Unravelling the developmental and functional significance of an ancient Argonaute duplication

Arie Fridrich[1,3], Vengamanaidu Modepalli[1,2,3], Magda Lewandowska[1], Reuven Aharoni[1] and Yehu Moran[1,*]

[1]Department of Ecology, Evolution and Behavior, Alexander Silberman Institute of Life Sciences, Faculty of Science, The Hebrew University of Jerusalem, Jerusalem, 9190401 Jerusalem, Israel.

[2]The Marine Biological Association of the United Kingdom, Citadel Hill, Plymouth, United Kingdom.

[3]These authors contributed equally to this work.

*corresponding author: yehu.moran@mail.huji.ac.il

Abstract

microRNAs (miRNAs), base-pair to messenger RNA targets and guide Argonaute proteins to mediate their silencing. This target regulation is considered crucial for animal physiology and development. However, this notion is based exclusively on studies in bilaterians, which comprise almost all lab model animals. To fill this glaring phylogenetic gap, we characterized the functions of two Argonaute paralogs in the sea anemone *Nematostella vectensis* of the phylum Cnidaria, which is separated from bilaterians by ~600 million years. Using genetic manipulation, Argonaute-immunoprecipitations and high-throughput sequencing we provide experimental evidence for the developmental importance of miRNAs in a non-bilaterian animal. Additionally, we uncover unexpected differential distribution of distinct miRNAs between the two Argonautes and the ability of one of them to load additional types of small RNAs. This enables us to postulate a novel model for evolution of miRNA precursors in sea anemones and their relatives, revealing alternative trajectories for metazoan miRNA evolution.

Introduction

microRNAs (miRNAs) are pivotal players in post-transcriptional gene regulation in animals and plants. As such, they carry important roles in many developmental and physiological processes in both kingdoms[1-4]. Whereas their mode of action and functional roles are extensively-studied in bilaterian animals such as vertebrates, nematodes and arthropods, they are vastly understudied in non-bilaterian animals such as cnidarians and sponges despite being identified in these animal phyla more than a decade ago[5]. Markedly, until now, there is only little experimental evidence supporting any developmental or physiological role for miRNAs in organisms other than bilaterian animals and land plants, despite their wide phyletic distribution[6,7]. Our previous studies revealed that the miRNAs of Cnidaria (sea anemones, corals, jellyfish and hydroids), the sister group of Bilateria, exhibit remarkable similarities to plant miRNA biogenesis and mode of action[8-10]. This was unexpected, as many contemporary studies argue that the vast differences between the plant and animal miRNA systems point towards convergent evolution and lack of miRNAs in

the last common ancestor of animals and plants[6,11,12]. Among the similarities found between plant and cnidarian miRNAs, arguably the most striking one is the frequent tendency of the miRNA-guided RNA-induced Silencing Complex (RISC) of cnidarians to cleave ('slice') its targets[9]. The key proteins of RISC that perform the slicing in all miRNA-bearing organisms are the Argonautes (AGOs)[13-15]. A bioinformatic analysis revealed that members of the Hexacorallia cnidarian subclass (sea anemones and reef-building corals) carry two AGOs with intact RNAse catalytic sites[10]. As hexacorallians diverged roughly 540 million years ago (MYA)[16,17], this gene duplication seems to be extremely stable and suggests that each of the two AGOs specialized in a different role[10].

To test the role of miRNAs in Cnidaria and the hypothesis of AGO specialization in Hexacorallia, in our current work we took advantage of genetic-manipulation tools available for the sea anemone *Nematostella vectensis* as well as immunoprecipitation and high-throughput sequencing techniques. Our results provide direct evidence that indeed miRNAs play a significant role in the development of a cnidarian and reveal the specialization of the two AGO proteins of Hexacorallia. Moreover, the results also enable us to postulate a new hypothesis regarding how novel miRNAs are born and further evolve in Hexacorallia.

Results and Discussion

Nematostella development requires two AGO paralogs that duplicated >500 MYA.

To get a broader insight into the evolutionary history and fate of AGO duplications in Metazoa with a focus on Cnidaria we constructed AGO phylogeny in which we included bilaterian as well as cnidarian representatives (**Fig. 1a**). This phylogeny revealed that while the first *Nematostella* AGO, NveAGO1, clusters with AGO1 proteins of other hexacorallians, the second *Nematostella* AGO, NveAGO2, is positioned in a separate clade, together with the hexacorallian AGO2. Interestingly, we could detect at least one AGO1 ortholog and one AGO2 ortholog in most hexacorallian transcriptomes or genomes we surveyed. Our phylogeny strongly suggests that the two AGO paralogs duplicated before the split of sea anemones and reef-building corals and co-existed for more than 500 million years. Such long co-existence of paralogs strongly hints towards sub- or neo- functionalization, otherwise we would expect a frequent loss of one of the paralogs over hundreds of millions of years. Our previous results suggested that destabilization of miRNAs and inhibition of their biogenesis by knocking down the methyltransferase NveHEN1 or the RNAse III NveDicer1, respectively, results in severe developmental defects in *Nematostella* early development[8]. However, these components also play a role in the stabilization and biogenesis of other small RNAs and may be involved in additional pathways[6,8]. Thus, we decided to test the potential specialization of the *Nematostella* AGOs and the role of miRNAs in *Nematostella* development by inhibiting the expression of each AGO separately. To this end, we microinjected Morpholino Antisense Oligos (MOs) against NveAGO1 and NveAGO2 to *Nematostella* zygotes. AGO

morphants, but not zygotes injected with a control MO, failed to reach the primary polyp stage nine days post fertilization (dpf) (**Fig. 1b and 1c**). These morphological phenotypes are grossly similar to the phenotypes that we previously obtained when we inhibited other components involved in the miRNA pathway[8]. The specificity of each of the knockdowns (KD) was assayed by Western blot with specific antibodies raised against NveAGO1 and NveAGO2 that were generated for this study (see methods). The bands that correspond to the expected size of NveAGO1 and NveAGO2 are absent from morphants (**Fig. 1d**). The specificity of the antibodies was further verified by immunoprecipitation (IP) performed on lysates from primary polyps (9-days-old) followed by liquid chromatography coupled to tandem mass spectrometry (LC-MS-MS) (**Fig. 1e, Supplementary table 1**). The LC-MS/MS analysis revealed that in NveAGO1 IP, NveAGO1 peptides (but not NveAGO2) were enriched compared to the control sample incubated with Rabbit IgG. Similarly, NveAGO2 IP samples were enriched for NveAGO2 (but not NveAGO1) peptides (**Fig. 1e**, **Supplementary table 1**). The differences were statistically significant ($P \leq 0.001$) and these results support the specificity of the custom antibodies.

While the failure in metamorphosis of the morphants of each AGO KD is an indication for functional specialization of each paralog[18], we could not exclude the possibility that the AGOs are redundant in their function but their high dosage is critical for development[19]. Thus, we performed RNA-seq of three days old control and morphant animals and looked at the differences in their transcriptomic signatures (**Fig. 1f, Supplementary table 2**). Strikingly, while the controls of all experiments clustered together, NveAGO1 and NveAGO2 morphants greatly differed from one another, pointing towards specialization rather than redundancy and dosage-dependency. Since the main biological function of AGOs in animals is to guide small RNAs (sRNAs) to target other RNA transcripts[20,21], we hypothesized that these transcriptomic differences might stem from different small RNA populations that are carried by these two AGOs, affecting different sets of RNA targets. Notably, hexacorallian AGO1 forms a sister group to the known bilaterian miRNA-AGOs. Contrastingly, the hexacorallian AGO2 is positioned outside these two sister groups (**Fig. 1a**), leading us and others to predict it will not bind miRNAs but another group of sRNAs[22]. However, because phylogeny alone is insufficient to determine the differences between the NveAGO1 and NveAGO2 sRNA cargos, we continued with an experimental approach to test that.

AGO IP uncovers novel *Nematostella* miRNAs

First, we confirmed by western blot that the antibodies αNveAGO1 and αNveAGO2 are suitable for a specific IP in three distinct developmental stages (planula, primary polyp and adult) (**Fig. 2a**). For each developmental stage, we immunoprecipitated each NveAGO in two biological replicates, and included two replicates of rabbit IgG IP as negative controls. Next, we generated sRNA libraries from IP-extracted, size-selected RNA (see methods) using a modified protocol suitable for lower sRNA quantities[23].

mirDeep2[24] was used for de-novo identification and quantification of novel and known *Nematostella* miRNAs. To minimize false positives, we used stringent criteria, which was previously suggested for miRNA annotation in bilaterians[25]. Specifically, a novel bona fide miRNA was considered if it exhibits: (a) >3-fold enrichment in the IP sample compared to the negative controls, (b) a clear signature for strand selection, with a dominant guide strand, that contains a homogeneous 5′ end, (c) guide/star ratio higher than two, (d) presence in both biological replicates with a minimum of 70 reads from the guide strand in at list one replicate, (e) a minimum of 16 hybridized nucleotides in the predicted guide/star duplex (see example in **Fig. 2b**). Since strand selection in *Nematostella* is significantly stronger on average than in bilaterians[5], starless miRNAs that passed the described criteria were also considered, however, classified as a separate category (**Supplementary table 3**). In all libraries, the fraction of reads that mapped to miRNAs was higher than in the negative controls (**Fig. 2c**). Altogether, we identified 86 novel miRNAs (of which 19 were classified as "starless") (**Fig 2b**, **Supplementary table 3**), increasing the total number to 138, with 42 previously identified miRNAs being revoked for not passing the precursor structure and homogeneity criteria, or for not being enriched in AGO IP compared to the IgG IP. Of the 52 previously identified remaining miRNAs, 49 overlapped with the list of Moran et al.[9] and 31 with the list of Grimson et al.[5] (see the updated list of miRNAs in **Supplementary table 3**).

