# A draft genome sequence of the miniature parasitoid wasp, *Megaphragma amalphitanum*

Artem V. Nedoluzhko[1,2*], Fedor S. Sharko[3], Brandon M. Lê[4], Svetlana V. Tsygankova[2], Eugenia S. Boulygina[2], Sergey M. Rastorguev[2], Alexey S. Sokolov[3], Fernando Rodriguez[4], Alexander M. Mazur[3], Alexey A. Polilov[5], Richard Benton[6], Michael B. Evgen'ev[7], Irina R. Arkhipova[4], Egor B. Prokhortchouk[3,5*], Konstantin G. Skryabin[2,3,5]

[*]Correspondence: nedoluzhko@gmail.com, prokhortchouk@gmail.com

[1]Nord University, Faculty of Biosciences and Aquaculture, Bodø, 8049, Norway.

[2]National Research Center "Kurchatov Institute", Moscow, 123182, Russia

[3]Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, 117312, Russia

[4]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, Massachusetts 02543

[5]Lomonosov Moscow State University, Faculty of Biology, Moscow, 119234, Russia

[6]Center for Integrative Genomics, Faculty of Biology and Medicine, Génopode Building, University of Lausanne, CH-1015 Lausanne, Switzerland

[7]Institute of Molecular Biology RAS, Moscow, 119991, Russia

Full correspondence address: Dr. Artem V. Nedoluzhko, Nord University, Faculty of Biosciences and Aquaculture, Universitetsalléen 11, 8049, Bodø, Norway.

Phone: +79166705595

**ABSTRACT**

Body size reduction, also known as miniaturization, is an important evolutionary process that affects a number of physiological and phenotypic traits and helps animals to conquer new ecological niches. However, this process is poorly understood at the molecular level. Here, we report genomic and transcriptomic features of arguably the smallest known insect – the parasitoid wasp, *Megaphragma amalphitanum* (Hymenoptera: Trichogrammatidae). In contrast to expectations, we find that the genome and transcriptome sizes of this parasitoid wasp are comparable to other members of the Chalcidoidea superfamily. Moreover, the gene content of *M. amalphitanum* compared to other chalcid wasps is remarkably conserved. Among the very rare cases of apparent gene loss is *centrosomin*, which encodes an important centrosome component; the absence of this protein might be related to the large number of anucleate neurons in *M. amalphitanum.* Intriguingly, we also observed significant changes in *M. amalphitanum* transposable element dynamics over time, whereby an initial burst was followed by suppression of activity, possibly due to a recent reinforcement of the genome defense machinery. Thus, while the *M. amalphitanum* genomic data reveal certain features that may be linked to the unusual biological properties of this organism, miniaturization is not associated with a large decrease in genome complexity.

## INTRODUCTION

Miniaturization in animals is an evolutionary process that is frequently accompanied by structural simplification and size reduction of organs, tissues and cells [1, 2]. The parasitoid wasp *Megaphragma amalphitanum* (Hymenoptera: Trichogrammatidae, subfamily Oligositinae) is one of the smallest known insects, whose size (250 μm adult length) is comparable with unicellular eukaryotes and even some bacteria (**Figure 1**). Parasitoids from the genus *Megaphragma* parasitize greenhouse thrips *Heliothrips haemorrhoidalis* (Thysanoptera: Thripidae) developing on the shrubs *Viburnum tinus* (Adoxaceae) and *Myrtus communis* (Myrtaceae) [3], and possibly *Hercinothrips femoralis* (Thysanoptera: Thripidae) [4]. The wasp spends most of its life cycle in host eggs, while the imago stage is very short and lasts only a few days [3, 4]. *M. amalphitanum* belongs to chalcid wasps, which represent one of the largest insect superfamilies (~23,000 described species)[5]. The higher-level taxonomic relationships of Trichogrammatidae, Chalcidoidea and Hymenoptera have been investigated in several recent studies [6-10] that helped to establish the placement of this unique taxon that related to Mymaridae and Pteromalidae.

**Figure 1.** Size comparison of the parasitoid wasp *M. amalphitanum* and bacterium *Thiomargarita namibiensis*. (A) An adult stage of the parasitoid wasp *M. amalphitanum* (image adapted from [5]), (B) *T. namibiensis* – the largest known bacterium (modified from Schulz et al. 1999) [11].

Amongst notable anatomical features of *M. amalphitanum*, this species has only ~4600 neurons in its brain, which is substantially fewer than in the brains of other

wasps, e.g. the parasitoid chalcid wasp *Trichogramma pretiosum* (Trichogrammatidae: Trichogrammatinae) (~18,000 neurons), *Hemiptarsenus sp*. (Chalcidoidea: Eulophidae) (~35,000 neurons), and the honey bee *Apis mellifera* (Apidae) (~850,000-1,200,000 neurons). Moreover, by the final stage of *M. amalphitanum* development, up to 95 percent of the neurons of the central nervous system lose their nuclei [12, 13]. Nevertheless, adult wasps, which have an average lifespan of 5 days, still preserve the basic functional traits of hymenopteran insects including flight, mating and oviposition in hosts [14].

In this study, we present a draft *M. amalphitanum* genome, and adult transcriptome, and compare these with several parasitoid wasp species of different body sizes from the Chalcidoidea and Ichneumonoidea hymenopteran superfamilies. We performed a comprehensive analysis of its coding potential, including both general gene ontology and pathway analyses as well as specific gene categories of interest, such as chemosensory receptors, venom components etc. Additionally, we investigated transposable element (TE) content and dynamics across several parasitoid wasp species and analyzed the major components of the genome defense machinery. As body size reduction and loss of physiological or phenotypic traits is often correlated with genome size diminution [15, 16] and/or gene networks reduction [17], including chromatin diminution from the somatic tissues during embryogenesis[18, 19], we initially anticipated that the *M. amalphitanum* genome would be greatly simplified during miniaturization.

## MATERIAL AND METHODS

Detailed information is presented in ***Supplementary Information***

*Nucleic acid extraction and library construction. M. amalphitanum* individuals were reared in the laboratory conditions from eggs of *Heliothrips haemorrhoidalis* (Thysanoptera: Thripidae) collected in Santa Margherita, Northern Italy. DNA was extracted from ten individuals (males and females) using NucleoSpin Tissue XS kit (Macherey-Nagel, Germany) for each DNA-library. Three DNA libraries (DNA-library1 – whole insects; DNA-library2 – thorax and abdomen; DNA-library3 – head) were constructed using Ovation Ultralow Systems V2 kit (NuGEN, USA). Limited amount of biological material and low quantity of starting material (1 – 3 ng) did not permit construction of mate-paired libraries. Genome was sequenced used Illumina HiSeq 1500 (Illumina, USA) with 150 bp paired-end reads. RNA was extracted from ten *M. amalphitanum* individuals (males and females) using the Trizol reagent (Thermo Fisher Scientific, USA) by a standard protocol, and cDNA libraries were constructed using Ovation RNA-Seq System V2 kit (NuGEN, USA) with poly(A) enrichment.

*Genome de novo assembly*. The output from Illumina sequencing of the genomic DNA library (source format *.fastq) was used for *de novo* genome assembly. To assemble the complete genome of *M. amalphitanum*, we used 102,188,833 paired-end reads. Genome assemblies have been constructed using different assembly algorithms, and their performance was compared to each other (**Figure S1**). Additionally, genomic DNA-libraries from thorax and abdomen (DNA-library2) of *M. amalphitanum* (SRR5982987) and from head (DNA-library3) of *M. amalphitanum* (SRR5982986) were prepared. Totally, 79,317,970 (paired-end sequencing: 2×100 bp) and 85,409,775 (single--end

sequencing: 50 bp) DNA reads were sequenced and were used for *M. amalphitanum* coverage increase and as additional evidence during the search for missing genes (**Table S1**).

*Transcriptome de novo assembly.* Illumina RNA sequencing generated a total of 59,790,973 paired-end reads. Transcriptome *de novo* assembly was conducted using the default k-mer size in the Trinity software package (v. 2.4.0) [20], which combines three assembly algorithms: Inchworm, Chrysalis and Butterfly. Annotation of the *M. amalphitanum* transcriptome assembly was performed using the Trinotate pipeline (https://trinotate.github.io/).

*Transposable element (TE) de novo identification and analysis.* For *de novo* TE library construction, we used the REPET package [21] which combines three mutually complementing repeat identification tools (RECON, GROUPER and PILER), yielding a combined repeat library with the average consensus sequence length of 1.66 kb (ranging from 157 to 14,640 bp). The outputs were subject to additional classification with the RepeatClassifier tool from the RepeatMasker package (www.repeatmasker.org), which was also used to build the corresponding TE landscape divergence plots.

