The miRNA biogenesis in marine bivalves.

Rosani U[1], Pallavicini A[2], Venier P[1]

[1] Dept. of Biology, University of Padua, via U. Bassi 58/b, 35121 Padua, ITALY

[2] Department of Life Sciences, University of Trieste, via L. Giorgeri 5, 34127 Trieste, ITALY

Corresponding author: Paola Venier

Abstract

Small non-coding RNAs include powerful regulators of gene expression, transposon mobility and virus activity. Among the various categories, mature microRNAs (miRNAs) guide the translational repression and decay of several targeted mRNAs. The biogenesis of miRNAs depends on few gene products, essentially conserved from basal to higher metazoans, whose protein domains allow specific interactions with dsRNA. Here, we report the identification of key genes responsible of the miRNA biogenesis in 32 bivalves, with particular attention to the aquaculture species *Mytilus galloprovincialis* and *Crassostrea gigas*. We have identified and phylogenetically compared eight evolutionary conserved proteins: DROSHA, DGCR8, EXP5, RAN, DICER TARBP2, AGO and PIWI. In mussels, we have also identified several other proteins participating in the miRNA biogenesis or in the subsequent RNA silencing. According to digital expression analysis, these genes display low and not inducible expression levels in adult mussels and oysters whereas they are considerably expressed during development. As miRNAs play an important role also in the antiviral responses, knowledge on their biogenesis and regulatory effects can shed light on essential molecular processes and provide new hints for disease prevention in bivalves.

Steven Roberts 12/24/2015 12:30 AM
Deleted: In detail, w

Introduction

A number of different types of non-coding RNAs (ncRNAs) have gained attention for their powerful regulatory action on eukaryotic genes and other genetic elements (Carninci et al. 2005; Mortimer et al. 2014). The process known as RNA interference (RNAi) exemplifies an evolutionary conserved mechanism of gene silencing based on small guide RNAs and specific interacting proteins (Tomoyasu et al. 2008; Gammon and Mello 2015). Silencing RNAs (siRNAs) and microRNAs (miRNAs) take part to the same silencing machinery whereas Piwi-interacting RNAs (piRNAs) peculiarly silence germ-line transposons, among other roles (Théron et al. 2014; Iwasaki et al. 2015). Long noncoding RNAs (lncRNAs) in their normal or mutated forms can widely influence physiological and pathological processes, as multiple lines of evidence indicate their involvement in chromosome inactivation and epigenetic modifications, control of mRNA decay and translation, and DNA sequestration of transcription factors (Huarte 2015; Ruan 2015). More recently, circular RNAs have been identified as a group of competing endogenous RNAs whose effects in the miRNA function and transcriptional/post-transcriptional regulation are now matter of study (Qu et al. 2015).

miRNAs are single-stranded RNA molecules of around 22 nucleotides, presenting conserved structural features and able to modulate the expression of eukaryotic genes by inhibition of mRNA translation or enhancement of mRNA decay (Ambros 2003; Bartel 2004; Tarver et al. 2012). Up to now, diversified sets of miRNAs have been detected in five eukaryotic taxa (eumetazoans, silicisponges, vascular plants, *Clamydomonas* and *Ectocarpus* spp.) while they are apparently absent in protists (Grimson et al. 2008; Tarver et al. 2012). In humans, the 2588 putative miRNAs described in the miRBase v. 21 (Kozomara and Griffiths-Jones 2014) could target 30-60 % of the transcribed genes (John et al. 2004; Sand et al. 2012), with implications in cell differentiation (Berezikov et al. 2005), cell death (Xu et al. 2015), stress responses (Mendell and Olson 2012) and diseases (Huang et al. 2014; Min and Chan 2015).

miRNA biogenesis starts from pri-miRNA transcripts, mostly generated from RNA polymerase II in form of long non-coding RNAs and able to form a hairpin subsequently recognized by the so called microprocessor complex. DROSHA, a double-stranded RNA-specific ribonuclease III, and the RNA binding protein *Di-George syndrome Critical Region gene 8* (DGCR8) are the microprocessor's core proteins which allow interactions with the DDX5 helicase, the RNA binding protein Lin-28 and hnRNP A1, among other elements (Jean-Philippe et al. 2013; Hong et al. 2013). During the recognition of pri-miRNAs at the dsRNA-ssRNA junction, DGCR8 acts as a crucial molecular anchor and directs DROSHA to cleave 11 bp away from the junction, with consequent release of hairpin-shaped pre-miRNAs (Denli et al. 2004). Pre-miRNAs are firstly exported to the cytoplasm via the *Exportin5* (XPO5) by interaction with the small GTPase RAN; then, they are further processed by the RISC loading complex, composed by the endoribonuclease DICER, the RNA binding protein TARBP2 and Argonaute proteins (MacRae et al. 2008; Miyoshi et al. 2009). The evolutionary conserved Argonaute proteins are

Steven Roberts 12/24/2015 12:33 AM
Deleted: In the time, various

Steven Roberts 12/24/2015 12:36 AM
Deleted: the

Steven Roberts 12/24/2015 12:36 AM
Deleted: the

Steven Roberts 12/24/2015 12:37 AM
Deleted: In detail, the

specialized in binding small RNAs and exist in several isoforms, with AGO and PIWI representing two distinct subclades (Tolia and Joshua-Tor 2007; Ender and Meister 2010).

AGOs select the 'guide' miRNA strand necessary for targeted gene silencing and, therefore, are responsible for final miRNA maturation. Several other proteins have been demonstrated to cooperate in miRNA processing and functions (Ender and Meister 2010). In fact, AGOs operate transcriptional repression and cause mRNA decay by interacting with the GW-rich N-terminal region of GW182, a protein associated with cellular P-bodies (van Kouwenhove et al. 2011). Other proteins involved in the mRNA turnover (CAF1, PABPC1, eIF4G; CCR4-NOT and PAN2-PAN3 deadenylation complexes; also decapping complex DCP1-DCP2 in human somatic cells and at least four helicases, DDX5, DDX6, DDX17 and DDX42) may also cooperate with the AGO-GW182 complex to reduce the mRNA translation efficiency (Nottrott et al. 2006; Fabian and Sonenberg 2012).

Unlike AGOs, the PIWI proteins specifically interact with piRNAs to participate in the germline specification, gametogenesis, transposon silencing and in the maintenance of genome integrity (Carmell et al. 2007; Malone and Hannon 2009; Ghildiyal and Zamore 2009; Siomi et al. 2011). The piRNA mechanism of action is not so well defined but probably involves the *arginine methyl-transferase* PRMT5, *tudor domain-containing proteins* (TDRDs) and the *Maelstrom* protein (MAEL) (Sokolova et al. 2011).