The reason so many miRNAs have not been identified earlier, stems from the fact that in previous studies, sRNAs were collected from whole animals. In *Nematostella*, unlike most bilaterians, such libraries are predominated by another class of sRNAs: the piRNAs[5,9]. Unlike most bilaterians, *Nematostella* piRNAs are not restricted to the germline, but are rather expressed broadly in somatic tissues[26]. Hence NveAGO IP enabled us for the first time to overcome this limitation and reliably identify a significantly larger set of lower expressed novel miRNAs.

NveAGO1 and NveAGO2 exhibit differential miRNA sorting as well as siRNA specialization

With the expanded list of miRNAs, we next continued with the examination of the differences between NveAGO1 and NveAGO2 sRNA cargo. Independent AGO duplications occurred several times in bilaterians and a frequent outcome for such duplications is specialization in carrying either miRNAs or siRNAs (but not both) as seen in insects and nematodes[27,28]. This enables a separation of the endogenous gene regulation pathway from the antiviral one, as target silencing of these two target types requires different modulation levels and kinetics as well as different partner proteins[29].

Our initial expectation was to identify a similar type of specialization in Hexacorallia as it is known that endogenous siRNAs (endo-siRNAs) are present in *Nematostella*[9,30]. Using ShortStack[31,32] we identified endo-siRNAs in our libraries (**Supplementary table 3**) and quantified their normalized read counts in NveAGO1 and NveAGO2 IP from three developmental stages (**Figs. 3a-c**). These results

confirmed our hypothesis as NveAGO2 but not NveAGO1 is enriched with endo-siRNAs throughout development. Surprisingly, when we looked at the miRNA abundance in the two *Nematostella* AGOs, we found both to heavily load miRNAs (**Figs. 3a-c**, lower panels), meaning that NveAGO2, is a carrier of both miRNAs and siRNAs. Non-specialized AGOs carrying both miRNAs and siRNAs are present in at least some mammals[33,34], but this seems to be a rare phenomenon in bilaterians. Some animals possess more than one miRNA-specific AGO. In vertebrates (most of which lack a widely functional siRNA pathway) the specialization of each of the four miRNA-AGOs is not fully understood. They all seem to carry the same miRNAs[35], and seem to somewhat differ in their functions. For example, vertebrate AGO2, but not the others, contains an active catalytic domain that enables to promote slicing of fully complement targets[36] and to enable non-canonical maturation of some miRNAs[37-39]. In *C. elegans*, Alg5 was shown to carry some unique miRNAs that are not carried by the other two miRNA-specific AGOs of this species[40]. In this case, the specificity is enabled via tissue specific co-expression of these particular miRNAs with Alg5 in the germline. In *Nematostella,* the two AGOs are expressed in all tissues and developmental stages[26] (**Supplementary Fig. 1a**, **Supplementary table 4**). Nonetheless, when we tested the relative preference of individual miRNAs in NveAGO1 and NveAGO2 (calculated as explained in the Methods section), we were surprised to find two populations of miRNAs. The first set of miRNAs is preferentially loaded into NveAGO1, while the other is enriched in NveAGO2 (**Fig. 3d-g**). NveAGO1 miRNAs were consistently under-represented in NveAGO2 and vice versa with the exception of two miRNAs: (a) miR-9446 is enriched in NveAGO1 in adult males, and in NveAGO2 in primary polyps (b) miR-2040a is enriched in NveAGO1 in early planulae and primary polyps, and in NveAGO2 in adult males (**Fig. 3d-f, Supplementary table 3**). These exceptions suggest that some developmentally-regulated elements may contribute to the AGO-selectivity. We characterized the differential expression of NveAGO1 and NveAGO2 miRNAs across the three developmental stages and similarly to the previously described *Nematostella* miRNAs[9], novel miRNAs also exhibit differential temporal expression (**Supplementary Fig. 1b**), further supporting the role of the miRNA pathway in *Nematostella* development. Interestingly, some miRNAs exhibited opposite strand selection in the two AGO proteins (**Supplementary Fig. 2**).

To elucidate the effect of AGO-depletion on sRNAs, we generated sRNA libraries from planulae injected with MOs against either NveAGO1 or NveAGO2. Strand selection (ratio between guide to star reads) was significantly disrupted for AGO1 miRNAs in AGO1 KDs but not for AGO2 miRNAs in AGO2 KDs (**Supplementary Fig. 3**, **Supplementary table 5**). These results show that when the carrier of NveAGO1 miRNAs is depleted, there is a relative reduction in miRNA guide levels and accumulation of star reads. When we examined the read counts of guide strands of miRNAs that strongly occupy each AGO (>90% preference), levels of 6 out of 7 analyzed AGO1 miRNAs were reduced in AGO1 KD, however this result is not statistically significant (P=0.06) (**Fig. 3h**, upper panel). AGO2 miRNAs were significantly

downregulated in AGO2 KD (P=0.013) and significantly upregulated in AGO1 KD (P=0.0012) (**Fig. 3h**, lower panel).

These results uncover a previously uncharacterized novel type of AGO subfunctionalization: a wide-scale differential miRNA sorting between the two AGOs, with NveAGO2 serving as a dual-functioning carrier of both miRNAs and endo-siRNAs.

Characterization of NveAGO1 and NveAGO2 miRNAs

The molecular basis for sRNA sorting into AGOs was uncovered in *Drosophila* where it was shown that central mismatches in miRNA duplexes (positions 9-11) enable their loading into the *Drosophila* miRNA-carrier: DmeAGO1, while siRNA duplexes (which usually do not contain these mismatches) are guided into the specialized siRNA carrier: DmeAGO2[41-43]. In *Arabidopsis*, mismatches at positions 9-12 facilitate miRNAs sorting[44] and additionally the identity of the first 5′ nucleotide of the guide strand plays an important role in miRNA sorting[45]. Thus, we generated NveAGO1 and NveAGO2 miRNA sequence signatures for miRNAs with >70% preference levels to a single AGO and found that unlike plants, the identity of the 5′ terminal nucleotide for both *Nematostella* miRNA populations is U (**Fig. 4a**). Next, we aligned the miRNAs of each NveAGO and calculated for each position of the guide-strand the frequency of mismatches between guide and star. For NveAGO1 miRNAs, there is a noticeable tendency to be enriched with mismatches in positions 10-12 compared to NveAGO2 miRNAs (**Fig. 4b**, **Supplementary table 3**). 65% of NveAGO1 miRNAs (26/40) tend to have at least one central mismatch, compared to only 18% of NveAGO2 miRNAs (11/60). This is comparable to *Arabidopsis* where 65.7% of AthAGO1 enriched miRNAs (23/35) contain central mismatches in their duplexes compared to 20% of AthAGO2 miRNAs (3/15)[44]. In *Drosophila* ~80% of surveyed miRNA duplexes (80/102) exhibited central mismatches[46]. This result suggests that the molecular basis for discerning miRNAs from siRNAs in bilaterians and plants is similar to the one of *Nematostella*. However, additionally, in this species the mechanism also enables to segregate two large sets of miRNAs. Notably, NveAGO1 miRNAs are slightly but significantly longer than NveAGO2 miRNAs (average of 22.2 and 21 nt, respectively, P<0.001 Student's *t*-test) (**Supplementary table 3**). miRNAs that present no strong preference (<70%) to a specific AGO paralog exhibit an intermediate average length of 21.5 nt.

In mammals, approximately one third of the miRNA encoding loci are positioned within introns of protein coding genes[47], while most of the rest are encoded from independent intergenic units. In plants intronic miRNA loci are rare[11]. We find that in *Nematostella*, similarly to mammals and unlike in plants, about one third of miRNA precursors are positioned within introns of protein coding genes, and combined with intergenic loci, comprise the vast majority of precursors (**Fig. 4c**). Very few precursors are positioned within the exons of protein coding genes. Interestingly, some miRNA precursors are positioned within

repetitive elements (REs). Many of these precursors (11/20) have predicted targets that contain signatures of the REs that host the miRNA precursor (**Fig. 4d**, **Supplementary table 6**). A similar mechanism was previously described in mammals where some protein coding genes can be regulated by piRNAs, via retrotransposition of repetitive elements to their 3′ UTRs[48]. Our results suggest that in *Nematostella* new miRNA targets can originate from the transposition of REs that host miRNA-precursors into the UTRs of protein coding genes. A deeper look at predicted targets of NveAGO1 and NveAGO2 miRNAs enabled us to identify a surprising signature that might serve as a clue for understanding the evolutionary origin of miRNA precursors in Hexacorallia, as described below.

