PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Review

Cite this article: Rosani U, Bortoletto E, Bai C-M, Novoa B, Figueras A, Venier P, Fromm B. 2021 Digging into bivalve miRNAomes: between conservation and innovation. *Phil. Trans. R. Soc. B* **376**: 20200165. https://doi.org/10.1098/rstb.2020.0165

Accepted: 18 January 2021

One contribution of 15 to a Theo Murphy meeting issue 'Molluscan genomics: broad insights and future directions for a neglected phylum'.

Subject Areas:
genomics, evolution

Keywords:
microRNA, Bivalvia, mussel, blood clam, bivalves

Authors for correspondence:
Umberto Rosani
e-mail: umberto.rosani@unipd.it
Bastian Fromm
e-mail: bastianfromm@gmail.com

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5324913.

THE ROYAL SOCIETY PUBLISHING

Digging into bivalve miRNAomes: between conservation and innovation

Umberto Rosani[1], Enrico Bortoletto[1], Chang-Ming Bai[2], Beatriz Novoa[3], Antonio Figueras[3], Paola Venier[1] and Bastian Fromm[4]

[1]Department of Biology, University of Padova, 35121 Padova, Italy
[2]Key Laboratory of Maricultural Organism Disease Control, Ministry of Agriculture; Qingdao Key Laboratory of Mariculture Epidemiology and Biosecurity; Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266237, People's Republic of China
[3]Institute of Marine Research (IIM), Spanish National Research Council (CSIC), Eduardo Cabello, 6, 36208 Vigo, Spain
[4]Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, 10691 Stockholm, Sweden

UR, 0000-0003-0685-1618

Bivalves are a diverse mollusc group of economic and ecological importance. An evident resilience to pollution, parasites and extreme environments makes some bivalve species important models for studying adaptation and immunity. Despite substantial progress in sequencing projects of bivalves, information on non-coding genes and gene-regulatory aspects is still lacking. Here, we review the current repertoire of bivalve microRNAs (miRNAs), important regulators of gene expression in Metazoa. We exploited available short non-coding RNA (sncRNA) data for *Pinctada martensii*, *Crassostrea gigas*, *Corbicula fluminea*, *Tegillarca granosa* and *Ruditapes philippinarum*, and we produced new sncRNA data for two additional bivalves, the Mediterranean mussel *Mytilus galloprovincialis* and the blood clam *Scapharca broughtonii*. We found substantial heterogeneity and incorrect annotations of miRNAs; hence, we reannotated conserved miRNA families using recently established criteria for *bona fide* microRNA annotation. We found 106 miRNA families missing in the previously published bivalve datasets and 89 and 87 miRNA complements were identified in the two additional species. The overall results provide a homogeneous and evolutionarily consistent picture of miRNAs in bivalves and enable future comparative studies. The identification of two bivalve-specific miRNA families sheds further light on the complexity of transcription and its regulation in bivalve molluscs.

This article is part of the Theo Murphy meeting issue 'Molluscan genomics: broad insights and future directions for a neglected phylum'.

1. Introduction

The Bivalvia include approximately 10 000 described species, mostly adapted to the marine environment and relevant for human nutrition as low-carbon-footprint meat [1–3]. Owing to their sessile and filter-feeding behaviour, some bivalves represent interesting models for studying biological adaptations to extreme environments, innate immune responses and host–pathogen interactions following a genomic approach [4]. The current advancements in sequencing technologies have resulted in a significant increase of genomic and transcriptomic datasets and, although limited to relatively few species, they are revealing fascinating aspects of the bivalve resilience to pathogens, such as the expansion of immune-related genes and an amazing gene presence–absence variation [5,6]. In this enlarging bubble of bivalve sequence data, small non-coding RNAs (sncRNAs) have gained attention as they are expected to act as gene regulators of many essential processes, including biomineralization [7,8]. Non-coding (nc) RNAs are RNA molecules not translated

into protein products but contributing to the complexity of gene expression processes [9]. Despite initially being regarded as a part of the 'junk' matter [10], thousands of functional ncRNAs participate in physiological and pathological processes [11,12]. Currently, deep sequencing is contributing to enlarge our view of different types of ncRNAs, and it progressively reveals their importance in gene silencing and in the control of viruses and transposons [13]. Among ncRNAs, microRNAs (miRNAs) are of paramount importance in various cellular and developmental processes, such as immunity, cell behaviour and host–microorganism interactions [14–16]. miRNAs have also been proposed as diagnostic markers of a condition in various diseases [17]. Mature miRNAs are short sequences (approx. 22 nt), which originate from hairpin precursor RNAs (pre-miRNAs) and, different from other ncRNAs such as the silencing RNAs, they commonly act as post-transcriptional repressors by targeting 3′ untranslated regions (UTRs) of messenger RNAs (mRNAs). After the first discovery of *lin-4* in *Caenorhabditis elegans* [18], miRNAs have been traced in animals, plants and viruses, and at least four independent origins have been suggested and debated [19,20]. Evolutionary studies showed how miRNA families have been continuously added along with the bilaterian evolution, whereas loss events seem to be uncommon and mature miRNA sequences are retained under strict negative selection [21,22]. Therefore, given the evolutionary conservation of the miRNA biogenesis mechanism and the small probability of homoplasy, miRNAs have also been proposed as powerful phylogenetics markers in Metazoa [23]. Nevertheless, the available knowledge on bivalve miRNAs is fragmentary and miRNA annotations are often completely missing in bivalve genomes and sequence databases, such as miRBase. Indeed, only three molluscs (*Haliotis rufescens*, *Lottia gigantea* and *Melibe leonina*) are represented in the last miRBase release, which includes 48 860 mature miRNAs from 271 different organisms [24]. In addition, peculiar evolutionary patterns of mollusc miRNAs, like the presence of typical vertebrate miRNA families in the gastropod *Patella vulgata*, have been reported [25]. The incomplete and error-prone annotations of many miRNAs present in miRBase have stalled progress in comparative analyses, leading to severe misinterpretations of putative loss patterns and to an undue devaluation of miRNAs in phylogenetic studies [21,26]. New databases, covering underrepresented species and implementing the miRNA annotation with up-to-date criteria, were in fact necessary and have been established [27]. The example of a curated miRNA gene database is MirGeneDB, initially developed from miRBase to provide reliable datasets of miRNAs and then substantially extended to species not present in miRBase [28], such as the Pacific oyster *Crassostrea gigas*, the brachiopod *Lingula anatina* and another five non-mollusc species. It is this kind of database that opens comparative miRNA analyses to non-model organisms of interest.

