INVITED REVIEW

A renaissance of microRNAs as taxonomic and phylogenetic markers in animals

Bastian Fromm 🆔

The Arctic University Museum of Norway, UiT – The Arctic University of Norway, Tromsø, Norway

Correspondence
Bastian Fromm, The Arctic University Museum of Norway, UiT – The Arctic University of Norway, Tromsø, Norway.
Email: bastian.fromm@uit.no

Abstract

Molecular markers for tracing animal sample origins and compositions are critical for applications such as parasite detection, contamination screening, and sample authentication. Among these, microRNAs have emerged as promising candidates due to their deep conservation, near-hierarchical evolution, and stability. I here review the suitability of microRNAs as taxonomic and also phylogenetic markers and show how careful annotation efforts and the establishment of the curated microRNA gene database MirGeneDB and tools like MirMachine have revitalized microRNA research. These advancements enable accurate phylogenetic and taxonomic studies, highlighting microRNAs' potential in resolving long-standing questions in animal relationships and extending to applications in ancient DNA and environmental RNA analysis. Future research must focus on expanding microRNA complements across all Metazoa and further improving annotation methodologies.

KEYWORDS

Metazoa, microRNAs, phylogenetic markers, taxonomic markers

1 | INTRODUCTION

Molecular markers to easily trace the origin or the composition of a given animal sample are highly sought after: may it be, for example, to test for parasites in the blood of livestock, to screen for biological contaminants in experimental data, or to detect animal traces from environmental samples. Moreso, it is often desired to confirm the authenticity of samples, their exact taxonomic origin down to the species level. This may, for instance, be the case in samples that are very old and likely contaminated such as those from permafrost or from scientific museum collections. Suitable molecular markers are also interesting to understand whether groups of organisms belong to the same taxonomic rank, i.e. if groups of organisms belong to the same species or if they belong to the same management unit or not. Such questions of relatedness are of course interesting at all taxonomic and, ultimately, phylogenetic levels, where, after centuries of research, many open questions of the relationships of animals remain (Bleidorn, 2019; Cannon et al., 2016; Dunn et al., 2008; Edgecombe et al., 2011; Laumer et al., 2018; Marlétaz et al., 2008; Najle et al., 2023; Philippe et al., 2011; Ryan et al., 2013; Schiffer et al., 2022; Simion et al., 2017).

Suitable markers should on the one hand be deeply conserved so they can be detected across geological time scales and on the other hand, in a clock-like manner, for instance by sequence-changes or the evolution of novel

genes, trace both evolutionary deep and more recent splits of e.g. phyla and species without evolving convergently. Importantly, such markers should be readily detectable and ideally be very stable from a range of relevant samples. Beginning with Carl Woese, the pioneer of metagenomics using ribosomal RNA to differentiate (and discover!) domains of life (Woese & Fox, 1977), to mitochondrial markers such as cytochrome oxidase (Brown et al., 1979) (or ITS in plants [Baldwin et al., 1995]) for taxonomic barcoding and ultimately phylogenomics using full or near complete complements of all protein coding gene sequences for a set of species, each could only fulfil one or the other criteria (Delsuc et al., 2005). Today, no marker is established that is used widely as a taxonomic and phylogenetic marker in animals, or any other organismic group.

With the discovery of microRNAs in 1993 (Lee et al., 1993) and the realization that they are highly conserved in all animals in 2000 (Pasquinelli et al., 2000) and 2001 (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee & Ambros, 2001), respectively, a new field of research was established (see Zamore, 2020). Shortly after, a wave of initially very promising studies appeared showing the deep conservation of microRNAs and their patterns of near-hierarchical evolution (Grimson et al., 2008; Hertel et al., 2006; Sempere et al., 2006). microRNAs are short non-coding gene-regulators that regulate the translation of messenger RNAs (mRNAs) and can act both as either switches or with more subtle effects (Bartel, 2018). Today, microRNAs are widely studied because of their roles in development and disease with potential as clinical biomarkers. Within the exponentially expanding non-coding RNA field, microRNAs are most studied, and more than 100,000 papers have been published by now on microRNAs alone. However, although there was an initial euphoria for microRNA-based phylogenetics (Campbell et al., 2011; Devor & Peek, 2008; Erwin et al., 2011; Fromm et al., 2013; Sperling & Peterson, 2009; Tarver et al., 2013; Wiegmann et al., 2011) and taxonomic studies (Fromm et al., 2014; Helm et al., 2012), the suitability of microRNAs for addressing such questions came into question due to supposed evidence for homoplastic (convergent) evolution of microRNAs in some groups and a proclaimed relatively high rate of independent losses in some animal lineages (Dunn, 2014; Thomson et al., 2014). Furthermore, there was only a relatively small community that advocated for the usefulness of microRNAs as phylogenetic and taxonomic markers, and this marked a substantial decline in the popularity of microRNAs for phylogenetics and taxonomy. However, it also turned out that the correct identification and annotation of microRNAs was an issue of concern and a set of carefully conducted studies was initiated that identified the source of the contradicting reports: while previous pioneering studies had taken a labour

intensive manual annotation approach to microRNAs, the new study used a public repository of published microRNA complements (miRBase [Kozomara et al., 2019]) at face value without curation (Tarver et al., 2018). Indeed, when carefully reanalyzing the same datasets after curating the data, the raised concerns were not supported by the data, and it became apparent that an effort for addressing the underlying issues in the microRNA annotation field was needed. For years, a major concern in microRNA research had already been the quality of miRBase with estimates of 2/3 false-positive entries (Axtell & Meyers, 2018; Castellano & Stebbing, 2013; Chiang et al., 2010; Fromm et al., 2015; Fromm, Domanska, et al., 2020; Fromm, Høye, et al., 2022; Fromm, Keller, et al., 2020; Guo et al., 2020; Jones-Rhoades, 2012; Langenberger et al., 2011; Ludwig et al., 2017; Meng et al., 2012; Tarver et al., 2012; Taylor et al., 2014; Wang & Liu, 2011). It is important to note that miRBase hosts published microRNA sequences, and is, thus, highly dependent on the quality of published studies. Furthermore, the database is heavily biased toward organisms that have been studied most. Therefore, miRBase annotations are often incomplete for non-model species, which, in turn, may have been interpreted as absence or loss of specific microRNAs and lead to the '*critical appraisal*' (Thomson et al., 2014). More so, miRBase uses an outdated and inconsistent nomenclature which leads to problems in identifying homologous genes in different organisms. Through the years this led to numerous scattered efforts to develop new naming systems (Budak et al., 2016) and study-specific reannotations of microRNA complements (Chiang et al., 2010; Fromm et al., 2013; Grimson et al., 2008; Jan et al., 2011; Ruby et al., 2006, 2007). Unfortunately, such progress never led to an update of miRBase.