Origin of miRNA precursors in Hexacorallia

miRNA precursors are believed to emerge and evolve differently in plants and animals (bilaterians) (reviewed by Axtell et al.[11]). In plants, but not in bilaterians, the loci of some miRNA precursors exhibit homology to the target of the miRNA, which extends beyond the recognition site by the miRNA-guide strand. Hence, it was suggested that in plants new miRNA precursors are born from their own targets, when these coding genes go through inverted duplications, which upon transcription can generate hairpin precursors that can be recognized by sRNA biogenesis machinery, and thus generate sRNA that target with high complementarity the gene from which it originated[11,49]. In bilaterians, where miRNA-target recognition is mostly restricted to a short stretch of nucleotides at the 5′ of the miRNA, miRNAs that are highly complementary to their targets expose their unprotected 3′ end which results in the degradation of the miRNA by exonucleases[50]. In plants, where high complementarity of miRNAs to their targets is frequent, the 3′ ends of miRNAs are methylated and thus protected from exonucleolytic activity[51,52].

We postulate, that a plausible explanation for the difference between plant and bilaterian evolutionary origins of miRNA precursors stems from the selection against highly complement miRNA-target interactions in bilaterians. Hence we decided to test whether in *Nematostella* (where high complementarity is frequent[9], and miRNA are stabilized by methylation[8]) miRNA precursors exhibit a signature of origin from their own targets.

To this end, we examined the list of miRNA precursors with extended homology to their predicted targets (see methods). 8 out of 40 NveAGO1 miRNAs and 19 out of 60 NveAGO2 miRNAs exhibited such signatures (**Fig. 5a**, **Supplementary table 6 and Supplementary table 7**). Together with the weakly expressed miRNAs and those that show no preference to a single AGO paralog, 35 miRNAs show this signature. Therefore, we propose that in cnidarians, similarly to plants and unlike bilaterians, miRNA precursors can originate from their own targets. This scenario implies that unlike the conservative view that animal and plant miRNA-precursors originate differently, the origin of miRNAs from targets is an ancestral mechanism that was utilized by plants and early animals, and was later lost in bilaterians. This observation

piles with previous examples for commonalities between plant and cnidarian miRNAs[6,8,9], challenging the traditional view that the last common ancestor of plants and animals lacked a miRNA pathway.

Conclusions and the evolutionary journey of a hexacorallian miRNAs

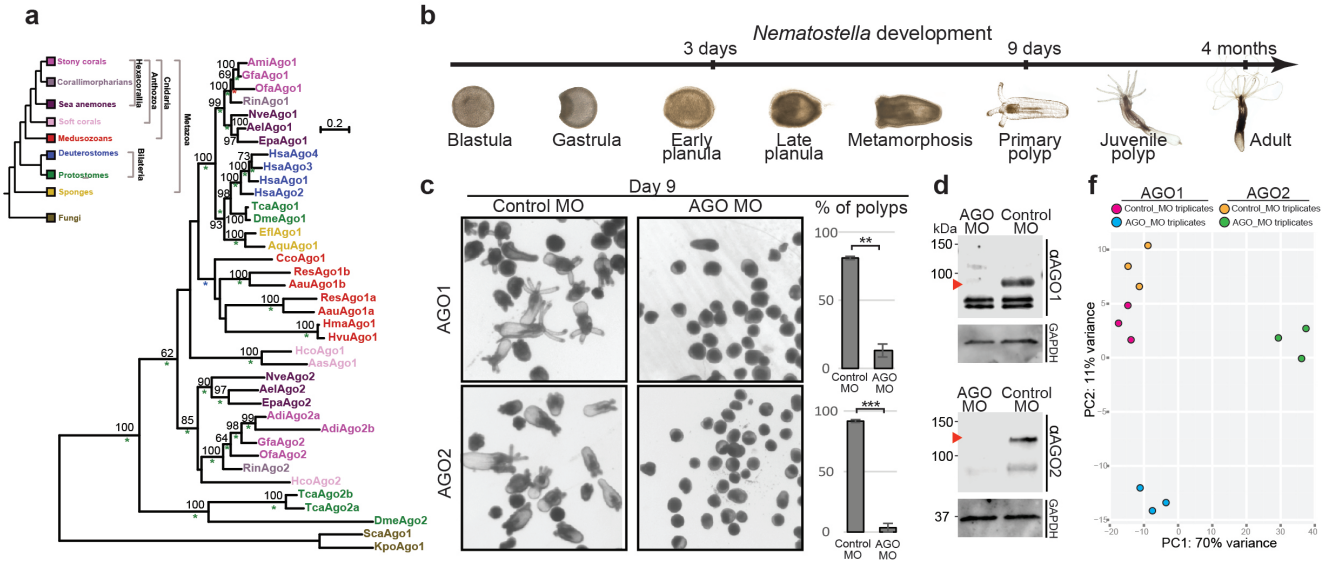
Using IP, knockdowns and high-throughput sequencing we revealed the developmental importance and functional differences of two hexacorallian AGOs that duplicated at least 500 million years ago at the split of sea anemones and reef-building corals. This independent duplication within the sister group of Bilateria emphasizes the recurrent trend of AGO duplications and subfunctionalizations in eukaryotes, reflecting presumably its advantageous value. In most bilaterians, such duplications resulted in specialization to carry either siRNAs or miRNAs. Evolutionary analysis of AGOs from many dipteran species[28] uncovered that once such subfunctionalization is established, the miRNA specialized AGOs do not seem to duplicate any further. On the contrary, siRNA AGOs in these species exhibit frequent duplications as well as longer branch lengths, which might correlate with the constant need to change in the evolutionary arms race against viruses and transposons. The hexacorallian AGO duplication occurred independently of the duplications in bilaterians and yet exhibits a similar case of shorter branches for miRNA AGOs compared to that of the siRNA AGOs (**Fig. 1a**). Although the specialization in siRNAs of NveAGO2 was previously suggested[22] our work is the first to functionally show that NveAGO2 indeed carries this sRNA type. Additionally, and unlike in any previous predictions, this study uncovered a dual function for NveAGO2 as we found that it loads high levels of miRNAs lacking central mismatches (**Fig. 3**). In *Nematostella*, the dual role of NveAGO2 is persistent throughout development (**Supplementary Fig. 1a**) and based on phylogeny (**Fig. 1a**) might be conserved in other sea anemones and reef-building corals.

The broad differential miRNA sorting into NveAGO1 and NveAGO2 is another unique specialization that does not typify the outcomes of AGO subfunctionalization in bilaterians. The same molecular mechanism for sorting miRNAs from siRNAs in bilaterians seems to enable the separation of one set of miRNAs from another in *Nematostella* (**Figs 3 and 4**) and brings us to suggest a new model for the evolution of newly born miRNA precursors in Hexacorallia (**Fig. 5**). Our results indicate that unlike other animals, cnidarian miRNA precursors might emerge via inverted duplications of their own targets, similarly to plants (**Fig. 5b**). Transcription of such inverted duplicated genes can generate hairpins from which highly complementary endo-siRNAs are formed and loaded into NveAGO2 (**Figs. 3a-c**, **Fig. 5b**). Next, such precursors accumulate mismatches that make them resemble miRNA precursors (e.g., 5′ homogeneity[25]), that still load into NveAGO2. Finally, mutations that facilitate NveAGO1 loading (such as central mismatches, **Fig. 4b**) enable these miRNAs to preferentially load into NveAGO1 that does not carry endo-siRNAs.

This model suggests a long evolutionary journey until a miRNA is able to load into NveAGO1. The model is further supported by the fact that the fraction of miRNA targets with extended homology to their miRNA precursor loci is higher for NveAGO2 miRNAs compared to NveAGO1 (**Supplementary table 7**), and finally, by the fact that the majority of the miRNAs conserved among hexacorallians are carried by NveAGO1 (**Fig. 5c**). Thus, we propose that the origin of miRNAs from their targets represents an ancestral phenomenon that is shared between plants and animals such as cnidarians, but was lost in bilaterians with their acquisition of seed-restricted target recognition and modulation[6].

Acknowledgements

The authors are grateful to the following researchers of the core facilities of the Alexander Silberman Institute of Life Sciences, The Hebrew University: Dr. William Breuer (Interdepartmental Unit) for his help with mass spectrometry and Dr. Michal Bronstein and Ms. Adi Turjeman (The Center for Genomic Technologies) for their help with high-throughput sequencing. The authors would like to thank Prof. Liran Carmel (Department of Genetics) for his advice regarding statistical analysis, and Dr. Joachim Surm (Department of Ecology, Evolution and Behavior) for his help with bioinformatic analysis. This work was funded by the European Research Council Starting Grant (CNIDARIAMICRORNA, 637456) to Y.M.



e

	NveAGO1 Avg LFQ	NveAGO1 Avg Unique peptides	NveAGO2 Avg LFQ	NveAGO2 Avg Unique peptides
AGO1 IP	309M|1063M|1315M	36|34|32	0|6M|0	2|2|2
IgG IP	3.91M|0M|2.5M	2|0|2	0|0|0	0|0|1
AGO2 IP	0|0|0	0|0|0	90M|234M|222M	8|8|13
IgG IP	0|0|0	0|0|0	0|0|0	0|0|2

