RESEARCH ARTICLE

The Identification of MicroRNAs in Calcisponges: Independent Evolution of MicroRNAs in Basal Metazoans

JOURNAL OF EXPERIMENTAL ZOOLOGY
PART B
MOLECULAR AND DEVELOPMENTAL EVOLUTION

JEFFREY M. ROBINSON[1], ERIK A. SPERLING[2], BRITH BERGUM[3], MARCIN ADAMSKI[3], SCOTT A. NICHOLS[4], MAJA ADAMSKA[3], AND KEVIN J. PETERSON[1]*

[1]Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire
[2]Department of Geology and Geophysics, Yale University, New Haven, Massachusetts
[3]Sars International Centre for Marine Molecular Biology, Bergen, Norway
[4]Department of Biological Sciences, University of Denver, Denver, Colorado

ABSTRACT

We present the discovery of microRNAs (miRNAs) in the calcisponges *Sycon* and *Leucosolenia* (phylum Calcarea), and potential miRNAs in the homoscleromorph *Oscarella carmela* (Phylum Homoscleromorpha), expanding the complement of poriferan miRNAs previously known only from the siliceous sponges (demosponges and hexactinellids). Comparison of these miRNAs with those previously described from silicisponges and eumetazoans reveals that these newly described miRNAs are novel, with each metazoan lineage (Silicea, Calcarea, Homoscleromorpha, and Eumetazoa) characterized by a unique and non-overlapping repertoire of miRNAs (or potential miRNAs as in the case of the homoscleromorphs). Because each group is characterized by a unique repertoire of miRNAs, miRNAs cannot be used to help resolve the contentious issue of sponge mono- versus paraphyly. Further, because all sponges are characterized by a similar repertoire of tissue types and body plan organisation, we hypothesize that the lack of conserved miRNAs amongst the three primary sponge lineages is evidence that cellular differentiation and cell type specificity in sponges are not dependent upon conserved miRNAs, contrary to many known cases in eumetazoans. Finally, we suggest that miRNAs evolved multiple times independently not only among eukaryotes, but even within animals, independently evolved miRNAs representing molecular exaptations of RNAi machinery into pre-existing gene regulatory networks. The role(s) miRNAs play though in sponge biology and evolution remains an open question. *J. Exp. Zool. (Mol. Dev. Evol.) 320B:84–93, 2013.* © 2013 Wiley Periodicals, Inc.

J. Exp. Zool. (Mol. Dev. Evol.) 320B:84–93, 2013

How to cite this article: Robinson JM, Sperling EA, Bergum B, Adamski M, Nichols SA, Adamska M, Peterson KJ. 2013. The identification of microRNAs in calcisponges: Independent evolution of microRNAs in basal metazoans. J. Exp. Zool. (Mol. Dev. Evol.) 320B:84–93.

Sponges represent the oldest (Love et al., 2009; Sperling et al., 2010; Brain et al., 2012) and most basal metazoan taxon (Philippe et al., 2009; Pick et al., 2010), and are therefore in a key phylogenetic position for investigations into the origin of complex tissues and cellular/developmental processes within the animal kingdom (Srivastava et al., 2010; Adamska et al., 2011). However, their potential to elucidate myriad cellular, developmental, and evolutionary problems is mitigated by the fact that their phylogenetic status as a "natural" or monophyletic group is

Grant sponsor: NASA Astrobiology Institute; Grant sponsor: Sars Centre; Grant sponsor: Agouron Geobiology Fellowship.

The present address of Erik A. Sperling is Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA

*Correspondence to: Kevin J. Peterson, Department of Biological Sciences, Hanover, NH 03755. E-mail: kevin.j.peterson@dartmouth.edu

Received 30 July 2012; Revised 29 October 2012; Accepted 17 December 2012

Published online 24 January 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/jez.b.22485

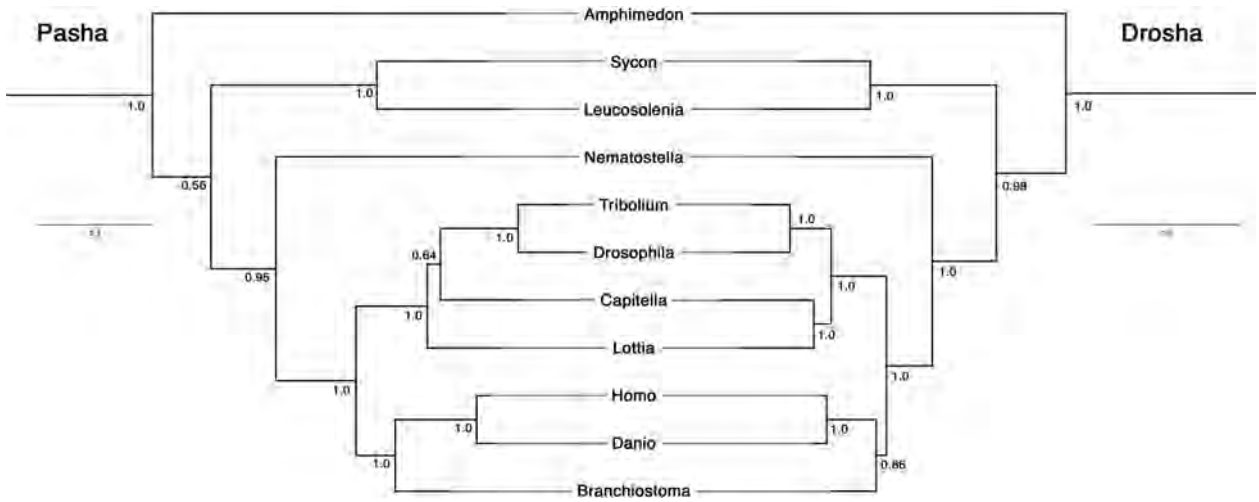

Figure 1. Bayesian phylogenetic analysis of metazoan Pasha (left) and Drosha (right) orthologs rooted on respective family members (see text). Branch lengths (black lines) are shown as a phylogram (see scales). Posterior probabilities are shown at the nodes. Note that there is unequivocal support for the identification of Drosha and Pasha orthologs in the two calcisponges examined, *Sycon* and *Leucosolenia*.

contentious. Since the mid-1990s various molecular phylogenetic studies have suggested that sponges are a paraphyletic grade at the base of the Metazoa (e.g., Cavalier-Smith et al., '96; Collins, '98; Borchiellini et al., 2001; Manuel et al., 2003; Sperling et al., 2009), whereas other molecular studies (e.g., Philippe et al., 2009; Pick et al., 2010) as well as morphological cladistics analyses (e.g., Zrzavy et al., '98; Peterson and Eernisse, 2001) are consistent with sponge monophyly.

