Identification and expression profiling of microRNAs in *Hymenolepis*

Natalia Macchiaroli[1]*, Marcela Cucher[1], Laura Kamenetzky[1], Cristian Yones[2], Leandro Bugnon[2], Matt Berriman[3], Peter Olson[4], Mara Rosenzvit[1]*

[1]Instituto de Investigaciones en Microbiología y Parasitología Médicas (IMPaM), Facultad de Medicina, Universidad de Buenos Aires (UBA)-Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Buenos Aires, Argentina.

[2]Research Institute for Signals, Systems and Computational Intelligence, (sinc(i)), FICH-UNL)-Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Santa Fe, Argentina.

[3]Parasite Genomics Group, Wellcome Trust Sanger Institute, Hinxton, United Kingdom.

[4]Department of Life Sciences, The Natural History Museum, London, UK.

*Corresponding authors:

Mara Cecilia Rosenzvit and Natalia Macchiaroli

mrosenzvit@fmed.uba.ar

Note: Supplementary data associated with this article

ABSTRACT

Tapeworms (cestodes) of the genus *Hymenolepis* are the causative agents of hymenolepiasis, a zoonotic neglected disease. *Hymenolepis nana* is the most prevalent human tapeworm, especially affecting children. The genomes of *Hymenolepis microstoma* and *H. nana* have been recently sequenced and assembled. MicroRNAs (miRNAs), a class of small non-coding RNAs, are principle regulators of gene expression at the post-transcriptional level and are involved in many different biological processes. In previous work, we experimentally identified miRNA genes in the cestodes *Echinococcus, Taenia* and *Mesocestoides.* However, current knowledge about miRNAs in *Hymenolepis* is limited. In this work we described for the first time the expression profile of the miRNA complement in *H. microstoma,* and discovered miRNAs in *H. nana.* We found a reduced complement of 37 evolutionarily conserved miRNAs, putatively reflecting their low morphological complexity and parasitic lifestyle. We found a high expression of a few miRNAs in the larval stage of *H. microstoma* that is conserved in other cestodes suggesting that these miRNAs may have important roles in development, survival and for host-parasite interplay. We performed a comparative analysis of the identified miRNAs across the Cestoda and showed that most of the miRNAs in *Hymenolepis* are located in intergenic regions implying that they are independently transcribed. We found a *Hymenolepis*-specific cluster composed of three members of the mir-36 family. Also, we found that one of the neighboring genes of mir-10 was a Hox gene as in most bilaterial species. This study provides a valuable resource for further experimental research in cestode biology that might lead to improved detection and control of these neglected parasites. The comprehensive identification and expression analysis of *Hymenolepis* miRNAs can help to identify novel biomarkers for diagnosis and/or novel therapeutic targets for the control of hymenolepiasis.

Keywords: *Hymenolepis microstoma*, *Hymenolepis nana*, Cestoda, tapeworm, microRNA, hymenolepiasis; miRNA discovery

1. Introduction

Tapeworms (cestodes) of the genus *Hymenolepis* are the causative agents of hymenolepiasis, a zoonotic neglected disease transmitted by rodents. Two cosmopolitan species of *Hymenolepis* infect humans, the rat tapeworm *H. diminuta* and, particularly, the dwarf tapeworm *H. nana* that is the most prevalent human tapeworm worldwide, especially affecting children in temperate areas (Soares Magalhaes et al., 2013). Whereas most *H. nana* infections are asymptomatic, heavy infections contribute to increased morbidity in children and symptoms including severe diarrhea, abdominal pain, decreased appetite and reduced growth (Soares Magalhaes et al., 2013). Although drugs such as praziquantel are available and effective against adult tapeworm infections, they do not prevent re-infection in endemic areas with poor hygiene and sanitation where the frequency of transmission is high (Thompson, 2015). Infections in such areas frequently co-occur with other intestinal helminths (Soares Magalhaes et al., 2013) and diseases such as HIV-AIDS. The latter poses a particular threat as *H. nana* infections in immunocompromised individuals have been shown to give rise to invasive, tapeworm-derived tumors (Olson et al., 2003; Muehlenbachs et al., 2015). Finally, praziquantel resistance could become a problem in large scale deworming campaigns (Olson et al., 2012) and thus novel strategies for the control of hymenolepiasis are needed. The mouse bile-duct tapeworm *H. microstoma* that is prevalent in rodents worldwide, is a laboratory model for the human parasite *H. nana* and for other tapeworms causing neglected tropical diseases such as *Echinococcus* and *Taenia* for which complete life cycles cannot be maintained in the lab. The genome of *H.*

microstoma and other cyclophyllidean cestodes have been recently sequenced and assembled (Olson et al., 2012; Tsai et al., 2013). Importantly, the laboratory strain of *H. microstoma* used for genome sequencing has been previously characterized (Cunningham and Olson, 2010). In addition, the draft genome of *H. nana* is available as part of the Helminth Genomes Initiative, a collaborative project that aims to survey the genomes of parasitic helminths that are either of medical or veterinary importance, or are used as models for those (International Helminth Genomes Consortium, 2017). These unique genomic resources will enable the discovery of novel biomarkers for diagnosis and/or therapeutic targets for the control of the infections they cause.

MicroRNAs (miRNAs), a class of small non-coding RNAs, are principle regulators of gene expression at the post-transcriptional level and are involved in many different biological processes (Bartel, 2004; He and Hannon, 2004; Filipowicz et al., 2008). MiRNAs are transcribed by RNA polymerase II as long, primary miRNAs (pri-miRNAs) from miRNA genes or from introns of protein coding genes (Filipowicz et al., 2008). The primary miRNA is cleaved by Drosha to produce a ~70-nt long stem-loop precursor miRNA (pre-miRNA) that is further cleaved by Dicer to generate both the mature miRNA and antisense miRNA products. The mature miRNA is incorporated into the RNA-induced silencing complex (RISC) and bind to complementary sequences of target genes. In most bilaterian animals, target recognition is primarily through Watson-Crick pairing between miRNA nucleotides 2-7 (miRNA seed) and sites located in the 3' untranslated regions (3'UTRs) of target mRNAs (Bartel, 2018). This promotes the repression of protein translation and/or degradation of the target mRNA (Hutvagner and Zamore, 2002; Filipowicz et al., 2008; Bartel, 2009; Ghildiyal and Zamore, 2009). MiRNAs have been identified in a range of organisms such as viruses, plants and metazoans including free-living and parasitic helminths, with an increase in the number of

miRNA families correlated with an increase in morphological complexity (Niwa and Slack, 2007; Berezikov, 2011). Recently, it has been suggested that miRNAs might be potential therapeutic targets for the control of parasitic helminths (Britton et al., 2014). Furthermore, worm-derived miRNAs have shown promise as markers for the early detection of helminth infections (Cai et al., 2016).

The recent availability of the genomes of parasitic helminths of medical and veterinary importance (Howe et al., 2017), including cestodes, has provided a platform for the identification of miRNAs using both computational and experimental approaches. In previous work, we identified miRNA genes through deep sequencing in the cestodes *Echinococcus canadensis* (Cucher et al., 2015; Macchiaroli et al., 2015), *E. granulosus* s. s. (Macchiaroli et al., 2015), *E. multilocularis* (Cucher et al., 2015), *Mesocestoides corti* (Basika et al., 2016), *Taenia solium* and *T. crassiceps* (Perez et al., 2017). In addition, miRNA genes have been computationally identified in *H. microstoma* by Jin et al. (2013). However, a comprehensive identification of the miRNA repertoire and their expression profile are still lacking in *H. microstoma*, and there is no previous report of miRNAs in *H. nana*. The aims of this study were to analyze the miRNA expression profile in *H. microstoma* larvae and to discover miRNAs in *H. nana* from recently available genome data. The comprehensive identification and expression analysis of *Hymenolepis* miRNAs can help to identify biomarkers for diagnosis and/or novel therapeutic targets for the control of hymenolepiasis.

2. Material and methods

2.1. Parasite material

Five-day old (ie. mid-metamorphosis) larvae were collected from the haemocoel of beetles. The Nottingham strain (Cunningham and Olson, 2010) of *H. microstoma* was maintained *in vivo* using flour beetles (*Tribolium confusum*) and outbred conventional BALB/c mice (*Mus musculus*). To produce mid-metamorphosis larval samples, beetles were starved for five days and then exposed to freshly macerated, gravid proglottides of *H. microstoma* for ~ six hours. Gravid tissues were removed and the beetles were allowed to feed on flour ad libitum. Beetles were dissected five days post-exposure and the resulting larvae collected from the haemocoel into conditioned water. Morphologically the larva is elongated and well differentiated at both poles (ie. stage 3 according to Voge's system). All larvae were approximately half way through metamorphosis from the oncosphere to cysticercoid stages, albeit some variation in maturity was seen among individuals and thus the samples included 'stages' 2-4 as defined by Voge (Voge, 1964) with the majority representing stage 3. Approximately 550 individuals were combined in each of the three biological replicate larval samples. The samples were then transferred live to RNAlater (Ambion) and stored at -80 C until RNA extraction.