With the widespread and cost-effective use of Next Generation Sequencing (NGS) technologies, miRNAs have been deeply explored in non-model organisms, including bacteria (Xu et al. 2014), plants (Rhee et al. 2015) and viruses (Kincaid and Sullivan 2012; Diebel et al. 2015). The basic set of genes involved in the miRNA biogenesis and the main protein interactions are well known in mammals (Lau and MacRae 2009), and also in other metazoans like *Cnidaria* (Moran et al. 2013), *Platyhelminthes* (Resch and Palakodeti 2012) and insects (Lucas and Raikhel 2013; Hussain and Asgari 2014). Regarding mollusks, lists of miRNAs have been reported for a few species (Jiao et al. 2014; Chen et al. 2014; Martín-Gómez et al. 2014; Zhou et al. 2014), miRNA families have been investigated in the limpet and polychaete genomes (Kenny et al. 2015) and one phylogenetic study has included bivalve DICER sequences (Gao et al. 2014). Since a general overview on the bivalve miRNA biogenesis complements is still lacking, we took advantage of several genomic and transcriptomic datasets available for *Lophotrochozoa* (GIGA Community of Scientists et al. 2014) to identify and characterize the core elements of the miRNA biogenesis pathway, with particular attention to *Mytilus* and *Crassostrea spp.*

Steven Roberts 12/24/2015 12:40 AM
Deleted: of the

Steven Roberts 12/24/2015 12:41 AM
Deleted: the

Steven Roberts 12/24/2015 12:42 AM
Deleted: also

Steven Roberts 12/24/2015 12:43 AM
Deleted: About

Steven Roberts 12/24/2015 12:44 AM
Comment [1]: not a mollusk

Steven Roberts 12/24/2015 12:44 AM
Deleted: , so far, only

Steven Roberts 12/24/2015 12:45 AM
Deleted: some

Materials & Methods

Sequences coding for proteins centrally involved in the miRNA pathway, namely DROSHA, DGCR8, XPO5, RAN, DICER, TARBP2, AGO and PIWI, have been methodically identified in the genomes and transcriptomes of *M. galloprovincialis* (Mg) and *C. gigas* (Cg) as well as in other bivalve and non-bivalve species (66 species, listed in Table 1).

Sequence retrieval and analysis

The Mg WGS project (ID APJB000000000.1 (Nguyen et al. 2014)) and the Cg genome draft (GCA_000297895 (Zhang et al. 2012)) were retrieved from GenBank, whereas the oyster genome annotations were obtained from Ensembl Metazoa release 29 (http://metazoa.ensembl.org/Crassostrea_gigas/Info/Index). A Mg reference transcriptome was produced using 18,788 ESTs of mixed tissues previously obtained by Sanger sequencing (Venier et al. 2009) and 453 million reads obtained by paired-end (2 ×100 bp) Illumina Hiseq2000 sequencing of digestive gland from North Adriatic Sea mussels (ID: PRJNA88481) (Gerdol et al. 2014), and from haemocytes, gills, mantle, and muscle of Spanish mussels (ID: SRP033481) (Moreira et al. 2015). The quality of the sequencing readout was evaluated by the FastQC suite (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) discarding the reads with PHRED quality below 20 and presenting more than 2 ambiguous nucleotides. D*e-novo* assembly was performed with Trinity, release 2013-08-14 (Grabherr et al. 2011) setting the minimum contig length at 200 bp and using default settings. Subsequently, protein coding sequences (cds) were predicted with Transdecoder (Grabherr et al. 2011). Transcriptomic reads of 30 bivalve *spp*. (Cg plus other 29 species) were retrieved from the SRA archive and assembled as described above (details in Supplementary File 1). The protein predictions of further 33 organisms were directly retrieved from public repositories or extracted from the corresponding genome releases. The NCBI transcriptome shotgun assembly (TSA) database was interrogated to retrieve hits for two additional cnidarians, *Porites australiensis* and *Anthopleura elegantissima* (Table 1).

Protein domain searches

To investigate the presence of eight key proteins of miRNA biogenesis (DROSHA, DGCR8, XPO5, RAN, DICER, TARBP2, AGO and PIWI), we downloaded their predictive HMM from PFAM v.27 (listed in Table 2) and we scanned the sequence datasets with HMMer v3.1 (Eddy 2011) applying a cut-off E-value of 0.01. To achieve a meaningful comparison of proteins from different organisms, we retained only hits presenting all diagnostic domains. Moreover, we identified several mussel transcripts related to protein interactions occurring in the miRNA biogenesis. To identify such proteins, we retrieved from PFAM the diagnostic domains of human homologs (listed in Table 3) and we scanned their presence in the Mg transcriptome as described above. Protein domain organization was reconstructed using SMART (Letunic et al. 2012)

Gene structure analysis

Steven Roberts 12/24/2015 12:47 AM
Deleted: el

Steven Roberts 12/24/2015 12:48 AM
Deleted: apart other

Steven Roberts 12/24/2015 12:50 AM
Deleted: the

We used the transcript sequences of DROSHA, DGCR8, XPO5, DICER and TARBP2 as blast queries against all Mg genomic contigs (*blastn*) in order to recover the related gene structures. Positive hits having e-value lower than 10^{-20} were extracted and assembled on the corresponding transcript, used as backbone. RNA-seq read mappings with adapted parameters (CLC Genomic Workbench *large gap mapping tool*, with similarity and length fraction set at 0.9) allowed us to ascertain the correct gene assembly. Homolog gene structures were retrieved by interrogating genomic browsers, like Metazome v.3 (for C*. intestinalis, B. floridae, D. rerio, S. kowalevskii, S. purpuratus, N. vectensis, T. castaneum, L. gigantea, O. bimaculoides, C. elegans* and *H. sapiens*) and Ensembl Metazoa v.29 (for *C. gigas*, *C. quinquefasciatus, D. melanogaster, N. vitripennis, A. mellifera, A. queenslandica, P. bachei, M. leidyi, T. adhaerens, N. vectensis, D. pulex, S. mansoni, S. mediterranea, A. vaga, L. anatine, H. robusta* and *C. telata*) or by local *blastn* against the downloaded genomes (A*. pisum* and *L. albipes*).

Phylogenetic analysis

The inferred protein sequences were aligned using MUSCLE, release 2014-05-29 (Edgar 2004). Subsequently, the fasta alignments were analyzed using Gblocks v.0.91 (Castresana 2000) to extract conserved positions (positions common to 51 % of the locally aligned sequences). Trees were built using neighbor joining or maximus likelihood clustering methods with 1000 bootstrap replicates. Bayesian phylogenies were reconstructed using MrBayes v.3.2.5 (Ronquist et al. 2012), with GTR substitution evolutionary model with gamma-distributed rate variation across sites, evaluating the convergence after 1,000,000 runs (0.5 was considered as cut-off value). Trees were visualized and edited with FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Digital expression analysis

To analyze the expression of the selected genes in Cg and Mg RNA datasets, we retrieved all available RNA-seq samples from the NCBI SRA archive. For Cg, we analyzed 123 Illumina RNA-seq samples related to adult tissues or developmental stages. For Mg, we analyzed 13 RNA samples from gills (1), digestive gland (6), haemocytes (2), mantle (2) and muscle (2). Overall, we included in the expression analysis 2,271 and 453M reads for Cg and Mg, respectively (Supplementary File 2). The trimmed reads were mapped to Cg and Mg genes using the CLC Genomics Workbench v.8.0 (Qiagen, Germany) mapping tool, with length and similarity fractions set at 0.75 and 0.95, respectively, and mismatch/insertion/deletion penalties at 3/3/3. The number of uniquely mapped reads of each dataset were counted and used to calculate digital expression values as TPM (Transcripts Per Kilobase Million mapped reads) as described by (Wagner et al. 2013), considering 3 TPMs as lower detection limit.

