Cnidarian microRNAs frequently regulate targets by cleavage

Yehu Moran

University of Vienna

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_sp

Part of the Ecology and Evolutionary Biology Commons, and the Genomics Commons

Repository Citation


This material is brought to you by eScholarship@UMassChan. It has been accepted for inclusion in GSBS Student Publications by an authorized administrator of eScholarship@UMassChan. For more information, please contact Lisa.Palmer@umassmed.edu.
Cnidarian microRNAs frequently regulate targets by cleavage

Yehu Moran,1,5,6 David Fredman,1,5,7 Daniela Praher,1 Xin Z. Li,2 Liang Meng Wee,2 Fabian Rentzsch,3 Phillip D. Zamore,2,8 Ulrich Technau,1,8 and Hervé Seitz4,8

1Department for Molecular Evolution and Development, Center for Organismal Systems Biology, Faculty of Life Sciences, University of Vienna, 1090 Vienna, Austria; 2RNA Therapeutics Institute, Department of Biochemistry & Molecular Pharmacology, and Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA; 3Sars Centre for Marine Molecular Biology, University of Bergen, N-5008 Bergen, Norway; 4Institute of Human Genetics, UPR 1142, CNRS, 34396 Montpellier Cedex 5, France

In bilaterians, which comprise most of extant animals, microRNAs (miRNAs) regulate the majority of messenger RNAs (mRNAs) via base-pairing of a short sequence (the miRNA “seed”) to the target, subsequently promoting translational inhibition and transcript instability. In plants, many miRNAs guide endonucleolytic cleavage of highly complementary targets. Because little is known about miRNA function in nonbilateral animals, we investigated the repertoire and biological activity of miRNAs in the sea anemone Nematostella vectensis, a representative of Cnidaria, the sister phylum of Bilateria. Our work uncovers scores of novel miRNAs in Nematostella, increasing the total miRNA gene count to 87. Yet only a handful are conserved in corals and hydras, suggesting that microRNA gene turnover in Cnidaria greatly exceeds that of other metazoan organisms. We further show that Nematostella miRNAs frequently direct the cleavage of their mRNA targets via nearly perfect complementarity. This mode of action resembles that of small interfering RNAs (siRNAs) and plant miRNAs. It appears to be common in Cnidaria, as several of the miRNA target sites are conserved among distantly related anemone species, and we also detected miRNA-directed cleavage in Hydra. Unlike in bilaterians, Nematostella miRNAs are commonly coexpressed with their target transcripts. In light of these findings, we propose that post-transcriptional regulation by miRNAs functions differently in Cnidaria and Bilateria. The similar, siRNA-like mode of action of miRNAs in Cnidaria and plants suggests that this may be an ancestral state.

[Supplemental material is available for this article.]

MicroRNAs (miRNAs), small RNAs 20–24 nucleotides (nt) long, regulate messenger RNA (mRNA) expression in plants and animals (Ghildiyal and Zamore 2009). In animals, sequential processing of most primary miRNAs by the RNase III enzymes Drosha and Dicer produces miRNA/miRNA* duplexes (Grishok et al. 2001; Hultvagner et al. 2001; Ketting et al. 2001; Knight and Bass 2001; Lee et al. 2003). Subsequently, these duplexes are loaded into a member of the Argonaute (AGO) protein family and the miRNA* strand is ejected, producing an active RNA-induced silencing complex. miRNAs direct AGO proteins to repress expression of partially complementary miRNAs, thereby modulating diverse biological processes such as organogenesis, developmental timing, and cell proliferation (Bartel 2009). In bilaterian animals, miRNAs typically recognize their targets via a short nucleotide sequence, the “seed” (miRNA nucleotides 2–8), triggering translational inhibition and transcript decay by mechanisms such as deadenylation (Bartel 2009; Huntzinger and Izaurralde 2011). In plants, by contrast, most known targets exhibit near perfect complementarity to the miRNA, permitting AGO proteins to cleave the miRNA target (Baumberger and Baulcombe 2005; Axtell et al. 2011). Target cleavage directed by plant miRNAs resembles the action of small interfering RNAs (siRNAs), which are produced by Dicer alone and are found in most eukaryote lineages, suggesting that they represent ancestral small silencing RNAs (Ghildiyal and Zamore 2009).

The biogenesis and action of small silencing RNAs are well studied in Bilateria, the group that includes the majority of extant animals such as vertebrates, arthropods, and nematodes. In contrast, our knowledge about small silencing RNAs in early-branching, nonbilaterian lineages remains limited. Multiple phylogenomic studies position Cnidaria (sea anemones, corals, hydrozoa, and jellyfish) as the sister group of Bilateria (Hejnol et al. 2009; Philippe et al. 2011). Hence, comparison of cnidarians and bilaterians may shed light on the biology of their common ancestor that lived more than 600 million yr ago. A pioneering study on small RNAs in basally branching animals found that the model sea anemone, Nematostella vectensis, produces miRNAs and a second class of small RNAs, PIWI-interacting RNAs (piRNAs). Intriguingly, only a single miRNA, miR-100, appears to be conserved between bilaterians and cnidarians (Grimson et al. 2008). Here, we identify many new miRNAs in Nematostella and show by high-throughput sequencing of nine developmental stages that the majority of miRNAs are developmentally regulated. Nematostella miRNAs are frequently
coexpressed with their targets and reveal extensive complementarity between most miRNAs and their targets. Further, this complementarity commonly leads to slicing of the target mRNA, like siRNAs and plant miRNAs. Our findings suggest that cnidarian miRNAs retain a potentially ancestral, RNAi-like mode of small RNA-mediated post-transcriptional regulation, unlike the vast majority of miRNAs in bilaterians.