To tackle this problem, MirGeneDB, a database of manually curated microRNA genes, was launched in 2015, and has since become the new gold standard for microRNA identification, annotation and analysis (Fromm et al., 2015; Fromm, Domanska, et al., 2020; Fromm, Høye, et al., 2022). Building upon earlier studies in the field MirGeneDB has been instrumental to facilitate a new era of microRNA-based research, including, but not limited to phylogenetic and taxonomic studies.

2 | microRNAs AS PHYLOGENETIC MARKERS

microRNAs are well suited as phylogenetic markers in animals due to their (1) deep conservation, (2) their near-hierarchical evolution and (3) rare secondary loss events (Figure 1). An in depth review by Tarver et al. (2013) lists additional features, such as the rarity
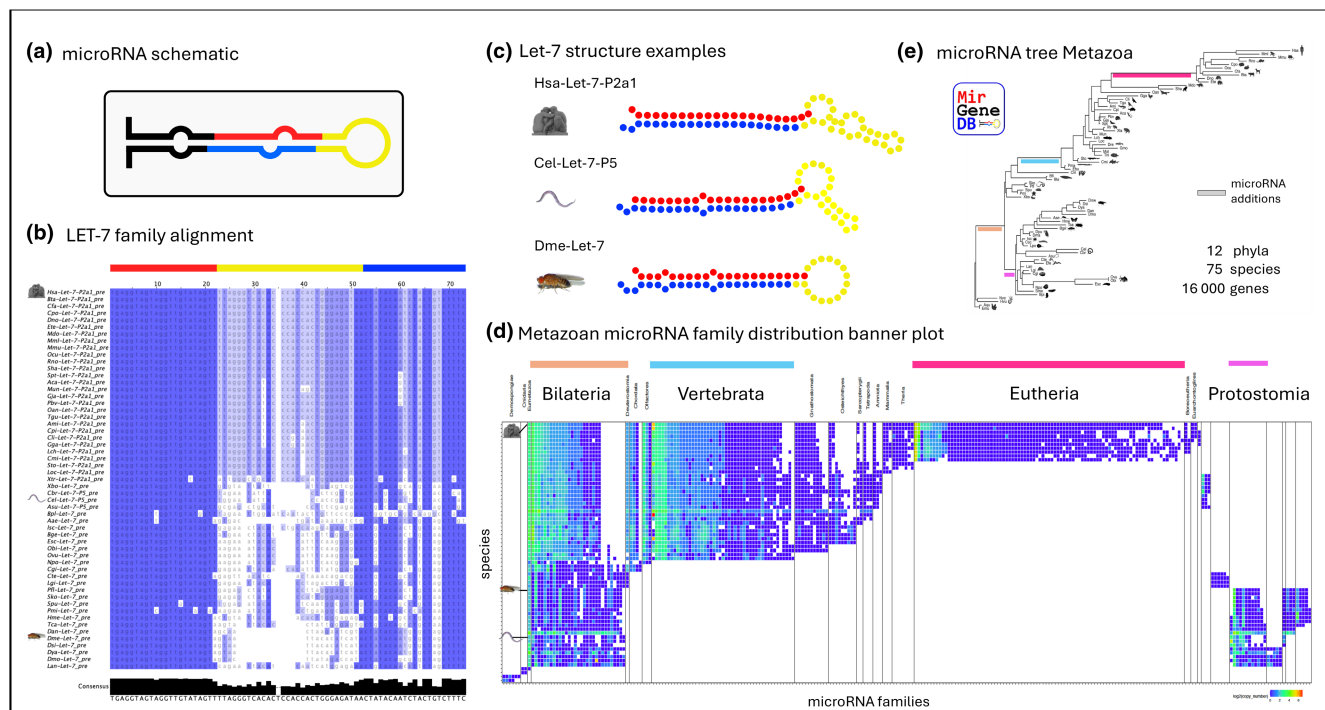
FIGURE 1 Deep conservation and hierarchical evolution of microRNAs in Metazoa. (a) microRNA schematic highlighting the primary microRNA hairpin like structure (2D view) composed of precursor microRNA (pre-microRNA) which gets processed into mature (red), star (blue) and loop part (yellow) (Note that there also exist 3′ mature and co-mature microRNAs but are not shown here). (b) Multisequence alignment of LET-7 family orthologues across Metazoa. Blue coloration of sequences indicates sequence similarity. Note that LET-7s are 5′ matures and how different mature, loop and star sequences are conserved. Icons highlight human, *C. elegans* and *D. melanogaster* sequences that are fully conserved for their mature microRNA sequences. (c) Secondary structures of LET-7 examples of the same species. (d) Banner plot of 75 Metazoan microRNA complements from MirGeneDB 2.1 highlights the strong conservation even at paralogue level (heat-function) and the near-hierarchical evolution. See Figure S1 for a fully detailed version of the banner plot). Coloured bars on top depict selected phylogenetic nodes of origin that can also be found in (e) Reconstructed phylogenetic tree of Metazoan species based on the presence and absence of microRNA families. Branch lengths indicate changes in microRNA complements highlighting specific bursts of microRNA evolution (colours like in d).

of nucleotide substitution in mature microRNAs and the seed in particular, which readily allows to identify orthologues between species. Combined with only very rarely observed convergent evolution of microRNAs (see Wang et al., 2024, for the very few examples), this allows for the computational reconstruction of phylogenetic trees both based on the presence and absence of microRNAs, i.e. the microRNA complements, as well as their concatenated nucleotide sequences, including flanking regions (Kenny et al., 2015). The phylogenetic analysis of microRNA complements is particularly useful for addressing higher level systematic issues, whereas phylogenetic analysis at the nucleotide sequence level may also be useful at lower taxonomic ranks such as the species, genus and family or order levels.