Resolution of this mono- versus paraphyly question is essential in order to polarize characters at the base of the Metazoa (Sperling et al., 2007). For example, if sponges are paraphyletic, then this is strong evidence that the last common ancestor (LCA) of all living animals was constructed like a modern sponge complete with a water canal system, and thus this LCA was a benthic micro-suspension feeder. If, however, sponges are monophyletic, then it is difficult to specify much about the biology of the metazoan LCA beyond that it most likely possessed choanocytes given that this cell type is present in unicellular outgroups (Nichols et al., 2009). Sponge monophyly is strongly supported with large-scale EST data sets (Philippe et al., 2009; Pick et al., 2010). However, Sperling et al. (2009) showed that although they were able to recover sponge monophyly, it was only possible under unrealistic evolutionary models, if poorly performing phylogenetic methods were used, or in situations where the potential for the generation of tree reconstruction artifacts was artificially exacerbated. Thus, it remains unclear whether sponges are a monophyletic clade or a paraphyletic grade.

One method of distinguishing between historical signal versus phylogenetic artifact is to change data sets and reanalyze the problem. MicroRNAs (miRNAs) provide one such alternative phylogenetic dataset. miRNAs (for reviews see Bartel, 2004, 2009; Carthew and Sontheimer, 2009; Berezikov, 2011; Huntzinger and Izaurralde, 2011; Starega-Roslan et al., 2011) are a class of endogenously transcribed small RNAs ~22 nt long that play roles in many cellular process through the transcriptional regulation of messenger RNA targets. Eumetazoan miRNAs are transcribed from either intergenic regions or from introns as a long primary transcript (pri-miRNA). Because of base pair complementarity, the pri-miRNA transcript folds into a characteristic hairpin-like structure (see Figs. 2 and 3), which is recognized by an enzyme complex including the proteins Drosha and Pasha. These enzymes form the core of the microprocessor, which cleaves the pri-miRNA into a ~70 nucleotide precursor miRNA (pre-miRNA) that is then exported into the cytoplasm where it is further processed by the RNaseIII enzyme Dicer to form a 22 ± 2 nt long RNA duplex. This duplex then separates into two distinct strands and the mature gene product(s) are loaded into an Argonaut-protein containing complex, which then regulates target mRNA(s). This regulation of mRNAs by miRNAs is primarily effected through mRNA de-adenylation and posttranslational repression, and is achieved through non-perfect complementary base-pairing to the 3′ UTR of mRNAs.

In eumetazoans, miRNAs function in the regulation of numerous cellular processes including pluripotency and differentiation (Li and He, 2012). Many miRNAs exhibit phylogenetically conserved, tissue specific expression (Christodoulou et al., 2010; Heimberg et al., 2010). miRNAs also notably function in stress response (Mendell and Olson, 2012), and perform a role in

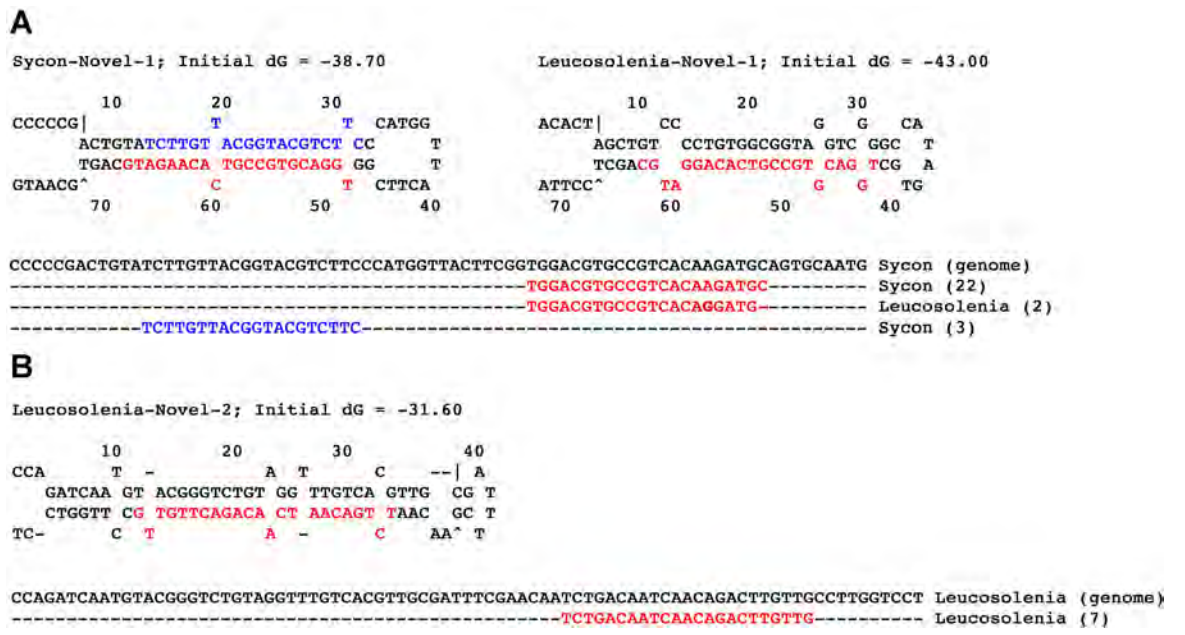

Figure 2. MicroRNAs in calcisponges. **A**: *Sycon* and *Leucosolenia* Novel-1 (note: the proper miRBase numerical designation will be assigned shortly after this article goes to publication). On the top are the putative precursor sequences with the cloned mature sequences in red, and the star sequence cloned in *Sycon* shown in blue (a star sequence in *Leucosolenia* was not cloned in our 454 library). Note that all the hallmarks of bona fide miRNAs are apparent (Ambros et al., 2003; Kozomara and Griffiths-Jones, 2011; Tarver et al., 2012) including the requisite number of base pairs between the mature and star sequences (≥16), low delta G (<∼20 kcal/mol), and two nucleotide offset between the mature and star, indicative of two sequential RNaseIII cuts (i.e., Drosha and Dicer). Shown at the bottom are the reads found in our library mapped to the Novel-1 precursor sequence. Note the 5′ homogeneity of the mature sequence (see Kozomara and Griffiths-Jones, 2011; Tarver et al., 2012 for discussion and counter examples). **B**: *Leucosolenia* putative Novel-2. Although this miRNA shows the requisite structural and 5′ homogeneity requirements, a star was not cloned so the two nucleotide offset cannot be evaluated, and as it was not found in *Sycon* phylogenetic conservation cannot be used to establish it as a *bona fide* miRNA.