Fig. 1: The hexacorallian AGO duplication, and its importance for *Nematostella* development

a, A phylogenetic relationship of metazoan AGOs. The tree was constructed with the LG model (+I, +G). Bootstrap support values above 50% are indicated above branches. Posterior probability values of a Bayesian tree of the same topology are indicated by asterisks. A green asterisk represents a value of 1.0, blue represents values greater than 0.95 and lower than 1.0 and a value of 0.95 is indicated by a red asterisk. The tree was rooted by the AGOs of the fungi *Saccharomyces castellii* (Sca) and *Kluyveromyces polysporus* (Kpo). Abbreviations of other species names are: Aas, *Acanthogorgia aspera*; Aau, *Aurelia aurita*; Ael, *Anthopleura elegantissima*; Adi, *Acropora digitifera*; Ami, *Acropora millepora*; Aqu, *Amphimedon queenslandica*; Cco, *Craterolophus convolvulus*; Dme, *Drosophila melanogaster*; Efl, *Ephydatia fluviatilis*; Epa, *Exaiptasia pallida*; Gfa, *Galaxea fascicularis*; Hco, *Heliopora coerulea*; Hsa, *Homo sapiens*; Hvu, *Hydra vulgaris*; Nve, *Nematostella vectensis*; Ofa, *Orbicella faveolata*; Res, *Rhopilema esculentum*; Rin, *Rhodactis indosinensis*; Tca, *Tribolium castaneum*. Protein sequences are available in Supplementary file 1. **b**, Timeline of *Nematostella* development. **c**, Animals injected with control MO (left panels) 9 dpf. Animals injected αNveAGO1 and αNveAGO2 MOs (middle upper and lower panels, respectively) 9 dpf. Majority of NveAGO1 and NveAGO2 depleted animals did not reach primary polyp stage (right upper and lower panels, respectively) at 9 dpf, three independent biological replicates. **P=0.0069, ***P=0.0005 (Students *t*-test), (**Supplementary table 5**). **d**, Western blot validation of knockdowns with NveAGO1 and NveAGO2 custom antibodies. **e**, LC-MS/MS analysis on three technical replicates. In NveAGO1 IP, an average of 34 specific peptides corresponding to NveAGO1, yielding an average label-free quantification (LFQ) value of ~895 million, whereas the control sample incubated with

Rabbit IgG rather than αNveAGO1 yielded 1.3 peptides and an average LFQ value of 2.1 million. NveAGO2 IP samples provided an average of 9.6 peptides specific for NveAGO2 and an average LFQ value of ~182 million vs. 0.6 peptides and LFQ of 0 for the IgG control. Green shaded pairs indicate statistical significance (P=0.0003 for NveAGO1, and P=0.001 for NveAGO2, see methods). **f**, Principal component analysis (PCA) plot exhibiting differences in transcriptomic signatures for 3 distinct biological replicates of NveAGO1 and NveAGO2 knockdowns.

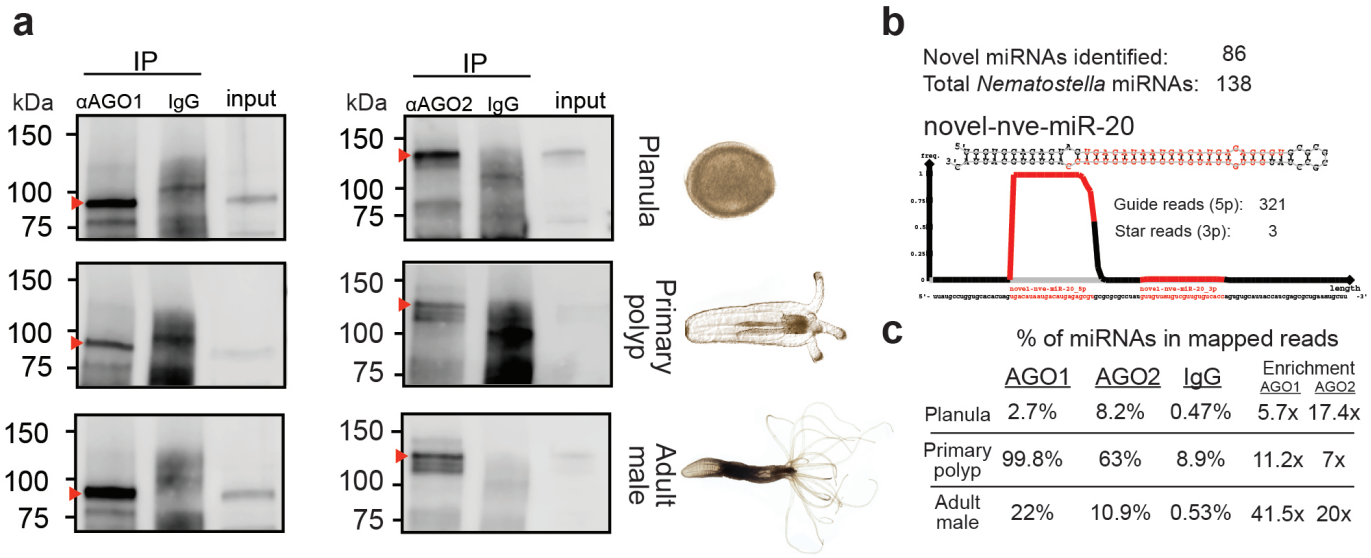


Fig. 2: Novel miRNAs identification by AGO-IP

a, Western blot validations with αNveAGO1 and αNveAGO2 for presence of NveAGO1(left, ~96 kDa) and NveAGO2 (right, ~122 kDa) following the IPs from three distinct developmental stages (bands corresponding to AGOs are indicated by red arrowheads). **b**, Number of *Nematostella* miRNAs including novel miRNAs, and an example of a newly identified miRNA. **c**, Average percentage of miRNA reads in sRNA libraries generated from extracts of NveAGO1, NveAGO2, and control IgG IPs. miRNA percentage was calculated from reads that mapped to the *Nematostella* genome. On the right is the enrichment in each of the NveAGO-IPs compared to control IgG.

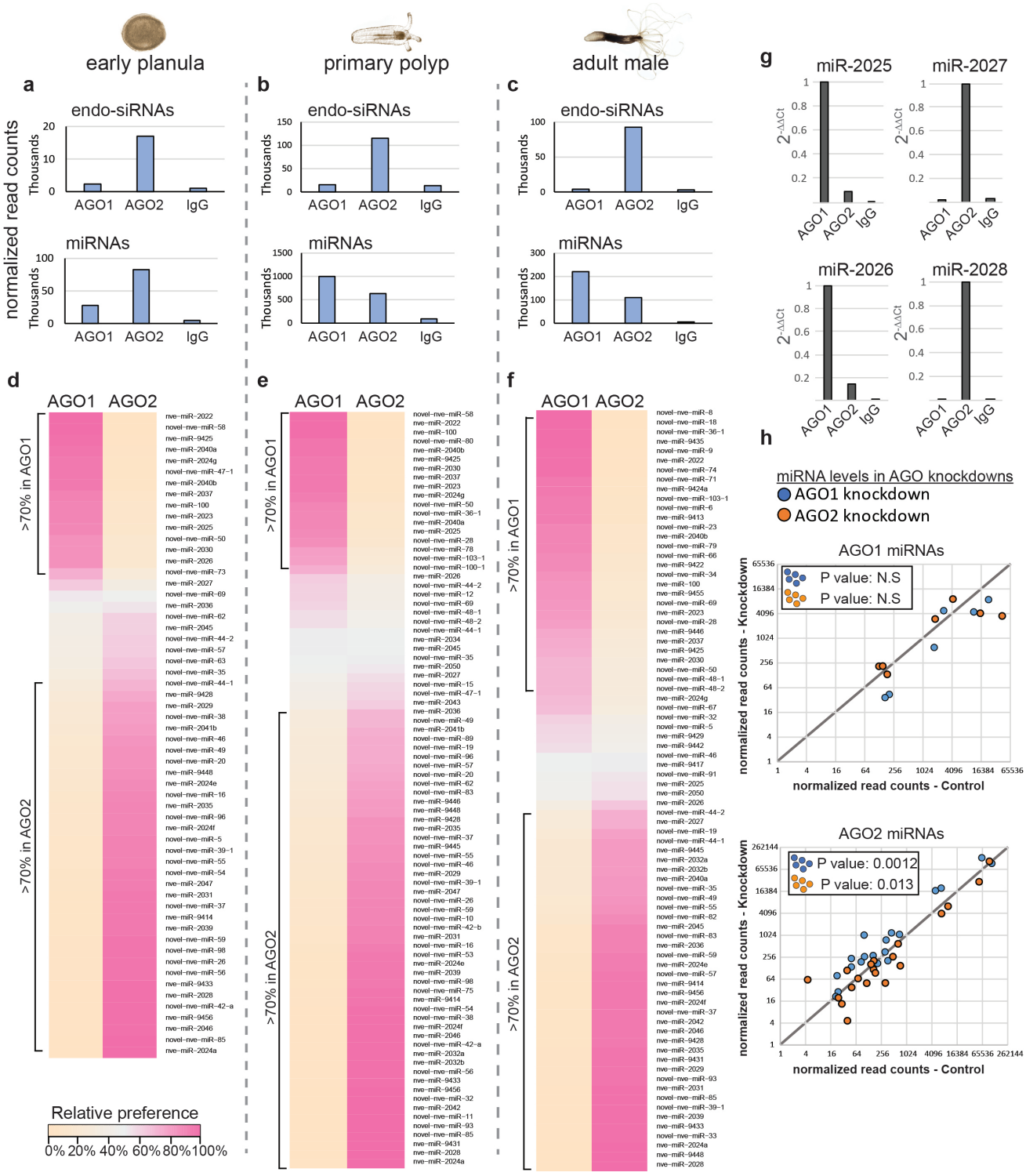


Fig. 3: miRNA and endo-siRNA distribution between the two *Nematostella* AGOs

a-c, Normalized read-counts of miRNAs (lower panels) and endo-siRNAs (upper panels) in sRNA libraries generated from *Nematostella* AGO IPs. Each bar represents the average of two distinct biological replicates normalized as described in the methods section. **d-f**, Heat map representing the relative preference of individual miRNA to NveAGO1 or NveAGO2, validated by qPCR (**g**) for two NveAGO1 and two NveAGO2 miRNAs. **h**, Levels of miRNAs with >90% preference for NveAGO1 (upper panel) and for

NveAGO2 (lower panel) in AGO1 knockdowns (compared to control, blue dots) and AGO2 knockdowns (orange dots). A binomial significance test was used to determine whether miRNA levels change following AGO knockdown (compared to their levels in controls, **Supplementary table 5**).

