bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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] Martin-Luther-Universität Halle-Wittenberg, Institut für Biologie, Hoher Weg 8, Halle (Saale), 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, 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

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

37 **ABSTRACT**

38 Evolutionary transitions to a social lifestyle in insects are associated with lineage-specific
39 changes in gene expression, but the key nodes that drive these regulatory changes are largely
40 unknown. We tested the hypothesis that changes in gene regulatory functions associated with
41 social evolution are facilitated by lineage-specific function of microRNAs (miRNAs). Genome
42 scans across 12 bee species showed that miRNA copy number is highly conserved and is not
43 associated with variation in social organization. However, deep sequencing of small RNAs of six
44 bee species that exhibit varying types of sociality revealed a substantial proportion (20-35%) of
45 detected miRNAs with lineage-specific expression in the brain, 24-72% of which did not have
46 any known homologs in other species. Lineage-specific miRNAs disproportionately target
47 lineage-specific genes, and have lower expression levels than more evolutionarily conserved
48 miRNAs. Consistent with our hypothesis, the predicted targets of lineage-specific miRNAs are
49 enriched for genes related to social behavior, such as caste-biased genes, in social species, but
50 they are either not enriched for or significantly depleted of genes under positive selection.
51 Together, these results suggest that novel miRNAs may contribute to lineage-specific patterns of
52 molecular evolution associated with the origins and elaborations of eusociality. Our analyses also
53 lend support to the earlier hypothesis that many new miRNAs are quickly purged by selection
54 due to slightly deleterious effects on mRNA targets, and suggest genome structure is not as
55 influential in regulating bee miRNA evolution as has been shown for mammalian miRNAs.

56

57 **Keywords:** Gene regulation; small, non-coding RNA; microRNA targets; eusociality; lineage-
58 specific

59

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

60 **INTRODUCTION**

61 Eusociality, the most complex form of social organization, has evolved several times in
62 insects from the order Hymenoptera. In its most basic form, this lifestyle involves reproductive
63 queens living with their worker daughters who forego direct reproduction of their own to
64 cooperatively defend the nest, care for their siblings, and forage for the colony (Wilson 1971).
65 Due to the complex nature of this lifestyle, the evolution of eusociality likely requires
66 modification of molecular pathways related to development, behavior, neurobiology, physiology,
67 and morphology (Robinson and Ben-Shahar 2002; Toth and Robinson 2007; Bloch and
68 Grozinger 2011; Sumner et al. 2018). The evolution of eusociality is thus expected to involve
69 both genetic changes as well as changes in the way the genome responds to the environment
70 (Robinson and Ben-Shahar 2002; Johnson and Linksvayer 2010). It is therefore unsurprising that
71 recent studies aimed at identifying the genomic signatures of eusocial evolution in insects have
72 found that social species share an increased capacity for gene regulation (Simola et al. 2013;
73 Kapheim et al. 2015). Evidence for this comes from signatures of rapid evolution of genes
74 involved in transcription and translation, gene family expansions of transcription factors, and
75 increasing potential for DNA methylation and transcription factor binding activity in conserved
76 genes. Interestingly, while these types of regulatory changes are common to independent origins
77 and elaborations of eusociality, the specific genes and regulatory elements involved are unique to
78 each lineage in which eusociality evolved (Simola et al. 2013; Kapheim et al. 2015; Warner et al.
79 2019). This suggests that lineage-specific processes are influential in generating new patterns of
80 gene regulation that contribute to social behavior.
81

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

82 Small, non-coding RNAs such as microRNAs (miRNAs) may be an important source of
83 regulatory novelty associated with the evolution of phenotypic complexity, including eusociality.
84 MiRNAs are short (~21-22 nt) noncoding RNAs that regulate protein-coding genes through post-
85 transcriptional binding to the 3' UTR region of messenger RNA (mRNA) transcripts, in most
86 cases preventing translation or causing mRNA degradation (Bartel 2018). Each miRNA can
87 target dozens to hundreds of mRNAs, and therefore miRNAs have enormous pleiotropic
88 potential to regulate gene networks (Filipowicz et al. 2008; Friedman et al. 2009; Bartel 2018).
89 Like mRNAs, the majority of miRNAs are generated via Pol II transcription, and therefore, are
90 spatially- and temporally-specific in their expression patterns. Thus, complex changes in gene
91 regulation can be achieved with relatively minor changes in miRNA expression. This can result
92 in major phenotypic shifts or fine-tuning of phenotypic optimization (Bartel 2018). Novel
93 miRNAs can originate in a variety of genomic features, including exons and introns of protein-
94 coding and non-coding RNA genes, transposable elements, pseudogenes, or intergenic regions,
95 and thus emerge and disappear over relatively rapid timescales (Chen and Rajewsky 2007; Lu et
96 al. 2008; Berezikov 2011; Zhu et al. 2012; Meunier et al. 2013). It is thus not surprising that
97 expansion of the miRNA repertoire is associated with the evolution of morphological complexity
98 across the tree of life (Grimson et al. 2008; Heimberg et al. 2008; Wheeler et al. 2009;
99 Christodoulou et al. 2010; Berezikov 2011).
100
101 There is accumulating evidence for a role of miRNAs in regulating the social lives of
102 insects. MiRNAs have been identified in the genomes of most major insect groups (Asgari 2013;
103 Quah et al. 2015; Ylla et al. 2016), including several ant and bee species (Weaver et al. 2007;
104 Bonasio et al. 2010; Kocher et al. 2013; Patalano et al. 2015; Sadd et al. 2015). Bioinformatic

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

105 scans of insect genomes have identified candidate miRNAs present in the genomes of social, but
106 not solitary, insects (Søvik et al. 2015). While most miRNAs seem to be conserved in major
107 lineages of insects (Søvik et al. 2015; Ylla et al. 2016), expression levels have been found to vary
108 across individuals performing different social functions, such as between workers performing
109 different tasks in honey bees (*Apis mellifera*) (Behura and Whitfield 2010; Greenberg et al. 2012;
110 Liu et al. 2012). MiRNAs may also play a role in caste determination, as queen- and worker-
111 destined larvae express different sets of miRNAs throughout development in honey bees
112 (Weaver et al. 2007; Shi et al. 2015; Ashby et al. 2016) and bumble bees (*Bombus terrestris*)
113 (Collins et al. 2017). Additionally, miRNAs play a role in regulating some of the physiological
114 correlates of social behavior in honey bees, including activation of ovaries in queens and workers
115 (MacEdo et al. 2016) and response to the reproductive protein *vitellogenin* (Nunes et al. 2013).
116 Together, these studies suggest that miRNAs could play a role in the evolution of eusociality
117 through their effects on gene regulatory networks that are involved in traits important for social
118 behavior. A rigorous test of this hypothesis requires comparisons of the presence, expression,
119 and function of miRNAs across related species that vary in social organization. However, none
120 of the previous studies of insect miRNAs have included solitary species that are closely related
121 to eusocial insects, and thus assumed to be more representative of the ancestors from which
122 sociality evolved.
123
124 Here we present a comprehensive comparative analysis of miRNAs across bee species
125 that vary in social organization. We first looked for miRNA repertoire expansions associated
126 with eusociality by scanning the genomes of 12 bee species for known miRNAs, and statistically
127 evaluating copy number of each miRNA type with regard to differences in sociality in a

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

128 phylogenetic model. We then described and compared miRNAs expressed in the brains of six
129 bee species from three families that include repeated origins of eusociality. In our analysis, we
130 identified shared and lineage-specific miRNAs, their evolutionary histories, and their predicted
131 gene targets. We then tested the hypothesis that changes in gene regulatory function associated
132 with social evolution are facilitated by lineage-specific miRNA regulatory function. We tested
133 three predictions of this hypothesis: (1) miRNAs that play a role in social behavior should target
134 different genes in solitary and social species. (2) If lineage-specific miRNAs are assimilated into
135 ancestral gene networks, then their predicted target genes should be ancient and conserved. (3) If
136 lineage-specific miRNAs play a role in social evolution, then their predicted targets should be
137 enriched for genes that function in social behavior (e.g., caste-biased expression) or genes that
138 are under selection in social species.
139

140 ## MATERIALS AND METHODS

141 ### Sample Acquisition

142 We used adult females from six bee species for our study. *Megalopta genalis* samples
143 were collected on Barro Colorado Island, Panama in 2015 and exported to the U.S.A. under
144 permit SEX/A-37-15. *Nomia melanderi* samples were collected in Touchet, WA, U.S.A. with
145 permission from private land owners. *Megachile rotundata* samples were collected from Logan,
146 UT, U.S.A. on the Utah State University campus. *Bombus impatiens* samples were collected
147 from a commercial colony purchased from BioBest. *Bombus terrestris* samples were collected
148 from colonies obtained from Pollination Services Yad-Mordechai, Kibbutz Yad-Mordechai,
149 Israel. *Apis mellifera* samples were collected from typical field hives in Urbana-Champaign, IL

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

150 or the Tyson Research Field Station, MO, U.S.A. All samples were flash-frozen in liquid
151 nitrogen upon collection and stored at -80 °C until dissection.
152

153 **RNA Isolation and Sequencing**

154 Head capsules from the *B. impatiens*, *M. genalis*, and *N. melanderi* samples were
155 dissected after incubation in RNALater ICE (Ambion) to remove the entire brain. We used the
156 mirVana miRNA Isolation kit with phenol (Ambion) to isolate total RNA from individual brains.
157 Total RNA was sent to the University of Illinois Roy J. Carver Biotechnology Center for library
158 preparation and sequencing. Libraries were prepared with the Illumina TruSeq Small RNA
159 Sample Preparation kit. Libraries were pooled, quantitated by qPCR, and sequenced on one lane
160 for 51 cycles on a HiSeq 2500, using TruSeq SBS sequencing kit version 2. Sequencing yielded
161 a mean of 19,524,877 (± 2,545,208 s.d.) reads per sample.
162

163 Whole brains of *A. mellifera*, *B. terrestris*, and *M. rotundata* were dissected from frozen
164 heads. Total RNA from individual brains were isolated by using the TRIzol reagent (Thermo
165 Fisher Scientific). All subsequent small-RNA sequencing steps were performed by the Genome
166 Technologies Access Center at Washington University, using their Illumina TruSeq pipeline.
167 Total RNA samples were size fractionated and multiplexed. Single-end small RNA libraries
168 were prepared using the SMARTer kit (Clontech). Up to 12 barcoded libraries from a single
169 species were run on a single Illumina HiSeq 2500 lane.
170

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

171 **miRNA Discovery and Quantification**

172 We used miRDeep2 (Friedländer et al. 2012) to identify and quantify miRNAs expressed
173 in the brains of each species. We used a three-step process of miRNA detection in order to
174 identify homologous miRNAs between species. For the first step, we generated a set of mature
175 miRNA sequences previously described in other insect species (Table S1). Reads for each
176 sample were quality filtered (minimum length 18, removal of reads with non-standard bases),
177 adapter-trimmed, and aligned to the species genome (Table S2) with the mapper.pl script.
178 Approximately 60-84% of reads successfully mapped.

