Evidence for a conserved queen-worker genetic toolkit across slave-making ants and their ant hosts

B. Feldmeyer[1], C. Gstöttl[2], J. Wallner [2], E. Jongepier[3,5], A. Séguret[5], D.A. Grasso[4], E. Bornberg-Bauer[5], S. Foitzik[6], J. Heinze[2]

[1] Senckenberg Biodiversity and Climate Research Centre, 60325 Frankfurt am Main, Germany

[2] Zoology / Evolutionary Biology, University of Regensburg, 93040 Regensburg, Germany

[3]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1098 XH Amsterdam, The Netherlands

[4]Department of Chemistry, Life Sciences and Environmental Sustainability, Università di Parma, 43124 Parma, Italy

[5]Institute for Evolution and Biodiversity, Westfälische Wilhelms University, 48149 Münster, Germany

[6]Institute of Molecular and Organismic Evolution, Johannes Gutenberg University, 55128 Mainz, Germany

Abstract

The ecological success of social Hymenoptera (ants, bees, wasps) depends on the division of labour between the queen and workers. Each caste is highly specialised in its respective function in morphology, behaviour and life-history traits, such as lifespan and fecundity. Despite strong defences against alien intruders, insect societies are vulnerable to social parasites, such as workerless inquilines or slave-making (dulotic) ants. Here, we investigate whether gene expression varies in parallel ways between lifestyles (slave-making versus host ants) across five independent origins of ant slavery in the *"Formicoxenus*-group" of the ant tribe Crematogastrini. As caste differences are often less pronounced in slave-making ants than non-parasitic ants, we also compare the transcriptomes of queens and workers in these species. We demonstrate a substantial overlap in expression differences between queens and workers across taxa, irrespective of lifestyle. Caste affects the transcriptomes much more profoundly than lifestyle, as indicated by 37 times more genes being linked to caste than to lifestyle and by multiple caste-associated gene modules with strong connectivity. However, several genes and one gene module are linked to the slave-making lifestyle across the independent origins, pointing to some evolutionary convergence. Finally, we do not find evidence for an interaction between caste and lifestyle, indicating that caste differences remain consistent even when species switch to a parasitic lifestyle. Our findings are a strong indication for the existence of a core set of genes whose expression is linked to the queen and worker caste in this ant taxon, supporting the "genetic toolkit" hypothesis.

Key words: Social parasitism, caste, transcriptomes, gene networks, slave-making ants, dulosis, selection

Introduction

The ecological success of social insects is based on the efficient division of labour between reproductives and non-reproductives, i.e., in the social Hymenoptera, the queens and workers (Hölldobler & Wilson, 2009; Wilson, 1971). Instead of producing their own offspring, workers help to raise the offspring of their mother or other related queens. The altruism of workers is explained by their relatedness to the recipients of their help, which is maintained by the closure of the society against unrelated freeloaders or parasites (Hamilton, 1964, 1987). Nevertheless, several species have evolved sophisticated ways to infiltrate and usurp social insect colonies (Rabeling, 2020), and among these are the charismatic "slave-making" or "dulotic" ants (Hölldobler & Wilson, 1990; Buschinger, 2009; D'Ettorre & Heinze, 2001; Mori et al., 2001; Visicchio et al., 2001). Freshly mated, young queens of slave-making ants invade the nests of closely related, non-parasitic ant species, where they kill or expel the resident queen(s) and often also the adult workers. Host workers emerging from the conquered host brood take care of the slave-maker queen and her offspring, maintain the nest, and forage for food. Slave-maker workers do not engage in normal worker chores (Hölldobler & Wilson 1990; Buschinger, 2009) but instead raid neighbouring host nests and pillage their brood, thus replenishing or increasing the host force.

Ten independent origins of slavery in ants are documented (Stoldt & Foitzik, 2021), and at least five convergent origins lie within the "*Formicoxenus*-group" of the myrmicine tribe Crematogastrini (Blaimer et al., 2018). This allows investigation of whether convergent changes in gene expression occurred among those evolutionary switches to slave-making. There are several morphological, physiological and behavioural similarities among slave-making species of independent origin. For example, slave-maker workers are often heavily armed with strong mandibles and associated muscles, which leads to enlarged heads (Hölldobler & Wilson, 1990). Given that workers of slave-making species no longer take care of the daily duties in the colony, they have become more queen-like in their task repertoire. They neither forage for food nor do they engage in brood care (Buschinger, 2009). Moreover, slave-maker workers tend to have increased reproductive potential.