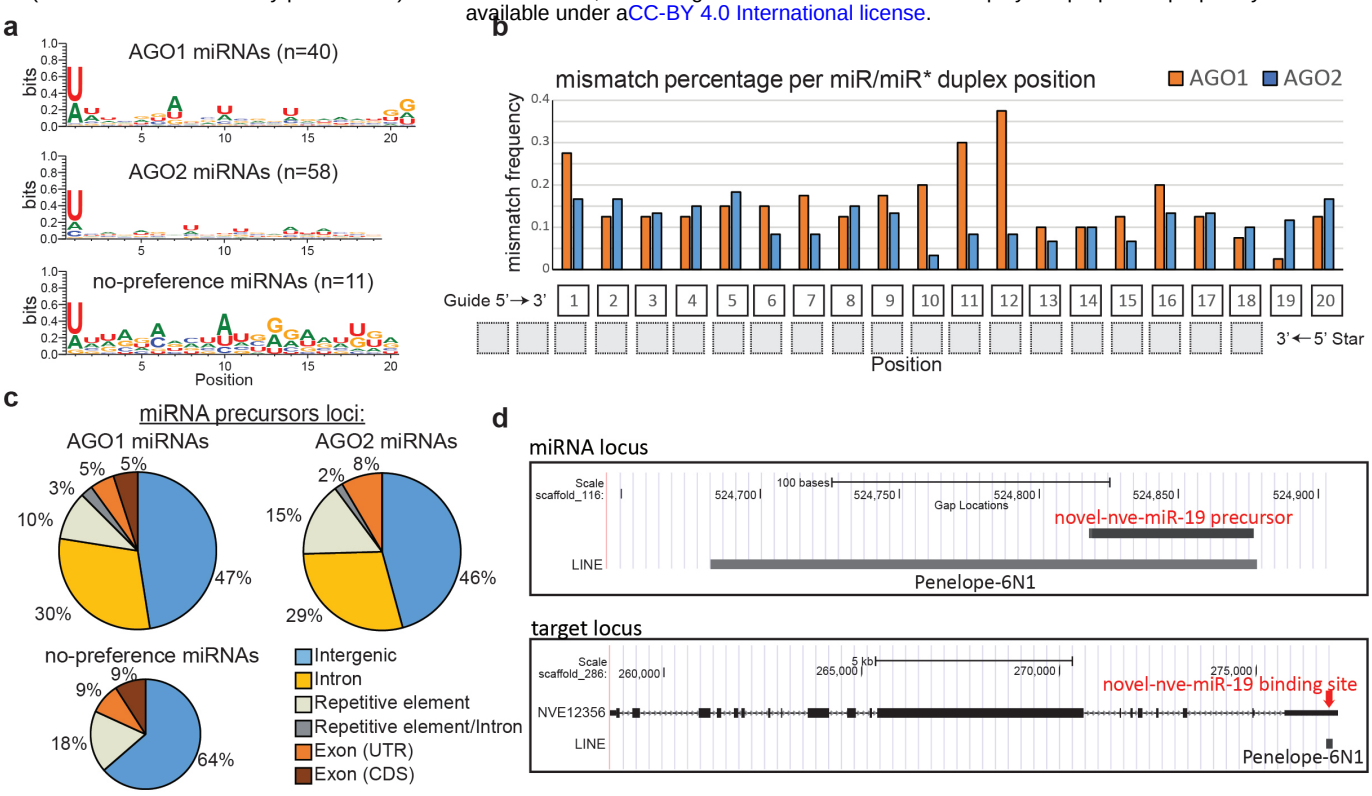


Fig. 4: Characteristics of NveAGO1 and NveAGO2 miRNAs and origin of some miRNA targets

a, Signature of miRNAs sequences that have > 70% preference for NveAGO1, NveAGO2, or no such preference. Sequence logos were generated using WebLogo3[53]. In NveAGO2 miRNAs ≥19nt were analyzed (58/60) **b**, Mismatch frequency for miR-guide's strand positions when it forms a hybridized duplex with the star strand. **c**, Genomic loci of miRNA precursors. **d**, An example of a miRNA precursor positioned within a repetitive element (upper panel), and its putative target's binding site which shows integration signature of the same repetitive element (lower panel).

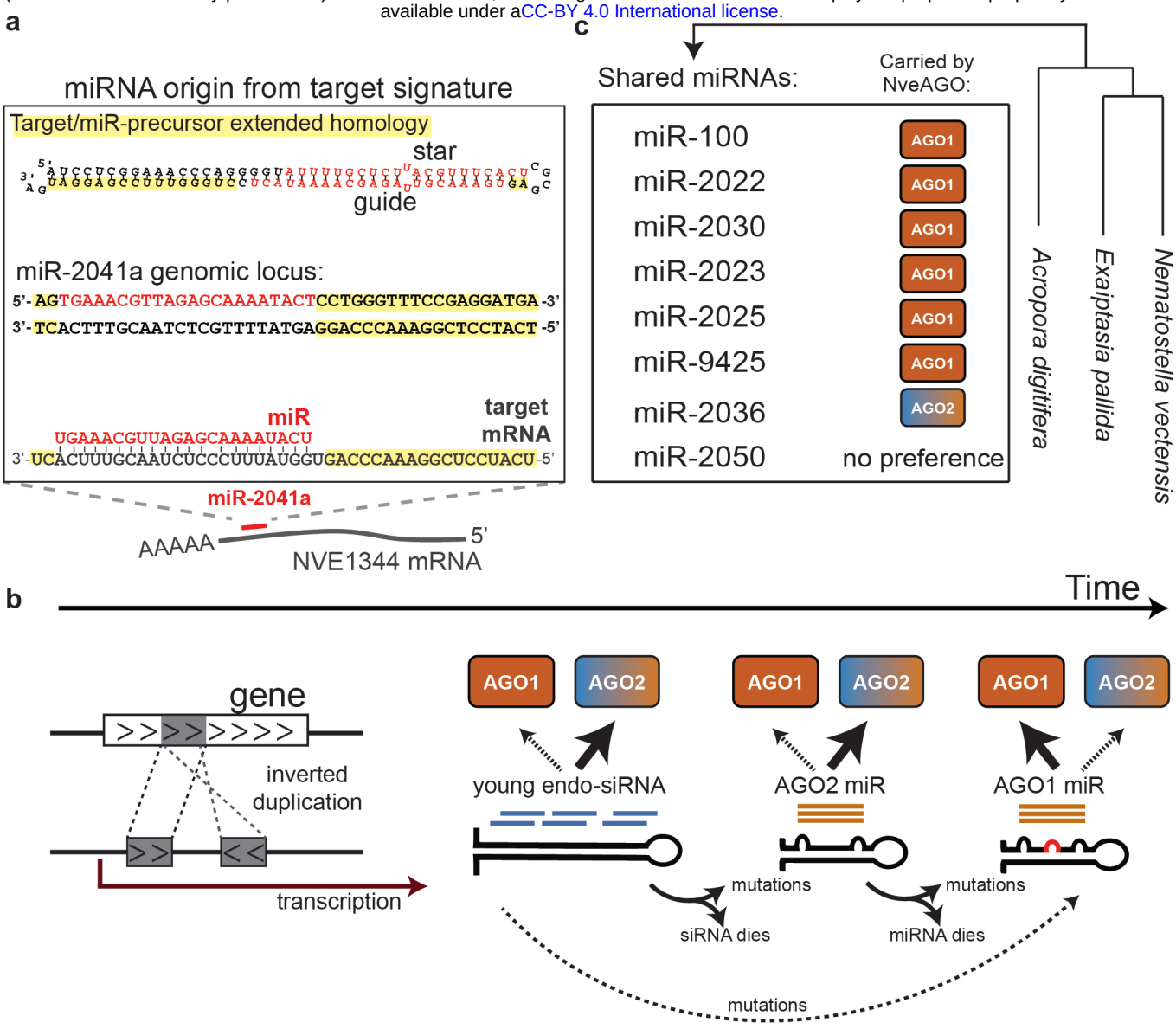


Fig. 5: Model for origin and evolution of miRNA-precursors in Hexacorallia

a, Extended homology between miRNA precursors and their targets exemplified for miR-2041a and its predicted target NVE1344. **b**, The model suggests endo-siRNA origin (that load into hexacorallian AGO2) from inverted duplications of coding genes. Next, precursors acquire mutations that result in processing by the miRNA biogenesis machinery, and such precursors still load into hexacorallian AGO2. Additional mutations such as a central mismatch in duplexes might result in the miRNA being directed to hexacorallian AGO1. **c**, miRNAs shared by sea anemones and reef-building corals and their AGO preference in *Nematostella* as discovered in this study.