Results

Discovery of Nematostella miRNAs and siRNAs

A large proportion (>93% in all libraries and developmental stages) (Table 1) of the sequenced Nematostella small RNAs exhibit sequence signatures typical of piRNAs: 23- and 24-mers displaying an adenosine bias at position 10 and longer reads exhibiting a uridine bias at position 1). We focused on those small RNAs with homogeneous 5’ ends that mapped to presumptive pre-miRNA structures (58-nt sequences folding into an unbranched hairpin, with a predicted folding free energy ≤ −14 kcal mol−1). Among these, we identified 128 pre-miRNAs that produce 86 distinct mature miRNAs belonging to 84 seed-sequence families (Supplemental Table S1; Supplemental Fig. S1). In order to estimate the robustness of our miRNA annotations, we evaluated each of the 86 identified candidates by running the annotation program on random subsets of the 18 deep-sequencing libraries. When 10 subsets (each covering 90% of the actual library depth) were reanalyzed, 52 candidates were recovered in all 10 subsets, 65 candidates were recovered in at least half of the subsets, and 75 candidates were recovered in at least one subset. Supplemental Table S2 gives the confidence level (assessed by bootstrap score) for each candidate. Poorly recovered candidates tended to be those least expressed, and it is expected that they would have been missed by a shallower sequencing effort. Reciprocally, additional, yet-to-be-discovered miRNAs have probably escaped identification, but we expect that their expression levels would be very low (i.e., ~10 reads per million). Our initial miRNA annotation included 28 of the 40 miRNAs that were annotated previously (Grimson et al. 2008). To assess why the remaining 12 had not been recovered, we reviewed each case and found that 11 miRNAs did not qualify as authentic miRNAs, either because their low abundance did not permit unambiguous annotation, or because they did not pass our precursor structure and sequence homogeneity criteria. Instead, they were represented by a complex set of slightly shifted small RNA sequences with heterogeneous 5’ ends, probably originating from imprecise nucleolytic processing (Fig. 1C; Supplemental Fig. S2). These small RNAs resemble both siRNAs and miRNAs derived from short hairpin precursors; we note that this is in contrast to the situation in

<table>
<thead>
<tr>
<th>Description</th>
<th>untreated libraries</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Library depth</td>
<td>miRNA read count</td>
<td>siRNA read count</td>
<td>piRNA read count</td>
</tr>
<tr>
<td>Unfertilized eggs</td>
<td>2,613,163</td>
<td>2,197</td>
<td>1,402</td>
<td>2,566,854</td>
</tr>
<tr>
<td>Blastula</td>
<td>2,637,518</td>
<td>4,440</td>
<td>1,450</td>
<td>2,587,622</td>
</tr>
<tr>
<td>Gastrula</td>
<td>4,497,317</td>
<td>8,366</td>
<td>2,828</td>
<td>4,363,042</td>
</tr>
<tr>
<td>Early planula</td>
<td>4,608,287</td>
<td>66,548</td>
<td>2,423</td>
<td>4,401,101</td>
</tr>
<tr>
<td>Late planula</td>
<td>2,417,870</td>
<td>19,652</td>
<td>1,586</td>
<td>2,322,038</td>
</tr>
<tr>
<td>Metamorphosing</td>
<td>2,125,638</td>
<td>45,732</td>
<td>1028</td>
<td>2,032,277</td>
</tr>
<tr>
<td>Primary polyps</td>
<td>4,173,973</td>
<td>167,348</td>
<td>2316</td>
<td>3,900,870</td>
</tr>
<tr>
<td>Adult males</td>
<td>2,890,922</td>
<td>89,589</td>
<td>3754</td>
<td>2,726,302</td>
</tr>
<tr>
<td>Adult females</td>
<td>2,745,716</td>
<td>77,286</td>
<td>2669</td>
<td>2,596,434</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description</th>
<th>oxidized libraries</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Library depth</td>
<td>miRNA read count</td>
<td>siRNA read count</td>
<td>piRNA read count</td>
</tr>
<tr>
<td>Unfertilized eggs</td>
<td>2,498,767</td>
<td>1893</td>
<td>1,402</td>
<td>2,435,869</td>
</tr>
<tr>
<td>Blastula</td>
<td>2,685,304</td>
<td>2092</td>
<td>1,334</td>
<td>2,628,297</td>
</tr>
<tr>
<td>Gastrula</td>
<td>4,504,795</td>
<td>3764</td>
<td>2,462</td>
<td>4,382,743</td>
</tr>
<tr>
<td>Early planula</td>
<td>4,576,031</td>
<td>11,911</td>
<td>2795</td>
<td>4,458,313</td>
</tr>
<tr>
<td>Late planula</td>
<td>2,719,635</td>
<td>20,791</td>
<td>1,419</td>
<td>2,644,275</td>
</tr>
<tr>
<td>Metamorphosing</td>
<td>2,458,096</td>
<td>22,713</td>
<td>1,056</td>
<td>2,392,391</td>
</tr>
<tr>
<td>Primary polyps</td>
<td>4,232,194</td>
<td>119,837</td>
<td>1746</td>
<td>4,013,714</td>
</tr>
<tr>
<td>Adult males</td>
<td>2,620,395</td>
<td>38,579</td>
<td>1396</td>
<td>2,529,642</td>
</tr>
<tr>
<td>Adult females</td>
<td>2,937,274</td>
<td>38,666</td>
<td>850</td>
<td>2,859,465</td>
</tr>
</tbody>
</table>