2.2. Small RNA isolation, library construction and sequencing

Larval samples were mechanically homogenized in Trizol (Invitrogen) for 10 s. Then, 200 µl of chloroform:isoamyl alcohol (24:1) was added and mixed thoroughly. Phase separation was carried out by centrifugation at maximum speed at 4 °C. Then, 0.5× isopropanol and 4 µl of glycogen (5 mg/ml) were added to the aqueous phase and the RNA was pelleted by centrifugation at maximum speed at 4 °C for 30 min. The resulting pellet was washed with 70%

ethanol, air dried, and re-suspended in nuclease-free water. The amount and integrity of total RNA was determined using a 2100 BioAnalyzer (Agilent, USA). RNA was concentrated by ethanol precipitation at -20°C overnight after elimination of polyadenylated mRNA using oligo-dT dynabead. The resulting pellet was re-suspended in 6µl nuclease free water and used as the input material.

Small RNA libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina. The 3' and 5' adapters were sequentially annealed with the annealing of the RT primer prior to 5' ligation to reduce the frequency of adapter dimer formation. First strand synthesis was performed followed by PCR enrichment of the libraries during which the index sequences were introduced post PCR, the small RNA libraries were quantified using the Agilent Bioanalyzer and the libraries pooled based on the concentration of the 147 bp peak (small RNA and adapters). The resulting pool was cleaned up through columns and sized selected using the Pippin Prep and the settings detailed in the NEB manual. Libraries were paired-end sequenced using an Illumina sequencing platform (HiSeq 2500) for 100 cycles. Small RNA libraries were constructed from three independent samples in order to count with biological replicates. For each sample, three technical replicates were sequenced. A total of nine libraries were sequenced. The small RNAseq data are available in ArrayExpress under accession code E-ERAD-236 (samples ERS353237, ERS353255 and ERS353262) (https://www.ebi.ac.uk/arrayexpress/experiments/E-ERAD 236/samples/?s_page=1%26s_pagesize=25).

2.3. *Source of genome assemblies and annotations*

The ~182 Mb *Hymenolepis microstoma* genome assembly (PRJEB124) and the gene annotations of 12,368 coding genes (Tsai et al., 2013) were downloaded from the WormBase Parasite

database (Howe et al., 2017). Also, the ~163 Mb *Hymenolepis nana* genome assembly (PRJEB508) and the corresponding 13,777 coding genes were retrieved from the WormBase Parasite database (Howe et al., 2017).

2.4. *MiRNA identification in* Hymenolepis microstoma *larvae*

To identify conserved and novel miRNAs from the small RNA libraries, the miRDeep2 software package (Friedländer et al., 2012) was used. The unique sequences were mapped to *H. microstoma* genome and used as input for miRNA prediction as previously described (Cucher et al., 2015; Macchiaroli et al., 2015). The initial miRDeep2 output list of candidate miRNA precursors of each library was manually curated to generate a final high confidence set of miRNAs retaining only candidate precursors with i) miRDeep2 score ≥ 4 ii) mature reads in more than one biological sample iii) star reads and/or seed conservation iv) no match to rRNA, tRNA, mRNA. The secondary structures of putative precursors and clusters and the minimum free energy were predicted using the mfold web server (Zuker, 2003) and RNAfold software (Gruber et al., 2008), respectively. MiRNA annotation and classification of the small RNAseq reads into RNA types (miRNAs, rRNA, tRNA and mRNAs) were performed as previously described (Cucher et al., 2015; Macchiaroli et al., 2015). In addition, mature miRNA sequences were compared by BLASTN against an in-house database of all previously reported cestode miRNA sequences obtained by deep sequencing (Bai et al., 2014; Cucher et al., 2015; Macchiaroli et al., 2015; Basika et al., 2016; Perez et al., 2017). Nucleotide sequence data reported in this paper have been submitted to the miRBase database.

178

179 **2.5. *MiRNA discovery in* Hymenolepis nana *from genome-wide data***

180 To discover miRNAs in *Hymenolepis nana* genome, the miRDeep2 software package
181 (Friedländer et al., 2012) was used. The small RNAseq reads of *H. microstoma* were mapped to
182 *H. nana* genome and miRNAs predicted as previously described (Macchiaroli et al., 2015). In
183 addition, an independent approach that required the *H. nana* genome and was based on the
184 combination of three methods i) miRNA-SOM (Kamenetzky et al., 2016), ii) deepSOM
185 (Stegmayer et al., 2017) and iii) miRNAss (Yones et al., 2018) was used. A unique list of best
186 candidates was obtained as the intersection of the three methods. Briefly, sequences with a
187 minimum free energy threshold of -20 kcal/mol and single-loop folded sequences were selected
188 according to the miRNA biogenesis model (Bartel, 2004). After that, the best candidates to
189 precursor miRNA sequences in *H. nana* were identified as those sequences more similar to high
190 confidence miRNAs of *H. microstoma* in the feature space (Yones et al., 2015). These methods
191 were recently used by our group for genome-wide discovery of miRNA precursor sequences in
192 *E. multilocularis* (Kamenetzky et al., 2016) and *T. solium* (Perez et al., 2017).

193 **2.6. *MiRNA expression profiling in* Hymenolepis microstoma *larvae***

194 The number of reads obtained in a small RNA sequencing experiment can be used as an
195 indicator of the abundance of a given miRNA at a particular life cycle stage (Kato et al., 2009).
196 To analyze miRNA expression, read counts of each individual miRNA in a sample (biological
197 replicate) were normalized to the total number of mature miRNA read counts in that sample as
198 described in Macchiaroli et al. (2015). Then, normalized miRNA reads were averaged between
199 the three biological replicates and the most expressed miRNAs in *H. microstoma* larvae were

determined. Correlation analyses between pairs of independent biological replicates were performed.

2.7. *Evolutionary conservation analysis of* Hymenolepis *miRNAs*

To identify miRNA families within *Hymenolepis*, all-against-all pairwise sequence alignments were computed using BLASTN and all sequences sharing the seed region (position 2-7 of the mature miRNA) were considered to belong to the same *Hymenolepis* miRNA family. To analyze conservation of all expected miRNA families in *Hymenolepis* (Fromm et al., 2013, 2017) mature miRNA sequences were compared to those previously reported present in miRBase 22 for selected phyla; Cnidaria, Nematoda, Arthropoda, Annelida and the subphylum Vertebrata, using only a seed match criteria. To analyze conservation of *Hymenolepis* miRNA sequences across Platyhelminthes, the species used for comparative analysis were selected on the following criteria: i) genome available; ii) data deposited in miRBase v22; iii) the species with a more complete miRNA complement within a same genera; iv) study based on high throughput sequencing. Selected species were *E. canadensis, E. multilocularis, E. granulosus, M. corti, T. solium, T. crassiceps, Schistosoma mansoni, Gyrodactylus salaris and Schmidtea mediterranea.*

2.8. Cluster identification and genomic location of Hymenolepis *miRNAs*

To identify miRNA clusters in both *Hymenolepis* species, the genomic arrangement of the miRNAs identified in this study was assessed. Precursor miRNA sequences were considered to be grouped in clusters if they were in the same scaffold/contig less than 10 kb apart and on the same strand. Alignment of precursor miRNA sequences were performed using MUSCLE (Edgar, 2004) followed by RNAalifold (Bernhart et al., 2008) using default parameters. The phylogenetic analysis of the cluster mir-4989/277 was conducted in MEGA7 (Kumar et al.,

2016). The phylogenetic trees were inferred by Maximum Likelihood using the Tamura 3-parameter model. The topology of the tree with the highest log-likelihood value is shown. The percentage of trees in which the sequences clustered together is shown next to the branches. The support for the node was assessed using 2000 bootstrap replicates. The genomic location (intronic, exonic and intergenic) of all miRNAs identified in this study and the genomic context of mir-10 was assessed by BLAST searches against current annotation of *Hymenolepis* genomes available in WormBase Parasite database v. WBPS9 (WS258). For intronic miRNAs, only miRNAs located in introns of coding genes with a predicted functional annotation were considered. For the analysis of the genomic context of mir-10 only the two neighboring genes were consider (i.e. the closest protein coding gene upstream/downstream with a functional annotation).