179

180 We then identified known and novel miRNAs in each sample with the miRDeep2.pl
181 script, using our curated set of insect miRNAs (Table S1) as known mature sequences. We
182 followed this with quantification of the miRNAs using the quantifier.pl script. This generated a
183 set of known and novel miRNAs in each sample, along with quantified expression information
184 for each. We then filtered the novel miRNAs in each species according to the following criteria:
185 no rRNA/tRNA similarities, minimum of five reads each on the mature and star strands of the
186 hairpin sequence, and a significant randfold p-value ($p < 0.05$). Randfold describes the RNA
187 secondary structure of potential pre-miRs (Friedländer et al. 2012).

188

189 We used these filtered miRNAs in a second run of detection and quantification. We
190 added the mature sequences of the novel miRNAs from each species to our set of known
191 miRNAs, and repeated the pipeline above. This allowed detection of homologous miRNAs
192 (based on matching seed sequences) that are not represented in miRBase across our species. We
193 applied the same set of filtering criteria as for our first run.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

194

195 Some of the novel miRNAs may exist in the genomes of other bees, even if they are not
196 expressed. We used blastn (-perc_identity 50 -evalue 1e-5) to search for homologous precursor
197 miR (pre-miR) sequences in 12 bee genomes (Table S2) for each of the novel miRNAs without a
198 matching seed sequence.

199

200 **miRNA Localization**

201 We characterized the location of each known and novel miRNA in its respective genome
202 assembly in relation to genes and transposable elements. We used bedtools intersect (Quinlan
203 and Hall 2010) to find overlap with predicted gene models (Table S3), and repetitive element
204 repeatmasker (Smit et al. 2013) annotations from previously established repeat libraries
205 (Kapheim et al., unpublished.; Elsik et al. 2014; Kapheim et al. 2015; Sadd et al. 2015; Kapheim
206 et al. 2019).

207

208 **Target Prediction**

209 We used computational methods to predict targets of each miRNA in each species. We
210 used bedtools flank and getfasta (Quinlan and Hall 2010) to extract a 500 bp region downstream
211 from each gene model, following previous studies (Ashby et al. 2016) and an average 3' UTR
212 region of 442 nt in *Drosophila melanogaster* (Grün et al. 2005). We used these as potential target
213 sites in miRanda (Enright et al. 2004) and RNAhybrid (Krüger and Rehmsmeier 2006) target
214 prediction analyses. miRanda v3.3 was run with a minimum energy threshold of -20, a minimum
215 score of 140, and strict alignment in the seed region (parameters -en -20 -sc 140 -strict). We also
216 used RNAhybrid v2.1.2 with a minimum free energy threshold of -20 and the fly 3' UTR set was

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

217 used to estimate xi and theta values (These are the position and shape parameters of the value
218 distribution from which p-values are calculated). We kept only miRNA-target gene pairs that
219 were predicted by both programs with $p < 0.01$.
220

221 **Target Age and Functional Enrichment**

222 We explored the taxonomic ages and putative functions of predicted target genes. Gene
223 ages were determined using orthogroups from OrthoDB v9 (Zdobnov et al. 2017), which
224 includes *A. mellifera*, *B. impatiens*, *B. terrestris*, and *M. rotundata*. OrthoDB delineates
225 orthologs by clustering gene best reciprocal hits (BRHs) between all species pairs, first with
226 triangulation of BRHs and then addition of in-paralogous groups and genes to build clusters of
227 orthologs. Gene sets of *M. genalis* and *N. melanderi* were mapped to Metazoa-level (330
228 species) orthogroups following the same procedure as for BRH clustering allowing genes to join
229 existing orthogroups when all BRH criteria are met. The age of each gene from the six bee
230 species was inferred from the taxonomic breadth of all species in each orthogroup, from
231 Vertebrata (at least one of 172 vertebrates), to Metazoa (at least one of 25 non-arthropod and
232 non-vertebrate metazoans), to Arthropoda (at least one of 17 non-insect arthropods), to Insecta
233 (at least one of 16 non-holometabolous insects), to Holometabola (at least one of 68 non-
234 hymenopteran holometabolous insects), to Hymenoptera (at least one of 7 non-Aculeata
235 hymenopterans), to Aculeata (at least one of 13 non-Apoidea Aculeata), to Apoidea (at least one
236 of 11 other Apoidea), and finally, genes without identifiable orthologs were labeled 'Unique'.
237

238 Gene Ontology (GO) terms for each species were derived from a previous study
239 (Kapheim et al. 2015), with the exception of *B. impatiens*, for which GO terms were assigned

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

240 based on reciprocal blastp (evalue < $1e^{-5}$) between two sets of gene models (OGS v1.2 and OGS
241 v1.0). Functional enrichment was performed with the GOstats package (Gentleman and Falcon
242 2013) in R (R Core Team 2016). We included all terms enriched at a value of unadjusted $p < 0.1$
243 to allow for more inclusivity in our cross-species comparisons.
244

245 **Enrichment tests of lineage-specific miRNA targets with previous studies**

246 For each species, brain or head gene expression datasets related to socially relevant
247 phenotypes (e.g., caste) and genes under selection were compared against targets of lineage-
248 specific miRNAs. Differential expression lists for *A. mellifera* included those from Grozinger et
249 al. (2007) (queen vs. worker, reproductive worker vs. sterile worker, queen vs. reproductive
250 worker), Whitfield et al. (2003) (nurse vs. forager), Alaux et al. (2009) (nurse vs. forager), and
251 Wheeler et al. (2013) (*vitellogenin* RNAi vs. control). For *B. terrestris,* differential expression
252 lists included three pairwise comparisons from Harrison et al. (2015) (queen vs. worker,
253 reproductive worker vs. sterile worker, queen vs. reproductive worker), as well as a comparison
254 between reproductive and sterile workers from Marshall et al. (2019). (The Harrison et al. dataset
255 was from whole body including brain, rather than just head or brain.) We also included lists of
256 differential expression between *B. terrestris* nurse and forager workers (Porath et al. 2019). For
257 *M. genalis* caste data, RNAseq reads from Jones et al. (2017) (NCBI PRJNA331103) were
258 trimmed using Trimmomatic (v. 0.36) and aligned to an unpublished genome assembly of *M.*
259 *genalis* (NCBI PRJNA494872) using STAR (v. 2.5.3). Reads were mapped to gene features
260 using the featureCounts function of the Subread package (v. 1.5.2). Remaining differential
261 expression analysis followed the methods of Jones et al. (2017) using edgeR (Robinson et al.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

262 2010) with tagwise dispersion estimates and FDR correction. The complete list of included
263 studies and gene lists are in Table S4.
264
265 In addition to gene expression comparisons, multiple datasets identifying genes under
266 selection in bee species or across multiple social lineages of bees were tested for enrichment of
267 lineage-specific miRNA targets. Species-specific selection datasets were used for *A. mellifera*
268 (Harpur et al. 2014), *B. impatiens* (Harpur et al. 2017), *M. genalis* (Kapheim et al. unpublished),
269 and *N. melanderi* (Kapheim et al. 2019). In addition, genes under selection in social relative to
270 solitary lineages identified by Woodard et al. (2011) and Kapheim et al. (2015) were tested
271 against each species' set of lineage-specific miRNA targets. The complete list of included studies
272 and gene lists are in Table S4.
273
274 In cases when we needed to identify orthologous genes across species, we used reciprocal
275 blastp (evalue < $10e^{-5}$). Only genes with putative orthologs were included in the final tested sets
276 of genes. Hypergeometric tests (using phyper in R) were used to test for significance of over- or
277 under-enrichment between each pair of lists. The representation factor (RF) given represents the
278 degree of overlap relative to random expectation (RF=1).
279
280 **miRNA Diversification**
281 We explored the diversification of miRNAs that have been previously implicated in
282 social behavior (miR-13b, miR-276, miR-6001-3p) or which are expressed in social bees, but not
283 solitary bees (miR-305). We performed multiple sequence alignment with the web version of

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

284 Clustal Omega with default settings (Sievers et al. 2011), and generated a Neighbour-joining
285 phylogenetic tree in Newick format.
286
287 We also performed genome scans for small RNAs across 12 bee genomes (Table S2)
288 using covariance models implemented with Infernal cmsearch using the gathering threshold for
289 inclusion (--cut_ga) (Cui et al. 2016) to find all Rfam accessions in each bee genome. We used
290 bedtools intersect to identify overlap between small RNAs identified through cmsearch and gene
291 models. We then used Spearman rank regressions to test for significant associations between
292 miRNA copy number and social biology. We categorized each species as either solitary,
293 facultative basic eusociality, obligate basic eusociality, or obligate complex eusociality following
294 Kapheim et al. (2015). We used the ape package (Paradis et al. 2004) in R (R Core Team 2016)
295 to calculate phylogenetic independent contrasts for both social organization and miRNA copy
296 number, cor.test to implement the Spearman's rank correlation, and p.adjust with the Benjamini-
297 Hochberg method (method = "BH") to correct for multiple comparisons.
298