While the ovaries of non-parasitic workers have fewer ovarioles than the queen's ovaries and rarely contain mature eggs in the presence of a fertile queen, the ovaries of slave-maker workers often have the same number of ovarioles as the queen (Heinze, 1996b). They form reproductive hierarchies (Franks and Scovell, 1983; Bourke, 1988; Heinze, 1996a) and frequently lay male-destined eggs even in the queen's presence (Foitzik & Herbers, 2001; Brunner et al., 2005; Suefuji & Heinze, 2014). Gene expression differs strongly between workers and queens of most species, reflecting their divergent function, behaviour, and physiology (Gstöttl et al., 2020; Korb et al., 2021; Morandin et al., 2019a, b; Feldmeyer et al., 2013). For example, queen transcriptomes are characterised by the expression of genes associated with fecundity (e.g. vitellogenins), immunity, DNA repair and response to oxidative stress functionalities linked to their long lifespan (Stoldt et al., 2021). However, given the described similarities between slave-maker queens and workers, we expected their transcriptomes to differ less than those of queens and workers of non-parasitic species.

We therefore investigated the influence of caste (queen vs worker) and lifestyle (slave-maker vs non-parasitic) on gene expression of adult individuals and the interaction between those two parameters, which would indicate that caste is affecting gene expression differently in non-parasitic versus slave-making species. Here we use a protocol which is aimed at maximising reproducibility and minimising confounding effects by using all five available and evolutionary closely related pairs of slave-maker and host in the myrmicine tribe Crematogastrini. For each species, we sequenced the transcriptomes of six pooled workers and three pooled queens, respectively, taking slave-making species vs. host species as replicates. We investigated gene regulatory network properties according to caste and lifestyle and constructed orthologue clusters to investigate putative parallel selection patterns in genes associated with the slave-maker versus host lifestyle, including two related non-host taxa and two samples of the distantly related ant *Cardiocondyla obscurior* as outgroup.

Material & Methods

Sampling and sequencing

Colonies of 15 myrmicine ant species of the "*Formicoxenus*-group" (genera *Harpagoxenus*, *Leptothorax*, and *Temnothorax*, including the previously synonymized genera of slave-making ants: *Chalepoxenus*, *Myrmoxenus*, and *Protomognathus* (Ward et al., 2015, but see Seifert et al., 2016) were collected between 2016-2018 from various locations across Germany, Italy, and the US (Supplement_coordinates). Colonies were either brought to the lab in Regensburg, Mainz, or Münster, and kept under standard conditions in incubators (12 h 25°/ 12 h 25°C day-night cycles) before six workers and three queens were pooled per species for RNA extraction. For each species, workers and queens originated from three colonies. We generated one queen and one worker transcriptome for seven slave-making ants and their host species (Supplement Table S1). RNA was extracted using the Nucleo-Spin Mini kit (Macherey-Nagel). Samples were shipped to StarSEQ (Mainz) for library preparation and 100bp paired end sequencing on an Illumina HiSeq. In total, we obtained 14mio reads on average per sample (Supplement Table S2).

Gene expression analyses

Raw reads were quality checked using FastQC v.0.11.8 (Andrews, 2010), and adapters trimmed with Trimmomatic v.2.8.4 (Bolger et al., 2014). HiSat2 v.2.1.0 (Kim et al., 2015) was used to map the reads to the *T. longispinosus* genome v.1 (GenBank accession: GCA_004794745.1_tlon_1.0; Kaur et al., 2019). We chose to use a single species genome as reference for all species to be able to later directly compare expression patterns between slave-making ants and their hosts as well as between queens and workers across species, with species as replicates. The counts table was created with HTSeq (Anders et al., 2015). To prevent spurious results due to low read counts, we removed from the counts matrix genes with less than 10 reads in at least four samples before the subsequent differential gene expression analysis with DESeq2 (Love et al., 2014). We started with the full model ~Caste+Lifestyle+Caste:Lifestyle. The interaction turned out to be nonsignificant (no differentially expressed genes for the interaction), and we thus based all follow-up analyses on the two main factors only. All p-values were adjusted by false discovery rate (FDR) correction as implemented in

DESeq2. To determine whether the significant slave-maker differentially expressed genes (DEGs) are more numerous than we would expect by chance, we ran 1000 permutations on the DEGs analysis in R.

