Changes in brain microRNAs are associated with social evolution in bees

Karen M. Kapheim[1*], Beryl M. Jones[2ϕ], Eirik Søvik[3], Eckart Stolle[4], Robert M. Waterhouse[5], Guy Bloch[6], Yehuda Ben-Shahar[7]

[1] Department of Biology, Utah State University, Logan, UT 84322, USA

[2] Program in Ecology, Evolution, and Conservation Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] Department of Science and Mathematics, Volda University College, 6100 Volda, Norway

[4] Centre of Molecular Biodiversity Research, Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany

[5] Department of Ecology and Evolution, University of Lausanne and Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

[6] Department of Ecology, Evolution and Behavior, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, 91904, Israel

[7] Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA

[ϕ] Current affiliation: Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA

* Author for Correspondence:
Karen M. Kapheim
Biology Department
Utah State University
5305 Old Main Hill
Logan, UT 84322, U.S.A.
karen.kapheim@usu.edu
+1-435-797-0685

ABSTRACT

Evolutionary transitions to a social lifestyle in insects are associated with lineage-specific changes in gene expression, but the key nodes that drive these regulatory changes are largely unknown. We tested the hypothesis that changes in gene regulation associated with social evolution are facilitated by lineage-specific function of microRNAs (miRNAs). Genome scans across 12 bee species showed that miRNA copy-number is mostly conserved and not associated with sociality. However, deep sequencing of small RNAs in six bee species revealed a substantial proportion (20-35%) of detected miRNAs had lineage-specific expression in the brain, 24-72% of which did not have homologs in other species. Lineage-specific miRNAs disproportionately target lineage-specific genes, and have lower expression levels than shared miRNAs. The predicted targets of lineage-specific miRNAs are enriched for genes related to social behavior in social species, but they are not enriched for genes under positive selection. Together, these results suggest that novel miRNAs may contribute to lineage-specific patterns of social evolution. Our analyses also support the hypothesis that many new miRNAs are purged by selection due to deleterious effects on mRNA targets, and suggest genome structure is not as influential in regulating bee miRNA evolution as has been shown for mammalian miRNAs.

Keywords: Gene regulation; small non-coding RNA; microRNA targets; eusociality; lineage-specific

INTRODUCTION

Eusociality has evolved several times in the hymenopteran insects. In its most basic form, this lifestyle involves reproductive queens living with their worker daughters who forego direct reproduction to cooperatively defend the nest, care for their siblings, and forage for the colony. Due to the complex nature of this lifestyle, the evolution of eusociality likely requires modification of molecular pathways related to development, behavior, neurobiology, physiology, and morphology [1]. The evolution of eusociality is thus expected to involve both genetic changes as well as changes in the way the genome responds to the environment [2]. It is therefore unsurprising that recent studies aimed at identifying the genomic signatures of eusocial evolution in insects have found that social species share an increased capacity for gene regulation [3,4]. Evidence for this comes from signatures of rapid evolution of genes involved in transcription and translation, gene family expansions of transcription factors, and increasing potential for DNA methylation and transcription factor binding activity in conserved genes. Interestingly, while these types of regulatory changes are common to independent origins and elaborations of eusociality, the specific genes and regulatory elements involved are unique to each lineage [3–5]. This suggests that lineage-specific processes are influential in generating new patterns of gene regulation that contribute to social behavior.

Small, non-coding RNAs such as microRNAs (miRNAs) may be an important source of regulatory novelty associated with the evolution of phenotypic complexity, including eusociality. MiRNAs are short (~21-22 nt), noncoding RNAs that regulate protein-coding genes through post-transcriptional binding to the 3’ UTR region of messenger RNA (mRNA) transcripts, in most cases preventing translation or causing mRNA degradation [6]. Each miRNA can target

dozens to hundreds of mRNAs, and may therefore regulate multiple gene networks [6,7]. Like mRNAs, the majority of miRNAs are generated via Pol II transcription, and are spatially- and temporally-specific in their expression patterns. Thus, complex changes in gene regulation can be achieved with relatively minor changes in miRNA expression. This can result in major phenotypic shifts or fine-tuning of phenotypic optimization [6]. Novel miRNAs can originate in a variety of genomic features, including exons and introns of protein-coding and non-coding RNA genes, transposable elements, pseudogenes, or intergenic regions, and thus emerge and disappear over relatively rapid timescales [8–11]. It is thus not surprising that expansion of the miRNA repertoire is associated with the evolution of morphological complexity across the tree of life [9,12–14].

There is accumulating evidence for a role of miRNAs in regulating the social lives of insects. While most miRNAs seem to be conserved in major lineages of insects [15,16], expression levels vary across individuals performing different social functions, such as between workers performing different tasks in honey bees [17–19]. MiRNAs may also play a role in caste determination, as queen- and worker-destined larvae express different sets of miRNAs throughout development in honey bees [20–22] and bumble bees [23]. Additionally, miRNAs play a role in regulating some physiological correlates of social behavior in honey bees, including activation of ovaries in queens and workers [24] and response to the reproductive protein *vitellogenin* [25]. Together, these studies suggest that miRNAs could play a role in the evolution of eusociality through their effects on gene regulatory networks involved in socially-relevant traits. A rigorous test of this hypothesis requires comparisons of the presence, expression, and function of miRNAs across related species that vary in social organization.

Here we present a comprehensive comparative analysis of miRNAs across bee species with variable social organization. These solitary and social species shared a common ancestor ~75-110 mya [26]. Previous comparative studies of miRNAs associated with eusociality have relied on the parasitoid wasp, *Nasonia vitripennis*, as a solitary comparison [16]. This is a far more distant relative to the social insects, sharing a last common ancestor with bees nearly 200 mya [26]. Moreover, the parasitoid lifestyle of *N. vitripennis* is different from that of eusocial bees in nearly every way. The lifestyle of solitary bees, such as the ones we include in this study, share many features of their natural history with the presumed ancestors from which eusociality evolved.

We first looked for miRNA repertoire expansions associated with eusociality by scanning 12 bee genomes for known miRNAs, and statistically evaluating copy-number of each miRNA type with regard to differences in sociality in a phylogenetic model. We then described and compared miRNAs expressed in the brains of six bee species from three families that include repeated origins of eusociality. We tested the hypothesis that changes in gene regulatory function associated with social evolution are facilitated by lineage-specific miRNA regulatory function with two predictions: (1) If lineage-specific miRNAs are assimilated into ancestral gene networks, their predicted target genes should be ancient and conserved. (2) If lineage-specific miRNAs play a role in social evolution, their predicted targets should be enriched for genes that function in social behavior (e.g., caste-biased expression) or genes that are under selection in social species.

126 **MATERIALS AND METHODS**

127 **Sample Acquisition**

128 We used adult females from six bee species for our study (Fig. 1). These species include
129 both eusocial and solitary species with well-studied behavior from three families. *Megalopta*
130 *genalis* samples were collected on Barro Colorado Island, Panama in 2015 and exported to the
131 U.S.A. (permit SEX/A-37-15). *Nomia melanderi* samples were collected in Touchet, WA,
132 U.S.A. with permission from land owners. *Megachile rotundata* samples were collected from
133 Logan, UT, U.S.A. on the Utah State University campus. *Bombus impatiens* samples were
134 collected from a commercial colony purchased from BioBest. *Bombus terrestris* samples were
135 collected from colonies obtained from Pollination Services Yad-Mordechai, Kibbutz Yad-
136 Mordechai, Israel. *Apis mellifera* samples were collected from hives in Urbana-Champaign, IL or
137 the Tyson Research Field Station, MO, U.S.A. *A. mellifera* and *Bombus* samples were workers.
138 *M. genalis* samples were lab-reared females. All other samples were reproductive females. All
139 samples were collected into liquid nitrogen and stored at -80 °C until dissection.
140

141 **RNA Isolation and Sequencing**

142 Head capsules from *B. impatiens*, *M. genalis*, and *N. melanderi* samples were dissected
143 after incubation in RNALater ICE (Ambion) to remove the entire brain. We used the mirVana
144 miRNA Isolation kit with phenol (Ambion) to isolate total RNA from individual brains. Total
145 RNA was sent to the University of Illinois Roy J. Carver Biotechnology Center for library
146 preparation with the Illumina TruSeq Small RNA Sample Preparation kit and sequencing.
147 Libraries were pooled, quantitated by qPCR, and sequenced on one lane for 51 cycles on a HiSeq
148 2500.