As indicated before, the establishment of MirGeneDB marked a huge progress for microRNA research. Very recently, the development of the microRNA prediction algorithm MirMachine (Umu et al., 2023) has further substantially alleviated the methodological portfolio. Briefly,

MirMachine is based on the application of covariance models of microRNA families trained on the manually annotated microRNA complements in MirGeneDB, and allows for the prediction of conserved microRNA complements from reference genomes without a need for specific smallRNA sequencing data (see also the similar tool ncOrtho, which uses microRNA covariance models of human microRNAs from MirGeneDB [Langschied et al., 2023]). MirMachine enables the fast annotation of microRNAs, their family annotation and sequence identification for comparative, e.g. phylogenetic studies. Such studies are relevant as only a relatively small proportion of the relationship of major animal groups has been solved (Figure 2).

It is likely that MirMachine derived conserved microRNA sequences will be able to solve many of these outstanding questions, however, given that more than half of the animal phyla have not been yet sampled for their microRNA complements, there is a probability that MirMachine models might not be accurate or
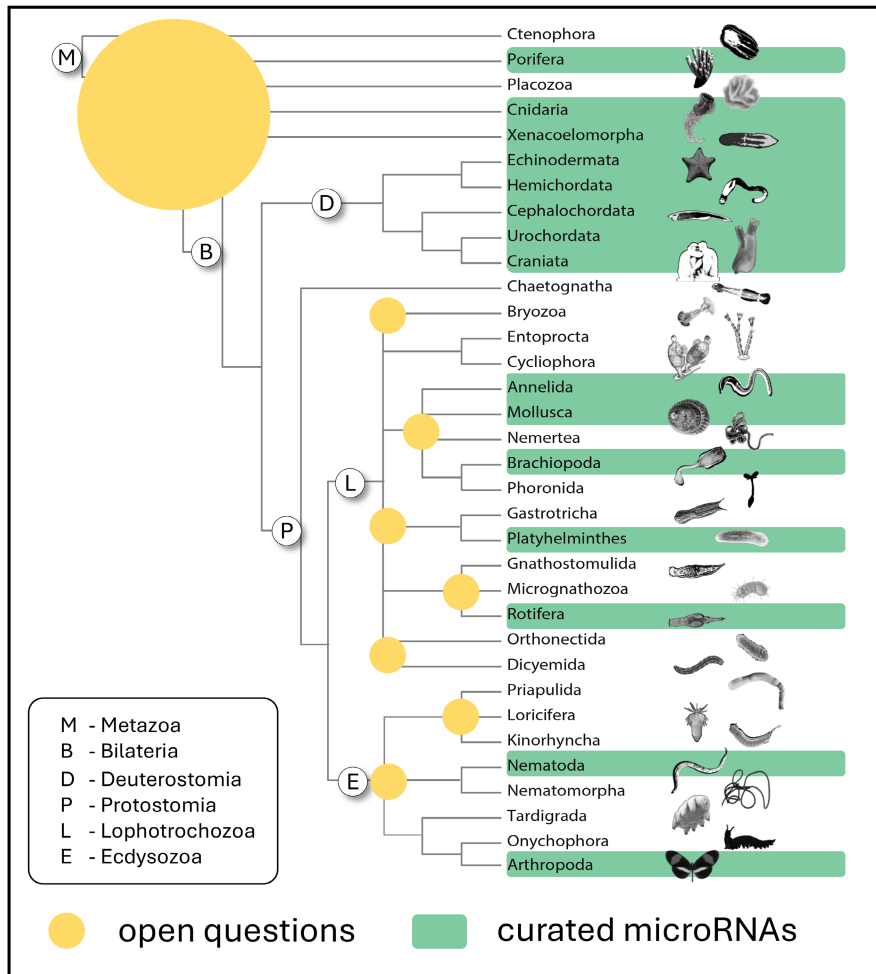
FIGURE 2 Simplified phylogenetic tree of Metazoan phyla highlighting open questions in our understanding of their relationship (yellow circles). The green highlighted phyla have microRNA complements annotated from at least one representative species (MirGeneDB).

sensitive enough to capture all conserved microRNAs in these genomes. Importantly, clade specific microRNAs might have to be discovered de novo and, hence, novel microRNA prediction using smallRNA sequencing data, might be required in addition to a near-mandatory MirMachine step. For the annotation of novel microRNAs, a range of tools has been published over the years that have been used widely, such as miRDeep (Friedländer et al., 2008, 2012), miRanalyzer (sRNAbench) (Aparicio-Puerta et al., 2019; Hackenberg et al., 2009), or MirMiner (Wheeler et al., 2009). Today, there is no, one-fits-all tool and each of the tools come with connected challenges: (1) they require command-line skills, (2) careful manual curation of outputs and (3) are not necessarily designed to process today's ultra-high throughput datasets (see Fromm, 2016; Fromm, Zhong, et al., 2022) (Figure 3a). For smallRNA sequencing-based analyses, the tool miRTrace was especially developed for microRNA quality control that should be included in a first step for the de novo prediction of microRNAs and has, in addition, a feature that makes it very relevant for taxonomic questions: taxonomic tracing with microRNAs (Kang et al., 2018).