conferring robustness to gene expression (Hornstein and Shomron, 2006; Herranz and Cohen, 2010; Ebert and Sharp, 2012). miRNAs have therefore been hypothesized to contribute to the evolution of new cell types and ultimately to morphological complexity (Peterson et al., 2009; Kosik, 2010; Berezikov, 2011). Most major metazoan taxa have been sampled for their respective miRNA repertoires (Grimson et al., 2008; Wheeler et al., 2009), and an interesting observation from these studies is that there is no overlap between the miRNAs found in silicisponges and those found in eumetazoans. The demosponge *Amphimedon queenslandica* has only eight miRNAs (Grimson et al., 2008) and all eight of these miRNAs were present in the LCA with the democlavid demosponges (e.g., *Suberites*), with one present in the hexactinellids *Aphrocallistes* and *Rhabdocalyptus*, supporting the monophyly of the Silicea (Sperling et al., 2010). Hence, these eight miRNAs had evolved by ∼650 Ma ago, and no additional miRNAs have evolved in the *A. queenslandica* lineage since the Cryogenian. Eumetazoans, on the other hand, show continual addition of highly conserved miRNAs, with the rate of acquisition broadly correlated with relative morphological complexity (Hertel et al., 2006; Sempere et al., 2006; Heimberg et al., 2008; Peterson et al., 2009; Wheeler et al., 2009; Berezikov, 2011). None of the thousands of miRNAs thus far discovered in eumetazoans are present in the genome *A. queenslandica* and none of the eight silicisponge-specific miRNAs have been described in any eumetazoan (or any other eukaryotic group for that matter).

This non-overlap of miRNA complements between silicisponges and eumetazoans mirrors the situation found in the plant lineage given that the complements between the chlorophyte green algae (*Chlamydomonas*) and embryophyte land plants (e.g., *Arabidopsis*) are distinct. Indeed, given our current knowledge it appears that there are five distinct repertoires of miRNAs in eukaryotes: Eumetazoa, Silicispongia, Chlorophyta, Embryophyta, and Chromalveolata (the kelp *Ectocarpus siliculosus*) (Tarver et al., 2012). What is of potential phylogenetic and biological interest is whether the silicisponge and/or the eumetazoan miRNA repertoires are

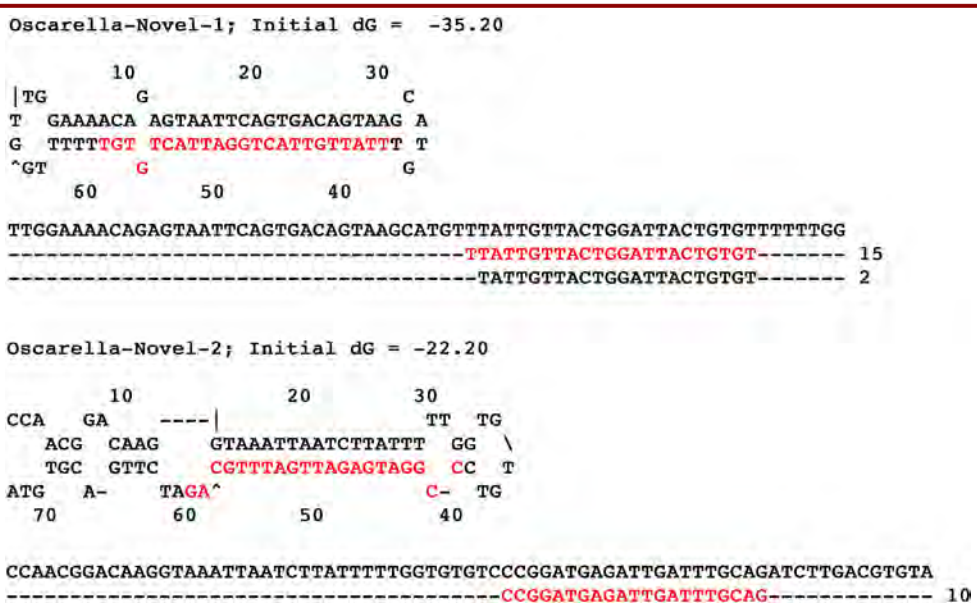

Figure 3. Putative miRNAs from the homoscleromorph *Oscarella carmela*. Similar to *Leucosolenia*-Novel-2, these two miRNAs show the requisite structural and 5′ homogeneity requirements, but again because star sequences were not cloned, the two nucleotide offset requirement cannot be evaluated, and neither were found in the *Corticium* sp. or *Plakinestrella* sp. small RNA libraries.

present in the other groups of sponges, the calcisponges and/or the homoscleromorphs. All sponges are built on a shared body plan, consisting principally of the water-canal system, and a minimal complement of histologically identifiable cell types (in particular choanocytes and pinacocytes) organized into the pinacoderm and choanoderm, and mesohyle containing an extracellular matrix of spongin, collagen, and often spicules made of either silicia (silicosponges and homoscleromorphs) or calcium carbonate (calcisponges) produced by various cell types of uncertain developmental origin (Simpson, '84; Boury-Esnault and Rutzler, '97). Given the morphological and cellular similarities among these three sponge groups, according to the miRNA hypothesis of cellular and morphological novelty one might predict that at least one, if not more, silicisponge-"specific" miRNAs would be found in these other sponge groups, especially if sponges are monophyletic. On the other hand, if sponges are paraphyletic, then it is possible that calcisponges and/or homoscleromorphs possess one or more eumetazoan-specific miRNAs. Of course it is also possible that like other groups of eukaryotes, these sponge groups have their own unique repertoire(s) of miRNAs, have a combination of both silicisponge- and eumetazoan-specific miRNAs, or even lack miRNAs all together.

