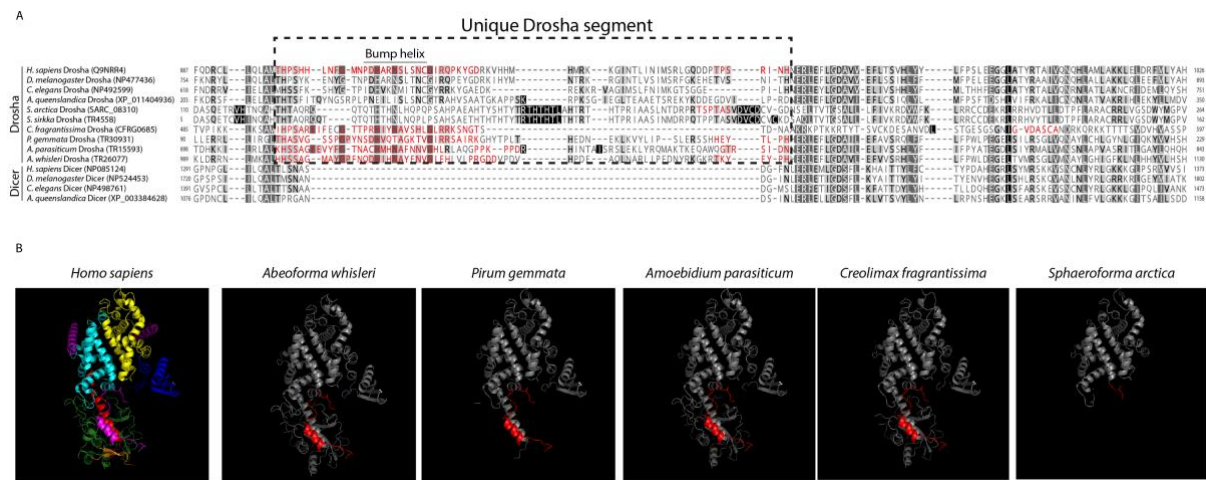


**Current Biology, Volume 28**

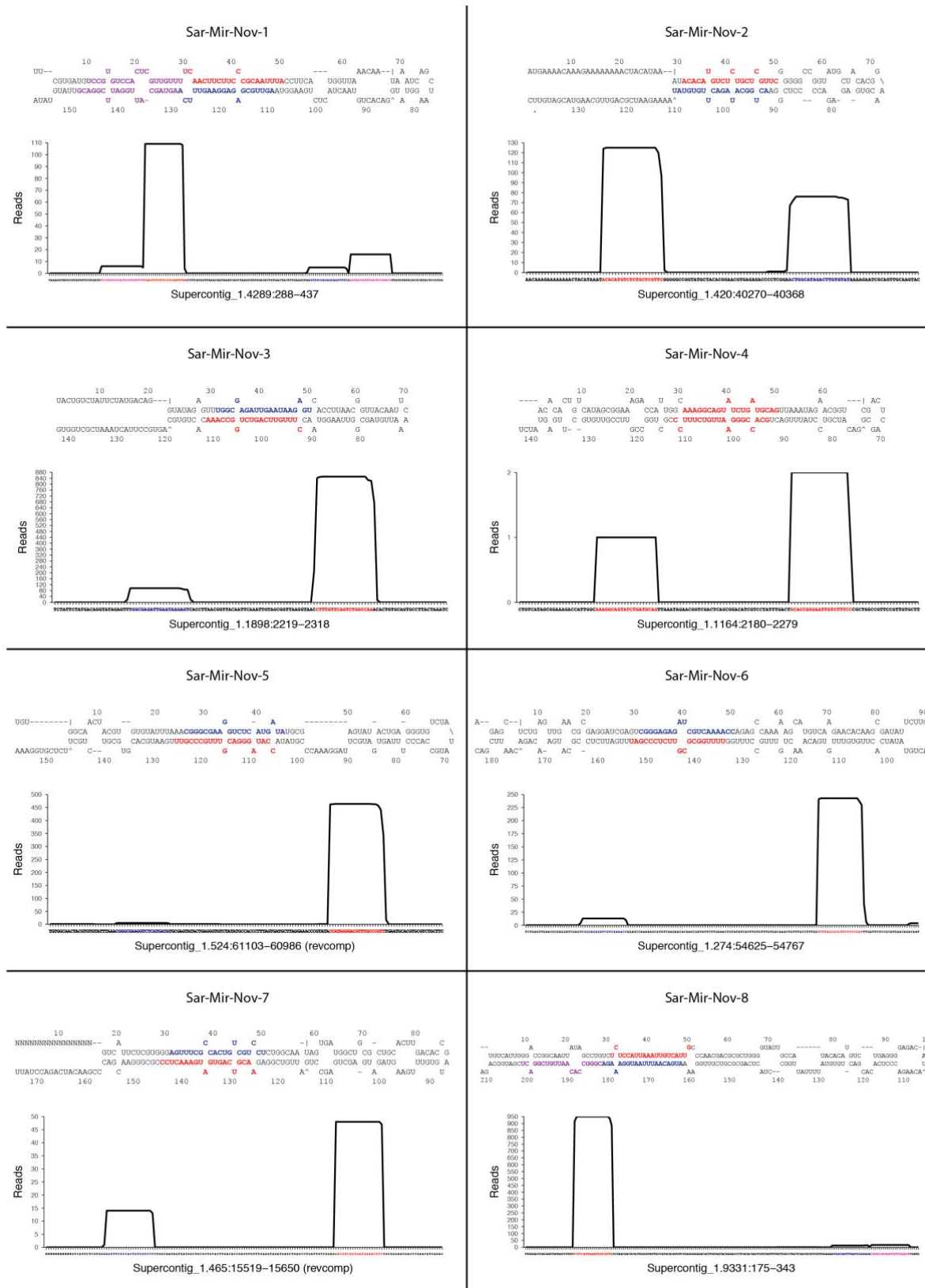
**Supplemental Information**

**Unicellular Origin  
of the Animal MicroRNA Machinery**

**Jon Bråte, Ralf S. Neumann, Bastian Fromm, Arthur A.B. Haraldsen, James E. Tarver, Hiroshi Suga, Philip C.J. Donoghue, Kevin J. Peterson, Iñaki Ruiz-Trillo, Paul E. Grini, and Kamran Shalchian-Tabrizi**