## RESULTS AND DISCUSSION

### Whole genome and transcriptome sequencing and assembly of *M. amalphitanum*

To gain insight into the genomic signatures of miniaturization that would distinguish *M. amalphitanum* from other Hymenoptera, we performed whole-genome shotgun sequencing of DNA (DNA-library1) isolated from ten adult individuals (males and females), using the Illumina platform **(Table S1).** The resulting genome assembly (PRJNA344956) has a cumulative length of 346 megabases (Mb), with a scaffold N50 of 10,296 bp. The total genome coverage is 88.6-fold. Thus, the genome of *M. amalphitanum* is comparable in size with other Chalcidoidea wasps, such as *Copidosoma floridanum*, *T. pretiosum* or *Nasonia vitripennis* [22, 23]. The best-performing combination of assembly software yielded contig N50 of 4,285 bp and allowed us to assemble 94,687 scaffolds from the quite low amounts of starting DNA material (**Table S2-S3; Figure S1**).

The *M. amalphitanum* genome assemblies were evaluated with the BUSCO (benchmarking universal single-copy orthologs) arthropod gene set [24], which uses 2,675 near-universal single-copy orthologs to assess the relative completeness of genome assemblies. Through this analysis, 7.55% of the conserved genes were initially identified in the *M. amalphitanum* assembly as missing (**Table S4**). More detailed information on our extensive search for the missing genes in *M. amalphitanum* genome is presented below.

We also performed whole-body transcriptome analysis using RNA extracted from ten *M. amalphitanum* individuals (males and females). Transcriptome *de novo* assembly (PRJNA344956) was performed using the Trinity software [20]. A total of

46,841 contigs were assembled with a mean length of 586 bp and an N50 of 633 bp from the quite low amounts of starting RNA material (**Table S5)**. The Illumina paired-end RNA-Seq data from *M. amalphitanum* were mapped to the previously assembled genome using bowtie2 [25]. Inspection of the alignments revealed that 79.95% of reads could be mapped to the genome. The BUSCO statistics for the transcriptome assembly is also presented in **Table S4.**

**Gene ontology analysis**

We used Gene Ontology (GO) analysis terms to describe characteristics of *M. amalphitanum* gene products in three independent categories: biological processes (**Figure S2**), molecular function (**Figure S3**), and cellular components (**Figure S4**). BLASTX outputs were used to retrieve the associated gene names and GO terms in all three categories (**Table 1**).

**Table 1.** Basic Gene Ontology (GO) analysis terms for *M. amalphitanum* gene products

| GO assignments of the transcripts | Transcript counts and percentage from total | | |
|---|---|---|---|
| Biological processes | 8,812 counts, 49.72% | | |
| | Transcription | Regulation of transcription | DNA integration |
| | 15 % | 10 % | 8 % |
| Cellular components | 4,802 counts, 27.10% | | |
| | Nucleus and cytoplasm components | Integral membrane components | Plasma membrane components |
| | 18 % | 9 % | 7% |
| Molecular functions | 4,108 counts, 23.18% | | |
| | ATP binding | Metal ion binding | Zinc ion binding |
| | 17 % | 12 % | 10 % |

All *M. amalphitanum* transcripts were matched to the Clusters of Orthologous Groups (COG) database to predict and classify their functions. In total, 8,810 genes were assigned to 25 COG functional categories. One of the largest groups is represented by the cluster for post-translational modification, protein turnover, and chaperones (988 counts; 10.7%), followed by intracellular trafficking, secretion, and vesicular transport (659 counts; 7.2%), DNA replication, recombination and repair (606 counts; 6.6%),

signal transduction mechanisms (599 counts, 6.5%) and transcription (587; 6.4%) (**Figure S5**).

To better understand incorporation of genes into diverse pathways, all annotated transcripts were mapped against the KEGG database for pathway-based analysis. As a result, 6,130 transcripts out of a total of 46,841 were assigned to a KEGG pathway, and were present in 328 different KEGG pathways. The KEGG pathway distribution is summarized in **Figure S6.** The top five pathways are metabolism (479 counts; 7.8%), biosynthesis of secondary metabolites (150 counts; 2.4%), RNA transport (100 counts; 1.6%), biosynthesis of antibiotics (95 counts; 1.5%), spliceosome (94 counts; 1.5%).

The annotation of *M. amalphitanum* and the available transcriptome assemblies of other parasitoid wasps from the families Trichogrammatidae (*T. pretiosum*, a lepidopteran egg parasitoid) and Braconidae including *Cotesia vestalis* (a diamondback moth parasitoid), *Diachasma alloeum* (an apple maggot parasitoid) and *Fopius arisanus* (tephritid fruit fly parasitoid) were used for comparative analysis of the most represented gene functions in parasitoids. We also used transcriptome assemblies from the Agaonidae fig wasp, *Ceratosolen solmsi*. We found significant similarities between *M. amalphitanum*, *T. pretiosum* and *C. vestalis* major GO enrichment categories (**Figure S7-S9**). At the same time, a significant number of transcripts related to DNA integration relative to other parasitoid wasps was found in *D. alloeum* and *M. amalphitanum* (**Figure S7**) (see below). Complete information about reference datasets used for *M. amalphitanum* genome and transcriptome data analysis is shown in **Table S6.** The Trinotate statistics for annotation of *M. amalphitanum*, *C. solmsi*, *D. alloeum*, *F. arisanus*, *C. vestalis* and *T. pretiosum* transcriptome assemblies is presented in **Table S7.**

**Missing genes and missing or rapidly evolving gene clusters in the *M. amalphitanum* genome**

We clustered gene orthologs and identified gene clusters for each hymenopteran taxa (Chalcidoidea: *M. amalphitanum*, *T. pretiosum*, *C. solmsi*, C. *floridanum*, and *N. vitripennis*; Ichneumonoidea: *D. alloeum* and *F. arisanus*; Apoidea: *A. mellifera*) using OrthoMCL [26]. The core gene set of all the hymenopteran species was composed of 6278 gene clusters, 122 gene clusters were unique to the chalcid clade. 262 gene clusters were not detected in any of the chalcids analysed (**Supplementary Dataset 3**; NCBI BioProject: PRJNA344956), but found in all the other hymenoptera, consistent with a similar recent analysis [27]. Our findings suggest that that these gene losses occurred in the last common ancestor of chalcids or point to the possibility of parallel genome evolution across these species. Interestingly, the missing/rapidly evolving genes include homologs of genes that have important roles in embryonic patterning and development in other insects (e.g., *krueppel-1*, *knirps* or *short gastrulation* [27]).

To determine whether miniaturization in *M. amalphitanum* is associated with gene loss, genomic data of six larger hymenopteran species (*T. pretiosum*, *C. vestalis*, *C. floridanum*, *F. arisanus*, *N. vitripennis*, and *N. giraulti*) – as well as the well-annotated genome of the honeybee (*A. mellifera*) as reference – were used (body sizes are presented in **Table S6)**. We mapped the *M. amalphitanum* (DNA-library1), *T. pretiosum*, *C. vestalis*, *C. floridanum*, *F. arisanus*, *N. vitripennis*, *N. giraulti* DNA reads on the *A. mellifera* genome sequence (PRJNA13343, PRJNA10625) (**Figure S10-S11**), and detected 115 genes that were not represented by *M. amalphitanum* sequencing reads but were present in other parasitoid wasps. We next increased the coverage of the *M. amalphitanum* genome to 146.8-fold by adding the reads from additional libraries

(DNA-library2 and DNA-library3) (**Table S1**) and observed the apparent absence of 114 of the 115 genes. An additional TBLASTX search identified 36 of these genes as present, yielding a total of 78 missing genes (**Table S8**). However, querying the *M. amalphitanum* genome with the corresponding amino acid sequences from the closest wasp ortholog (*N. vitripennis* or *T. pretiosum*) in TBLASTN searches reduced the number of missing genes to just five: centrosomin, phosphoglycerate mutase 5, phosphoglycerate mutase 5-2, 26S proteasome complex subunit DSS1, and mucin-1/nucleoporin NSP1-like. We detected short *M. amalphitanum* genome sequences encoding protein fragments (~8-23 amino acid residues) with some similarity to four of them, suggesting that they may be in the process of degeneration in this species. Despite a careful search, we were unable to find any homologous sequence related to *centrosomin (cnn)* gene either in the assembled genome or our cDNA libraries (**Figure S15**). Although *cnn* is regarded as rapidly evolving [28], sequence homology can be readily discerned and orthologs are present in every other insect, including the parasitoid *T. pretiosum*, suggesting that this is a *bona fide* absent gene. Globally, however, these analyses indicate that there has been little gene loss in *M. amalphitanum.*

**Chemosensory genes in the *M. amalphitanum* genome**

Chemosensory receptors are encoded by some of the largest gene families in insect genomes, reflecting their important and wide-ranging roles in detection of environmental odors and tastants. We asked how these gene families have evolved in *M. amalphitanum*, whose central and peripheral nervous systems are highly reduced [2, 14]. The highly divergent sequences of chemosensory receptors and relatively short genomic contig lengths available for *M. amalphitanum* precluded accurate annotation of

full-length sequences in this species for the majority of loci. Nevertheless, comparison with chemosensory receptor repertoires of other insects allowed us to define probable orthologous relationships with receptors of known function in other species and obtain initial estimates of the size of each family.