Sequencing depth measures the total number of genome-matching reads not matching any abundant noncoding RNA (rRNAs, tRNAs, snRNAs, and snoRNAs). miRNAs were annotated as described in the Methods section. siRNAs were defined as 19- to 22-nt-long RNAs originating from long RNA hairpins. Longer reads (23–30 nt) not matching miRNAs were flagged as piRNAs. They exhibit the typical piRNA sequence signatures (with 23- and 24-mers displaying an adenosine bias at position 10 and longer reads exhibiting a uridine bias at position 1).
bilaterians, where siRNAs and miRNAs are well-separated classes (Ghildiyal and Zamore 2009). In this respect, the case of the miR-2024 variants in *Nematostella* is of special interest: Our analysis suggests that four miR-2024 variants (miR-2024a, e, f, and g) are bona fide miRNAs, whereas the three remaining variants do not present all the features of canonical miRNAs. The overhangs of the miR-2024b/miR-2024b* duplex are unusually long, and miR-2024c is imperfectly processed, while the single-stranded flanks of the pre-miR-2024d hairpin liberate more reads than the miRNA*. The unusual features of pre-miRs 2024b, c, and d and their low abundance relative to other miR-2024 variants suggest that they may be degenerating, providing unprecedented examples of pre-miRNAs transitioning into sources of siRNAs. We found the conserved and experimentally confirmed miR-2022 (Chapman et al. 2010; Krishna et al. 2013; this study) to be a false negative in our initial annotation, because it did not pass one of the many stringent criteria (it folded less stably than another RNA hairpin located in close proximity). After the addition of miR-2022, the final set of 87 miRNAs includes 29 of the 40 previously annotated miRNAs (Grimson et al. 2008).

Of the 87 *Nematostella* miRNAs, miR-100 remains the only miRNA conserved between cnidarians and bilaterians (Fig. 1D). Thus, the previous observation indicating that the vast majority of miRNAs evolved independently in Cnidaria and Bilateria was not biased by sequencing depth or the samples used (Grimson et al. 2008). Comparing our annotated *Nematostella* miRNAs to the genome of the reef-building coral *Acropora digitifera* and to sequenced small RNAs from *Hydra magnipapillata* (Shinzato et al. 2011; Krishna et al. 2013) enabled us to reconstruct the gain and loss history of miRNAs in Cnidaria (Fig. 1D). Interestingly, only two miRNAs seem to be shared among the three cnidarian groups, compared to 31 ancestral miRNA families in bilaterians (Fig. 1D; Sperling et al. 2009). Notably, not only the mature sequences of these two miRNAs are conserved, but sequence homology can also be found in their hairpin structure (Fig. 1E).

**Spatiotemporal regulation of *Nematostella* miRNAs**

Most miRNAs (66 of 87) are barely detectable before the early planula stage (<10 ppm in untreated libraries) (Fig. 2A; Supplementary Table S1), and the expression of most peaks either at the primary polyp or adult stage (Fig. 2A). These temporal profiles indicate that the expression of the majority of *Nematostella* miRNAs is developmentally regulated. Eighteen miRNAs are male-specific and just one (miR-9459) is female-specific (enrichment in one...
sex > 10, with Fisher’s exact test P-value < 0.01). Sex-specific miRNAs have also been reported in mammals and flatworms and may exist in fish and birds (Wienholds et al. 2005; Ciaudo et al. 2009; Hao et al. 2010; Luo et al. 2012; Marco et al. 2013). In situ hybridization experiments suggest that many Nematostella miRNAs have highly specific expression patterns (Fig. 2B) as documented for plant and bilaterian miRNAs (Flynt and Lai 2008; Axtell et al. 2011). For example, miR-100, miR-2022, and miR-2030 are expressed in groups of cells at the oral end; miR-2025 is expressed in the aboral endoderm; miR-2026 expression is confined to one side of the endoderm along the oral-aboral axis, and miR-2023 expression is restricted to cells in the aboral ectoderm. These data support a role for miRNAs in defining various cell or body region identities in Nematostella.

**Nematostella** miRNAs mediate target cleavage

In plants, miRNAs are 2’-O-methylated at their 3’ ends, whereas most animal miRNAs are not (Yu et al. 2005; Axtell et al. 2011). We assayed whether Nematostella miRNAs are methylated by sequencing oxidized small RNAs, a strategy that excludes unmethylated RNAs from the sequencing library (Ghildiyal et al. 2008; Axtell et al. 2011). For example, miR-100, miR-2022, and miR-2030 are expressed in groups of cells at the oral end; miR-2025 is expressed in the aboral endoderm; miR-2026 expression is confined to one side of the endoderm along the oral-aboral axis, and miR-2023 expression is restricted to cells in the aboral ectoderm. These data support a role for miRNAs in defining various cell or body region identities in Nematostella.

With our recent observation that the RNA methyltransferase Hen1 of Nematostella is expressed throughout the animal (Moran et al. 2013b), unlike in vertebrates where Hen1 expression is restricted to the germline where it methylates piRNAs (Kamminga et al. 2010). To estimate the efficiency of the oxidation reaction, we spiked eight of the libraries (oxidized and untreated libraries of unfertilized eggs, blastula, adult males, and adult females) with methylated and unmethylated RNA oligos. Unmethylated RNA oligos were depleted at >98% in the oxidized libraries, while methylated RNA oligos were detected at similar levels in untreated and oxidized libraries, demonstrating the high efficiency of the treatment (Supplemental Fig. S3; Supplemental Table S1). We classified miRNAs by their relative levels of methylation (see Methods); 43 miRNAs were overmethylated, 26 undermethylated, and 18 undetermined. Interestingly, overmethylated miRNAs were, on average, significantly longer than miRNAs with lower average methylation levels (P-value = 0.018, Wilcoxon test) (Fig. 3), whereas miRNAs whose methylation level could not be reliably determined displayed an intermediate size distribution. This may suggest that the methylation machinery in Nematostella has preference for longer small RNAs such as long miRNAs and piRNAs.