3. Results and Discussion

3.1. *MiRNA identification in* Hymenolepis microstoma *and miRNA discovery in* Hymenolepis nana *from genome-wide data*

To identify the repertoire of miRNAs expressed in *H. microstoma* larvae, we sequenced small RNA libraries from three biological replicates of *H. microstoma* larvae. After trimming and collapsing, between 8.4 and 83.9 million reads per sample were mapped to *H. microstoma* genome, representing about ~96% of reads. The general results of the Illumina deep sequencing are shown in **Supplementary Table S1.** We predicted a high confidence repertoire of 37 conserved miRNAs in *H. microstoma*, providing for the first time experimental evidence of miRNA expression in *H. microstoma* (**Table 1**). Of the 37 miRNAs, 26 precursors have

previously been bioinformatically predicted from the *H. microstoma* genome data (Jin et al., 2013) but 11 are new to this study. The latter precursor miRNAs comprised four mir-36 (hmi-mir-36b, hmi-mir-36c, hmi-mir-36d, hmi-mir-36e), two mir-3479 (hmi-mir-3479a and hmi-mir-3479b), hmi-mir-210, hmi-mir-307, hmi-mir-7b, hmi-mir-124a and hmi-mir-277b. The secondary structures of all precursor miRNA sequences identified in *H. microstoma* are shown in **Supplementary Figure S1**. Thus, we expanded the miRNA repertoire of *H. microstoma* highlighting the potential of the deep sequencing approach for miRNA discovery. The repertoire of precursor miRNA sequences and their genomic location are shown in **Supplementary Table S2.**

To discover miRNA precursors in the *H. nana* genome by using small RNAseq data of *H. microstoma*, we first mapped the small RNAseq reads of *H. microstoma* to *H. nana* genome. We obtained a high percentage of genome mapping (average 90%, **Supplementary Table S1**). Then, we predicted a high confidence repertoire of 37 conserved miRNAs in the *H. nana* genome. All precursor miRNAs identified in *H. nana* were conserved in *H. microstoma,* with 81% (mir-210) to 100% (bantam, mir-71, mir-2162) of sequence identity (average 95%**)**. We did not find *H. nana*-specific precursor miRNAs in the genome. The precursor miRNA sequences of *H. nana* and their genomic location are shown in **Supplementary Table S3**.

In addition, we performed a genome-wide discovery of miRNAs in *Hymenolepis nana* by using a different approach based on the combination of three methods. We found a unique list of 36 best candidates to miRNA precursors by the intersection of the three methods (**Supplementary Figure S2).** All best candidates were previously found by using small RNAseq data of *H. microstoma*. This is the first time that miRNAs are described in *H. nana*.

3.2. *MiRNA expression profiling in* Hymenolepis microstoma *larvae*

To analyze miRNA expression profiling in *H. microstoma* larvae the normalized reads per million of each mature miRNA were averaged among the three biological replicates. Correlation analyses between pairs of biological replicates indicated high technical reproducibility and low biological variation ($r > 0.83$). The mature miRNA repertoire and larval expression levels for the 37 mature *H. microstoma* miRNAs are shown in **Table 1**.

For almost all precursor sequences identified (32/37), we detected the corresponding antisense miRNA sequences consistent with the miRNA biogenesis model, adding confidence to the predictions obtained **(Supplementary Figure S1)**. Also, we found that most mature miRNAs in *H. microstoma* (~ 65%) are processed from the 3' arm (**Table 1, Supplementary Figure S1**). This bias was also observed in *Echinococcus* (Cucher et al., 2015; Macchiaroli et al., 2015), *T. crassiceps* (Perez et al., 2017) and *M. corti* (Basika et al., 2016). In addition, this bias toward 3' arm processing was observed in nematodes, fruit fly and plants (de Wit et al., 2009). For three precursor miRNAs (hmic-mir-36b, hmic-mir-210 and hmic-mir-3479b) the antisense miRNA sequence was abundantly expressed (>30% with respect to the mature sequence) (**Supplementary Figure S1)**. Since the antisense miRNA sequence has a different seed sequence that is the principal determinant of the interaction between miRNA and mRNA target, it may indicate additional functions for the same miRNA gene.

Regarding miRNA expression profile, a few miRNAs showed very high expression levels (**Table 1**). The most expressed miRNAs identified in *H. microstoma* larvae were hmic-miR-9-5p, hmic-miR-71-5p and hmic-miR-10-5p, which accounted for about 60% of total miRNA expression. The expression of these miRNAs was followed by hmic-4989-3p and hmic-bantam-3p (**Fig. 1)**. These top five miRNAs accounted for about 70% of the total miRNA expression.

The high expression of these miRNAs in *Hymenolepis larvae* is conserved in the larval stages of the cestode parasites *Echinococcus* (Cucher et al., 2015; Macchiaroli et al., 2015), *M. corti* (Basika et al., 2016) and *T. cracisseps* (Perez et al., 2017), suggesting an essential function in the biology of the parasites. Interestingly, miR-9 and miR-10 are highly conserved miRNAs across metazoans with known roles in neural development (Yuva-Aydemir et al., 2011) and Hox regulation (Lund, 2010; Tehler et al., 2011), respectively. MiR-71, a bilaterian miRNA absent in vertebrates, is known to be involved in lifespan regulation and stress response in *Caenorhabditis elegans* (Zhang et al., 2011; Boulias and Horvitz, 2012) . Recently, miR-71 and miR-10, were predicted to target developmental pathways such as MAPK and Wnt in *Echinococcus* (Macchiaroli et al., 2017). In addition, most miRNAs identified in *H. microstoma* larvae (24/37) showed very low expression levels, less than 1% of total miRNA reads (**Table 1**) and the low the expression is conserved in the larval stages of the cestode parasites *Echinococcus* (Cucher et al., 2015; Macchiaroli et al., 2015), *M. corti* (Basika et al., 2016) and *T. cracisseps* (Perez et al., 2017).

3.3. *Evolutionary origin and conservation analysis of* Hymenolepis *miRNA families*

The 37 miRNAs identified in both *Hymenolepis* species were classified into 27 miRNA families according to the conservation of their seed regions (positions 2–7 of the mature miRNAs). Among them, we found that six miRNA families had multiple members: miR-2 (miR-2a, miR-2b, miR-2c), miR-7 (miR-7a, miR-7b), miR-36 (miR-36b, miR-36c, miR-36d, miR-36e), miR-124 (miR-124a, miR-124b), miR-277 (miR-277a, miR-277b, miR-4989) and miR-3479 (miR-3479a, miR-3479b). All miRNA families identified in *Hymenolepis* were conserved across evolution and the protostomian-specific miR-36 family is the largest in *Hymenolepis*. Regarding their evolutionary origin, we found one eumetazoan-specific miRNA family, 18 bilaterian-

317 specific miRNA families and 8 protostomian-specific miRNA families in *Hymenolepis* (**Fig. 2**).
318 We did not find either lophotrochozoan-specific miRNA families or platyhelminth-specific
319 miRNA families **(Fig. 2)**. The phylogenetic distribution of all expected miRNA families (Fromm
320 et al., 2013, 2017) was based on the classification of miRNA families by Wheeler et al. (2009)
321 and Tarver et al. (2013) and was confirmed by homology searches in miRBase v.22.
322 Interestingly, miR-71 is a bilaterian-specific miRNA that is absent in the subphylum Vertebrata
323 (**Fig. 2**).
324 We found that the small number of conserved miRNA families identified in *Hymenolepis* (~27)
325 is similar to that found in species of *Echinococcus* and *Taenia*, as well as *M. corti*, the trematode
326 *S. mansoni*, the monogenean *Gyrodactylus salaris* and the planarian *S. mediterranea* (**Table 2**).
327 These results agree with the loss of flatworm miRNAs proposed by Fromm et al. (2013) and
328 reflect the low morphological complexity of platyhelminths compared with other metazoans
329 consistent with previous knowledge (Niwa and Slack, 2007; Berezikov, 2011).
330 However, the total number of conserved miRNAs is ~ two-fold lower in parasitic platyhelminths
331 compared to the free-living *S. mediterranea* (**Table 2**). This may be due to a reduction in the
332 number of members of almost all miRNA families in parasitic platyhelminths compared to free-
333 living *S. mediterranea* (with exception of miR-3479 and miR-36). The reduced complement of
334 evolutionarily conserved miRNAs found in both *Hymenolepis* species may reflect their parasitic
335 lifestyle as previously described for other cestodes (Macchiaroli et al., 2015). Many microRNA
336 families are deeply conserved in bilaterian animals and display similar tissue specificities
337 between divergent species, suggesting a role in the evolution of tissue identity (Christodoulou et
338 al., 2010). The loss of some of these deeply conserved miRNA families in tapeworms may be
339 related to the loss or reduction of tissues and organs. For example, miRNAs associated with
340 locomotor related cilia (miR-29), gut (miR-216/miR-283, miR-278) and sensory organs (miR-

2001) were specifically lost in tapeworms, probably reflecting the reduction or loss of these cells and organs.

Some differences between the miRNA repertoires of *Hymenolepis* and other parasitic platyhelminths were found. Interestingly, we found four members of the miR-36 miRNA family in *Hymenolepis*, whereas only two members of miR-36 were described in the cestodes *Echinococcus*, *Taenia* and *M. corti* and the trematode *S. mansoni.* We found members of miR-210 family in *Hymenolepis* that are not found in *Echinococcus* and *Schistosoma*. Recently, two members of the miRNA mir-210 family were identified by us for the first time in *M. corti:* mco-miR-12065-3p and mco-miR-12066-3p (Basika et al., 2016) and one member in *Taenia* (Perez et al., 2017).