In this review, to build an updated foundation for future comparative studies on miRNAs, we have considered all bivalve miRNA studies published so far and have searched for conserved miRNA families using *Pinctada fucata*, *Tegillarca granosa*, *Corbicula fluminea and Ruditapes philippinarum* datasets. To expand the available miRNAome data, we investigated the miRNA repertoire of the mussel *Mytilus galloprovincialis* and the blood clam *Scapharca (Anadara) broughtonii* by sequencing small non-coding RNA (sncRNA-seq) of haemocytes. Based on up-to-date annotation criteria, we identified 80 mussel and 87 blood clam conserved miRNAs, respectively, and we showed a strict conservation of all the expected miRNA families across bivalves. Comparing our findings with previous reports, we have substantially improved miRNA annotations for bivalves, including the identification of two miRNA families conserved in and specific to bivalves.

2. Methods

(a) Retrieval and production of sncRNA-seq data

Short non-coding RNA datasets of bivalve species were retrieved from the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra, table 1). New sncRNA-seq data from *M. galloprovincialis* haemocytes were produced as described in Moreira *et al.* [51] and deposited in the NCBI SRA archive (SAMN09104631 to SAMN09104634). New sncRNA-seq data from *S. broughtonii* haemocytes were produced as described in electronic supplementary material, file S1 and deposited in the SRA archive (Bioproject ID PRJNA668611). A total of 29 bivalve genomes were retrieved from public repositories (electronic supplementary material, file S2), whereas the *M. galloprovincialis* genome (version Mg3) was obtained from Prof. A. Figueras [52]. The miRBase v. 22.1 entries were retrieved from http://www.mirbase.org/ftp.shtml [24], MirGeneDB 2.0 data were downloaded from http://mirgenedb.org/ [53], and previously reported bivalve miRNAs were retrieved from the electronic supplementary material files of a number of papers [29,31,36,39].

(b) sncRNA-seq data analysis

All sncRNA data were converted into FASTQ format and *cutadapt* v. 1.18 [54] was used to trim the raw reads, setting the minimal quality threshold to PHRED25, removing adaptor sequences and applying a size range of 18–40 nt. Bivalve miRNAs obtained from published studies were BLASTed (*blastn*) against miRBase and MirGeneDB databases. Limited to mussel and clam sncRNA-seq data, *miRTrace* v.1.0.0 [55] was used to group similar read sequences into clusters, to verify the quality of each dataset, miRNA size distribution and the presence of possible contaminants, namely miRNAs of different lineages. *MirMiner* [22] was applied to identify *bona fide* miRNAs and to provide a phylogenetic classification of known miRNAs following up-to-date annotation criteria. In detail: (i) the presence of coverage for both arms of the miRNA sequences, (ii) the distance between the mature and star sequences being lower than 40 nt, (iii) the absence of reads mapped in the surroundings of the annotated miRNAs, (iv) 5′ homogeneity of the mature miRNA, (v) 2 nt overhang and (vi) a reduced free energy. The genomic position of each *bona fide* mussel and clam miRNA was localized using *blastn*.

3. Result and discussion

(a) Only few Bivalvia species are represented in sncRNA-seq data currently available

Deep sequencing data of bivalve sncRNAs (sncRNA-seq) currently have results from *Pinctada martensii*, *Crassostrea gigas*, *C. hongkongensis*, *Tegillarca granosa* and *Ruditapes philippinarum*. Other studies reporting sncRNA-seq data for three additional species, *Chlamys farreri*, *Hyriopsis cumingii* and *Mytilus galloprovincialis*, have been published with no raw data available (table 1). Altogether, these studies cover a

Table 1. Summary of bivalve miRNA studies. Species, study target, number of produced samples, considered tissue, aim of the research and reference are reported. The underlined species refer to experiments with no available datasets.

species	target(s)	no. samples	sequenced tissue	aim of the research	reference
Pinctada martensii	sncRNA-seq - Solexa	1	multiple	biomineralization	Jiao *et al.*, [29]
Chlamys farreri	sncRNA-seq - Solexa	2	haemocytes	host–pathogen interactions	Chen *et al.*, [30]
Crassostrea gigas	sncRNA-seq – Illumina	3	haemocytes	immunity	Zhou *et al.*, [31]
Crassostrea gigas	sncRNA-seq – Illumina	21	multiple + developmental stages	miRNA discovery	Xu *et al.*, [32]
Tegillarca granosa	sncRNA-seq – Illumina	2	haemocytes	metal exposure	Bao *et al.*, [33]
Ostrea edulis	microarray	\	haemocytes	immunity	Martin-Gomez *et al.*, [34]
Crassostrea gigas	sncRNA-seq – Illumina	6		neuromodulation	Chen *et al.*, [35]
Pinctada martensii	bioinformatic	\	multiple	biomineralization	Zheng *et al.*, [36]
Pinctada martensii	*Mir-2305*	\	multiple	biomineralization	Jiao *et al.*, [37]
Pinctada martensii	*Mir-29a*	\	multiple	immunity	Tian *et al.*, [38]
C. gigas + *C. hongkongensis*	sncRNA-seq – Illumina	4	gills	osmotic stress	Zhao *et al.*, [39]
Crassostrea gigas	*Mir-92d*	\		immunity	Chen *et al.*, [40]
Crassostrea gigas	*Mir-2d*	\		immunity	Chen *et al.*, [41]
Crassostrea gigas	*Mir-2d*	\		neuromodulation	Chen *et al.*, [42]
Crassostrea gigas	scaffold42648_5080	\	haemocytes	immunity	Chen *et al.*, [43]
Crassostrea gigas	*Mir-365*	\	haemocytes	desiccation	Chen *et al.*, [44]
Ruditapes philippinarum	sncRNA-seq – Illumina	6		cell regulation	Pozzi *et al.*, [45]
Crassostrea gigas	sncRNA-seq – Illumina	1	gills	host–pathogen interactions	Rosani *et al.*, [46]
Crassostrea hongkongensis	sncRNA-seq – Illumina	2	gonads	gonad development	Wei *et al.*, [47]
Hyriopsis cumingii	sncRNA-seq – Illumina	2	mantle	biomineralization	Chen *et al.*, [48]
Hyriopsis cumingii	*Mir-4504*	\	mantle	shell colour	Chen *et al.*, [48]
Mytilus galloprovincialis	sncRNA-seq – Illumina	3	whole body	miRNA discovery	Yu *et al.*, [49]
Crassostrea gigas	sncRNA-seq – Illumina	6	gills	host–pathogen interactions	Rosani *et al.*, [50]
Mytilus galloprovincialis	sncRNA-seq – Illumina	5	haemocytes	immunity	Moreira *et al.*, [51]
Scapharca broughtonii	sncRNA-seq – Illumina	27	haemocytes	immunity	in preparation