299 RESULTS

300 Low levels of miRNA copy number variation among bee genomes

301 Our genome scans revealed very little variation in copy number of most miRNAs among
302 bee genomes. Of the 50 miRNA Rfam accessions, half had the same number of copies in all 12
303 bee genomes (1 or 2 copies) (Table S5). The mean copy number across all miRNAs in all bee
304 genomes was 1.19 ± 0.74. Seven of the Rfam miRNAs were detected in a single bee species, but
305 mostly at low copy numbers (1-3). One exception was miR-1122, for which we found 70 copies

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

306 in *M. genalis*, but no copies in any of the other species. We did not find any significant
307 associations between miRNA copy number and social organization (Table S5).
308

309 **Expressed miRNA diversity in bee brains**

310 We identified 97-245 known and novel miRNAs expressed in the brains of each of our
311 six species (Table S6). The majority of these were located in intergenic regions or in introns
312 (Table 1). Each species had at least one miRNA that originated from exons of protein-coding
313 genes and repetitive DNA (Table 1). Most of the overlap between miRNA precursors and
314 repetitive DNA corresponded to uncharacterized repeat elements, with very few overlaps with
315 well characterized transposons or retrotransposons (Table 1).
316

317 Most of the detected miRNAs in each species had known homologs in at least one other
318 species. However, each species had a substantial proportion (20-35%) of detected miRNAs with
319 lineage-specific expression in the brain (Table 1; Fig. 1A), 24-72% of which did not have any
320 known homologs in other species (Table 1). We defined lineage-specific miRNAs as those
321 miRNAs with lineage-specific expression and for which no seed match with a known mature
322 miRNA was identified (columns 6 and 7 in Table 1), because these show the most evidence of
323 being real miRNAs that are unique to a particular species. (Sequence similarity of pre-miRs in
324 the genome of other bee species is not sufficient evidence that a mature miRNA is transcribed.)
325 Lineage-specific miRNAs had significantly lower expression levels in each species (t-tests: *A.*
326 *mellifera*, $p = 3.81e^{-05}$, *B. impatiens*, $p = 0.003$, *B. terrestris*, $p = 0.006$, *M. genalis*, $p = 0.0003$,
327 *M. rotundata*, $p = 8.00e\text{-}^{05}$, *N. melanderi*, $p = 0.02$).
328

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

329 **Table 1.** Localization of miRNAs in the genomes of six bee species. Numbers not in parentheses
330 represent features on the same strand as the pre-miR. Numbers in parentheses indicate strand
331 mismatch. Some pre-miRs overlapped with one or more genes on both the same and opposite
332 strands, and are thus counted twice (*A. mellifera* and *M. genalis* – 1, *B. impatiens* – 5, *B.*
333 *terrestris* – 4, *N. melanderi* – 3). Seed match – Mature miR had a seed match with a known miR;
334 Pre-miR – Successful blastn hit to the pre-miR sequence in at least one other bee genome;
335 Unique – No homolog was found in other species (seed match to mature or blastn hit to pre-
336 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

337

338 Lineage-specific miRNAs were localized both within genes and intergenically. The
339 proportion of lineage-specific miRNAs that were intra- or intergenic was similar to miRNAs
340 with homologs for every species except *N. melanderi*, for which a disproportionate number of
341 lineage-specific miRNAs were intragenic ($\chi^2 = 4.78$, $p = 0.03$). Genes that serve as hosts for
342 intragenic lineage-specific miRNAs were not significantly older than would be expected by
343 chance (i.e., belong to orthogroups shared with vertebrates) in any species (hypergeometric tests:
344 $p = 0.14$-0.76). Across all species, genes that serve as hosts for intragenic lineage-specific
345 miRNAs were not significantly older than genes that host miRNAs with known homologs (χ^2
346 tests: $p = 0.05$-0.89).

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

347

348 Of the miRNAs with mature miRNA homologs, most were expressed in all six species,
349 and we detected few family-specific expression patterns of miRNAs (Table S6). For example,
350 miR-3049 and miR-3786 were only detected in the bees from the family Apidae (*A. mellifera*, *B.*
351 *impatiens*, *B. terrestris*). miR-3049 is predicted to target a neurotransmitter-gated ion-channel in
352 all three species (OG EOG091G0R20), and a histone acetyltransferase (OG EOG091G00D2), a
353 zinc-finger protein (OG EOG091G0N0A), a leucine-rich repeat (OG EOG091G01ZI), and a
354 sodium-potassium-calcium exchanger (OG EOG091G0M5C) in both *Bombus* species. The two
355 *Bombus* species shared a cytochrome P450 (OG EOG091G06KN) as a predicted target of miR-
356 3786, but did not share predicted targets with *A. mellifera*. We identified a miRNA (nmel-
357 scaffold2759_cov63_18669) without a known seed match that was only expressed in the two
358 halictid bees (*N. melanderi*, *M. genalis*) and is predicted to target kinase associated proteins in
359 both species.

360

361 We identified one miRNA (miR-305) that was expressed in the brains of each of the
362 social, but not the solitary, species. Although we did not detect expression of miR-305 in the two
363 solitary species, *M. rotundata* and *N. melanderi*, genome scans of each species against the Rfam
364 database suggested all bee species have one copy of this miRNA (Table S5). Predicted targets of
365 miR-305 differed across species. *Oxysterol* (OG EOG091G0FV2) was a common target among
366 the (social) Apidae bees, but was not among the targets for *M. genalis*. However,
367 *arylformamidase* (OG EOG091G0KT8), which is also involved in lipid metabolism and
368 transport, was a predicted target in *M. genalis*. *Synaptobrevin* (OG EOG091G0MPE), which is

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

369 involved in synaptic plasticity and neurotransmitter release, was a predicted target of miR-305 in
370 *B. impatiens*.
371

372 **miRNAs associated with honey bee social behavior are conserved across bee species**

373 As existing gene networks become co-opted for social evolution, we should expect novel
374 regulatory connections between existing miRNAs and protein-coding genes. If social evolution
375 involves miRNA-mediated changes in gene regulatory networks, then miRNAs with functions in
376 social behavior should target different genes in social and solitary species. We tested this
377 prediction by comparing the diversification and targets of three miRNAs that have been
378 repeatedly associated with social behavior in bees. miR-6001-3p is upregulated in queen-
379 destined larvae, compared to worker-destined larvae, in both *A. mellifera* (Shi et al. 2015) and *B.*
380 *terrestris* (Collins et al. 2017). However, we only detected brain expression of this miRNA in *M.*
381 *rotundata*, despite the fact that the honey bee miR-6001-3p was included in our known set of
382 miRNAs. Because this miRNA does not have an accession in Rfam, we were unable to verify
383 presence or copy number of this miRNA across the bee species. Additional targeted sequencing
384 is needed to determine the role of this miRNA in adult bee brains.
385

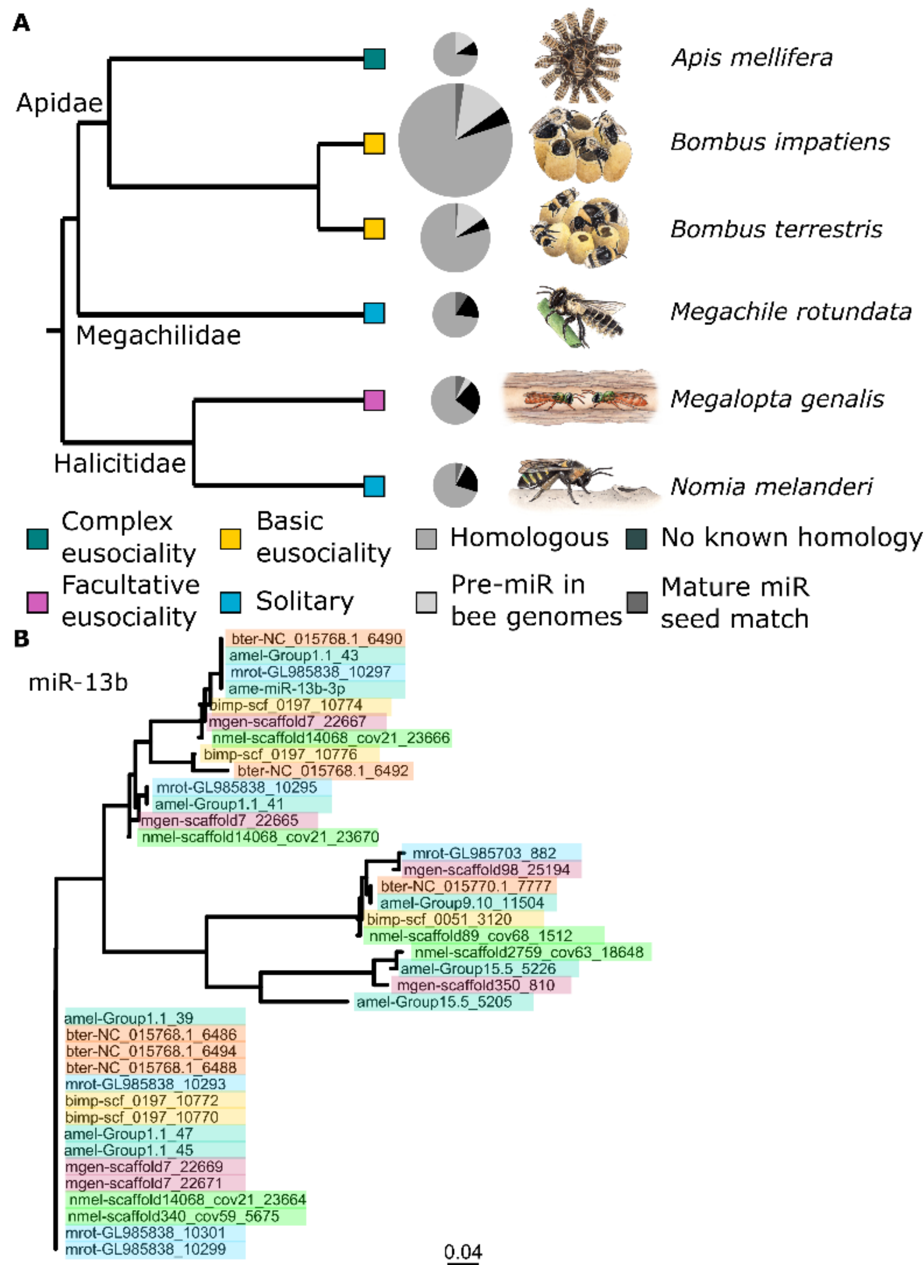
386 MiR-276 is upregulated in honey bee worker ovaries (MacEdo et al. 2016) and in queen-
387 destined larvae, but it is unknown if or how it functions in the brain (Shi et al. 2015; Ashby et al.
388 2016). Our genome scans revealed that each bee species has a single copy of this miRNA, and
389 we detected expression in the brains of each of our six focal species. Multiple alignment of the
390 expressed mature miRNA revealed almost no sequence variation across the species. There were
391 few shared gene targets across the six species, but the functional enrichment of predicted miR-