In conjunction with the genome sequencing projects of the homoscleromorph sponge *Oscarella carmela* (Nichols et al., 2012) and the calcaronean calcisponges *Sycon ciliatum* and *Leucosolenia complicata*, we sampled and sequenced the small-RNA complement of *O. carmela*, *S. ciliatum* and a second species/sample of *Sycon*, a second species/sample of *Leucosolenia*, and two additional homoscleromorph taxa, *Corticium* sp. and *Plakinestrella* sp. In addition to showing that calcisponges possess at least the key enzymes for processing miRNAs in the nucleus (Drosha and Pasha), we find that the two calcisponge genera share a single miRNA. We also show that *Leucosolenia* might possess a second miRNA. These miRNAs are novel and not present in any other metazoan genome or small RNA library sequenced to date. Further, these two miRNAs were not found in any of the homoscleromorph samples. Instead, two potential miRNAs were found in *Oscarella*, but none of these two were present in the two other homoscleromorph taxa. Because each of the metazoan miRNA repertoires are unique, miRNAs cannot help resolve the poriferan mono- versus paraphyly problem, or the relationships between the sponge classes. We argue that despite sharing a similar body plan, conserved cell types and morphological structures, at best each of the three sponge groups possess unique repertoires of miRNAs, and thus poriferan miRNAs do not function as conserved factors in establishing conserved cell types and morphological structures. Finally, our data suggest that there are multiple parallel inventions of miRNAs in animals.

MATERIALS AND METHODS

Drosha/Pasha Phylogenetics

The Pasha amino acid sequence from *A. queenslandica* was obtained from the supplementary material of Grimson et al.

(2008); the C-terminal RNAseIII domain region (residues 414–781, accession XP_003387755) of Drosha was obtained by a BLAST search of the NCBI database. *Drosophila melanogaster*, *Tribolium castaneum*, *Danio rerio*, and *Homo sapiens* sequences were obtained from the NCBI Genbank database. *Capitella teleta*, *Lottia gigantea*, *Branchiostoma floridae*, and *Nematostella vectensis* orthologous protein sequences were obtained from the US Joint Genome Institute genome portal (http://genome.jgi-psf.org/) using reciprocal best BLAST hit method with fly and human orthologs as reference sequences. *Sycon* and *Leucosolenia* sequences were predicted from the *Sycon* and *Leucosolenia* genome and transcriptome sequencing projects. Sequences were aligned using the MUSCLE algorithm (http://www.ebi.ac.uk/Tools/msa/muscle), and manually edited using JalView 2.6.1. The alignment is available upon request.

Bayesian phylogenetic analysis used Phylobayes (v. 3.3) (Lartillot et al., 2009) and employed the LG (Le and Gascuel, 2008) model. A consensus tree from two chains was constructed with the "maxdiff" <0.3 (Drosha = 0.066; Pasha = 0.071) with a burnin of 10,000 and a subsampling frequency of 100 for both the Drosha and Pasha trees. The Drosha tree was rooted on the human and *Arabidopsis* Dicer1 protein sequence, whereas the Pasha tree was rooted on protein sequences from two other related dsRNA binding domain family members, Loquacious and Staufen, from human and fly.

MicroRNA Libraries and Deep Sequencing

Total RNA was obtained from three calcisponges, two species of *Sycon* and *Leucosolenia* sp., and three homoscleromorphs, *O. carmela*, *Plakinestrella* sp., and *Corticium* sp. Two types of material were used to generate the *S. ciliatum* library. To obtain RNA from juveniles, adult specimens were collected from fjords located near Bergen, Norway (27′33 N60°, 56′1 W4°) in June 2009. The sponges were brought to the laboratory and kept in plastic cups until larvae were released (1–4 days). The larvae were transferred to standard Petri dishes containing filtered seawater, were they settled and metamorphosed. Juvenile sponges were collected from the dishes several days after reaching olynthus stage (an asconoid grade organization), between 1 and 2 weeks after settlement. The collection was performed under dissecting microscope, with care taken to avoid possible contamination, especially of eumetazoan origin. Approximately 8,000 juveniles were used to isolate total RNA for the miRNA library. To obtain RNA from adults, small specimens were collected in August 2009, cleaned under the dissecting scope to remove visible contaminations, and placed in RNAlater, which dissolves the skeleton and allows for further cleaning of the material before isolation of total RNA. The RNA from these two preparations was pooled to generate a single *S. ciliatum* library. Small RNAs from adults of two other species of calcisponges were also analysed - *Sycon* sp., collected at Friday Harbor Laboratories, San Juan Islands, WA and sequenced in (Sperling et al., 2010), and *Leucosolenia* sp., purchased from the Marine Biological Laboratories, Woods Hole, MA. Before processing for total RNA the adult animals were cleaned under a dissecting scope. *O. carmela* was collected from sea tables at the Joseph Long Marine Laboratory at the University of California, Santa Cruz. Adult tissue was detached from a glass aquarium surface, washed in sterile seawater and cleaned of macroscopic contaminants. Clean tissue was flash-frozen in liquid nitrogen, ground to a fine powder using a mortar and pestle, and was sequenced as reported in Sperling et al. (2010). While this tissue possibly (likely) contained contaminating organisms, including microscopic animals and other eukaryotes, the reference genome used here to validate detected miRNAs was derived from a single, contaminant free larva (Nichols et al., 2012). Specimens of *Plakinestrella* sp., and *Corticium* sp. were kindly given to us by Prof. D. Lavrov (Iowa St. University). Total RNA procurement and library construction for *Leucosolenia* sp., *Plakinastrella* sp., and *Corticium* sp. used standard procedures as described in Wheeler et al. (2009). The *S. ciliatum* RNA was sequenced following standard protocols on an Illumina Genome Analyzer II at the Yale Center for Genome Analysis; all other sponges (including *Sycon* sp.) were sequenced using 454 technology at 454 Life Sciences (Branford, CT, USA).

Total reads between 17 and 25 nt were collapsed into a set of non-redundant reads and then blasted against miRBase (v. 15) using the algorithm miRMiner (Wheeler et al., 2009). Once the libraries were filtered for both known miRNAs and non-miRNA sequences (tRNAs, rRNAs, etc.) miRMiner allows the user to take the remaining sequences and look for shared sequences between two or more taxa. Both shared reads and unique reads for *Sycon*, *Leucosolenia*, and *Oscarella*, were then blasted against the available genomic traces for these three taxa. All hits without any mismatch between the library read and the genomic fragment were then checked to see if they indeed are derived from a miRNA locus. A 200 bp fragment was extracted manually and folded using the algorithm mfold (v. 3.2; http://mobyle.pasteur.fr) using standard secondary structural information to distinguish bona fide miRNAs from other types of small RNAs (Ambros et al., 2003; Kozomara and Griffiths-Jones, 2011; Tarver et al., 2012).

RESULTS

To explore the potential miRNA repertoire of calcisponges and homoscleromorphs, we addressed both the presence of conserved nuclear processing enzymes—specifically the enzymes Drosha and Pasha, which forms the core of the nuclear processing machine for miRNAs in metazoans—and the presence of reads in small RNA libraries that emanate from hairpin sequences within the genome. To address whether calcisponges possess orthologs of the key nuclear processing enzymes Drosha and Pasha, we searched the sequenced genomes and transcriptomes of both *S. ciliatum* and *L. complicata* for each of these two genes using *Drosophila* protein sequences and BlastX search. A Bayesian phylogenetic analysis strongly suggests that both calcisponge species possess orthologs of both Pasha and Drosha (Fig. 1).