small number of bivalve families, mainly Pteriomorphia, with an evident bias towards a few species relevant in aquaculture. A total of 51 sncRNA datasets, namely 963 million sncRNA reads, were obtained from 12 different experiments using Illumina or Solexa technologies. In particular, the Pacific oyster *C. gigas* was the most represented species with 655 million reads, followed by the clam *R. philippinarum* with 224 million reads (electronic supplementary material, figure S1). The size profiles of total and genome-mapped sncRNA reads suggested miRNA enrichment in most of the analysed datasets (electronic supplementary material, figure S2). Additional peaks were evident at 26 nt in the sncRNA profiles from gonadal tissues and from early developmental stages of *C. gigas* (electronic supplementary material, figure S2, panel C) and at 30 nt in the sncRNA profiles from *R. philippinarum* gonads (electronic supplementary material, figure S2, panel B). As regards, the unusual size distribution of sncRNAs from *C. gigas* developmental stages, it can arise from piwi-interacting RNAs (piRNAs). Although not limited to germ line [56], a high transcription of piRNAs in both the germ line and in early embryo stages was also reported for molluscs, consistent with a possible role in the control of transposable elements and maintenance of genomic stability [57]. The atypical distribution observed for *R. philippinarum* sncRNAs,

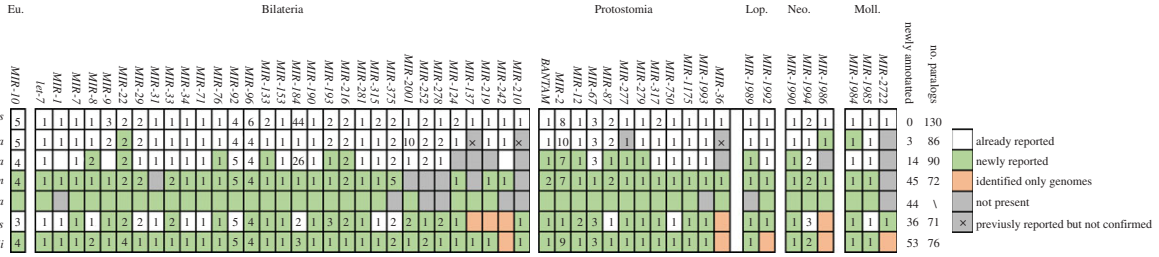

Figure 1. miRNA families conserved in bivalves. For the analysed species, the presence of Eumetazoa (Eu.), Bilateria, Protosomia, Lophotrochozoa (Lop.), Neotrochozoa (Neo.) and Mollusca (Moll.) conserved miRNA families is reported with different colours, if already mentioned (white) or newly reported (green, orange or blue). For the species with a sequenced genome, the number of paralogs for each miRNA family is also indicated.

with the main peak shifted to 30 nt, derived from highly transcribed small mitochondrial RNAs, which could modulate the expression of coding genes through RNA interference (RNAi) [45].

Given the evident enrichment in miRNAs of most of the analysed bivalve samples, we proceeded with the grouping of similar sequences into clusters, their classification using available databases (e.g. miRBase/MirGeneDB) and the identification of novel putative miRNAs using dedicated tools (e.g. miRDeep2 and miRPara). Overall, the available studies have reported 30–258 miRNAs per species and suggested that a considerable fraction of bivalve miRNAs is different from available database entries and, as such, it can only be classified as 'unknown'. Considering all the resulting bivalve miRNAs, we showed that they encompass 840 conserved and 437 novel miRNAs and, after removing redundancy because of similar sequences of different species, we obtained 717 different mature miRNA sequences. Only a small fraction of these miRNAs (15.8%) matched to at least one miRBase entry, mostly referring to gastropod molluscs (*Lottia gigantea*, *Melibe leonina* and *Haliotis rufescens*) and to miRNAs of the sponge *Amphimedon queenslandica*. Conversely, the percentage of annotated bivalve miRNAs increased to 31.8% when considering MirGeneDB, mostly because of the inclusion of *C. gigas* miRNA complements in this database. Altogether, the data regarding bivalve miRNAs remain fragmentary and not all the expected miRNA families have been identified in the considered species (figure 1). While at least 1 miRNA complement per family was identified in *C. gigas* (MirGeneDB) and 50 out of 53 in *P. fucata*, only 39 miRNA families have been identified in *T. granosa*. Although data are available, miRNAs have not previously been classified for *R. philippinarum* and *C. fluminea*.

(b) The miRNAomes of the Mediterranean mussel and the blood clam in comparison with other bivalves

To increase the current knowledge on bivalve miRNAomes, we performed sncRNA sequencing of haemocyte samples from two species belonging to distantly related families and having poor miRNA information, namely the Mediterranean mussel *Mytilus galloprovincialis* and the blood clam *Scapharca (Anadara) broughtonii*. The genomes of both species have been recently sequenced [52,58] and can be used for the identification of *bona fide* miRNAs. We analysed in total 138.4 million sncRNA reads for *M. galloprovincialis* and 452.7 million sncRNA reads for *S. broughtonii* (electronic supplementary material, figure S1). *miRTrace* [55] showed high-quality values for most of the reads, typical size profiles indicating enrichment of miRNA reads and low levels of possible contamination (electronic supplementary material, figure S3). The evidence of some contaminating miRNAs, especially in the clam reads, might arise from the filter-feeder behaviour of bivalves or from their open circulatory system, possibly contributing to the inclusion of water column microorganisms. Although *miRTrace* identified non-lophotrochozoan hits, most of the reads (99.99 and 96.2% for mussel and clam, respectively) referred to lophotrochozoan miRNAs, indicating a minimal impact of possible biological contaminants on the analysis. *MirMiner* was then applied to identify *bona fide* miRNAs for conserved families. To complete the fragmented patchwork of miRNAs obtained from the literature, we also reanalysed with *MirMiner* the datasets of *P. fucata*, *T. granosa*, *C. fluminea and R. philippinarum* using the reference genome when available or only the sncRNA reads for species without a high-quality sequenced genome (*C. fluminea*). Following a taxonomic classification of miRNAs, 80 known and 76 unknown hits resulted from *M. galloprovincialis*, whereas 87 known miRNAs and 135 unknown hits resulted from *S. broughtonii*. In both datasets, we could identify miRNA belonging to a total of 48 and 47 conserved families for mussel and clam, respectively. While the Eumetazoa-conserved *MIR-10* is present in both species, 28 out of 31 Bilateria-conserved miRNA families were retrieved from mussel, compared to 30 from clam (figure 1). The protostome-conserved *MIR-36* was absent in the mussel sncRNA data and the lophotrochozoan-conserved *MIR-1992* was absent in the clam data. *MIR-1986*, the representative of a neotrochozoa-conserved miRNA family, was absent from both datasets and *MIR-2722*, a Mollusca-conserved miRNA, was absent in the clam reads. The lack of some miRNA families in the sncRNA data suggested that either they were not expressed in our samples or, if absent also from the genome, that they might have been lost. However, we showed that at least one member of all the miRNA families absent from our sncRNA-seq data is still present in the mussel and clam genomes, which is expected given that hemocytes are quite a specific tissue and miRNAs often show cell- and tissue-specific expression.