The most deeply conserved family of chemosensory receptors in insects are the Ionotropic Receptors (IRs), which are distantly related to ionotropic glutamate receptors [29, 30]. IRs function in heteromeric protein complexes comprising more broadly-expressed co-receptors with selectively expressed "tuning" IRs that determines sensory specificity. We identified orthologs of each of the co-receptors (*Ir8a*, *Ir25a* (two paralogs), *Ir93a* and *Ir76b*), as well as four genes encoding tuning IRs related to acid-sensing receptors in other species. We also identified orthologs of IR68a, which functions in hygrosensation [31] and IR21a, which functions in cool temperature-sensing [32, 33]. Overall, the repertoire of IRs in *M. amalphitanum* is therefore very similar in size and content to that of *N. vitripennis [30]*.

Insects possess a second superfamily of chemosensory ion channels – distinguished by a heptahelical protein structure – comprising Odorant Receptor (OR) and Gustatory Receptor (GR) subfamilies, which generally function in detection of volatile and non-volatile stimuli, respectively [34-37]. Similar to IRs, ORs function in heteromeric complexes of a conserved co-receptor (ORCO) and a tuning OR. We identified an *M. amalphitanum* ortholog of *Orco* and 83 additional *Or*-related sequences. We caution that many of these *Or* sequences are small fragments (often located near the end of the assembled contigs), so it is currently difficult to determine whether these are intact genes or pseudogenes. Within the GR repertoire, we identified genes encoding proteins related to GR43a, a sensor of both external and internal

fructose [38], two others similar to other insect sugar-sensing GRs [39], and 25 additional *Gr* gene fragments. The sizes of these repertoires are smaller than in *N. vitripennis* (300 *Or*s (including 76 pseudogenes) and 58 *Gr*s (including 11 pseudogenes) [40]), but similar to non-miniaturized parasitoid wasps *Meteorus pulchricornis* and *Macrocentrus cingulum* [41, 42]. However, precise comparison with the latter two species is difficult, as receptors in these wasps were identified from antennal transcriptomes, thereby representing only one of these insects' chemosensory organs.

In sum, these analyses reveal that despite drastic nervous system reduction, *M. amalphitanum* has retained the conserved chemosensory receptors of larger wasps (and other insects), and appears to have numerous additional order- or species-specific receptors to allow detection of environmental chemical cues.

**Venom components in the *M. amalphitanum* transcriptomic data**

Parasitoid wasps often use venom to modify the metabolism of their hosts; toxins and their known or presumed biological functions are described in various species [43]. We investigated the presence of homologs of *N. vitripennis* toxin constituents in *M. amalphitanum* and other parasitoid wasps (*Megastigmus spermotrophus*, *N. vitripennis*, *C. solmsi*, *T. pretiosum*), using previously published venom data [44, 45] and the transcriptomes of chalcid wasps (**Table S6).** We identified 28 transcripts encoding putative venom proteins (**Figure 2; Table S9**); homologs of these are found in all investigated Chalcidoidea species (**Table 2**). Assuming that most of these candidates are truly conserved venom proteins among Chalcidoids, *M. amalphitanum* venom's diversity does not seem to have been significantly affected by size reduction.

298 **Figure 2.** A Venn diagram showing *Nasonia vitripennis* venom components in other
299 Chalcidoidea species: *M. spermotrophus, C. solmsi, T. pretiosum,* and *M.*
300 *amalphitanum.*

301 **Table 2**. Number of homologs of *N. vitripennis* venom (*N. vitripennis* toxin
302 constituents) in *M. amalphitanum* and other Chalcidoidea species based on Universal
303 Chalcidoidea Database (http://www.nhm.ac.uk/our-science/data/chalcidoids/database/)
304

| Parasitoid wasp species | Families of Chalcidoidea | Number of *N. vitripennis* venom constituents | Body size, mm | Approximate number of hosts |
|---|---|---|---|---|
| *M. amalphitanum* | Trichogrammatidae | 37 | 0.25 | 2 insect species from one order |
| *C. solmsi* | Agaonidae | 38 | 2.7 | 2 plant species from one family |
| *M. spermotrophus* | Torymidae | 41 | 2.8 | 13 plant species from one family |
| *T. pretiosum* | Trichogrammatidae | 45 | 0.5 | >140 insect species from 4 orders |
| *N. vitripennis* | Pteromalidae | 64 | 2.2 | >110 insect species from 8 orders |

305 ***M. amalphitanum* transposable elements and genome defense**

306 Transposable elements (TEs) constitute a measurable fraction of virtually all
307 eukaryotic genomes, and can play important roles in their function and evolution. In
308 insects, TE activity has been implicated in evolution of eusociality, based on
309 comparison of ten bee genomes with increasing degrees of social complexity [46]. We
310 performed *de novo* TE identification and comparative analysis of TE dynamics in *M.*
311 *amalphitanum* and in a representative set of larger wasp genomes for which TE content
312 has previously been reported: the parasitoid *N. vitripennis* and two primitively eusocial
313 aculeate wasps *Polistes canadensis* and *Polistes dominula* [12, 23, 47]. Additionally, we
314 analyzed TEs in the genomes of parasitoid wasps *T. pretiosum* from the family
315 Trichogrammatidae and *D. alloeum* from the family Braconidae, in which TE content
316 has not been studied.

317 For uniformity of measurements, we applied the same workflow to all genomes,
318 without relying on pre-existing repeat libraries. We employed the REPET package for
319 *de novo* TE identification (also used in [46]), and RepeatMasker for repeat classification
320 and construction of TE landscape divergence plots. Comparison of the overall repeat
321 content across six wasp species did not reveal substantial differences between four
322 species (18.5% in *M. amalphitanum vs.* 18.1%, 17.7% and 14.2% in *P. canadensis*, *P.*
323 *dominula* and *T. pretiosum*, respectively), while the *N. vitripennis* genome was found to
324 be 32.5% repetitive, in close agreement with the published estimate [23], and *D.*
325 *alloeum* was highly repetitive at 52.8% (pie charts in **Figure 3; Figure S12**).
326 Surprisingly, TE dynamics over time, which is shown on the corresponding TE
327 landscape divergence plots, was found to differ substantially for *M. amalphitanum*,

328 which displayed a pronounced decline in recent TE activity after an initial increase, a
329 pattern that was not observed in any other wasp (**Figure 3**).
330
331 **Figure 3.** Comparison of TE landscape divergence plots and TE genome fraction pie
332 charts in four parasitoid wasp species: *M. amalphitanum*, *T. pretiosum*, *N. vitripennis*,
333 and *D. alloeum*.
334
335 While TE dynamics may be affected by different factors, the observed drop in
336 active TE content in *M. amalphitanum* may be relevant to the unique biology of this
337 highly miniaturized insect. Its closest relative, *T. pretiosum*, is about 2-fold larger in
338 body length. The spike in recent TE activity in *T. pretiosum* may have been caused by
339 *Wolbachia* infection, which typically results in abandonment of sexual reproduction
340 [48] and the concomitant TE proliferation in a non-recombining genome [49]. Other
341 wasps do not display notable drops or spikes in current TE activity; TE inactivation was
342 reported in two asexual mites [50], however it appears to be ancient and may have
343 occurred prior to the abandonment of sex. Overall, the continued decline in *M.*
344 *amalphitanum* TE activity over the span of several million years – not observed in *T.*
345 *pretiosum* which shares the most recent common ancestor with *M. amalphitanum* –
346 represents a highly unusual genomic feature in every hymenopteran we examined,
347 including ants (not shown). Interestingly, no traces of *Wolbachia* infection or other
348 representatives of the Rickettsiaceae family were found in *M. amalphitanum* individuals
349 [51], while the sequenced *T. pretiosum* carries the *Wolbachia* symbiont [52]; the
350 sequenced *Nasonia* strain was maintained on antibiotics to cure it of infection.

To gain insights into possible reasons for reduction in TE activity after the initial burst, we investigated the major components of the genome defense machinery in *M. amalphitanum*, including Dicer (Dcr)-like and Argonaute (Ago)/Piwi-like protein-coding genes. In insects, *Ago-1* and *Dcr-1* homologs represent the key components of the miRNA pathway; *Ago-2* and *Dcr-2* mediate antiviral RNA interference; and *Piwi* and *Ago-3*/*Aub* suppress TE activity in the germline [53]. Both *M. amalphitanum* and *T. pretiosum* possess equal numbers of *Dcr-1* and *Dcr-2* homologs, as well as *Ago-2* and *Ago-3* homologs (**Figure S13**). However, in *M. amalphitanum*, the *Ago-1* and the *Piwi/Aub* homologs underwent a relatively recent duplication in comparison to *T. pretiosum* (**Figure 4**). This may indicate additional layers of enforcement in the miRNA and piRNA pathways of *M. amalphitanum*, both of which should result in suppression of TE activity.