Methods

Animal culture and knockdowns

Nematostella polyps were grown in the dark at 18°C in 16‰ artificial seawater and fed with freshly hatched *Artemia salina* nauplii three times a week. Induction of spawning was performed as previously described[54]. The gelatinous sack surrounding the eggs was removed using 4% L-Cysteine (Merck Milipore, USA) and followed by microinjecting the zygotes with Morpholino antisense oligonucleotide (MO). Next, zygotes were cultured at 22°C in 16 ‰ artificial seawater in the dark. The MO sequences were designed and synthesized by Gene Tools, LLC (USA).

NveAGO1 translation blocking MO: GATTCACTGGATTATTAGAAGCCAT

NveAGO2 translation blocking MO: TTAACAGCCTTTTGATGCTTTACGA

Standard control MO: CCTCTTACCTCAGTTACAATTTATA

1 mM stock solution of each MO was prepared with nuclease-free water. Equal concentrations of the MOs of different treatments and controls were injected the same day into zygotes from the same batch to generate one biological replicate. In total, three biological replicates of ~300 injected animals were generated for each NveAGO MO and control MO. Samples for RNA or protein extraction were flash frozen in liquid nitrogen and stored at -80°C until processed. Additionally, three independent biological replicates were generated for morphological analysis and the images were collected 9 dpf.

Antibody generation

For NveAGO1 and NveAGO2 IP and Western blots, custom polyclonal antibodies were generated in rabbits against recombinant fragments corresponding to regions that significantly differ between the two AGOs (GenScript, USA).

NveAGO1 C-terminal region:

QVGQEQRHTYLPLEVCNIVPGQRCVKKLTDTQTSKMIRATARSAPDREREIRGLVKKANFDEDA
YVKDFSISIGKNMVELQGRVLPPPKLVYGGKQSSPITPKGGVWDMRGRQLFHGIEIRTWAIACFV
KQQMCTEDSLRRFSNQLMKISVEQGMPISCPPVFFRYARNPDEVERMFRRLKEAHPDLQMILVU
LPGKTP

NveAGO2 N-terminal:

MPKKSKGRGRGRGRGNPHHEKOQTLVGQQATSRNNEHKQLPKPTQTQQETSLSQQPYTCSSAA
EAQGPLTPPNNTQGGSLTNAEPAQADTLGQKFETQLNLSPQSGKEQGAIPKTGARLKGNSLAPGS
QGNGQFSSKNLLAQMQRTPRSASESQNEASSKSQQAHNNQSQAAAHQQTQAGPQQTPARSQQT
PAGPQQTQAGPQQTPAGPQQTQAGPQQTQEGPQQTPARSQQTPAIIGSTTTADEHLARHRQENM
EL

Specifically, each recombinant fragment was injected into three rabbits. After the third round of immunization, pre-immune and post-immune sera were sent to us for screening by Western blot against

Nematostella lysate to identify sera specifically positive for NveAGO1 and NveAGO2 (bands of ~96 and ~122 kDa respectively). Finally, the antigens were used by the company for affinity purification from the relevant rabbits.

AGO immunoprecipitations

100 µl of magnetic beads (SureBeads™ Protein A Magnetic Beads, Bio-Rad, USA) were washed in 1 ml of 1×PBS for 5 times. 5 µg of antibodies/total IgG was added to the washed beads with 1.3 ml of 1×PBS, and left rotating overnight at 4ºC. Animals that correspond to a volume of 100 µl were frozen in liquid nitrogen and lysed (with homogenizer) using 1 ml of the following lysis buffer: 25 mM Tris-HCl (pH 7.4), 150 mM KCl, 25 mM EDTA, 0.5% NP-40, 1 mM DTT, Protease inhibitor cOmplete ULTRA tablets (Roche) and Protease Inhibitor Cocktail Set III, EDTA-Free (Merck Millipore, USA). Murine RNAse inhibitor (New England Biolabs, USA) was added whenever samples were used for downstream RNA-based applications. The DTT, Protease inhibitor and RNAse inhibitor were added fresh just before use. After 2 h rotation in 4ºC the samples were centrifuged at 16000 × *g*, 15 min, 4ºC and supernatant was collected and stored in -80ºC.

Next, the lysate was precleared as following: 100 µl of magnetic beads were washed in 1 ml of 1×PBS for 3 times and 1 time in lysis buffer and the lysate was added to the washed beads. Lysis buffer with RNAse inhibitor was added to make up 1.2 ml and the samples were incubated at 4ºC rotation for one hour. Next, the pre-cleared lysate was collected and added to the antibody-bound beads (that were pre-incubated with the antibody overnight). These samples were incubated for 2 h in rotation at 4ºC. After incubation the lysate was removed, and the beads were washed 5 times with the following wash buffer: 50mM Tris-HCl (pH 7.4), 300mM NaCl, 5mM $MgCl_2$, 0.05% NP-40, Protease inhibitor cOmplete ULTRA tablets (Roche, Switzerland) and Protease Inhibitor Cocktail Set III, EDTA-Free (Merck Millipore). Murine RNAse inhibitor was added whenever samples were used for downstream RNA-based applications The protease and RNAse inhibitors were added fresh just before use. Finally, for RNA extraction 1,000 µl of Tri-reagent (Sigma-Aldrich) was added followed by RNA isolation according to the manufacturer's protocol with the addition of 1.5 µl of RNA-grade glycogen (Thermo Fisher Scientific, USA) when precipitating with isopropanol.

For Western blot, 40 µl of filtered double-distilled water and 20 µl of blue SDS sample buffer (New England Biolabs) were added to the beads. The samples were heated at 100°C for 8 min and placed on ice for 1 min, then centrifuged 1 min at 23,000 × *g* at 4°C, and the supernatant was collected for Western blot.

Western blot

The samples ran on polyacrylamide gradient gel (4-15%; Bio-Rad) followed by blotting to a nitrocellulose membrane (Bio-Rad). Next, the membrane was washed with TBST buffer (20mM Tris pH 7.6, 150mM NaCl, 0.1% Tween 20) and blocked (5% skim milk in TBST) for 1 hour at room temperature. Purified polyclonal antibody against NveAGO1 or NveAGO2 was added to 5 ml of TBST containing 5% BSA (1:1000) (MP Biomedicals, USA), to the membrane in a sealed sterile plastic bag and incubated at 4°C overnight. Next the membrane was washed three times with TBST and incubated for 1 hour with Goat αrabbit IgG conjugated to horseradish peroxidase (Jackson's ImmunoResearch, USA) diluted to a concentration of 0.08 µg/ml in 5% skim milk in TBST. Finally, the membrane was washed three times with TBST and detection was performed with the Clarity™ ECL kit (Bio-Rad) according to the manufacturer's instructions and visualized with a CCD camera of the Odyssey Fc imaging system (Li-COR Biosciences, USA). Size determination was carried out by simultaneously running Precision Plus Protein™ Dual Color Protein Ladder (Bio-Rad) and scanning the same blot on the same system at 700 nm.

Small-RNA sequencing

Animals injected with NveAGO1, NveAGO2 and Control MOs were collected three days post fertilization. Three distinct biological replicates (of ~150 animals each) were generated. Total RNA was extracted and library preparation was carried out as described in the Zamore's lab, Illumina TruSeq small RNA Cloning Protocol April 2014 (http://www.umassmed.edu/zamore/resources/protocols/). Briefly, sRNAs were ligated to 3′ and 5′ adapters that contain four random nucleotides for minimizing ligation biases. SuperScriptIII (Thermo Fisher Scientific) was used to reverse-transcribe ligated products and KAPA Real-Time Library Amplification Kit (Roche) was used for cDNA amplification. The amplified samples were run and cut from 2% low-melt agarose gels (Bio-Rad) and followed by sRNA sequencing with NextSeq500 (Illumina, USA) with read lengths of 50 nucleotides. Prior to this library preparation, four synthetic spike-ins mimicking bilaterian miRNA sequences absent from *Nematostella* genome were added to the total RNA. Two were 2′-o-methylated at their 3′ ends (mmu-miR-125a-5p and mmu-miR-148a-3p), and the other two were not methylated (cel-lin-4-5p and hsa-miR-659-5p). These four miRNAs were further used for normalization of *Nematostella* miRNA read counts in the bioinformatic analysis.

Library preparation for RNA extracted from NveAGO1 and NveAGO2 IP was size selected for 18-30 nucleotides on 15% denaturing urea polyacrylamide gel (Bio-Rad) followed by overnight RNA elution in 0.3M NaCl. Next, library preparation was carried out using NEBNext Multiplex Small RNA Library Prep Set for Illumina kit (New England Biolabs) with modified protocol[23] for improving low-input samples and sequenced with NextSeq500 (Illumina) as described above. This procedure was carried out on two distinct biological replicates for each AGO in each of the developmental stages. Raw data was deposited at NCBI GEO submission (GSE144203).

