In sickness and in health: viral infections correlate with the quality of ant bodyguarding behavior in an Amazonian ant-plant symbiosis

María C. Tocora[1,2], Christopher Reid[1], Jesse Huisken[1,3], Haoran Xue[1,4], Megan E. Frederickson[1]

1. Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S 3B2, Canada
2. Department of Biological Sciences, Texas Tech University, Lubbock, TX, USA
3. Department of Biology, York University, Toronto, ON M3J 1P3, Canada
4. Institute for Biochemistry and Biology, University of Potsdam, 14476, Potsdam-Golm, Germany

Keywords: ant behavior, ant-plant mutualism, symbiosis, transcriptomics, viral infection

Abstract

Ant-plant symbioses are classic examples of mutualism in which ant "bodyguards" defend myrmecophytic plants against enemies in exchange for nest sites and often food. We used RNA-Seq to profile the transcriptomes of *Allomerus octoarticulatus* ant workers, which aggressively defend the Amazonian plant *Cordia nodosa* against herbivores, but to varying degrees. Field behavioral assays with herbivores in the Peruvian Amazon showed striking variation among colonies in the relative zeal with which *A. octoarticulatus* workers defend their host plant. Highly effective and ineffective bodyguards differed in their gene expression profiles, which revealed viral infections significantly associated with ant bodyguarding behavior. Transcripts from eight new positive-sense single-stranded RNA viruses were differentially expressed between colonies with high- or low-quality bodyguards. Colonies of high- and low-quality bodyguards were infected by distinct viruses, including viruses clustering phylogenetically with viruses known to cause aggression or reduced locomotion, respectively, in bees. Gene expression, including of immunity-related genes, also differed between broodcare workers and

bodyguard ants, suggesting bodyguarding is a distinct worker task. Ant colony health and viral infections may influence ant cooperation with plants in ant-plant mutualisms.

Significance statement

Many animals have evolved to cooperate with plants in plant-animal mutualisms, including in pollination, seed dispersal, and defensive ant-plant interactions. Although the ecology of these interactions is well studied, little is known about the molecular mechanisms involved in animal cooperation with plants. Studying an Amazonian ant-plant symbiosis, we combined field assays of ant behavior with RNA-Seq to profile the transcriptomes of ants that vary in their quality as defensive plant "bodyguards." Bodyguard ants that effectively defended plants against herbivores expressed different genes than both ineffective bodyguards and broodcare workers. Surprisingly, many genes differentially expressed between high- and low-quality bodyguards were from viruses, and not ants. Thus, we linked viral infections to mutualist quality in this ant-plant symbiosis.

Introduction

Most plants and animals are involved in at least one mutualism; they depend on other species for nutrition, protection, or dispersal (Bronstein, 2015). Yet how mutualism evolves remains poorly understood, especially in insects, as we know little about the genes that encode mutualistic behaviors, or how these behaviors are regulated at the molecular level and evolve in insect genomes. In insects and many other animals, cooperation is typically a behavioral trait, and many studies have explored the molecular basis of social behavior (i.e., within-species cooperation) in insects such as ants (Kohlmeier et al., 2019; Manfredini et al., 2013) and bees

(Jones et al., 2020). In contrast, very little is known about the molecular mechanisms that give rise to cooperative animal behaviors directed at other species in mutualisms. Here, we take a transcriptomic approach to understand the defensive "bodyguard" behavior of symbiotic ants that defend their myrmecophytic hosts against herbivores.

Many theoretical models explore mutualism evolution, often with the goal of explaining how mutualism originates and persists despite the potential for underlying fitness conflicts between partners (Sachs & Simms, 2006). Nonetheless, in plants and microbes, where research has made great strides towards unraveling the "molecular dialogue" between partners, the evolution of genes known to encode mutualistic traits does not always match theoretical predictions (e.g., Epstein et al., 2023). For example, theory predicts that fitness conflicts select for "cheaters" that drive the breakdown or dissolution of mutualisms, but cheating is unlikely to be an important source of variation in mutualist quality in many systems (Frederickson, 2013). Instead, symbionts often vary in effectiveness because they are mismatched to their host's genotype (Batstone et al., 2020) or environment (Kiers et al., 2007). Failure to cooperate might also occur because of stress or illness. Disease might trigger a decrease in mutualistic behavior, making symbiotic partners "defective, but not defectors" (Friesen, 2012; Riehl & Frederickson, 2016). Although many studies have explored how pathogens modulate host behavior to increase parasite fitness and transmission (de Bekker, et al., 2018), how parasites or pathogens impact the beneficial behaviors of mutualists has received little attention outside of pollination mutualisms (Gillespie & Adler, 2013).

Approximately 4,000 species of angiosperms interact mutualistically with ants (Weber & Keeler, 2013). These associations are classic examples of defensive mutualisms, in which plants provide food or housing to "bodyguard" ants that defend them against herbivores or other

enemies. Macroevolutionary analyses suggest that ant-plant mutualisms spurred plant diversification (Weber & Agrawal, 2014). However, despite ample ecological study (Mayer et al., 2014), few studies have applied 'omics approaches to ant-plant interactions. On the plant side, metabolomic and transcriptomic comparisons have assessed the effects of ant colonization on *Tococa* ant-plants (Müller et al., 2022). On the ant side, Malé et al. (2017) suggested that the cGMP-PKG signaling pathway modulates ant protection of plants, while Rubin and Moreau (2016) found higher rates of molecular evolution genome-wide in mutualistic acacia-ants compared to their non-mutualistic close relatives, but there has been little additional research.

Here, we studied the mutualism between the ant species *Allomerus octoarticulatus* and the Amazonian myrmecophyte *Cordia nodosa* (Fig. 1A). *Cordia nodosa* has hollow stem swellings called domatia where ants nest, with a single *A. octoarticulatus* colony nesting across many domatia on a single *C. nodosa* tree. We combined bioassays in the field with transcriptomics to unravel the molecular mechanisms underlying mutualistic outcomes. We first characterized bodyguard behavior in ant colonies in the Peruvian Amazon by placing herbivores (grasshoppers) on plants and monitoring worker recruitment (Supplementary Video 1 and Fig. 1). Most measured variables, including herbivore size, age, and species (Supplementary Fig. 1), did not significantly affect worker recruitment, but ant activity in these bioassays did significantly predict herbivory on young *C. nodosa* leaves. We then identified transcriptional signatures of mutualist quality by analyzing differential gene expression between high- and low-quality ant bodyguards. We also used RNA-Seq on broodcare workers and bodyguards to identify task-specific transcripts that vary between ants working inside and outside the nest. Our results show that highly effective ant bodyguards express different genes than both ineffective ant bodyguards in other colonies and broodcare workers from the same nest. We also found underlying viral

infections and variation in the expression of immunity-related genes that correlated with bodyguarding activity, suggesting that ant pathogens can affect not only the degree of cooperation exhibited by workers, but also the fitness of the host plants that they inhabit. Finally, we investigated the phylogenetic placement of eight new positive-sense single-stranded RNA viruses that were differentially associated with illness-related behavioral changes in bodyguard ants.

Results

Ant colonies vary in their effectiveness as plant bodyguards

To characterize the bodyguarding behavior of *Allomerus octoarticulatus* colonies, we conducted behavioral bioassays on 60 naturally occurring colonies nesting in *C. nodosa* at the Los Amigos Biological Station (12°34'9" S, 70°6'0.40" W) in Madre de Dios, Peru from December 2018 to January 2019. Ant colonies varied considerably in bodyguarding behavior, from almost no ants attacking herbivores (even though all plants were occupied by *A. octoarticulatus* colonies) to sustained high recruitment to herbivores (Fig. 1, Supplementary Video 1). The ambient temperature, plant size (a proxy for ant colony size, Frederickson et al., 2012), and herbivore age (e.g., juvenile or adult), herbivore species, and herbivore size all had non-significant effects on bodyguarding behavior, but significantly fewer ants attacked herbivores in the afternoon than in the morning (Supplementary Fig. 1). The mean number of ants that attacked herbivores in bioassays significantly predicted herbivory over three weeks on the associated young leaves; plants with more active ant bodyguards experienced less herbivory (Fig. 1, linear regression, $R^2 = 0.094$, $p = 0.019$).

Gene expression correlates with variation in ant bodyguarding behavior across colonies

We used RNA-Seq to assemble a *de novo* transcriptome for *A. octoarticulatus* (see Methods) and measured gene expression in ant bodyguards from 14 colonies, seven colonies each of high-quality bodyguards and low-quality bodyguards. For six colonies (Supplementary Table 1), we also generated RNA-Seq data from both ant bodyguards attacking herbivores on leaves and broodcare workers collected from inside the nearest domatium (i.e., ant nest).

We found 40 genes that were differentially expressed between high- and low-quality bodyguards, comprising 15 genes upregulated in low-quality bodyguards and 25 genes upregulated in high-quality bodyguards (Fig. 2). These genes have expression patterns correlated with better or worse ant defense of plants, meaning that their expression either generates or responds to changes in the mutualistic behavior of ants towards their host plants. Only one sample (s8), characterized as highly aggressive (average activity of 27.2 attacks/min), did not show the same expression pattern for most genes as the rest of the high-activity bodyguards (Fig. 2A). At the transcript level, we found 13 genes with differential transcript usage (DTU) between high- and low-quality bodyguards participating in different metabolic pathways (Supplementary Fig. 2, Supplementary Table 2). Manual annotation revealed only a few DEGs with functions studied in other insects, including cytoskeleton proteins and protease genes overexpressed in low-activity bodyguards (Fig. 2B and 2C, Supplementary Table 3). We used Gene Set Enrichment Analysis (GSEA) (FDR < 0.05) to test for gene sets that differed between high- and low-activity bodyguards, but we did not find any statistically significant gene sets (Supplementary Table 4). However, we would not necessarily expect GSEA to identify distinct expression profiles between high- and low-quality bodyguards if the behavioral differences among bodyguard ants result from other factors such as viral infections.