Methylation of small RNAs has been proposed to protect small RNAs—such as anti-viral siRNAs in insects—from degradation when they bind to extensively complementary target RNAs (Ameres et al. 2010). The presence of miRNAs with modified...
miRNA-mediated target cleavage in Cnidaria

miRNA target sites and the slicing mechanism are conserved in Cnidaria

If the cleavage of targets by miRNAs in Nematostella is under selective pressure, we would expect some of the miRNA target sites to be conserved in other cnidian species. Degradome analysis identified NvHoxD (also called Anthox) as a target of miR-2026. NvHoxD, a homeobox-containing transcription factor localized to one side of the directive axis of Nematostella (Fig. 5A), is thought to be involved in the differentiation along this axis (Finnerty et al. 2004). HoxD has been reported to exist as two closely related paralogs (Chourrout et al. 2006), but our exhaustive cloning of HoxD transcripts as well as sequencing of the corresponding genomic region suggests that Nematostella possesses a single HoxD gene bearing duplicated 3’ exons that produce several mRNA isoforms via alternative splicing (Fig. 5B). Interestingly, the miR-2026 binding site is found in both copies of the exon containing the 3’ UTR of NvHoxD, despite otherwise low sequence conservation of this region, suggesting purifying selection to retain the miR-2026 binding site. RLM-RACE showed that both 3’ UTR copies are cleaved (Fig. 5B). Moreover, the miR-2026 binding site in the 3’ UTR of HoxD is conserved in the HoxD ortholog of the sea anemone Metridium senile, which is separated from Nematostella vectensis by about 250–450 million yr (Erwin et al. 2011; Park et al. 2012). This time range is comparable to that of the deepest miRNA target conservations known in plants (Addo-Quaye et al. 2009).

Another unexpected finding regarding the targeting of HoxD by miR-2026 is the genomic location of the miRNA gene: In Bilateria, several miRNA genes reside in the Hox cluster upstream of the Hox gene they regulate (Mansfield and McGlinn 2012). This arrangement is thought to reinforce the posterior prevalence mediated by Hox genes (Yekta et al. 2008), a phenomenon that causes posterior phenotypes when both anterior and posterior genes are coexpressed. Nematostella lacks bilaterian-like Hox clusters, but several of its Hox genes are found in the same genomic region (Chourrout et al. 2006). Notably, the miR-2026 gene is located gene-specific, 5’ RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE), a gene-specific and high-sensitivity approach used previously in plants, green algae, and mammals (Llave et al. 2002; Zhao et al. 2007; Karginov et al. 2010). Among 25 potential miRNA targets assayed, seven could either not be detected or yielded shorter PCR products than expected, an inconclusive finding that might reflect further degradation of cleavage products. Of the 18 conclusive cases, 13 (70%) supported miRNA-directed target cleavage (Table 2; Supplemental Table S4). Interestingly, we could validate a few cases of cleavage for moderately expressed miRNA targets, like the targets of miR-2026, which were barely detected by the degradome approach (Fig. 4D,E). This suggests that the degradome sequencing missed a portion of cleavage events, possibly due to downstream effects of exonucleases, as was shown in plants and mammals, and due to the use of polyA selection in the degradome sequencing protocol which hinders the detection of sliced products whose poly-A tails have been degraded (German et al. 2008; Shin et al. 2010). In sum, we demonstrated by RLM-RACE the specific cleavage of 13 different target miRNAs by 11 different miRNAs belonging to nine miRNA families. Together, the common occurrence of high complementarity of miRNAs to specific targets and our RLM-RACE and degradome data suggest that target slicing by miRNAs is a common mechanism in Nematostella.
between the HoxC (Anthox7) and HoxD genes. Another unexpected similarity between miRNA-mediated regulation of Hox genes in Nematostella and bilaterians is the slicing of targets: One of the best-studied cases of miRNA-directed transcript slicing in bilaterians is the cleavage of HoxB8 mRNA by the Hox cluster-embedded miR-196 (Yekta et al. 2004). Since miR-196 and miR-2026 share no sequence motif, and because HoxD and HoxB8 are not orthologs, it is very likely that this reflects the convergent evolution of Hox gene regulation by neighboring miRNAs in Bilateria and Cnidaria.

Another example of a miRNA target of the homeobox transcription factor family in Nematostella is NvSix3/6. Six3/6 specifies the anterior-most part of the bilaterian brain and epidermis and has a role as an upstream regulator of the apical domain in the larva of Nematostella (Steinmetz et al. 2010; Sinigaglia et al. 2013). The Nematostella Six3/6 binding site for miR-2025 in the coding sequence (CDS) is conserved in the orthologous transcripts from the anemone species Anemonia viridis and Anthopleura japonica, despite poor overall sequence conservation of the mRNAs (Fig. 5C). The miRNA binding sites in the Six3/6 transcripts of A. viridis and A. japonica and the HoxD transcript of M. senile were the only conserved sites we detected between Nematostella and these three species. However, this might be explained by the limited tran-
Coherent and incoherent modes of regulation

miRNAs can be involved in either coherent or incoherent regulation (Shkumatava et al. 2009; Ebert and Sharp 2012). In coherent regulation, a transcription factor (TF) drives expression of a miRNA and concurrently inhibits, directly or indirectly, the expression of the miRNA target. Consequently, the miRNA reinforces the down-stream effect of the upstream regulatory system. In contrast, in the incoherent regulation topology, the miRNA and its target are activated by the same regulators, and therefore the miRNA effect counteracts the effect of the upstream system. In coherent regulation, the expression domains of the miRNA and its target are nearly mutually exclusive in space or time, whereas in incoherent regulation they must overlap temporally and spatially (Fig. 5D; Shkumatava et al. 2009; Ebert and Sharp 2012) and suggest that miR-2025 might help in setting sharper boundaries between cell layers.