We did not identify miR-8 and miR-1992 orthologs in *Hymenolepis*, consistent with a previous report from *H. microstoma* (Jin et al., 2013) and *M. corti* (Basika et al., 2016). Also, miR-8 was not identified in *Taenia* (Perez et al., 2017). However, these two miRNAs were found to be expressed in *Echinococcus* (Bai et al., 2014; Cucher et al., 2015; Macchiaroli et al., 2015), suggesting that miR-8 and miR-1992 may have been lost during evolution in *H. microstoma* **(Table 2).** We also searched for mir-8 and mir-1992 orthologs bioinformatically within the genome of *H. microstoma* but we did not find them. However, we cannot rule out that these miRNAs that are not expressed in the larval stage here analyzed could be present in the genomes. It would be interesting to analyze small RNAseq data of other stages in order to confirm if these two miRNAs and other known cestode miRNA families are identified, especially miR-1992 that is the only lophotrochozoan-specific miRNA present in cestodes.

3.4. *MiRNA clusters in* Hymenolepis

MiRNAs can be grouped into clusters in the genome if they are less than 10 kb apart (miRBase v22) suggesting co-expression as a single transcriptional unit (Bartel, 2004). MiRNA clusters have been found in the genomes of many species, including helminth parasites (miRBase 22). To investigate the presence of miRNA clusters in *Hymenolepis*, the genomic arrangement of the 37 miRNAs identified in this study was assessed. We found three miRNA clusters conserved in both *Hymenolepis* species: mir-71/2c/2b, mir-277a/4989 and mir-36c/36d/36e. Each miRNA cluster comprised a genomic region of up to 320 bp and located in intergenic regions (**Supplementary Table S4**). The predicted secondary structure of the three clusters found in *Hymenolepis* is shown in **Fig. 3.** The miRNA clusters mir-71/2c/2b and mir-277a/4989 were previously reported in *H. microstoma* (Jin et al., 2013), whereas the miRNA cluster mir-36c/36d/36e was identified for the first time in the class Cestoda in this study. All three miRNA clusters were described for the first time in *H. nana.*

It has been proposed that mir-1 and mir-133 form another cluster in *H. microstoma* (Jin et al., 2013). Although miR-1 and miR-133 clustering is highly conserved across metazoan species (Campo-Paysaa et al., 2011), we found that mir-1 is located 15 kb far from mir-133 suggesting that these miRNAs are not co-expressed as a single transcriptional unit in *Hymenolepis*, consistent with the situation reported in *Echinococcus* (Cucher et al., 2015; Macchiaroli et al., 2015), *M. corti* (Basika et al., 2016) and *T. solium* (Perez et al., 2017).

It is estimated that more than 40% of human miRNAs and more than 30% in worms and flies are found in cluster with other miRNAs (Griffiths-Jones et al., 2008). Here, we found a smaller percentage of miRNAs in clusters in both *Hymenolepis* genomes ~ 20% (8/37) and other cestodes analyzed. The fact that miRNA clusters are conserved across evolution suggests evolutionary and functional importance.

3.4.1. *The miRNA cluster mir-71/mir-2c/mir-2b in* Hymenolepis

The miRNA cluster mir-71/mir-2c/mir-2b is a protostomian-specific miRNA cluster highly conserved within platyhelminths. The mir-71 cluster with members of the mir-2 family is conserved in *Schistosoma* (Huang et al., 2009; de Souza Gomes et al., 2011) and *S. mediterranea* (Palakodeti et al., 2006), among other species within platyhelminths. Also, clustering of mir-71 and mir-2 has been found in nematodes indicating wider evolutionary conservation (Winter et al., 2012). Phylogenetic analyses have been done recently in other platyhelminths that showed multiple copies of mir-71 cluster (Fromm et al., 2013; Jin et al., 2013). We did not find additional copies of this cluster in *Hymenolepis* concurring with previous studies from *Echinococcus*, *Mesocestoides corti* and *Taenia solium* (Cucher et al., 2015; Macchiaroli et al., 2015; Basika et al., 2016; Perez et al., 2017). Regarding the evolutionary origin of the members of this cluster, miR-71 is a bilaterian-specific miRNA family absent in vertebrata, whereas miR-2 is a protostomian-specific miRNA family. Whether the conserved genomic arrangement among platyhelminths means functional conservation of this cluster within this lineage remains to be investigated.

Regarding expression of *H. microstoma* miRNAs encoded in clusters, we found different levels of expression between members of the cluster mir-71/2c/2b. This result agree with previous observations in *Echinococcus* (Cucher et al., 2015), *M. corti* (Basika et al., 2016) and *T. crassiceps* (Perez et al., 2017). This could be explained by the fact that miRNA themselves are subject to sophisticated regulation through control of miRNA processing, RNA editing or miRNA decay (Krol et al., 2010).

3.4.2. *The miRNA cluster mir-277a/4989 in* Hymenolepis

The miRNA cluster mir-277a/4989 is a lophotrochozoan-specific cluster highly conserved across Platyhelminthes. The alignment of the *Hymenolepis* cluster mir-277/4989 and its orthologs across Plathyhelminthes is shown in **Fig. 4A**. The mature miRNA sequences miR-4989 and miR-277a are located in the 3'arm of their corresponding precursor sequences. The alignment showed a high level of nucleotide conservation in the mature miRNA sequences of all platyhelminths analyzed, especially in the seed regions, and the whole sequences are highly conserved only in the Cestoda lineage. Although the two members of the cluster mir-277a/4989 belong to the protostomian miRNA family mir-277, their genomic organization in cluster was only found in lophotrochozoan species (miRBase v22). One additional copy of a cluster that contains two members of the mir-277 family was described in the free living *S. mediterranea* (Palakodeti et al., 2006). Recently, cluster mir-277/4989 was found in the trematode *Schistosoma mansoni* (Protasio et al., 2017). Interestingly, the results of that study suggest that the cluster mir-277/4989 might play a dominant role in post transcriptional regulation during development of juvenile worms in *Schistosoma mansoni*. Whether this cluster plays a developmental role in *Hymenolepis* remains to be investigated. We also performed for the first time a phylogenetic analysis of the cluster miR-277a/4989 that confirmed the relationship of the parasitic platyhelminths highlighting the potential of miRNAs as phylogenetic markers (**Fig. 4B).**

3.4.3. *The miRNA cluster mir-36c/36d/36e in* Hymenolepis

The miRNA cluster mir-36c/36d/36e was identified for the first time in the class Cestoda in this study. Although this cluster is not conserved in Platyhelminthes (Lophotrochozoa), members of the mir-36 family were found to be in cluster with members of the same or other miRNA family

in *S. mediterranea* (sme-mir-36c/36a and sme-mir-36b/let7b) and *S. mansoni* (sma-mir-36b/8406). In addition, the miRNA family miR-36 has a protostomian origin, and also their arrangement in cluster with members of the same family. Interestingly, this cluster is conserved in Nematoda (Ecdysozoa), where this miRNA family has multiple members organized in clusters. In the free-living *C. elegans* the miR-36 cluster is composed of seven members of this family (mir-35 to mir-41). In this model organism, deletion of the miR-36 cluster produces embryonic and larval lethality suggesting an essential role in early development (Alvarez-Saavedra and Horvitz, 2010). In parasitic nematodes such as *Brugia malayi*, four members of the miR-36 family have been identified (Poole et al., 2014) with three of them (miR-36c/36d/36c) organized in a cluster with five members of other families. However, the whole cluster is contained in a genomic region of ~12,000 bp (miRBase v.22). Recently, miR-36 from *S. mansoni*, among others, was found to be secreted *in vitro* (Samoil et al., 2018).

3.5. *Genomic arrangement of* Hymenolepis *miRNAs*

The genomic arrangement of the 37 miRNAs identified in both *Hymenolepis* species was analyzed. This analysis showed that most precursor miRNAs 92% (34/37) were located in intergenic regions distant from annotated genes, and 8% (3/37) were located in introns of protein coding genes in *Hymenolepis*. The bias in the genomic location found in this study was also observed in other platyhelminths such as *E. multilocularis* (Cucher et al., 2015) and *S. japonicum* (Cai et al., 2011) where 81% and 90% of the miRNA complement, respectively, was located in intergenic regions. The three intronic miRNAs found in this study in *Hymenolepis* were mir-190, mir-96 and mir-3479b (**Supplementary Table S2 and Supplementary Table S3**).

3.5.1. The intronic miRNAs in *Hymenolepis*

We also performed a comparative analysis of the genomic location of the three intronic miRNAs identified in each *Hymenolepis* species across selected platyhelminths with available genomes: *S. mediterranea, S. mansoni, E. multilocularis, E. granulosus, E. canadensis, T. solium* and *M. corti*. The results of this analysis are shown in **Table 3**.