To gain a deeper understanding for the functionality of DEGs, we 1) conducted a functional enrichment analysis, 2) inferred pathways in which the DEGs are involved, and 3) used a word mining approach based on longevity and fecundity terms. In detail, we ran Interproscan v.5.39-77.0 (Jones et al., 2014) locally to obtain GO information using the *T. longispinosus* predicted proteome (GenBank accession: GCA_004794745.1_tlon_1.0; Kaur et al., 2019) as query. The GO enrichment analysis was performed with the R package TopGO (Alexa & Rahnenführer, 2016), using the 'parentchild' algorithm and the Fishers exact test for significance. Furthermore, the proteome was annotated with KEGG functional ortholog numbers using the BlastKOALA web utility (last accessed: 11.02.2021; Kanehisa et al., 2016) with 'Eukaryotes, Animals' specified as Taxonomy group. KEGG pathway affiliation of genes that were differentially expressed in each of the four groups (queens, workers, slave-makers and hosts) was assessed using the online utility of KEGG mapper with *Apis mellifera* as reference species (Kanehisa & Sato, 2020), and visualized including the log2-fold change using pathview (Luo & Brouwer, 2013) in R v.3.6.3 (R Core team, 2020). KEGG pathway enrichment was assessed with the enrichKEGG function implemented in clusterProfiler v.3.14.3 (Yu et al., 2012) where the universe was defined as the total set of *T. longispinosus* protein predictions with a KEGG functional ortholog annotation and organism 'ko'. For the text mining approach, we used gene annotations based on a BlastP search of the proteome versus the RefSeq invertebrate database, as query for a UniProt search. More specifically, we extracted the gene function information for each gene with entries from *C. elegans, D. melanogaster, A. mellifera* and searched for terms related to fecundity and longevity (Supplement Table S3), both traits which we assumed to differ between queens and workers (script available from Negroni et al., 2021). We conducted an enrichment analysis by conducting a Fisher's exact test to test whether the number of genes with terms in fecundity or longevity within differentially expressed genes was higher than expected with respect to

144 the complete proteome. The online tool Venny v.2.0.2 was used to generate the Venn diagram
145 (https://bioinfogp.cnb.csic.es/tools/venny).
146
147 *Network analysis*
148 To identify networks of co-expressed genes (modules), we constructed a weighted gene co-
149 expression network analysis using the WGCNA (Langenfelder & Horwath, 2008) package in R. We
150 used all genes that had passed the quality filtering step for the expression analysis (N = 8,327), thus
151 not only the differentially expressed genes. Gene counts were normalized using the
152 *varianceStabilizingTransformation* function from DESeq2 (Love et al., 2014). Following the WGCNA
153 guidelines, we picked a soft-thresholding power of 8 for adjacency calculation. To associate modules
154 to either caste or lifestyle, we first calculated the modules' eigengene using the *moduleEigengenes*
155 function and tested for module trait correlation using the *corPvalueStudent* function. The hub gene
156 of each module (i.e., the gene with the highest connectivity within a module) was determined using
157 the *chooseTopHubInEachModule* function. Moreover, for modules associated with caste or lifestyle,
158 we tested whether the connectivity of genes with caste-specific expression and genes differed from
159 the connectivity of those that were not differentially expressed. We ran linear models with
160 connectivity as response and caste-specificity (queen, worker, no expression difference) as
161 explanatory variable in R.
162 To determine the relevance of fecundity and longevity associated genes in lifestyle and caste
163 associated modules, we downloaded a list of 123 genes that were assembled based on information
164 from *Drosophila* and are part of the TI-J-LiFe pathways (TOR/IIS-JH-Lifespan and Fecundity) (Korb et
165 al. 2021). We used the Flybase identifiers to download the corresponding *Drosophila* sequences and
166 used a BlastX search to identify *T. longispinosus* proteins with a blast hit e-value < e-10. In cases
167 where we had two or three hits, we selected the protein with the longest match, which mostly
168 corresponded to the lowest e-value. In cases where we had >3 blast hits, we created a sequence
169 alignment using MACSE v.2.03 (Ranwez et al., 2011). A Maximum Likelihood phylogenetic tree with

1000 bootstrap replicates was constructed with RAxML with PROTGAMMAWAG as protein substitution model. Based on the tree topology we chose the *T. longispinosus* sequence with the closest relationship to the target *Drosophila* sequence.