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

392 276 targets was similar among species. Target genes in each species were enriched for cell
393 signaling, though the specific pathways differed. *A. mellifera*, *B. impatiens*, and *M. rotundata*
394 targets were enriched for insulin receptor binding, and the two halictids (*Megalopta genalis* and
395 *N. melanderi*) were enriched for G-protein coupled receptor signaling. All species except *N.*
396 *melanderi* shared enrichment for functions related to protein translation, including translation,
397 protein-containing complex binding, structural constituent of ribosome, peptide biosynthetic
398 process, peptide metabolic process, and amide biosynthetic process.
399

400 MiR-13b is differentially expressed across honey bee castes in different life stages and
401 tissues. It is upregulated in worker ovaries compared to queen ovaries (MacEdo et al. 2016),
402 downregulated in queen-destined larvae compared to worker-destined larvae (Shi et al. 2015;
403 Ashby et al. 2016), and downregulated in workers performing nursing tasks compared to
404 foragers (Liu et al. 2012). This miRNA has expanded and diversified among the bees, with three
405 major clades containing at least one copy from each species (Fig. 1B). As with miR-276, there
406 was very little overlap in predicted targets among species. Just one gene, the *beta subunit of*
407 *nuclear transcription factor Y*, is predicted to be a target of miR-13b in each of the social
408 species. However, there were more similarities across species at the functional level (Table S7).
409 All species had enrichment for transcription factor activity, and all species except the two
410 halictid bees had significant enrichment for genes involved in ecdysone-mediated signaling.
411 Within the family Halictidae, the predicted targets of miR-13b for *M. genalis* were enriched for
412 genes related to neurotransmitter-gated ion channel activity, while *N. melanderi* targets were
413 enriched for phospholipid metabolic processes.
414

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



415

416 **Fig. 1. Diversity of miRNAs expressed in the brains of six bee species.** (A) Homologous and
417 lineage-specific miRNAs expressed in the brain of each species. The three types of homology
418 (shades of grey) correspond to those in Table 1. Black – has not been previously detected in other
419 species. Pie size corresponds to number of miRNAs detected from small RNA sequencing. Boxes
420 indicate social organization (green – complex eusociality, yellow – basic eusociality, pink –
421 facultative eusociality, blue – solitary). Phylogenetic relationships are following previous studies
422 (Cardinal and Danforth 2011; Sadd et al. 2015; Branstetter et al. 2017). (B) Diversification of miR-
423 13b across six bee species. Each miRNA has a matching seed sequence to ame-miR-13b-3p
424 (miRBase v21). Tip labels are the name of each miRNA expressed in the brain, beginning with a
425 four letter code for species. Different species are also indicated by colors. Scale indicates
426 substitution rate.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

429 If lineage-specific changes in gene regulatory function associated with social evolution
430 are facilitated by novel miRNAs inserted into existing gene networks, then the predicted targets
431 of lineage-specific miRNAs should be highly conserved and enriched for genes with a known
432 function in social evolution. Most of the predicted mRNA targets of lineage-specific miRNAs
433 were highly conserved and belonged to orthogroups shared by vertebrates (Fig. 2A; Table S8).
434 However, most of the genes in each genome are also highly conserved, and there was not a
435 significant enrichment for conserved genes among predicted targets of lineage-specific miRNAs,
436 beyond what would be expected by chance ($p > 0.99$). We did, however, find a significant
437 enrichment for lineage-specific genes that are unique to each species among the predicted targets
438 of lineage-specific miRNAs ($p = 0.02 - 1.48e^{-12}$), indicating that novel miRNAs are more likely
439 to target novel genes than would be expected by chance (Fig. 2A; Table S8).

440

441 We found mixed support for the prediction that novel miRNAs should target genes that
442 function in social behavior and evolution. The predicted targets of lineage-specific miRNAs
443 were enriched for genes that are differentially expressed between castes in the social Apidae (*A.*
444 *mellifera* and *B. terrestris*), but not Halictidae (*M. genalis*) (Fig. 2B; Table S4). In *A. mellifera*,
445 this included genes that are upregulated in the brains of reproductive workers, compared with
446 sterile workers (RF = 3.4, $p = 0.007$) and queens (RF = 1.6, $p = 0.015$) (Grozinger et al. 2007), as
447 well as genes upregulated in the brains of foragers compared with nurses (RF = 2.8, $p = 0.011$)
448 (Whitfield et al. 2003). However, there was not significant enrichment for genes differentially
449 expressed between nurse and forager honey bee brains in a later study ($p = 0.09$) (Alaux et al.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

450 2009). In *B. terrestris*, we find significant overlap between the predicted targets of lineage-
451 specific miRNAs and genes that are upregulated in workers, compared to queens (whole body,
452 including brain; RF = 2, p = 0.013). We did not find significant overlap with genes differentially
453 expressed in the brains of nurses and foragers (p = 0.103) (Porath et al. 2019) or between
454 reproductive and sterile worker brains (p = 0.39) (Marshall et al. 2019), but these were much
455 more limited gene sets. To our knowledge, there are no studies of gene expression differences
456 between *B. impatiens* castes, so we could not evaluate target overlap with caste-biased genes in
457 this species. We do not find significant enrichment for caste-biased genes in the brains of the
458 facultatively eusocial halictid *M. genalis* (p = 0.25).
459
460 Contrary to our prediction, targets of lineage-specific miRNAs were not significantly
461 enriched for genes under selection in any species. We assessed overlaps between genes
462 undergoing positive directional selection in *A. mellifera* (Harpur et al. 2014), *B. impatiens*
463 (Harpur et al. 2017), *M. genalis* (Kapheim et al., unpublished), and *N. melanderi* (Kapheim et al.
464 2019) and the predicted targets of lineage-specific miRNAs in each of these species. There was
465 no significant enrichment for predicted targets of lineage-specific miRNAs with genes under
466 positive directional selection in any species (Table S4). In fact, genes under selection in the
467 halictid bees were significantly depleted for targets of lineage-specific miRNAs (*M. genalis* – RF
468 = 0.2, p = $4.28e^{-10}$; *N. melanderi* – RF = 0.3, p = $5.59e^{-4}$). We also assessed overlaps with genes
469 previously found to be under selection in social species, compared to solitary species (Woodard
470 et al. 2011; Kapheim et al. 2015), but found only marginally significant overlap (Kapheim et al.
471 2015) or depletion (Woodard et al. 2011) with predicted targets of lineage-specific genes in one
472 species (*M. genalis* – RF = 1.9, p = 0.053; RF = 0, p = 0.05; Table S4).