We found that the bilaterian miRNA mir-190 was located within the intron of the protein coding gene *talin* in both *Hymenolepis* species. Also, we report for the first time that the genomic location of miR-190 was conserved in *M. corti, T. solium, E. granulosus* and *E. canadensis.* These findings are consistent with previous studies, where mir-190 has been found in the intronic region of the *talin* gene in the platyhelminths *H. microstoma, E. multilocularis* and *S. mansoni* (**Table 3)** and in higher metazoans including *Homo sapiens* (Campo-Paysaa et al., 2011). Recently, mir-190 was also found to be located in an intron of the gene encoding *talin* protein in three Opisthorchiids (Ovchinnikov et al., 2015). Several functions have been proposed for miR-190 in mammals, for example it regulates *neurogenic differentiation 1* (NeuroD) activity and can also interact with other transcription factors that regulate neurogenesis, such as Pax6 (Zheng et al., 2012). *Talin* is an adhesion plaque protein that links the integrin-mediated cell–matrix contacts to the actin cytoskeleton. These interactions play an important role in regulating synapse morphology and number, neuron-neuron and neuromuscular synaptic transmission, and neuroplasticity that modulates neuronal cell proliferation, migration, and differentiation (Venstrom and Reichardt, 1993).

In addition, we found that the bilaterian miRNA mir-96 was located within the intron of the protein coding gene *Fras1-related extracellular matrix protein (Frem1)* in *H. microstoma* consistent with Jin et al. (2013). Also, in this study, we found that this arrangement was conserved in *E. granulosus,* concurring with previous reports from *E. multilocularis* and *S.*

mediterranea (**Table 3)**. Also, in this study, we found that this arrangement was conserved in *E. granulosus*, concurring with previous reports from *E. multilocularis* and *S. mediterranea* (**Table 3)**. Unlike miR-190, the intronic location of miR-96 is not conserved in mammals. Recently, it was shown that miR-96 is a sensory organ-specific miRNA expressed in the mammalian cochlea that regulates the progression of differentiation of inner and outer hair cells during development (Kuhn et al., 2011). The extracellular matrix protein *Frem1* plays a role in epidermal differentiation and is essential for epidermal adhesion during embryonic development in mice (Smyth et al., 2004).

Here we found for the first time that the bilaterian miRNA mir-3479b (family miR-92/25, seed AUUGCA), was located within the intron of the protein coding gene *minichromosome maintenance complex component 2* (*mcm2*) in *H. microstoma*. Also, in this study, we found that this arrangement was conserved in *E. granulosus* consistent with previous reports in *E. multilocularis* (Cucher et al., 2015). Interestingly, the human miRNA mir-25 (family miR-92/25, seed AUUGCA) is a member of the miRNA cluster mir-25/93/106b that is located in the thirteenth intron of the gene *mcm-7* in *Homo sapiens* (Rodriguez et al., 2004). Recently, the miR-25-93-106b cluster was shown to regulate tumor metastasis and immune evasion (Cioffi et al., 2017). Also, miR-92 was shown to be part of a cancer miRNA signature composed by a large portion of overexpressed miRNAs (Volinia et al., 2006). The protein encoded by the *mcm7* belongs to the highly conserved *mcm* protein family of DNA helicases that are essential for the initiation of genome replication in eukaryotes. High expression levels of this protein promote cancer progression (Qu et al., 2017).

Interestingly, the functions of both the intronic miRNAs and their host genes suggest that they may be functionally related in higher organisms. It remains to be determined whether the intronic miRNAs are functionally related with their host gene in *Hymenolepis* and in other flatworms.

The three intronic miRNAs found here are bilaterian-specific and are located in the same orientation of their corresponding host genes suggesting that they may depend on their host gene promoter for transcription resulting in a coordinated expression. Also, the small length of mir-190 and mir-3479b harboring introns in all species supports this hypothesis. The only exception was mir-190 harboring intron in *S. mansoni* (**Table 3**). In addition, we found that all intronic miRNAs were located in the last introns of their corresponding host genes and this structure is conserved among the platyhelminths analyzed (**Table 3**). It would be interesting to determine whether this location has implications for the regulation of their expression. Interestingly, the three intronic miRNAs and their host genes are expressed at very low levels; less than 1% (**Table 1)** and less than 100 FPKM, respectively. Finally, the lack of conservation of the genomic location of mir-96 and mir-3479b in *H. nana* genome could be due to the draft nature of the genome assembly in this species. The *H. nana* genome is highly fragmented (16,212 scaffolds in the WBPS9 release available from Wormbase Parasite compared to 3,643 scaffolds for *H. microstoma*) making the genome localization of miRNAs and the analysis of their genomic context more difficult. The same interpretation may be valid for *E. canadensis*.

3.5.2. The mir-10 genomic organization in *Hymenolepis*

Mir-10 is one of the most ancient miRNAs that is present in all species of metazoans. In most bilateral animal species, mir-10 is encoded within Hox clusters (Campo-Paysaa et al., 2011). In this work, we found that one of the neighboring genes of mir-10 in *H. microstoma* was a Hox gene (HmN_000772500) and was located 41 kb apart from mir-10. We also found that this genome arrangement was conserved in *E. granulosus* and *E. canadensis* (**Supplementary Table S5**). These findings are consistent with our previous results in the cestodes *E. multilocularis* (Cucher et al., 2015) and *T. solium* (Perez et al., 2017) where the neighboring genes in these

species were the Homeobox protein Hox B4a (EmuJ_000813900) and the Homeobox protein mab 5 (TsM_000864600), respectively. These genes were found to be located 25 kb far from mir-10 in both species (**Supplementary Table S5**). In addition, we found that one of the neighboring genes of mir-10 in the trematode *S. mansoni* was a Gsx family Homeobox protein (Smp_081620) and was located 84 kb far from mir-10, consistent with previous results (de Souza Gomes et al., 2011). Recently, miR-10 was predicted to target Hox genes and transcription factors in *T. solium* (Perez et al., 2017). Also, miR-10 was predicted to target one homeobox containing protein from the Meis family in all *Echinococcus* species (Macchiaroli et al., 2017). In mammals, mir-10a resides upstream from Hoxb4 and mir-10b upstream from Hoxd4 (Lund, 2010). We found that the flatworm Homeobox proteins above mentioned are all orthologs of Hox4 in vertebrates (Tsai et al., 2013). Here, we showed a strong conservation of the genomic organization of mir-10 and Hox4 gene in flatworms. However, the distance between mir-10 and the corresponding Hox4 gene is greater than in vertebrates (i.e. whole HoxA cluster ~100 kb) (Santini et al., 2003). Whether the genomic organization of mir-10 in *Hymenolepis* and related flatworms is functionally linked with the Hox genes remains to be investigated.

In many species, miR-10 is co-expressed with Hox genes (Lund, 2010; Tehler et al., 2011). Interestingly, we found that miR-10 was one of the most expressed miRNAs in *H. microstoma* larvae. It would be interesting to determine whether the neighboring Hox gene is expressed in a similar pattern in this parasite stage.

Parasitic flatworms, including *Hymenolepis,* have the smallest complement of Homeobox genes of any studied bilaterian animal (Tsai et al., 2013). In addition, Hox genes are at least partially dispersed within the genome and flatworms and may not exhibit temporal colinearity in the expression patterns characteristic of Hox genes of many other animal groups (Olson, 2008). As

mentioned above, parasitic flatworms have only one copy of miR-10 in the genome, whereas mammals and zebrafish have two and five copies, respectively (Tehler et al., 2011).

The other neighboring gene of mir-10 in *H. microstoma* was HmN_002012500, a protein coding gene that is the ortholog of the nuclear hormone receptor protein nhr-25 of *C. elegans*. Thus, mir-10 is flanked by two transcription factors in *H. microstoma*, and this genomic organization is conserved in *E. multilocularis* and *E. granulosus*, but not in *S. mansoni*.

3.6. Hymenolepis *miRNAs as potential biomarkers*

In this work, we identified some mature miRNAs in *Hymenolepis* that could represent potential biomarkers or therapeutic targets. Some parasite miRNAs are protostome-specific such as miR-277 and bantam or bilaterian-specific but absent in the vertebrate host such miR-71. Others are bilateria-specific but divergent at the sequence level from their host orthologs such as miR-3479 (miR-92 family). In addition, other miRNAs are highly conserved across metazoans such as miR-10. Interestingly, several recent works have shown that members of these families are secreted in parasitic helminths and can be detected in the serum of the host.