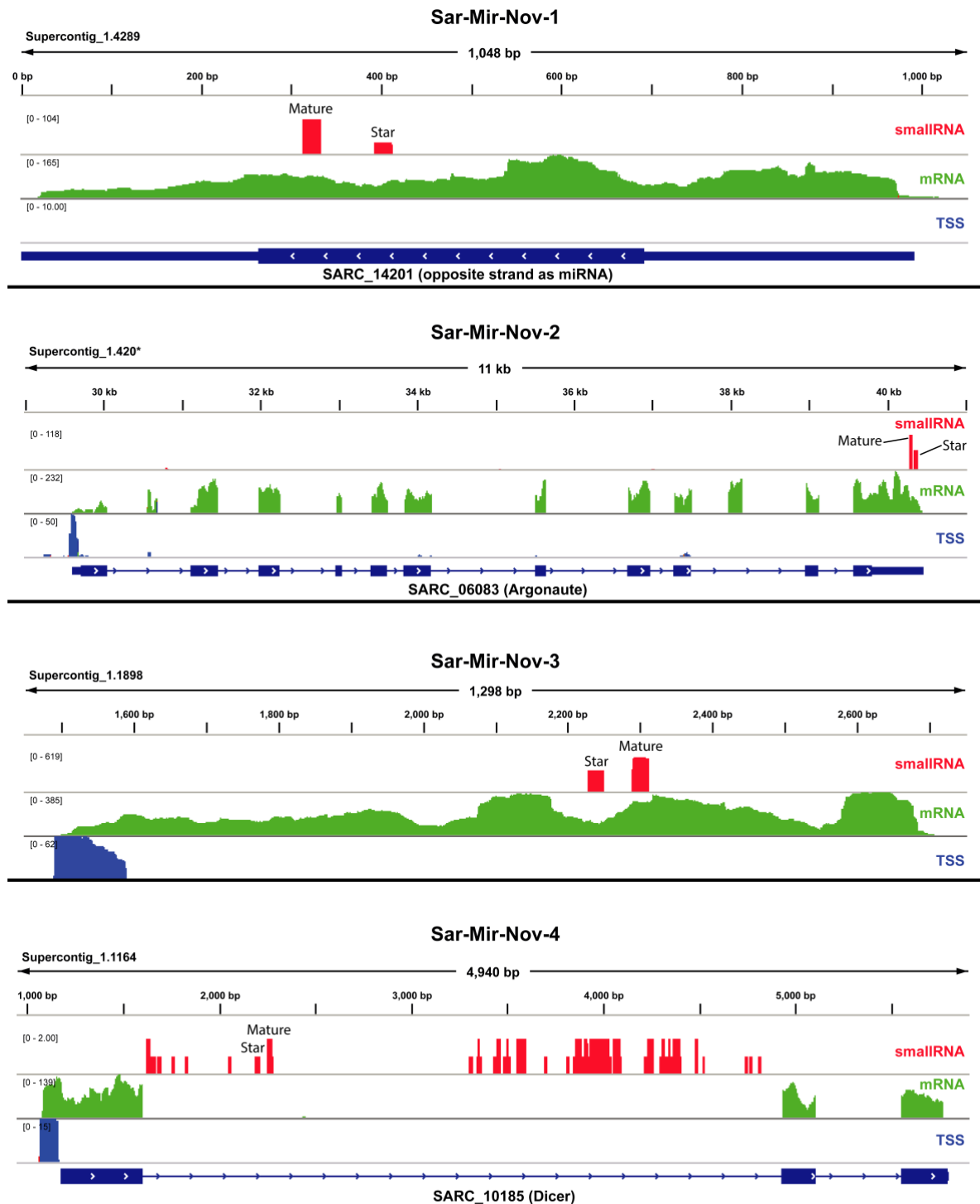


**Figure S1. Alignment and secondary structures of the unique Drosha insert, related to Figure 3.** A) Alignment of animal *Drosha* and *Dicer* sequences together with the novel *Drosha* sequences from ichthyosporans discovered in this study. The dashed box indicates the unique insertion found only in *Drosha* sequences. Amino acids highlighted in red corresponds to the amino acids highlighted in red in the secondary structures in (B). Not all amino acids of the insert are highlighted red in the alignment because they are missing from the reference *Drosha* structure. This is because not all residues were successfully modelled. B) Modeled secondary structures of the ichthyosporan *Drosha* sequences made by template based modeling to the reference *Drosha* structure from *Homo sapiens* (pdb entry: 5B16) which was identified as the most similar structure by Phyre2. Red amino acids correspond to the unique insert.