Steven Roberts 12/24/2015 12:57 AM
Deleted:

Results

Mussel transcripts related to the miRNA biogenesis

We identified *Mytilus galloprovincialis* transcripts involved in the miRNA biogenesis by systematic searches of diagnostic domains (Table 2) in a transcriptome assembly produced from 453 million Illumina reads. Thus, we recovered nine transcripts coding for DROSHA, DGCR8, XPO5, RAN, DICER, TARBP2 and for three Argonaute genes (one Ago and two Piwi-like proteins, Table 3). We also identified 21 mussel proteins expected to play a role in the miRNA maturation or involved in RNAi processes (Supplementary File 3).

Steven Roberts 12/24/2015 1:14 AM
Comment [2]: This is Table 3 in version I see in PeerJ

Figure 1 relates the general process of eukaryotic miRNA biogenesis to the mussel proteins identified in this work. MgDROSHA and MgDGCR8 are expected to start the maturation of pri-miRNAs produced by RNA polymerase II. MgDROSHA codes for a 1377 aa length protein containing all the canonical domains (2 RIBOc domains in positions 959-1093 and 1139-1271 and one DSRM domain in position 1278-1351) whereas MgDGCR8 is a 728 aa length protein having one WW domain in position 229-258, necessary for the interaction with DROSHA, and two DSRM domains (positions 472-536 and 578-642) necessary for pri-miRNA binding. MgXPO5 is expected to cooperate with MgRAN in the pre-miRNA cytoplasmic translocation. MgRAN encodes a 214 aa protein whereas MgXPO5 has a length of 1201 aa and includes two 5' conserved domains (IBN_N and Xpo1) and one conserved region necessary for the interaction with interleukin enhancer-binding factor 3 (position 525-562). In mussels, the RISC complex uploading pre-miRNAs is defined by the endoribonuclease MgDICER (1850 aa) and MgTARBP2 (321 aa). Like in *Lophotrococzoa*, mussel DICER is encoded by a unique gene and contains the seven canonical domains, namely two helicase domains, one DICER-dimer domain, one PAZ, two RIBOc and a final DSRM domain. MgTARBP2 displays three DSRM domains in positions 9-73, 101-166 and 249-314. Moreover, *M. galloprovincialis* possess three argonaute proteins ranging from 861 to 941 aa in length and representative of one AGO (DUF1785, PAZ and PIWI domains) and two PIWI-like proteins (PAZ and PIWI domains). We considered the above mentioned gene products as the key complement of the miRNA biogenesis.

Steven Roberts 12/24/2015 1:22 AM
Deleted: ify

Among the possible interacting proteins, we identified MgGW182, a transcript encoding a protein shorter than the human counterparts but holding all the features considered significant for its interaction with AGOs and the CCR4-NOT complex. In fact, MgGW182 possesses 19 N-terminal GW stretches, followed by one UBA domain, a Q-rich region (M domain) and a C-terminal RNA recognition motif (RRM domain). Moreover, we identified a C-terminal conserved site known as PAM2 (Kozlov et al. 2010), expected to interact with the poly(A) binding protein 1 (MgPABPC1) through the MLLE motif and inhibit the mRNA translation by interfering with the mRNA circularization process (Piao et al. 2010; van Kouwenhove et al. 2011). In the mussel transcriptome, we also identified putative homologs for a number of CNOT complex proteins (CNOT1, 2, 3, 6, 7, 9, and 10), for the eukaryotic translation initiation factor 4 gamma, 1 eIF4G, PAB-dependent poly(A)-specific ribonuclease subunits PAN2, PAN3, the decapping complex proteins DCP1 and DCP2, and several RNA helicases

demonstrated to be crucial in the miRNA maturation (DDX5) and RNAi (DDX5- 6- 20 and 42). Finally, we recognized the putative mussel homologs of protein arginine methyltransferase 5 (MgPRMT5), tudor domain containing protein (MgTDRD-11) and maelstrom spermatogenic transposon silencer (MgMAEL).

Mussel genes related to the miRNA biogenesis

Taking advantage of a mussel genome sequencing project (Nguyen et al. 2014) we investigated the organization of the main genes involved in the mussel miRNA biogenesis. Fragmentation of the genomic mussel assembly (2.3 million contigs; 700 bp on average) and considerable dimension of the analyzed genes (9.6 – 17.6 kbp gene size in the case of Cg) prevented the recovery of the full gene sequences. Nevertheless, we can describe the complete gene structures of DROSHA, DGCR8, EXP5, DICER and TARBP2 (i.e. five of eight searched sequences) whose length varies between 7.5 and 27 kbp, confirmed by the back-mapping of 115,377 Illumina paired reads (Figure 2, Supplementary File 4). Moreover, these mussel genes showed a remarkable conservation in terms of exon number when compared with a selection of homolog genes from deuterostome and protostome organisms (Table 4).

Transcripts related to the miRNA biogenesis in bivalve *spp*.

To identify the miRNA biogenesis complements in marine mollusks, we used homologous genes retrieved from the genomes of *C. gigas, L. gigantea* and *A. californica*. Since the *C. gigas* genome includes annotations only for the cds regions, we exploited full-length transcripts obtained from a locally assembled oyster transcriptome to expand the genome annotations in this species. In particular, we updated the annotation of CgDGCR8 and CgDICER and we added new annotations for CgPIWI-1 (CGI_10008757: genomic contig JH815696, position 184178-187825) and CgTARBP2 (JH818440, 414703-419857).

Since many marine bivalve *spp*. do not have at present a sequenced genome, we used publicly available RNA-seq data to build 29 specie-specific transcriptome assemblies and retrieve the homologous sequences of interest. After domain searching, we carefully considered the high number of positive hits to retain only proteins including all the expected protein features. Thus, we retrieved 132 complete hits from marine mollusks: 10 DROSHAs, 9 DGCR8s, 14 XPO5s, 34 RANs, 7 DICERs, 13 TARBP2s and 45 Argonaute-like proteins, the latter classified in 13 AGO and 32 PIWI proteins by phylogenetic analysis (Table 1, Supplementary File 5).

Phylogenetic analysis of the miRNA biogenesis proteins

The inferred sequences of single miRNA biogenesis proteins were aligned together with those retrieved from 34 sequenced genomes. Here, we report the phylogenetic analysis of the five proteins centrally involved in the miRNA biogenesis, namely DROSHA, DGCR8, DICER, TARBP2 and AGOs (Figure 3, A-D, Supplementary File 5 for all considered protein sequences).