140

141 *Viral infections are associated with mutualistic behavior*

142 Eight of the 40 genes differentially expressed between high- and low-activity bodyguards were
143 viral sequences, suggesting that viral infection may influence *A. octoarticulatus* behavior, or *vice*
144 *versa*. These transcripts came from eight new positive-sense single-stranded RNA (+ssRNA)
145 viruses, one matching homologous sequence from the family Picornaviridae, three matching
146 Iflaviridae, and three others matching Dicistroviridae viruses (Supplementary Table 5). We also
147 identified one virus in the family Narnaviridae (Supplementary Table 6).

148 To further examine the evolutionary history of these potential new Picorna-like and
149 Narna viruses, we reconstructed a phylogeny comparing their RNA-dependent RNA polymerase
150 (RdRp) sequences, a highly conserved viral domain, against RdRp regions of representative
151 viruses within every known viral family (Fig. 3, Supplementary Tables 7 and 8). Phylogenetic
152 analyses supported the placement of these viruses within the picorna- and narna-like clades.
153 However, the viruses associated with *A. octoarticulatus* appear to be unique lineages that may
154 represent new taxa. DGE analyses revealed different levels of viral expression between high- and
155 low-activity bodyguards (Fig. 4A-H).

156

157 *Gene expression correlates with ant task within colonies*

158 We found 155 DEGs between broodcare workers and high-activity bodyguards; 133 were
159 upregulated in brood-care workers and 22 were upregulated in high-activity bodyguards (Fig.
160 5A, Supplementary Table 9). We observed similar patterns of gene expression when comparing
161 broodcare workers and low-activity bodyguards, which identified 197 DEGs, 144 upregulated in
162 broodcare workers and 53 genes upregulated in low-activity bodyguards (Fig. 5B,

163 Supplementary Table 10). Seventy-eight genes were shared between these two comparisons
164 (Supplementary Table 11). However, none of those genes were viral transcripts, unlike in the
165 high- versus low-activity bodyguard comparison. At the transcript level, we obtained 28 genes
166 with significant DTU between broodcare workers and high-activity bodyguards, and 40 DTU
167 genes between broodcare workers and low-activity bodyguards (Supplementary Table 2). Finally,
168 we compared the DEGs obtained in every comparison and found that most of the genes are
169 uniquely highly expressed in broodcare workers in comparison to low- and high-activity
170 bodyguards (Fig. 2D and Supplementary Fig. 3).

171 Again, we used GSEA (FDR < 0.05) to reveal molecular functions that differ between
172 bodyguard and broodcare workers (Fig. 5C, and Supplementary Table 12). We found DEGs
173 upregulated in broodcare workers significantly enriched for two gene ontology (GO) terms or
174 gene functional categories involving chitin binding and structural molecule activity. Another
175 three GO terms were enriched in bodyguards (Fig. 5C). We manually verified the GO annotation
176 of every DEG (Supplementary Tables 9 and 10) and found genes associated with juvenile
177 hormone (JH) (Fig. 5D), vitellogenin pathways (Fig. 5E), and major royal jelly protein 1-like
178 (Fig. 5F) highly expressed in broodcare workers. Our results also revealed the overexpression of
179 serine-threonine protein kinase genes (Fig. 5). Phylogenetic analysis of those kinase genes
180 showed they cluster within the GATA zinc finger domain-containing protein (GATA) (Fig. 5G),
181 and dual-specificity tyrosine-phosphorylation-regulated kinase (DYRK1A) (Fig. 5H) clades
182 (Supplementary Fig. 3). Finally, we also identified genes known to be involved in insect
183 immunity against viruses that were differentially expressed between broodcare workers and
184 bodyguards (Fig. 4I-J).
185

186

187 **Discussion**

188 Using field behavioral assays and RNA-Seq, we explored gene expression associated with
189 bodyguarding behavior in *Allomerus octoarticulatus,* an obligately symbiotic ant species that
190 defends its hosts against herbivores. We found differentially expressed genes that either generate
191 or respond to variation in the quality of *A. octoarticulatus* colonies as plant mutualists. Many of
192 the DEGs were viral transcripts, suggesting that viral infections explain why some *A.*
193 *octoarticulatus* colonies are highly effective plant bodyguards, while others are slow to recruit to
194 herbivores, allowing damage to accumulate on their host plants.

195 The high expression of viral transcripts suggests that colonies had active viral infections
196 that modulated their bodyguarding behavior. Inactive bodyguards were infected by viruses
197 phylogenetically related to viruses causing poor locomotion and paralysis in Hymenoptera.
198 AOV-2, a picornavirus highly expressed in ineffective *A. octoarticulatus* bodyguards, is closely
199 related to viruses in the acute bee paralysis (ABP) viral complex, namely acute bee paralysis
200 virus (ABPV), Israeli acute paralysis virus (IAPV), and Kashmir bee virus (KBV). These viruses
201 are well-known and widespread honeybee RNA viruses that induce paralysis and increase bee
202 mortality (Amiri et al., 2019). We also identified AOV-1, a narna-like virus highly expressed in
203 low-activity bodyguards that belongs to the same family as another ant-associated virus,
204 *Linepithema humile* narna-like virus 1 (LhuNLV1), that has previously been described as
205 influencing *L. humile* immunity (Lester et al., 2019). These viruses might reduce ant foraging, as
206 previously observed in colonies of invasive fire ants infected with the dicistrovirus *Solenopsis*
207 *invicta* virus 1 (SINV-1) (Hsu et al., 2018). In other words, ineffective bodyguards might be poor
208 symbionts because they are sick.

209 However, we also found several viruses overexpressed in high-activity bodyguards
210 (AOV-3, AOV-4, AOV-5, AOV-6, AOV-7, and AOV-8), and these viruses could potentially be
211 inducing greater activity or aggression in their ant hosts to promote their own transmission. One
212 of the best-known examples of parasite manipulation of host behavior occurs in ("zombie") ants,
213 in which the fungal parasite *Ophiocordyceps* manipulates the behavior of its ant hosts to promote
214 its own onward transmission (de Bekker 2019). Similarly, Hsu et al. (2019) report aggressive
215 behavior in the ant *Anoplolepis gracilipes* being shaped by viral disease transmission triggered
216 by a dicistrovirus. In bees, iflaviruses such as Sacbrood virus and Deformed wing virus (DWV)
217 increase aggressiveness in infected colonies of *Apis cerana* (Hasegawa et al., 2023) and *A.*
218 *mellifera* (Fujiyuki et al., 2004; Terio et al., 2008, but see Rortais et al 2006). Thus, the viruses
219 that were highly expressed in very active *A. octoarticulatus* bodyguards may also be
220 manipulating the behavior of their ant hosts, with potential consequences for both their own
221 transmission and for the degree of protection afforded by the ants. The new viruses we identified
222 also increase the known viral diversity associated with ants (Baty et al., 2020).
223 Changes in aggressiveness might result from viral infections that trigger immune
224 responses in ant workers. As part of ants' humoral immunity, we found *hymenoptaecin* (*hym*)
225 genes highly expressed in low-activity bodyguards. Previous studies described high expression
226 of this infection-inducible antimicrobial peptide after ant workers are immune-challenged (Gupta
227 et al., 2015; Yek et al., 2013; Negroni et al., 2020; Von et al., 2016). Expression of *hym* genes in
228 ants has also been reported to be higher in foragers than in broodcare workers (Qui et al., 2017),
229 as the former are exposed to more viruses from spending more time outside the nest where they
230 may more regularly encounter viruses and their vectors. We also found higher expression of
231 genes encoding matrix metalloproteinases (MMPs) in ants that are poor plant bodyguards.

232 MMPs can be expressed in response to viral infections in Lepidoptera (Liu et al., 2022), and
233 participate in insect innate immune regulation (Vilcinskas, 2022). Finally, *chymotrypsin-1-like*, a
234 protease gene, was highly expressed in low-activity bodyguards. Proteolytic cascades are
235 pathogen-inducible enzymes participating in the immediate response of the innate immune
236 system in insects (Cerenius et al., 2010, Kanost & Jiang, 2015). These results further suggest that
237 low-activity workers might be diseased.

238 Comparing between bodyguards and broodcare workers in the same colony, we observed
239 patterns of task-specific gene expression that align with those previously reported in other social
240 insects (Buttstedt et al., 2014; Hawkings et al., 2019; Malé et al., 2017; Das & Bekker, 2022; Qiu
241 et al., 2022). We identified serine-threonine protein kinase (STK) genes over-expressed in
242 broodcare workers. Alleman et al. (2019) report STK genes involved in learning and memory
243 formation in ants. Previous research has also identified other kinase genes, especially cGMP-
244 dependent protein kinase (PKG) genes, that mediate behavioral shifts in ants (Lucas et al., 2009;
245 Ingram et al., 2011), including in *A. octoarticulatus* (Malé et al., 2017), although we did not find
246 PKGs among the DEGs in this study. Overall, some genes highly expressed in bodyguards also
247 follow previous reports of transcriptomic signatures of environmental stress in ants. For instance,
248 we found the accumulation of digestive enzymes in bodyguards, as previously reported in
249 foragers of other ant species exposed to temperature stress (Vatanparast et al., 2021).