Whole brains of *A. mellifera*, *B. terrestris*, and *M. rotundata* were dissected from frozen heads. Total RNA from individual brains was isolated using TRIzol reagent (Thermo Fisher Scientific). All subsequent small-RNA sequencing steps were performed by the Genome Technologies Access Center at Washington University, using their Illumina TruSeq pipeline. Total RNA samples were size fractionated and multiplexed. Single-end small RNA libraries were prepared using the SMARTer kit (Clontech). Up to 12 barcoded libraries from a single species were run on a single Illumina HiSeq 2500 lane.

miRNA Discovery and Quantification

We used miRDeep2 [27] to identify and quantify miRNAs expressed in the brains of each species, with a three-step process of miRNA detection to identify homologous miRNAs between species. First, we gathered a set of mature miRNA sequences previously described in other insect species (Table S1). Reads for each sample were quality filtered (minimum length 18, removal of reads with non-standard bases), adapter-trimmed, and aligned to the species' genome (Table S2) with the mapper.pl script. Approximately 60-84% of reads successfully mapped.

We then identified known and novel miRNAs in each sample with the miRDeep2.pl script, using our curated set of insect miRNAs (Table S1) as known mature sequences. We followed this with the quantifier.pl script to generate sets of known and novel miRNAs in each sample, along with quantified expression information for each. We then filtered novel miRNAs in each species according to the following criteria: no rRNA/tRNA similarities, minimum of five

171 reads each on the mature and star strands of the hairpin sequence, and a randfold p-value < 0.05.
172 Randfold describes the RNA secondary structure of potential pre-miRs [27].
173
174 We used these filtered miRNAs in a second run of detection and quantification, adding
175 the mature sequences of novel miRNAs from each species to our set of known miRNAs, and
176 repeated the pipeline above. This allowed detection of homologous miRNAs (based on matching
177 seed sequences) that are not represented in miRBase across our species. We applied the same set
178 of filtering criteria as for our first run.
179
180 Some of the novel miRNAs may exist in the genomes of other bees, even if they are not
181 expressed. We used blastn (-perc_identity 50 -evalue 1e-5) to search for homologous precursor
182 miR (pre-miR) sequences in 12 bee genomes (Table S2) for each of the novel miRNAs without a
183 matching seed sequence.
184
185 **miRNA Localization**
186 We used bedtools intersect [28] to find overlap of miRNAs with predicted gene models
187 (Table S3), and repetitive element repeatmasker [29] annotations from previously established
188 repeat libraries [4,30–33].
189
190 **Target Prediction**
191 We extracted potential target sites 500 bp downstream from each gene model using
192 bedtools flank and getfasta [28], following previous studies [21] and an average 3' UTR region
193 of 442 nt in *Drosophila melanogaster* [34]. Target prediction was run with miRanda v3.3 [35]

(minimum energy threshold -20, minimum score 140, strict alignment to the seed region [-en -20 -sc 140 –strict]) and RNAhybrid v2.12 [36] (minimum free energy threshold -20). We kept only miRNA-target gene pairs that were predicted by both programs with $p < 0.01$.

Target Age and Functional Enrichment

Gene ages were determined using orthogroups from OrthoDB v9 [37], which includes *A. mellifera*, *B. impatiens*, *B. terrestris*, and *M. rotundata*. Gene sets of *M. genalis* and *N. melanderi* were mapped to Metazoa-level (330 species) orthogroups. Gene sets of *M. genalis* and *N. melanderi* were mapped to Metazoa-level (330 species) orthogroups. Gene ages were inferred from the taxonomic breadth of all species in each orthogroup: Vertebrata (≥ one vertebrate), Metazoa (≥ one non-arthropod and non-vertebrate metazoans), Arthropoda (≥ one non-insect arthropods), Insecta (≥ one non-holometabolous insects), Holometabola (≥ one non-hymenopteran holometabolous insects), Hymenoptera (≥ one non-Aculeata hymenopterans), Aculeata (≥ one non-Apoidea Aculeata), Apoidea (≥ one other Apoidea). Genes without identifiable orthologs were labeled 'Unique'.

Gene Ontology (GO) terms for each species were derived from a previous study [4], with the exception of *B. impatiens*, for which GO terms were assigned based on reciprocal blastp (evalue < $1e^{-5}$) between two sets of gene models (OGS v1.2 and OGS v1.0). Functional enrichment was performed with the GOstats package [38] in R [39]. We included terms enriched at an unadjusted $p < 0.1$.

Enrichment tests of lineage-specific miRNA targets with previous studies

For each species, brain or head gene expression datasets related to socially relevant phenotypes (e.g., caste) and genes under positive selection were compared against targets of lineage-specific miRNAs. The complete list of included studies and gene lists are in Table S4. For *M. genalis* caste data, RNAseq reads from Jones et al. [40] (NCBI PRJNA331103) were trimmed using Trimmomatic (v. 0.36) [41] and aligned to an unpublished genome assembly of *M. genalis* (NCBI PRJNA494872) using STAR (v. 2.5.3) [42]. Reads were mapped to gene features using featureCounts in the Subread package (v. 1.5.2) [43]. Remaining differential expression analysis followed the methods of Jones et al. [40] using edgeR [44].

We also tested datasets identifying genes under selection in bee species or across social lineages of bees for enrichment of lineage-specific miRNA targets (Table S4). When necessary, we used reciprocal blastp (evalue < $10e^{-5}$) to identify orthologous genes across species, and only genes with putative orthologs were included in the analysis. Hypergeometric tests (using phyper in R) were used to test for significance of over- or under-enrichment between each pair of lists. The representation factor (RF) given represents the degree of overlap relative to random expectation (RF=1). RF is calculated as RF=x/E, where x is the number of genes in common between two lists and E is the expected number of shared genes (E = nD/N, where n is the number of genes in list 1, D is the number of genes in list 2, and N is the total number of genes.)

miRNA Diversification

We performed genome scans for small RNAs across 12 bee genomes (Table S2) using covariance models implemented with Infernal cmsearch using the gathering threshold for

inclusion (--cut_ga) [45] to find all Rfam accessions in each bee genome. We used Spearman rank regressions to test for significant associations between miRNA copy-number and social biology. We categorized each species as solitary, facultative basic eusocial, obligate basic eusocial, or obligate complex eusocial following Kapheim et al. [4]. We used the ape package [46] in R [39] to calculate phylogenetic independent contrasts for both social organization and miRNA copy-number, cor.test to implement the Spearman's rank correlation, and p.adjust with the Benjamini-Hochberg method to correct for multiple comparisons.

RESULTS

Low levels of miRNA copy-number variation among bee genomes

Our genome scans revealed very little variation in copy-number of most miRNAs among bee genomes. Of the 50 miRNA Rfam accessions, half had the same number of copies (1 or 2) in all 12 bee genomes (Table S5). The mean copy-number across all miRNAs in all bee genomes was 1.19 ± 0.74. One exception was miR-1122, for which we found 70 copies in *M. genalis*, but no copies in the other species. We did not find any significant associations between miRNA copy-number and social organization (Table S5).

Expressed miRNA diversity in bee brains

We identified 97-245 known and novel miRNAs expressed in the brains of each of our six species (Table S6). The majority of these were located in intergenic regions or introns (Table 1). Each species had at least one miRNA that originated from exons of protein-coding genes and repetitive DNA (Table 1). Most of the overlap between miRNA precursors and repetitive DNA

corresponded to uncharacterized repeat elements, with very few overlaps with well-characterized transposons or retrotransposons (Table 1).