We back-traced the presence of a canonical DROSHA up to Cnidaria, although we found only incomplete hits in Porifera and Placozoa and the genomes of *Ctenophora spp.* lack of both DROSHA and DGCR8, as reported by other authors (Maxwell et al. 2012). The DROSHA sequences from Cnidaria's appeared as general outgroup whereas those of Chordata clustered as outgroup of the other protostomes. DROSHAs from Mollusca and Arthropoda clustered consistently with the different *taxa* whereas those from Platyhelmintes, Rotifera, Brachiopoda and Annelida grouped together, with DROSHA from *Caenorhabditis elegans* (Nematoda) being the most far-related (Figure 3A). Conversely to DROSHA, we identified a complete DGCR8 also in the *Porifera Amphimedon queenslandica*, suggesting that also DROSHA should be present in this taxa. Following phylogenetic analysis, we highlighted Cnidaria and Porifera proteins as outgroup, with mollusks (and Annelida) clustering with Arthropoda and more distantly *Platyhelmintes* and *Rotifera* hits. The *Chordata* sequences clustered as a separate group (Figure 3B).

The finding of putative DICER sequences in Ctenophora *spp*. supports the presence of this gene through the whole Opisthokonta evolution. Also plants possess DICER homologues which occur in different copy number among taxa: two genes in Porifera, Placozoa, Cnidaria and Arthropoda; four genes in plants like *A. thaliana* and *P. trichocarpa* and one gene in Ctenophora, Rotifera, Cephalopoda Mollusca and Chordata. Moreover, the presence of DICER was reported in some Protozoa and fungi (Mukherjee et al. 2013). Phylogenetic analyses, firstly separate insect DICER-2 and plant DICERs from the DICER-1 and DICERs from Arthropoda, Mollusca and Chordata (Figure 3C). Likewise, the phylogenetic tree regarding TARBP2 displays a clear cut-off between the proteins of mollusks, chordates and arthropods (Figure 3D). We back-traced the miRNA cytoplasm export complex composed by RAN and XPO5 in all analyzed metazoans. Both RAN and XPO5 represent widely expressed sequences since we found them also in transcriptome assemblies, although with suboptimal sequence coverage.

Several AGO and Piwi proteins can be present in individual organisms and, in fact, we identified a total of 235 proteins. Whereas humans possess eight proteins, we found four proteins in the majority of the analyzed insect *spp.* (with the exception of 15 proteins in *A. pisum*) and three or four different proteins in bivalve *spp*. Also basal Metazoa possess Argonaute-like sequences: four in the genomes of Ctenophora and Cnidaria *spp.*, one in the Placozoa *T. adhaerens* and two in *A. queenslandica*. The case of *C. elegans* is remarkable since it holds several Argonaute gene families and at least 24 proteins (Hoogstrate et al. 2014). In agreement with other phylogenetic studies (Swarts et al. 2014), the Argonaute proteins from plants and the majority of those from *C. elegans* formed distinct clades and, moreover, a clear separation was evident between AGO and PIWI proteins. Bivalve protein sequences clustered always separately forming one cluster for AGO-like hits and two clusters for PIWI-like proteins (Figure 4).

Digital expression analysis of mussel and oyster miRNA biogenesis genes

We used the 13 Mg and 124 Cg RNA-seq samples to evaluate the expression levels of miRNA biogenesis genes in different tissues and conditions. Based on total mapped reads, we computed TPM values and we used

Steven Roberts 12/24/2015 1:38 AM
Deleted: (13)

elongation factor 1 α (El1α) as normalizer housekeeping gene to compare the expression level of the different genes in each sample.

For Mg, the sequence analysis indicated a scarce basal expression of the genes mentioned above in five adult tissues: gill, digestive gland, haemolymph, muscle and mantel (below 2 % of El1α, except for DDX5, RAN and CNOT9). Mantel and muscle appeared the most responsive tissues whereas haemolymph was the least responsive one. In particular, the genes that we considered as the core components of miRNA biogenesis were expressed at levels below 0.5 % of El1α (Supplementary File 6).

For Cg, we analyzed a considerable number of RNA-seq libraries representative of adult tissues (85) and developmental stages (39). In adult oysters we observed low basal expression, as detected in the mussel samples. In fact, none of the experimental conditions reported for the analyzed RNA-seq samples influenced substantially the expression of the core miRNA pathway genes (expression levels below 2 % of El1α), with the exception of the high levels of CgPIWI-1 levels in male and female gonads (around 3.5 %, Figure 5). Conversely, most of the miRNA biogenesis genes were expressed at remarkable levels during the early stages of the oyster development: mainly from two cells to the rotary movement and, for some genes, also in the next developmental stages until *D-shaped* larvae, with no detectable signals afterward in spat and juveniles. Hence, these genes are particularly active in the early development, in particular one AGO (CGI_10020511) and two PIWI transcripts from the egg to trocophora (Figure 5). In the same developmental stages we also noticed a remarkable expression of the key miRNA genes whereas the co-expression of DROSHA and DGCR8 was evident in all the analyzed samples.

Discussion

Small RNAs are important regulators of the gene expression, as recognized in various model and non-model organisms (Kim Myung-Chul et al. 2014; Martini et al. 2014; Hussain and Asgari 2014; Sahoo et al. 2014; Britton et al. 2014; Poole et al. 2014; Solofoharivelo et al. 2014), including some bivalves (Jiao et al. 2014; Zhou et al. 2014). In addition to the identification of miRNAs, a general comprehension of the miRNA biogenesis in itself is also significant (Grimson et al. 2008; Wu et al. 2011; Moran et al. 2013). However, the main genes involved in miRNA biogenesis in bivalves have not been described and characterized so far. In this study, we have provided an overview on the miRNA biogenesis complements in bivalves *spp*., with particular attention to *M. galloprovincialis* and *C. gigas*. To the best of our knowledge, we report for the first time the presence of a complete miRNA biogenesis pathway in *M. galloprovincialis,* the full-length transcript sequences of DICER, DGCR8, XPO5, RAN, DROSHA, TARBP2, three Argonaute genes and the identification of many other components that are candidate miRNA complement-interacting proteins such as MgGW182. By using local transcriptome assemblies, we identified these genes also in many other marine bivalves. The general low expression levels of these transcripts in the adult tissues of both *M. galloprovincialis* and *C. gigas,* and the considerable gene size, have probably prevented a previous identification of full-length sequences in not-well-covered bivalve transcriptomes. In fact, we obtained complete transcript sequences only from sequenced genomes or highly-covered transcriptomes whereas in other transcriptome assemblies we retrieved only few complete sequences. Overall, we have analyzed 508 miRNA complement sequences, 145 of them belonging to marine mollusks and displaying a consistent sequence clustering (*Ostreoida* and *Mytiloida* proteins generated two distinct clades, located always as sister group of arthropods).