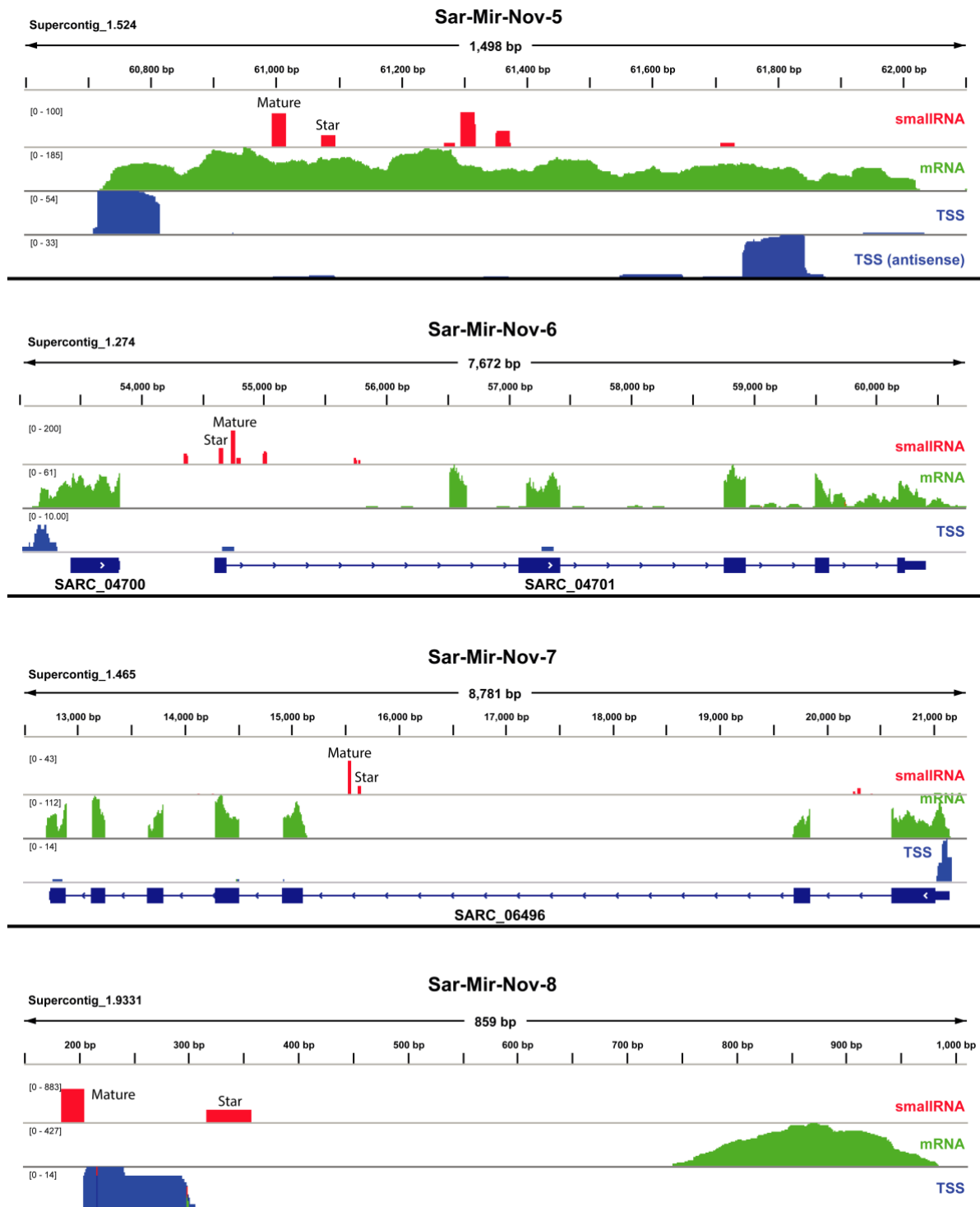
**Figure S2. Secondary structure and small RNA mapping for the 8 miRNAs detected in *Sphaeroforma arctica*, related to Figure 4. The mature strand is indicated in red font and the**

star strand is indicated in blue. For Sar-Mir-Nov-4 two mature strands (co-mature) are identified.



**Figure S3. Genomic localization and expression of *S. arctica* miRNA Sar-Mir-Nov1-4, related to Figure 4.** Mapping of small RNAs (red track), mRNAs (green track) and transcription start site (TSS) reads (blue tracks) at the genomic regions surrounding the miRNAs in *Sphaeroforma arctica*. \*For Supercontig\_1.420 we merged it with

Supercontig\_1.3651 which contain the middle domain of AGO. These two contigs are expressed together in the transcriptome as confirmed by mapping the mRNA reads to the genome. Note that Sar-Mir-Nov-2 is lacking a TSS signal but this is probably an artifact of the genome assembly (the contig is very short and the mRNAs map almost from beginning to the end).