Selected genes analysis

For this analysis, we added transcriptomes of four taxa that are not known to be parasitized by slave-making ants, which acted as a biological control for comparisons of selection intensity between slave-makers and hosts: as outgroup, two populations of *Cardiocondyla obscurior,* and the two non-host species of the *Formicoxenus*group, *Temnothorax nylanderi* and *T. rugatulus* (Supplement Table S1). We used Trinity v.2.8.6 (Grabherr et al., 2011) with standard settings to construct species-specific transcriptomes using both worker and queen transcripts. Nucleotide sequences were translated into amino acid sequences with TransDecoder (https://github.com/TransDecoder). As *de novo* transcriptomes are known to contain many transcript fragments as well as isoforms, we constructed orthogroups across all species with OrthoFinder (Emms & Kelly, 2015), including protein sequences derived from the *Temnothorax longispinosus* genome (Kaur et al., 2019), and retained orthogroups with a tlon-v1 ortholog only. After orthogroup construction, the *T. longispinosus* protein sequences derived from the genome were removed from the orthogroups from all downstream summary statistics and analyses. Since we only obtained few single copy orthologs, we used an inhouse script (Supplement S_script) to retain only a single sequence per species. In short, based on the pair-wise blast results from OrthoFinder the sequence with the highest sum of bit scores (i.e. best match to all other sequences) within each orthogroup was chosen as “centroid”. From each species the sequence with the best match to the centroid was chosen to create the single copy orthogroup. We used Clustal-omega v.1.2.4 (Goujon et al., 2010) to construct sequence alignments for each orthogroup, which were trimmed with TrimAL v.1.4.1 (Capella-Gutierrez et al., 2009) and the following settings: -gappyout -resoverlap 0.75 -seqoverlap 75 -backtrans. To test for signatures of positive selection we used the codeml implementation in ete3 v.3.1.1 (Huerta-Cepas et al., 2016) running the branch site

196 test of selection as follows: ete3 evol --models bsA bsA1 --tests bsA, bsA1 --leaves --internals. The
197 cluster-specific tree topology as inferred by RAxML v.8.2.12 (Stamatakis, 2014) was used as input
198 tree, and each species was coded as foreground branch consecutively. Finally, we blasted the *T.*
199 *longispinosus* sequence of each ortholog cluster with signatures of selection versus the *T.*
200 *longispinosus* genome and only retained ortholog clusters with sequences mapping to a single
201 location, as further means of preventing putative paralogs. We further ran a GO enrichment analysis
202 to test for overrepresented functions among the selected genes (details above), and used the KEGG
203 annotations to investigate the pathways in which genes with signature of selection are involved.
204
205 **Results**
206 We set out to investigate the effect of lifestyle and caste on gene expression patterns across 15
207 different species of ants contrasting slave-making ants with their hosts, and queens with workers.
208 We were mainly interested in determining whether there are common toolsets of genes
209 characteristic for a specific lifestyle or caste. We additionally tested for genes under selection in the
210 above-mentioned slave-maker-host pairs plus four additional non-host outgroup species.
211 In general, gene expression patterns seem to be most similar amongst phylogenetically close
212 species. The principal component analysis revealed a very strong phylogenetic effect with 71% of
213 variance explained by PC1, which separated the *Harpagoxenus / Leptothorax* group from the
214 *Temnothorax* species (Supplement Figure S1), while PC2, explaining 8% of the variance, separated
215 queens from workers. Caste had a much stronger effect on gene expression compared to lifestyle, as
216 also evident in the heatmap dendrogram, where samples cluster according to caste rather than
217 lifestyle (Figure 1). As our analysis did not reveal an interaction between lifestyle and caste (0 DEGs
218 associated to the interaction term), we present the gene expression plus gene network results
219 according to the main effects, caste and lifestyle, and finally the results from the selection analysis.
220
221 *Lifestyle*

We detected 62 differentially expressed genes (DEGs) between lifestyles (host versus slave-maker species) cumulative across the five origins of slavery (see Supplement Table 1 DEGs). A permutation test with 1000 iterations revealed that this is more than expected by chance (p = 0.04). Six genes were consistently higher expressed in all seven slave-maker species compared to hosts, of which three were annotated: “protein DVR-1 homolog”, “sulfotransferase family cytosolic 1B member 1-like”, and “ATP-dependent DNA helicase II subunit 1-like”. Only a single gene, “NADPH oxidase 5”, showed higher expression in all seven host species compared to slave-makers (Figure 2). Various metabolic processes were significantly enriched among the differentially expressed genes between lifestyles (total = 21 processes). Among genes up-regulated in slave-makers, 14 processes were enriched, and three functions were enriched in genes up-regulated in hosts, including “cell death” (Supplement_GO). In addition to the functional level, we also investigated whether the genes in the lifestyle DEGs set were overrepresented in specific pathways. Due to the low number of lifestyle-associated DEGs, and only about half of these with KEGG annotation, the enrichment analyses for the complete set of lifestyle DEGs, and the separate host and slave-maker gene-sets, did not result in any overrepresented functions.