3 | microRNAs AS TAXONOMIC MARKERS

Because microRNAs evolve near-hierarchical, they cannot only help to solve relationships between organism's groups based on shared microRNAs, but can be used, in a barcode kind of way, to classify species of unknown relatedness to known taxonomic groups, such as phyla or, ideally, down to species level (see Table S1 for exhaustive list of taxonomic node specific microRNA family gains and losses). For instance, in a study on *Octopus* (Zolotarov et al., 2022) the microRNA complements of two *Octopus* species and two other Cephalopods were annotated and compared to known microRNA complements in MirGeneDB (Figure 3b,c). The cephalopods all shared eumetazoan, bilaterian and protostome microRNA families in addition to Lophotrochozoan and Platyrochozoan, but most importantly to molluscan microRNA families, clearly identifying them as Molluscs. Using MirMiner (Wheeler et al., 2009), a large number of novel microRNA families was identified and shared among all cephalopods, a large number in coleoid cephalopods and a similar number in both *Octopus* species. Based on the
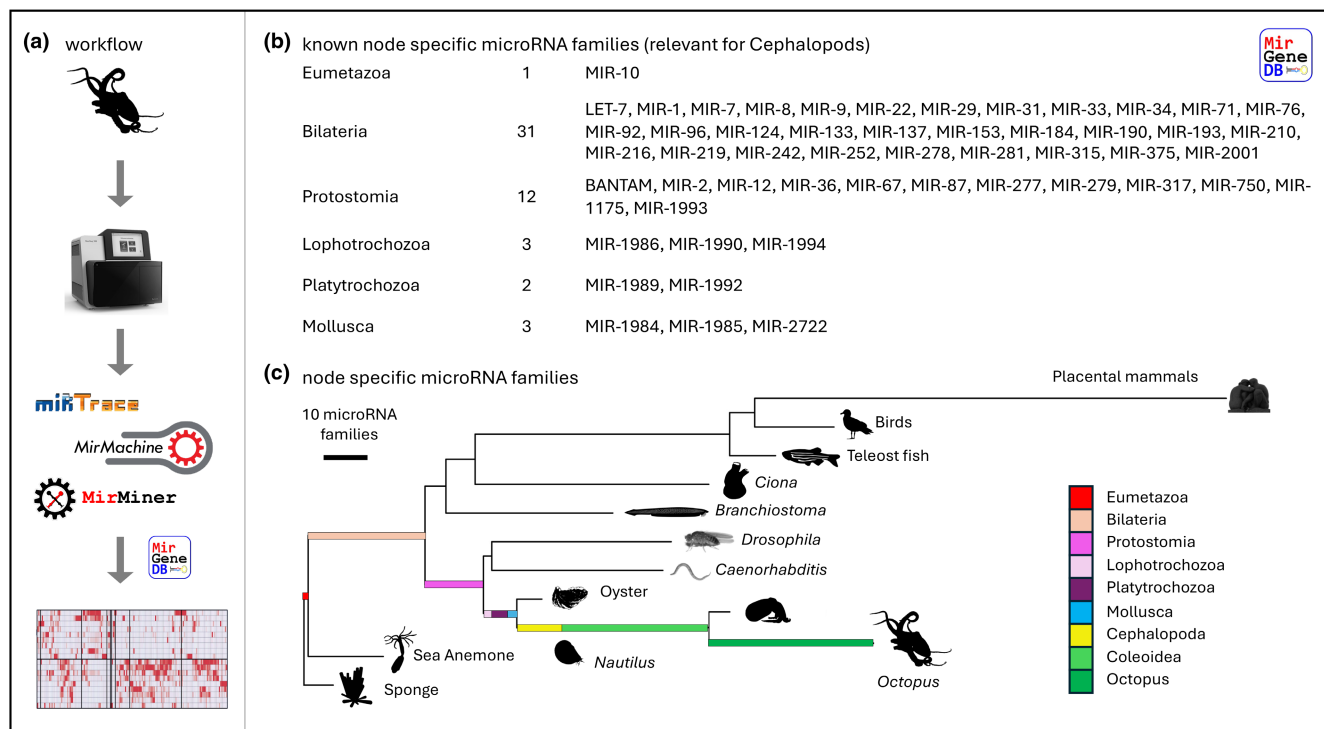
FIGURE 3 Pipeline for microRNA study and example for microRNAs as taxonomic markers. (a) RNA samples ideally from several tissues should be extracted and smallRNA libraries sequenced in single end mode for approximately 20 million reads per sample. For quality control of smallRNA sequencing data, miRTrace (Kang et al., 2018) should be used as a first step. miRTrace can already give basic taxonomic information on likely origin and the presence of contaminants. The next step is a MirMachine analysis (Umu et al., 2023) which requires the presence of a reference genome and will predict conserved microRNAs. Tools such as MirMiner (Fromm et al. in prep; Wheeler et al., 2009) that use smallRNA sequencing data, reference genomes can be used for the verification of MirMachine results and the prediction of novel microRNAs. (b) Relevant part of known microRNA family origins in the evolution of Metazoans. The detection of these microRNAs enables the association of samples to taxonomic groupings. In the example of the study of smallRNAs in Octopus vulgaris (Zolotarov et al., 2022), we detected Eumetazoan, Protostome, Lophotrochozoan, Platytrochozoan and molluscan microRNA families that clearly authenticated the sequencing experiment. (c) Expansion of Octopus microRNA complements relative to selected other metazoans. Note the coloration of annotated nodes of origin and how the detection of them from other samples, e.g. environmental or ancient, would facilitate the identification of corresponding node association for any sample.

presence of any of these families in a dataset, we can today confidently trace the origin of any sample.

In the mentioned miRTace tool (Kang et al., 2018) part of this information is implemented at a relatively coarse level and is used to e.g. test for contamination in small-RNA sequencing data in a quality control step and to get a first idea of data quality and composition after sequencing. Efforts to incorporate annotation information and phylogenetic information on microRNAs from MirGeneDB are under way and will substantially improve the resolution of miRTrace and massively expand potential applications.

Another example of using microRNAs as taxonomic markers is the study of sequence variation in the usually very strongly conserved microRNAs to test for differences between populations of species or different species. Examples from the flatworm field are the use of microRNAs in the salmonid ectoparasites *Gyrodactylus salaris* and *Gyrodactylus thymalli* that showed that the supposed two species show stronger difference within the species

than between, which suggests a synonymization of the two, with wide reaching consequences for management (Fromm et al., 2014). In a study on the *Opisthorchis—Clonorchis* species complex, an important group of endoparasitic trematodes (liver flukes), microRNAs were shown to be useful to differentiate between the three species, and importantly between the *O. viverrini* species, which is classified as bio carcinogen and the others.

Such differences of microRNAs between pathogens species of different severity, and between pathogen and hosts, are being implemented in studies to detect pathogens from blood by using microRNAs as taxonomic blood markers (Cheng et al., 2013; Ghalehnoei et al., 2020; Quintana et al., 2015; Tritten et al., 2014), although more systematic approaches using ultra deep sequencing or targeted assays, are currently missing.