250 Variation in gene expression has been associated with partner quality in pollination (Lüthi
251 et al., 2022) and insect-microbe interactions (Koga et al., 2022;), but it has not yet been explored
252 in defensive mutualisms. By combining field studies with transcriptomics, we report changes in
253 gene expression associated with how ants cooperate with plants and find that viral infections may
254 explain why some ants are better mutualists than others.

Methods

Field bioassays and sample collection

We measured bodyguard effectiveness in 60 naturally occurring colonies of *A. octoarticulatus* in the Peruvian Amazon. We collected live grasshoppers, almost entirely in the genus *Orphulella* (Supplementary Fig. 1), at our field site and attached these grasshoppers to *C. nodosa* leaves at a consistent distance to the domatium (i.e., nest) entrance using entomological pins (Supplementary Video 1). To measure ant recruitment, we counted the number of ants interacting with (i.e., stinging, biting, or walking on) the grasshopper each minute for five minutes and then averaged these counts to get a measure of bodyguarding activity (Fig. 1B, Supplementary Table 1). We also measured many factors that could potentially influence ant behavior in the bioassay, including time of day, ambient temperature, herbivore size, age, and species, and host plant size (i.e., number of domatia, a proxy for ant colony size, Frederickson et al. 2012), but found that they mostly had non-significant effects on ant behavior (Supplementary Fig. 1). Only time of day significantly predicted ant recruitment, with fewer ants recruiting to grasshoppers in the afternoon than the morning (Supplementary Fig. 1). We also measured herbivory on each plant in the field; three weeks before the bioassays, we marked up to three domatia per plant that had attached leaves with minimal herbivore damage, and then after conducting the bioassays three weeks later, we collected all the leaves from the marked domatia, grouped them by plant, and ranked the plants in order of herbivore damage, with a rank of 1 indicating the least herbivore damage and a rank of 60 indicating the most damage (Fig. 1C).

We collected bodyguards and broodcare workers from 14 colonies for transcriptomic analyses, evenly split between colonies with high- and low-activity bodyguards (Supplementary

278 Table 1). Both high- and low-activity colonies were sampled in both the morning and the
279 afternoon (Supplementary Fig. 1). Immediately after each bioassay, we collected 10-15
280 bodyguard workers on or near the grasshopper bait. We also collected up to 15 broodcare
281 workers from six of these ant colonies by opening domatia with secateurs and extracting workers
282 tending brood. We collected specimens in 1.5ml RNase-free tubes with 1ml of RNAlater
283 Stabilization Solution (Thermo Fisher Scientific), transported them back to the University of
284 Toronto (Toronto, Canada), and stored them at -80ºC until RNA extractions. We obtained permits
285 for field sampling and experiments from Peru's Servicio Nacional Forestal y de Fauna Silvestre
286 (SERFOR, authorization N° AUT-IFS-2018-061).
287
288 *RNA-Seq library construction and sequencing*
289 We submitted samples for RNA-Seq to the Donnelly Sequencing Center at the University of
290 Toronto (Toronto, Canada). The samples had a median RNA Integrity Number (RIN) of 8.3. The
291 sequencing facility processed 500 ng per sample using the NEBNext Ultra II Directional RNA
292 Library Prep Kit for Illumina, including PolyA selection using the NEBNext Poly(A) mRNA
293 Magnetic Isolation. Fragmentation was 15 minutes at 94 °C and samples were amplified for 9
294 cycles. The facility ran 1 µL top stock of each purified final library on an Agilent Bioanalyzer
295 dsDNA High Sensitivity chip. The libraries were quantified using the Quant-iT High-Sensitivity
296 dsDNA Assay Kit and pooled at equimolar ratios after size adjustment. The final pool was run on
297 an Agilent Bioanalyzer dsDNA High Sensitivity chip and quantified using a NEBNext Library
298 Quant Kit for Illumina. The quantified pool was hybridized at a final concentration of 2.15 pM
299 and sequenced with 2x50 bp reads on the Illumina NovaSeq 6000 platform using an SP flowcell,

300 generating an average of 28.2 million read pairs per sample. Read quality scores were 36 and
301 sequence length after trimming ranged from 30 to 51 bp (Supplementary Table 13).
302
303 De novo *transcriptome assembly, validation, and completeness*
304 Because there was no reference genome or transcriptome for any *Allomerus* species before our
305 study, we generated a *de novo* transcriptome assembly for *Allomerus octoarticulatus.* We
306 removed low-quality sequences and adapters using Trimmomatic v0.39 (Bolger et al., 2014), and
307 assessed the quality of the trimmed reads with FastQC v0.11.2 (Supplementary Table 13). Using
308 the rnaSPAdes pipeline (Bushmanova et al., 2019) and CD-HIT (Li & Godzik, 2006; Huang et
309 al., 2010), we assembled the high-quality RNA-Seq reads into a transcriptome with 67,613
310 unigenes and 93,122 transcripts. Then we assessed the quality of the transcriptome with
311 rnaQUAST (Bushmanova, et al., 2019). The assembled transcripts had an average length of 1478
312 bp, and the N50 was 3232 bp (Supplementary Table 14). Assessed with BUSCO (Manni et al.,
313 2021), the transcriptome had completeness scores of 95.62% (using the Hymenoptera database)
314 and 99.34% (using the Insecta database). We deposited the *de novo* transcriptome assembly at
315 GenBank (NCBI database) under the submission number (available upon acceptance:
316 XXXXXXXXXX).
317
318 *Differential expression analysis*
319 For transcript quantification, we implemented the Salmon v1.9.0 pipeline (Patro et al., 2017). We
320 built a Salmon index based on a decoy-aware transcriptome to avoid false mapping (Srivastava et
321 al., 2020). We used the genome of a close relative of *A. octoarticulatus, Wasmannia*
322 *auropunctata* (GenBank: GCA_000956235.1) as the 'decoy.' Then we aligned the RNA-Seq

323 reads to the Salmon index. In this step, we implemented the Gibbs sampling option (--
324 numGibbsSamples 20) and reduced autocorrelation in the Gibbs samples (thinningFactor 100).
325 Finally, we obtained transcript-level counts for each sample with tximport (Soneson et al., 2015).
326 To identify DEGs, we used the DESeq2 workflow (Love et al., 2014) in v4.2.1 (R Core
327 Team, 2022). We extracted the final list of significant genes with the *results* function and set a
328 false discovery rate (FDR) < 0.01 and fold change (FD) > 1.5. We made heatmaps with the R
329 library *pheatmap.*
330 For the phylogenetic tree of serine-threonine protein kinase genes, we obtained
331 Formicidae orthologues of the GATA zinc finger domain-containing proteins (OrthoDB Group
332 32050at36668) and dual specificity tyrosine-phosphorylation-regulated kinases (OrthoDB Group
333 21499at36668). We also downloaded PKG orthologues for ants using OrthoDB (Group
334 EOG8PP0WX) and aligned those genes to PKG-encoding genes characterized in *A.*
335 *octoarticulatus*: *Aofor* (GenBank KX809576) and *Aopkg* (GenBank KX809574). We aligned
336 sequences using MAFFT v7.505 (Katoh, et al., 2005) with the algorithm L-INS-I and the default
337 scoring matrix (BLOSUM62) and gap penalty (1.53) values. We then generated a maximum
338 likelihood tree based on amino-acid sequences with IQ-Tree v1.6.12 software (Nguyen et al.,
339 2015) The amino acid substitution model implemented was JTT+F+G4 (Kalyaanamoorthy, et al.,
340 2017), and we evaluated node support with 1000 bootstrap replicates (Hoang, et al., 2017).
341
342 *Differential transcript usage analysis*
343 We implemented the Dirichlet-multinomial framework with the R library *DRIMSeq* (Nowicka &
344 Robinson, 2016) following Love et al (2018). First, we scaled transcripts-per-million (TPM)
345 abundance estimates per sample with the *scaledTPM* method to generate a final matrix with

346 count-scale data. Along with that matrix, we imported to *DRIMSeq* a transcript-to-gene mappings
347 file to create a *dmDSdata* object containing sample and count information. We filtered data to
348 minimal expression and proportion levels of 10 and 0.1, respectively. We used the functions
349 *dmPrecision* and *dmFit* for the precision and proportion estimations, and *dmTest* to test for DTU.
350 Finally, we performed two-stage testing with the R library *stageR* (Van den Berge et al., 2017) by
351 implementing screening and confirmation stages to identify the final list of genes participating in
352 DTU. For the latter stage, we used a target overall false discovery rate (OFDR) level of 5%
353 (alpha = 0.05).
354
355 *Functional annotation and enrichment analysis*
356 For annotation of the *de novo* reference transcriptome, we used the blastx function in Blast
357 v2.11.0 (Camacho et al., 2009). We used the *de novo* transcriptome as a query against a high-
358 quality manually annotated and non-redundant protein sequence database -UniProtKB (UniProt
359 Consortium, 2021), with an e-value cut-off of 1e-5. We then loaded this XML file into the
360 OmicsBox program v2.1 (OmicsBox, 2019) to perform GO mapping and annotation (Gotz et al.,
361 2018) against the Gene Ontology database implemented in the program. For GO assignment, we
362 set the annotation cut-off to 55 and the e-value hit filter to 1.0E-6. To obtain a more complete
363 functional annotation for the subset of DE genes obtained in the DEG and DTU analyses, we
364 used the software package InterProScan (Jones et al., 2014) in addition to BLAST (Altschul et
365 al., 1990) in OmicsBox. We implemented these annotation steps to allow for the sequences to be
366 screened against more databases. We performed sequence annotation with the public EMBL-EBI
367 InterPro web service by loading the fasta sequences and running the analysis under the default
368 setting. We then ran a blast search against the non-redundant NCBI database with an e-value of