Host colonies are generally larger than slave-maker colonies, *i.e.*, host queens are more fecund than slave-maker queens. Moreover, fecundity is often positively correlated with longevity in social insects (Korb & Heinze, 2016; Negroni et al., 2016). We therefore conducted a word mining approach based on UniProt entries to investigate whether genes with known fecundity or longevity functionality were overrepresented in the sets of DEGs between hosts and slave-makers (Table 1). We indeed recovered more DEGs putatively associated with fecundity than expected by chance (N = 31; Fisher’s exact test, p-value = 6.619e-09), but not for the longevity associated genes (Table 1).

We obtained 10 modules of co-expressed genes (named after colours as given by the *WGCNA* package), of which one module, red, was positively associated with the slave-making lifestyle (Figure 5). This module contained 191 genes, of which 19 were up-regulated in slave-makers and none in hosts. Genes associated with this lifestyle module were enriched for “reactive oxygen species

metabolic processes", "oxidation reduction processes", or "fatty acid biosynthetic process" (Supplement_WGCNA). Genes up-regulated in slave-makers had the highest connectivity in this module associated with lifestyle (Supplement Figure S2).

Caste

Caste had a much stronger effect on gene expression than lifestyle (χ^2= 2496.8, df = 1, p-value < 2.2e-16), with 2,321 differentially expressed genes (DEGs) between queens and workers. Around half of all caste-specific DEGs showed consistently higher expression in queens (N=738 out of 1,295) or in workers (N = 450 out of 1,026) across all species. For example, genes for an *insulin-like growth factor 2*, *mRNA-binding protein 1*, *maternal protein exuperantia*, *histone deacetylases*, and several *serine/threonine-protein kinases* were more highly expressed in queens of all 15 species, while *pro-corazonine like* had higher counts in workers of all but one species. The number of genes with consistent expression across taxa was significantly lower for the lifestyle comparison (χ^2= 36.865, df = 1, p-value = 1.266e-09).

As queens and workers strongly differ in fecundity and longevity, we used a word mining approach based on UniProt entries to investigate whether genes with known fecundity or longevity functionality were overrepresented in the sets of caste-specific DEGs (Table 1). Indeed, we detected more fecundity-associated terms (N = 1076; Fisher's exact test, p-value = 2.2e-16), and more longevity-associated terms (N = 412; Fisher's exact test, p-value < 0.0003) than expected by chance.

In the caste-specific DEG set, 39 gene ontology (GO) functions were significantly overrepresented, many of which were associated with metabolic and biosynthetic processes (Supplement_GO). Genes up-regulated in queens were enriched for 60 functions linked to various metabolic processes or stress responses. Among the worker up-regulated genes, 78 functions were significantly enriched, also belonging to metabolic and biosynthetic processes, but also oxidation-reduction. In addition to the functional enrichment, we conducted a KEGG-pathway enrichment analysis. In total, we found 27 pathways enriched amongst the caste-specific DEGs, 12 for genes up-

regulated in queens, and 61 genes up-regulated in workers (Supplement_KEGG). Multiple putatively reproduction-associated and repair pathways, such as "meiosis-yeast", "cell cycle", "DNA replication", "RNA transport," or "ribosome biogenesis," were enriched in queens, and pathways such as "olfactory transduction" and "longevity regulating pathway" (Figure 3) were enriched in genes up-regulated in workers (Supplement_KEGG).

The gene network analysis resulted in five modules associated with caste, three of which were positively associated with queens (yellow, green, magenta), and two with workers (black and blue) (Figure 5). There was a trend for the pink module to also be associated with caste ($p = 0.07$), and it was included in the following analyses (for details see: Supplement_WGCNA). In all three worker associated modules (black, blue, pink), connectivity was highest for worker up-regulated genes, lowest for queen up-regulated genes, and intermediate for genes that were not differentially expressed. The same pattern was observed in queen-specific modules (yellow, green, magenta), in which the queen up-regulated genes had the highest connectivity (Supplement_WGCNA).