To ask if potential mature miRNA reads are expressed in these two taxa, we next constructed and sequenced small RNA libraries. Sponge tissue was cleaned and processed, total RNA procured, and the resulting small RNA libraries sequenced with either Illumina (*S. ciliatum*) or 454 (*Sycon* sp. and *Leucosolenia*) technology (see the Materials and Methods Section). The Illumina run resulted in 77,552 non-redundant reads from *S. ciliatum*, whereas 16,044 non-redundant reads were sequenced from *Sycon* sp. using 454 technology (Sperling et al., 2010). These reads were processed by miRMiner (Wheeler et al., 2009), and potential mature miRNA reads from both *Sycon* species were then mapped to the *S. ciliatum* genome. Resulting hits were folded to determine whether or not the surrounding genomic sequence meets the minimal structural requirements for miRNA annotation (Ambros et al., 2003; Kozomara and Griffiths-Jones, 2011; Tarver et al., 2012), and all reads derived from the locus mapped to the putative pre-sequences to confirm that the locus meets the minimal processing requirements for bona fide miRNAs, specifically the presence of both mature and star reads with the requisite two nucleotide offsets, and a consistent 5′ end of the mature gene product (Kozomara and Griffiths-Jones, 2011; Tarver et al., 2012).

No silicisponge-specific miRNAs were identified in either calcisponge. miRMiner did identify several eumetazoan-specific miRNAs in both the *Sycon* and *Leucosolenia* libraries (e.g., let-7, miR-184). However, none of these genes are present in the *Sycon* genome (which has been processed to eliminate contamination by labeling scaffolds of sponge origin using genomic sequences obtained from laboratory-grown, carefully inspected juveniles), and are therefore likely to be contaminants. Hence there is no evidence for the presence of either silicisponge or eumetazoan-specific miRNAs in the calcisponge genome. One sequence could confidently be identified as a miRNA, Sycon-Novel-1 (Fig. 2A). In addition, the mature sequence of Novel-1 was identified in our *Leucosolenia* library, and a precursor structure identified in the *L. complicata* genomic traces (Fig. 2A), providing further evidence for this sequence being correctly identified as a miRNA. No additional miRNA candidates were identified in *Sycon*. A second putative miRNA was discovered in *Leucosolenia*; although the mature sequence was found in our *Leucosolenia* sp. library, and a putative precursor structure found in the *L. complicata* genome traces, a star sequence was not identified (which is not that unusual in our 454 libraries given the sequencing depth) (Fig. 2B); nonetheless until a star sequence is identified or evolutionary conservation demonstrated this candidate should be considered at best a potential miRNA.

An interesting observation then is that both silicisponges and calcisponges appear to possess taxon-specific miRNAs; silicisponges and calcisponges do not share any miRNAs with one another, nor do they share miRNAs with any other eukaryotic group including eumetazoans. To ask if any of these sponge and/or eumetazoan miRNAs are shared with the third group of sponges —the homoscleromorphs—we constructed and sequenced small RNA libraries from three homoscleromorph taxa, *O. carmela* (11,564 non-redundant reads), *Plakinestrella* sp. (6,988 non-redundant reads) and *Corticium* sp. (1,705 non-redundant reads), and analysed them in conjunction with a previously sequence library from O. *carmela* (11,546 non-redundant reads; Sperling et al., 2010), using the recently sequenced genome from *O. carmela* to confirm the presence of bona fide miRNAs. Two potential miRNAs were found in *O. carmela* (Fig. 3), and although both meet the minimal structure requirements for miRNA annotation (Tarver et al., 2012), star sequences were not sequenced for either of the two candidates, and neither of the mature sequences were conserved with the other homoscleromorph species sequenced. Although a partial fragment of Pasha was found in Oscarella, it was not complete enough for phylogenetic analysis, and no orthologue of Drosha was identified in the current genomic assembly or EST data sets. Because of this, we cannot confidently state that homoscleromorphs possess miRNAs. Nonetheless, again we were not able to identify either reads or loci for any of the silicisponge- or calcisponge-specific miRNAs, and the few reads from eumetazoan-specific miRNAs all appear again to be exogenous as none are present in the genome of *O. carmela*. Thus, if they possess them at all, like the other two sponge groups homoscleromorphs possess their own distinct repertoire of miRNAs.

DISCUSSION

To date, five eukaryotic taxa are known to possess miRNAs: eumetazoans, silicisponges, vascular plants, *Chlamydomonas*, and the brown kelp *Ectocarpus*; descriptions of miRNAs in other protist groups such as *Giardia* are erroneous according to accepted criteria for identifying miRNA (Tarver et al., 2012). What is interesting is that despite the conservation of structural and processing "rules" there is no overlap in sequence among any of the miRNAs from one of these five groups to the next—each of these five groups is characterized by a unique repertoire of miRNAs (Tarver et al., 2012). Now we have added a sixth and possibly a seventh group to this list, calcisponges and homoscleromorphs as each of these two sponge groups is again characterized by a unique set (or potential set) of miRNAs (Figs. 2 and 3). It might not be surprising that there is no overlap amongst the sponge miRNA sequences with those in eumetazoans if sponges are monophyletic—indeed this might even be a prediction of the monophyly hypothesis. However, miRNAs are unable to contribute to the mono- versus paraphyly debate given that there is no overlap of miRNAs in each of the three major sponge groups (silicisponges, calcisponges, and homoscleromorphs). The demosponge *A. queenslandica* was originally described as having eight miRNAs (Grimson et al., 2008; Wheeler et al., 2009); one of these miRNAs (miR-2019) is shared across the silicisponges (hexactinellids + demosponges), with the remaining seven evolving by the time *A. queenslandica* split from the democlavid demosponges some 650 Ma ago (Sperling et al., 2010). None of these eight

miRNAs were found in any of the calcisponge species or homoscleromorph species analyzed herein, and none of the novel miRNAs or potential miRNAs reported herein are found in any of our silicisponge libraries or in the genome sequence of *A. queenslandica* (Wheeler et al., 2009; Sperling et al., 2010). Further, the calcisponge miRNAs were not found in any of our homoscleromorph data sets, which included small RNA libraries made from three different homoscleromorph species and the genomic traces of *O. carmela*. And finally, the potential miRNAs found in *O. carmela* are not found in any of our calcisponge data sets (three small RNA libraries and two genome sequences). Thus, although it is possible that we missed miRNAs(s) based on selective expression during time points not examined in our study, it remains unlikely that miRNAs are shared across sponge groups given that we have representative genomic sequences from all three major groups.