**Figure 4.** Maximum likelihood analysis of phylogenetic relationships between Piwi/Argonaute coding sequences. Percent identity is indicated for the duplicated Ago-1 and Piwi/Aub homologs, both at the DNA and protein (cds) level. Phylogeny analysis and notations are as in Supplementary Figure S13.

The drop in TE activity is also evident from the transcriptome analysis. The GO radar plot (**Figure S7**) shows a substantial number of short contigs related to DNA integration, most of which upon inspection were found to represent separate fragments of gypsy-like and copia-like LTR retrotransposons, and a few belong to Polinton, P and Ginger DNA TEs. Transcriptionally active copies fall into two groups: first, those which apparently proliferated during the burst of TE activity and have since

accumulated debilitating mutations making them incapable of transposition, but still retain a certain level of transcriptional activity; second, those that originate from recent infections by retrovirus-like TEs and contain uninterrupted ORFs, but are not actively proliferating and are present at very few genomic loci. Box plot of per cent identity between BLASTN hits for *M. amalphitanum* integrase-related TE transcripts showed that high-copy hits represent MITEs (**Figure S14**). We hypothesize that actively proliferating TE copies can represent recent arrivals, possibly brought about by viruses or host-parasite interactions [54].

Our study provides a first view of the genomic content of one of the smallest insects currently known, the parasitoid wasp *M. amalphitanum*. In contrast to the expectation that the small body size, in combination with the parasitic lifestyle, should lead to significant reduction in the amount of genomic DNA and in gene content, we do not observe a drastic reduction in the overall genome size or in the number of expressed genes in comparison with larger parasitic wasps.

Of the rare cases of genes that were apparently lost from – or highly degenerated in – the *M. amalphitanum* genome, the most intriguing is *centrosomin (cnn),* which is universally present in insects, including *T. pretiosum.* In *Drosophila melanogaster*, Cnn has important roles at the centrosome in mitotic spindle formation, cytoskeleton organization and neuronal morphogenesis [55, 56], although these functions may not be indispensable because this species (and possibly other insects) possesses centrosome-independent mechanisms for spindle nucleation [57]. A fungal homolog of Cnn, called *anucleate primary sterigmata*, is involved in nuclear migration [58-60]. We speculate that the loss of this gene is linked to the unique biological feature of widespread neuron denucleation in this tiny parasitoid wasp.

399 Surprisingly, transposable element dynamics over time was found to differ
400 greatly between the analyzed species, with *M. amalphitanum* displaying a relatively
401 recent dramatic decline in TE activity preceded by a burst, a pattern not observed in
402 other parasitoid wasps. The decline in TE activity may have been associated with
403 evolution of additional *Ago* and *Piwi* copies, not present in *T. pretiosum*, which could
404 have reinforced the genome defense machinery to prevent uncontrolled TE expansion.
405 The relationship between body size and genome size has been discussed for a
406 long time. Significant correlations of these values have been described for flatworms
407 and copepods [16]; by contrast, such correlations were not found in ants [61]. Our
408 results show that body size reduction in hymenopterans, while being associated with
409 certain genomic adaptations, is not accompanied by greatly decreased transcriptomic
410 and genomic complexity. This observation begs the question of how miniaturization is
411 encoded genetically. We hypothesize that changes in regulatory sequences, rather than
412 gene content, were important in the process of body size reduction, similar to
413 mechanisms of morphological evolution that have driven adaptive diversification in all
414 animals, great or small [62].
415

**Acknowledgements**

This work has been carried out using computing resources of the federal collective usage center Complex for Simulation and Data Processing for Mega-science Facilities at NRC "Kurchatov Institute" (ministry subvention under agreement RFMEFI62117X0016), http://ckp.nrcki.ru/. Support of this project was provided by the Russian Scientific Foundation (RSF) grant #14-24-00175 and RSF grant #14-50-00060. B.L. was supported by the Rosenthal Brown-MBL internship and the REU supplement to NSF MCB-1121334 to I.A. R.B. was supported by a European Research Council Consolidator Grant (615094).

**Data availability**

All genome and transcriptome assemblies and all SRA sequence data are publicly available at the NCBI BioProject: PRJNA344956.

**Competing interests**

The authors declare that they have no competing interests.

433 **References**

434 1. Hanken J, Wake DB. Miniaturization of Body-Size - Organismal Consequences
435 and Evolutionary Significance. Annu Rev Ecol Syst. 1993;24:501-19. doi: DOI
436 10.1146/annurev.es.24.110193.002441. PubMed PMID: ISI:A1993MJ37100018.
437 2. Polilov AA. Small is beautiful: features of the smallest insects and limits to
438 miniaturization. Annu Rev Entomol. 2015;60:103-21. Epub 2014/10/24. doi:
439 10.1146/annurev-ento-010814-020924. PubMed PMID: 25341106.
440 3. Bernardo U, Viggiani, G. Biological data on Megaphragma amalphitanum
441 Viggiani and Megaphragma mymaripenne Timberlake (Hymenoptera:
442 Trichogrammatidae), egg-parasitoid of H. haemorrhoidalis Bouché) (Thysanoptera:
443 Thripidae) in southern Italy. Bollettino del Laboratorio di Entomologia Agraria Filippo
444 Silvestri. 2002;58:77-85.
445 4. Pintureau B, Lassabliere F, Khatchadourian C, Daumal J. Eggs parasitoids and
446 symbionts of two European Thrips. Annales de la Société Entomologique de France.
447 1999;35:416-20. PubMed PMID: ISI:000085892500074.
448 5. Polilov AA. Anatomy of adult Megaphragma (Hymenoptera:
449 Trichogrammatidae), one of the smallest insects, and new insight into insect
450 miniaturization. PLoS One. 2017;12(5):e0175566. Epub 2017/05/04. doi:
451 10.1371/journal.pone.0175566
452 PONE-D-16-42043 [pii]. PubMed PMID: 28467417; PubMed Central PMCID:
453 PMC5414980.
454 6. Branstetter MG, Danforth BN, Pitts JP, Faircloth BC, Ward PS, Buffington ML,
455 et al. Phylogenomic Insights into the Evolution of Stinging Wasps and the Origins of

456 Ants and Bees. Current Biology. 2017;27(7):1019-25. doi: 10.1016/j.cub.2017.03.027.
457 PubMed PMID: ISI:000398061700023.
458 7. Nedoluzhko AV, Sharko FS, Boulygina ES, Tsygankova SV, Sokolov AS,
459 Mazur AM, et al. Mitochondrial genome of Megaphragma amalphitanum
460 (Hymenoptera: Trichogrammatidae). Mitochondrial DNA A DNA Mapp Seq Anal.
461 2016;27(6):4526-7. Epub 2015/12/01. doi: 10.3109/19401736.2015.1101546. PubMed
462 PMID: 26617282.
463 8. Owen AK, George J, Pinto JD, Heraty JM. A molecular phylogeny of the
464 Trichogrammatidae (Hymenoptera : Chalcidoidea), with an evaluation of the utility of
465 their male genitalia for higher level classification. Syst Entomol. 2007;32(2):227-51.
466 doi: 10.1111/j.1365-3113.2006.00361.x. PubMed PMID: ISI:000245745800003.
467 9. Heraty JM, Burks RA, Cruaud A, Gibson GAP, Liljeblad J, Munro J, et al. A
468 phylogenetic analysis of the megadiverse Chalcidoidea (Hymenoptera). Cladistics.
469 2013;29(5):466-542. doi: 10.1111/cla.12006. PubMed PMID: ISI:000324622500003.
470 10. Peters RS, Niehuis O, Gunkel S, Blaser M, Mayer C, Podsiadlowski L, et al.
471 Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera:
472 Chalcidoidea) reveals a history of rapid radiations, convergence, and evolutionary
473 success. Mol Phylogenet Evol. 2018;120:286-96. Epub 2017/12/17. doi: S1055-
474 7903(17)30336-6 [pii]
475 10.1016/j.ympev.2017.12.005. PubMed PMID: 29247847.
476 11. Schulz HN, Brinkhoff T, Ferdelman TG, Marine MH, Teske A, Jorgensen BB.
477 Dense populations of a giant sulfur bacterium in Namibian shelf sediments. Science.
478 1999;284(5413):493-5. Epub 1999/04/16. PubMed PMID: 10205058.

12. Standage DS, Berens AJ, Glastad KM, Severin AJ, Brendel VP, Toth AL. Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. Mol Ecol. 2016;25(8):1769-84. Epub 2016/02/10. doi: 10.1111/mec.13578. PubMed PMID: 26859767.