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

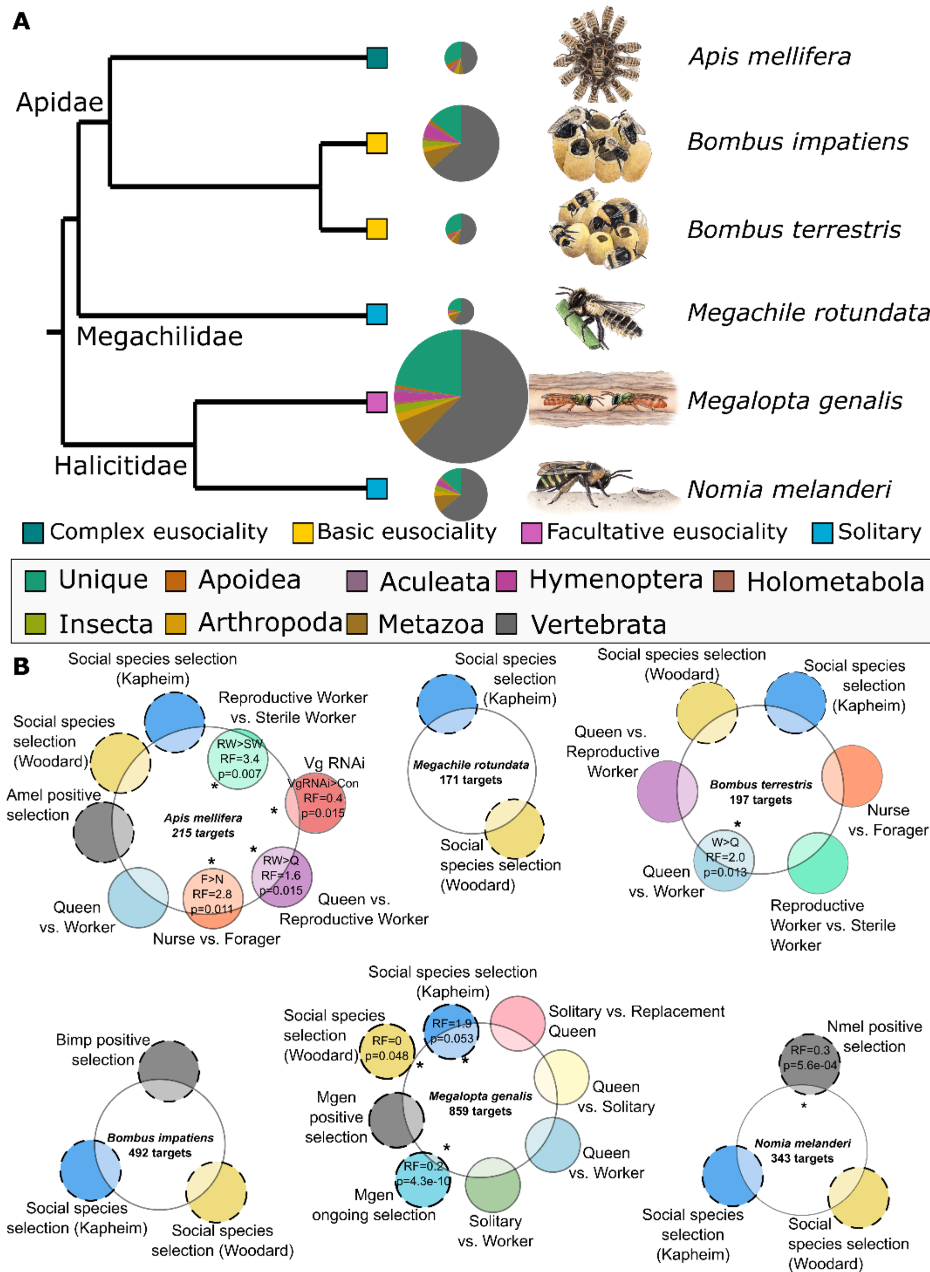


Fig. 2. Predicted targets of lineage-specific miRNAs in relationship to social behavior. (A) Genes predicted to be targeted by lineage-specific miRNAs are more likely to be unique to each species than predicted by chance. Pie chart size is scaled to number of predicted target genes. Color slices indicate orthogroup age for each predicted gene. (B) Genes that are both predicted targets of lineage-specific miRNAs and genes with differential expression in a social context (solid outlines) or genes under selection (dashed outlines) are represented by overlapping circles for each study and species. Numbers of lineage-specific miRNA targets are given for each species. Colors indicate different studies. Overlaps not significantly different from random (representation factor, RF=1) are unlabeled, while significant over- or under-enrichments are marked with asterisks with RF and p-value as indicated.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

485 **DISCUSSION**

486 Eusociality is a major evolutionary innovation that requires regulatory changes in a wide
487 range of molecular pathways (Robinson and Ben-Shahar 2002; Bloch and Grozinger 2011;
488 Sumner et al. 2018). We tested the hypothesis that miRNAs play a role in the evolution of
489 eusociality via their regulatory effects on gene networks. Our results provide several lines of
490 support for this hypothesis.

491

492 First, miRNAs that have been previously implicated in social behavior (i.e., miR-276,
493 miR-13b) do not show significant differences in copy number or diversification patterns
494 consistent with social evolution, but the predicted targets of these miRNAs are largely disparate
495 across species. This is consistent with what has been shown in vertebrates, flies, and nematodes,
496 in which miRNAs are highly conserved at the sequence level, but miRNA-target relationships
497 are highly divergent across clades (Chen and Rajewsky 2006; Chen and Rajewsky 2007). In our
498 analysis, there were similarities in the broad functional categories of predicted targets across
499 closely related species, and these categories were largely consistent with gene networks expected
500 to play a role in social evolution (e.g., hormone and signal transduction pathways). Future
501 research experimentally evaluating the function of these miRNAs in other social species will be
502 necessary to understand the degree to which miRNA function has diversified during social
503 evolution.

504

505 Second, we identified a single miRNA (miR-305) that was expressed exclusively in the
506 brains of the social bees in our study. The presence of this miRNA in the solitary bee genomes
507 suggests that an evolutionary shift in expression pattern has accompanied at least two

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

508 independent origins of eusociality in bees. This miRNA coordinates Insulin and Notch signaling
509 in *D. melanogaster*, and both of these pathways are important regulators of social dynamics in
510 insects (Wheeler et al. 2006; Ament et al. 2008; de Azevedo and Hartfelder 2008; Nilsen et al.
511 2011; Wang et al. 2013; Hartfelder et al. 2015; Duncan et al. 2016; Chandra et al. 2018;
512 Hartfelder et al. 2018). Some of the predicted targets of miR-305 in bees (e.g., *oxysterol*,
513 *synaptobrevin*) interact with Insulin and Notch pathways in *D. melanogaster* (Stewart et al.
514 2001; Obniski et al. 2018). This miRNA was also found to be upregulated in worker-destined
515 honey bee larvae, compared to queen-destined larvae, and may thus play a role in caste
516 differentiation (Shi et al. 2015). Interestingly, miR-305 also promotes aging in *D. melanogaster*
517 (Ueda et al. 2018). Extreme differences in lifespan are one of the defining characteristics of
518 social insect castes, whereby queens of some ant and bee species live orders of magnitude longer
519 than workers (Wilson 1971). The mechanisms underlying these differences are a topic of
520 interest, and a role for a miRNA that modulates insulin signaling is consistent with current
521 understanding of this phenomenon (Corona et al. 2007; Münch et al. 2008; Remolina and
522 Hughes 2008). Additional investigation is necessary to determine how this miRNA may
523 influence social behavior across species.
524
525 We focused additional attention on miRNAs for which no mature miRNAs with seed
526 matches were detected in any other species, because these have the potential to influence the
527 lineage-specific patterns of gene regulatory changes previously shown to influence social
528 evolution (Simola et al. 2013; Kapheim et al. 2015). We hypothesized that if novel miRNAs are
529 inserted into existing gene networks that become co-opted for social evolution, they should target
530 genes that are highly conserved across species. Instead, we find that the targets of lineage-

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

531 specific miRNAs are enriched for lineage-specific genes, while genes belonging to ancient
532 orthogroups were not more likely to be targets than expected by chance. This could suggest that
533 novel miRNAs co-evolve with novel genes, as has been shown for the evolution of cognitive
534 function in humans (Barbash et al. 2014). Previous work in honey bees has shown that
535 taxonomically-restricted (i.e., younger) genes play an important role in social evolution.
536 Expression of taxonomically-restricted genes is significantly biased toward glands with
537 specialized functions for life in a social colony (e.g., the hypopharyngeal gland and the sting
538 gland) (Jasper et al. 2015), and toward genes that are upregulated in workers (Johnson and
539 Tsutsui 2011). Thus, it is reasonable to expect that new miRNAs targeting new genes could have
540 important social functions in honey bees.
541
542 Alternatively, it is possible that new miRNAs that target lineage-specific genes are
543 transient and more likely to be purged by natural selection because they are less likely to
544 integrate into existing gene networks (Chen and Rajewsky 2007; França et al. 2016; França et al.
545 2017). Emergent miRNAs are expected to initially have limited expression to mitigate potential
546 deleterious effects on the protein-coding genes that they target. Thus, lineage-specific miRNAs
547 with low levels of expression may be in the process of being purged and may not have
548 accumulated gene targets with important functions (Chen and Rajewsky 2007; Berezikov 2011).
549 Evidence for this model comes from primates (Berezikov et al. 2006) and flies (Lu et al. 2008;
550 Tang et al. 2010). Likewise, we find that lineage-specific miRNAs are expressed at significantly
551 lower levels than those with at least one homolog in another species. A purging process like this
552 could explain why there are large differences in the numbers of miRNAs detected in even closely
553 related species (e.g., the two *Bombus* species). It is also possible that some of this variation

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

554 comes from sequencing stochasticity. Functional analysis of lineage-specific genes in additional
555 tissues and life stages will help to resolve their role in social evolution.
556
557 We find support for the prediction that lineage-specific miRNAs should target genes with
558 social function in the Apidae (e.g., honey bees and bumble bees), but not the Halictidae (*M.*
559 *genalis*). There are several potential explanations for this pattern. One explanation is technical.
560 We define genes with social functions as those that are differentially expressed among castes.
561 The genetic basis of social behavior has been much better studied in honey bees and bumble bees
562 than in any other species, and the sets of genes known to function in sociality is thus richer for
563 apids than for halictids. Further, not all genes that function in social behavior are expected to be
564 differentially expressed in the brains of different castes, and our analysis is thus likely to exclude
565 some important genes. Nonetheless, our results reflect differences in the degree of social
566 complexity, and thus caste-biased gene expression patterns, between apid and halictid bees.
567 Unlike for honey bees and bumble bees, which cannot live outside of social colonies, eusociality
568 is facultative in *M. genalis*. As such, caste traits are not fixed during development, and females
569 who once served as non-reproductive workers can transform into reproductive queens if given
570 the opportunity (Smith et al. 2009; Kapheim et al. 2012). This flexibility is reflected in the
571 magnitude of differences in brain gene expression patterns between queen and worker honey
572 bees (thousands of genes; Grozinger et al. 2007) and *M. genalis* (dozens of genes; Jones et al.
573 2017). Previous research suggests that miRNAs increase their functional influence over
574 evolutionary time (Berezikov et al. 2006; Chen and Rajewsky 2007; Lu et al. 2008; Roux et al.
575 2012; França et al. 2016; Patel and Capra 2017). Thus, emergent miRNAs are more likely to
576 target genes with social function due to chance alone in species with increased social complexity