Hence, we find evidence for both coherent and incoherent modes of regulation in Nematostella. miR-2030 and its target is a particularly interesting case of incoherent regulation: The precursor of this miRNA resides in an intron of NVE19315, a gene coding for an unknown protein (Fig. 5E). The transcript of this gene has two sites that are nearly perfectly complementary to miR-2030, and we show by RLM-RACE and degradome sequencing that both of them are active cleavage sites (Table 2; Supplemental Table S4; Fig. 5E). Since miR-2030 shows a similar expression pattern to its host/target gene, and there is no support in the transcriptome for independent transcription of the intronic region, it is clear that the expression of the miRNA and the hosting target gene are tightly linked (Fig. 5A; unpubl.). Thus, miR-2030 is part of a feed-forward incoherent circuit, a rare topology for miRNAs and their targets in bilaterians (Shkumatava et al. 2009). A similar case to miR-2030 and its host/target gene is miR-2028, as its precursor is also embedded in the intron of its own target (Supplemental Fig. S5).

Discussion

By deep sequencing of small RNAs from nine developmental stages of Nematostella, we increased the number of miRNAs to 87. Of these miRNAs, only two are conserved throughout Cnidaria (Fig. 1D). This low number could be due to acquisition of miRNAs after the divergence of these species from a common ancestor that evolved only a few miRNAs to begin with. Alternatively, this could be explained by a high turnover rate of miRNAs in the cnidarian lineage, in contrast to miRNA evolution in bilaterians, where losses of miRNA families are suggested to be rare (Wheeler et al. 2009). Interestingly, despite differences in sequencing depth, small RNA

Table 2. Summary of RLM-RACE results in Nematostella

<table>
<thead>
<tr>
<th>microRNA</th>
<th>Target gene</th>
<th>Genomic coordinates of miRNA site</th>
<th># Clones positive/total</th>
<th>Cleavage site transcript region</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-2022</td>
<td>NVE18870/NR2</td>
<td>scaffold_499: 72404–72425</td>
<td>9 of 10</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-2022</td>
<td>NVE16498</td>
<td>scaffold_171: 181267–181286</td>
<td>7 of 8</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-2024</td>
<td>NVE26059</td>
<td>scaffold_98: 36367–36408</td>
<td>10 of 10</td>
<td>5’ UTR</td>
</tr>
<tr>
<td>miR-2024</td>
<td>NVE11132</td>
<td>scaffold_257: 262069–262088</td>
<td>7 of 8</td>
<td>3’ UTR</td>
</tr>
<tr>
<td>miR-2025</td>
<td>NVE12346/Six3/6</td>
<td>scaffold_286: 145620–145641</td>
<td>6 of 8</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-2025</td>
<td>NVE8472</td>
<td>scaffold_201: 263493–263512</td>
<td>5 of 8</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-2026</td>
<td>NVE21156/HoxD isoform 1</td>
<td>scaffold_61: 677442–677462</td>
<td>7 of 7</td>
<td>3’ UTR</td>
</tr>
<tr>
<td>miR-2026</td>
<td>NVE21279/HoxD isoform 3</td>
<td>scaffold_61: 681906–681926</td>
<td>3 of 4</td>
<td>3’ UTR</td>
</tr>
<tr>
<td>miR-2028</td>
<td>NVE20913</td>
<td>scaffold_63: 333111–333131</td>
<td>2 of 6</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-2029</td>
<td>NVE6251/vasa2</td>
<td>scaffold_17: 1521160–1521178</td>
<td>9 of 10</td>
<td>3’ UTR</td>
</tr>
<tr>
<td>miR-2030</td>
<td>NVE19315</td>
<td>scaffold_507: 91565–91586</td>
<td>3 of 6</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-9414</td>
<td>NVE17823</td>
<td>scaffold_45: 200459–200478</td>
<td>6 of 6</td>
<td>CDS</td>
</tr>
<tr>
<td>miR-9426</td>
<td>NVE947</td>
<td>scaffold_106: 163363–163384</td>
<td>1 of 4</td>
<td>5’ UTR</td>
</tr>
</tbody>
</table>

The table presents miRNA targets supported by RLM-RACE analysis, the genomic coordinates of the miRNA/mRNA match, the number of clones supporting the cleavage events of miRNAs, and the location of the match in the target transcript. Note that the miR-2024 variants target a family of genes, all highly similar to NVE11132 and which include multiple binding sites for miR-2024.
In recent years, it was shown that \textit{Nematostella} has gene families, gene structure, and genome architecture surprisingly reminiscent of those of vertebrates and slowly evolving protostomes (Kusserow et al. 2005; Technau et al. 2005; Putnam et al. 2007). Moreover, \textit{cis}-regulation of transcription is very similar in \textit{Nematostella} and bilaterians (Schwaiger et al. 2014). These findings are in striking contrast to the vastly different body plans and cell type compositions of Cnidaria and Bilateria.

An attractive hypothesis is that differences in post-transcriptional regulation may explain this apparent contradiction. Our findings suggest that miRNAs play substantially different regulatory roles in these two groups as the miRNA-directed target cleavage (slicing) appears far more common in cnidarians than in bilaterians, where only very few cases have been reported (Shin et al. 2010). For example, a recent comprehensive study employed degradome sequencing in the nematode \textit{C. elegans} and, strikingly, found only a single case of miRNA-directed slicing (Park et al. 2013). Slicing requires extensive base-pairing between a miRNA and its target and generally has a stronger effect on target levels, but also greatly limits the number of possible targets. Indeed, each plant miRNA is known to have only a handful of targets compared to hundreds of potential targets modulated by a single animal miRNA, as well as several miRNAs targeting the same mRNA (Bartel 2009; Axtell et al. 2011). Moreover, translational inhibition is usually less potent than target slicing (e.g., Mulkopardov et al. 2012). Such differences can considerably alter the role of miRNAs in regulatory gene networks. However, the finding of many miRNA-directed slicing events in \textit{Nematostella} does not mean that bilaterian-like interactions between the miRNA seed and other transcript targets do not occur in parallel. Moreover, extended