Recently the new field of paleotranscriptomics, the study of historic or ancient samples for their RNA expression profiles, has become a new area of interest for

accurate—and stable—markers to verify the taxonomic authenticity of the data (Smith et al., 2019; Smith & Gilbert, 2019). In a detailed analysis on the 14,300 year old Tumat puppy microRNA complement from small RNA sequencing, we have shown that despite the fact that the majority of microRNA reads are identical between human and dog, and hence not taxonomically informative, we can detect microRNAs that either show canid specific sequence differences, or are truly canid specific (Fromm, Tarbier, et al., 2020). We followed up on this study by sequencing historical museum samples of the extinct Tasmanian tiger (thylacine) and could, among many conserved families and microRNAs specific to marsupials, identify microRNAs specific to Thylacine and novel to science (Mármol-Sánchez et al., 2023). Those microRNAs are especially exciting as they actually represent extinct genes which could not have been found without RNA sequencing data.

4 | FUTURE PERSPECTIVES

microRNAs currently experience a renaissance as suitable markers for various fields of research, among them as taxonomic and phylogenetic studies. Today, the limiting factors for microRNA-based research is a lack of a complete picture of microRNA complements from all animal phyla and the lack of more user-friendly and scalable computational tools for the prediction and annotation of novel microRNAs sequences from smallRNA sequencing data. Given the extensive number of studies producing smallRNA sequencing data in the past, it is surprising and worth mentioning that, currently, a substantial number of datasets has neither been publicly released or, much more common, are very hard to find in public repositories. Adherence to FAIR principles (Wilkinson et al., 2016) and more accessible depositories are warranted.

Although the number of reference databases continues to increase in number (see Fromm, Keller, et al., 2020), the focus should be on the use of highly curated databases that contain bona fide microRNA annotations (see Fromm, Zhong, et al., 2022). For MirGeneDB, the most limiting factors are the availability of smallRNA sequencing data for available nuclear genomes of missing phyla, as well as the laborious time-consuming manual annotation of microRNAs, especially for the correct paralogue identification. Improvement of available genome annotation and automated microRNA annotation algorithms using genome references only, such as MirMachine, might pave the way to a de novo prediction of full complements with accurate naming and significantly reduced need for manual curational work.

Exemplified in the paleotranscriptomics field, microRNAs have proven to be the molecule of choice as taxonomic markers because they are apparently also very stable due to their relatively small size, the fact they are capped and their association to proteins (see Friedländer & Gilbert, 2024). The ability to accurately differentiate species of parasites with severely divergent pathogenicity, the detection of parasitic species in cattle or livestock species or the detection of novel taxonomically informative microRNAs in species that haven't been studied yet holds a huge potential for humanity and a unique chance to solve outstanding scientific questions in biosystematics. For the latter, microRNAs will continue to contribute to open questions of Metazoan Phylogeny in e.g. gastrotrich paraphyly (Fromm et al., 2019), internal molluscan (Rosani et al., 2021) and ecdysozoan relationships (Campbell et al., 2011) and the Xenacoelomorpha discussion (Philippe et al., 2011; Schiffer et al., 2022) and, depending on available data on hitherto uncharted animal phyla (Figure 2). While the focus of this review is animals, there is clearly a large potential for microRNAs as taxonomic and phylogenetic markers in plants (as shown before by e.g. Taylor et al. (2014)), but there is an even larger backlog of annotation work and a less clear phylogenetic resolution in them.

We are only beginning to understand the full potential of microRNAs as taxonomic and phylogenetic markers in other fields. For instance, it was recently suggested to use RNA as environmental markers (see Stevens & Parsley, 2023), i.e. to extract RNA from sea water or sediment samples to reconstruct information on taxonomic distribution. Due to RNA, especially microRNAs, being often not single stranded but either protected by associated proteins (Argonaute), or structured in hairpins of other tertiary structures, RNA might be more stable than DNA and very well suited for this task. This is further supported by the fact that there are many thousands of copies of a range of RNA molecules in each cell, enhancing the probability of RNA detection over often diploid DNA copies from the cells by several orders of magnitude.

Finally, it is exciting to think that microRNAs can not only be used to distinguish between species as phylogenetic or taxonomic markers but show tissue and cell-type specific expression in addition to clear expression differences between developmental stages. This could theoretically give insights on developmental stage, age or injuries from ancient samples. Today such approaches are being established already in forensics (Sauer et al., 2017).

ORCID
Bastian Fromm https://orcid.org/0000-0003-0352-3037

REFERENCES

Aparicio-Puerta, E., Lebrón, R., Rueda, A., Gómez-Martín, C., Giannoukakos, S., Jaspez, D., Medina, J. M., Zubkovic, A., Jurak, I., Fromm, B., Marchal, J. A., Oliver, J., & Hackenberg, M. (2019). sRNAbench and sRNAtoolbox 2019: Intuitive fast small RNA profiling and differential expression. *Nucleic Acids Research*, *47*, W530–W535.

Axtell, M. J., & Meyers, B. C. (2018). Revisiting criteria for plant microRNA annotation in the era of big data. *The Plant Cell*, *30*, 272–284.

Baldwin, B. G., Sanderson, M. J., & Porter, J. M. (1995). The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden*, *82*(2), 247–277.

Bartel, D. P. (2018). Metazoan microRNAs. *Cell*, *173*, 20–51.

Bleidorn, C. (2019). Recent progress in reconstructing lophotrochozoan (spiralian) phylogeny. *Organisms, Diversity and Evolution*, *19*, 557–566.

Brown, W. M., George, M., Jr., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *76*, 1967–1971.

Budak, H., Bulut, R., Kantar, M., & Alptekin, B. (2016). MicroRNA nomenclature and the need for a revised naming prescription. *Briefings in Functional Genomics*, *15*, 65–71.

Campbell, L. I., Rota-Stabelli, O., Edgecombe, G. D., Marchioro, T., Longhorn, S. J., Telford, M. J., Philippe, H., Rebecchi, L., Peterson, K. J., & Pisani, D. (2011). MicroRNAs and phylogenomics resolve the relationships of Tardigrada and suggest that velvet worms are the sister group of Arthropoda. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 15920–15924.

Cannon, J. T., Vellutini, B. C., Smith, J., 3rd, Ronquist, F., Jondelius, U., & Hejnol, A. (2016). Xenacoelomorpha is the sister group to Nephrozoa. *Nature*, *530*, 89–93.