Most of the detected miRNAs in each species had known homologs in at least one other species. However, each species had a substantial proportion (20-35%) of detected miRNAs with lineage-specific expression in the brain (Table 1; Fig. 1A), 24-72% of which did not have any known homologs in other species (Table 1). We defined lineage-specific miRNAs as those with lineage-specific expression and for which no seed match with a known mature miRNA was identified (Table 1, columns 6-7), because these show the most evidence of being real miRNAs that are unique to a particular species. (Sequence similarity of pre-miRs in the genome of other bee species is not sufficient evidence that a mature miRNA is transcribed.) Lineage-specific miRNAs had significantly lower expression levels compared with homologous miRNAs in each species (t-tests: *A. mellifera*, $p = 3.81e^{-05}$, *B. impatiens*, $p = 0.003$, *B. terrestris*, $p = 0.006$, *M. genalis*, $p = 0.0003$, *M. rotundata*, $p = 8.00e^{-05}$, *N. melanderi*, $p = 0.02$).

Lineage-specific miRNAs were localized both within genes and intergenically. The proportion of lineage-specific miRNAs that were intra- or intergenic was similar to miRNAs with homologs for every species except *N. melanderi*, for which a disproportionate number of lineage-specific miRNAs were intragenic ($\chi^2 = 4.78$, $p = 0.03$). Genes that serve as hosts for intragenic lineage-specific miRNAs were not significantly older than would be expected by chance (i.e., belong to orthogroups shared with vertebrates) in any species (hypergeometric tests: $p = 0.14$-0.76). Across all species, genes that serve as hosts for intragenic lineage-specific

283 miRNAs were not significantly older than genes hosting miRNAs with known homologs (χ^2
284 tests: p = 0.05-0.89).

285

286 **Table 1.** Localization of miRNAs in the genomes of six bee species. Numbers not in parentheses
287 represent features on the same strand as the pre-miR. Numbers in parentheses indicate strand
288 mismatch. Some pre-miRs overlapped with one or more genes on both the same and opposite
289 strands, and are thus counted twice (*A. mellifera* and *M. genalis* – 1, *B. impatiens* – 5, *B.*
290 *terrestris* – 4, *N. melanderi* – 3). Seed match – Mature miR had a seed match with a known miR;
291 Pre-miR – Successful blastn hit to the pre-miR sequence in at least one other bee genome;
292 Unique – No homolog was found in other species (seed match to mature or blastn hit to pre-
293 miR).

			miRs with lineage-specific expression in the brain				Location in the genome				
Species	**Sociality**	**Expressed miRs**	**Total**	**Seed match**	**Pre-miR**	**Unique**	**Intergenic**	**Exon**	**Intron**	**Transposable element**	**Uncharacterized repetitive DNA**
Apis mellifera	Complex eusocial	97	25	0	15	10	45	5	38 (10)	0	0
Bombus impatiens	Basic eusocial	245	49	6	31	12	129	4 (1)	89 (27)	7	32
Bombus terrestris		150	31	2	21	8	76	1 (1)	56 (20)	13	36
Megalopta genalis	Facultative eusocial	105	37	7	5	25	63	3	30 (10)	2	28
Megachile rotundata	Solitary	99	27	9	0	18	48	8 (1)	37 (5)	2	15
Nomia melanderi	Solitary	97	29	5	3	21	50	8	34 (8)	2	27

294

295 Of the miRNAs with homologs, most were expressed in all six species, but we detected
296 one miRNA (miR-305) that was expressed in the brains of each of the social, but not the solitary,
297 species. Although we did not detect expression of miR-305 in the two solitary species, *M.*
298 *rotundata* and *N. melanderi*, genome scans of each species against the Rfam database suggested
299 all bee species have one copy of this miRNA (Table S5). Predicted targets of miR-305 differed
300 across species. *Oxysterol* (OG EOG091G0FV2) was a common target among the (social) Apidae

bees, but was not among the targets for *M. genalis*. However, *arylformamidase* (OG EOG091G0KT8), which is also involved in lipid metabolism and transport, was a predicted target in *M. genalis*. *Synaptobrevin* (OG EOG091G0MPE), which is involved in synaptic plasticity and neurotransmitter release, was a predicted target of miR-305 in *B. impatiens*.

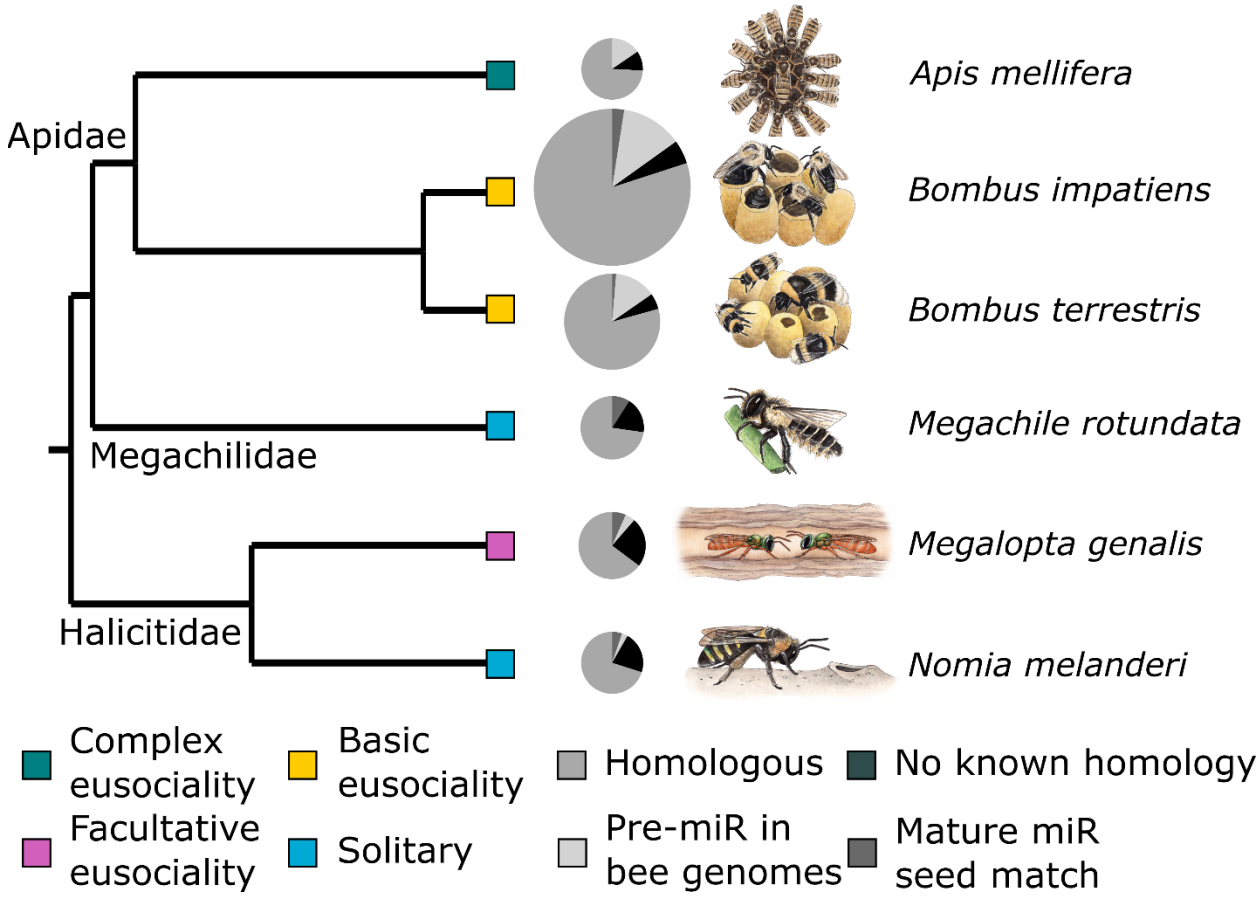


Fig. 1. Diversity of miRNAs expressed in the brains of six bee species. The three types of homology (shades of grey) correspond to those in Table 1. Black – has not been previously detected in other species. Pie size corresponds to number of miRNAs detected from small RNA sequencing. Boxes indicate social organization (green – complex eusociality, yellow – basic eusociality, pink – facultative eusociality, blue – solitary). Phylogenetic relationships are following previous studies [26,30,47].