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Regulation of miRNA targets and target conservation. (A) Expression of microRNAs and their targets as determined by double in situ hybridization. (B) A scheme showing the structure of the partially duplicated \textit{HoxD} gene and the position of the two binding sites of miR-2026. (C) The duplicated 3' UTRs of \textit{HoxD} in \textit{Nematostella} (Nv) are very derived, yet the two binding sites of miR-2026 are highly conserved, and both transcript variants are cleaved. The site is also conserved in the sea anemone \textit{Metridium senile} (Ms). Similarly, the miR-2025 binding site in the coding sequence of the \textit{Six3/6} transcript is conserved in \textit{Nematostella}, \textit{Anemonia viridis} (A), and \textit{Anthopleura japonica} (A). In panels B and C, cleavage positions are indicated by an arrow, and ratios of positive clones appear in brackets. (D) Schematic representation of coherent and incoherent circuits involving a transcription factor (TF) that regulates the expression of a miRNA and its targets and the degree of overlap between the expression domains of a miRNA and its targets that would indicate each of the two topologies. (E) miR-2030 and its host/target gene NVE19315 represent a clear case of incoherent regulatory circuit. The precursor of miR-2030 is located in an intron of NVE19315 as indicated by the green arrow. miR-2030 targets two sites in the NVE19315 transcript (indicated by magenta arrows) as supported by the degradome analysis (Site 1) and RLM-RACE (Site 2; cleavage position is indicated by an arrow, and ratios of positive clones appear in brackets).
\end{figure}
and miR-2030, represent a topology where the miRNA is embedded within the intron of, and cotranscribed with, the gene encoding the target, which is sliced with the guidance of the miRNA (Fig. 6; Supplemental Fig. S5). Strikingly, the exact same topology was shown experimentally to confer robustness of protein concentration to variation in transcription levels in synthetic genetic circuits (Blier et al. 2011). These observations suggest that the incoherently regulated targets we detected in *Nematostella* are genes whose product amounts should be constant and resistant to fluctuation in transcription levels.

miRNAs are thought to have evolved independently in plants and animals because of the different biogenesis, the lack of miRNA sequence homology and, as discussed above, the different mode of action (Axtell et al. 2011). However, cnidarians and other non-bilaterians have recently been shown to possess a homolog of HYL1, previously considered to be a plant-specific protein required for the miRNA biogenesis (Moran et al. 2013b). Thus, HYL1 was present in the common ancestor of plants and animals and was lost in Bilateria. As in plants, a large fraction of miRNAs in *Nematostella* is methylated, presumably by HEN1. While *Nematostella* shares only one miRNA (miR-100) with Bilateria, another *Nematostella* miRNA has significant sequence identity to miR-156 (Fasta E-value = 0.0094), a miRNA conserved from mosses to higher plants (Arazi et al. 2005). The *Nematostella* miRNA is identical to *Arabidopsis* miR-156a in 16 of its positions, including the seed sequence (Fig. 6). While miRNAs are so short that such average sequence identity to a much larger gene would be expected by chance alone, these findings suggest that miRNAs in plants and animals might have a common ancestry and that miRNAs in the common ancestor of the two kingdoms acted via slicing. miRNA-mediated slicing recalls the siRNA slicing mechanism, which is widespread among many eukaryotic lineages and this implies that slicing is the ancestral mechanism of miRNAs. However, the alternative hypothesis that the miRNA pathways and slicing originated convergently from siRNA mechanisms in the plant and cnidarian lineages cannot be rejected at this point.

In sum, our findings suggest that (1) the diversity of miRNAs in Cnidaria is comparable to that in some bilaterians, (2) miRNA composition is vastly different between different cnidarian groups, (3) miRNAs may play regulatory roles in the development of Cnidaria, (4) a large proportion of *Nematostella* miRNAs direct slicing of their mRNA targets, a mechanism with consequences different from the primary mode of action of bilaterian miRNAs, and (5) miRNAs participate in regulatory circuits of two opposite topologies in Cnidaria, including one that is rare in bilaterians. We propose that differences in post-transcriptional regulation like those reported here may contribute to the evident phenotypic differences between Cnidaria and Bilateria.

**Methods**

**Sea anemone culture**

*N. vectensis* was cultured under lab conditions (15 Promille sea water at 18°C) and spawning was induced as described (Genikhovich and Technau 2009b). The developmental stages studied were unfertilized eggs, blastula, gastrula, early planula, late planula, metamorphosing planula, primary polyps, adult male, and adult female.

**High-throughput sequencing**

Size-selection, oxidation, and sequencing were performed as described (Ghildiyal et al. 2008; Seitz et al. 2008). The resulting libraries were sequenced at 2 pM on each Illumina lane (UMass Medical School Deep Sequencing Core). Read counts were normalized to the library sequencing depth, excluding abundant noncoding RNAs (rRNAs, tRNAs, and snRNAs). Abundant noncoding RNAs (rRNAs, tRNAs, snRNAs, and snoRNAs) were annotated to the *N. vectensis* genome using BLASTN (version 2.2.18) to identify homologs of (1) *Drosophila melanogaster* annotated rRNAs, tRNAs, and snoRNAs (downloaded from FlyBase: http://flybase.org/ and GenBank: http://www.ncbi.nlm.nih.gov/Genbank/index.html), (2) the most abundant rRNA variants found in previous studies involving sequencing of fly small RNAs (Ghildiyal et al. 2008; Seitz et al. 2008), (3) *Homo sapiens* rRNAs, snoRNAs, and snoRNAs (downloaded from GenBank, and the LBME human snoRNA database: http://www-snorna.biotoul.fr/), and (4) the complete set of tRNAs found in the Genomic tRNA database (http://lowelab.ucsc.edu/GtRNAdb/). *N. vectensis* genomic sequences with a significant BLAST hit to those databases (E-value < 10⁻⁴) were flagged as "abundant noncoding RNAs," and deep-sequencing reads showing perfect identity to one of those were discarded.