However, the copy number of Argonaute genes somewhat differs among bivalves, as *C. gigas* and *A. californica* genomes coding for four proteins (2 AGO and 2 PIWI proteins) whereas *M. galloprovincialis* and *L. gigantea* possess three proteins (1 AGO and 2 PIWIs). We also highlighted the over-expression of the miRNA biogenesis complements during the first phases of the oyster development. A genome protection mechanism based on piRNA expression during early developmental stages is well known in mammals (Malone and Hannon 2009; Kim et al. 2014) but such mechanism has not been reported in bivalves and additional investigations are necessary.

Finally, the identification of several mussel proteins either necessary or cooperative in the miRNA biogenesis, supports the existence of a complete and functional miRNA pathway in mussels and, probably, in other bivalves. Up to now, protein-protein or protein-RNA interaction data are not available for bivalve *spp*. and these topics may represent a direction of work in the future. Meanwhile, the expression analyses of miRNA biogenesis genes coupled with the identification of the miRNAs expressed in naturally infected and laboratory-treated bivalves could provide both validation and new insights on these interesting processes.

Supplementary Materials

Supplementary File 1. Details of the bivalve transcriptome assemblies. Organism ID, project ID (SRA archive), species name and order, origin of the sequenced sample, sequencing platform, millions of reads, number of assembled contig and number of predicted proteins are reported.

Supplementary File 2. Details on the analyzed oyster and mussel RNA-seq samples. Tissue, SRA ID, description and number of reads (in million) are reported.

Supplementary File 3. Graphical view of the mussel proteins interacting during miRNA biogenesis or subsequent interactions. **A**, CNOT1; **B**, CNOT2; **C**, CNOT3; **D**, CNOT6; **E**, CNOT7; **F**, CNOT9; **G**, CNOT10; **H**, DCP1, **I**, DCP2; **J**, PAN2; **K**, PAN3; **L**, eIF4G; **M**, PABP; **N**, PRMT5; **O**, DDX5; **P**, DDX6; **Q**, DDX20; **R**, DDX42; **S**, TUDOR-11 and **T**, MAEL. Domain organization of each protein is shown, green bars represent coiled-coil regions and purple bars disordered regions. Length is expressed as amino acid scale above the figure.

Supplementary File 4. Mussel gene sequences in fasta format.

Supplementary File 5. All considered protein sequences in fasta format.

Supplementary File 6. Digital Expression values of the miRNA biogenesis genes measured by RNA-seq analysis for Mg and Cg.

Figure and table captions

Figure 1. A. Graphical reconstruction of mussel miRNA biogenesis process (modified from (Kapinas and Delany 2011)). **B.** Conserved domains of the mussel miRNA complements.

Figure 2. Mussel gene structures of DROSHA (A), DGCR8 (B), EXP5 (C), DICER (D) and TARBP2 (E) Green boxes represent exons, length is reported as base pair scale.

Figure 3. Phylogenetic relationships of four miRNA biogenesis proteins. **A.** DROSHA, **B.** DGCR8, **C.** DICER and **D.** TARBP2. Inferred protein sequences were aligned using MUSCLE, conserved positions were extracted using Gblocks and subjected to MrBayes analysis.

Figure 4. Phylogenetic relationships of Argonaute-like proteins. Proteins were aligned using MUSCLE and tree was generated using Neighbor Joining algorithm with 1000 bootstrap replicates. Plant proteins are highlighted in green, whereas *C. elegans* hits are reported in grey. Blue lines represent mollusk hits, red lines represent hits from basal metazoans.

Figure 5. Digital expression analysis in oyster. The expression of the 8 miRNA biogenesis genes were computed in tissue-specific RNA libraries and in RNA libraries from different developmental phases. **A.** Expression values represented as percentage of El1α. **B.** Cumulative TPM expression values of the 8 genes in the same samples.

Table 1. Organisms included in the present work. Kingdom, phylum, organism name, sequence origin and reference, ID used in phylogenetic trees and identified sequences are reported. In green, Protostomia; in orange, Deuterostomia. Numbers in bold indicate novel protein sequences.

Table 2. Key proteins of the miRNA biogenesis with their structural domains.

Table 3. miRNA biogenesis proteins of *Mytilus galloprovincialis.* Protein name, GenBank ID, transcript (bp) and protein length (aa), identified domains and annotation (first hit, e-value and percentage of similarity) are reported.

Table 4. Number of exons of five key miRNA biogenesis genes. Metazome 3.0 and Ensembl Metazoa v.29 genome browsers were interrogated with the previously analyzed hits for each organism. La and Ap genomes were downloaded and analyzed locally. Mg gene structures were retrieved as described in Methods. In green are reported Protostomia; in orange Deuterostomia.

References

Ambros, Victor. 2003. MicroRNA Pathways in Flies and Worms: Growth, Death, Fat, Stress, and Timing. *Cell* 113 (6): 673–76. doi:10.1016/S0092-8674(03)00428-8.

Bartel, David P. 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* 116 (2): 281–97. doi:10.1016/S0092-8674(04)00045-5.

Berezikov, Eugene, Victor Guryev, José van de Belt, Erno Wienholds, Ronald H. A. Plasterk, and Edwin Cuppen. 2005. Phylogenetic Shadowing and Computational Identification of Human microRNA Genes. *Cell* 120 (1): 21–24. doi:10.1016/j.cell.2004.12.031.

Britton, Collette, Alan D. Winter, Victoria Gillan, and Eileen Devaney. 2014. microRNAs of Parasitic Helminths - Identification, Characterization and Potential as Drug Targets. *International Journal for Parasitology: Drugs and Drug Resistance* 4 (2): 85–94. doi:10.1016/j.ijpddr.2014.03.001.

Carmell, Michelle A., Angélique Girard, Henk J. G. van de Kant, Deborah Bourc'his, Timothy H. Bestor, Dirk G. de Rooij, and Gregory J. Hannon. 2007. MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. *Developmental Cell* 12 (4): 503–14. doi:10.1016/j.devcel.2007.03.001.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, et al. 2005. The Transcriptional Landscape of the Mammalian Genome. *Science* 309 (5740): 1559–63. doi:10.1126/science.1112014.

Castresana, J. 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution* 17 (4): 540–52.

Chen, Guofu, Chunyun Zhang, Fengjuan Jiang, Yuanyuan Wang, Zhong Xu, and Chongming Wang. 2014. Bioinformatics Analysis of Hemocyte miRNAs of Scallop Chlamys Farreri against Acute Viral Necrobiotic Virus (AVNV). *Fish & Shellfish Immunology* 37 (1): 75–86. doi:10.1016/j.fsi.2014.01.002.

Denli, Ahmet M., Bastiaan B. J. Tops, Ronald H. A. Plasterk, René F. Ketting, and Gregory J. Hannon. 2004. Processing of Primary microRNAs by the Microprocessor Complex. *Nature* 432 (7014): 231–35. doi:10.1038/nature03049.