**Figure S4. Genomic localization and expression of *S. arctica* miRNA Sar-Mir-Nov5-8, related to Figure 4. Mapping of small RNAs (red track), mRNAs (green track) and**

transcription start site (TSS) reads (blue tracks) at the genomic regions surrounding the miRNAs in *Sphaeroforma arctica*.



Type of sequencing	Nr. of reads
<i>Sphaeroforma arctica</i>	
small RNA (round 1)	2,515,574
small RNA (round 2)	10,128,170
TSS (tap +)	26,383,940
TSS (tap -)	17,898,227
mRNA	28,487,927
<i>Sphaeroforma sirkka</i>	
small RNA	6,464,007
mRNA	14,258,902
<i>Sphaeroforma napiecek</i>	
small RNA	6,569,325
mRNA	14,239,644
<i>Creolimax fragrantissima</i>	
small RNA (round 1)	4,000,000
small RNA (round 2)	6,197,622
mRNA	7,445,610
<i>Capsaspora owczarzaki</i>	
small RNA	1,781,653

**Table S1. Sequencing and mapping statistics, related to Figure 4.** Number of RNA-Seq reads before trimming and the mapping of reads to the genomes. For *S. arctica* and *C. fragrantissima* small RNAs were isolated and sequenced twice (round 1 and round 2).

	Argonaute	Exportin 5	Dicer	Drosha	Pasha
<b>Cnidaria</b>					
<i>Acropora digitifera</i>	4	1	3	1	1
<i>Nematostella vectensis</i>	3	1	2	1	1
<b>Placozoa</b>					
<i>Trichoplax adhaerens</i>	1	1	5	1	- <sup>a</sup>
<b>Porifera</b>					
<i>Amphimedon queenslandica</i>	2	1	2	1	1
<i>Sycon ciliatum</i>	2	1	2	1	1
<b>Ctenophora</b>					
<i>Mnemiopsis leidyi</i>	4	1	1	-	-
<i>Pleurobrachia bachei</i>	2	1	1	-	-
<b>Choanoflagellata</b>					
<i>Acanthoeca spectabilis</i>	-	-	-	-	-
<i>Acanthoeca sp.</i>	-	-	-	-	-
<i>Monosiga brevicollis</i>	-	1	-	-	-
<i>Salpingoeca pyxidium</i>	-	1	-	-	-
<i>Salpingoeca rosetta</i>	-	1	-	-	-
<b>Filasterea</b>					
<i>Capsaspora owczarzaki</i>	-	1	-	-	-
<i>Ministeria vibrans</i>	-	-	-	-	-
<b>Ichthyosporea</b>					
<i>Abeoforma whisleri</i>	2	1	2	1	1
<i>Amoebidium parasiticum</i>	1	1	2 <sup>b</sup>	1	1
<i>Creolimax fragrantissima</i>	1	1	1	1	1
<i>Ichthyophonus hoferi</i>	2	1	1 <sup>c</sup>	-	-
<i>Pirum gemmata</i>	2	1	2 <sup>d</sup>	1	1
<i>Sphaeroforma arctica</i>	2	1	2	1	1
<i>Sphaeroforma napiecek</i>	2	1	2	1 <sup>e</sup>	1
<i>Sphaeroforma sirkka</i>	2	1	2	1	1
<i>Sphaerothecum destruens</i>	-	1	-	-	-
<b>Corallochytrrea</b>					
<i>Corallochytrium limacisporum</i>	-	1	-	-	-
<b>Amoebozoa</b>					
<i>Dictyostelium discoideum</i>	0	1	2	-	-
<b>Nucleariidae</b>					
<i>Fonticula alba</i>	-	1	-	-	-
<i>Nuclearia sp.</i>	-	1	-	-	-
<b>Fungi</b>					
<i>Allomyces macrogynus</i>	8	1	1	-	-
<i>Mortierella verticillata</i>	6	1	4	-	-
<i>Rozella allomycis</i>	2	1	2	-	-
<i>Spizellomyces punctatus</i>	6	1	3	-	-

**Table S2. Presence of genes in the miRNA biogenesis pathway across Opisthokonta, related to Figure 1.** The presence (the number indicates how many copies of the gene) or absence (-) of key genes of the animal miRNA biogenesis pathway in selected species

covering all the major groups of opisthokonts. Shaded rows show results new to this study.