Extending the *MirMiner* analysis to other bivalve species, we could identify previously unreported miRNAs, namely 44 for *C. fluminea*, 45 for *R. philippinarum*, 14 for *T. granosa* and 3 for *P. fucata*, whose families overall showed a good conservation (electronic supplementary material, file S3). Notably, *MIR-210* and *MIR-2722* were absent from the sncRNA reads of all these four species and from the blood clam (*MIR-2722*), whereas *MIR-1986* was absent from mussel and from the two species belonging to the Arcidae family. Although repetitive absence might suggest a gene lost, we should carefully interpret this data, making additional

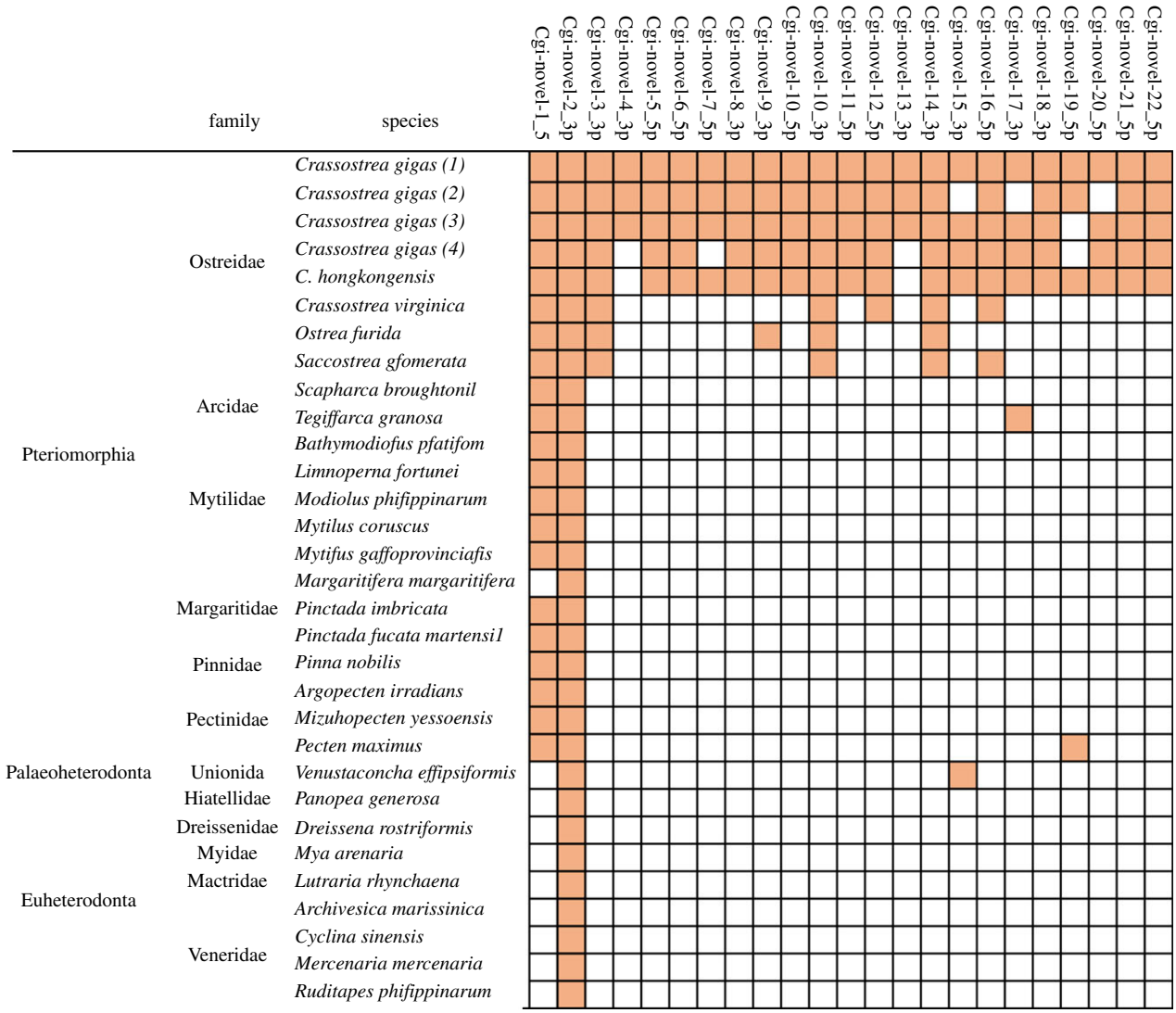

Figure 2. Conservation among bivalves of miRNA complements previously reported as oyster novelties. The conservation of the 23 novel oyster miRNAs retrieved from MirGeneDB was investigated in the genomes of 28 bivalve species. Four *C. gigas* genome assemblies have been considered, (1) oyster_v9, (2) ASM29789v2, (3) ASM1103280v1 and (4) cgigas_uk_roslin_v1.

studies on genome assembly completeness in order to ascertain this point.

Both mussel and clam presented an expanded number of *MIR-184*, *MIR-2*, *MIR-92*, *MIR-96*, *MIR-22*, *MIR-10*, *MIR-67*, *MIR-242*, *MIR-375* and *MIR-1994*, although some differences in their number were noticed also in the comparison with the other analysed bivalves (electronic supplementary material, figure S4). Interestingly, only *R. philippinarum* (the only Palaeoheterodonta analysed) possesses 5 *MIR-375* copies.