Castellano, L., & Stebbing, J. (2013). Deep sequencing of small RNAs identifies canonical and non-canonical miRNA and endogenous siRNAs in mammalian somatic tissues. *Nucleic Acids Research*, *41*, 3339–3351.

Cheng, G., Luo, R., Hu, C., Cao, J., & Jin, Y. (2013). Deep sequencing-based identification of pathogen-specific microRNAs in the plasma of rabbits infected with *Schistosoma japonicum*. *Parasitology*, *140*, 1751–1761.

Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., Johnston, W. K., Russ, C., Luo, S., Babiarz, J. E., Blelloch, R., Schroth, G. P., Nusbaum, C., & Bartel, D. P. (2010). Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes & Development*, *24*, 992–1009.

Delsuc, F., Brinkmann, H., & Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nature Reviews. Genetics*, *6*, 361–375.

Devor, E. J., & Peek, A. S. (2008). miRNA profile of a Triassic common mammalian ancestor and PremiRNA evolution in the three mammalian reproductive lineages. *The Open Genomics Journal*, *1*, 22–32.

Dunn, C. W. (2014). Reconsidering the phylogenetic utility of miRNA in animals. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 12576–12577.

Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. A., Seaver, E., Rouse, G. W., Obst, M., Edgecombe, G. D., Sørensen, M. V., Haddock, S. H. D., Schmidt-Rhaesa, A., Okusu, A., Kristensen, R. M., Wheeler, W. C., Martindale, M. Q., & Giribet, G. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*, *452*, 745–749.

Edgecombe, G. D., Giribet, G., Dunn, C. W., Hejnol, A., Kristensen, R. M., Neves, R. C., Rouse, G. W., Worsaae, K., & Sørensen, M. V. (2011). Higher-level metazoan relationships: Recent progress and remaining questions. *Organisms, Diversity and Evolution*, *11*, 151–172.

Erwin, D. H., Laflamme, M., Tweedt, S. M., Sperling, E. A., Pisani, D., & Peterson, K. J. (2011). The Cambrian conundrum: Early divergence and later ecological success in the early history of animals. *Science*, *334*, 1091–1097.

Friedländer, M. R., Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., & Rajewsky, N. (2008). Discovering microRNAs from deep sequencing data using miRDeep. *Nature Biotechnology*, *26*, 407–415.

Friedländer, M. R., & Gilbert, M. T. P. (2024). How ancient RNA survives and what we can learn from it. *Nature Reviews. Molecular Cell Biology*, *25*, 417–418. https://doi.org/10.1038/s41580-024-00726-y

Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Research*, *40*, 37–52.

Fromm, B. (2016). microRNA discovery and expression analysis in animals. In A. M. Aransay & J. L. L. Trueba (Eds.), *Field guidelines for genetic experimental designs in high-throughput sequencing* (pp. 121–142). Springer International Publishing.

Fromm, B., Billipp, T., Peck, L. E., Johansen, M., Tarver, J. E., King, B. L., Newcomb, J. M., Sempere, L. F., Flatmark, K., Hovig, E., & Peterson, K. J. (2015). A uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. *Annual Review of Genetics*, *49*, 213–242.

Fromm, B., Burow, S., Hahn, C., & Bachmann, L. (2014). MicroRNA loci support conspecificity of *Gyrodactylus salaris* and *Gyrodactylus thymalli* (Platyhelminthes: Monogenea). *International Journal for Parasitology*, *44*, 787–793.

Fromm, B., Domanska, D., Høye, E., Ovchinnikov, V., Kang, W., Aparicio-Puerta, E., Johansen, M., Flatmark, K., Mathelier, A., Hovig, E., Hackenberg, M., Friedländer, M. R., & Peterson, K. J. (2020). MirGeneDB 2.0: The metazoan microRNA complement. *Nucleic Acids Research*, *48*, D132–D141.

Fromm, B., Høye, E., Domanska, D., Zhong, X., Aparicio-Puerta, E., Ovchinnikov, V., Umu, S. U., Chabot, P. J., Kang, W., Aslanzadeh, M., Tarbier, M., Mármol-Sánchez, E., Urgese, G., Johansen, M., Hovig, E., Hackenberg, M., Friedländer, M. R., & Peterson, K. J. (2022). MirGeneDB 2.1: Toward a complete sampling of all major animal phyla. *Nucleic Acids Research*, *50*, D204–D210.

Fromm, B., Keller, A., Yang, X., Friedlander, M. R., Peterson, K. J., & Griffiths-Jones, S. (2020). Quo vadis microRNAs? *Trends in Genetics*, *36*, 461–463.

Fromm, B., Tarbier, M., Smith, O., Marmol-Sanchez, E., Dalen, L., Gilbert, T. P., & Friedlander, M. R. (2020). Ancient microRNA profiles of a 14,300-year-old canid samples confirm taxonomic origin and give glimpses into tissue-specific gene regulation from the Pleistocene. *RNA*, *27*, 324–334.

Fromm, B., Tosar, J. P., Aguilera, F., Friedländer, M. R., Bachmann, L., & Hejnol, A. (2019). Evolutionary implications of the

microRNA- and piRNA complement of _Lepidodermella squamata_ (Gastrotricha). _Non-coding RNA_, _5_, 19. https://doi.org/10.3390/ncrna5010019

Fromm, B., Worren, M. M., Hahn, C., Hovig, E., & Bachmann, L. (2013). Substantial loss of conserved and gain of novel MicroRNA families in flatworms. _Molecular Biology and Evolution_, _30_, 2619–2628.

Fromm, B., Zhong, X., Tarbier, M., Friedländer, M. R., & Hackenberg, M. (2022). The limits of human microRNA annotation have been met. _RNA_, _28_, 781–785.

Ghalehnoei, H., Bagheri, A., Fakhar, M., & Mishan, M. A. (2020). Circulatory microRNAs: Promising non-invasive prognostic and diagnostic biomarkers for parasitic infections. _European Journal of Clinical Microbiology & Infectious Diseases_, _39_, 395–402.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degnan, B. M., Rokhsar, D. S., & Bartel, D. P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. _Nature_, _455_, 1193–1197.