<sup>a</sup>*Pasha* was recently discovered in the Placozoan lineage *Trichoplax* sp. H2 [S1]. And we have also confirmed by Blasp search that the gene is also present in the newly sequenced Placozoan genome of *Hoilungia hongkongensis* [S2] (previously *Trichoplax* sp. H13).

<sup>b</sup>For *Amoebidium parasiticum* one full-length Dicer sequences was identified, and one fragment containing a single RNase III domain, but which gave Dicer as the top Blast hit. <sup>c</sup>For *Ichthyoponus hoferi*, only one fragment containing a single RNase III domain was identified. This gave Dicer as the nearest Blast hit. <sup>d</sup>For *Pirum gemmata* four sequences each containing a single RNase III domain were identified in the transcriptome assembly. Each sequence gave Dicer as the nearest Blast hit and showed homology to two a-domains and two b-domains. <sup>e</sup>In *Sphaeroforma napiiecek* we identified a gene with obvious homology to the other *Sphaeroforma* Droscha genes but which did not cover the RNase III domains, probably due to incomplete transcriptome assembly. This sequence also gave Droscha as the nearest Blast hit.

Species	Gene name	Accession number
<i>Homo sapiens</i>	Argonaute-1	NP_001304051
	Argonaute-2	NP_036286
	Argonaute-3	NP_079128
	Argonaute-4	NP_060099
	Exportin 5	NP_065801
	Dicer 1	NP_085124
	Ribonuclease 3	NP_037367
	DGCR8	NP_073557
	DGCR8	NP_001177255
	<i>Drosophila melanogaster</i>	Argonaute-1
Argonaute-2		ABB54719
Argonaute-3		ABO27430
RanBP21		AF222746
Dicer 1		NP_524453
Dicer 2		NP_523778
Drosha		NP_477436
Pasha		NP_651879
Pasha		NP_001263149
<i>Nematostella vectensis</i>		Argonaute-1
	Argonaute-2	AGW15595
	NEMVEDRAFT_v1g144913	XP_001621438
	Dicer 1	AGW15597
	Dicer 2	AGW15596
	Drosha	AGW15598
	Pasha/DGCR8	AGW15599
	piwi-like protein 1	XP_011409849
<i>Amphimedon queenslandica</i>	argonaute-2-like	XP_003385988
	exportin-5-like	XP_011402610
	Dicer-like	XP_003384628
	Dicer-like	XP_011402849
	Dicer-like	XP_003391904
	Dicer-like	XP_003384628
	Ribonucelase-3-like	XP_011404936
	sms-2	XP_958586
<i>Neurospora crassa</i>	QDE2	AAF43641
	NCU02387	XP_959707
	Dicer-like protein 1	XP_961898
	Dicer-like-2	XP_963538

**Table S3. Genes used as queries for the detection of miRNA processing genes, related to Figure 1.**

## Supplemental References

- S1. Kamm, K., Osigus, H.-J., Stadler, P.F., DeSalle, R., and Schierwater, B. (2018). Trichoplax genomes reveal profound admixture and suggest stable wild populations without bisexual reproduction. *Sci. Rep.* 8, 11168.
- S2. Eitel, M., Francis, W.R., Varoqueaux, F., Daraspe, J., Osigus, H.-J., Krebs, S., Vargas, S., Blum, H., Williams, G.A., Schierwater, B., *et al.* (2018). Comparative genomics and the nature of placozoan species. *PLOS Biol.* 16, e2005359.