(c) Bivalve-specific miRNA families are few

Crassostrea gigas was the unique bivalve species for which the miRNA genomic landscape was investigated [32], taking advantage of the early availability of the genome [59]. This revealed the conservation of the *MIR-17-92* cluster, the presence of duplicated *MIR-2* and *MIR-9* clusters, as well as the *MIR-1991* cluster (*MIR-10* family) [32]. A dataset of 153 manually curated oyster miRNAs is available in MirGeneDB [28] and includes at least one complement for each of the conserved miRNA families expected in agreement with the phylogenetic position of molluscs. Twenty-three miRNA complements represent oyster novelties and most of them are characterized by a very narrowed expression. One exception is represented by *Cgi-novel-16*, which showed broader expression levels in the analysed oyster samples whereas sequence reanalysis revealed a match with the *MIR-96* family. The next MirGeneDB release will fix this issue, demonstrating the importance of manually curated and updated databases. We further tested if any of these novel miRNAs were conserved in the genomes of 28 bivalve species (electronic supplementary material, file S2). As a result, *Cgi-novel-2* was found conserved in all the analysed bivalve genomes, and therefore it should be considered as a new bivalve-specific miRNA family (figure 2). Similarly, *Cgi-novel-1* appeared to be conserved in all Pteriomorphia species, except in *Margaritifera margaritifera*, probably because of incomplete genome sequencing. *Cgi-novel-3*, *Cgi-novel-10*, *Cgi-novel-14* and *Cgi-novel-16* (*Cgi-Mir-96*) appeared to be conserved at the family level, while *Cgi-novel-12* was conserved among *Crassostrea* species. All the other miRNAs (except for *Cgi-novel-4* and *Cgi-novel-13*) were shared between *C. gigas* and *C. hongkonkiensis*, supporting the similarity between these two oyster species (figure 2). Notably, a few differences were also detected among different genome assemblies of *C. gigas*.

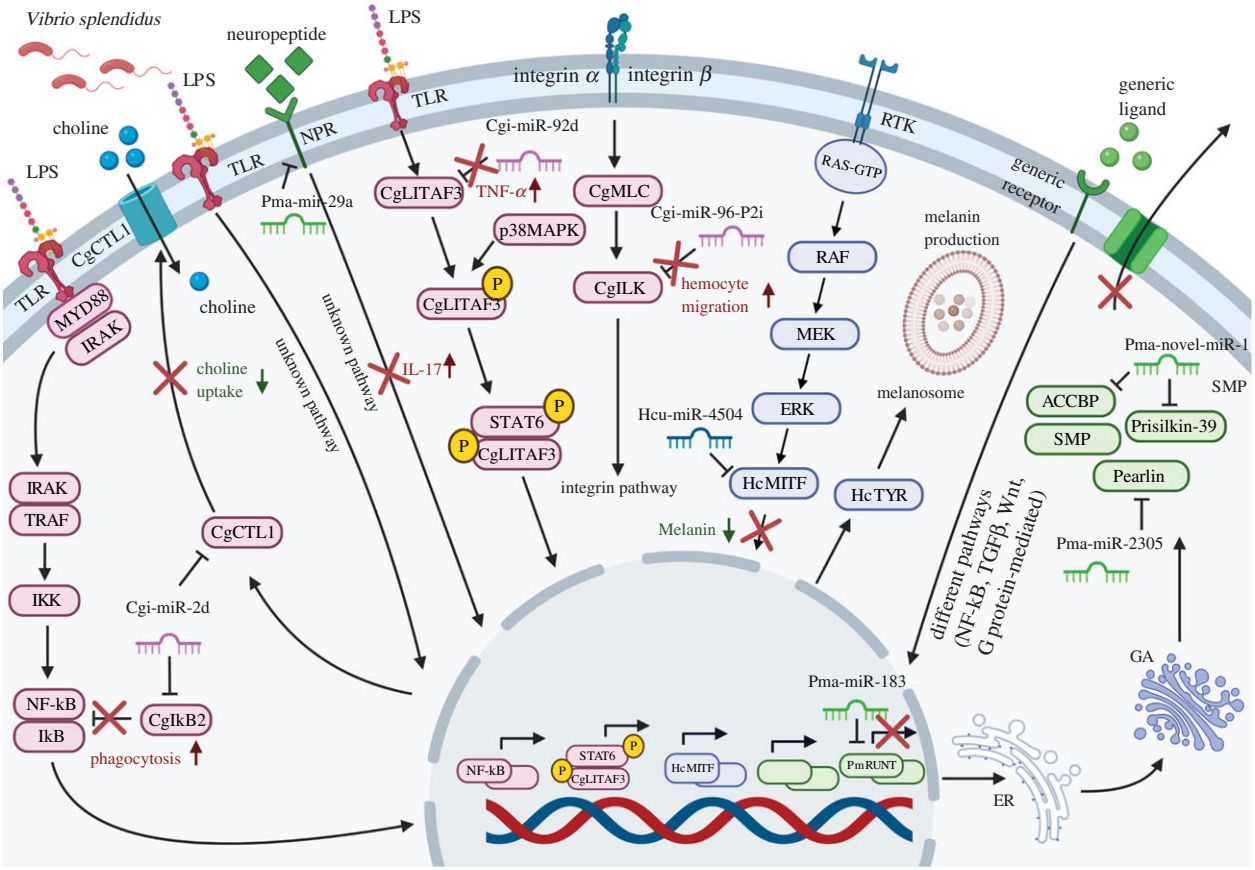

Figure 3. Gene pathways regulated by miRNAs in bivalves. According to published data on validated miRNA-mRNA interactions, the gene pathways regulated by *C. gigas* miRNAs (Cgi-miR, in pink), *P. fucata* (Pma-miR, in green) and *H. comingii* (Hcu-miR, in blue) are depicted. The picture was drawn with BioRender.com.