**miRNA and siRNA annotation**

Sequencing reads 20–23 nt long were mapped to the *Nematostella* genome using an exhaustive search allowing no mismatches and every 201-nt segment (centered on the first nucleotide of the read) was extracted. Stable RNA hairpins from these loci were identified using a sliding 58-nt window to scan the 201-nt segments, selecting those predicted to fold into unbranched hairpins with an RNAfold predicted ΔG ≤ -14 kcal·mol⁻¹ (https://www.tbi.univie.ac.at/~ronny/RNA/RNAfold.html), parameters that optimally identify known pre-miRNAs in bilateral animals. For each hairpin, we identified the nucleotide where read abundance reaches 25% of the maximal read abundance of that hairpin and the nucleotide where it reaches 95%. Small RNA 5’ ends were annotated as “homogeneous” if the distance between these two nucleotides was not larger than 1 nt. Small RNAs with homogeneous 5’ ends and derived from an arm rather than the loop or the flanks of the hairpin were retained. Conserved pre-miRNA genes were identified using BLAST with the 201-nt genomic segment as query, then screened for conserved secondary structure, conserved miRNA sequence, and perfectly conserved seed sequence. miRNAs exhibiting

**Figure 6.** *Nematostella* miR-9422 exhibits significant sequence identity to the plant miR-156 family. Positions identical to the *Nematostella* miRNA are indicated by a black background. Abbreviations of species names are *Ath*, *Arabidopsis thaliana* (thale cress); *Nve*, *Nematostella vectensis*; *Osa*, *Oryza sativa* (rice); *Ppt*, *Physcomitrella patens* (moss); *Smo*, *Selaginella moellendorffii* (spikemoss).
the same seed sequence (from nt 2 to nt 8) were grouped in families. To detect conserved miRNA genes in Cnidaria, *Nematostella* pre-miRNA sequences were aligned on the *A. digitifera* genome and the *H. magnipapillata* genome (both available *Hydra* genome assemblies, ABRM01 and ACZU01, gave the same result) using a low-stringency BLAST search (word size = 4, E-value cutoff = 10), followed by a stringent conservation screen for hairpin sequence and secondary structure. Identified orthologs of *g2–g13* counting double (scoremismatch cutoff = 10) were called “undetermined.” miRNAs for which the ratio was <1 in all libraries were classified as undermethylated, and miRNAs for which the ratio was ≥1 in at least one comparison were annotated as overmethylated.

**Transcriptome data sets**

We used the following data sets for all analyses. *H. sapiens*: All miRNA from NCBI RefSeq (accessed 01-08-2013). *D. melanogaster*: FlyBase release 5.46 (FB2012_04, dated 2012-07-06). *A. thaliana*: TAIR 10 gene models (ftp://ftp.arabidopsis.org/home/tair/Genes/TAIR10_genome_release/TAIR10_gff3). *N. vectensis*: NvGeneR2.0 (http://www.cnidariangenomes.org/). Any precursor miRNA transcripts were removed from these sequence sets by intersection with miRNA annotation prior to target search.

**Target site prediction**

Mature miRNAs and shuffled sequences were mapped to the transcriptome and scored as described (Fahlgen et al. 2007). Briefly, we mapped sequences with FASTA v36 (Pearson and Lipman 1988) using the parameters -n -H -Q -f -16 -r +15/-10 -g 1000 using the 50-nt read protocol (Illumina), and two biological products 250–500 bp were purified for sequencing on the HiSeq 2000 using alternative miRNA sets for human (used for Fig. 4; Meunier et al. 2013) and *Nematostella* miRNAs from miRBase 19 (not shown); and second, by partitioning the *Nematostella* miRNAs and the miRBase miRNAs by relative abundance level in each species (Supplemental Fig. S4).

**Methylation levels of miRNAs**

For each miRNA, the ratio of normalized read count in oxidized and untreated libraries was calculated for each sample, requiring a minimum of 10 reads in each sample. If the abundance requirement was not met in any comparison, the miRNA was labeled “undetermined.” miRNAs for which the ratio was <1 in all libraries were classified as undermethylated, and miRNAs for which the ratio was ≥1 in at least one comparison were annotated as overmethylated.

**miRNA target site conservation**

We downloaded ESTs from *Metridium senile* and *Anemonea viridis* from NCBI dbEST (accessed 11-15-2013), removed sequence redundancy by clustering using cd-hit-est-2d (Li and Godzik 2006) with default parameters, and obtained orthologous groups with *Nematostella* genes by pairwise bidirectional BLAST. We identified conserved miRNA target sites using the same *Nematostella* miRNA target prediction method as elsewhere to find matches to *Nematostella* miRNAs indicative of putative target sites in the nonredundant EST sets. Finally, we intersected ortholog information and predicted targets to obtain the final list of three conserved sites described in the text.

