WoRMS
1020 Editorial Board, 2021). Annotation information can be found in Table 1 for *Mim. varia*,
1021 and Additional file 3 for all other species.

1022 **Additional files**

1023 **File name:** Additional File 1.

1024 **File format:** PDF Document (.pdf)

1025 **Title of data:** Putative secondary structures of mitochondrial tRNAs in *Mimachlamys*
1026 *varia*.

1027 **Description of data:** Amino acids are represented by their one-letter code. (* :
1028 structure predicted by RNAfold, all other structures were predicted by MITOS2)

1029 **File name:** Additional File 2.

1030 **File format:** PDF Document (.pdf)

1031 **Title of data:** Putative secondary structure of the Major Non-coding Region of
1032 *Mimachlamys varia*.

1033 **Description of data:** There is a 100 basepair overlap kept between parts. 1: 5484–
1034 6439; 2: 6339–7805; 3: 7705–9689.

1035 **File name:** Additional file 3.

1036 **File format:** Microsoft Word Document (.docx)

1037 **Title of data:** Annotation tables for species involved in this study.

1038 **File name:** Additional file 4.

1039 **File format:** PDF Document (.pdf)

1040 **Title of data:** Common interval analysis results presented on a Bayesian tree, where
1041 each tip represents a unique gene order.

1042 **Description of data:** Internal nodes are named as A_n , where n is a number between
1043 0 and 7, and represent ancestral gene orders. C and P indicate consistency and
1044 parsimony values, respectively. The colors code consistent (green; highest level of
1045 certainty), k-consistent (yellow; intermediate level of certainty) and inconsistent (red;
1046 low level of certainty) nodes. Mutation types are indicated on corresponding
1047 branches, T: transposition, TDRL: tandem duplication and random loss. Arrows point
1048 to diagrams showing the two gene orders, and the mutation steps leading from the
1049 more ancestral to the younger node. Frames connect gene blocks affected by a
1050 rearrangement event, and parts of the mitogenome that is affected by a
1051 rearrangement event.

1052 **File name:** Additional file 5.

1053 **File format:** Microsoft Word Document (.docx)

1054 **Title of data:** Tandem repeats detected in pectinid mitogenomes.