miRNA biogenesis between plants and animals is clearly convergent, with notable differences in processing machinery (reviewed in Axtell et al., 2011; Tarver et al., 2012), therefore the lack of overlap amongst miRNA loci in plants and animals is not surprising. However, the monophyly of the Metazoa presents profound similarities between sponges and eumetazoans at least in terms of their genetic repertoires (Harcet et al., 2010; Srivastava et al., 2010; Adamska et al., 2011) including deep conservation of transcription factors (Larroux et al., 2008; Degnan et al., 2009; Fortunato et al., 2012) signaling pathways (Nichols et al., 2006; Richards and Degnan, 2009, 2012; Adamska et al., 2010), cellular differentiation and pluripotency pathways (Funayama, 2010), and even the small RNA processing pathway (Grimson et al., 2008) including the presence of the key metazoan-specific microprocessor enzymes Drosha and Pasha (Fig. 1).

In the Eumetazoa, miRNAs are key players in the processes of pluripotency, differentiation, and cell cycle regulation (Ivey and Srivastava, 2010), and many of these miRNAs are deeply conserved including let7, the miR-100 family, miR-200, and miR-34. For example, the miR-100 family has conserved expression associated with the transcription factor *Brachyury* around the developing mouth in cnidarians and bilaterian invertebrates (Christodoulou et al., 2010), while let-7 and miR-200 are widely expressed in differentiated tissues in many different bilaterians and act essentially as differentiation factors, interacting with pluripotency, differentiation, and stem cell regulatory networks (Peter, 2009; Martinez and Gregory, 2010). Other deeply conserved miRNAs have conserved tissue-specific expression (Christodoulou et al., 2010), with miR-1, for example, universally expressed in bilaterian muscle tissue, a bilaterian-specific cell type (Steinmetz et al., 2012); other examples of tissue specificity include miR-9 in the bilaterian brain and miR-122 in the vertebrate liver (Heimberg et al., 2010). Therefore, expression profiles of eumetazoan miRNAs are often directly associated with cellular phenotype and tissue identity, in both normal and pathological systems (Lu et al., 2005; Kosik, 2010). This evidence taken together is indicative of a functional role for miRNAs in the evolution of novel eumetazoan cells and tissues.

This conservation of miRNA expression presumably homologous cell and tissue types in eumetazoans is in striking contrast to sponges. Sponges have approximately 12 cell types (Harrison and De Vos, '91), and there is very little difference amongst at least the choanocytes and the pinacocytes in the three groups of sponges, an observation that plays a major role in current discussions concerning sponge mono- versus paraphyly (Philippe et al., 2009; Pick et al., 2010). Thus, in contrast to what is seen in eumetazoans, sponges show conserved cell types but non-conservation of miRNAs, suggesting that the origin, or at least the maintenance, of cell types is not predicated upon the conservation of any one particular miRNA. In fact, it remains to be shown exactly what role sponge miRNAs play in sponge biology—it is possible, and even maybe likely, that sponge miRNAs are not involved in cellular differentiation, but function in some other manner distinct from that currently understood in eumetazoans.

Indeed, it is actually unclear whether "miRNAs" are homologous between sponges and eumetazoans. Grimson et al. (2008) hypothesized that miRNAs are homologous within the Metazoa based on the presence of orthologs of key microprocessor enzymes Drosha and Pasha in *A. queenslandica* (see Fig. 1); orthologs of both of these genes are absent in non-metazoan taxa. This appears to be a straightforward conclusion: ancestrally conserved genotype (Drosha and Pasha) gives rise to the "metazoan miRNA" phenotype, and thus miRNA biogenesis—and hence miRNAs—was present in the LCA of all living metazoans. Wheeler et al. (2009) though came to a different conclusion—these authors were more impressed by the non-overlap of miRNA families between demosponges and eumetazoans, and particularly the structural differences between demosponge and eumetazoan miRNAs. miRNA biogenesis processing depends on structure, not on sequence, as no sequence is universally conserved to all human pri-miRNAs (Berezikov et al., 2005). In eumetazoans, Pasha is responsible for the binding of the microprocessor complex to the pri-miRNA (primary miRNA transcript) as Pasha recognizes both the single stranded flanks and the double stranded stem, and then acts as a ruler guiding Drosha to cleave the molecule in the correct place, 11 base pairs up the stem from the single stranded RNA to double-stranded RNA junction (Han et al., 2006). What is interesting is that in most eumetazoan miRNA sequences Drosha cuts approximately 22–24 nt from the loop resulting in the mature sequence within about two nucleotides from the loop (Kim, 2005). In demosponges though, the miRNA precursor structures resemble plant miRNAs (Reinhart et al., 2002; Jones-Rhoades et al., 2006) in that the mature gene sequence is often 30 or more nucleotides away from the loop, and no miRNAs were found in sponges whose mature sequence was within 10 nt of the loop (Grimson et al., 2008; Wheeler et al., 2009). This pronounced difference in the pre-miRNA structures between demosponges and eumetazoans, led Wheeler et al. (2009) to propose that miRNAs were not