Comparison to other species

We identified 84 genes associated with one of the 10 WGCNA modules that have also previously been linked to fecundity or longevity in Drosophila. We obtained these from the "TI-J-LiFe" list containing 123 candidate genes (Korb et al., 2021). About half of these (N = 60) were found in the turquoise module, which was neither associated to caste nor lifestyle, and eleven were found in the blue, caste-linked module (Supplement_WGCNA). To determine whether overrepresented gene functions linked to caste associated modules from our Formicoxenus-group data set could be extrapolated to other ants or even termites, we compared our modules to Morandin et al. (2016) and Lin et al. (2021). Only three functions were shared between enriched GO-terms of caste associated modules in our slave-making data set (N = 46) and enriched GO-functions in Formica caste associated modules (N = 155; Morandin et al., 2016), namely "cellular protein modification process", "protein modification process", "monovalent inorganic cation transport" (Table 2). Eight terms were

shared with caste associated modules in termites (N = 277; Lin et al., 2021) (Table 2). No terms were shared amongst the three studies. However, these results should be taken with caution as most enrichment algorithms take the hierarchical structure of GO terms into account. Thus one may have a lower or higher term in the hierarchy, which does not mean they are essentially different.

As queen and workers differ in fecundity and longevity, we investigated whether the same set of fecundity-longevity associated genes of the "TI-J-LiFe" list (N = 84) which were found in the termites (Lin et al., 2021) could be found in our data set . There were two genes, Ras64B and Kr-h1, associated with a queen-worker co-expression module in termites and in Crematogastrini.

Selected genes analysis

The transcriptomes of the 19 taxa, including four samples of non-host species, consisted of 67,150-155,629 transcripts with 32,584-51,743 open reading frames and 92-96% DOGMA completeness. In total, we obtained 10,699 ortholog clusters of which 5,826 clusters contained at least one transcript per species. 1,398 clusters remained after trimming and filtering for single copies per species. We identified 660 signatures of selection across all species, that are associated with 424 ortholog clusters, and mapped to a single location on the genome (Supplement_codeml). Between 6-92 genes showed signatures of selection in each species (Supplement_codeml), but there was no difference in the number of selected genes between lifestyles (quasipoisson model, $X^2 = 1.33$, d.f. = 2, $p = 0.51$). In slave-makers, genes with signatures of selection were significantly enriched in functions such as protein modification, demethylation and energy maintenance (Figure 4a, Supplement_codeml). In hosts and non-hosts, several regulatory and metabolic functions are significantly overrepresented in genes with signatures of selection (Figure 4b+c).

We identified 114 genes that were both differentially expressed and additionally showed signatures of selection (Supplement_overlap). Of these, a single gene was differentially expressed between castes and lifestyles and positively selected in *T. unifasciatus* (TLON_06615-RA; pleckstrin homology domain-containing family F member 2 isoform X1). All remaining 113 were differentially expressed

only between castes. 35 of these caste-specific DEGs showed signatures of selection in hosts, 23 in non-hosts, and 57 in slave-makers. Interestingly, many nucleotide metabolism pathways were included in the overlapping gene lists of differentially expressed and selected genes, including purine, alanine, aspartate, valine, and the "mTOR signalling pathway" (Supplement_overlap).

Discussion

Eusocial insects are characterised by their sophisticated division of labour, which led to the evolution of different castes. Ant queens are the main reproductives, known for their long life of up to several decades and high fecundity (Keller & Genoud, 1997), while the mostly infertile and short-lived workers take care of the brood, foraging, and nest defence (Hölldobler & Wilson, 1990). In slave-making species, however, workers do not perform the "general" worker chores. They raid host nests in summer and often are permitted to lay male-destined eggs in the presence of the queen. Our study shows that whole-body transcriptomes of seven species pairs of the "*Formicoxenus* group" of Crematogastrini (Blaimer et al., 2018; *Harpagoxenus*, *Leptothorax,* and *Temnothorax*) differ considerably more between castes (queen vs. worker) than with lifestyles (slave-maker vs. non-parasitic species). Caste polyphenism is based on differential gene expression, and our transcriptome analyses show that the expression of 1,188 genes consistently differs between queens and workers of all studied species, regardless of lifestyle. Although the evolution of a parasitic lifestyle is associated with considerable changes in morphology and behaviour, only few transcriptomic shifts from host to slave-making species were consistent across the five independent origins of slave-making.