(d) About the functional roles of bivalve miRNAs

The role of bivalve miRNAs has mostly been investigated in relation to biomineralization and immunity (table 1). The validation of miRNAs putatively identified by sncRNA sequencing is often based on stem-loop RT-qPCR or on the dual-luciferase assay, the latter aimed at verifying the miRNA interactions with their putative target genes. Below, we provide a comparative overview of the bivalve miRNA functions investigated so far (figure 3). Specific miRNAs have been demonstrated to play a role during the immune response of *C. gigas* against *Vibrio splendidus*. Following bacterial infection, the invertebrate-specific miRNA *Cgi-Mir-2d* is upregulated and leads to the enhancement of Nf-kB-mediated gene transcription. Consequently, the phagocytosis rate and apoptosis rate of haemocytes are significantly increased and reduced, respectively [41]. The same *Cgi-Mir-2d* was reported to play a role in the repression of acetylcholine synthesis/release by targeting Choline Transporter-Like proteins (CgCTL1) and by reducing the coline uptake in hemocytes, thus promoting their defense activity [42]. Conversely, *Cgi-Mir-92d* and *scaffold42648-5080* (renamed *Cgi-Mir-96-P2i* in MirGeneDB) are both downregulated after oyster injection with *V. splendidus*. Members of the *MIR-17-92* family are able to modulate the expression of cytokines [60]. In detail, *Cgi-Mir-92* can indirectly regulate the expression of CgTNF by interacting with the lipopolysaccharide induced TNF-α factor (CgLITAF3) CDS region [40]. *Cgi-Mir-96-P2i* plays a regulatory role in the integrin pathway and its under-expression has the effect of increasing hemocyte migration, allowing a more effective immune reaction [43]. Several miRNAs whose functional role has been confirmed, participate in the biomineralization of *P. fucata*. Biomineralization is a sophisticated process, finely tuned by many cellular molecules in turn acting in different pathways. In the pearl oyster, miRNAs, such as *novel-Mir-1* and *Pma-Mir-2305*, target mRNAs for shell matrix proteins (SMPs) which are considered key components in the shell and pearl formation [37,61,62]. Also, the expression of SMPs can be indirectly modulated by the interaction of miRNAs with components located upstream in the biomineralization-related pathways, such as pmRUNT-1 (by *Pma-Mir-183* [63]) or the neuropeptide receptor Y2R (by *Pma-Mir-29a* [38]). As regards *Hyriopsis cumingii*, a peculiar miRNA (*Hcu-Mir-4504*) participates in the regulation of pigment biosynthesis. Intriguingly, the binding of such a miRNA to the melanocyte inducing transcription factor (HcMITF) can impair melanin production. In fact, MITF regulates genes involved in melanin biogenesis, interacting via a typical basic helix-loop-helix leucine zipper (bHLH Zip) to E-box promoters of genes such as tyrosinase (TYR), tyrosinase-related protein 1 (TRP-1) and dopachrome tautomerase (DCT) [64]. Melanin is a key component in the nacre and colour formation but at the same time it plays an important role in encapsulation, a process blocking pathogens into a melanin wall by the activation of biomineralization-like pathways. So the regulation of melanin production could impact both biomineralization and immunity.

4. Conclusion

The available data on bivalve miRNAs made evident a biased representation of only a few species, which impairs phylogenetic investigations on the distribution and conservation of miRNA families. By re-evaluating and reannotating conserved miRNAs reported in previous studies, and expanding miRNAome data with two additional species, we showed that previous annotations were incomplete, and that there is a high conservation of miRNA families, consistent with the taxonomic position of bivalves. While this finding was not unexpected, it opens novel paths to study the evolution of bivalves with miRNAs and, *vice versa*, miRNA evolution in bivalves, as we have identified distinct patterns of miRNA expansion, putative losses, and bivalve-specific miRNA families. Beyond the scope of this review, reannotation of novel or species-specific miRNAs could help the understanding of how the regulatory role of miRNAs contributes to the adaptation of individual species to their habitats.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material.

Authors' contributions. U.R. and B.F. conceptualized the paper and analysed the sncRNA data; C.M.B., A.F. and B.N. performed the sncRNA-seq experiments; E.B. analysed the functional pathways. U.R. wrote the manuscript; P.V. and B.F. substantially contributed to the discussion. All the authors have revised and approved the final manuscript.

Competing interests. We declare we have no competing interests

Funding. The production of *S. broughtonii* sncRNA data was supported by the National Natural Science Foundation of China (grant numbers U1706204 and 32073014). The work was also partially supported by the European H2020 project VIVALDI (GA 678589).