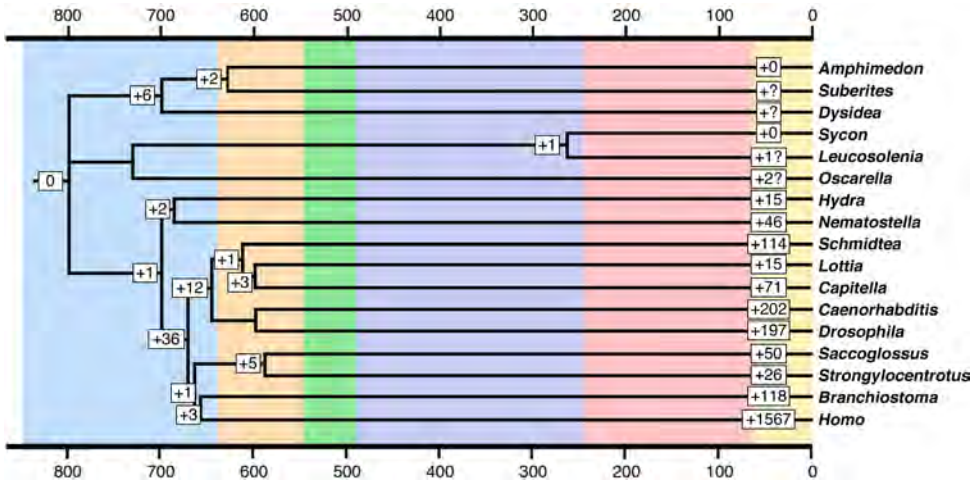

Figure 4. Evolutionary dynamics of miRNAs in metazoans. Shown are the major model systems that have been queried for their miRNA complements using miRBase (release 19). Shown in boxes are the estimated number of miRNA genes that evolved at that point in the lineage's history. Divergence times are taken from Erwin et al. (2011) with the important geological stages shown in color (Cryogeneian—Blue; Ediacaran—Brown; Cambrian—Green; the remainder of the Paleozoic—Violet; Mesozoic—Red; and Cenozoic—Yellow). Porifera is represented by a polytomy given the uncertainty regarding is mono- versus paraphyly (see text). Note the paucity of miRNAs in sponges as compared to the rest of the Metazoa.

homologous between demosponges and eumetazoans, and hence that the metazoan LCA did not possess miRNAs.

Calcisponge miRNAs, and the candidate miRNAs from *Oscarella*, show eumetazoan pre-like structures with the mature gene sequences within ~2 nucleotides from the loop (Figs. 2 and 3), although of course the sample size is very small. Nonetheless, that the few miRNAs discovered resemble virtually all known miRNA pre-miRNA structures from eumetazoans and none of the known demosponge pre-miRNAs suggests the possibility that structural–functional molecular mechanisms that maintain eumetazoan mature sequences within a few nucleotides of the loop may be present in calcisponges and potentially homoscleromorphs, but not silicisponges (nor plants, *Chlamydomonas* or *Ectocarpus*).

It could be argued that the absence of sequence overlap amongst the calcisponge and homoscleromorph miRNAs repertoires either with one another or with eumetazoan miRNAs is *because* sponge miRNAs have a high turn-over, if their LCA did in fact possess miRNAs. We observe however, that the miRNA repertoire of demosponges and calcisponges are conserved for hundreds of millions of years, which casts a strict interpretation of this into some doubt (Fig. 4).

Therefore it remains unclear whether the LCA of all animals—irrespective of sponge mono- versus paraphyly—possessed miRNAs. Because of the presence of Drosha and Pasha orthologs in demosponges and calcisponges, resolving this issue will require elucidation of miRNA biogenesis mechanisms in sponges, as well as their biological function, and further characterization of potential miRNA repertoires in protist species near the fungal/metazoan divergence. We suggest based on the evidence presented herein that the multiple groups of miRNAs in the three or four metazoan groups characterized by the possession of miRNAs represent discrete examples of independent molecular inventions—molecular exaptations—of a more primitive RNAi machinery. And finally, we suggest that the role miRNAs play in establishing cell and tissue specificity in the Eumetazoa is not conserved in sponges.

ACKNOWLEDGEMENTS

We want to thank D. Lavrov for homoscleromorph material and for providing input into the *Sycon* genome sequencing, and B. King for bioinformatic support. K.J.P. is supported by the NASA Astrobiology Institute; M.A. is supported by the Sars Centre. E.A.S. is supported by an Agouron Geobiology Fellowship. The majority of S. *ciliatum* and *L. complicata* sequencing has been performed at The Norwegian High-Throughput Sequencing Centre funded by the Research Council of Norway.

LITERATURE CITED

Adamska M, Degnan BM, Green K, Zwafink C. 2011. What sponges can tell us about the evolution of developmental processes. Zoology 114:1–10.

Adamska M, Larroux C, Adamski M, et al. 2010. Structure and expression of conserved Wnt pathway components in the demosponge *Amphimedon queenslandica*. Evol Dev 12:494–518.

Ambros V, Bartel B, Bartel DP, et al. 2003. A uniform system for microRNA annotation. RNA 9:277–279.

Axtell MJ, Westholm JO, Lai EC. 2011. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. Genome Biol 12:221.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. Cell 136:215–233.

Berezikov E. 2011. Evolution of microRNA diversity and regulation in animals. Nat Rev Genet 12:846–860.

Berezikov E, Guryev V, van de Belt J, et al. 2005. Phylogenetic shadowing and computational identification of human microDNA genes. Cell 120:21–24.

Borchiellini C, Manuel M, Alivon E, et al. 2001. Sponge paraphyly and the origin of Metazoa. J Evol Biol 14:171–179.

Boury-Esnault N, Rutzler K, editors. 1997. Thesaurus of Sponge Morphology. Smithsonian Contributions to Zoology 596.

Brain CK, Prave AR, Hoffmann KH, Fallick AE, Botha A, Herd DA, Sturrock C, Young I, Condon DJ, Allison SG. 2012. The first animals: ca. 760-million-year-old sponge-like fossils from namibia. S Afr J Sci 108(1-2):658.

Carthew RW, Sontheimer EJ. 2009. Origins and mechanisms of miRNAs and siRNAs. Cell 136:642–655.

Cavalier-Smith T, Allsopp MTEP, Chao EE, Boury-Esnault N, Vacelet J. 1996. Sponge phylogeny, animal monophyly, and the origin of the nervous system: 18S rRNA evidence. Can J Zool 74:2031–2045.

Christodoulou F, Raible F, Tomer R, et al. 2010. Ancient animal microRNAs and the evolution of tissue identify. Nature 463:1084–1088.

Collins AG. 1998. Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. Proceedings of the National Academy of Sciences, USA 95, 15458–15463.

Degnan BM, Vervoort M, Larroux C, Richards GS. 2009. Early evolution of metazoan transcription factors. Curr Opin Genet Dev 19:591–599.

Ebert MS, Sharp PA. 2012. Roles for microRNAs in conferring robustness to biological processes. Cell 149:515–524.

Erwin DH, LaFlamme M, Tweedt SM, et al. 2011. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science 334:1091–1097.

Funayama N. 2010. The stem cell system in demosponges: insights into the origin of somatic stem cells. Dev Growth Differ 52:1–14.

Fortunato S, Adamski M, Bergum B, et al. 2012. Genome-wide analysis of the sox family in the calcareous sponge *Sycon ciliatum*: multiple genes with unique expression patterns. EvoDevo 3:14.

Grimson A, Srivastava M, Fahey B, et al. 2008. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature 455:1193–1197.

Han J, Lee Y, Yeom K-H, et al. 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell 125: 887–901.