Approximately 40x more genes were differentially expressed between castes than between lifestyles (2,321 vs. 62), and six out of ten co-expression modules were caste-associated in contrast to a single lifestyle module. This difference might in part reflect the single evolutionary origin of caste diphenism in ants, whereas slavery evolved repeatedly (Beibl et al., 2005; Feldmeyer et al., 2017; Prebus, 2017). Nevertheless, as slave-making workers are often fertile and do not take over normal

worker chores, we had expected to find a considerable interaction between caste and lifestyle in gene expression as well as in gene connectivity, as caste differences are less pronounced in slave-making species. The lack of the interaction indicates that slave-maker and host queens and workers are rather similar on a molecular level. Furthermore, it corroborates the result of a previous study where caste differences also exceeded differences due to other traits, such as worker sterility, queen number, or invasiveness (Morandin et al., 2016).

While in this study, half of the genes that were differentially expressed between queens and workers show the same expression pattern across all species, another study identified only a single gene (the myosin light chain) similarly expressed between queens and workers a set of 16 species of multiple genera (Morandin et al., 2016). The reason for this discrepancy may be explained by the different species relationships as well as the underlying data basis. We studied species within a single, closely related clade and used the genome of a single species as reference. In contrast, Morandin et al. (2016) investigated species from five genera in different subfamilies based on de novo assembled transcriptomes. Additionally, there is evidence for similarities in pathways across different lineages of social insects (ants, bees, wasps) rather than a "common toolkit" of genes responsible for the caste phenotype (Berens et al., 2015). In our phylogenetically more restricted data set however, we identified 1,188 genes representing the "core set" of queen-worker differences across the 15 species. For example, the gene "*corazonin*" has been shown to control social behaviour and caste identity in ants (Gospocic et al., 2017). The gene "*maternal protein exuperantia*", a maternal effect gene which is needed for proper localisation of the *bicoid* RNA during oocye formation (de Olivereira et al., 2017; McDonald et al., 1991), but also plays a role in *Drosophila* spermatogenesis (Hazelrigg et al., 1990) was up-regulated in all queens. Also, "G1/S-specific cyclin-D2" (cycD) was more strongly expressed in queens compared to workers in all but one species. This gene is involved in cell cycle regulation and part of the "FoxO signalling" and "Wnt signalling pathway", both important regulators of longevity. It could thus be associated with the lifespan

differences between these two castes. More generally, half of the genes differentially expressed between caste were associated with the UniProt functionalities “fecundity” and “longevity.”

Reflecting the convergent evolution of ant slavery (Beibl et al., 2005; Feldmeyer et al., 2017; Prebus, 2017), the different parasitic species not only show pronounced differences in their morphology, but also in raiding behaviour (Brandt et al., 2006; Johnson, 2008; Kleeberg & Foitzik, 2016). Species-specific raiding patterns are mirrored by species-specific gene expression patterns (Alleman et al., 2018) and genes under selection (Feldmeyer et al., 2017). Gene expression differences between pairs of slave-makers and their hosts showed much more variation across species (representing different origins of slaver-making) than the differences between queens and workers, *i.e.*, there are many more idiosyncrasies in how slave-makers and hosts differ. Nevertheless, we found 62 genes that varied with lifestyle, including genes for fatty acid synthases. This could be indicative either of differences in fat synthesis and maybe storage between the two lifestyles, but these synthases may also be involved in the synthesis of cuticular hydrocarbons which are used as communication signals to discriminate species and castes (Leonhard et al., 2016).

As slave-makers are the derived lifestyle with species-specific behaviours and traits, we expected to find more genes under selection in slave-makers than in hosts and/or non-host species, however the number of selected genes did not differ between lifestyles. Genes with signatures of selection were significantly enriched in functions such as protein modification, demethylation, and energy maintenance. An interesting candidate among the genes that showed signatures of selection in only one or a few of the species is venom protease-like in *T. muellerianus*. Though at present nothing is known on the composition of the venom used by *T. muellerianus* to kill host ants during slave-raids, many Hymenopteran venoms are proteases (Touchard et al., 2016). On a higher level, many metabolism pathways were included in this overlapping gene list, including purine, alanine, aspartate, valine, to name just a few, and the “mTOR signalling pathway”.

Conclusion

Social parasitism represents a derived state with clear differences to host and non-host species from morphology to behaviour. In most social parasites, queens and workers are more similar than in host species. Examining gene expression patterns and gene regulatory networks of 15 different ant species spanning five origins of slave-making, we expected to find an interaction between caste and lifestyle effects on gene expression patterns. Despite the phenotypic differences, between slavemaker and host castes, gene expression profiles were remarkably similar. Our study corroborates previous results indirectly indicating species-specific expression and selection patterns with respect to lifestyle, thus little common difference between host and slave-maker species. However, we observed a very strong and reliable effect of caste on gene expression. Within our broad taxonomic species spectrum, we were able to identify a core-set of 1,188 caste-specific genes, which show a consistent expression pattern across all species irrespective of lifestyle, pointing to a "genetic toolkit" in this set of related ant species.