In recent work from our group, miR-71 and miR-277 from the cestode *T. crassiceps* were found to be secreted *in vitro* (Ancarola et al., 2017). Recently, miR-71, bantam and miR-3479 from the trematode *S. mansoni* were found to be secreted *in vitro*, and also miR-71 and bantam were detected in serum of infected mice (Samoil et al., 2018). In a previous study, miR-277, miR-3479 and bantam from *S. mansoni* were detected in the serum of infected mice and human patients and could distinguish infected individuals with high specificity and sensitivity (Hoy et al., 2014). In addition, miR-10 was only found in serum of mice infected with *S. mansoni* (Hoy et al., 2014). Also, miR-100 (miR-10 family), bantam and miR-71 from the filarial nematode

574 *Litomosoides sigmodontis* were among the most abundant miRNAs detected in the serum of
575 infected mice, thus confirming *in vivo* secretion of parasite miRNAs (Buck et al., 2014). Also,
576 miR-10 and miR-71 from the filarial nematode *Dirofilaria immitis* were detected in plasma of
577 infected dogs (Tritten et al., 2014). Among others, miR-71, miR-100 and bantam from the filarial
578 nematode *Onchocerca volvulus* were detected in serum of infected humans (Quintana et al.,
579 2015).
580 Taken together, these results suggest that parasite miRNAs might be evaluated as novel
581 biomarkers for detecting helminth infection. It will be important to determine whether these
582 miRNAs can also be secreted by *Hymenolepis* and to assess whether they may be detected in
583 serum of infected humans. Since miRNAs are main components of the eukaryotic transcriptome
584 they require further investigation in *Hymenolepis*. Parasite miRNAs could complement existing
585 diagnostic techniques to improve diagnosis and may provide a platform for further research in
586 the area of therapeutic targets of neglected parasites.
587

588 **4. Conclusions**

589 The recent availability of the genome assemblies of *Hymenolepis* and the limited knowledge
590 about miRNAs in these zoonotic neglected parasites encouraged us to identify and characterize
591 these small, non-coding RNAs that have recently emerged as potential biomarkers and
592 therapeutic targets of infections. In this work we described for the first time the expression
593 profile of the miRNA complement in *H. microstoma,* and discovered miRNA genes at the
594 genome-wide level in *H. nana* using two different approaches. The high expression of a few
595 miRNAs in the larval stage of *H. microstoma* is conserved in other cestodes suggesting that these
596 miRNAs may have important roles in development, survival and for host-parasite interplay. We
597 found a reduced complement of evolutionarily conserved miRNAs in both *Hymenolepis* species,

putatively reflecting their low morphological complexity and parasitic lifestyle. We performed a comparative analysis of the identified miRNAs and examined their genomic arrangement across the Cestoda providing new insights about their post-transcriptional mechanisms. Our results showed that most of the miRNAs in *Hymenolepis* are located in intergenic regions implying that they are independently transcribed. Interestingly, we found a *Hymenolepis*-specific cluster composed of three members of the mir-36 family. Also, we found that one of the neighboring genes of mir-10 in *H. microstoma* was a Hox gene as in most bilaterial species. In addition, some *Hymenolepis* miRNAs are protostome-specific or bilaterian-specific, but divergent from host orthologs and therefore could represent novel biomarkers of *Hymenolepis* infection. This study provides a valuable resource for further experimental research in cestode biology that might lead to improved detection and control of these neglected parasites.

Conflict of interest

All authors declare no conflict of interest.

Funding

This work was supported by the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), Argentina (grant number PICT 2013 N° 2121) and the Wellcome Trust (grant WT 098051). We thank Gabriel Lichtenstein from IMPaM for technical assistance on bioinformatic 616 analyses. All bioinformatic analysis were performed in a local server at Instituto de Investigaciones en Microbiología y Parasitología Médicas (IMPaM) that is part of Sistema Nacional de Computación de Alto Desempeño (SNCAD) ID 924 of Ministerio de Ciencia, Tecnología e Innovación Productiva (MINCyT) and in the Research Institute for Signals, Systems and Computational Intelligence, (sinc(i)), FICH-UNL/CONICET, Santa Fe, Argentina.

References

Alvarez-Saavedra, E., Horvitz, H.R., 2010. Many families of C. elegans microRNAs are not essential for development or viability. Curr. Biol. 20, 367–73. doi:10.1016/j.cub.2009.12.051

Ancarola, M.E., Marcilla, A., Herz, M., Macchiaroli, N., Perez, M., Asurmendi, S., Brehm, K., Poncini, C., Rosenzvit, M., Cucher, M., 2017. Cestode parasites release extracellular vesicles with microRNAs and immunodiagnostic protein cargo. Int. J. Parasitol. 47, 675–686. doi:10.1016/j.ijpara.2017.05.003

Bai, Y., Zhang, Z., Jin, L., Kang, H., Zhu, Y., Zhang, L., Li, X., Ma, F., Zhao, L., Shi, B., Li, J., McManus, D.P., Zhang, W., Wang, S., 2014. Genome-wide sequencing of small RNAs reveals a tissue-specific loss of conserved microRNA families in Echinococcus granulosus. BMC Genomics 15, 736. doi:10.1186/1471-2164-15-736

Bartel, D.P., 2018. Metazoan MicroRNAs. Cell 173, 20–51. doi:10.1016/j.cell.2018.03.006

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233. doi:10.1016/j.cell.2009.01.002

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–97.

Basika, T., Macchiaroli, N., Cucher, M., Espínola, S., Kamenetzky, L., Zaha, A., Rosenzvit, M., Ferreira, H.B., 2016. Identification and profiling of microRNAs in two developmental stages of the model cestode parasite Mesocestoides corti. Mol. Biochem. Parasitol. 210, 37–49. doi:10.1016/j.molbiopara.2016.08.004

Berezikov, E., 2011. Evolution of microRNA diversity and regulation in animals. Nat. Rev. Genet. 12, 846–860. doi:10.1038/nrg3079

Bernhart, S.H., Hofacker, I.L., Will, S., Gruber, A.R., Stadler, P.F., 2008. RNAalifold: improved consensus structure prediction for RNA alignments. BMC Bioinformatics 9, 474. doi:10.1186/1471-2105-9-474

Boulias, K., Horvitz, H.R., 2012. The C. elegans microRNA mir-71 acts in neurons to promote germline-mediated longevity through regulation of DAF-16/FOXO. Cell Metab. 15, 439–50. doi:10.1016/j.cmet.2012.02.014

Britton, C., Winter, A.D., Gillan, V., Devaney, E., 2014. microRNAs of parasitic helminths - Identification, characterization and potential as drug targets. Int. J. Parasitol. Drugs drug Resist. 4, 85–94. doi:10.1016/j.ijpddr.2014.03.001

Buck, A.H., Coakley, G., Simbari, F., McSorley, H.J., Quintana, J.F., Le Bihan, T., Kumar, S., Abreu-Goodger, C., Lear, M., Harcus, Y., Ceroni, A., Babayan, S.A., Blaxter, M., Ivens, A.,

Maizels, R.M., 2014. Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. Nat. Commun. 5, 5488. doi:10.1038/ncomms6488

Cai, P., Gobert, G.N., McManus, D.P., 2016. MicroRNAs in Parasitic Helminthiases: Current Status and Future Perspectives. Trends Parasitol. 32, 71–86. doi:10.1016/j.pt.2015.09.003

Cai, P., Hou, N., Piao, X., Liu, S., Liu, H., Yang, F., Wang, J., Jin, Q., Wang, H., Chen, Q., 2011. Profiles of Small Non-Coding RNAs in Schistosoma japonicum during Development. PLoS Negl. Trop. Dis. 5, e1256. doi:10.1371/journal.pntd.0001256

Campo-Paysaa, F., Sémon, M., Cameron, R.A., Peterson, K.J., Schubert, M., 2011. microRNA complements in deuterostomes: origin and evolution of microRNAs. Evol. Dev. 13, 15–27. doi:10.1111/j.1525-142X.2010.00452.x

Christodoulou, F., Raible, F., Tomer, R., Simakov, O., Trachana, K., Klaus, S., Snyman, H., Hannon, G.J., Bork, P., Arendt, D., 2010. Ancient animal microRNAs and the evolution of tissue identity. Nature 463, 1084–8. doi:10.1038/nature08744

Cioffi, M., Trabulo, S.M., Vallespinos, M., Raj, D., Kheir, T.B., Lin, M.-L., Begum, J., Baker, A.-M., Amgheib, A., Saif, J., Perez, M., Soriano, J., Desco, M., Gomez-Gaviro, M.V., Cusso, L., Megias, D., Aicher, A., Heeschen, C., 2017. The miR-25-93-106b cluster regulates tumor metastasis and immune evasion via modulation of CXCL12 and PD-L1. Oncotarget 8, 21609–21625. doi:10.18632/oncotarget.15450

Coghlan, A., Mitreva, M., Berriman, M., 2017. Comparative genomics of the major parasitic worms. bioRxiv.