Guo, Z., Kuang, Z., Wang, Y., Zhao, Y., Tao, Y., Cheng, C., Yang, J., Lu, X., Hao, C., Wang, T., Cao, X., Wei, J., Li, L., & Yang, X. (2020). PmiREN: A comprehensive encyclopedia of plant miRNAs. _Nucleic Acids Research_, _48_, D1114–D1121.

Hackenberg, M., Sturm, M., Langenberger, D., Falcón-Pérez, J. M., & Aransay, A. M. (2009). miRanalyzer: A microRNA detection and analysis tool for next-generation sequencing experiments. _Nucleic Acids Research_, _37_, W68–W76.

Helm, C., Bernhart, S. H., Siederdissen, C. H., Nickel, B., & Bleidorn, C. (2012). Deep sequencing of small RNAs confirms an annelid affinity of Myzostomida. _Molecular Phylogenetics and Evolution_, _64_, 198–203.

Hertel, J., Lindemeyer, M., Missal, K., Fried, C., Tanzer, A., Flamm, C., Hofacker, I. L., Stadler, P. F., & Students of Bioinformatics Computer Labs 2004 and 2005. (2006). The expansion of the metazoan microRNA repertoire. _BMC Genomics_, _7_, 25.

Jan, C. H., Friedman, R. C., Ruby, J. G., & Bartel, D. P. (2011). Formation, regulation and evolution of _Caenorhabditis elegans_ 3′UTRs. _Nature_, _469_, 97–101.

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. _Plant Molecular Biology_, _80_, 3–16.

Kang, W., Eldfjell, Y., Fromm, B., Estivill, X., Biryukova, I., & Friedländer, M. R. (2018). miRTrace reveals the organismal origins of microRNA sequencing data. _Genome Biology_, _19_, 213.

Kenny, N. J., Sin, Y. W., Hayward, A., Paps, J., Chu, K. H., & Hui, J. H. L. (2015). The phylogenetic utility and functional constraint of microRNA flanking sequences. _Proceedings. Biological Sciences/The Royal Society_, _282_, 20142983.

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: From microRNA sequences to function. _Nucleic Acids Research_, _47_, D155–D162.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., & Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. _Science_, _294_, 853–858.

Langenberger, D., Bartschat, S., Hertel, J., Hoffmann, S., Tafer, H., & Stadler, P. F. (2011). _MicroRNA or not microRNA?_ 1–9.

Langschied, F., Leisegang, M. S., Brandes, R. P., & Ebersberger, I. (2023). ncOrtho: Efficient and reliable identification of miRNA orthologs. _Nucleic Acids Research_, _51_, e71.

Lau, N. C., Lim, L. P., Weinstein, E. G., & Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in _Caenorhabditis elegans_. _Science_, _294_, 858–862.

Laumer, C. E., Gruber-Vodicka, H., Hadfield, M. G., Pearse, V. B., Riesgo, A., Marioni, J. C., & Giribet, G. (2018). Support for a clade of placozoa and Cnidaria in genes with minimal compositional bias. _eLife_, _7_, e36278. https://doi.org/10.7554/eLife.36278

Lee, R. C., & Ambros, V. (2001). An extensive class of small RNAs in _Caenorhabditis elegans_. _Science_, _294_, 862–864.

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. Elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. _Cell_, _75_, 843–854.

Ludwig, N., Becker, M., Schumann, T., Speer, T., Fehlmann, T., Keller, A., & Meese, E. (2017). Bias in recent miRBase annotations potentially associated with RNA quality issues. _Scientific Reports_, _7_, 5162.

Marlétaz, F., Gilles, A., Caubit, X., Perez, Y., Dossat, C., Samain, S., Gyapay, G., Wincker, P., & Le Parco, Y. (2008). Chaetognath transcriptome reveals ancestral and unique features among bilaterians. _Genome Biology_, _9_, R94.

Mármol-Sánchez, E., Fromm, B., Oskolkov, N., Pochon, Z., Kalogeropoulos, P., Eriksson, E., Biryukova, I., Sekar, V., Ersmark, E., Andersson, B., Dalén, L., & Friedländer, M. R. (2023). Historical RNA expression profiles from the extinct Tasmanian tiger. _Genome Research_, _33_, 1299–1316.

Meng, Y., Shao, C., Wang, H., & Chen, M. (2012). Are all the miRBase-registered microRNAs true? A structure- and expression-based re-examination in plants. _RNA Biology_, _9_, 249–253.

Najle, S. R., Grau-Bové, X., Elek, A., Navarrete, C., Cianferoni, D., Chiva, C., Cañas-Armenteros, D., Mallabiabarrena, A., Kamm, K., Sabidó, E., Gruber-Vodicka, H., Schierwater, B., Serrano, L., & Sebé-Pedrós, A. (2023). Stepwise emergence of the neuronal gene expression program in early animal evolution. _Cell_, _186_, 4676–4693.e29.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Müller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E., & Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. _Nature_, _408_, 86–89.

Philippe, H., Brinkmann, H., Copley, R. R., Moroz, L. L., Nakano, H., Poustka, A. J., Wallberg, A., Peterson, K. J., & Telford, M. J. (2011). Acoelomorph flatworms are deuterostomes related to Xenoturbella. _Nature_, _470_, 255–258.

Quintana, J. F., Makepeace, B. L., Babayan, S. A., Ivens, A., Pfarr, K. M., Blaxter, M., Debrah, A., Wanji, S., Ngangyung, H. F., Bah, G. S., Tanya, V. N., Taylor, D. W., Hoerauf, A., & Buck, A. H. (2015). Extracellular Onchocerca-derived small RNAs in host nodules and blood. _Parasites & Vectors_, _8_, 58.

Rosani, U., Bortoletto, E., Bai, C.-M., Novoa, B., Figueras, A., Venier, P., & Fromm, B. (2021). Digging into bivalve miRNAomes: Between conservation and innovation. _Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences_, _376_, 20200165.

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., Ge, H., & Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in _C. elegans_. _Cell_, _127_, 1193–1207.

Ruby, J. G., Stark, A., Johnston, W. K., Kellis, M., Bartel, D. P., & Lai, E. C. (2007). Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. _Genome Research_, _17_, 1850–1864.

Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A.-D., Moreland, R. T., Simmons, D. K., Koch, B. J., Francis, W. R., Havlak, P., Comparative Sequencing Program, N. I. S. C., Smith, S. A., Putnam, N. H., Haddock, S. H. D., Dunn, C. W., Wolfsberg, T. G., Mullikin, J. C., Martindale, M. Q., & Baxevanis, A. D. (2013). The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science*, *342*, 1242592.