369 1.0E-5. We then used the same previously described approach for GO annotation and mapping of
370 the DE genes. We also implemented Gene Set Enrichment Analysis (GSEA) in OmicsBox using
371 an FDR value of 0.05. Finally, we retrieved Kyoto Encyclopedia of Genes and Genomes
372 (KEGG) pathways (Kanehisa, 2000) by annotating and predicting orthologues with the tool
373 EggNOG-Mapper (Huerta-Cepas et al., 2019) in OmicsBox.
374
375 *Virus characterization and phylogenetic analysis*
376 We identified genes with GO terms corresponding to viral-related processes. The genes were
377 manually annotated by aligning them against the NCBI non-redundant protein sequence database
378 using blastx. We found eleven sequences of eight genes from picornavirales and Naranavirus that
379 were homologous to the viral genes we identified. We further extracted potential open reading
380 frames (ORF) and conserved domains by using those genes as queries on NCBI's ORF Finder
381 and Conserved Domain Database (CDD), respectively.
382 For phylogenetic reconstruction, we used the full-length amino acid sequences of the
383 non-structural RNA-dependent RNA polymerase (RdRp), a highly conserved gene suitable for
384 phylogenetic reconstructions of viruses (Venkataraman et al., 2018). We extracted RdRp
385 sequences from genes identified as viral from CDD. Unclassified and classified representative
386 amino acid sequences of species of each family were used in the phylogeny as suggested by
387 Xavier et al. (2021). We also included viral sequences with the most similar matches based on
388 the blastx search. We used noumavirus and three species in the family Secoviridae as outgroups
389 for the phylogenetic placement of picorna-like viruses and narna-like viruses, respectively. We
390 aligned viral protein sequences using MAFFT v7.505 (Katoh, et al., 2005). We generated a
391 maximum likelihood phylogenetic tree based on amino acid sequences with the software

program IQ-Tree v1.6.12 software (Nguyen et al., 2015), and the amino acid substitution model Blosum62+R5 (Henikoff and Henikoff, 1992). Finally, we evaluated the support for each node with 1000 bootstrap replicates.

Acknowledgments

The authors thank the Centro de Investigación y Capacitación Río Los Amigos (CICRA) for their support of the field research, especially station manager Arianna Basto Eyzaguirre. MEF acknowledges funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant program, CR from a National Geographic Early Career Grant, a Sigma Xi Grant in Aid of Research, and a University of Toronto Centre for Global Change Science Graduate Scholarship, and MCT from a Vanier Canada Graduate Scholarship. We also thank Alberto Escudero for his help with fieldwork, Ina Anreiter for helping with RNA extractions, Kyle Turner and the Donnelly Sequencing Center for sequencing samples, Santiago Sanchez and Pierre-Jean Malé for bioinformatic support, Gianpiero Fiorentino and Patricio Picon for help with graphical design, and Naomi Pierce for feedback that improved the manuscript.

Author contributions

MCT, CR, JH, and MEF designed and performed the research. MCT, CR, HX, and MEF analyzed the data. MCT wrote the article with input from all the authors.

Declaration of interests

The authors declare no competing interests.

415 **Data availability**

416 Lists of DEGs and DTUs are available in the supplementary information. Raw sequencing and
417 transcriptome assembly data will be deposited in NCBI databases. Field data and code will be
418 archived on Dryad and Zenodo, respectively, on acceptance, as well as in the GitHub repository:
419 https://github.com/mariatocora/Transcriptomic-analysis-ant-plant.

420

421 **References**

422

423 1. A. Alleman, M. Stoldt, B. Feldmeyer, S. Foitzik, Tandem-running and scouting
424 behaviour are characterized by up-regulation of learning and memory formation genes
425 within the ant brain. *Molecular Ecology* **28**, 2342–2359 (2019).

426 1. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool.
427 *Journal of Molecular Biology* **215**, 403–410 (1990).

428 2. E. Amiri, *et al.*, Israeli Acute Paralysis Virus: Honey Bee Queen–Worker Interaction and
429 Potential Virus Transmission Pathways. *Insects* **10**, 9 (2019).

430 3. R. T. Batstone, L. T. Burghardt, K. D. Heath, Phenotypic and genomic signatures of interspecies
431 cooperation and conflict in naturally occurring isolates of a model plant symbiont. *Proc. R. Soc.*
432 *B.* **289**, 20220477 (2022).

433 4. R. T. Batstone, A. M. O'Brien, T. L. Harrison, M. E. Frederickson, Experimental evolution
434 makes microbes more cooperative with their local host genotype. *Science* **370**, 476–478 (2020).

435 5. BioBam Bioinformatics, OmicsBox – Bioinformatics Made Easy, BioBam (2019).

436 6. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data.
437 *Bioinformatics* **30**, 2114–2120 (2014).

438 7. Bronstein, JL, *Mutualism.* (Oxford University Press, New York., 2015).

439 8. E. Bushmanova, D. Antipov, A. Lapidus, A. D. Prjibelski, rnaSPAdes: a de novo transcriptome
440 assembler and its application to RNA-Seq data. *GigaScience* **8**, giz100 (2019).
441 9. A. Buttstedt, R. F. A. Moritz, S. Erler, Origin and function of the major royal jelly proteins of the
442 honeybee (*Apis mellifera*) as members of the *yellow* gene family: Honeybee's major royals. *Biol*
443 *Rev* **89**, 255–269 (2014).
444 10. C. Camacho, *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
445 11. L. Cerenius, S. Kawabata, B. L. Lee, M. Nonaka, K. Söderhäll, Proteolytic cascades and their
446 involvement in invertebrate immunity. *Trends in Biochemical Sciences* **35**, 575–583 (2010).
447 12. B. Das, C. De Bekker, Time-course RNASeq of Camponotus floridanus forager and nurse ant
448 brains indicate links between plasticity in the biological clock and behavioral division of labor.
449 *BMC Genomics* **23**, 57 (2022).
450 13. C. De Bekker, Ophiocordyceps–ant interactions as an integrative model to understand the
451 molecular basis of parasitic behavioral manipulation. *Current Opinion in Insect Science* **33**, 19–
452 24 (2019).
453 14. de Bekker, C, Will, I, Das, B, Adams, R.M.M, The ants (Hymenoptera: Formicidae) and their
454 parasites: effects of parasitic manipulations and host responses on ant behavioral ecology. **28**, 1–
455 24 (2018).
456 15. B. Epstein, *et al.*, Combining GWAS and population genomic analyses to characterize
457 coevolution in a legume-rhizobia symbiosis. *Molecular Ecology* **32**, 3798–3811 (2023).
458 16. M. E. Frederickson, Rethinking Mutualism Stability: Cheaters and the Evolution of Sanctions.
459 *The Quarterly Review of Biology* **88**, 269–295 (2013).
460 17. M. E. Frederickson, *et al.*, The Direct and Ecological Costs of an Ant-Plant Symbiosis. *The*
461 *American Naturalist* **179**, 768–778 (2012).
462 18. M. L. Friesen, Widespread fitness alignment in the legume–rhizobium symbiosis. *New*
463 *Phytologist* **194**, 1096–1111 (2012).

464 19. T. Fujiyuki, *et al.*, Novel Insect Picorna-Like Virus Identified in the Brains of Aggressive Worker
465 Honeybees. *J Virol* **78**, 1093–1100 (2004).
466 20. S. D. Gillespie, L. S. Adler, Indirect effects on mutualisms: parasitism of bumble bees and
467 pollination service to plants. *Ecology* **94**, 454–464 (2013).
468 21. S. Gotz, *et al.*, High-throughput functional annotation and data mining with the Blast2GO suite.
469 *Nucleic Acids Research* **36**, 3420–3435 (2008).
470 22. S. K. Gupta, *et al.*, Scrutinizing the immune defence inventory of Camponotus floridanus
471 applying total transcriptome sequencing. *BMC Genomics* **16**, 540 (2015).
472 23. N. Hasegawa, *et al.*, Evolutionarily diverse origins of deformed wing viruses in western honey
473 bees. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2301258120 (2023).
474 24. C. Hawkings, T. L. Calkins, P. V. Pietrantonio, C. Tamborindeguy, Caste-based differential
475 transcriptional expression of hexamerins in response to a juvenile hormone analog in the red
476 imported fire ant (Solenopsis invicta). *PLoS ONE* **14**, e0216800 (2019).
477 25. S. Henikoff, J. G. Henikoff, Amino acid substitution matrices from protein blocks. *Proc. Natl.*
478 *Acad. Sci. U.S.A.* **89**, 10915–10919 (1992).
479 26. D. T. Hoang, O. Chernomor, A. Von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the
480 Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**, 518–522 (2018).
481 27. H.-W. Hsu, M.-C. Chiu, C.-C. Lee, C.-Y. Lee, C.-C. S. Yang, The Association between Virus
482 Prevalence and Intercolonial Aggression Levels in the Yellow Crazy Ant, Anoplolepis Gracilipes
483 (Jerdon). *Insects* **10**, 436 (2019).
484 28. H.-W. Hsu, M.-C. Chiu, D. Shoemaker, C.-C. S. Yang, Viral infections in fire ants lead to
485 reduced foraging activity and dietary changes. *Sci Rep* **8**, 13498 (2018).
486 29. Y. Huang, B. Niu, Y. Gao, L. Fu, W. Li, CD-HIT Suite: a web server for clustering and
487 comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).