Cucher, M., Macchiaroli, N., Kamenetzky, L., Maldonado, L., Brehm, K., Rosenzvit, M., 2015. High-throughput characterization of Echinococcus spp. metacestode miRNomes. Int. J. Parasitol. 45, 253–267. doi:10.1016/j.ijpara.2014.12.003

Cucher, M., Macchiaroli, N., Kamenetzky, L., Maldonado, L., Brehm, K., Rosenzvit, M.C., 2015. High-throughput characterization of Echinococcus spp. metacestode miRNomes. Int. J. Parasitol. 45, 253–267. doi:10.1016/j.ijpara.2014.12.003

Cunningham, L.J., Olson, P.D., 2010. Description of Hymenolepis microstoma (Nottingham strain): a classical tapeworm model for research in the genomic era. Parasit. Vectors 3, 123. doi:10.1186/1756-3305-3-123

de Souza Gomes, M., Muniyappa, M.K., Carvalho, S.G., Guerra-Sá, R., Spillane, C., 2011. Genome-wide identification of novel microRNAs and their target genes in the human parasite Schistosoma mansoni. Genomics 98, 96–111. doi:10.1016/j.ygeno.2011.05.007

de Wit, E., Linsen, S.E. V, Cuppen, E., Berezikov, E., 2009. Repertoire and evolution of miRNA genes in four divergent nematode species. Genome Res. 19, 2064–2074. doi:10.1101/gr.093781.109

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. doi:10.1093/nar/gkh340

Filipowicz, W., Bhattacharyya, S.N., Sonenberg, N., 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. 9, 102–114. doi:10.1038/nrg2290

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 Res. 40, 37–52. doi:10.1093/nar/gkr688

Fromm, B., Ovchinnikov, V., Høye, E., Bernal, D., Hackenberg, M., Marcilla, A., 2017. On the presence and immunoregulatory functions of extracellular microRNAs in the trematode *Fasciola hepatica*. Parasite Immunol. 39, e12399. doi:10.1111/pim.12399

Fromm, B., Worren, M.M., Hahn, C., Hovig, E., Bachmann, L., 2013. Substantial loss of conserved and gain of novel MicroRNA families in flatworms. Mol. Biol. Evol. 30, 2619–28. doi:10.1093/molbev/mst155

Ghildiyal, M., Zamore, P.D., 2009. Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94–108. doi:10.1038/nrg2504.Small

Griffiths-Jones, S., Saini, H.K., van Dongen, S., Enright, A.J., 2008. miRBase: tools for microRNA genomics. Nucleic Acids Res. 36, D154–8. doi:10.1093/nar/gkm952

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neubock, R., Hofacker, I.L., 2008. The Vienna RNA websuite. Nucleic Acids Res. 36, W70–4. doi:10.1093/nar/gkn188

He, L., Hannon, G.J., 2004. MicroRNAs: small RNAs with a big role in gene regulation. Nat. Rev. Genet. 5, 522–531. doi:10.1038/nrg1379

Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P., Berriman, M., 2017. WormBase ParaSite - a comprehensive resource for helminth genomics. Mol. Biochem. Parasitol. 215, 2–10. doi:10.1016/j.molbiopara.2016.11.005

Hoy, A.M., Lundie, R.J., Ivens, A., Quintana, J.F., Nausch, N., Forster, T., Jones, F., Kabatereine, N.B., Dunne, D.W., Mutapi, F., Macdonald, A.S., Buck, A.H., 2014. Parasite-derived microRNAs in host serum as novel biomarkers of helminth infection. PLoS Negl. Trop. Dis. 8, e2701. doi:10.1371/journal.pntd.0002701

Huang, J., Hao, P., Chen, H., Hu, W., Yan, Q., Liu, F., Han, Z.-G., 2009. Genome-wide identification of Schistosoma japonicum microRNAs using a deep-sequencing approach. PLoS One 4, e8206. doi:10.1371/journal.pone.0008206

Hutvagner, G., Zamore, P.D., 2002. A microRNA in a multiple-turnover RNAi enzyme complex. Science 297, 2056–2060. doi:10.1126/science.1073827

Jin, X., Lu, L., Su, H., Lou, Z., Wang, F., Zheng, Y., Xu, G.-T., 2013. Comparative analysis of known miRNAs across platyhelminths. FEBS J. 280, 3944–51. doi:10.1111/febs.12395

Kamenetzky, L., Stegmayer, G., Maldonado, L., Macchiaroli, N., Yones, C., Milone, D.H., 2016. MicroRNA discovery in the human parasite Echinococcus multilocularis from genome-wide data. Genomics 107, 274–280. doi:10.1016/j.ygeno.2016.04.002

Kato, M., de Lencastre, A., Pincus, Z., Slack, F.J., 2009. Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during Caenorhabditis elegans development. Genome Biol. 10, R54. doi:10.1186/gb-2009-10-5-r54

Krol, J., Loedige, I., Filipowicz, W., 2010. The widespread regulation of microRNA biogenesis, function and decay. Nat. Rev. Genet. 11, 597–610. doi:10.1038/nrg2843

Kuhn, S., Johnson, S.L., Furness, D.N., Chen, J., Ingham, N., Hilton, J.M., Steffes, G., Lewis, M.A., Zampini, V., Hackney, C.M., Masetto, S., Holley, M.C., Steel, K.P., Marcotti, W., 2011. miR-96 regulates the progression of differentiation in mammalian cochlear inner and outer hair cells. Proc. Natl. Acad. Sci. U. S. A. 108, 2355–60. doi:10.1073/pnas.1016646108

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874. doi:10.1093/molbev/msw054

Lund, A.H., 2010. miR-10 in development and cancer. Cell Death Differ. 17, 209–14. doi:10.1038/cdd.2009.58

Macchiaroli, N., Cucher, M., Zarowiecki, M., Maldonado, L., Kamenetzky, L., Rosenzvit, M.C., 2015. microRNA profiling in the zoonotic parasite Echinococcus canadensis using a high-throughput approach. Parasit. Vectors 8, 83. doi:10.1186/s13071-015-0686-8

Macchiaroli, N., Maldonado, L.L., Zarowiecki, M., Cucher, M., Gismondi, M.I., Kamenetzky, L., Rosenzvit, M.C., 2017. Genome-wide identification of microRNA targets in the neglected disease pathogens of the genus Echinococcus. Mol. Biochem. Parasitol. 214, 91–100. doi:10.1016/j.molbiopara.2017.04.001

Muehlenbachs, A., Bhatnagar, J., Agudelo, C.A., Hidron, A., Eberhard, M.L., Mathison, B.A., Frace, M.A., Ito, A., Metcalfe, M.G., Rollin, D.C., Visvesvara, G.S., Pham, C.D., Jones, T.L., Greer, P.W., Vélez Hoyos, A., Olson, P.D., Diazgranados, L.R., Zaki, S.R., 2015. Malignant Transformation of *Hymenolepis nana* in a Human Host. N. Engl. J. Med. 373, 1845–1852. doi:10.1056/NEJMoa1505892

Niwa, R., Slack, F.J., 2007. The evolution of animal microRNA function. Curr. Opin. Genet. Dev. 17, 145–150. doi:10.1016/j.gde.2007.02.004

Olson, P.D., 2008. Hox genes and the parasitic flatworms: new opportunities, challenges and lessons from the free-living. Parasitol. Int. 57, 8–17. doi:10.1016/j.parint.2007.09.007

Olson, P.D., Yoder, K., Fajardo L-G, L.F., Marty, A.M., van de Pas, S., Olivier, C., Relman, D. a, 2003. Lethal invasive cestodiasis in immunosuppressed patients. J. Infect. Dis. 187, 1962–1966. doi:10.1086/375357

Olson, P.D., Zarowiecki, M., Kiss, F., Brehm, K., 2012. Cestode genomics - progress and prospects for advancing basic and applied aspects of flatworm biology. Parasite Immunol. 34, 130–150. doi:10.1111/j.1365-3024.2011.01319.x

Ovchinnikov, V.Y., Afonnikov, D.A., Vasiliev, G. V, Kashina, E. V, Sripa, B., Mordvinov,

V.A., Katokhin, A. V, 2015. Identification of microRNA genes in three opisthorchiids. PLoS Negl. Trop. Dis. 9, e0003680. doi:10.1371/journal.pntd.0003680

Palakodeti, D., Smielewska, M., Graveley, B.R., 2006. MicroRNAs from the Planarian Schmidtea mediterranea: A model system for stem cell biology. RNA 12, 1640–1649. doi:10.1261/rna.117206

Perez, M.G., Macchiaroli, N., Lichtenstein, G., Conti, G., Asurmendi, S., Milone, D.H., Stegmayer, G., Kamenetzky, L., Cucher, M., Rosenzvit, M.C., 2017. microRNA analysis of Taenia crassiceps cysticerci under praziquantel treatment and genome-wide identification of Taenia solium miRNAs. Int. J. Parasitol. 47, 643–653. doi:10.1016/j.ijpara.2017.04.002

Poole, C.B., Gu, W., Kumar, S., Jin, J., Davis, P.J., Bauche, D., McReynolds, L.A., 2014. Diversity and Expression of MicroRNAs in the Filarial Parasite, Brugia malayi. PLoS One 9, e96498. doi:10.1371/journal.pone.0096498

Protasio, A. V, van Dongen, S., Collins, J., Quintais, L., Ribeiro, D.M., Sessler, F., Hunt, M., Rinaldi, G., Collins, J.J., Enright, A.J., Berriman, M., 2017. MiR-277/4989 regulate transcriptional landscape during juvenile to adult transition in the parasitic helminth Schistosoma mansoni. PLoS Negl. Trop. Dis. 11, e0005559. doi:10.1371/journal.pntd.0005559