314 **Lineage-specific miRNAs preferentially target lineage-specific genes and genes with caste-**
315 **biased expression, but not genes under selection**

316 If lineage-specific changes in gene regulatory function associated with social evolution
317 are facilitated by novel miRNAs inserted into existing gene networks, then predicted targets of
318 lineage-specific miRNAs should be highly conserved and enriched for genes with known
319 functions in social evolution. Most of the predicted mRNA targets of lineage-specific miRNAs
320 were highly conserved and belonged to orthogroups shared by vertebrates (Fig. 2; Table S8).
321 However, most genes in each genome are also highly conserved, and there was not a significant
322 enrichment for conserved genes among predicted targets of lineage-specific miRNAs, beyond
323 what would be expected by chance (hypergeometric test: $p > 0.99$). We did, however, find a
324 significant enrichment for genes unique to each species among the predicted targets of lineage-
325 specific miRNAs (hypergeometric tests: *A. mellifera* – RF = 1.51, $p = 5.44e^{-5}$; *B. impatiens* – RF
326 = 1.28, $p = 0.02$; *B. terrestris* – RF = 1.78, $p = 1.90e^{-6}$; *M. rotundata* – RF = 1.79, $p = 0.0002$; *M.*
327 *genalis* – RF = 1.62, $p = 1.48e^{-12}$; *N. melanderi* – RF = 1.78, $p = 9.02e^{-5}$), indicating that novel
328 miRNAs are more likely to target novel genes than would be expected by chance (Fig. 2; Table
329 S8).
330

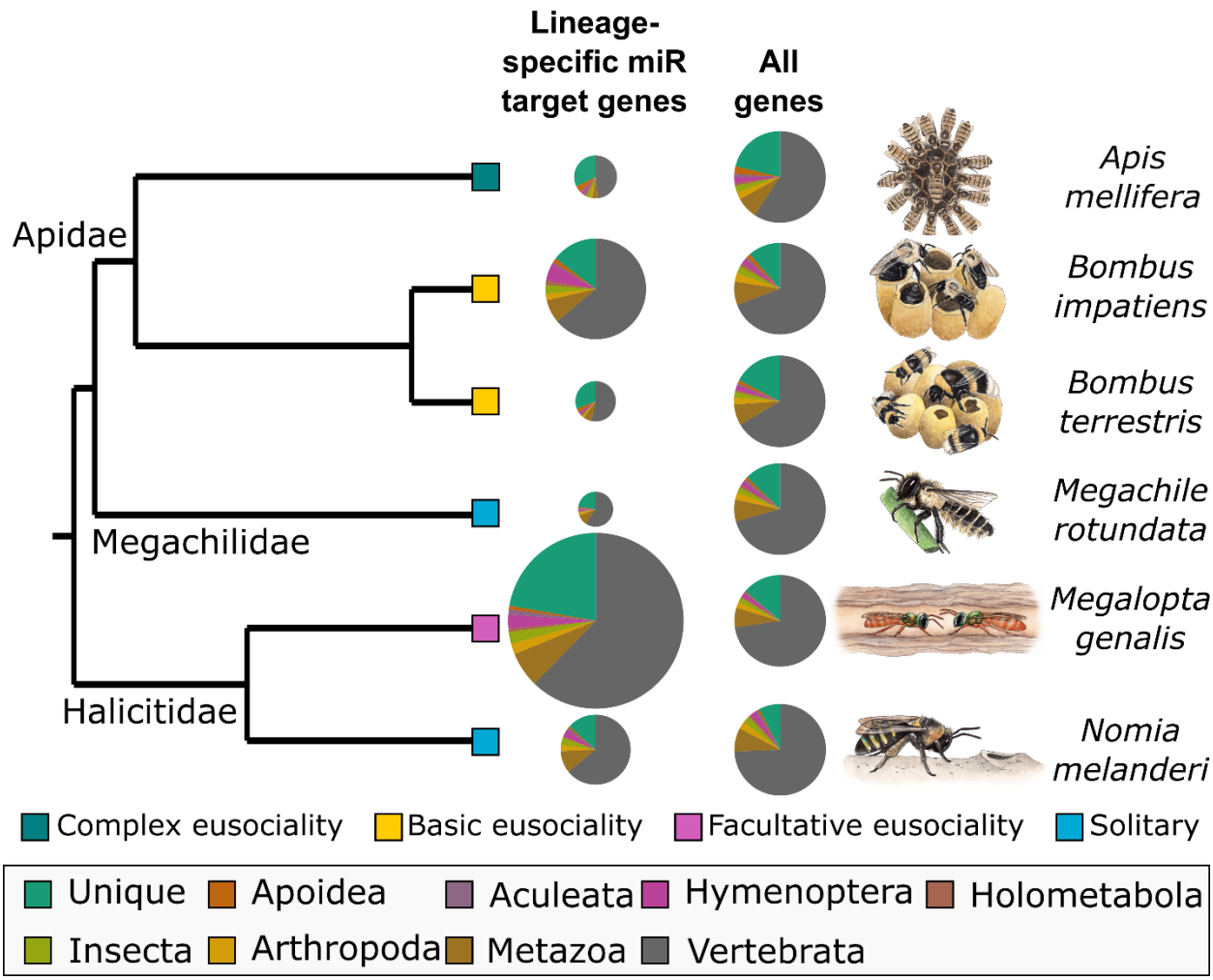


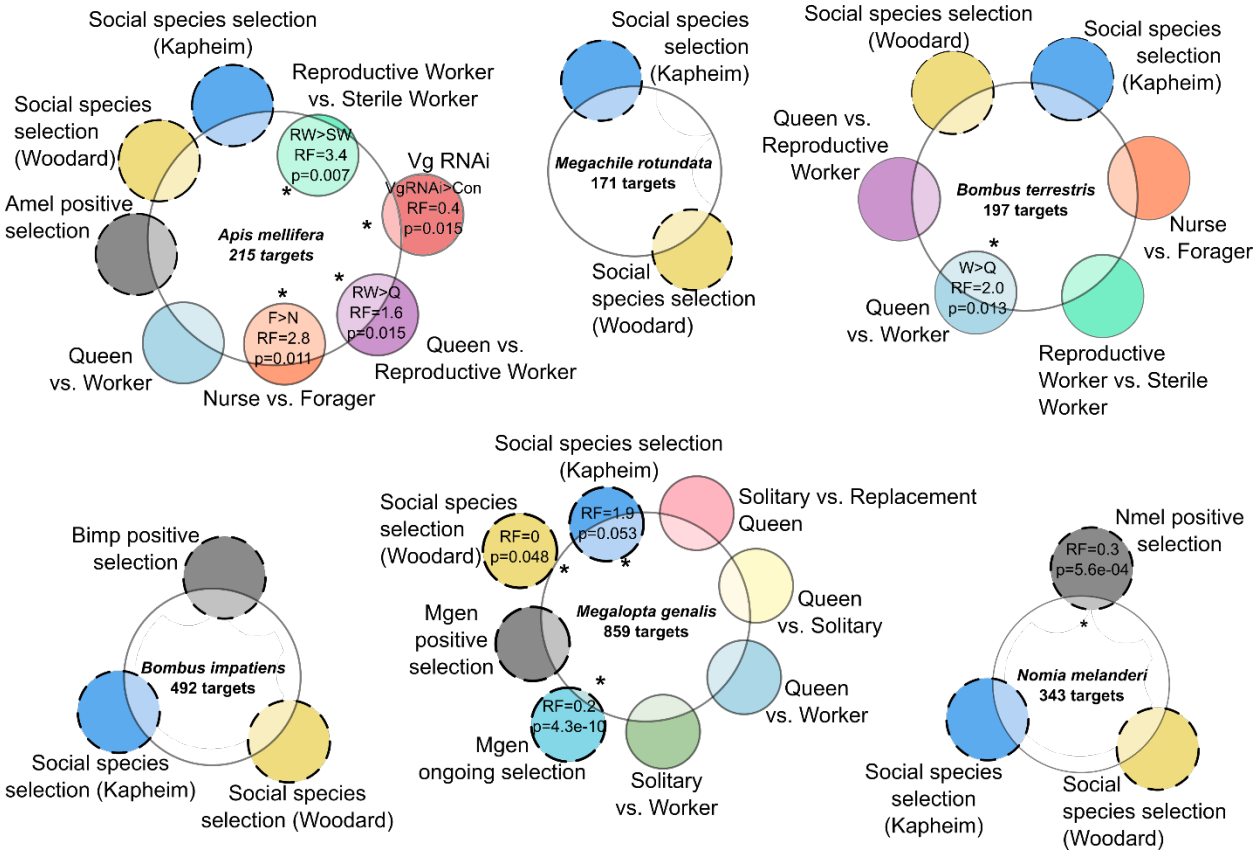
Fig. 2. Age of genes targeted by lineage-specific miRNAs. Genes predicted to be targeted by lineage-specific miRNAs are more likely to be unique to each species than predicted by chance. Pie chart size is scaled to number of predicted target genes for lineage-specific miRNAs, but not for all genes. Color slices indicate orthogroup age for each predicted gene. The green slice (lineage-specific genes) is larger for the set of genes predicted to be targeted by lineage-specific miRNAs than for all genes.