Diebel, Kevin W., Lauren M. Oko, Eva M. Medina, Brian F. Niemeyer, Cody J. Warren, David J. Claypool, Scott A. Tibbetts, Carlyne D. Cool, Eric T. Clambey, and Linda F. van Dyk. 2015. Gammaherpesvirus Small Noncoding RNAs Are Bifunctional Elements That Regulate Infection and Contribute to Virulence in Vivo. *mBio* 6 (1): e01670–01614. doi:10.1128/mBio.01670-14.

Eddy, Sean R. 2011. Accelerated Profile HMM Searches. *PLoS Comput Biol* 7 (10): e1002195. doi:10.1371/journal.pcbi.1002195.

Edgar, Robert C. 2004. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Research* 32 (5): 1792–97. doi:10.1093/nar/gkh340.

Ender, Christine, and Gunter Meister. 2010. Argonaute Proteins at a Glance. *Journal of Cell Science* 123 (11): 1819–23. doi:10.1242/jcs.055210.

Fabian, Marc R., and Nahum Sonenberg. 2012. The Mechanics of miRNA-Mediated Gene Silencing: A Look under the Hood of miRISC. *Nature Structural & Molecular Biology* 19 (6): 586–93. doi:10.1038/nsmb.2296.

Gammon, Don B., and Craig C. Mello. 2015. RNA Interference-Mediated Antiviral Defense in Insects. *Current Opinion in Insect Science* 8 (April): 111–20. doi:10.1016/j.cois.2015.01.006.

Gao, Zeqian, Miao Wang, David Blair, Yadong Zheng, and Yongxi Dou. 2014. Phylogenetic Analysis of the Endoribonuclease Dicer Family. *PLoS ONE* 9 (4): e95350. doi:10.1371/journal.pone.0095350.

Gerdol, Marco, Gianluca De Moro, Chiara Manfrin, Anna Milandri, Elena Riccardi, Alfred Beran, Paola Venier, and Alberto Pallavicini. 2014. RNA Sequencing and de Novo Assembly of the Digestive Gland Transcriptome in Mytilus Galloprovincialis Fed with Toxinogenic and Non-Toxic Strains of Alexandrium Minutum. *BMC Research Notes* 7 (1). doi:10.1186/1756-0500-7-722.

Ghildiyal, Megha, and Phillip D. Zamore. 2009. Small Silencing RNAs: An Expanding Universe. *Nature Reviews Genetics* 10 (2): 94–108. doi:10.1038/nrg2504.

GIGA Community of Scientists, Heather Bracken-Grissom, Allen G. Collins, Timothy Collins, Keith Crandall, Daniel Distel, Casey Dunn, et al. 2014. The Global Invertebrate Genomics Alliance (GIGA): Developing Community Resources to Study Diverse Invertebrate Genomes. *The Journal of Heredity* 105 (1): 1–18. doi:10.1093/jhered/est084.

Grabherr, Manfred G., Brian J. Haas, Moran Yassour, Joshua Z. Levin, Dawn A. Thompson, Ido Amit, Xian Adiconis, et al. 2011. Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. *Nature Biotechnology* 29 (7): 644–52. doi:10.1038/nbt.1883.

Grimson, Andrew, Mansi Srivastava, Bryony Fahey, Ben J. Woodcroft, H. Rosaria Chiang, Nicole King, Bernard M. Degnan, Daniel S. Rokhsar, and David P. Bartel. 2008. Early Origins and Evolution of microRNAs and Piwi-Interacting RNAs in Animals. *Nature* 455 (7217): 1193–97. doi:10.1038/nature07415.

Hong, Sungguan, Hyangsoon Noh, Haoming Chen, Ravi Padia, Zhixing K. Pan, Shi-Bing Su, Qing Jing, Han-Fei Ding, and Shuang Huang. 2013. Signaling by p38 MAPK Stimulates Nuclear Localization of the Microprocessor Component p68 for Processing of Selected Primary MicroRNAs. *Science Signaling* 6 (266). doi:10.1126/scisignal.2003706.

Hoogstrate, Suzanne W, Rita JM Volkers, Mark G Sterken, Jan E Kammenga, and L Basten Snoek. 2014. Nematode Endogenous Small RNA Pathways. *Worm* 3 (March). doi:10.4161/worm.28234.

Huang, Jing-Tao, Jin Wang, Vibhuti Srivastava, Subrata Sen, and Song-Mei Liu. 2014. MicroRNA Machinery Genes as Novel Biomarkers for Cancer. *Frontiers in Oncology* 4 (May). doi:10.3389/fonc.2014.00113.

Huarte, Maite. 2015. The Emerging Role of lncRNAs in Cancer. *Nature Medicine* 21 (11): 1253–61. doi:10.1038/nm.3981.

Hussain, Mazhar, and Sassan Asgari. 2014. MicroRNAs as Mediators of Insect Host–pathogen Interactions and Immunity. *Journal of Insect Physiology*. Accessed September 16. doi:10.1016/j.jinsphys.2014.08.003.

Iwasaki, Yuka W., Mikiko C. Siomi, and Haruhiko Siomi. 2015. PIWI-Interacting RNA: Its Biogenesis and Functions. *Annual Review of Biochemistry*, March. doi:10.1146/annurev-biochem-060614-034258.

Jean-Philippe, Jacques, Sean Paz, and Massimo Caputi. 2013. hnRNP A1: The Swiss Army Knife of Gene Expression. *International Journal of Molecular Sciences* 14 (9): 18999–24. doi:10.3390/ijms140918999.

Jiao, Yu, Zhe Zheng, Xiaodong Du, Qingheng Wang, Ronglian Huang, Yuewen Deng, Shangli Shi, and Xiaoxia Zhao. 2014. Identification and Characterization of MicroRNAs in Pearl Oyster Pinctada Martensii by Solexa Deep Sequencing. *Marine Biotechnology* 16 (1): 54–62. doi:10.1007/s10126-013-9528-x.

John, Bino, Anton J Enright, Alexei Aravin, Thomas Tuschl, Chris Sander, and Debora S Marks. 2004. Human MicroRNA Targets. *PLoS Biol* 2 (11): e363. doi:10.1371/journal.pbio.0020363.

Kapinas, Kristina, and Anne M Delany. 2011. MicroRNA Biogenesis and Regulation of Bone Remodeling. *Arthritis Research & Therapy* 13 (3): 220. doi:10.1186/ar3325.

Kenny, Nathan J., Erica K. O. Namigai, Ferdinand Marlétaz, Jerome H. L. Hui, and Sebastian M. Shimeld. 2015. Draft Genome Assemblies and Predicted microRNA Complements of the Intertidal Lophotrochozoans Patella Vulgata (Mollusca, Patellogastropoda) and Spirobranchus (Pomatoceros) Lamarcki (Annelida, Serpulida). *Marine Genomics*, August. doi:10.1016/j.margen.2015.07.004.

Kim, Myung-Chul, Seung-Woo Lee, Doug-Young Ryu, Feng-Ji Cui, Jong Bhak, and Yongbaek Kim. 2014. Identification and Characterization of MicroRNAs in Normal Equine Tissues by Next Generation Sequencing. *PLoS ONE* 9 (4): e93662. doi:10.1371/journal.pone.0093662.