488 30. J. Huerta-Cepas, *et al.*, eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated
489 orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research* **47**,
490 D309–D314 (2019).
491 31. K. K. Ingram, L. Kleeman, S. Peteru, Differential regulation of the foraging gene associated with
492 task behaviors in harvester ants. *BMC Ecol* **11**, 19 (2011).
493 32. James W. Baty, Mariana Bulgarella, Jana Dobelmann, Antoine Felden, Philip J. Lester, Viruses
494 and their effects in ants (Hymenoptera: Formicidae) *Myrmecol News* 30, 213-228 (2020)
495 https:/doi.org/10.25849/MYRMECOL.NEWS_030:213 (February 12, 2024).
496 33. B. M. Jones, *et al.*, Individual differences in honey bee behavior enabled by plasticity in brain
497 gene regulatory networks. *eLife* **9**, e62850 (2020).
498 34. P. Jones, *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**,
499 1236–1240 (2014).
500 35. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. Von Haeseler, L. S. Jermiin, ModelFinder:
501 fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587–589 (2017).
502 36. M. Kanehisa, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**,
503 27–30 (2000).
504 37. M. R. Kanost, H. Jiang, Clip-domain serine proteases as immune factors in insect hemolymph.
505 *Current Opinion in Insect Science* **11**, 47–55 (2015).
506 38. K. Katoh, MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic*
507 *Acids Research* **33**, 511–518 (2005).
508 39. E. T. Kiers, M. G. Hutton, R. F. Denison, Human selection and the relaxation of legume defences
509 against ineffective rhizobia. *Proc. R. Soc. B.* **274**, 3119–3126 (2007).
510 40. R. Koga, *et al.*, Single mutation makes Escherichia coli an insect mutualist. *Nat Microbiol* **7**,
511 1141–1150 (2022).

512 41. P. Kohlmeier, A. R. Alleman, R. Libbrecht, S. Foitzik, B. Feldmeyer, Gene expression is more
513 strongly associated with behavioural specialization than with age or fertility in ant workers.
514 *Molecular Ecology* **28**, 658–670 (2019).
515 42. P. J. Lester, K. H. Buick, J. W. Baty, A. Felden, J. Haywood, Different bacterial and viral
516 pathogens trigger distinct immune responses in a globally invasive ant. *Sci Rep* **9**, 5780 (2019).
517 43. W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or
518 nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
519 44. T. Liu, *et al.*, The dual roles of three MMPs and TIMP in innate immunity and metamorphosis in
520 the silkworm, *Bombyx mori*. *The FEBS Journal* **289**, 2828–2846 (2022).
521 45. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-
522 seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
523 46. M. I. Love, C. Soneson, R. Patro, Swimming downstream: statistical analysis of differential
524 transcript usage following Salmon quantification. *F1000Res* **7**, 952 (2018).
525 47. C. Lucas, M. B. Sokolowski, Molecular basis for changes in behavioral state in ant social
526 behaviors. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6351–6356 (2009).
527 48. M. N. Lüthi, A. E. Berardi, T. Mandel, L. B. Freitas, C. Kuhlemeier, Single gene mutation in a
528 plant MYB transcription factor causes a major shift in pollinator preference. *Current Biology* **32**,
529 5295-5308.e5 (2022).
530 49. P.-J. G. Malé, *et al.*, An ant–plant mutualism through the lens of cGMP-dependent kinase genes.
531 *Proc. R. Soc. B.* **284**, 20170896 (2017).
532 50. F. Manfredini, *et al.*, Sociogenomics of Cooperation and Conflict during Colony Founding in the
533 Fire Ant Solenopsis invicta. *PLoS Genet* **9**, e1003633 (2013).
534 51. M. Manni, M. R. Berkeley, M. Seppey, E. M. Zdobnov, BUSCO: Assessing Genomic Data
535 Quality and Beyond. *Current Protocols* **1**, e323 (2021).
536 52. V. E. Mayer, M. E. Frederickson, D. McKey, R. Blatrix, Current issues in the evolutionary
537 ecology of ant–plant symbioses. *New Phytologist* **202**, 749–764 (2014).

538 53. A. T. Müller, *et al.*, Combined –omics framework reveals how ant symbionts benefit the
539 Neotropical ant-plant Tococa quadrialata at different levels. *iScience* **25**, 105261 (2022).
540 54. M. A. Negroni, F. H. I. D. Segers, F. Vogelweith, S. Foitzik, Immune challenge reduces gut
541 microbial diversity and triggers fertility-dependent gene expression changes in a social insect.
542 *BMC Genomics* **21**, 816 (2020).
543 55. L.-T. Nguyen, H. A. Schmidt, A. Von Haeseler, B. Q. Minh, IQ-TREE: A Fast and Effective
544 Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and*
545 *Evolution* **32**, 268–274 (2015).
546 56. M. Nowicka, M. D. Robinson, DRIMSeq: a Dirichlet-multinomial framework for multivariate
547 count outcomes in genomics. *F1000Res* **5**, 1356 (2016).
548 57. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-
549 aware quantification of transcript expression. *Nat Methods* **14**, 417–419 (2017).
550 58. B. Qiu, *et al.*, Canalized gene expression during development mediates caste differentiation in
551 ants. *Nat Ecol Evol* **6**, 1753–1765 (2022).
552 59. R Core Team, R: A Language and Environment for Statistical Computing (2017).
553 60. C. Riehl, M. E. Frederickson, Cheating and punishment in cooperative animal societies. *Phil.*
554 *Trans. R. Soc. B* **371**, 20150090 (2016).
555 61. A. Rortais, *et al.*, Deformed wing virus is not related to honey bees' aggressiveness. *Virol J* **3**, 61
556 (2006).
557 62. B. E. R. Rubin, C. S. Moreau, Comparative genomics reveals convergent rates of evolution in
558 ant–plant mutualisms. *Nat Commun* **7**, 12679 (2016).
559 63. J. Sachs, E. Simms, Pathways to mutualism breakdown. *Trends in Ecology & Evolution* **21**, 585–
560 592 (2006).
561 64. C. Soneson, M. I. Love, M. D. Robinson, Differential analyses for RNA-seq: transcript-level
562 estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015).

563 65. A. Srivastava, *et al.*, Alignment and mapping methodology influence transcript abundance
564 estimation. *Genome Biol* **21**, 239 (2020).
565 66. V. Terio, *et al.*, Detection of a honeybee iflavirus with intermediate characteristics between
566 kakugo virus and deformed wing virus. *New Microbiol* **31**, 439–444 (2008).
567 67. The UniProt Consortium, *et al.*, UniProt: the universal protein knowledgebase in 2021. *Nucleic*
568 *Acids Research* **49**, D480–D489 (2021).
569 68. K. Van Den Berge, C. Soneson, M. D. Robinson, L. Clement, stageR: a general stage-wise
570 method for controlling the gene-level false discovery rate in differential expression and
571 differential transcript usage. *Genome Biol* **18**, 151 (2017).
572 69. M. Vatanparast, Y. Park, Comparative RNA-Seq Analyses of Solenopsis japonica (Hymenoptera:
573 Formicidae) Reveal Gene in Response to Cold Stress. *Genes* **12**, 1610 (2021).
574 70. S. Venkataraman, B. Prasad, R. Selvarajan, RNA Dependent RNA Polymerases: Insights from
575 Structure, Function and Evolution. *Viruses* **10**, 76 (2018).
576 71. A. Vilcinskas, Matrix metalloproteinases and their inhibitors – pleiotropic functions in insect
577 immunity and metamorphosis. *The FEBS Journal* **289**, 2805–2808 (2022).
578 72. K. Von Wyschetzki, H. Lowack, J. Heinze, Transcriptomic response to injury sheds light on the
579 physiological costs of reproduction in ant queens. *Mol Ecol* **25**, 1972–1985 (2016).
580 73. M. G. Weber, A. A. Agrawal, Defense mutualisms enhance plant diversification. *Proc. Natl.*
581 *Acad. Sci. U.S.A.* **111**, 16442–16447 (2014).
582 74. M. G. Weber, K. H. Keeler, The phylogenetic distribution of extrafloral nectaries in plants.
583 *Annals of Botany* **111**, 1251–1261 (2013).
584 75. C. A. D. Xavier, M. L. Allen, A. E. Whitfield, Ever-increasing viral diversity associated with the
585 red imported fire ant Solenopsis invicta (Formicidae: Hymenoptera). *Virol J* **18**, 5 (2021).
586 **76.** S. H. Yek, J. J. Boomsma, M. Schiøtt, Differential gene expression in *Acromyrmex* leaf-cutting
587 ants after challenges with two fungal pathogens. *Mol Ecol* **22**, 2173–2187 (2013)
588

589 **Figures**

590 **Figure 1. A.** Summary of a typical ant-plant defensive mutualism. (1) Host plants produce
591 specialized structures (domatia) that (2) attract ants that benefit from the food and shelter the
592 plant provides. (3) As they forage for food, ant 'bodyguards' defend the plant against attacking
593 enemies (i.e., herbivores, pathogens, parasites). (4) This reduces the frequency and duration of
594 herbivore visits to the host plant, and (5) ultimately benefits the plants by reducing tissue damage
595 caused by herbivory. **B.** Spaghetti plot of behavioral trials across all 60 colonies. Blue and pink
596 lines show colonies used for RNA-Seq. Grey lines indicate additional colonies that were
597 measured but not used in transcriptomic analyses. **C.** Linear model of herbivory vs. bodyguard
598 activity. Herbivory was measured by ranking plants based on estimated leaf area missing after
599 three weeks in the field (here, higher numbers represent greater herbivory). **D.** Herbivory on
600 plants occupied by the low-activity (pink) and high-activity (blue) bodyguards used for RNA-
601 Seq.