Qu, K., Wang, Z., Fan, H., Li, J., Liu, J., Li, P., Liang, Z., An, H., Jiang, Y., Lin, Q., Dong, X., Liu, P., Liu, C., 2017. MCM7 promotes cancer progression through cyclin D1-dependent signaling and serves as a prognostic marker for patients with hepatocellular carcinoma. Cell Death Dis. 8, e2603. doi:10.1038/cddis.2016.352

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. Parasit. Vectors 8, 58. doi:10.1186/s13071-015-0656-1

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., Bradley, A., 2004. Identification of mammalian microRNA host genes and transcription units. Genome Res. 14, 1902–10. doi:10.1101/gr.2722704

Samoil, V., Dagenais, M., Ganapathy, V., Aldridge, J., Glebov, A., Jardim, A., Ribeiro, P., 2018. Vesicle-based secretion in schistosomes: Analysis of protein and microRNA (miRNA) content of exosome-like vesicles derived from Schistosoma mansoni. Sci. Rep. 8, 3286. doi:10.1038/s41598-018-21587-4

Santini, S., Boore, J.L., Meyer, A., 2003. Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. Genome Res. 13, 1111–22. doi:10.1101/gr.700503

Smyth, I., Du, X., Taylor, M.S., Justice, M.J., Beutler, B., Jackson, I.J., 2004. The extracellular matrix gene Frem1 is essential for the normal adhesion of the embryonic epidermis. Proc. Natl. Acad. Sci. U. S. A. 101, 13560–5. doi:10.1073/pnas.0402760101

Soares Magalhaes, R.J., Fancony, C., Gamboa, D., Langa, A.J., Sousa-Figueiredo, J.C., Clements, A.C.A., Vaz Nery, S., 2013. Extending helminth control beyond STH and

schistosomiasis: the case of human hymenolepiasis. PLoS Negl. Trop. Dis. 7, e2321. doi:10.1371/journal.pntd.0002321

Stegmayer, G., Yones, C., Kamenetzky, L., Milone, D.H., 2017. High Class-Imbalance in pre-miRNA Prediction: A Novel Approach Based on deepSOM. IEEE/ACM Trans. Comput. Biol. Bioinforma. 14, 1316–1326. doi:10.1109/TCBB.2016.2576459

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. Mol. Biol. Evol. 30, 2369–82. doi:10.1093/molbev/mst133

Tehler, D., Høyland-Kroghsbo, N.M., Lund, A.H., 2011. The miR-10 microRNA precursor family. RNA Biol. 8, 728–734. doi:10.4161/rna.8.5.16324

Thompson, R.C.A., 2015. Neglected zoonotic helminths: Hymenolepis nana, Echinococcus canadensis and Ancylostoma ceylanicum. Clin. Microbiol. Infect. doi:10.1016/j.cmi.2015.01.004

Tritten, L., Burkman, E., Moorhead, A., Satti, M., Geary, J., Mackenzie, C., Geary, T., 2014. Detection of Circulating Parasite-Derived MicroRNAs in Filarial Infections. PLoS Negl. Trop. Dis. 8, e2971. doi:10.1371/journal.pntd.0002971

Tsai, I.J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sanchez-Flores, A., Brooks, K.L., Tracey, A., Bobes, R.J., Fragoso, G., Sciutto, E., Aslett, M., Beasley, H., Bennett, H.M., Cai, J., Camicia, F., Clark, R., Cucher, M., De Silva, N., Day, T.A., Deplazes, P., Estrada, K., Fernández, C., Holland, P.W.H., Hou, J., Hu, S., Huckvale, T., Hung, S.S., Kamenetzky, L., Keane, J.A., Kiss, F., Koziol, U., Lambert, O., Liu, K., Luo, X., Luo, Y., Macchiaroli, N., Nichol, S., Paps, J., Parkinson, J., Pouchkina-Stantcheva, N., Riddiford, N., Rosenzvit, M., Salinas, G., Wasmuth, J.D., Zamanian, M., Zheng, Y., Cai, X., Soberón, X., Olson, P.D., Laclette, J.P., Brehm, K., Berriman, M., 2013. The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496, 57–63. doi:10.1038/nature12031

Venstrom, K.A., Reichardt, L.F., 1993. Extracellular matrix. 2: Role of extracellular matrix molecules and their receptors in the nervous system. FASEB J. 7, 996–1003.

VOGE, M., 1964. DEVELOPMENT OF HYMENOLEPIS MICROSTOMA (CESTODA: CYCLOPHYLLIDEA) IN THE INTERMEDIATE HOST TRIBOLIUM CONFUSUM. J. Parasitol. 50, 77–80.

Volinia, S., Calin, G.A., Liu, C.-G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R.L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C.C., Croce, C.M., 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc. Natl. Acad. Sci. 103, 2257–2261. doi:10.1073/pnas.0510565103

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. Evol. Dev. 11, 50–68. doi:10.1111/j.1525-142X.2008.00302.x

Winter, A.D., Weir, W., Hunt, M., Berriman, M., Gilleard, J.S., Devaney, E., Britton, C., 2012.

Diversity in parasitic nematode genomes: the microRNAs of Brugia pahangi and Haemonchus contortus are largely novel. BMC Genomics 13, 4. doi:10.1186/1471-2164-13-4

Yones, C., Stegmayer, G., Milone, D.H., Sahinalp, C., 2018. Genome-wide pre-miRNA discovery from few labeled examples. Bioinformatics 34, 541–549. doi:10.1093/bioinformatics/btx612

Yones, C.A., Stegmayer, G., Kamenetzky, L., Milone, D.H., 2015. miRNAfe: A comprehensive tool for feature extraction in microRNA prediction. Biosystems. 138, 1–5. doi:10.1016/j.biosystems.2015.10.003

Yuva-Aydemir, Y., Simkin, A., Gascon, E., Gao, F.-B., 2011. MicroRNA-9: functional evolution of a conserved small regulatory RNA. RNA Biol. 8, 557–64. doi:10.4161/rna.8.4.16019

Zhang, X., Zabinsky, R., Teng, Y., Cui, M., Han, M., 2011. microRNAs play critical roles in the survival and recovery of Caenorhabditis elegans from starvation-induced L1 diapause. Proc. Natl. Acad. Sci. U. S. A. 108, 17997–8002. doi:10.1073/pnas.1105982108

Zheng, H., Law, P.-Y., Loh, H.H., 2012. Non-Coding RNAs Regulating Morphine Function: With Emphasis on the In vivo and In vitro Functions of miR-190. Front. Genet. 3, 113. doi:10.3389/fgene.2012.00113

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–15.

Figure legends

Fig. 1. The top five most expressed miRNAs identified in *Hymenolepis microstoma* larvae. The normalized reads per million of each mature miRNA were averaged among biological replicates. The average proportion of miRNA reads between samples is shown in the pie chart.

Fig. 2. Evolutionary origin of all expected *Hymenolepis* miRNA families and their conservation among selected phyla. Blocks from left to right represent the evolutionary origin of all expected miRNA families in *Hymenolepis*: eumetazoan, bilaterian, protostomian and lophotrochozoan. Shaded grey blocks represent presence in the corresponding phyla. Light grey blocks represent absence in the corresponding phyla.

878 **Fig. 3**. Secondary structures of miRNA clusters identified in *Hymenolepis microstoma* and
879 *Hymenolepis nana* predicted with RNAfold. hmi-mir-71/2c/2b (A), hmi-mir-277a/4989 (B),
880 hmi-mir-36c/36d/36e (C), hna-mir-71/2c/2b (D), hna-mir-277a/4989 (E), hna-mir-36c/36d/36e
881 (F). Precursor miRNA sequences are indicated and the minimun free energy (MFE) of each
882 cluster is shown.

883 **Fig. 4.** Alignment and phylogenetic analysis of the *Hymenolepis* miRNA cluster mir-4989/277a
884 with its orthologs across Plathyhelminthes: *Schmidtea mediterranea* (sme), *Schistosoma mansoni*
885 (sma), *Mesocestoides corti* (mco), *Taenia solim* (tsol), *Echinococcus multilocularis* (emu),
886 *Echinococcus canadensis* (eca), *Echinococcus granulosus* (egr). Alignment of precursor miRNA
887 sequences was performed using MUSCLE followed by RNAalifold. The conserved nucleotides
888 are highlighted in grey scale and the mature miRNA are indicated with a box. The level of
889 nucleotide identity is indicated with grey color below the alignment (A). Phylogenetic tree was
890 inferred using the Maximum Likelihood method based on the Tamura 3-parameter model in
891 MEGA7. The sequence alignment used as input is shown above. The percentage of trees in
892 which the sequences clustered together is shown next to the branches. The support for the node
893 was assessed using 2000 bootstrap replicates (B).

894 **Supplementary Figure legends**

895 **Supplementary Fig. S1**. Secondary structures of all precursor miRNA sequences identified in
896 *Hymenolepis microstoma* predicted with mfold.

897 **Supplementary Fig. S2**. Genome-wide discovery of *Hymenolepis nana* miRNAs based on the
898 combination of three methods, miRNA-SOM, deepSOM and miRNAss. The number of miRNA
899 candidates discovered by each method is shown.