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

577 and a larger set of caste-biased genes. Consistent with this explanation, regulatory relationships
578 between miRNAs and genes with caste-biased expression were not found among two other social
579 insect species with reduced social complexity (Patalano et al. 2015).
580
581 An additional explanation for these differences in the function of lineage-specific
582 miRNAs between bee families concerns the role of miRNAs in gene regulatory networks. One of
583 these roles is to stabilize regulatory relationships in the face of environmental variation, thus
584 canalizing phenotypes during development (Hornstein and Shomron 2006; Peterson et al. 2009;
585 Wu et al. 2009; Berezikov 2011). This is likely to be more important in species with obligate
586 eusociality, such as the honey bees and bumble bees for which caste determination is canalized,
587 than in species like *M. genalis*, where plasticity of phenotypes related to eusociality are
588 maintained in totipotent females.
589
590 Contrary to their effects on genes with socially-differentiated expression patterns,
591 lineage-specific miRNAs showed no evidence for preferential targeting of genes under positive
592 selection – either within individual species or across social species. In contrast, we find these
593 emergent miRNAs are less likely than expected by chance to target genes under positive
594 selection in the two halictid bees. A potential explanation for this pattern is that genes that are
595 adaptively targeted by miRNAs tend to be under purifying selection, because this maintains the
596 regulatory relationship between the miRNA and target, thus preventing gene mis-expression
597 (Chen and Rajewsky 2006; Saunders et al. 2007; Sarasin-Filipowicz et al. 2009; Franchini et al.
598 2016). However, this selective constraint is likely to be most significant in the 3’ UTR region,
599 where miRNA binding sites are located.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

600

601 A more likely explanation involves the hypothesized pattern of miRNA origins and
602 assimilation, as proposed by Chen and Rajewsky (2007). This model suggests that newly
603 emerged miRNAs are likely to have many targets throughout the genome due to chance. Most of
604 these initial miRNA-target regulatory relationships are likely to have slightly deleterious effects,
605 and would likely be purged through purifying selection very quickly. These deleterious effects
606 could be particularly strong for target genes undergoing positive selection, because changes in
607 the functional regulation of these genes are likely to have significant fitness consequences. Also,
608 genes under positive selection are undergoing rapid evolution, and thus may be more likely to
609 "escape" control by errant miRNAs. Indeed, it is easier for mRNAs to lose miRNA target
610 binding sites, which typically require exact sequence matches, than to gain them (Chen and
611 Rajewsky 2007). Thus, emergent miRNAs may not be expected to target adaptively or fast
612 evolving genes, regardless of their role in social evolution.

613

614 The evolution of eusociality depends on many different tissues and physiological
615 processes, and brain-specific expression patterns are not likely to be representative of the
616 complete role of individual miRNAs in social behavior. Some or all of the predicted miRNA-
617 gene relationships we identified may have evolved to support traits in other cell types or
618 processes unrelated to sociality. Additional sequencing of miRNA and mRNA across tissue-
619 types and stages of development in social and solitary species is necessary to provide a
620 comprehensive assessment of the role of emergent miRNAs in social traits. Nonetheless, the
621 brain is a major focus of research in social evolution because it is the primary source of

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

622 behavioral and neuroendocrine output. Our results thus provide a good starting place for
623 evaluating the role of miRNAs in lineage-specific processes in the evolution of social behavior.
624
625 Our analyses reveal important differences in patterns of miRNA evolution between bees
626 and other species. For example, expansion in miRNA repertoire is associated with the evolution
627 of animal complexity in a wide range of species (Heimberg et al. 2008; Christodoulou et al.
628 2010; Berezikov 2011). The evolution of eusociality from a solitary ancestor is associated with
629 increases in phenotypic complexity, and considered to be one of the major transitions in
630 evolution (Maynard Smith and Szathmáry 1995). We therefore hypothesized that evolutionary
631 increases in social complexity would be associated with expansions in the number of miRNAs
632 found within bee genomes. To the contrary, we find that most bees have a single copy of
633 previously identified miRNAs in their genomes. This is consistent with results of comparative
634 genome scans across several ant species (Simola et al. 2013). A recent study of miRNA diversity
635 in insects found that morphological innovations such as holometabolous development was
636 accompanied by the acquisition of only three miRNA families (Ylla et al. 2016). This suggests
637 that insect evolution is not as reliant on major expansions of miRNA families as other taxonomic
638 groups.
639
640 Additionally, our characterization of lineage-specific miRNAs expressed in the brain of
641 each species reveals that genome structure is not as influential in regulating bee miRNA
642 evolution as has been shown for human miRNAs. Novel human miRNAs tend to arise in ancient
643 genes that have multiple functions and broad expression patterns (França et al. 2016). It is
644 hypothesized that this increases the expression repertoire of emergent miRNAs, and thus

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

645 facilitates persistence in the population (França et al. 2016; França et al. 2017). Only in one
646 species (*N. melanderi*) were lineage-specific miRNAs more likely to be localized intragenically
647 than previously identified miRNAs, while lineage-specific miRNAs did not differ from
648 previously identified miRNAs in their genomic locations in the other five species. This suggests
649 that emergence patterns for new miRNAs are unique to each lineage in bees. We also do not find
650 a consistent pattern between young, emerging miRNAs and host gene age. There was no
651 significant difference in the age of genes that serve as hosts for established versus lineage-
652 specific miRNAs across all species. This is despite the fact that a similar proportion of bee
653 miRNAs are located within introns (31-43%; Table 1), as compared to in vertebrates (36-65%)
654 (Meunier et al. 2013). However, the fact that 73-88% of miRNAs that are localized to genes are
655 encoded on the sense strand suggests that they would benefit from host transcription, as is
656 observed in vertebrates (Meunier et al. 2013). Additional research with insects will be necessary
657 to identify general patterns of miRNA evolution in relationship to genome structure.
658
659 Our study identifies patterns of miRNA evolution in a set of closely related bees that vary
660 in social organization. Our results highlight important similarities and differences in the
661 emergence patterns and functions of mammalian and insect genomes. We also find evidence that
662 emergent miRNAs function in lineage-specific patterns of social evolution, perhaps through co-
663 evolution of novel miRNAs and species-specific targets. We do not see an overall increase in the
664 number of miRNAs in the genome or expressed in the brains of species with more complex
665 eusociality. However, we do find evidence that the role of miRNAs in social evolution may
666 strengthen with increasing social complexity, perhaps due to an increased need for canalization
667 of caste determination or due to chance, as a function of an increased number of genes with

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

668 caste-biased expression. Empirical tests of miRNA function across additional species with
669 variable social organization will further improve our understanding of how gene regulatory
670 evolution gives rise to eusociality.
671

672 **REFERENCES**

673 Alaux C, Le Conte Y, Adams HA, Rodriguez-Zas S, Grozinger CM, Sinha S, Robinson GE.
674 2009. Regulation of brain gene expression in honey bees by brood pheromone. Genes, Brain
675 Behav. 8:309–319.
676 Ament SA, Corona M, Pollock HS, Robinson GE. 2008. Insulin signaling is involved in the
677 regulation of worker division of labor in honey bee colonies. Proc. Natl. Acad. Sci. U. S. A.
678 105:4226–4231.
679 Asgari S. 2013. MicroRNA functions in insects. Insect Biochem. Mol. Biol. 43:388–397.
680 Ashby R, Forêt S, Searle I, Maleszka R. 2016. MicroRNAs in honey bee caste determination.
681 Sci. Rep. 6:1–15.
682 de Azevedo SV, Hartfelder K. 2008. The insulin signaling pathway in honey bee (*Apis mellifera*)
683 caste development - differential expression of insulin-like peptides and insulin receptors in
684 queen and worker larvae. J. Insect Physiol. 54:1064–1071.
685 Barbash S, Shifman S, Soreq H. 2014. Global coevolution of human microRNAs and their target
686 genes. Mol. Biol. Evol. 31:1237–1247.
687 Bartel DP. 2018. Metazoan microRNAs. Cell 173:20–51.
688 Behura SK, Whitfield CW. 2010. Correlated expression patterns of microRNA genes with age-
689 dependent behavioural changes in honeybee. Insect Mol. Biol. 19:431–439.
690 Berezikov E. 2011. Evolution of microRNA diversity and regulation in animals. Nat. Rev.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

691 Genet. 12:846–860.

692 Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk RHA.
693 2006. Diversity of microRNAs in human and chimpanzee brain. Nat. Genet. 38:1375–1377.

694 Bloch G, Grozinger CM. 2011. Social molecular pathways and the evolution of bee societies.
695 Philos. Trans. R. Soc. B Biol. Sci. 366:2155–2170.

696 Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, Qin N, Donahue G, Yang P, Li Q, Li C, et al.
697 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*.
698 Science 329:1068–1071.

699 Branstetter MG, Danforth BN, Pitts JP, Faircloth BC, Ward PS, Buffington ML, Gates MW,
700 Kula RR, Brady SG. 2017. Phylogenomic insights into the evolution of stinging wasps and
701 the origins of ants and bees. Curr. Biol. 27:1019–1025.

702 Cardinal S, Danforth BNBN. 2011. The antiquity and evolutionary history of social behavior in
703 bees. PLoS One 6:e21086.

704 Chandra V, Fetter-Pruneda I, Oxley PR, Ritger AL, McKenzie SK, Libbrecht R, Kronauer DJC.
705 2018. Social regulation of insulin signaling and the evolution of eusociality in ants. Science
706 361:398–402.

707 Chen K, Rajewsky N. 2006. Deep conservation of microRNA-target relationships and 3′UTR
708 motifs in vertebrates, flies, and nematodes. In: Cold Spring Harbor Symposia on
709 Quantitative Biology. Vol. 71. p. 149–156.

710 Chen K, Rajewsky Nikolaus. 2006. Natural selection on human microRNA binding sites inferred
711 from SNP data. Nat. Genet. 38:1452–1456.