We found mixed support for the prediction that novel miRNAs should target genes that function in social behavior and evolution. The predicted targets of lineage-specific miRNAs were enriched for genes differentially expressed between castes in the social Apidae (*A. mellifera* and *B. terrestris*), but not Halictidae (*M. genalis*) (Fig. 3; Table S4). In *A. mellifera*, this included genes upregulated in the brains of reproductive workers, compared with sterile workers

(hypergeometric test: RF = 3.4, p = 0.007) and queens (hypergeometric test: RF = 1.6, p = 0.015) [48], as well as genes upregulated in the brains of foragers compared with nurses (hypergeometric test: RF = 2.8, p = 0.011) [49]. However, there was no significant enrichment for genes differentially expressed between nurse and forager honey bee brains in a later study (hypergeometric test: p = 0.09) [50]. In *B. terrestris*, we found significant overlap between the predicted targets of lineage-specific miRNAs and genes that are upregulated in workers, compared to queens (whole body, including brain; hypergeometric test: RF = 2, p = 0.013). We did not find significant overlap with genes differentially expressed in the brains of nurses and foragers (hypergeometric test: p = 0.103) [51] or between reproductive and sterile worker brains (hypergeometric test: p = 0.39) [52], but these were much more limited gene sets. To our knowledge, there are no studies of gene expression differences between *B. impatiens* castes, so we could not evaluate target overlap with caste-biased genes in this species. We did not find significant enrichment for caste-biased genes in the brains of the facultatively eusocial *M. genalis* (hypergeometric test: p = 0.25).

Contrary to our prediction, targets of lineage-specific miRNAs were not significantly enriched for genes under selection in any species. We assessed overlaps between genes undergoing positive directional selection in *A. mellifera* [53], *B. impatiens* [54], *M. genalis* [33], and *N. melanderi* [32] and the predicted targets of lineage-specific miRNAs in each species. There was no significant enrichment for targets of lineage-specific miRNAs with genes under positive directional selection in any species (Table S4). In fact, genes under selection in the halictid bees were significantly depleted for targets of lineage-specific miRNAs (hypergeometric test: *M. genalis* – RF = 0.2, p = $4.28e^{-10}$; *N. melanderi* – RF = 0.3, p = $5.59e^{-4}$). We also assessed

368 overlaps with genes previously found to be under positive selection in social species, compared
369 to solitary species [4,55], but found only marginally significant overlap [4] or depletion [55] with
370 predicted targets of lineage-specific genes in one species (hypergeometric tests: *M. genalis* – RF
371 = 1.9, p = 0.053; RF = 0, p = 0.05; Table S4).
372



373

374 **Fig. 3.** Predicted targets of lineage-specific miRNAs in relation to social behavior. Genes that are
375 both predicted targets of lineage-specific miRNAs and genes with differential expression in a
376 social context (solid outlines) or genes under selection (dashed outlines) are represented by
377 overlapping circles for each study and species. Numbers of lineage-specific miRNA targets are
378 given for each species. Colors indicate different studies. Overlaps not significantly different from
379 random (representation factor, RF=1) are unlabeled, while significant over- or under-enrichments
380 are marked with asterisks with RF and p-value as indicated.
381

382 **DISCUSSION**

383 Eusociality is a major evolutionary innovation that requires regulatory changes in a wide
384 range of molecular pathways [1]. We tested the hypothesis that miRNAs play a role in the
385 evolution of eusociality via their regulatory effects on gene networks by comparing miRNA
386 expression in three eusocial and three solitary bee species from three families. Our results
387 provide several lines of support for this hypothesis.

388

389 We identified a single miRNA (miR-305) that was expressed exclusively in the brains of
390 the social bees in our study. The presence of this miRNA in the solitary bee genomes suggests
391 that an evolutionary shift in expression pattern has accompanied at least two independent origins
392 of eusociality in bees. This miRNA coordinates Insulin and Notch signaling in *D. melanogaster*,
393 and both of these pathways are important regulators of social dynamics in insects [56–60].
394 Interestingly, this miRNA is also upregulated in worker-destined compared to queen-destined
395 honey bee larvae, and may thus play a role in caste differentiation [22]. Further investigation
396 with additional social and solitary species is necessary to determine how this miRNA may
397 influence social behavior across species.

398

399 We focused attention on miRNAs for which no mature miRNAs with seed matches were
400 detected in any other species, because these have the potential to influence the lineage-specific
401 patterns of gene regulatory changes previously shown to influence social evolution [3,4]. We
402 hypothesized that if novel miRNAs are inserted into existing gene networks that become co-
403 opted for social evolution, they should target genes that are highly conserved across species.
404 Instead, we find that the targets of lineage-specific miRNAs are enriched for lineage-specific

405 genes, while genes belonging to ancient orthogroups were not more likely to be targets than
406 expected by chance. This suggests that novel miRNAs co-evolve with novel genes, as has been
407 shown for the evolution of cognitive function in humans [61]. Previous work in honey bees has
408 shown that taxonomically-restricted genes play an important role in social evolution. Expression
409 of taxonomically-restricted genes is significantly biased toward glands with specialized functions
410 for life in a social colony (e.g., the hypopharyngeal gland and the sting gland) [62], and toward
411 genes that are upregulated in workers [63]. Thus, it is reasonable to expect that new miRNAs
412 targeting new genes could have important social functions.
413
414 Alternatively, it is possible that new miRNAs targeting lineage-specific genes are
415 transient and will be purged by natural selection because they are less integrated into existing
416 gene networks [10,64,65]. Emergent miRNAs are expected to initially have limited expression to
417 mitigate potential deleterious effects on the protein-coding genes they target. Thus, lineage-
418 specific miRNAs with low levels of expression may be in the process of being purged and may
419 not have accumulated gene targets with important functions [9,10]. Evidence for this model
420 comes from primates [66] and flies [11,67]. Likewise, we find that lineage-specific miRNAs are
421 expressed at significantly lower levels than those with at least one homolog in another species. A
422 purging process could explain why there are large differences in the numbers of miRNAs
423 detected in even closely related species (e.g., the two *Bombus* species). Functional analysis of
424 lineage-specific genes in additional tissues and life stages will help to resolve their roles in social
425 evolution.
426

We find support for the prediction that lineage-specific miRNAs should target genes with social function in the Apidae (e.g., honey bees and bumble bees), but not the Halictidae (*M. genalis*). One explanation for this pattern is technical. We define genes with social functions as those that are differentially expressed among castes. The genetic basis of social behavior has been much better studied in honey bees and bumble bees than in any other species, and the sets of genes known to function in sociality is thus richer for apids than for halictids. Further, not all genes that function in social behavior are expected to be differentially expressed in the brains of different castes, and our analysis is thus likely to exclude some important genes.

Nonetheless, our results reflect differences in the antiquity and degree of social complexity, and thus caste-biased gene expression patterns, between apid and halictid bees. Eusociality has a deeper origin in the Apidae than in Halictidae [47,68], and thus more time has accumulated for associated changes in miRNA regulation to evolve. Unlike for honey bees and bumble bees, which cannot live outside of social colonies, eusociality is facultative in *M. genalis*. As such, caste traits are not fixed during development, and females who served as non-reproductive workers can become reproductive queens if given the opportunity [69]. This flexibility is reflected in the magnitude of differences in brain gene expression patterns between queen and worker honey bees (thousands of genes [48]) and *M. genalis* (dozens of genes [40]). Previous research suggests that miRNAs increase their functional influence over evolutionary time [10,11,65,66,70,71]. Thus, emergent miRNAs are more likely to target genes with social function due to chance alone in species with increased social complexity and a larger set of caste-biased genes. Consistent with this explanation, regulatory relationships between miRNAs

and genes with caste-biased expression were not found among two other social insect species with reduced social complexity [72].