Kim, Shinseog, Ufuk Günesdogan, Jan J. Zylicz, Jamie A. Hackett, Delphine Cougot, Siqin Bao, Caroline Lee, et al. 2014. PRMT5 Protects Genomic Integrity during Global DNA Demethylation in Primordial Germ Cells and Preimplantation Embryos. *Molecular Cell* 56 (4): 564–79. doi:10.1016/j.molcel.2014.10.003.

Kincaid, Rodney P., and Christopher S. Sullivan. 2012. Virus-Encoded microRNAs: An Overview and a Look to the Future. *PLoS Pathog* 8 (12): e1003018. doi:10.1371/journal.ppat.1003018.

Kozlov, Guennadi, Nozhat Safaee, Angelika Rosenauer, and Kalle Gehring. 2010. Structural Basis of Binding of P-Body-Associated Proteins GW182 and Ataxin-2 by the Mlle Domain of poly(A)-Binding Protein. *The Journal of Biological Chemistry* 285 (18): 13599–606. doi:10.1074/jbc.M109.089540.

Lau, Pick-Wei, and Ian J. MacRae. 2009. The Molecular Machines That Mediate microRNA Maturation. *Journal of Cellular and Molecular Medicine* 13 (1): 54–60. doi:10.1111/j.1582-4934.2008.00520.x.

Letunic, Ivica, Tobias Doerks, and Peer Bork. 2012. SMART 7: Recent Updates to the Protein Domain Annotation Resource. *Nucleic Acids Research* 40 (D1): D302–5. doi:10.1093/nar/gkr931.

Lucas, Keira, and Alexander S. Raikhel. 2013. Insect microRNAs: Biogenesis, Expression Profiling and Biological Functions. *Insect Biochemistry and Molecular Biology* 43 (1): 24–38. doi:10.1016/j.ibmb.2012.10.009.

MacRae, Ian J., Enbo Ma, Min Zhou, Carol V. Robinson, and Jennifer A. Doudna. 2008. In Vitro Reconstitution of the Human RISC-Loading Complex. *Proceedings of the National Academy of Sciences of the United States of America* 105 (2): 512–17. doi:10.1073/pnas.0710869105.

Malone, Colin D., and Gregory J. Hannon. 2009. Small RNAs as Guardians of the Genome. *Cell* 136 (4): 656–68. doi:10.1016/j.cell.2009.01.045.

Martín-Gómez, Laura, Antonio Villalba, Ron H. Kerkhoven, and Elvira Abollo. 2014. Role of microRNAs in the Immunity Process of the Flat Oyster Ostrea Edulis against Bonamiosis. *Infection, Genetics and Evolution* 27 (October): 40–50. doi:10.1016/j.meegid.2014.06.026.

Martini, Paolo, Gabriele Sales, Mattia Brugiolo, Alessandro Gandaglia, Filippo Naso, Cristiano De Pittà, Michele Spina, et al. 2014. Tissue-Specific Expression and Regulatory Networks of Pig MicroRNAome. *PLoS ONE* 9 (4): e89755. doi:10.1371/journal.pone.0089755.

Maxwell, Evan K., Joseph F. Ryan, Christine E. Schnitzler, William E. Browne, and Andreas D. Baxevanis. 2012. MicroRNAs and Essential Components of the microRNA Processing Machinery Are Not Encoded in the Genome of the Ctenophore Mnemiopsis Leidyi. *BMC Genomics* 13: 714. doi:10.1186/1471-2164-13-714.

Mendell, Joshua T., and Eric N. Olson. 2012. MicroRNAs in Stress Signaling and Human Disease. *Cell* 148 (6): 1172–87. doi:10.1016/j.cell.2012.02.005.

Min, Pil-Ki, and Stephen Y. Chan. 2015. The Biology of Circulating MicroRNAs in Cardiovascular Disease. *European Journal of Clinical Investigation*, June, n/a – n/a. doi:10.1111/eci.12475.

Miyoshi, Keita, Tomoko N. Okada, Haruhiko Siomi, and Mikiko C. Siomi. 2009. Characterization of the miRNA-RISC Loading Complex and miRNA-RISC Formed in the Drosophila miRNA Pathway. *RNA* 15 (7): 1282–91. doi:10.1261/rna.1541209.

Moran, Yehu, Daniela Praher, David Fredman, and Ulrich Technau. 2013. The Evolution of microRNA Pathway Protein Components in Cnidaria. *Molecular Biology and Evolution* 30 (12): 2541–52. doi:10.1093/molbev/mst159.

Moreira, Rebeca, Patricia Pereiro, Carlos Canchaya, David Posada, Antonio Figueras, and Beatriz Novoa. 2015. RNA-Seq in Mytilus Galloprovincialis: Comparative Transcriptomics and Expression Profiles among Different Tissues. *BMC Genomics* 16 (1): 728. doi:10.1186/s12864-015-1817-5.

Mortimer, Stefanie A., Mary Anne Kidwell, and Jennifer A. Doudna. 2014. Insights into RNA Structure and Function from Genome-Wide Studies. *Nature Reviews. Genetics* 15 (7): 469–79. doi:10.1038/nrg3681.

Mukherjee, Krishanu, Henry Campos, and Bryan Kolaczkowski. 2013. Evolution of Animal and Plant Dicers: Early Parallel Duplications and Recurrent Adaptation of Antiviral RNA Binding in Plants. *Molecular Biology and Evolution* 30 (3): 627–41. doi:10.1093/molbev/mss263.

Nguyen, Thuy T. T., Ben J. Hayes, and Brett A. Ingram. 2014. Genetic Parameters and Response to Selection in Blue Mussel (Mytilus Galloprovincialis) Using a SNP-Based Pedigree. *Aquaculture* 420–421 (January): 295–301. doi:10.1016/j.aquaculture.2013.11.021.

Nottrott, Stephanie, Martin J. Simard, and Joel D. Richter. 2006. Human Let-7a miRNA Blocks Protein Production on Actively Translating Polyribosomes. *Nature Structural & Molecular Biology* 13 (12): 1108–14. doi:10.1038/nsmb1173.

Piao, Xianghua, Xue Zhang, Ligang Wu, and Joel G. Belasco. 2010. CCR4-NOT Deadenylates mRNA Associated with RNA-Induced Silencing Complexes in Human Cells. *Molecular and Cellular Biology* 30 (6): 1486–94. doi:10.1128/MCB.01481-09.

Poole, Catherine B., Weifeng Gu, Sanjay Kumar, Jingmin Jin, Paul J. Davis, David Bauche, and Larry A. McReynolds. 2014. Diversity and Expression of MicroRNAs in the Filarial Parasite, Brugia Malayi. *PLoS ONE* 9 (5). doi:10.1371/journal.pone.0096498.