Total RNA sequencing

RNA was extracted from three biological replicates of 3 days old animals injected with control, NveAGO1 and NveAGO2 MOs. Library preparation was carried with SENSE Total RNA-seq Library Prep Kit (Lexogen, Austria) and 75 nt single end sequencing was carried out with NextSeq500 (Illumina).

Reverse transcription and quantitative real-time PCR (qPCR)

Reverse transcription (RT) of miRNAs was carried out using miRCURY LNA Universal RT microRNA PCR Kit (Exiqon-Qiagen, Denmark), as instructed in miRCURY LNA RT Kit manual in three technical replicates for each miRNA. Equal amounts of RNA spike-in (Uni-Sp6) were added to the RNA and later used as internal amplification control. RT mixture included template RNA, 5× miRCURY RT Reaction Buffer (2 µl), 10× miRCURY RT Enzyme Mix (1 µl) and nuclease-free water to make up 10 µl of total volume. The mixture was incubated at 42°C for 1 h, then incubated at 95°C for 5 min to inactivate the reverse transcriptase and immediately cooled to 4°C. Real-Time PCR was performed using miRCURY SYBR Green PCR Kit (Exiqon-Qiagen) according to the manufacturer's instructions with StepOnePlus Real-Time PCR System (ABI, Thermo Fisher Scientific). The qPCR mixture contained cDNA template (3 µl), 2× miRCURY SYBR Green Master Mix (5 µl), LNA primer set (1 µl) and nuclease-free water to make up 10 µl total volume. qPCR thermocycling conditions were 95°C for 2 min, 40 cycles of 95°C for 10 s, 56°C for 1 min. Melt curve analysis: 60–95°C for 15 min at a ramp-rate of 1.6°C/s. The expression levels of miRNAs were normalized to the RNA spike-in (Uni-Sp6).

Semi-quantitative LC-MS/MS analysis

Sample preparation for MS analysis: After the last step of immunoprecipitation, the beads were washed twice with 25 mM Tris-HCl pH 8.0. The packed beads were re-suspended in 100 µl 8M urea, 10 mM DTT, 25 mM Tris-HCl pH 8.0 and incubated for 30 min at 22°C. Next, Iodoacetamide (55 mM) was added and beads were incubated for 30 min (22°C, in the dark), followed by addition of DTT (20 mM). The Urea was diluted by the addition of 6 volumes of 25 mM Tris-HCl pH 8.0. Trypsin was added (0.3 µg per sample) and the beads were incubated overnight at 37°C with gentle agitation. The beads were spun down and the peptides were desalted on C18 home-made Stage tips[55]. Two-thirds of the eluted peptides were used for MS analysis.

nanoLC-MS/MS analysis: MS analysis was performed using a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) coupled on-line to a nanoflow UHPLC instrument, Ultimate 3000 Dionex (Thermo Fisher Scientific). Peptides were separated over a 60 min gradient run at a flow rate of 0.3 µl/min on a reverse phase 25-cm-long C18 column (75 µm ID, 2 µm, 100Å, Thermo PepMapRSLC). The survey scans (380–2,000 m/z, target value 3E6 charges, maximum ion injection times 50 ms) were acquired and followed by higher energy collisional dissociation (HCD) based fragmentation (normalized collision energy 25). A resolution of 70,000 was used for survey scans and up to 15 dynamically chosen most abundant precursor ions, with "peptide preferable" profile were fragmented (isolation window 1.6 m/z). The MS/MS

scans were acquired at a resolution of 17,500 (target value 1E5 charges, maximum ion injection times 120 ms). Dynamic exclusion was 60 sec. Data were acquired using Xcalibur software (Thermo Scientific). To avoid a carryover, the column was washed with 80% acetonitrile, 0.1% formic acid for 25 min between samples.

MS data analysis: Mass spectra data were processed using the MaxQuant computational platform, version 1.5.3.1254. Peak lists were searched against translated coding sequences of *Nematostella* gene models. The search included cysteine carbamidomethylation as a fixed modification and oxidation of methionine as variable modifications and allowed up to two miscleavages. The match-between-runs option was used. Peptides with a length of at least seven amino-acids were considered and the required FDR was set to 1% at the peptide and protein level. Protein identification required at least 2 unique or razor peptides per protein. Relative protein quantification in MaxQuant was performed using the label-free quantification (LFQ) algorithm[56]. LFQ in MaxQuant uses only common peptides for pair-wise ratio determination for each protein and calculates a median ratio to protect against outliers. It then determines all pair-wise protein ratios and requires a minimal number of two peptide ratios for a given protein ratio to be considered valid. Statistical analysis (n=3) was performed using the Perseus statistical package[57]. Only sample groups with at least 2 valid values were used. Protein contaminants and proteins identified by less than 2 peptides were excluded from the analysis. The procedure described above was carried out on three technical replicates for each AGO-IP. MS/MS raw files, as well as results of MaxQuant analysis were deposited to the ProteomeXchange Consortium via the PRIDE55 partner repository with the dataset identifier: PXD011644.

Bioinformatic analysis

mirDeep2[24] core algorithm was used to identify novel miRNAs in AGO-IP samples. For quantification of miRNA counts for AGO-preference analysis, mirDeep2 quantifier.pl module was used with default parameters, and the reads were normalized to number of reads mapping to the genome (and multiplied by 1 million) in each library. To reduce noise, only miRNAs that exceeded a threshold of minimum 20 normalized reads in at least one library were included in this analysis. The normalized read counts from two biological replicates in each AGO-IP were averaged and the relative preference levels of each miRNA in NveAGO1 and NveAGO2 were calculated by dividing the counts from each NveAGO by the miRNA's total counts in both AGOs (**Fig. 3**). miRNAs from the heatmap that showed no AGO-preference in any of the developmental stages were classified as "no-preference miRNAs" (Supplementary table 3). In NveAGO1 and NveAGO2 knockdown sRNA libraries, the quantifier.pl module was run allowing no mismatches in sRNA reads mapped to miRNA precursors. These counts were normalized to the average of the four miRNA spike-ins in each library.

For total RNA sequencing analysis, filtered fastq reads were aligned to the Bowtie2[58] (version 2.3.4.1) indexed *Nematostella* genome using TopHat[59] (version 2.1.1) The number of reads mapping to

each *Nematostella* gene models (https://figshare.com/articles/Nematostella_vectensis_transcriptome_and_gene_models_v2_0/807696) was extracted using featurecounts[60] 1.6.0.2. The differential gene expression analysis was carried out using DESeq2[61].

Putative miRNA targets were predicted using psRNATarget[62] with a maximum expectation score of 2.5, allowing no gaps. miRNA origin from target signature was defined for miRNAs when the stem outside the guide region exhibited extended homology (of at least five consecutive bases) to the targeted gene. Prediction of miRNA target origin via retro-transposition was carried out manually by examination of miRNA loci and their putative targets.

To identify counts of endogenous siRNAs from IP sRNA libraries we preprocessed the raw reads to trim the adapter and then reads were filtered from known *Nematostella* miRNAs and other non-coding RNAs using Bowtie2 version 2.3.4.1. The filtered reads were subjected to ShortStack[31] (version 3.8.5) analysis with default parameters. In the final outputs, the "DicerCall" column was filtered from value: "N" and MIRNA column was filtered from the value "Y".

Phylogenetic analysis

The AGO amino acid sequences were aligned using MUSCLE[63] and low certainty alignment regions were removed by TrimAl[64] using the –automatic1 for heuristic model selection. ProtTest was used to find the most suitable model for phylogeny reconstruction[65]. The maximum-likelihood (ML) phylogenetic trees were constructed using PhyML with the LG +I +G model and support values were calculated using 100 bootstrap replicates[66]. A Bayesian tree was constructed using MrBayes version 3.2.1[67] with the WAG +I +G model and the run lasted 5,000,000 generations with every 100th generation being sampled. The Bayesian analysis was estimated to reach convergence when the potential scale reduction factor (PSRF) reached 1.0. Sequences are available in supplementary file 1.

References

1 Flynt, A. S. & Lai, E. C. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat Rev Genet* **9**, 831-842, doi:10.1038/nrg2455 (2008).

2 Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20-51, doi:10.1016/j.cell.2018.03.006 (2018).

3 Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **136**, 669-687, doi:10.1016/j.cell.2009.01.046 (2009).

4 D'Ario, M., Griffiths-Jones, S. & Kim, M. Small RNAs: Big Impact on Plant Development. *Trends in plant science* **22**, 1056-1068, doi:10.1016/j.tplants.2017.09.009 (2017).

5 Grimson, A. *et al.* Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**, 1193-1197, doi:10.1038/nature07415 (2008).

6 Moran, Y., Agron, M., Praher, D. & Technau, U. The evolutionary origin of plant and animal microRNAs. *Nat Ecol Evol* **1**, 27, doi:10.1038/s41559-016-0027 (2017).

7 Fridrich, A., Hazan, Y. & Moran, Y. Too Many False Targets for MicroRNAs: Challenges and Pitfalls in Prediction of miRNA Targets and Their Gene Ontology in Model and Non-model Organisms. *Bioessays* **41**, e1800169, doi:10.1002/bies.201800169 (2019).