An additional explanation for these differences in the function of lineage-specific miRNAs concerns the role of miRNAs in gene regulatory networks. One of these roles is to stabilize regulatory relationships in the face of environmental variation, thus canalizing phenotypes during development [9,73–75]. This is likely to be more important in species with obligate eusociality, such as the honey bees and bumble bees for which caste determination is canalized, than in species like *M. genalis*, where plasticity of phenotypes related to eusociality are maintained in totipotent females.

Contrary to their effects on genes with socially-differentiated expression patterns, lineage-specific miRNAs showed no evidence for preferential targeting of genes under positive selection – either within or across species. In contrast, we find these emergent miRNAs are less likely than expected by chance to target genes under positive selection in the two halictid bees. A potential explanation for this pattern is that genes adaptively targeted by miRNAs tend to be under purifying selection to maintain the regulatory relationship between the miRNA and target, thus preventing gene mis-expression [76–78]. This selective constraint is likely to be most significant in the 3' UTR region, where miRNA binding sites are located.

A more likely explanation involves the hypothesized pattern of miRNA origins and assimilation, as proposed by Chen and Rajewsky [10]. This model suggests that new miRNAs are likely to have many targets throughout the genome due to chance. Most of these initial

472 miRNA-target regulatory relationships are likely to have slightly deleterious effects, and would
473 be quickly purged through purifying selection. These deleterious effects could be particularly
474 strong for target genes undergoing positive selection, because changes in the functional
475 regulation of these genes are likely to have significant fitness consequences. Also, genes under
476 positive selection are undergoing rapid evolution, and thus may be more likely to "escape"
477 control by errant miRNAs. Indeed, it is easier for mRNAs to lose miRNA target binding sites,
478 which typically require exact sequence matches, than to gain them [10]. Thus, emergent miRNAs
479 may not be expected to target adaptively or fast evolving genes, regardless of their role in social
480 evolution.
481
482 The evolution of eusociality depends on many different tissues and physiological
483 processes, and brain-specific expression patterns are not likely to be representative of the
484 complete role of individual miRNAs in social behavior. Some or all of the predicted miRNA-
485 gene relationships we identified may have evolved to support traits in other cell types or
486 processes unrelated to sociality. Additional sequencing of miRNA and mRNA across tissue-
487 types and stages of development in social and solitary species is necessary to provide a
488 comprehensive assessment of the role of emergent miRNAs in social traits. Nonetheless, the
489 brain is a major focus of research in social evolution because it is the primary source of
490 behavioral and neuroendocrine output. Our results thus provide a good starting place for
491 evaluating the role of miRNAs in lineage-specific processes in the evolution of social behavior.
492
493 Our analyses reveal important differences in patterns of miRNA evolution between bees
494 and other species. For example, expansion in miRNA repertoire is associated with the evolution

of animal complexity in a wide range of species [9,12,13]. The evolution of eusociality from a solitary ancestor is associated with increases in phenotypic complexity, and considered to be one of the major transitions in evolution [79]. We therefore hypothesized that evolutionary increases in social complexity would be associated with expansions in the number of miRNAs found within bee genomes. To the contrary, we find that most bees have a single copy of previously identified miRNAs in their genomes. This is consistent with results of comparative genome scans across several ant species [3]. A recent study of miRNA diversity in insects found that morphological innovations such as holometabolous development was accompanied by the acquisition of only three miRNA families [15]. This suggests that insect evolution is not as reliant on major expansions of miRNA families as other taxonomic groups.

Additionally, our characterization of lineage-specific miRNAs expressed in the brain of each species reveals that genome structure is not as influential in regulating bee miRNA evolution as has been shown for human miRNAs. Novel human miRNAs tend to arise within ancient genes that have multiple functions and broad expression patterns [65]. It is hypothesized that this increases the expression repertoire of emergent miRNAs, and thus facilitates persistence in the population [64,65]. Only in one species (*N. melanderi*) were lineage-specific miRNAs more likely to be localized intragenically than previously identified miRNAs, while lineage-specific miRNAs did not differ from previously identified miRNAs in their genomic locations in the other five species. This suggests emergence patterns for new miRNAs are unique to each lineage in bees. We also do not find a consistent pattern between young, emerging miRNAs and host gene age. There was no significant difference in the age of genes that serve as hosts for established versus lineage-specific miRNAs across species. This is despite the fact that a similar

518 proportion of bee miRNAs are located within introns (31-43%; Table 1), compared to in
519 vertebrates (36-65%) [8]. However, the fact that 73-88% of miRNAs localized to genes are
520 encoded on the sense strand suggests that they would benefit from host transcription, as is
521 observed in vertebrates [8]. Additional research with insects will be necessary to identify general
522 patterns of miRNA evolution in relationship to genome structure.
523
524 Our study identifies patterns of miRNA evolution in a set of closely related bees that vary
525 in social organization. Our results highlight important similarities and differences in the
526 emergence patterns and functions of mammalian and insect genomes. We find evidence that
527 emergent miRNAs function in lineage-specific patterns of social evolution, perhaps through co-
528 evolution of novel miRNAs and species-specific targets. We do not see an overall increase in the
529 number of miRNAs in the genome or expressed in the brains of species with more complex
530 eusociality. However, we do find evidence that the role of miRNAs in social evolution may
531 strengthen with increasing social complexity, perhaps due to an increased need for canalization
532 of caste determination or due to chance, as a function of an increased number of genes with
533 caste-biased expression. Empirical tests of miRNA function across additional species with
534 variable social organization will further improve our understanding of how gene regulatory
535 evolution gives rise to eusociality.
536
537 **REFERENCES**

538 1. Sumner S, Bell E, Taylor D. 2018 A molecular concept of caste in insect societies. *Curr.*
539 *Opin. Insect Sci.* **25**, 42–50. (doi:10.1016/j.cois.2017.11.010)
540 2. Whitfield CW, Band MR, Bonaldo MF, Kumar CG, Liu L, Pardinas JR, Robertson HM,

Soares MB, Robinson GE. 2002 Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Genome Res.* **12**, 555–566.

3. Simola DF *et al.* 2013 Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Res.* **23**, 1235–1247. (doi:10.1101/gr.155408.113)

4. Kapheim KM *et al.* 2015 Genomic signatures of evolutionary transitions from solitary to group living. *Science* **348**, 24–32. (doi:10.1126/science.aaa4788)

5. Warner MR, Qiu L, Holmes MJ, Mikheyev AS, Linksvayer TA. 2019 Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. *Nat. Commun.* **10**, 2651. (doi:10.1038/s41467-019-10546-w)

6. Bartel DP. 2018 Metazoan microRNAs. *Cell* **173**, 20–51. (doi:10.1016/j.cell.2018.03.006)

7. Friedman RC, Farh KK-H, Burge CB, Bartel DP. 2009 Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105. (doi:10.1101/gr.082701.108)

8. Meunier J, Lemoine F, Soumillon M, Liechti A, Weier M, Guschanski K, Hu H, Khaitovich P, Kaessmann H. 2013 Birth and expression evolution of mammalian microRNA genes. *Genome Res.* **23**, 34–45. (doi:10.1101/gr.140269.112)

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

10. Chen K, Rajewsky N. 2007 The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* **8**, 93–103. (doi:10.1038/nrg1990)

11. Lu J, Shen Y, Wu Q, Kumar S, He B, Shi S, Carthew RW, Wang SM, Wu C-I. 2008 The birth and death of microRNA genes in *Drosophila*. *Nat. Genet.* **40**, 351–355. (doi:10.1038/ng.73)

12. Christodoulou F *et al.* 2010 Ancient animal microRNAs and the evolution of tissue identity. *Nature* **463**, 1084–1088. (doi:10.1038/nature08744)

13. Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ. 2008 MicroRNAs and the advent of vertebrate morphological complexity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2946–50. (doi:10.1073/pnas.0712259105)

14. Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ. 2009 The deep evolution of metazoan microRNAs. *Evol. Dev.* **11**, 50–68. (doi:10.1111/j.1525-142X.2008.00302.x)