References

1. FAO. 2020 *The state of world fisheries and aquaculture 2020: sustainability in action*. Rome, Italy: FAO. See http://www.fao.org/documents/card/en/c/ca9229en.
2. Morton B, Machado FM. 2019 Predatory marine bivalves: a review. *Adv. Mar. Biol.* **84**, 1–98. (doi:10.1016/bs.amb.2019.10.001)
3. Tan Y, Fang L, Qiu M, Huo Z, Yan X. 2020 Population genetics of the Manila clam (*Ruditapes philippinarum*) in East Asia. *Sci. Rep.* **10**, 21890. (doi:10.1038/s41598-020-78923-w)
4. Fernández Robledo JA *et al.* 2019 From the raw bar to the bench: bivalves as models for human health. *Dev. Comp. Immunol.* **92**, 260–282. (doi:10.1016/j.dci.2018.11.020)
5. Takeuchi T *et al.* 2016 Bivalve-specific gene expansion in the pearl oyster genome: implications of adaptation to a sessile lifestyle. *Zool. Lett.* **2**, 3. (doi:10.1186/s40851-016-0039-2)
6. Gerdol M *et al.* 2020 Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. *Genome Biol.* **21**, 275. (doi:10.1186/s13059-020-02180-3)
7. Du X *et al.* 2017 The pearl oyster *Pinctada fucata martensii* genome and multi-omic analyses provide insights into biomineralization. *GigaScience* **6**, 1–12. (doi:10.1093/gigascience/gix089)
8. Figueras A, Moreira R, Sendra M, Novoa B. 2019 Genomics and immunity of the Mediterranean mussel *Mytilus galloprovincialis* in a changing environment. *Fish Shellfish Immunol.* **90**, 440–445. (doi:10.1016/j.fsi.2019.04.064)
9. Nguyen TC, Zaleta-Rivera K, Huang X, Dai X, Zhong SRNA. 2018 Action through interactions. *Trends Genet.* **34**, 867–882. (doi:10.1016/j.tig.2018.08.001)
10. Wright MW, Bruford EA. 2011 Naming 'junk': human non-protein coding RNA (ncRNA) gene nomenclature. *Hum. Genomics* **5**, 90–98. (doi:10.1186/1479-7364-5-2-90)
11. Kinoshita C, Okamoto Y, Aoyama K, Nakaki T. 2020 MicroRNA: a key player for the interplay of circadian rhythm abnormalities. *Sleep Disorders Neurodegener. Dis.* **2**, 282–307. (doi:10.3390/clockssleep2030022)
12. Saw PE, Xu X, Chen J, Song E-W. 2020 Non-coding RNAs: the new central dogma of cancer biology. *Sci. China Life Sci.* **64**, 22–50. (doi:10.1007/s11427-020-1700-9)
13. Gomes AQ, Nolasco S, Soares H. 2013 Non-coding RNAs: multi-tasking molecules in the cell. *Int. J. Mol. Sci.* **14**, 16 010–16 039. (doi:10.3390/ijms140816010)
14. Bhaskaran M, Mohan M. 2014 MicroRNAs: history, biogenesis, and their evolving role in animal development and disease. *Vet. Pathol.* **51**, 759–774. (doi:10.1177/0300985813502820)
15. Eddy SR. 2001 Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2**, 919–929. (doi:10.1038/35103511)
16. Pal AS, Kasinski AL. 2017 Animal models to study microRNA function. *Adv. Cancer Res.* **135**, 53–118. (doi:10.1016/bs.acr.2017.06.006)
17. Trzybulska D, Vergadi E, Tsatsanis C. 2018 miRNA and other non-coding RNAs as promising diagnostic markers. *EJIFCC* **29**, 221–226.
18. Lee RC, Feinbaum RL, Ambros V. 1993 The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854. (doi:10.1016/0092-8674(93)90529-Y)
19. Bartel DP. 2018 Metazoan microRNAs. *Cell* **173**, 20–51. (doi:10.1016/j.cell.2018.03.006)
20. Moran Y, Agron M, Praher D, Technau U. 2017 The evolutionary origin of plant and animal microRNAs. *Nat. Ecol. Evol.* **1**, 27. (doi:10.1038/s41559-016-0027)
21. Tarver JE *et al.* 2018 Well-annotated microRNAomes do not evidence pervasive miRNA loss. *Genome Biol. Evol.* **10**, 1457–1470. (doi:10.1093/gbe/evy096)
22. Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ. 2009 The deep evolution of metazoan microRNAs. *Evol. Dev.* **11**, 50–68. (doi:10.1111/j.1525-142X.2008.00302.x)
23. Tarver JE, Sperling EA, Nailor A, Heimberg AM, Robinson JM, King BL, Pisani D, Donoghue PCJ, Peterson KJ. 2013 miRNAs: small genes with big potential in metazoan phylogenetics. *Mol. Biol. Evol.* **30**, 2369–2382. (doi:10.1093/molbev/mst133)
24. Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019 miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**(D1), D155–D162. (doi:10.1093/nar/gky1141)
25. Kenny NJ, Namigai EKO, Marlétaz F, Hui JHL, Shimeld SM. 2015 Draft genome assemblies and predicted microRNA complements of the intertidal lophotrochozoans *Patella vulgata* (Mollusca, Patellogastropoda) and *Spirobranchus* (*Pomatoceros*) *lamarcki* (Annelida, Serpulida). *Mar. Genomics* **24**, 139–146. (doi:10.1016/j.margen.2015.07.004)
26. Thomson RC, Plachetzki DC, Mahler DL, Moore BR. 2014 A critical appraisal of the use of microRNA data in phylogenetics. *Proc. Natl Acad. Sci. USA* **111**, E3659–E3668. (doi:10.1073/pnas.1407207111)
27. Fromm B, Keller A, Yang X, Friedlander MR, Peterson KJ, Griffiths-Jones S. 2020 Quo vadis microRNAs? *Trends Genet.* **36**, 461–463. (doi:10.1016/j.tig.2020.03.007)
28. Fromm B *et al.* 2020 MirGeneDB 2.0: the metazoan microRNA complement. *Nucleic Acids Res.* **48**. (doi:10.1093/nar/gkz885)
29. Jiao Y, Zheng Z, Du X, Wang Q, Huang R, Deng Y, Shi S, Zhao X. 2014 Identification and characterization of microRNAs in pearl oyster *Pinctada martensii* by Solexa deep sequencing. *Mar. Biotechnol.* **16**, 54–62. (doi:10.1007/s10126-013-9528-x)
30. Chen G, Zhang C, Jiang F, Wang Y, Xu Z, Wang C. 2014 Bioinformatics analysis of hemocyte miRNAs of scallop *Chlamys farreri* against acute viral necrobiotic virus (AVNV). *Fish Shellfish Immunol.* **37**, 75–86. (doi:10.1016/j.fsi.2014.01.002)
31. Zhou Z, Wang L, Song L, Liu R, Zhang H, Huang M, Chen H. 2014 The identification and characteristics of immune-related microRNAs in haemocytes of

oyster *Crassostrea gigas*. *PLoS ONE* **9**, e88397. (doi:10.1371/journal.pone.0088397)