602

603

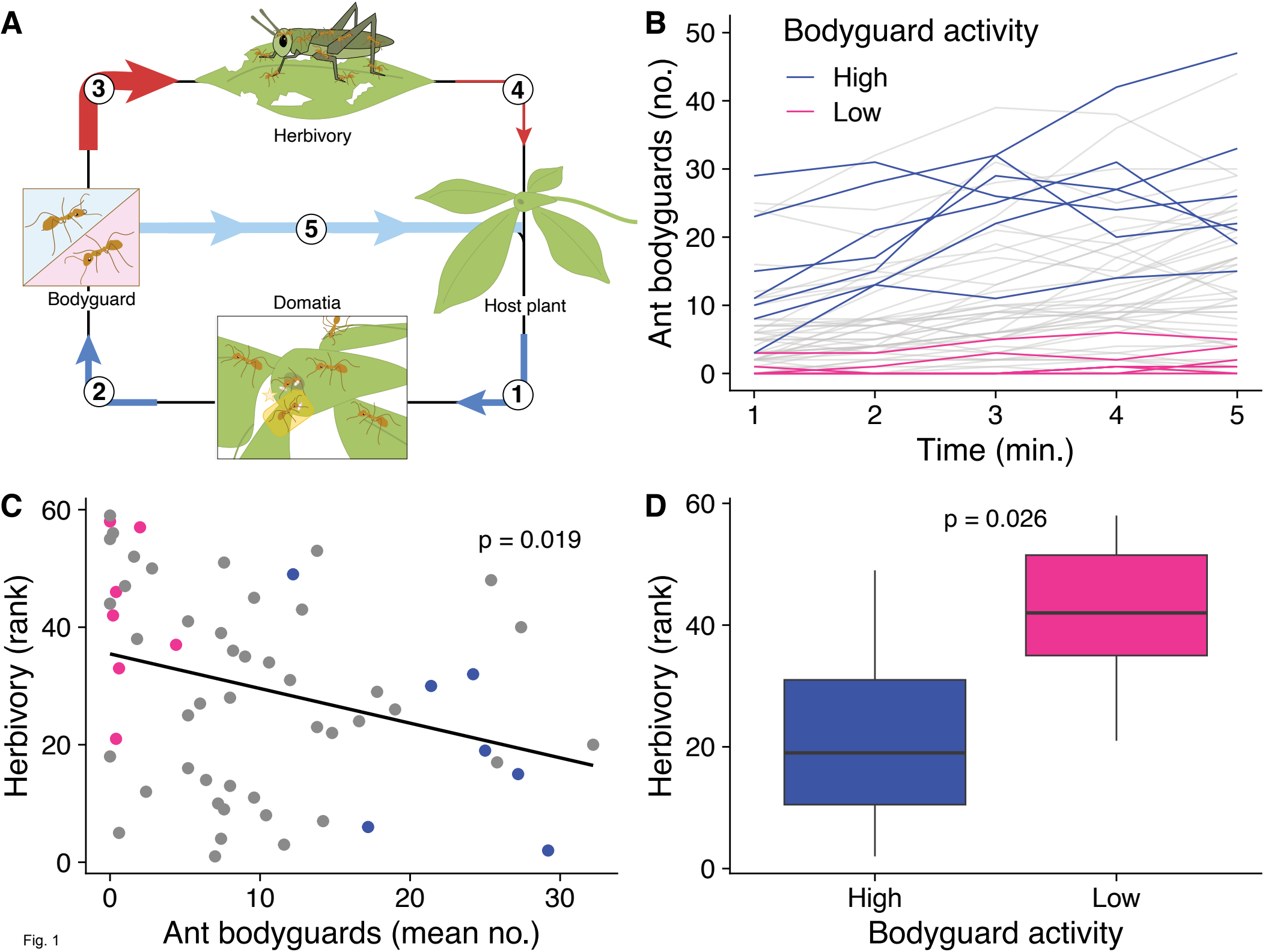
A
Herbivory
Bodyguard
Domatia
Host plant
1
2
3
4
5
B
Bodyguard activity
High
Low
Ant bodyguards (no.)
Time (min.)
C
p = 0.019
Herbivory (rank)
Ant bodyguards (mean no.)
D
p = 0.026
Herbivory (rank)
High
Low
Bodyguard activity

Fig. 1

604 **Figure 2. A.** Heatmap showing 40 DEGs, with 15 genes upregulated in low-activity bodyguards
605 and 25 genes upregulated in high-activity bodyguards. **B.** Kinase and **C.** protease genes highly
606 expressed in high and low-activity bodyguards, respectively. **D.** Venn diagram illustrating DEGs
607 shared among workers doing different tasks. Most genes are uniquely highly expressed in
608 broodcare workers.

609

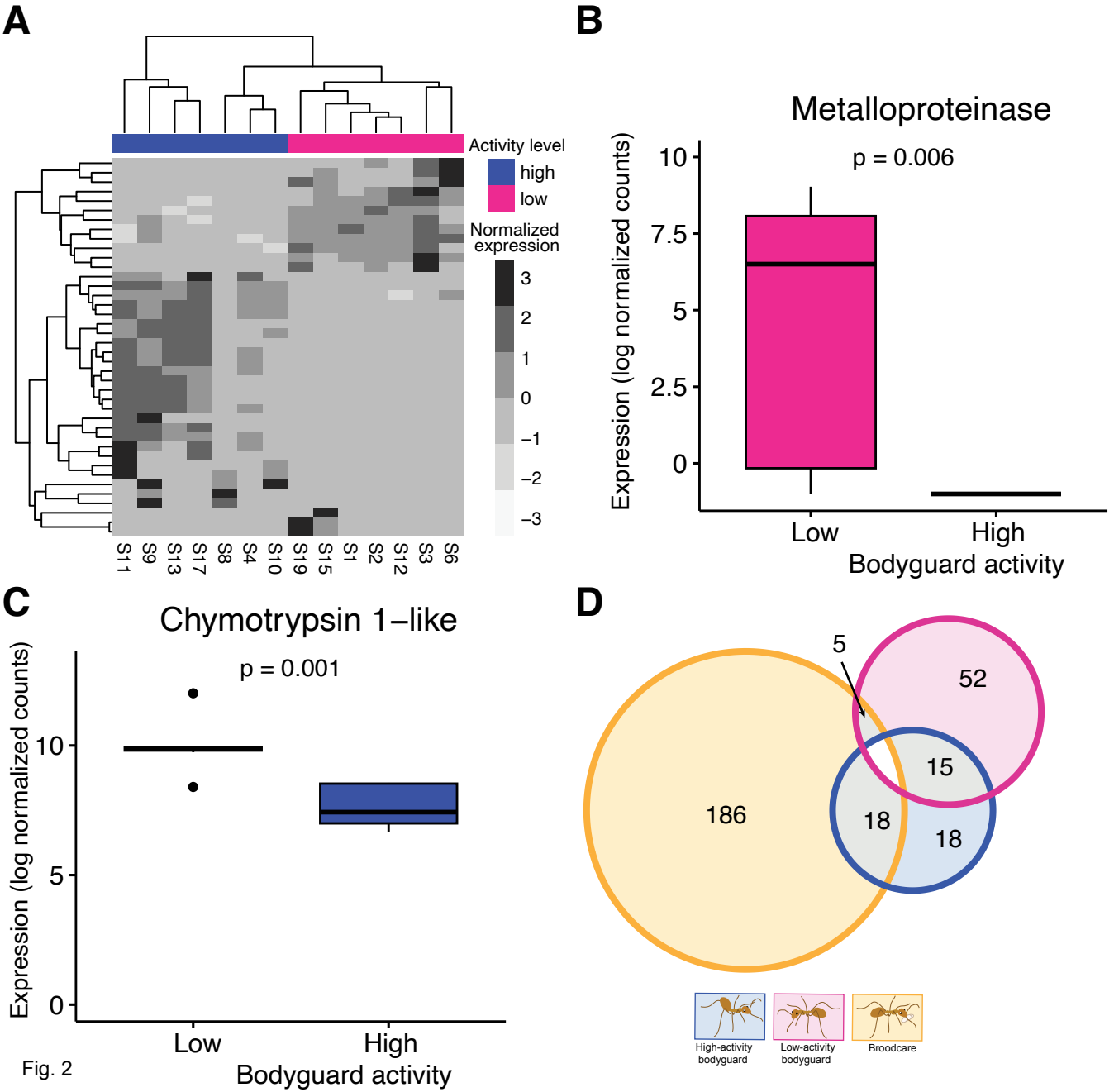
A
Activity level
high
low
Normalized expression
3
2
1
0
−1
−2
−3
S11
S9
S13
S17
S8
S4
S10
S19
S15
S1
S2
S12
S3
S6
B
Metalloproteinase
p = 0.006
Expression (log normalized counts)
10
7.5
5
2.5
0
Low
High
Bodyguard activity
C
Chymotrypsin 1−like
p = 0.001
Expression (log normalized counts)
10
5
0
Low
High
Bodyguard activity
D
5
52
15
186
18
18
High-activity bodyguard
Low-activity bodyguard
Broodcare