712 Chen K, Rajewsky N. 2007. The evolution of gene regulation by transcription factors and
713 microRNAs. Nat. Rev. Genet. 8:93–103.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

714 Christodoulou F, Raible F, Tomer R, Simakov O, Trachana K, Klaus S, Snyman H, Hannon GJ,
715 Bork P, Arendt D. 2010. Ancient animal microRNAs and the evolution of tissue identity.
716 Nature 463:1084–1088.
717 Collins DH, Mohorianu I, Beckers M, Moulton V, Dalmay T, Bourke AFG. 2017. MicroRNAs
718 associated with caste determination and differentiation in a primitively eusocial insect. Sci.
719 Rep. 7:1–9.
720 Corona M, Velarde RA, Remolina S, Moran-Lauter A, Wang Y, Hughes KA, Robinson GE.
721 2007. Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity.
722 Proc. Natl. Acad. Sci. U. S. A. 104:7128–7133.
723 Cui X, Lu Z, Wang S, Jing-Yan Wang J, Gao X. 2016. CMsearch: simultaneous exploration of
724 protein sequence space and structure space improves not only protein homology detection
725 but also protein structure prediction. Bioinformatics 32:i332–i340.
726 Duncan EJ, Hyink O, Dearden PK. 2016. Notch signalling mediates reproductive constraint in
727 the adult worker honeybee. Nat. Commun. 7:1–10.
728 Elsik CG, Worley KC, Bennett AK, Beye M, Camara F, Childers CP, de Graaf DC, Debyser G,
729 Deng J, Devreese B, et al. 2014. Finding the missing honey bee genes: Lessons learned
730 from a genome upgrade. BMC Genomics 15:1–29.
731 Enright AJ, John B, Gaul U, Tuschl T, Sander C. 2004. MicroRNA targets in *Drosophila*.
732 Genome Biol. 5:R1.
733 Filipowicz W, Bhattacharyya SN, Sonenberg N. 2008. Mechanisms of post-transcriptional
734 regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. 9:102–114.
735 França GS, Hinske LC, Galante PAFF, Vibranovski MD. 2017. Unveiling the impact of the
736 genomic architecture on the evolution of vertebrate microRNAs. Front. Genet. 8:1–8.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

737 França GS, Vibranovski MD, Galante PAF. 2016. Host gene constraints and genomic context
738 impact the expression and evolution of human microRNAs. Nat. Commun. 7.
739 Franchini P, Xiong P, Fruciano C, Meyer A. 2016. The role of microRNAs in the repeated
740 parallel diversification of lineages of midas cichlid fish from Nicaragua. Genome Biol.
741 Evol. 8:1543–1555.
742 Friedländer MR, MacKowiak SD, Li N, Chen W, Rajewsky N. 2012. MiRDeep2 accurately
743 identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic
744 Acids Res. 40:37–52.
745 Friedman RC, Farh KK-H, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved
746 targets of microRNAs. Genome Res. 19:92–105.
747 Gentleman R, Falcon S. 2013. Package "GOstats." 2.26.0.
748 Greenberg JK, Xia J, Zhou X, Thatcher SR, Gu X, Ament SA, Newman TC, Green PJ, Zhang W,
749 Robinson GE, et al. 2012. Behavioral plasticity in honey bees is associated with differences
750 in brain microRNA transcriptome. Genes, Brain Behav. 11:660–670.
751 Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, Degnan BM, Rokhsar
752 DS, Bartel DP. 2008. Early origins and evolution of microRNAs and Piwi-interacting RNAs
753 in animals. Nature 455:1193–1197.
754 Grozinger CM, Fan Y, Hoover SER, Winston ML. 2007. Genome-wide analysis reveals
755 differences in brain gene expression patterns associated with caste and reproductive status
756 in honey bees (*Apis mellifera*). Mol. Ecol. 16:4837–4848.
757 Grün D, Wang Y-L, Langenberger D, Gunsalus KC, Rajewsky N. 2005. microRNA target
758 predictions across seven *Drosophila* species and comparison to mammalian targets. PLoS
759 Comput. Biol. 1:e13.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

760 Harpur BA, Dey A, Albert JR, Patel S, Hines HM, Hasselmann M, Packer L, Zayed A. 2017.
761 Queens and workers contribute differently to adaptive evolution in bumble bees and honey
762 bees. Genome Biol. Evol. 9:2395–2402.
763 Harpur BA, Kent CF, Molodtsova D, Lebon JM, Alqarni AS, Owayss AA, Zayed A. 2014.
764 Population genomics of the honey bee reveals strong signatures of positive selection on
765 worker traits. Proc Natl Acad Sci U S A 111:2614–2619.
766 Harrison MC, Hammond RL, Mallon EB. 2015. Reproductive workers show queenlike gene
767 expression in an intermediately eusocial insect, the buff-tailed bumble bee *Bombus*
768 *terrestris*. Mol. Ecol. 24:3043–3063.
769 Hartfelder K, Guidugli-Lazzarini KR, Cervoni MS, Santos DE, Humann FC. 2015. Old threads
770 make new tapestry—rewiring of signalling pathways underlies caste phenotypic plasticity in
771 the honey bee, *Apis mellifera* L. In: Amro Z, Clement FK, editors. Advances in Insect
772 Physiology. Vol. 48. 1st ed. Elsevier Ltd. p. 1–36.
773 Hartfelder K, Tiberio GJ, Lago DC, Dallacqua RP, Bitondi MMG. 2018. The ovary and its
774 genes—developmental processes underlying the establishment and function of a highly
775 divergent reproductive system in the female castes of the honey bee, *Apis mellifera*.
776 Apidologie 49:49–70.
777 Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ. 2008. MicroRNAs and the
778 advent of vertebrate morphological complexity. Proc. Natl. Acad. Sci. U. S. A. 105:2946–
779 2950.
780 Hornstein E, Shomron N. 2006. Canalization of development by microRNAs. Nat. Genet.
781 38:S20–S24.
782 Jasper WC, Linksvayer TA, Atallah J, Friedman D, Chiu JC, Johnson BR. 2015. Large-scale

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

783 coding sequence change underlies the evolution of postdevelopmental novelty in honey
784 bees. Mol. Biol. Evol. 32:334–346.
785 Johnson B, Tsutsui N. 2011. Taxonomically restricted genes are associated with the evolution of
786 sociality in the honey bee. BMC Genomics 12:164.
787 Johnson BR, Linksvayer TA. 2010. Deconstructing the superorganism: social physiology,
788 groundplans, and sociogenomics. Q. Rev. Biol. 85:57–79.
789 Jones BM, Kingwell CJ, Wcislo WT, Robinson GE. 2017. Caste-biased gene expression in a
790 facultatively eusocial bee suggests a role for genetic accommodation in the evolution of
791 eusociality. Proc. R. Soc. B Biol. Sci. 284.
792 Kapheim KM, Jones BM, Pan H, Harpur BA, Kent CF, Zayed A, Ioannidis P, Waterhouse RM,
793 Kingwell C, Stolle E, et al. Developmental plasticity shapes the genomic signatures of
794 eusociality. unpublished.
795 Kapheim KM, Pan H, Li C, Blatti C, Harpur BA, Ioannidis P, Jones BM, Kent CF, Ruzzante L,
796 Sloofman L, et al. 2019. Draft genome assembly and population genetics of an agricultural
797 pollinator, the solitary alkali bee (Halictidae: *Nomia melanderi*). G3 9:625–634.
798 Kapheim KM, Pan H, Li C, Salzberg SL, Puiu D, Magoc T, Robertson HM, Hudson ME, Venkat
799 A, Fischman BJ, et al. 2015. Genomic signatures of evolutionary transitions from solitary to
800 group living. Science 348:24–32.
801 Kapheim KM, Smith AR, Ihle KE, Amdam G V, Nonacs P, Wcislo WT. 2012. Physiological
802 variation as a mechanism for developmental caste-biasing in a facultatively eusocial sweat
803 bee. Proc. R. Soc. B Biol. Sci. 279:1437–1446.
804 Kocher SD, Li C, Yang W, Tan H, Yi S V, Yang X, Hoekstra HE, Zhang G, Pierce NE, Yu DW.
805 2013. The draft genome of a socially polymorphic halictid bee, *Lasioglossum albipes*.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

806 Genome Biol. 14:R142.

807 Krüger J, Rehmsmeier M. 2006. RNAhybrid: microRNA target prediction easy, fast and flexible.
808 Nucleic Acids Res. 34:W451-4.

809 Liu F, Peng W, Li Z, Li W, Li L, Pan J, Zhang S, Miao Y, Chen S, Su S. 2012. Next-generation
810 small RNA sequencing for microRNAs profiling in *Apis mellifera*: Comparison between
811 nurses and foragers. Insect Mol. Biol. 21:297–303.

812 Lu J, Shen Y, Wu Q, Kumar S, He B, Shi S, Carthew RW, Wang SM, Wu C-I. 2008. The birth
813 and death of microRNA genes in *Drosophila*. Nat. Genet. 40:351–355.

814 MacEdo LMF, Nunes FMF, Freitas FCP, Pires C V., Tanaka ED, Martins JR, Piulachs MD,
815 Cristino AS, Pinheiro DG, Simões ZLP. 2016. MicroRNA signatures characterizing caste-
816 independent ovarian activity in queen and worker honeybees (*Apis mellifera* L.). Insect
817 Mol. Biol. 25:216–226.

818 Marshall H, Lonsdale ZN, Mallon EB. 2019. Methylation and gene expression differences
819 between reproductive and sterile bumblebee workers. Evol. Lett.

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

822 Meunier J, Lemoine F, Soumillon M, Liechti A, Weier M, Guschanski K, Hu H, Khaitovich P,
823 Kaessmann H. 2013. Birth and expression evolution of mammalian microRNA genes.
824 Genome Res. 23:34–45.

825 Münch D, Amdam G V, Wolschin F. 2008. Ageing in a eusocial insect: molecular and
826 physiological characteristics of life span plasticity in the honey bee. Funct. Ecol. 22:407–
827 421.