13. Makarova AA, Polilov AA. Peculiarities of the Brain Structure and Ultrastructure in Small Insects Related to Miniaturization. 1. The Smallest Coleoptera (Ptiliidae). Zool Zh. 2013;92(5):523-33. doi: 10.7868/S0044513413050073. PubMed PMID: ISI:000321472300003.

14. Polilov AA. The smallest insects evolve anucleate neurons. Arthropod Struct Dev. 2012;41(1):29-34. Epub 2011/11/15. doi: 10.1016/j.asd.2011.09.001 S1467-8039(11)00094-6 [pii]. PubMed PMID: 22078364.

15. Gregory TR. Synergy between sequence and size in large-scale genomics. Nat Rev Genet. 2005;6(9):699-708. Epub 2005/09/10. doi: nrg1674 [pii] 10.1038/nrg1674. PubMed PMID: 16151375.

16. Gregory TR, Hebert PDN, Kolasa J. Evolutionary implications of the relationship between genome size and body size in flatworms and copepods. Heredity. 2000;84(2):201-8. doi: DOI 10.1046/j.1365-2540.2000.00661.x. PubMed PMID: ISI:000086196500009.

17. Erives AJ. Genes conserved in bilaterians but jointly lost with Myc during nematode evolution are enriched in cell proliferation and cell migration functions. Dev Genes Evol. 2015;225(5):259-73. Epub 2015/07/16. doi: 10.1007/s00427-015-0508-1 10.1007/s00427-015-0508-1 [pii]. PubMed PMID: 26173873; PubMed Central PMCID: PMC4568025.

503 18. Akif'ev AP, Grishanin AK, Degtiarev SV. [Chromatin diminution is a key
504 process explaining the eukaryotic genome size paradox and some mechanisms of
505 genetic isolation]. Genetika. 2002;38(5):595-606. Epub 2002/06/19. PubMed PMID:
506 12068542.
507 19. Wang J, Mitreva M, Berriman M, Thorne A, Magrini V, Koutsovoulos G, et al.
508 Silencing of germline-expressed genes by DNA elimination in somatic cells. Dev Cell.
509 2012;23(5):1072-80. Epub 2012/11/06. doi: 10.1016/j.devcel.2012.09.020
510 S1534-5807(12)00431-5 [pii]. PubMed PMID: 23123092; PubMed Central PMCID:
511 PMC3620533.
512 20. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al.
513 Full-length transcriptome assembly from RNA-Seq data without a reference genome.
514 Nature Biotechnology. 2011;29(7):644-U130. doi: 10.1038/nbt.1883. PubMed PMID:
515 ISI:000292595200023.
516 21. Flutre T, Duprat E, Feuillet C, Quesneville H. Considering transposable element
517 diversification in de novo annotation approaches. PLoS One. 2011;6(1):e16526. Epub
518 2011/02/10. doi: 10.1371/journal.pone.0016526. PubMed PMID: 21304975; PubMed
519 Central PMCID: PMC3031573.
520 22. The i5K Initiative: advancing arthropod genomics for knowledge, human health,
521 agriculture, and the environment. J Hered. 2013;104(5):595-600. Epub 2013/08/14. doi:
522 10.1093/jhered/est050
523 est050 [pii]. PubMed PMID: 23940263; PubMed Central PMCID: PMC4046820.
524 23. Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, Colbourne JK, et
525 al. Functional and evolutionary insights from the genomes of three parasitoid Nasonia

526 species. Science. 2010;327(5963):343-8. Epub 2010/01/16. doi:
527 10.1126/science.1178028
528 327/5963/343 [pii]. PubMed PMID: 20075255; PubMed Central PMCID:
529 PMC2849982.
530 24. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM.
531 BUSCO: assessing genome assembly and annotation completeness with single-copy
532 orthologs. Bioinformatics. 2015;31(19):3210-2. Epub 2015/06/11. doi:
533 10.1093/bioinformatics/btv351
534 btv351 [pii]. PubMed PMID: 26059717.
535 25. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat
536 Methods. 2012;9(4):357-9. Epub 2012/03/06. doi: 10.1038/nmeth.1923
537 nmeth.1923 [pii]. PubMed PMID: 22388286; PubMed Central PMCID: PMC3322381.
538 26. Li L, Stoeckert CJ, Jr., Roos DS. OrthoMCL: identification of ortholog groups
539 for eukaryotic genomes. Genome Res. 2003;13(9):2178-89. Epub 2003/09/04. doi:
540 10.1101/gr.1224503
541 13/9/2178 [pii]. PubMed PMID: 12952885; PubMed Central PMCID: PMC403725.
542 27. Lindsey ARI, Kelkar YD, Wu X, Sun D, Martinson EO, Yan Z, et al.
543 Comparative genomics of the miniature wasp and pest control agent Trichogramma
544 pretiosum. BMC Biol. 2018;16(1):54. Epub 2018/05/20. doi: 10.1186/s12915-018-
545 0520-9
546 10.1186/s12915-018-0520-9 [pii]. PubMed PMID: 29776407; PubMed Central PMCID:
547 PMC5960102.

548 28. Eisman RC, Kaufman TC. Probing the boundaries of orthology: the
549 unanticipated rapid evolution of Drosophila centrosomin. Genetics. 2013;194(4):903-
550 26. Epub 2013/06/12. doi: 10.1534/genetics.113.152546
551 genetics.113.152546 [pii]. PubMed PMID: 23749319; PubMed Central PMCID:
552 PMC3730919.
553 29. Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. Variant Ionotropic
554 Glutamate Receptors as Chemosensory Receptors in Drosophila. Cell. 2009;136(1):149-
555 62. doi: 10.1016/j.cell.2008.12.001. PubMed PMID: ISI:000262318400022.
556 30. Rytz R, Croset V, Benton R. Ionotropic Receptors (IRs): Chemosensory
557 ionotropic glutamate receptors in Drosophila and beyond. Insect Biochem Molec.
558 2013;43(9):888-97. doi: 10.1016/j.ibmb.2013.02.007. PubMed PMID:
559 ISI:000323801100011.
560 31. Knecht ZA, Silbering AF, Cruz J, Yang L, Croset V, Benton R, et al. Ionotropic
561 Receptor-dependent moist and dry cells control hygrosensation in Drosophila. Elife.
562 2017;6. Epub 2017/06/18. doi: 10.7554/eLife.26654
563 e26654 [pii]. PubMed PMID: 28621663; PubMed Central PMCID: PMC5495567.
564 32. Knecht ZA, Silbering AF, Ni L, Klein M, Budelli G, Bell R, et al. Distinct
565 combinations of variant ionotropic glutamate receptors mediate thermosensation and
566 hygrosensation in Drosophila. Elife. 2016;5. Epub 2016/09/24. doi:
567 10.7554/eLife.17879
568 e17879 [pii]. PubMed PMID: 27656904; PubMed Central PMCID: PMC5052030.
569 33. Ni L, Klein M, Svec KV, Budelli G, Chang EC, Ferrer AJ, et al. The Ionotropic
570 Receptors IR21a and IR25a mediate cool sensing in Drosophila. Elife. 2016;5. Epub
571 2016/04/30. doi: 10.7554/eLife.13254

572 e13254 [pii]. PubMed PMID: 27126188; PubMed Central PMCID: PMC4851551.

573 34. Benton R. Multigene Family Evolution: Perspectives from Insect

574 Chemoreceptors. Trends Ecol Evol. 2015;30(10):590-600. Epub 2015/09/29. doi:

575 10.1016/j.tree.2015.07.009

576 S0169-5347(15)00188-3 [pii]. PubMed PMID: 26411616.

577 35. Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, et al.

578 Ancient protostome origin of chemosensory ionotropic glutamate receptors and the

579 evolution of insect taste and olfaction. PLoS Genet. 2010;6(8):e1001064. Epub

580 2010/09/03. doi: 10.1371/journal.pgen.1001064

581 e1001064 [pii]. PubMed PMID: 20808886; PubMed Central PMCID: PMC2924276.

582 36. Larsson MC, Domingos AI, Jones WD, Chiappe ME, Amrein H, Vosshall LB.

583 Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction.

584 Neuron. 2004;43(5):703-14. Epub 2004/09/02. doi: 10.1016/j.neuron.2004.08.019

585 S0896627304005264 [pii]. PubMed PMID: 15339651.

586 37. Joseph RM, Carlson JR. Drosophila Chemoreceptors: A Molecular Interface

587 Between the Chemical World and the Brain. Trends in Genetics. 2015;31(12):683-95.

588 doi: 10.1016/j.tig.2015.09.005. PubMed PMID: ISI:000366346000002.

589 38. Miyamoto T, Slone J, Song XY, Amrein H. A Fructose Receptor Functions as a

590 Nutrient Sensor in the Drosophila Brain. Cell. 2012;151(5):1113-25. doi:

591 10.1016/j.cell.2012.10.024. PubMed PMID: ISI:000311423500022.

592 39. Fujii S, Yavuz A, Slone J, Jagge C, Song X, Amrein H. Drosophila sugar

593 receptors in sweet taste perception, olfaction, and internal nutrient sensing. Curr Biol.