Fig. 2

610 **Figure 3.** Phylogenetic reconstructions of new viruses associated with *A. octoarticulatus* based
611 on comparisons of RNA-dependent RNA polymerase (RdRp) sequences of **A.** a narna-like virus
612 1 (AOV-1, in bold) from *A. octoarticulatus* and representative species within the family
613 Narnaviridae (order Wolframvirales), and **B**. other viruses (AOV-2 to AOV-8, in bold) from *A.*
614 *octoarticulatus*, and representative species within the families Dicistroviridae, Picornaviridae,
615 and Iflaviridae in the order Picornavirales. Three species of the family Secoviridae and an
616 oumavirus are the outgroups of the trees in **B** and **A**, respectively. Ultrafast bootstrap (ufBS)
617 support values are indicated by colored circles on respective nodes.
618
619
620

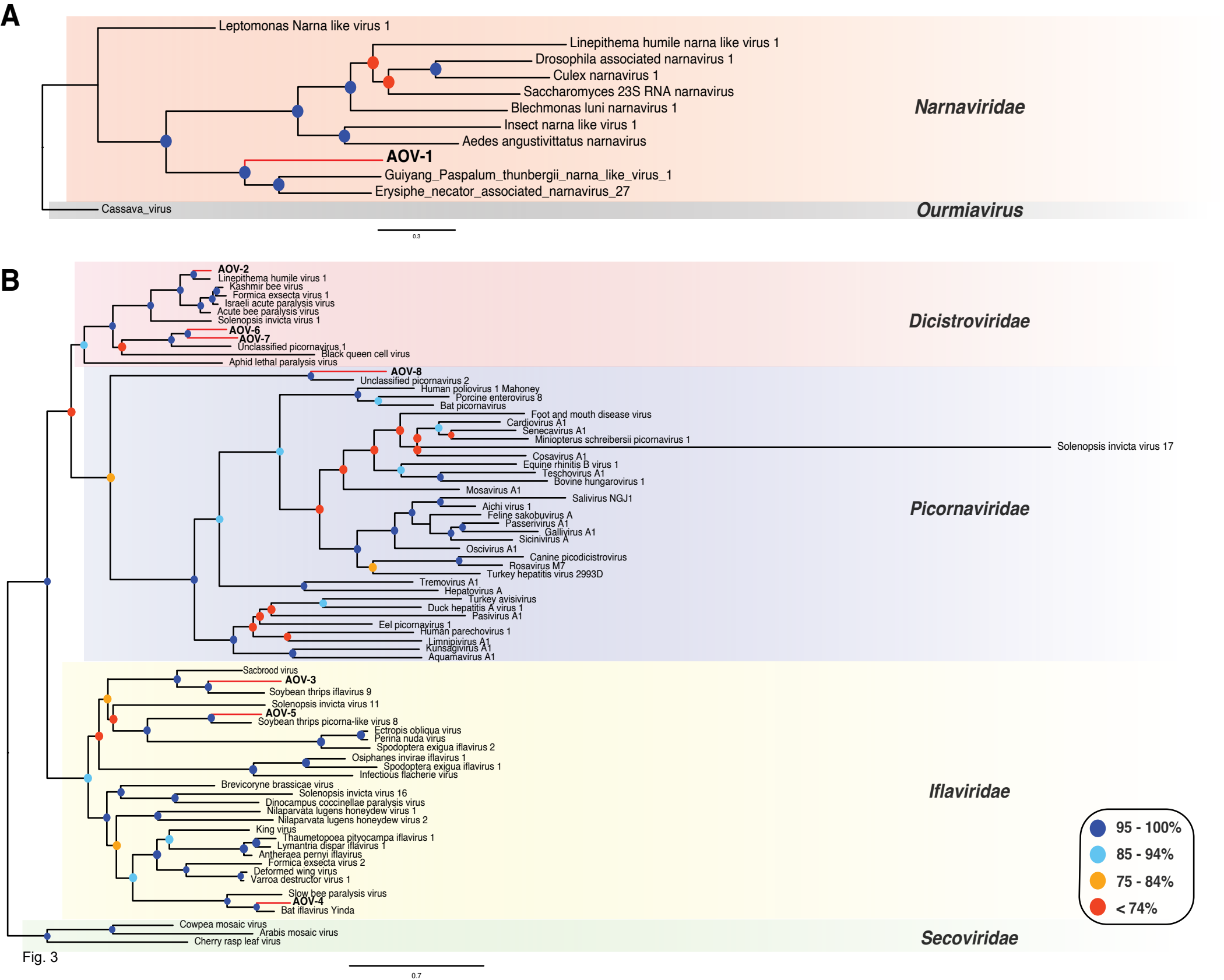
A
Leptomonas Narna like virus 1
Linepithema humile narna like virus 1
Drosophila associated narnavirus 1
Culex narnavirus 1
Saccharomyces 23S RNA narnavirus
Blechmonas luni narnavirus 1
Insect narna like virus 1
Aedes angustivittatus narnavirus
AOV-1
Guiyang_Paspalum_thunbergii_narna_like_virus_1
Erysiphe_necator_associated_narnavirus_27
Cassava_virus
Narnaviridae
Ourmiavirus
0.3
B
AOV-2
Linepithema humile virus 1
Kashmir bee virus
Formica exsecta virus 1
Israeli acute paralysis virus
Acute bee paralysis virus
Solenopsis invicta virus 1
AOV-6
AOV-7
Unclassified picornavirus 1
Black queen cell virus
Aphid lethal paralysis virus
AOV-8
Unclassified picornavirus 2
Human poliovirus 1 Mahoney
Porcine enterovirus 8
Bat picornavirus
Foot and mouth disease virus
Cardiovirus A1
Senecavirus A1
Miniopterus schreibersii picornavirus 1
Solenopsis invicta virus 17
Cosavirus A1
Equine rhinitis B virus 1
Teschovirus A1
Bovine hungarovirus 1
Mosavirus A1
Salivirus NGJ1
Aichi virus 1
Feline sakobuvirus A
Passerivirus A1
Gallivirus A1
Sicinivirus A
Oscivirus A1
Canine picodicistrovirus
Rosavirus M7
Turkey hepatitis virus 2993D
Tremovirus A1
Hepatovirus A
Turkey avisivirus
Duck hepatitis A virus 1
Pasivirus A1
Eel picornavirus 1
Human parechovirus 1
Limnipivirus A1
Kunsagivirus A1
Aquamavirus A1
Sacbrood virus
AOV-3
Soybean thrips iflavirus 9
Solenopsis invicta virus 11
AOV-5
Soybean thrips picorna-like virus 8
Ectropis obliqua virus
Perina nuda virus
Spodoptera exigua iflavirus 2
Osiphanes invirae iflavirus 1
Spodoptera exigua iflavirus 1
Infectious flacherie virus
Brevicoryne brassicae virus
Solenopsis invicta virus 16
Dinocampus coccinellae paralysis virus
Nilaparvata lugens honeydew virus 1
Nilaparvata lugens honeydew virus 2
King virus
Thaumetopoea pityocampa iflavirus 1
Lymantria dispar iflavirus 1
Antheraea pernyi iflavirus
Formica exsecta virus 2
Deformed wing virus
Varroa destructor virus 1
Slow bee paralysis virus
AOV-4
Bat iflavirus Yinda
Cowpea mosaic virus
Arabis mosaic virus
Cherry rasp leaf virus
Dicistroviridae
Picornaviridae
Iflaviridae
Secoviridae
0.7
95 - 100%
85 - 94%
75 - 84%
< 74%

Fig. 3

621 **Figure 4.** Differences in viral transcript expression between high- and low-activity bodyguards.
622 **A**. The Narnavirus: *A. octoarticulatus* virus 1 (AOV-1). **B.** A Picorna-like virus highly expressed
623 in low-activity bodyguards: *A. octoarticulatus* virus 2 (AOV-2). **C-H.** Picorna-like viruses highly
624 expressed in high-activity bodyguards: *A. octoarticulatus* virus 3 (AOV-3), *A. octoarticulatus*
625 virus 4 (AOV-4), *A. octoarticulatus* virus 5 (AOV-5), *A. octoarticulatus* virus 6 (AOV-6), and *A.*
626 *octoarticulatus* virus 7 (AOV-7), and *A. octoarticulatus* virus 8 (AOV-8). **I-J.** Expression of two
627 immunity-related genes among broodcare workers, low-activity bodyguards, and high-activity
628 bodyguards.
629
630