**Degradome library construction, mapping, and analysis**

Total RNA (150 μg) was extracted from primary polyps with TRIzol (Life Technologies), then purified using poly(A) selection (Dynabead miRNA Purification Kit; Life Technologies). Purified RNA was ligated to an RNA linker (5'-HO-GUUCAGAGUUCUA CAGUGCCAGGAUC-3') with T4 RNA ligase (Life Technologies) at 20°C for 3 h. The ligated RNA was isolated using Agencourt RNAclean XP beads and amplified by PCR (SuperScript III; Invitrogen) using a degenerate primer (5'-GCA CCGGAGAATTCGACAGAATTCCCA-3'). First-strand cDNA was purified with Agencourt RNAclean XP beads and amplified by PCR using primers 5'-CTACAGTTCAGATCTACAGTCCCA-3' and 5'-GCTTGGCCACCCCGGAATTTCCA-3' (five PCR cycles), then 5'-AAGATACGGCCGACCCCGGAATTTCCA-3' and 5'-CAAGCAGAAGAAGCGATATCGGATAATN6 GTGACTGCTTTTCCGCCGAGATTCCCA-3' (N6 indicates the 6-nucleotide barcode for multiplexing; 13 PCR cycles). PCR products 250–500 bp were purified for sequencing on the HiSeq 2000 using the 50-nucleotide read protocol (Illumina), and two biological replicates were sequenced.

Degradome reads were trimmed to remove the adapter sequence and then mapped to the nonredundant transcriptome using BWA 0.5.9 (Li and Durbin 2009). Two data sets were generated: reads aligning to the sense strand with no mismatches (52,301,385 and 49,642,119 reads in each library); and reads aligned to the sense strand with ≥6 consecutive 5’ end matches, a total alignment length ≥30 bp, and mapping quality ≥20
(8,761,317 and 8,440,491 reads, respectively). For each set, we compared the starting position of degradome reads on transcript sequences with the predicted miRNA-guided cleavage sites. For all statistical analyses, we required a minimum of two reads from each library starting at a predicted cleavage site to regard it as supported by the degradome. In addition, we mapped reads to the genome using the same procedure as detailed above for visualization purposes.

In situ hybridization

In situ hybridization (ISH) of RNA probes to detect pri-miRNA was performed as described (Genikhovich and Technau 2009a). Locked nucleic acid (LNA) probes labeled on both 5′ and 3′ ends with digoxygenin (Exiqon) were used to detect mature miRNAs. LNA ISH was performed as described except that the hybridization buffer contained 70% (v/v) formamide and 5% (w/v) buffered dextran sulfate (Sigma-Aldrich), and hybridization was at a temperature 30°C below the calculated Tm of that probe on a DNA substrate. LNA probe concentration (1–25 nM) was optimized experimentally. Simultaneous detection of pri-miRNAs and target mRNAs was as described (Genikhovich and Technau 2009a), whereas simultaneous detection of mature miRNAs and targets was performed under miRNA conditions to permit efficient hybridization of the LNA probe. Double ISH probes for miRNAs were labeled with fluorescein isothiocyanate (Roche), and miRNA probes were labeled with digoxygenin (Genikhovich and Technau 2009a). Probes were detected sequentially using alkaline phosphatase (AP) coupled to anti-fluorescein and anti-digoxygenin Fab antibody fragments (Roche) and nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP; Roche) and Fast Red (Sigma-Aldridge) as described (Moran et al. 2013a). Staining was documented with an Eclipse 80i fluorescence microscope and Digital Sight DS-U2 camera (Nikon).

RLM-RACE

The RLM-RACE procedure was performed as described (German et al. 2008) using the FirstChoice RLM-RACE kit (Life Technologies) according to the manufacturer’s instructions. Ligated RNA was reverse transcribed (Superscript III; Life Technologies) according to the manufacturer’s instructions using gene-specific primers at 55°C for 60 min. First-strand cDNAs were used as template for PCR using a nested gene-specific primer, the outer 5′ RACE primer from the FirstChoice RLM-RACE kit, and Advantage 2 polymerase mix (Clontech) with five cycles (30 sec at 94°C, 12 sec at 72°C), five cycles (30 sec at 94°C, 20 sec at 70°C, 12 sec at 72°C), 27 cycles (30 sec at 94°C, 20 sec at 68°C, 12 sec at 72°C), and then 5 min at 72°C. The PCR product was diluted 1:1000 and used as template for a nested PCR reaction with a nested, gene-specific primer, the outer 5′ RACE primer from the FirstChoice RLM-RACE kit and Advantage 2 polymerase mix. The reaction conditions were as follows: 30 cycles (30 sec at 94°C, 20 sec at 68°C, 20 sec at 72°C), with a final elongation of 5 min at 72°C. PCR products were cloned into the pGEM-T Easy vector (Promega), and clones were sequenced (Macrogen). Clones were scored positively only when the gene-specific sequence following the adapter started in a position across position 10 of the miRNA.

Data access

Deep-sequencing libraries have been submitted to the NCBI Sequence Read Archive (SRA; http://www.ncbi.nlm.nih.gov/sra) under accession number SRP000409.

Acknowledgments

This work was funded by a Marie Curie IEF fellowship and EMBO postdoctoral fellowship ALTF 1096-2009 to Y.M., FWF P22618-B17 grant to U.T., and HFSP Career Development Award #CD00017/2010-C and a CNRS ATIP-Avenir grant to H.S.

References


Received June 24, 2013; accepted in revised form January 22, 2014.
Cnidarian microRNAs frequently regulate targets by cleavage

Yehu Moran, David Fredman, Daniela Praher, et al.

*Genome Res.* 2014 24: 651-663 originally published online March 18, 2014
Access the most recent version at doi:10.1101/gr.162503.113

**Supplemental Material**
http://genome.cshlp.org/content/suppl/2014/02/04/gr.162503.113.DC1.html

**References**
This article cites 78 articles, 33 of which can be accessed free at:
http://genome.cshlp.org/content/24/4/651.full.html#ref-list-1

**Open Access**
Freely available online through the *Genome Research* Open Access option.

**Creative Commons License**
This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at http://creativecommons.org/licenses/by-nc/3.0/.

**Email Alerting Service**
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here.](#)