Sauer, E., Extra, A., Cachée, P., & Courts, C. (2017). Identification of organ tissue types and skin from forensic samples by microRNA expression analysis. *Forensic Science International. Genetics*, *28*, 99–110.

Schiffer, P. H., Natsidis, P., Leite, D. J., Robertson, H., Lapraz, F., Marlétaz, F., Fromm, B., Baudry, L., Simpson, F., Høye, E., Zakrzewski, A.-C., Kapli, P., Hoff, K. J., Mueller, S., Marbouty, M., Marlow, H., Copley, R. R., Koszul, R., Sarkies, P., & Telford, M. J. (2022). The slow evolving genome of the xenacoelomorph worm *Xenoturbella bocki*. *bioRxiv*, 2022.06.24.497508 https://doi.org/10.1101/2022.06.24.497508

Sempere, L. F., Cole, C. N., McPeek, M. A., & Peterson, K. J. (2006). The phylogenetic distribution of metazoan microRNAs: Insights into evolutionary complexity and constraint. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, *306*, 575–588.

Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D. J., Di Franco, A., Roure, B., Satoh, N., Quéinnec, É., Ereskovsky, A., Lapébie, P., Corre, E., Delsuc, F., King, N., Wörheide, G., & Manuel, M. (2017). A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Current Biology: CB*, *27*, 958–967.

Smith, O., Dunshea, G., Sinding, M.-H. S., Fedorov, S., Germonpre, M., Bocherens, H., & Gilbert, M. T. P. (2019). Ancient RNA from late Pleistocene permafrost and historical canids shows tissue-specific transcriptome survival. *PLoS Biology*, *17*, e3000166.

Smith, O., & Gilbert, M. T. P. (2019). Ancient RNA. In C. Lindqvist & O. P. Rajora (Eds.), *Paleogenomics: Genome-scale analysis of ancient DNA* (pp. 53–74). Springer International Publishing.

Sperling, E. A., & Peterson, K. J. (2009). microRNAs and metazoan phylogeny: big trees from little genes. In M. J. Telford & D. T. J. Littlewood (Eds.), *Animal evolution—genomes, trees and fossils*. Oxford University Press.

Stevens, J. D., & Parsley, M. B. (2023). Environmental RNA applications and their associated gene targets for management and conservation. *Environmental DNA*, *5*, 227–239.

Tarver, J. E., Donoghue, P. C., & Peterson, K. J. (2012). Do miRNAs have a deep evolutionary history? *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *34*, 857–866.

Tarver, J. E., Sperling, E. A., Nailor, A., Heimberg, A. M., Robinson, J. M., King, B. L., Pisani, D., Donoghue, P. C. J., & Peterson, K. J. (2013). miRNAs: Small genes with big potential in metazoan phylogenetics. *Molecular Biology and Evolution*, *30*, 2369–2382.

Tarver, J. E., Taylor, R. S., Puttick, M. N., Lloyd, G. T., Pett, W., Fromm, B., Schirrmeister, B. E., Pisani, D., Peterson, K. J., & Donoghue, P. C. J. (2018). Well-annotated microRNAomes do not evidence pervasive miRNA loss. *Genome Biology and Evolution*, *10*, 1457–1470.

Taylor, R. S., Tarver, J. E., Hiscock, S. J., & Donoghue, P. C. (2014). Evolutionary history of plant microRNAs. *Trends in Plant Science*, *19*, 175–182. https://doi.org/10.1016/j.tplants.2013.11.008

Thomson, R. C., Plachetzki, D. C., Mahler, D. L., & Moore, B. R. (2014). A critical appraisal of the use of microRNA data in phylogenetics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, E3659–E3668.

Tritten, L., O'Neill, M., Nutting, C., Wanji, S., Njouendoui, A., Fombad, F., Kengne-Ouaffo, J., Mackenzie, C., & Geary, T. (2014). Loa loa and *Onchocerca ochengi* miRNAs detected in host circulation. *Molecular and Biochemical Parasitology*, *198*, 14–17.

Umu, S. U., Paynter, V. M., Trondsen, H., Buschmann, T., Rounge, T. B., Peterson, K. J., & Fromm, B. (2023). Accurate microRNA annotation of animal genomes using trained covariance models of curated microRNA complements in MirMachine. *Cell Genomics*, *3*, 100348.

Wang, X., & Liu, X. S. (2011). Systematic curation of miRBase annotation using integrated small RNA high-throughput sequencing data for *C. elegans* and Drosophila. *Frontiers in Genetics*, *2*, 25.

Wang, Y., Tang, X., & Lu, J. (2024). Convergent and divergent evolution of microRNA-mediated regulation in metazoans. *Biological Reviews of the Cambridge Philosophical Society*, *99*, 525–545.

Wheeler, B. M., Heimberg, A. M., Moy, V. N., Sperling, E. A., Holstein, T. W., Heber, S., & Peterson, K. J. (2009). The deep evolution of metazoan microRNAs. *Evolution & Development*, *11*, 50–68.

Wiegmann, B. M., Trautwein, M. D., Winkler, I. S., Barr, N. B., Kim, J. W., Lambkin, C., Bertone, M. A., Cassel, B. K., Bayless, K. M., Heimberg, A. M., Wheeler, B. M., Peterson, K. J., Pape, T., Sinclair, B. J., Skevington, J. H., Blagoderov, V., Caravas, J., Kutty, S. N., Schmidt-Ott, U., ... Yeates, D. K. (2011). Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 5690–5695.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018.

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, *74*, 5088–5090.

Zamore, P. D. (2020). One small step for worms, one giant leap for small RNAs. *Nature Reviews. Molecular Cell Biology*, *21*, 565.

Zolotarov, G., Fromm, B., Legnini, I., Ayoub, S., Polese, G., Maselli, V., Chabot, P. J., Vinther, J., Styfhals, R., Seuntjens, E., Di Cosmo, A., Peterson, K. J., & Rajewsky, N. (2022). MicroRNAs are deeply linked to the emergence of the complex octopus brain. *Science Advances*, *8*, eadd9938.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.