594 2015;25(5):621-7. Epub 2015/02/24. doi: 10.1016/j.cub.2014.12.058

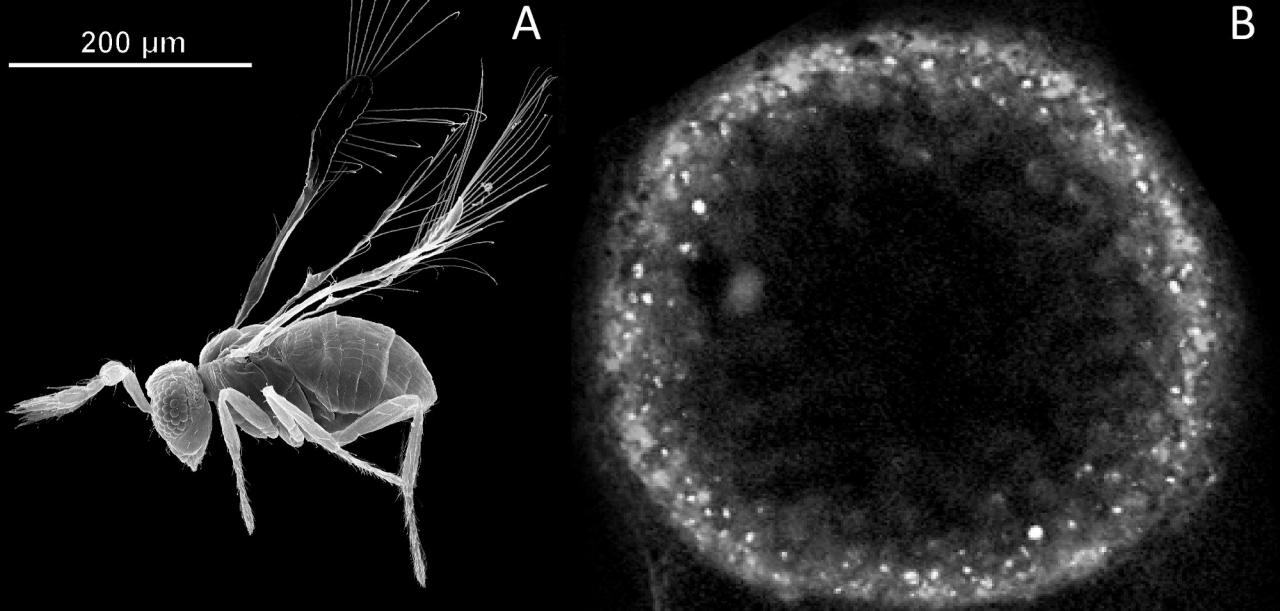
595 S0960-9822(14)01696-0 [pii]. PubMed PMID: 25702577; PubMed Central PMCID:
596 PMC4711800.
597 40. Robertson HM, Gadau J, Wanner KW. The insect chemoreceptor superfamily of
598 the parasitoid jewel wasp Nasonia vitripennis. Insect Mol Biol. 2010;19 Suppl 1:121-
599 36. Epub 2010/02/20. doi: 10.1111/j.1365-2583.2009.00979.x
600 IMB979 [pii]. PubMed PMID: 20167023.
601 41. Ahmed T, Zhang T, Wang Z, He K, Bai S. Gene set of chemosensory receptors
602 in the polyembryonic endoparasitoid Macrocentrus cingulum. Sci Rep. 2016;6:24078.
603 Epub 2016/04/20. doi: 10.1038/srep24078
604 srep24078 [pii]. PubMed PMID: 27090020; PubMed Central PMCID: PMC4835793.
605 42. Sheng S, Liao CW, Zheng Y, Zhou Y, Xu Y, Song WM, et al. Candidate
606 chemosensory genes identified in the endoparasitoid Meteorus pulchricornis
607 (Hymenoptera: Braconidae) by antennal transcriptome analysis. Comp Biochem Physiol
608 Part D Genomics Proteomics. 2017;22:20-31. Epub 2017/02/12. doi: S1744-
609 117X(17)30004-7 [pii]
610 10.1016/j.cbd.2017.01.002. PubMed PMID: 28187311.
611 43. Moreau SJM, Asgari S. Venom Proteins from Parasitoid Wasps and Their
612 Biological Functions. Toxins. 2015;7(7):2385-412. doi: 10.3390/toxins7072385.
613 PubMed PMID: ISI:000359191200004.
614 44. de Graaf DC, Aerts M, Brunain M, Desjardins CA, Jacobs FJ, Werren JH, et al.
615 Insights into the venom composition of the ectoparasitoid wasp Nasonia vitripennis
616 from bioinformatic and proteomic studies. Insect Mol Biol. 2010;19 Suppl 1:11-26.
617 Epub 2010/02/20. doi: 10.1111/j.1365-2583.2009.00914.x
618 IMB914 [pii]. PubMed PMID: 20167014; PubMed Central PMCID: PMC3544295.

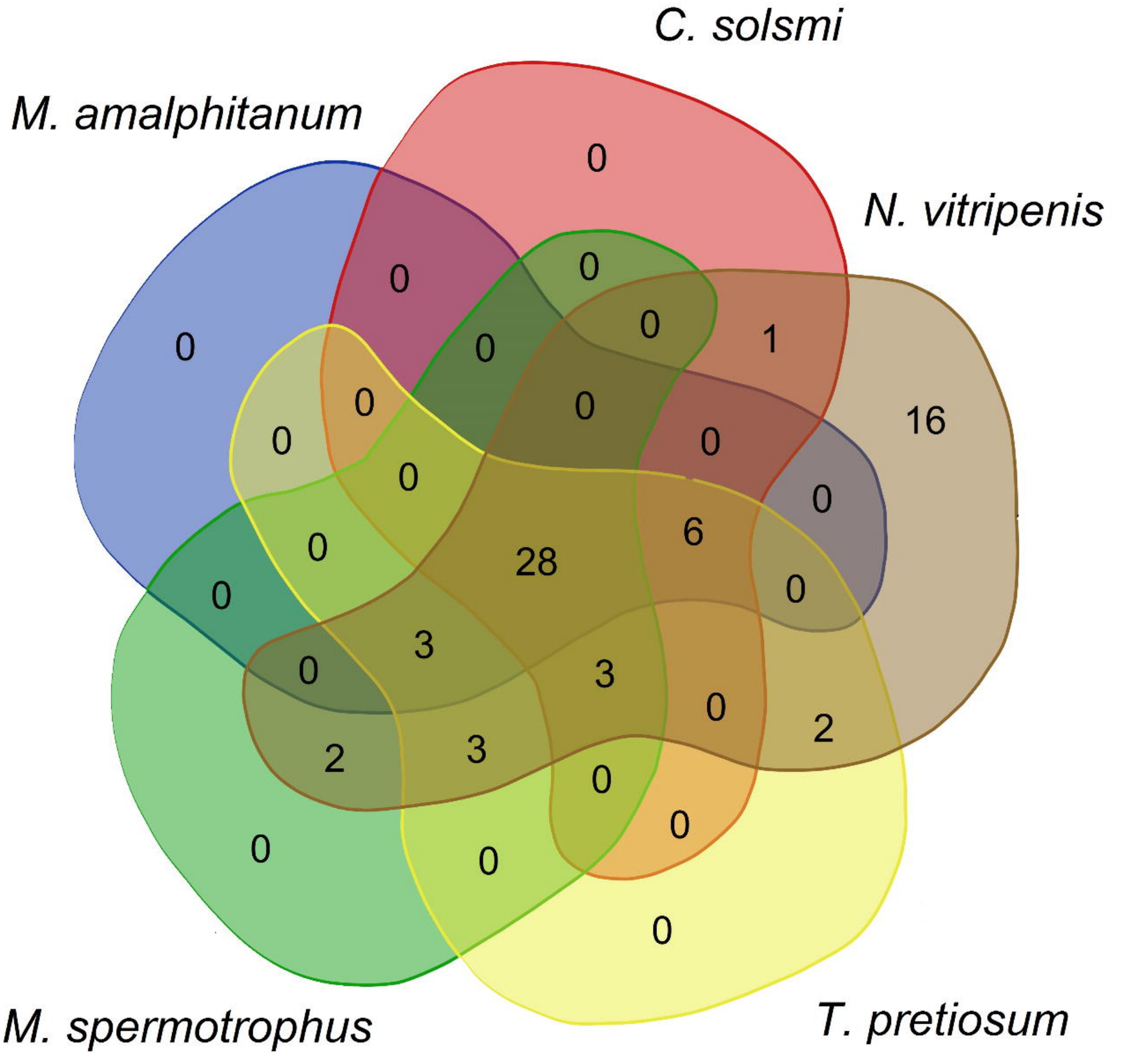
619 45. Paulson AR, Le CH, Dickson JC, Ehlting J, von Aderkas P, Perlman SJ.
620 Transcriptome analysis provides insight into venom evolution in a seed-parasitic wasp,
621 Megastigmus spermotrophus. Insect Molecular Biology. 2016;25(5):604-16. doi:
622 10.1111/imb.12247. PubMed PMID: ISI:000383345000008.
623 46. Kapheim KM, Pan H, Li C, Salzberg SL, Puiu D, Magoc T, et al. Social
624 evolution. Genomic signatures of evolutionary transitions from solitary to group living.
625 Science. 2015;348(6239):1139-43. Epub 2015/05/16. doi: 10.1126/science.aaa4788
626 science.aaa4788 [pii]. PubMed PMID: 25977371.
627 47. Patalano S, Vlasova A, Wyatt C, Ewels P, Camara F, Ferreira PG, et al.
628 Molecular signatures of plastic phenotypes in two eusocial insect species with simple
629 societies. Proc Natl Acad Sci U S A. 2015;112(45):13970-5. Epub 2015/10/21. doi:
630 10.1073/pnas.1515937112
631 1515937112 [pii]. PubMed PMID: 26483466; PubMed Central PMCID: PMC4653166.
632 48. Russell JE, Stouthamer R. The genetics and evolution of obligate reproductive
633 parasitism in Trichogramma pretiosum infected with parthenogenesis-inducing
634 Wolbachia. Heredity (Edinb). 2011;106(1):58-67. Epub 2010/05/06. doi:
635 10.1038/hdy.2010.48
636 hdy201048 [pii]. PubMed PMID: 20442735; PubMed Central PMCID: PMC3183844.
637 49. Kraaijeveld K, Bast J. Transposable element proliferation as a possible side
638 effect of endosymbiont manipulations. Mob Genet Elements. 2012;2(5):253-6. Epub
639 2013/04/04. doi: 10.4161/mge.22878
640 2012MGE0056 [pii]. PubMed PMID: 23550173; PubMed Central PMCID:
641 PMC3575435.