15. Ylla G, Fromm B, Piulachs MD, Belles X. 2016 The microRNA toolkit of insects. *Sci. Rep.* **6**, 1–13. (doi:10.1038/srep37736)

16. Søvik E, Bloch G, Ben-Shahar Y. 2015 Function and evolution of microRNAs in eusocial Hymenoptera. *Front. Genet.* **6**, 1–11. (doi:10.3389/fgene.2015.00193)

17. Greenberg JK *et al.* 2012 Behavioral plasticity in honey bees is associated with differences in brain microRNA transcriptome. *Genes, Brain Behav.* **11**, 660–670. (doi:10.1111/j.1601-183X.2012.00782.x)

18. Liu F *et al.* 2012 Next-generation small RNA sequencing for microRNAs profiling in *Apis mellifera*: Comparison between nurses and foragers. *Insect Mol. Biol.* **21**, 297–303. (doi:10.1111/j.1365-2583.2012.01135.x)

19. Behura SK, Whitfield CW. 2010 Correlated expression patterns of microRNA genes with age-dependent behavioural changes in honeybee. *Insect Mol. Biol.* **19**, 431–439. (doi:10.1111/j.1365-2583.2010.01010.x)

20. Weaver D *et al.* 2007 Computational and transcriptional evidence for microRNAs in the honey bee genome. *Genome Biol.* **8**, R97.

587 21. Ashby R, Forêt S, Searle I, Maleszka R. 2016 MicroRNAs in honey bee caste
588 determination. *Sci. Rep.* **6**, 1–15. (doi:10.1038/srep18794)
589 22. Shi YY, Zheng HJ, Pan QZ, Wang ZL, Zeng ZJ. 2015 Differentially expressed
590 microRNAs between queen and worker larvae of the honey bee (*Apis mellifera*).
591 *Apidologie* **46**, 35–45. (doi:10.1007/s13592-014-0299-9)
592 23. Collins DH, Mohorianu I, Beckers M, Moulton V, Dalmay T, Bourke AFG. 2017
593 MicroRNAs associated with caste determination and differentiation in a primitively
594 eusocial insect. *Sci. Rep.* **7**, 1–9. (doi:10.1038/srep45674)
595 24. MacEdo LMF *et al.* 2016 MicroRNA signatures characterizing caste-independent ovarian
596 activity in queen and worker honeybees (*Apis mellifera* L.). *Insect Mol. Biol.* **25**, 216–226.
597 (doi:10.1111/imb.12214)
598 25. Nunes FMF, Ihle KE, Mutti NS, Simoes ZLP, Amdam G V. 2013 The gene *vitellogenin*
599 affects microRNA regulation in honey bee (*Apis mellifera*) fat body and brain. *J Exp Biol*
600 **216**, 3724–3732. (doi:10.1242/jeb.089243)
601 26. Branstetter MG, Danforth BN, Pitts JP, Faircloth BC, Ward PS, Buffington ML, Gates
602 MW, Kula RR, Brady SG. 2017 Phylogenomic insights into the evolution of stinging
603 wasps and the origins of ants and bees. *Curr. Biol.* **27**, 1019–1025.
604 (doi:10.1016/j.cub.2017.03.027)
605 27. Friedländer MR, MacKowiak SD, Li N, Chen W, Rajewsky N. 2012 MiRDeep2
606 accurately identifies known and hundreds of novel microRNA genes in seven animal
607 clades. *Nucleic Acids Res.* **40**, 37–52. (doi:10.1093/nar/gkr688)
608 28. Quinlan AR, Hall IM. 2010 BEDTools: a flexible suite of utilities for comparing genomic
609 features. *Bioinformatics* **26**, 841–842. (doi:10.1093/bioinformatics/btq033)

29. Smith AF., Hubley R, Green P. 2013 RepeatMasker. See http://www.repeatmasker.org/.

30. Sadd BM *et al.* 2015 The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol.* **16**, 76. (doi:10.1186/s13059-015-0623-3)

31. Elsik CG *et al.* 2014 Finding the missing honey bee genes: Lessons learned from a genome upgrade. *BMC Genomics* **15**, 1–29. (doi:10.1186/1471-2164-15-86)

32. Kapheim KM *et al.* 2019 Draft genome assembly and population genetics of an agricultural pollinator, the solitary alkali bee (Halictidae: *Nomia melanderi*). *G3* **9**, 625–634. (doi:10.1534/g3.118.200865)

33. Kapheim KM *et al.* Developmental plasticity shapes the genomic signatures of eusociality. *unpublished*

34. Grün D, Wang Y-L, Langenberger D, Gunsalus KC, Rajewsky N. 2005 microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Comput. Biol.* **1**, e13. (doi:10.1371/journal.pcbi.0010013)

35. Enright AJ, John B, Gaul U, Tuschl T, Sander C. 2004 MicroRNA targets in *Drosophila*. *Genome Biol.* **5**, R1. (doi:10.1186/gb-2003-5-1-r1)

36. Krüger J, Rehmsmeier M. 2006 RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* **34**, W451-4. (doi:10.1093/nar/gkl243)

37. Zdobnov EM, Tegenfeldt F, Kuznetsov D, Waterhouse RM, Simão FA, Ioannidis P, Seppey M, Loetscher A, Kriventseva E V. 2017 OrthoDB v9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res.* **45**, D744–D749. (doi:10.1093/nar/gkw1119)

38. Gentleman R, Falcon S. 2013 *Package 'GOstats'*. 2.26.0.

39. Team RC. 2016 R: A language and environment for statistical computing.

633 40. Jones BM, Kingwell CJ, Wcislo WT, Robinson GE. 2017 Caste-biased gene expression in
634 a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of
635 eusociality. *Proc. R. Soc. B Biol. Sci.* **284**. (doi:10.1098/rspb.2016.2228)
636 41. Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina
637 sequence data. *Bioinformatics* **30**. (doi:10.1093/bioinformatics/btu170)
638 42. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S. 2013 STAR: ultrafast
639 universal RNA-seq aligner. *Bioinforma Oxf Engl* **29**. (doi:10.1093/bioinformatics/bts635)
640 43. Liao Y, Smyth GK, Shi W. 2014 featureCounts: an efficient general purpose program for
641 assigning sequence reads to genomic features. *Bioinformatics* **30**.
642 (doi:10.1093/bioinformatics/btt656)
643 44. Robinson MD, McCarthy DJ, Smyth GK. 2010 edgeR: a Bioconductor package for
644 differential expression analysis of digital gene expression data. *Bioinformatics* **26**.
645 (doi:10.1093/bioinformatics/btp616)
646 45. Cui X, Lu Z, Wang S, Jing-Yan Wang J, Gao X. 2016 CMsearch: simultaneous
647 exploration of protein sequence space and structure space improves not only protein
648 homology detection but also protein structure prediction. *Bioinformatics* **32**, i332–i340.
649 (doi:10.1093/bioinformatics/btw271)
650 46. Paradis E, Claude J, Strimmer K. 2004 APE: analyses of phylogenetics and evolution in R
651 language. *Bioinformatics* **20**, 289–290.
652 47. Cardinal S, Danforth BNBN. 2011 The antiquity and evolutionary history of social
653 behavior in bees. *PLoS One* **6**, e21086. (doi:10.1371/journal.pone.0021086)
654 48. Grozinger CM, Fan Y, Hoover SER, Winston ML. 2007 Genome-wide analysis reveals
655 differences in brain gene expression patterns associated with caste and reproductive status

in honey bees (*Apis mellifera*). *Mol. Ecol.* **16**, 4837–4848. (doi:doi:10.1111/j.1365-294X.2007.03545.x)

49. Whitfield CW, Cziko A-M-. M, Robinson GE. 2003 Gene expression profiles in the brain predict behavior in individual honey bees. *Science* **302**, 296–299. (doi:10.1126/science.1086807)

50. Alaux C, Le Conte Y, Adams HA, Rodriguez-Zas S, Grozinger CM, Sinha S, Robinson GE. 2009 Regulation of brain gene expression in honey bees by brood pheromone. *Genes, Brain Behav.* **8**, 309–319. (doi:10.1111/j.1601-183X.2009.00480.x)