Qu, Shibin, Xisheng Yang, Xiaolei Li, Jianlin Wang, Yuan Gao, Runze Shang, Wei Sun, Kefeng Dou, and Haimin Li. 2015. Circular RNA: A New Star of Noncoding RNAs. *Cancer Letters*, June. doi:10.1016/j.canlet.2015.06.003.

Resch, Alissa M., and Dasaradhi Palakodeti. 2012. Small RNA Pathways in Schmidtea Mediterranea. *The International Journal of Developmental Biology* 56 (1-3): 67–74. doi:10.1387/ijdb.113436ar.

Rhee, S., H. Chae, and S. Kim. 2015. PlantMirnaT: miRNA and mRNA Integrated Analysis Fully Utilizing Characteristics of Plant Sequencing Data. *Methods (San Diego, Calif.)*, April. doi:10.1016/j.ymeth.2015.04.003.

Ronquist, Fredrik, Maxim Teslenko, Paul van der Mark, Daniel L. Ayres, Aaron Darling, Sebastian Höhna, Bret Larget, Liang Liu, Marc A. Suchard, and John P. Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Systematic Biology* 61 (3): 539–42. doi:10.1093/sysbio/sys029.

Ruan, Xiangbo. 2015. Long Noncoding RNA Central of Glucose Homeostasis. *Journal of Cellular Biochemistry*, November. doi:10.1002/jcb.25427.

Sahoo, Ganesh, Md. Ansari, Manas Dikhit, Neha Gupta, Sindhuprava Rana, and Pradeep Das. 2014. Computational Identification of microRNA-like Elements in Leishmania Major. *MicroRNA* 2 (3): 225–30. doi:10.2174/2211536602666131203232422.

Sand, Michael, Marina Skrygan, Dimitrios Georgas, Daniel Sand, Thilo Gambichler, Peter Altmeyer, and Falk G. Bechara. 2012. The miRNA Machinery in Primary Cutaneous Malignant Melanoma, Cutaneous

Malignant Melanoma Metastases and Benign Melanocytic Nevi. *Cell and Tissue Research* 350 (1): 119–26. doi:10.1007/s00441-012-1446-0.

Siomi, Mikiko C., Kaoru Sato, Dubravka Pezic, and Alexei A. Aravin. 2011. PIWI-Interacting Small RNAs: The Vanguard of Genome Defence. *Nature Reviews Molecular Cell Biology* 12 (4): 246–58. doi:10.1038/nrm3089.

Sokolova, O. A., E. Iu Iakushev, A. D. Stoliarenko, E. A. Mikhaleva, V. A. Gvozdev, and M. S. Klenov. 2011. [The interplay of transposon silencing genes in the Drosophila melanogaster germline]. *Molekuliarnaia Biologiia* 45 (4): 633–41.

Solofoharivelo, M. C., A. P. van der Walt, D. Stephan, J. T. Burger, and S. L. Murray. 2014. MicroRNAs in Fruit Trees: Discovery, Diversity and Future Research Directions. *Plant Biology (Stuttgart, Germany)* 16 (5): 856–65. doi:10.1111/plb.12153.

Swarts, Daan C., Kira Makarova, Yanli Wang, Kotaro Nakanishi, René F. Ketting, Eugene V. Koonin, Dinshaw J. Patel, and John van der Oost. 2014. The Evolutionary Journey of Argonaute Proteins. *Nature Structural & Molecular Biology* 21 (9): 743–53. doi:10.1038/nsmb.2879.

Tarver, James E., Philip C. J. Donoghue, and Kevin J. Peterson. 2012. Do miRNAs Have a Deep Evolutionary History? *BioEssays* 34 (10): 857–66. doi:10.1002/bies.201200055.

Théron, Emmanuelle, Cynthia Dennis, Emilie Brasset, and Chantal Vaury. 2014. Distinct Features of the piRNA Pathway in Somatic and Germ Cells: From piRNA Cluster Transcription to piRNA Processing and Amplification. *Mobile DNA* 5 (1): 28. doi:10.1186/s13100-014-0028-y.

Tolia, Niraj H., and Leemor Joshua-Tor. 2007. Slicer and the Argonautes. *Nature Chemical Biology* 3 (1): 36–43. doi:10.1038/nchembio848.

Tomoyasu, Yoshinori, Sherry C. Miller, Shuichiro Tomita, Michael Schoppmeier, Daniela Grossmann, and Gregor Bucher. 2008. Exploring Systemic RNA Interference in Insects: A Genome-Wide Survey for RNAi Genes in Tribolium. *Genome Biology* 9 (1): R10. doi:10.1186/gb-2008-9-1-r10.

van Kouwenhove, Marieke, Martijn Kedde, and Reuven Agami. 2011. MicroRNA Regulation by RNA-Binding Proteins and Its Implications for Cancer. *Nature Reviews Cancer* 11 (9): 644–56. doi:10.1038/nrc3107.

Venier, Paola, Cristiano De Pittà, Filippo Bernante, Laura Varotto, Barbara De Nardi, Giuseppe Bovo, Philippe Roch, et al. 2009. MytiBase: A Knowledgebase of Mussel (M. Galloprovincialis) Transcribed Sequences. *BMC Genomics* 10 (February): 72. doi:10.1186/1471-2164-10-72.

Wagner, Günter P., Koryu Kin, and Vincent J. Lynch. 2013. A Model Based Criterion for Gene Expression Calls Using RNA-Seq Data. *Theory in Biosciences = Theorie in Den Biowissenschaften* 132 (3): 159–64. doi:10.1007/s12064-013-0178-3.

Wu, Baojun, Meng Jin, Jing Gong, Xiumin Du, and Zengliang Bai. 2011. Dynamic Evolution of CIKS (TRAF3IP2/Act1) in Metazoans. *Developmental & Comparative Immunology* 35 (11): 1186–92. doi:10.1016/j.dci.2011.03.027.

Xu, Chengshan, Ling Zhang, Hengheng Li, Zhihua Liu, Lianning Duan, and Chengrong Lu. 2015. MiRNA-1469 Promotes Lung Cancer Cells Apoptosis through Targeting STAT5a. *American Journal of Cancer Research* 5 (3): 1180–89.

Xu, Min-Jun, Shen-Ben Qiu, Alasdair J. Nisbet, Jing-Hua Fu, Chang-Chun Shao, and Xing-Quan Zhu. 2014. Global Characterization of microRNAs in Trichomonas Gallinae. *Parasites & Vectors* 7: 99. doi:10.1186/1756-3305-7-99.

Zhang, Guofan, Xiaodong Fang, Ximing Guo, Li Li, Ruibang Luo, Fei Xu, Pengcheng Yang, et al. 2012. The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation. *Nature* 490 (7418): 49–54. doi:10.1038/nature11413.

Zhou, Zhi, Lingling Wang, Linsheng Song, Rui Liu, Huan Zhang, Mengmeng Huang, and Hao Chen. 2014. The Identification and Characteristics of Immune-Related MicroRNAs in Haemocytes of Oyster Crassostrea Gigas. *PLoS ONE* 9 (2): e88397. doi:10.1371/journal.pone.0088397.