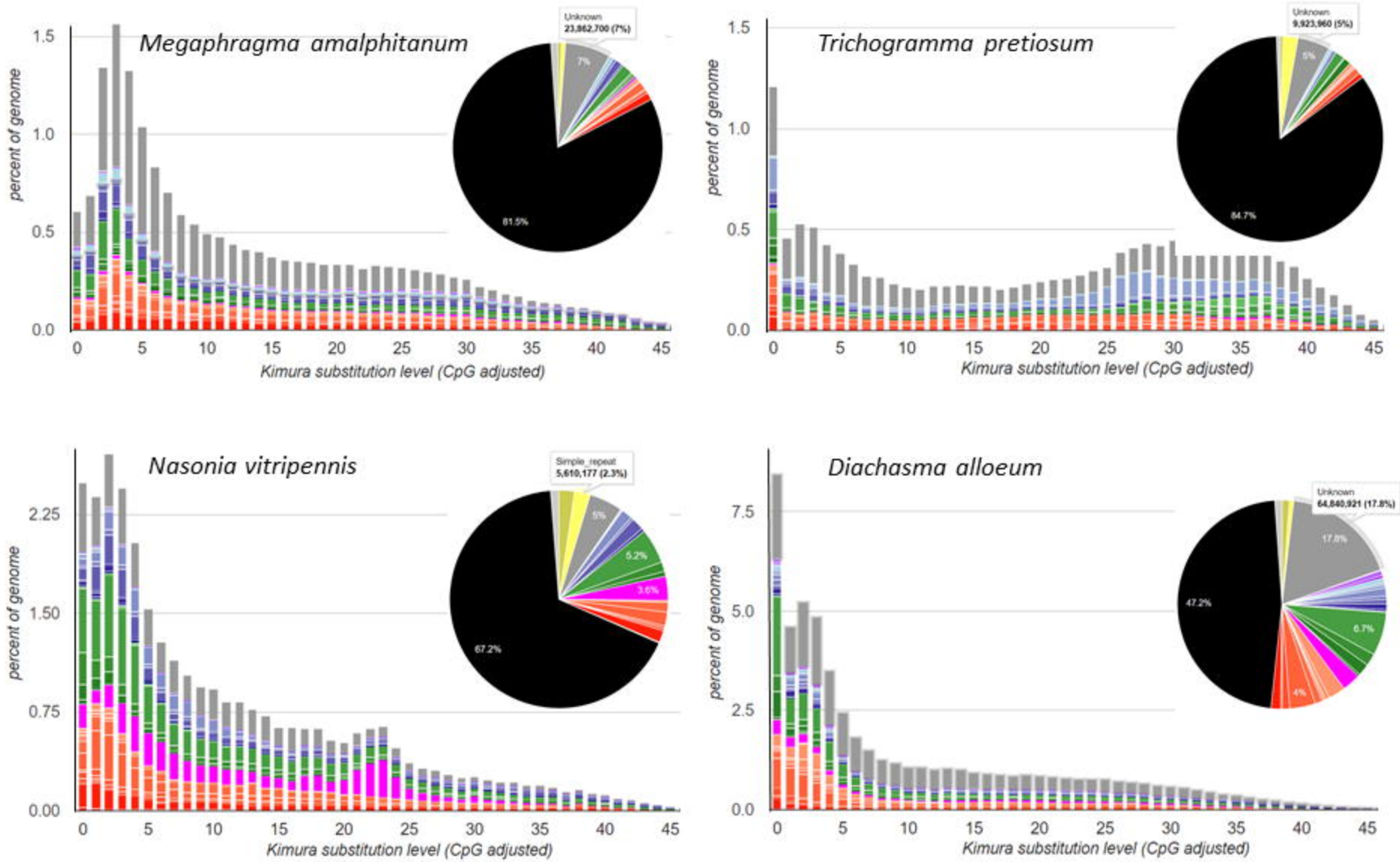
642 50. Bast J, Schaefer I, Schwander T, Maraun M, Scheu S, Kraaijeveld K. No
643 Accumulation of Transposable Elements in Asexual Arthropods. Molecular Biology
644 and Evolution. 2016;33(3):697-706. doi: 10.1093/molbev/msv261. PubMed PMID:
645 ISI:000371219500008.
646 51. Nedoluzhko AV, Sharko FS, Tsygankova SV, Boulygina ES, Sokolov AS,
647 Rastorguev SM, et al. Metagenomic analysis of microbial community of a parasitoid
648 wasp Megaphragma amalphitanum. Genom Data. 2017;11:87-8. Epub 2017/01/10. doi:
649 10.1016/j.gdata.2016.12.007
650 S2213-5960(16)30189-1 [pii]. PubMed PMID: 28066711; PubMed Central PMCID:
651 PMC5200880.
652 52. Lindsey AR, Werren JH, Richards S, Stouthamer R. Comparative Genomics of a
653 Parthenogenesis-Inducing Wolbachia Symbiont. G3 (Bethesda). 2016;6(7):2113-23.
654 Epub 2016/05/20. doi: 10.1534/g3.116.028449
655 g3.116.028449 [pii]. PubMed PMID: 27194801; PubMed Central PMCID:
656 PMC4938664.
657 53. Obbard DJ, Gordon KH, Buck AH, Jiggins FM. The evolution of RNAi as a
658 defence against viruses and transposable elements. Philos Trans R Soc Lond B Biol Sci.
659 2009;364(1513):99-115. Epub 2008/10/18. doi: 10.1098/rstb.2008.0168
660 NQ2U7841H17U6381 [pii]. PubMed PMID: 18926973; PubMed Central PMCID:
661 PMC2592633.
662 54. Schaack S, Gilbert C, Feschotte C. Promiscuous DNA: horizontal transfer of
663 transposable elements and why it matters for eukaryotic evolution. Trends Ecol Evol.
664 2010;25(9):537-46. Epub 2010/07/02. doi: 10.1016/j.tree.2010.06.001