828 Nilsen K-A, Ihle KE, Frederick K, Fondrk MK, Smedal B, Hartfelder K, Amdam G V. 2011.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

829 Insulin-like peptide genes in honey bee fat body respond differently to manipulation of
830 social behavioral physiology. J. Exp. Biol. 214:1488–1497.
831 Nunes FMF, Ihle KE, Mutti NS, Simoes ZLP, Amdam G V. 2013. The gene *vitellogenin* affects
832 microRNA regulation in honey bee (*Apis mellifera*) fat body and brain. J Exp Biol
833 216:3724–3732.
834 Obniski R, Sieber M, Spradling AC. 2018. Dietary lipids modulate Notch signaling and
835 influence adult intestinal development and metabolism in *Drosophila*. Dev. Cell 47:98-
836 111.e5.
837 Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R
838 language. Bioinformatics 20:289–290.
839 Patalano S, Vlasova A, Wyatt C, Ewels P, Camara F, Ferreira PG, Asher CL, Jurkowski TP,
840 Segonds-Pichon A, Bachman M, et al. 2015. Molecular signatures of plastic phenotypes in
841 two eusocial insect species with simple societies. Proc. Natl. Acad. Sci. 112:13970–13975.
842 Patel VD, Capra JA. 2017. Ancient human miRNAs are more likely to have broad functions and
843 disease associations than young miRNAs. BMC Genomics 18:672.
844 Peterson KJ, Dietrich MR, McPeek MA. 2009. MicroRNAs and metazoan macroevolution:
845 insights into canalization, complexity, and the Cambrian explosion. BioEssays 31:736–747.
846 Porath HT, Hazan E, Shpigler H, Cohen M, Band M, Ben-Shahar Y, Levanon EY, Eisenberg E,
847 Bloch G. 2019. RNA editing is abundant and correlates with task performance in a social
848 bumblebee. Nat. Commun. 10:1605.
849 Quah S, Hui JHL, Holland PWH. 2015. A burst of miRNA innovation in the early evolution of
850 butterflies and moths. Mol. Biol. Evol. 32:1161–1174.
851 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

852 features. Bioinformatics 26:841–842.
853 Remolina SC, Hughes KA. 2008. Evolution and mechanisms of long life and high fertility in
854 queen honey bees. Age (Omaha). 30:177–185.
855 Robinson GE, Ben-Shahar Y. 2002. Social behavior and comparative genomics: new genes or
856 new gene regulation? Genes, Brain Behav. 1:197–203.
857 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential
858 expression analysis of digital gene expression data. Bioinformatics 26.
859 Roux J, Gonzàlez-Porta M, Robinson-Rechavi M. 2012. Comparative analysis of human and
860 mouse expression data illuminates tissue-specific evolutionary patterns of miRNAs. Nucleic
861 Acids Res. 40:5890–5900.
862 Sadd BM, Barribeau SM, Bloch G, de Graaf DC, Dearden P, Elsik CG, Gadau J,
863 Grimmelikhuijzen CJ, Hasselmann M, Lozier JD, et al. 2015. The genomes of two key
864 bumblebee species with primitive eusocial organization. Genome Biol. 16:76.
865 Sarasin-Filipowicz M, Krol J, Markiewicz I, Heim MH, Filipowicz W. 2009. Decreased levels of
866 microRNA miR-122 in individuals with hepatitis C responding poorly to interferon therapy.
867 Nat. Med. 15:31–33.
868 Saunders MA, Liang H, Li W-H. 2007. Human polymorphism at microRNAs and microRNA
869 target sites. Proc. Natl. Acad. Sci. U. S. A. 104:3300–3305.
870 Shi YY, Zheng HJ, Pan QZ, Wang ZL, Zeng ZJ. 2015. Differentially expressed microRNAs
871 between queen and worker larvae of the honey bee (*Apis mellifera*). Apidologie 46:35–45.
872 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert
873 M, Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence
874 alignments using Clustal Omega. Mol. Syst. Biol. 7:539.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

875 Simola DF, Wissler L, Donahue G, Waterhouse RM, Helmkampf M, Roux J, Nygaard S, Glastad
876 KM, Hagen DE, Viljakainen L, et al. 2013. Social insect genomes exhibit dramatic
877 evolution in gene composition and regulation while preserving regulatory features linked to
878 sociality. Genome Res. 23:1235–1247.
879 Smith AF., Hubley R, Green P. 2013. RepeatMasker.
880 Smith AR, Kapheim KM, O'Donnell S, Wcislo WT. 2009. Social competition but not subfertility
881 leads to a division of labour in the facultatively social sweat bee *Megalopta genalis*
882 (Hymenoptera: Halictidae). Anim. Behav. 78:1043–1050.
883 Søvik E, Bloch G, Ben-Shahar Y. 2015. Function and evolution of microRNAs in eusocial
884 Hymenoptera. Front. Genet. 6:1–11.
885 Stewart BA, Mohtashami M, Zhou L, Trimble WS, Boulianne GL. 2001. SNARE-dependent
886 signaling at the *Drosophila* wing margin. Dev. Biol. 234:13–23.
887 Sumner S, Bell E, Taylor D. 2018. A molecular concept of caste in insect societies. Curr. Opin.
888 Insect Sci. 25:42–50.
889 Tang T, Kumar S, Shen Y, Lu J, Wu M-L, Shi S, Li W-H, Wu C-I. 2010. Adverse interactions
890 between micro-RNAs and target genes from different species. Proc. Natl. Acad. Sci. U. S.
891 A. 107:12935–12940.
892 Team RC. 2016. R: A language and environment for statistical computing.
893 Toth AL, Robinson GE. 2007. Evo-devo and the evolution of social behavior. Trends Genet.
894 23:334–341.
895 Ueda M, Sato T, Ohkawa Y, Inoue YH. 2018. Identification of *miR-305* , a microRNA that
896 promotes aging, and its target mRNAs in *Drosophila*. Genes to Cells 23:80–93.
897 Wang Y, Azevedo S V, Hartfelder K, Amdam G V. 2013. Insulin-like peptides (AmILP1 and

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

898 AmILP2) differentially affect female caste development in the honey bee (*Apis mellifera*
899 L.). J Exp Biol 216:4347–4357.
900 Warner MR, Qiu L, Holmes MJ, Mikheyev AS, Linksvayer TA. 2019. Convergent eusocial
901 evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes.
902 Nat. Commun. 10:2651.
903 Weaver D, Anzola J, Evans J, Reid J, Reese J, Childs K, Zdobnov E, Samanta M, Miller J, Elsik
904 C. 2007. Computational and transcriptional evidence for microRNAs in the honey bee
905 genome. Genome Biol. 8:R97.
906 Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ. 2009.
907 The deep evolution of metazoan microRNAs. Evol. Dev. 11:50–68.
908 Wheeler DE, Buck N, Evans JD. 2006. Expression of insulin pathway genes during the period of
909 caste determination in the honey bee, *Apis mellifera*. Insect Mol. Biol. 15:597–602.
910 Wheeler MM, Ament SA, Rodriguez-Zas SL, Robinson GE. 2013. Brain gene expression
911 changes elicited by peripheral *vitellogenin* knockdown in the honey bee. Insect Mol. Biol.
912 22:562–573.
913 Whitfield CW, Cziko A-M-. M, Robinson GE. 2003. Gene expression profiles in the brain
914 predict behavior in individual honey bees. Science 302:296–299.
915 Wilson EO. 1971. The Insect Societies. Cambridge, Massachusetts: Harvard University Press
916 Woodard SH, Fischman BJ, Venkat A, Hudson ME, Varala K, Cameron SA, Clark AG,
917 Robinson GE. 2011. Genes involved in convergent evolution of eusociality in bees. Proc
918 Natl Acad Sci U S A 108:7472–7477.
919 Wu C-I, Shen Y, Tang T. 2009. Evolution under canalization and the dual roles of microRNAs: a
920 hypothesis. Genome Res. 19:734–743.

bioRxiv preprint doi: https://doi.org/10.1101/730317; this version posted August 28, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

921 Ylla G, Fromm B, Piulachs MD, Belles X. 2016. The microRNA toolkit of insects. Sci. Rep.
922 6:1–13.
923 Zdobnov EM, Tegenfeldt F, Kuznetsov D, Waterhouse RM, Simão FA, Ioannidis P, Seppey M,
924 Loetscher A, Kriventseva E V. 2017. OrthoDB v9.1: cataloging evolutionary and functional
925 annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. Nucleic Acids
926 Res. 45:D744–D749.
927 Zhu Y, Skogerbø G, Ning Q, Wang Z, Li B, Yang S, Sun H, Li Y. 2012. Evolutionary
928 relationships between miRNA genes and their activity. BMC Genomics 13:718.
929

930 **DATA AVAILABILITY**

931 Sequences are deposited at NCBI SRA as BioProject PRJNA559906. Code is available upon
932 request.
933

934 **ACKNOWLEDGMENTS**

935 This work was supported by the USDA National Institute of Food and Agriculture [2018-67014-
936 27542 to K.M.K.]; the Utah Agricultural Experiment Station, Utah State University [Project
937 1297, journal paper number 9239 to K.M.K.]; the U.S.-Israel Binational Science Foundation
938 [BSF 2012807 to G.B. and Y.B.S.]; and the Swiss National Science Foundation
939 [PP00P3_170664 to R. M. W.]. G.B. thanks the Clark Way Harrison Visiting Professor in Arts
940 and Sciences that supported his stay in Washington University in St. Louis. Sequencing was
941 performed at the University of Illinois Roy J. Carver Biotechnology Center. We thank the
942 University of Utah High Performance Computing Center for computational time and assistance.
943 Illustrations were created by J. Johnson (LifeSciences Studios). We thank G. Robinson for
944 helpful feedback on an earlier draft of this manuscript.