Harcet M, Roller M, Cetkovi H, et al. 2010. Demosponge EST sequencing reveals a complex genetic toolkit of the simplest metazoans. Mol Biol Evol 27:2747–2756.

Harrison FW, De Vos L. 1991. Porifera. In: Harrison FW, Westfall JA, editors. Microscopic anatomy of the invertebrates. New York: Wily-Liss. p 29–89.

Heimberg AM, Cowper-Sal lari R, Sémon M, Donoghue PCJ, Peterson KJ. 2010. microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. Proceedings of the National Academy of Sciences, USA 107, 19379–19383.

Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ. 2008. MicroRNAs and the advent of vertebrate morphological complexity. Proceedings of the National Academy of Sciences, USA 105, 2946–2950.

Herranz H, Cohen SM. 2010. MicroRNAs and gene regulatory networks: managing the impact of noise in biological systems. Genes Dev 24:1339–1344.

Hertel J, Lendemeyer M, Missal K, et al. 2006. The expansion of the metazoan microRNA repertoire. BMC Genomics 7:25.

Hornstein E, Shomron N. 2006. Canalization of development by microRNAs. Nat Genet 38:S20–S24.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet 12:99–110.

Ivey KN, Srivastava D. 2010. MicroRNAs as regulators of differentiation and cell fate decisions. Cell Stem Cell 7:36–41.

Jones-Rhoades MW, Bartel DP, Bartel B. 2006. MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57:19–53.

Kim VN. 2005. MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 6:376–385.

Kosik KS. 2010. MicroRNAs and cellular phenotypy. Cell 143:21–26.

Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39:D152–D157.

Larroux C, Luke GN, Koopman P, et al. 2008. Genesis and expansion of metazoan transcription factor gene classes. Mol Biol Evol 25:980–996.

Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25:2286–2288.

Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. Mol Biol Evol 25:1307–1320.

Li MA, He L. 2012. microRNAs as novel regulators of stem cell pluripotency and somatic cell reprogramming. BioEssays 34:670–680.

Love GD, Grosjean E, Stalvies C, et al. 2009. Fossil steroids record the appearance of Demospongiae during the Cryogenian period. Nature 457:718–721.

Lu J, Getz G, Miska EA, et al. 2005. MicroRNA expression profiles classify human cancers. Nature 435:834–838.

Manuel M, Borchiellini C, Alivon E, et al. 2003. Phylogeny and evolution of calcareous sponges: monophyly of Calcinea and Calcaronea, high level of morphological homoplasy, and the primitive nature of axial symmetry. Syst Biol 52:311–333.

Martinez NJ, Gregory RI. 2010. MicroRNA gene regulatory pathways in the establishment and maintenance of ESC identity. Cell Stem Cell 7:31–35.

Mendell JT, Olson EN. 2012. MicroRNAs in stress signaling and human disease. Cell 148:1172–1187.

Nichols SA, Dayel MJ, King N. 2009. Genomic, phylogenetic, and cell biological insights into metazoan origins. In: Telford MJ, Littlewood DTJ, editors. Animal evolution: genes, fossils, and trees. Oxford: Oxford University Press. p 24–32.

Nichols SA, Dirks W, Pearse JS, King N. 2006. Early evolution of animal cell signaling and adhesion genes. Proc Natl Acad Sci USA 103:12451–12456.

Nichols SA, Roberts BW, Richter DJ, Fairclough SR, King N. 2012. Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/β-catenin complex. Proc Natl Acad Sci USA 109(32): 13046–13051.

Peter ME. 2009. Let-7 and miR-200 microRNAs. Cell Cycle 8:843–852.

Peterson KJ, Dietrich MR, McPeek MA. 2009. MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. BioEssays 31:736–747.

Peterson KJ, Eernisse DJ. 2001. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. Evol Dev 3:170–205.

Philippe H, Derelle R, Lopez P, et al. 2009. Phylogenomics revives traditional views on deep animal relationships. Curr Biol 19:706–712.

Pick KS, Philippe H, Schreiber F, et al. 2010. Improved phylogenomic taxon sampling noticably affects nonbilaterian relationships. Mol Biol Evol 27:1983–1987.

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. 2002. MicroRNAs in plants. Genes Dev 16:1616–1626.

Richards GS, Degnan BM. 2009. The dawn of developmental signaling in the Metazoa. Cold Spring Harb Symp Quant Biol 74:1–10.

Richards GS, Degnan BM. 2012. The expression of Delta ligands in the sponge *Amphimedon queenslandica* suggest an ancient role for Notch signaling in metazoan development. EvoDevo 3:15.

Sempere LF, Cole CN, McPeek MA, Peterson KJ. 2006. The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. J Exp Zool B Mol Dev Evol 306B: 575–588.

Simpson T. 1984. The Cell Biology of Sponges. New York/Berlin: Springer-Verlang.

Sperling EA, Peterson KJ, Pisani D. 2009. Phylogenetic-signal dissection of nuclear housekeeping genes supports the paraphyly of sponges and the monophyly of Eumetazoa. Mol Biol Evol 26:2261–2274.

Sperling EA, Pisani D, Peterson KJ. 2007. Poriferan paraphyly and its implications for Precambrian paleobiology. In: Vickers-Rich P, Komarower P, editors. The rise and fall of the ediacaran biota. London: The Geological Society. p 355–368. Special Publications.

Sperling EA, Robinson JM, Pisani D, Peterson KJ. 2010. Where's the glass? Biomarkers, molecular clocks, and microRNAs suggest a 200-Mry missing Precambrian fossil record of siliceous-sponge spicules. Geobiology 8:24–36.

Srivastava M, Simakov O, Chapman J, et al. 2010. The *Amphimedon queenslandica* genome and the evolution of animal complexity. Nature 466:720–726.

Starega-Roslan J, Koscianska E, Kozlowski P, Kryzyzosiak WJ. 2011. The role of the precursor structure in the biogenesis of microRNA. Cell Mol Life Sci 68:2859–2871.

Steinmetz PRH, Kraus JEM, Larroux C, et al. 2012. Independent evolution of striated muscles in cnidarians and bilaterians. Nature 487:231–234.

Tarver JE, Donoghue PCJ, Peterson KJ. 2012. Do miRNAs have a deep evolutionary history? BioEssays 34(10):857–866.

Wheeler BM, Heimberg AM, Moy VN, et al. 2009. The deep evolution of metazoan microRNAs. Evol Dev 11:50–68.

Zrzavy J, Mihulka S, Kepka P, Bezdek A, Tietz D. 1998. Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. Cladistics 14:249–285.