32. Xu F, Wang X, Feng Y, Huang W, Wang W, Li L, Fang X, Que H, Zhang G. 2014 Identification of conserved and novel microRNAs in the Pacific oyster *Crassostrea gigas* by deep sequencing. *PLoS ONE* **9**, e104371. (doi:10.1371/journal.pone.0104371)
33. Bao Y, Zhang L, Dong Y, Lin Z. 2014 Identification and comparative analysis of the *Tegillarca granosa* haemocytes microRNA transcriptome in response to Cd using a deep sequencing approach. *PLoS ONE* **9**, e93619. (doi:10.1371/journal.pone.0093619)
34. Martín-Gómez L, Villalba A, Kerkhoven RH, Abollo E. 2014 Role of microRNAs in the immunity process of the flat oyster *Ostrea edulis* against bonamiosis. *Infect. Genet. Evol.* **27**, 40–50. (doi:10.1016/j.meegid.2014.06.026)
35. Chen H *et al*. 2015 The comprehensive immunomodulation of NeurimmiRs in haemocytes of oyster *Crassostrea gigas* after acetylcholine and norepinephrine stimulation. *BMC Genom.* **16**, 942. (doi:10.1186/s12864-015-2150-8)
36. Zheng Z, Jiao Y, Du X, Tian Q, Wang Q, Huang R, Deng Y. 2016 Computational prediction of candidate miRNAs and their potential functions in biomineralization in pearl oyster *Pinctada martensii*. *Saudi J. Biol. Sci.* **23**, 372–378. (doi:10.1016/j.sjbs.2015.04.001)
37. Jiao Y, Zheng Z, Tian R, Du X, Wang Q, Huang R. 2015 MicroRNA, Pm-miR-2305, participates in nacre formation by targeting pearlin in pearl oyster *Pinctada martensii*. *Int. J. Mol. Sci.* **16**, 21 442–21 453. (doi:10.3390/ijms160921442)
38. Tian R, Zheng Z, Huang R, Jiao Y, Du X. 2015 miR-29a participated in nacre formation and immune response by targeting Y2R in *Pinctada martensii*. *Int. J. Mol. Sci.* **16**, 29 436–29 445. (doi:10.3390/ijms161226182)
39. Zhao X, Yu H, Kong L, Liu S, Li Q. 2016 High throughput sequencing of small RNAs transcriptomes in two *Crassostrea* oysters identifies microRNAs involved in osmotic stress response. *Sci. Rep.* **6**, 22687. (doi:10.1038/srep22687)
40. Chen H, Jiang S, Wang L, Wang L, Wang H, Qiu L, Song L. 2016 Cgi-miR-92d indirectly regulates TNF expression by targeting CDS region of lipopolysaccharide-induced TNF-α factor 3 (CgLITAF3) in oyster *Crassostrea gigas*. *Fish Shellfish Immunol.* **55**, 577–584. (doi:10.1016/j.fsi.2016.06.036)
41. Chen H, Zhou Z, Wang H, Wang L, Wang W, Liu R, Qiu L, Song L. 2016 An invertebrate-specific and immune-responsive microRNA augments oyster haemocyte phagocytosis by targeting CgIκB2. *Sci. Rep.* **6**, 1–10. (doi:10.1038/s41598-016-0001-8)
42. Chen H, Zhou Z, Wang L, Wang H, Liu R, Zhang H, Song L. 2016 An invertebrate-specific miRNA targeted the ancient cholinergic neuroendocrine system of oyster. *Open Biol.* **6**, 160059. (doi:10.1098/rsob.160059)
43. Chen H, Wang H, Jiang S, Xu J, Wang L, Qiu L, Song L. 2016 An oyster species-specific miRNA scaffold42648_5080 modulates haemocyte migration by targeting integrin pathway. *Fish Shellfish Immunol.* **57**, 160–169. (doi:10.1016/j.fsi.2016.08.032)
44. Chen H *et al.* 2017 A norepinephrine-responsive miRNA directly promotes CgHSP90AA1 expression in oyster haemocytes during desiccation. *Fish Shellfish Immunol.* **64**, 297–307. (doi:10.1016/j.fsi.2017.03.020)
45. Pozzi A, Plazzi F, Milani L, Ghiselli F, Passamonti M. 2017 SmithRNAs: could mitochondria 'bend' nuclear regulation? *Mol. Biol. Evol.* **34**, 1960–1973. (doi:10.1093/molbev/msx140)
46. Rosani U, Shapiro M, Venier P, Allam B. 2019 A needle in a haystack: tracing bivalve-associated viruses in high-throughput transcriptomic data. *Viruses* **11**, 205. (doi:10.3390/v11030205)
47. Wei P *et al*. 2019 Identification and characterization of microRNAs in the gonads of *Crassostrea hongkongensis* using high-throughput sequencing. *Comp. Biochem. Physiol. D* **31**, 100606. (doi:10.1016/j.cbd.2019.100606)
48. Chen X, Bai Z, Li J. 2019 The mantle exosome and microRNAs of *Hyriopsis cumingii* involved in nacre color formation. *Mar. Biotechnol.* **21**, 634–642. (doi:10.1007/s10126-019-09908-8)
49. Yu D, Wu H, Peng X, Ji C, Zhang X, Song J, Qu J. 2020 Profiling of microRNAs and mRNAs in marine mussel *Mytilus galloprovincialis*. *Comp. Biochem. Physiol. C* **230**, 108697. (doi:10.1016/j.cbpc.2019.108697)
50. Rosani U, Abbadi M, Green T, Bai C-M, Turolla E, Arcangeli G, Wegner KM, Venier P. 2020 Parallel analysis of miRNAs and mRNAs suggests distinct regulatory networks in *Crassostrea gigas* infected by Ostreid herpesvirus 1. *BMC Genomics* **21**, 620. (doi:10.1186/s12864-020-07026-7)
51. Moreira R, Romero A, Rey-Campos M, Pereiro P, Rosani U, Novoa B, Figueras A. 2020 Stimulation of *Mytilus galloprovincialis* hemocytes with different immune challenges induces differential transcriptomic, miRNomic, and functional responses. *Front. Immunol.* **11**, 606102. (doi:10.3389/fimmu.2020.606102)
52. Gerdol M *et al*. 2019 Massive gene presence/absence variation in the mussel genome as an adaptive strategy: first evidence of a pan-genome in Metazoa. *bioRxiv*. See https://www.biorxiv.org/content/10.1101/781377v1.
53. Fromm B *et al.* 2015 A uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. *Annu. Rev. Genet.* **49**, 213–242. (doi:10.1146/annurev-genet-120213-092023)
54. Martin M. 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12. (doi:10.14806/ej.17.1.200)
55. Kang W, Eldfjell Y, Fromm B, Estivill X, Biryukova I, Friedländer MR. 2018 miRTrace reveals the organismal origins of microRNA sequencing data. *Genome Biol.* **19**, 213. (doi:10.1186/s13059-018-1588-9)
56. Waldron FM, Stone GN, Obbard DJ. 2018 Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. *PLoS Genet*. **14**, e1007533. (doi:10.1371/journal.pgen.1007533)
57. Jehn J, Gebert D, Pipilescu F, Stern S, Kiefer JST, Hewel C, Rosenkranz D. 2018 PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation. *Commun. Biol.* **1**, 137. (doi:10.1038/s42003-018-0141-4)
58. Bai C-M, Xin L-S, Rosani U, Wu B, Wang Q-C, Duan X-K, Liu ZH, Wang CM. 2019 Chromosomal-level assembly of the blood clam, *Scapharca* (*Anadara*) *broughtonii*, using long sequence reads and Hi-C. *GigaScience* **8**, giz067.
59. Zhang G *et al.* 2012 The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**, 49–54. (doi:10.1038/nature11413)
60. Liu M *et al.* 2011 TNF-α is a novel target of miR-19a. *Int. J. Oncol.* **38**, 1013–1022. (doi:10.3892/ijo.2011.924)
61. Suzuki M, Nagasawa H. 2013 Mollusk shell structures and their formation mechanism. *Can. J. Zool.* **91**, 349–366. (doi:10.1139/cjz-2012-0333)
62. Zhu X, Chen Y, Zhang Z, Zhao S, Xie L, Zhang R. 2020 A species-specific miRNA participates in biomineralization by targeting CDS regions of Prisilkin-39 and ACCBP in *Pinctada fucata*. *Sci. Rep.* **10**, 8971. (doi:10.1038/s41598-020-65708-4)
63. Zheng Z, Du X, Xiong X, Jiao Y, Deng Y, Wang Q, Huang R. 2017 *PmRunt* regulated by *Pm-miR-183* participates in nacre formation possibly through promoting the expression of *collagen VI-like* and *Nacrein* in pearl oyster *Pinctada martensii*. *PLoS ONE* **12**, e0178561. (doi:10.1371/journal.pone.0178561)
64. Wang R, Chen T, Zhao B, Fan R, Ji K, Yu X, Wang X, Dong C. 2017 FGF21 regulates melanogenesis in alpaca melanocytes via ERK1/2-mediated MITF downregulation. *Biochem. Biophys. Res. Commun.* **490**, 466–471. (doi:10.1016/j.bbrc.2017.06.064)