8 Modepalli, V., Fridrich, A., Agron, M. & Moran, Y. The methyltransferase HEN1 is required in Nematostella vectensis for microRNA and piRNA stability as well as larval metamorphosis. *PLoS Genet* **14**, e1007590, doi:10.1371/journal.pgen.1007590 (2018).

9 Moran, Y. *et al.* Cnidarian microRNAs frequently regulate targets by cleavage. *Genome research* **24**, 651-663, doi:10.1101/gr.162503.113 (2014).

10 Moran, Y., Praher, D., Fredman, D. & Technau, U. The evolution of microRNA pathway protein components in Cnidaria. *Molecular biology and evolution* **30**, 2541-2552, doi:10.1093/molbev/mst159 (2013).

11 Axtell, M. J., Westholm, J. O. & Lai, E. C. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome biology* **12**, 221, doi:10.1186/gb-2011-12-4-221 (2011).

12 Tarver, J. E., Donoghue, P. C. & Peterson, K. J. Do miRNAs have a deep evolutionary history? *Bioessays* **34**, 857-866, doi:10.1002/bies.201200055 (2012).

13 Elkayam, E. *et al.* The structure of human argonaute-2 in complex with miR-20a. *Cell* **150**, 100-110, doi:10.1016/j.cell.2012.05.017 (2012).

14 Huntzinger, E. & Izaurralde, E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* **12**, 99-110, doi:10.1038/nrg2936 (2011).

15 Wilson, R. C. & Doudna, J. A. Molecular mechanisms of RNA interference. *Annu Rev Biophys* **42**, 217-239, doi:10.1146/annurev-biophys-083012-130404 (2013).

16 Park, E. *et al.* Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Molecular phylogenetics and evolution* **62**, 329-345, doi:10.1016/j.ympev.2011.10.008 (2012).

17 Shinzato, C. *et al.* Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature* **476**, 320-323, doi:10.1038/nature10249 (2011).

18 He, X. & Zhang, J. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics* **169**, 1157-1164, doi:10.1534/genetics.104.037051 (2005).

19 Birchler, J. A. & Veitia, R. A. Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc Natl Acad Sci U S A* **109**, 14746-14753, doi:10.1073/pnas.1207726109 (2012).

20 Cenik, E. S. & Zamore, P. D. Argonaute proteins. *Current biology : CB* **21**, R446-449, doi:10.1016/j.cub.2011.05.020 (2011).

21 Meister, G. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* **14**, 447-459, doi:10.1038/nrg3462 (2013).

22 Wynant, N., Santos, D. & Vanden Broeck, J. The evolution of animal Argonautes: evidence for the absence of antiviral AGO Argonautes in vertebrates. *Sci Rep* **7**, 9230, doi:10.1038/s41598-017-08043-5 (2017).

23 Plotnikova, A., Kellner, M. J., Schon, M. A., Mosiolek, M. & Nodine, M. D. MicroRNA Dynamics and Functions During Arabidopsis Embryogenesis. *Plant Cell* **31**, 2929-2946, doi:10.1105/tpc.19.00395 (2019).

24 Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic acids research* **40**, 37-52, doi:10.1093/nar/gkr688 (2012).

25 Fromm, B. *et al.* 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 (2015).

26 Praher, D. *et al.* Characterization of the piRNA pathway during development of the sea anemone Nematostella vectensis. *RNA Biol* **14**, 1727-1741, doi:10.1080/15476286.2017.1349048 (2017).

27 Rubio, M., Maestro, J. L., Piulachs, M. D. & Belles, X. Conserved association of Argonaute 1 and 2 proteins with miRNA and siRNA pathways throughout insect evolution, from cockroaches to flies. *Biochim Biophys Acta Gene Regul Mech* **1861**, 554-560, doi:10.1016/j.bbagrm.2018.04.001 (2018).

28 Lewis, S. H., Salmela, H. & Obbard, D. J. Duplication and Diversification of Dipteran Argonaute Genes, and the Evolutionary Divergence of Piwi and Aubergine. *Genome Biol Evol* **8**, 507-518, doi:10.1093/gbe/evw018 (2016).

29 Wee, L. M., Flores-Jasso, C. F., Salomon, W. E. & Zamore, P. D. Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell* **151**, 1055-1067, doi:10.1016/j.cell.2012.10.036 (2012).

30 Calcino, A. D., Fernandez-Valverde, S. L., Taft, R. J. & Degnan, B. M. Diverse RNA interference strategies in early-branching metazoans. *BMC Evol Biol* **18**, 160, doi:10.1186/s12862-018-1274-2 (2018).

31 Axtell, M. J. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* **19**, 740-751, doi:10.1261/rna.035279.112 (2013).

32 Shahid, S. & Axtell, M. J. Identification and annotation of small RNA genes using ShortStack. *Methods* **67**, 20-27, doi:10.1016/j.ymeth.2013.10.004 (2014).

33 Schuster, S., Miesen, P. & van Rij, R. P. Antiviral RNAi in Insects and Mammals: Parallels and Differences. *Viruses* **11**, doi:10.3390/v11050448 (2019).

34 Flemr, M. *et al.* A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* **155**, 807-816, doi:10.1016/j.cell.2013.10.001 (2013).

35 Burroughs, A. M. *et al.* Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol* **8**, 158-177, doi:10.4161/rna.8.1.14300 (2011).

36 Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-1441, doi:10.1126/science.1102513 (2004).

37 Cheloufi, S., Dos Santos, C. O., Chong, M. M. & Hannon, G. J. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**, 584-589, doi:10.1038/nature09092 (2010).

38 Cifuentes, D. *et al.* A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**, 1694-1698, doi:10.1126/science.1190809 (2010).

39 Yang, J. S. *et al.* Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc Natl Acad Sci U S A* **107**, 15163-15168, doi:10.1073/pnas.1006432107 (2010).

40 Brown, K. C., Svendsen, J. M., Tucci, R. M., Montgomery, B. E. & Montgomery, T. A. ALG-5 is a miRNA-associated Argonaute required for proper developmental timing in the Caenorhabditis elegans germline. *Nucleic acids research* **45**, 9093-9107, doi:10.1093/nar/gkx536 (2017).

41 Forstemann, K., Horwich, M. D., Wee, L., Tomari, Y. & Zamore, P. D. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* **130**, 287-297, doi:10.1016/j.cell.2007.05.056 (2007).

42 Tomari, Y., Du, T. & Zamore, P. D. Sorting of Drosophila small silencing RNAs. *Cell* **130**, 299-308, doi:10.1016/j.cell.2007.05.057 (2007).

43 Ameres, S. L., Hung, J. H., Xu, J., Weng, Z. & Zamore, P. D. Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins. *RNA* **17**, 54-63, doi:10.1261/rna.2498411 (2011).

44 Zhang, X. *et al.* ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. *Nat Commun* **5**, 5468, doi:10.1038/ncomms6468 (2014).

45 Mi, S. *et al.* Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**, 116-127, doi:10.1016/j.cell.2008.02.034 (2008).

46 Kawamata, T., Seitz, H. & Tomari, Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nature structural & molecular biology* **16**, 953-960, doi:10.1038/nsmb.1630 (2009).

47 Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome research* **14**, 1902-1910, doi:10.1101/gr.2722704 (2004).

48 Watanabe, T., Cheng, E. C., Zhong, M. & Lin, H. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome research* **25**, 368-380, doi:10.1101/gr.180802.114 (2015).

49 Allen, E. *et al.* Evolution of microRNA genes by inverted duplication of target gene sequences in Arabidopsis thaliana. *Nat Genet* **36**, 1282-1290, doi:10.1038/ng1478 (2004).

50 Ameres, S. L. *et al.* Target RNA-directed trimming and tailing of small silencing RNAs. *Science* **328**, 1534-1539, doi:10.1126/science.1187058 (2010).

51 Zhao, Y., Mo, B. & Chen, X. Mechanisms that impact microRNA stability in plants. *RNA Biol* **9**, 1218-1223, doi:10.4161/rna.22034 (2012).

52 Li, J., Yang, Z., Yu, B., Liu, J. & Chen, X. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Current biology : CB* **15**, 1501-1507, doi:10.1016/j.cub.2005.07.029 (2005).

53 Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome research* **14**, 1188-1190, doi:10.1101/gr.849004 (2004).

54 Genikhovich, G. & Technau, U. Induction of spawning in the starlet sea anemone Nematostella vectensis, in vitro fertilization of gametes, and dejellying of zygotes. *Cold Spring Harb Protoc* **2009**, pdb prot5281, doi:10.1101/pdb.prot5281 (2009).

55 Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906, doi:10.1038/nprot.2007.261 (2007).

56 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526, doi:10.1074/mcp.M113.031591 (2014).

57 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

58 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

59 Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).

60 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

61 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

62 Dai, X., Zhuang, Z. & Zhao, P. X. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic acids research* **46**, W49-W54, doi:10.1093/nar/gky316 (2018).

63 Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113, doi:10.1186/1471-2105-5-113 (2004).

64 Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973, doi:10.1093/bioinformatics/btp348 (2009).

65 Abascal, F., Zardoya, R. & Posada, D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104-2105, doi:10.1093/bioinformatics/bti263 (2005).

66 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321, doi:10.1093/sysbio/syq010 (2010).

67 Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**, 539-542, doi:10.1093/sysbio/sys029 (2012).