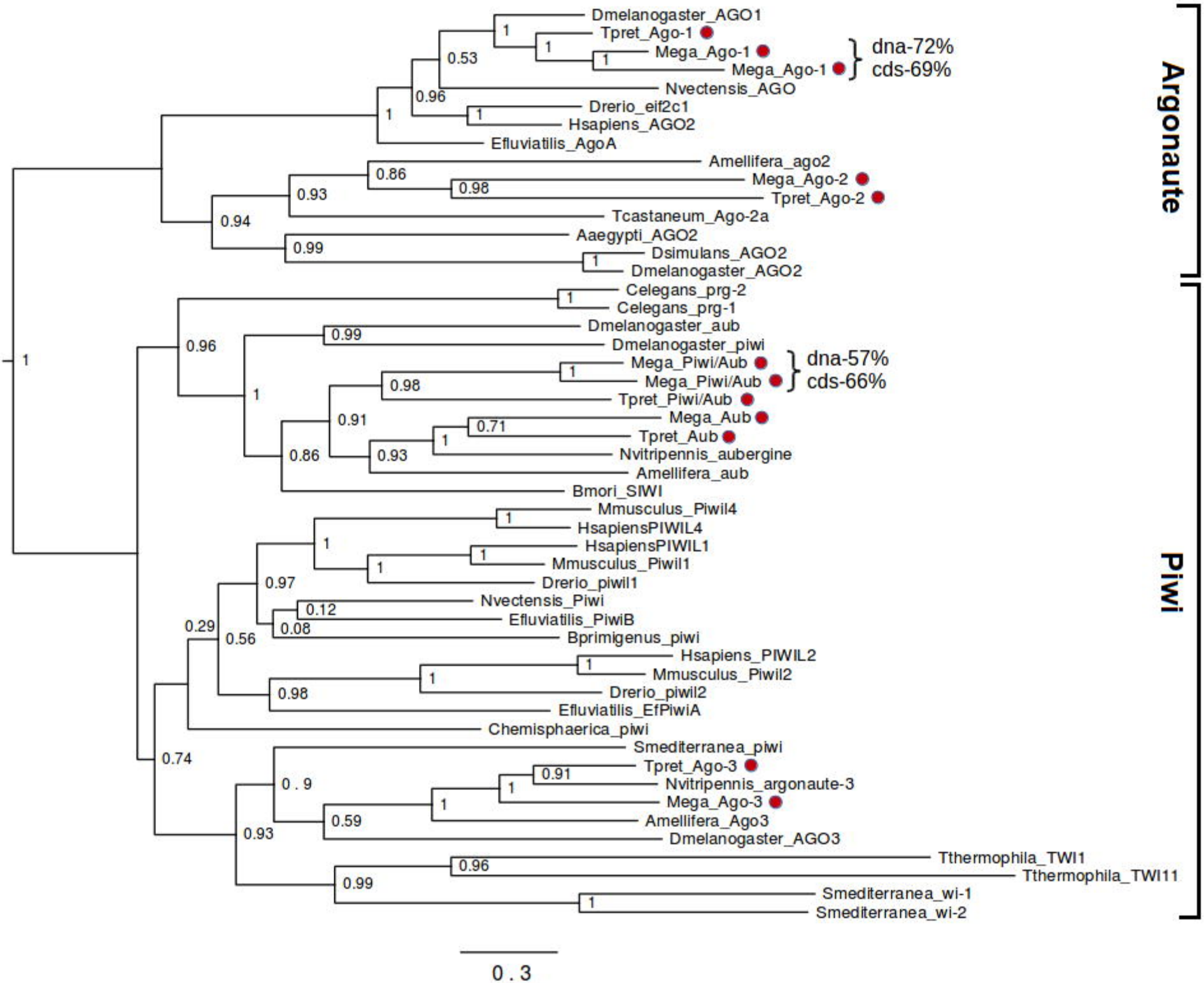
665 S0169-5347(10)00123-0 [pii]. PubMed PMID: 20591532; PubMed Central PMCID:
666 PMC2940939.
667 55. Feng Z, Caballe A, Wainman A, Johnson S, Haensele AFM, Cottee MA, et al.
668 Structural Basis for Mitotic Centrosome Assembly in Flies. Cell. 2017;169(6):1078-89
669 e13. Epub 2017/06/03. doi: S0092-8674(17)30591-3 [pii]
670 10.1016/j.cell.2017.05.030. PubMed PMID: 28575671; PubMed Central PMCID:
671 PMC5457487.
672 56. Yalgin C, Ebrahimi S, Delandre C, Yoong LF, Akimoto S, Tran H, et al.
673 Centrosomin represses dendrite branching by orienting microtubule nucleation. Nat
674 Neurosci. 2015;18(10):1437-45. Epub 2015/09/01. doi: 10.1038/nn.4099
675 nn.4099 [pii]. PubMed PMID: 26322925.
676 57. Basto R, Lau J, Vinogradova T, Gardiol A, Woods CG, Khodjakov A, et al.
677 Flies without centrioles. Cell. 2006;125(7):1375-86. doi: 10.1016/j.cell.2006.05.025.
678 PubMed PMID: ISI:000239104500019.
679 58. Zheng Z, Gao T, Hou Y, Zhou M. Involvement of the anucleate primary
680 sterigmata protein FgApsB in vegetative differentiation, asexual development, nuclear
681 migration, and virulence in Fusarium graminearum. FEMS Microbiol Lett.
682 2013;349(2):88-98. Epub 2013/10/15. doi: 10.1111/1574-6968.12297. PubMed PMID:
683 24117691.
684 59. Veith D, Scherr N, Efimov VP, Fischer R. Role of the spindle-pole-body protein
685 ApsB and the cortex protein ApsA in microtubule organization and nuclear migration in
686 Aspergillus nidulans. J Cell Sci. 2005;118(Pt 16):3705-16. Epub 2005/08/18. doi:
687 118/16/3705 [pii]
688 10.1242/jcs.02501. PubMed PMID: 16105883.

689 60. Suelmann R, Sievers N, Galetzka D, Robertson L, Timberlake WE, Fischer R.
690 Increased nuclear traffic chaos in hyphae of Aspergillus nidulans: molecular
691 characterization of apsB and in vivo observation of nuclear behaviour. Mol Microbiol.
692 1998;30(4):831-42. Epub 1999/03/27. PubMed PMID: 10094631.
693 61. Tsutsui ND, Suarez AV, Spagna JC, Johnston JS. The evolution of genome size
694 in ants. BMC Evol Biol. 2008;8:64. Epub 2008/02/28. doi: 10.1186/1471-2148-8-64
695 1471-2148-8-64 [pii]. PubMed PMID: 18302783; PubMed Central PMCID:
696 PMC2268675.
697 62. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory
698 of morphological evolution. Cell. 2008;134(1):25-36. Epub 2008/07/11. doi:
699 10.1016/j.cell.2008.06.030
700 S0092-8674(08)00817-9 [pii]. PubMed PMID: 18614008.


200 µm
A
B


C. solsmi
M. amalphitanum
N. vitripenis
M. spermotrophus
T. pretiosum
0
0
0
0
0
1
0
0
0
16
0
0
0
0
0
6
28
0
0
0
3
3
0
2
2
3
0
0
0
0
0


Megaphragma amalphitanum
Unknown
23,862,700 (7%)
7%
81.5%
percent of genome
Kimura substitution level (CpG adjusted)
Trichogramma pretiosum
Unknown
9,923,960 (5%)
5%
84.7%
Nasonia vitripennis
Simple_repeat
5,610,177 (2.3%)
5%
5.2%
3.6%
67.2%
Diachasma alloeum
Unknown
64,840,921 (17.8%)
17.8%
47.2%
6.7%
4%
Unknown
SINE/5S
SINE
LINE/Penelope
LINE/R2
LINE/Jockey-I
LINE/R1
LINE/LOA
LINE/L2
LINE/CR1
LINE/RTE
LINE
LTR
LTR/Gypsy
LTR/Copia
LTR/Pao
LTR/Ngaro
LTR/DIRS
RC/Helitron
DNA/TcMar
DNA/Sola
DNA/PiggyBac
DNA/P
DNA/MULE
DNA
DNA/Maverick
DNA/Kolobok
DNA/hAT
DNA/Harbinger
DNA/Ginger
DNA/Crypton
DNA/CMC


Dmelanogaster_AGO1
Tpret_Ago-1
Mega_Ago-1
Mega_Ago-1
dna-72%
cds-69%
Nvectensis_AGO
Drerio_eif2c1
Hsapiens_AGO2
Efluviatilis_AgoA
Amellifera_ago2
Mega_Ago-2
Tpret_Ago-2
Tcastaneum_Ago-2a
Aaegypti_AGO2
Dsimulans_AGO2
Dmelanogaster_AGO2
Argonaute
Celegans_prg-2
Celegans_prg-1
Dmelanogaster_aub
Dmelanogaster_piwi
Mega_Piwi/Aub
Mega_Piwi/Aub
dna-57%
cds-66%
Tpret_Piwi/Aub
Mega_Aub
Tpret_Aub
Nvitripennis_aubergine
Amellifera_aub
Bmori_SIWI
Mmusculus_Piwil4
HsapiensPIWIL4
HsapiensPIWIL1
Mmusculus_Piwil1
Drerio_piwil1
Nvectensis_Piwi
Efluviatilis_PiwiB
Bprimigenus_piwi
Hsapiens_PIWIL2
Mmusculus_Piwil2
Drerio_piwil2
Efluviatilis_EfPiwiA
Chemisphaerica_piwi
Smediterranea_piwi
Tpret_Ago-3
Nvitripennis_argonaute-3
Mega_Ago-3
Amellifera_Ago3
Dmelanogaster_AGO3
Tthermophila_TWI1
Tthermophila_TWI11
Smediterranea_wi-1
Smediterranea_wi-2
Piwi
0.3