51. Porath HT, Hazan E, Shpigler H, Cohen M, Band M, Ben-Shahar Y, Levanon EY, Eisenberg E, Bloch G. 2019 RNA editing is abundant and correlates with task performance in a social bumblebee. *Nat. Commun.* **10**, 1605. (doi:10.1038/s41467-019-09543-w)

52. Marshall H, Lonsdale ZN, Mallon EB. 2019 Methylation and gene expression differences between reproductive and sterile bumblebee workers. *Evol. Lett.* (doi:10.1002/evl3.129)

53. Harpur BA, Kent CF, Molodtsova D, Lebon JM, Alqarni AS, Owayss AA, Zayed A. 2014 Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc Natl Acad Sci U S A* **111**, 2614–2619. (doi:10.1073/pnas.1315506111)

54. Harpur BA, Dey A, Albert JR, Patel S, Hines HM, Hasselmann M, Packer L, Zayed A. 2017 Queens and workers contribute differently to adaptive evolution in bumble bees and honey bees. *Genome Biol. Evol.* **9**, 2395–2402. (doi:10.1093/gbe/evx182)

55. Woodard SH, Fischman BJ, Venkat A, Hudson ME, Varala K, Cameron SA, Clark AG, Robinson GE. 2011 Genes involved in convergent evolution of eusociality in bees. *Proc Natl Acad Sci U S A* **108**, 7472–7477. (doi:10.1073/pnas.1103457108)

679 56. Duncan EJ, Hyink O, Dearden PK. 2016 Notch signalling mediates reproductive
680 constraint in the adult worker honeybee. *Nat. Commun.* **7**, 1–10.
681 (doi:10.1038/ncomms12427)
682 57. Hartfelder K, Tiberio GJ, Lago DC, Dallacqua RP, Bitondi MMG. 2018 The ovary and its
683 genes—developmental processes underlying the establishment and function of a highly
684 divergent reproductive system in the female castes of the honey bee, *Apis mellifera*.
685 *Apidologie* **49**, 49–70. (doi:10.1007/s13592-017-0548-9)
686 58. Wang Y, Azevedo S V, Hartfelder K, Amdam G V. 2013 Insulin-like peptides (AmILP1
687 and AmILP2) differentially affect female caste development in the honey bee (*Apis*
688 *mellifera* L.). *J Exp Biol* **216**, 4347–4357. (doi:10.1242/jeb.085779)
689 59. Chandra V, Fetter-Pruneda I, Oxley PR, Ritger AL, McKenzie SK, Libbrecht R, Kronauer
690 DJC. 2018 Social regulation of insulin signaling and the evolution of eusociality in ants.
691 *Science* **361**, 398–402. (doi:10.1126/science.aar5723)
692 60. Ament SA, Corona M, Pollock HS, Robinson GE. 2008 Insulin signaling is involved in
693 the regulation of worker division of labor in honey bee colonies. *Proc. Natl. Acad. Sci. U.*
694 *S. A.* **105**, 4226–4231. (doi:10.1073/pnas.0800630105)
695 61. Barbash S, Shifman S, Soreq H. 2014 Global coevolution of human microRNAs and their
696 target genes. *Mol. Biol. Evol.* **31**, 1237–1247. (doi:10.1093/molbev/msu090)
697 62. Jasper WC, Linksvayer TA, Atallah J, Friedman D, Chiu JC, Johnson BR. 2015 Large-
698 scale coding sequence change underlies the evolution of postdevelopmental novelty in
699 honey bees. *Mol. Biol. Evol.* **32**, 334–346. (doi:10.1093/molbev/msu292)
700 63. Johnson B, Tsutsui N. 2011 Taxonomically restricted genes are associated with the
701 evolution of sociality in the honey bee. *BMC Genomics* **12**, 164.

64. França GS, Hinske LC, Galante PAFF, Vibranovski MD. 2017 Unveiling the impact of the genomic architecture on the evolution of vertebrate microRNAs. *Front. Genet.* **8**, 1–8. (doi:10.3389/fgene.2017.00034)

65. França GS, Vibranovski MD, Galante PAF. 2016 Host gene constraints and genomic context impact the expression and evolution of human microRNAs. *Nat. Commun.* **7**. (doi:10.1038/ncomms11438)

66. Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk RHA. 2006 Diversity of microRNAs in human and chimpanzee brain. *Nat. Genet.* **38**, 1375–1377. (doi:10.1038/ng1914)

67. Tang T, Kumar S, Shen Y, Lu J, Wu M-L, Shi S, Li W-H, Wu C-I. 2010 Adverse interactions between micro-RNAs and target genes from different species. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12935–40. (doi:10.1073/pnas.1007591107)

68. Brady SG, Sipes S, Pearson A, Danforth BN. 2006 Recent and simultaneous origins of eusociality in halictid bees. *Proc. R. Soc. B Biol. Sci.* **273**, 1643–1649. (doi:10.1098/rspb.2006.3496)

69. Smith AR, Kapheim KM, O'Donnell S, Wcislo WT. 2009 Social competition but not subfertility leads to a division of labour in the facultatively social sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Anim. Behav.* **78**, 1043–1050.

70. Patel VD, Capra JA. 2017 Ancient human miRNAs are more likely to have broad functions and disease associations than young miRNAs. *BMC Genomics* **18**, 672. (doi:10.1186/s12864-017-4073-z)

71. Roux J, Gonzàlez-Porta M, Robinson-Rechavi M. 2012 Comparative analysis of human and mouse expression data illuminates tissue-specific evolutionary patterns of miRNAs.

Nucleic Acids Res. **40**, 5890–5900. (doi:10.1093/nar/gks279)

72. Patalano S *et al.* 2015 Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. *Proc. Natl. Acad. Sci.* **112**, 13970–13975. (doi:10.1073/pnas.1515937112)

73. Wu C-I, Shen Y, Tang T. 2009 Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome Res.* **19**, 734–43. (doi:10.1101/gr.084640.108)

74. Hornstein E, Shomron N. 2006 Canalization of development by microRNAs. *Nat. Genet.* **38**, S20–S24. (doi:10.1038/ng1803)

75. Peterson KJ, Dietrich MR, McPeek MA. 2009 MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. *BioEssays* **31**, 736–747. (doi:10.1002/bies.200900033)

76. Chen K, Rajewsky N. 2006 Natural selection on human microRNA binding sites inferred from SNP data. *Nat. Genet.* **38**, 1452–1456. (doi:10.1038/ng1910)

77. Saunders MA, Liang H, Li W-H. 2007 Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3300–5. (doi:10.1073/pnas.0611347104)

78. Franchini P, Xiong P, Fruciano C, Meyer A. 2016 The role of microRNAs in the repeated parallel diversification of lineages of midas cichlid fish from Nicaragua. *Genome Biol. Evol.* **8**, 1543–1555. (doi:10.1093/gbe/evw097)

79. Maynard Smith J, Szathmáry E. 1995 *The major transitions in evolution*. Oxford, England: Oxford University Press.

DATA AVAILABILITY

Sequences are deposited at NCBI SRA as BioProject PRJNA559906. Code is available upon request.

ACKNOWLEDGEMENTS

This work was supported by the USDA National Institute of Food and Agriculture [2018-67014-27542 to K.M.K.]; the Utah Agricultural Experiment Station, Utah State University [Project 1297, journal paper number 9239 to K.M.K.]; the U.S.-Israel Binational Science Foundation [BSF 2012807 to G.B. and Y.B.S.]; and the Swiss National Science Foundation [PP00P3_170664 to R. M. W.]. G.B. thanks the Clark Way Harrison Visiting Professor in Arts and Sciences that supported his stay in Washington University in St. Louis. Sequencing was performed at the University of Illinois Roy J. Carver Biotechnology Center. We thank the University of Utah High Performance Computing Center for computational time and assistance. Illustrations were created by J. Johnson (LifeSciences Studios). We thank G. Robinson for helpful feedback on an earlier draft of this manuscript.

AUTHOR CONTRIBUTIONS

K.M.K. conceived of the study and designed the experiments. K.M.K., E.S., G.B., and Y.B-S. collected the data. K.M.K., B.M.J., E.S., and R.M.W. analyzed the data. K.M.K. wrote the initial draft of the manuscript. All authors edited and approved the article for publication.