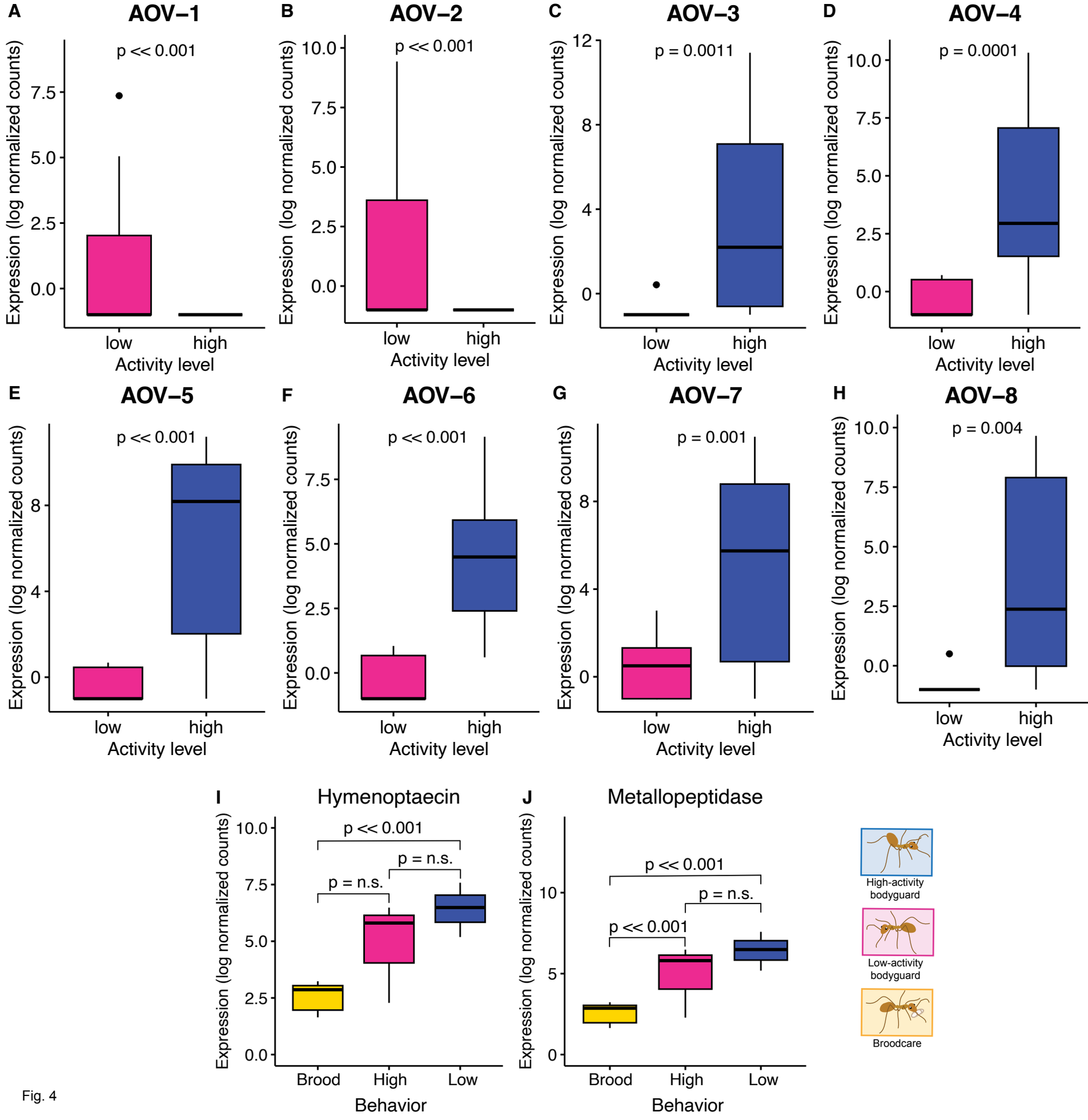
A
AOV−1
p << 0.001
B
AOV−2
p << 0.001
C
AOV−3
p = 0.0011
D
AOV−4
p = 0.0001
E
AOV−5
p << 0.001
F
AOV−6
p << 0.001
G
AOV−7
p = 0.001
H
AOV−8
p = 0.004
Expression (log normalized counts)
low
high
Activity level
I
Hymenoptaecin
p << 0.001
p = n.s.
p = n.s.
J
Metallopeptidase
p << 0.001
p = n.s.
p << 0.001
Brood
High
Low
Behavior
High-activity bodyguard
Low-activity bodyguard
Broodcare

Fig. 4

631 **Figure 5. A.** Heatmap showing 155 DEGs between broodcare workers (133 genes upregulated)
632 and high-activity bodyguards (22 genes upregulated). **B.** Heatmap comparing 197 DEGs between
633 broodcare workers (144 genes upregulated) and low-activity bodyguards (53 genes upregulated).
634 **C.** Gene Set Enrichment Analysis (GSEA) results show three GO terms (in bold) overrepresented
635 in bodyguards, and two in broodcare workers (FDR < 0.05). Ontologies: Biological Process
636 (BP), Cellular Component (CC), and Molecular Function (MF). Gene number is number of genes
637 in gene set. **D.** A subset of genes differentially expressed between workers performing different
638 tasks, including juvenile hormone (JH) biosynthesis, vitellogenin, and major royal jelly protein
639 1-like (MRJP1). **E.** Serine-threonine protein kinase genes highly expressed in broodcare workers,
640 including dual specificity tyrosine-phosphorylation-regulated kinase (DYRK1A) and GATA zinc
641 finger domain-containing protein (GATA).

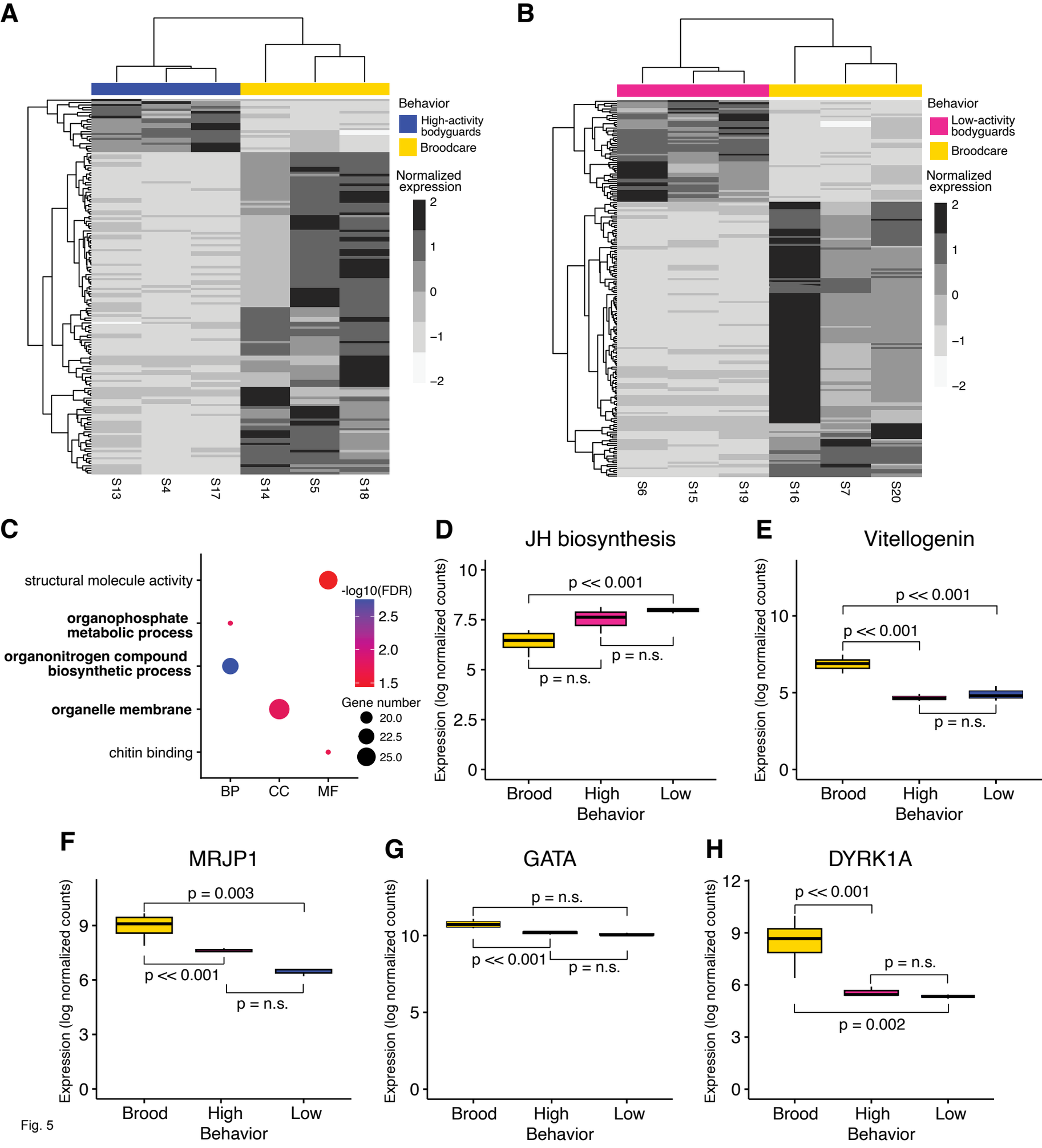
A
Behavior
High-activity bodyguards
Broodcare
Normalized expression
2
1
0
−1
−2
S13
S4
S17
S14
S5
S18
B
Behavior
Low-activity bodyguards
Broodcare
Normalized expression
2
1
0
−1
−2
S6
S15
S19
S16
S7
S20
C
structural molecule activity
organophosphate metabolic process
organonitrogen compound biosynthetic process
organelle membrane
chitin binding
BP
CC
MF
-log10(FDR)
2.5
2.0
1.5
Gene number
20.0
22.5
25.0
D
JH biosynthesis
Expression (log normalized counts)
10
7.5
5
2.5
0
p << 0.001
p = n.s.
p = n.s.
Brood
High
Low
Behavior
E
Vitellogenin
Expression (log normalized counts)
10
5
0
p << 0.001
p << 0.001
p = n.s.
Brood
High
Low
Behavior
F
MRJP1
Expression (log normalized counts)
9
6
3
0
p = 0.003
p << 0.001
p = n.s.
Brood
High
Low
Behavior
G
GATA
Expression (log normalized counts)
10
5
0
p = n.s.
p << 0.001
p = n.s.
Brood
High
Low
Behavior
H
DYRK1A
Expression (log normalized counts)
12
9
6
3
0
p << 0.001
p = n.s.
p = 0.002
Brood
High
Low
Behavior

Fig. 5