Acknowledgements

We would like to thank Tilman Shell for writing the ortholog cluster filter script for us. This work was supported by the Deutsche Forschungsgemeinschaft (grant numbers Bo 2544/12-1 to EBB, Fo 298/20 to FO, He 1623/40 to JH).

References

Alexa, A., & Rahnenführer, J. (2016). *Gene set enrichment analysis with topGO. R package version 2.24. 0*.

Alleman, A., Feldmeyer, B., & Foitzik, S. (2018). Comparative analyses of co-evolving host-parasite associations reveal unique gene expression patterns underlying slavemaker raiding and host defensive phenotypes. *Scientific Reports*, *8*(1), 1951. https://doi.org/10.1038/s41598-018-20262-y

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, S. (2010). Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. *URL: https://www. bioinformatics. babraham. ac. uk/projects/fastqc*.

Beibl, J., Stuart, R. J., Heinze, J., & Foitzik, S. (2005). Six origins of slavery in formicoxenine ants. *Insectes sociaux*, *52*(3), 291–297.

Berens, A. J., Hunt, J. H., & Toth, A. L. (2015). Comparative Transcriptomics of Convergent Evolution: Different Genes but Conserved Pathways Underlie Caste Phenotypes across Lineages of Eusocial Insects. *Molecular Biology and Evolution*, *32*(3), 690–703. https://doi.org/10.1093/molbev/msu330

Blaimer, B. B., Ward, P. S., Schultz, T. R., Fisher, B. L., & Brady, S. G. (2018). Paleotropical Diversification Dominates the Evolution of the Hyperdiverse Ant Tribe Crematogastrini (Hymenoptera: Formicidae). *Insect Systematics and Diversity*, *2*(5). https://doi.org/10.1093/isd/ixy013

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bourke, A. F. G. (1988). Dominance orders, worker reproduction, and queen-worker conflict in the slave-making ant *Harpagoxenus sublaevis*. *Behavioral Ecology and Sociobiology*, *23*(5), 323–333. https://doi.org/10.1007/BF00300579

Brandt, M., Heinze, J., Schmitt, T., & Foitzik, S. (2006). Convergent evolution of the Dufour's gland secretion as a propaganda substance in the slave-making ant genera *Protomognathus* and *Harpagoxenus*. *Insectes Sociaux*, *53*(3), 291–299.

Brunner, E., Trindl, A., Falk, K. H., Heinze, J., & D'Ettorre, P. (2005). Reproductive Conflict in Social Insects: Male Production by Workers in a Slave-Making Ant. *Evolution*, *59*(11), 2480–2482. https://doi.org/10.1111/j.0014-3820.2005.tb00957.x

Buschinger, A. (2009). Social Parasitism among Ants: A Review (Hymenoptera: Formicidae). *Myrmecological News*, *12*, 219–235.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford, England)*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

de Oliveira, J. L., Sobrinho-Junior, I. S., Chahad-Ehlers, S., & de Brito, R. A. (2017). Evolutionary coincidence of adaptive changes in exuperantia and the emergence of bicoid in *Cyclorrhapha* (Diptera). *Development Genes and Evolution*, *227*(5), 355–365. https://doi.org/10.1007/s00427-017-0594-3

D'Ettorre, P., & Heinze, J. (2001). Sociobiology of slave-making ants. *Acta Ethologica*, *3*(2), 67–82. https://doi.org/10.1007/s102110100038

462 Emms, D. M., & Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome
463 comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157.
464 https://doi.org/10.1186/s13059-015-0721-2

465 Feldmeyer, B., Elsner, D., Alleman, A., & Foitzik, S. (2017). Species-specific genes under selection
466 characterize the co-evolution of slavemaker and host lifestyles. *BMC Evolutionary Biology*, *17*(1), 237.
467 https://doi.org/10.1186/s12862-017-1078-9

468 Feldmeyer, B., Elsner, D., & Foitzik, S. (2013). Gene expression patterns associated with caste and
469 reproductive status in ants: Worker-specific genes are more derived than queen-specific ones.
470 *Molecular ecology*, *23*. https://doi.org/10.1111/mec.12490

471 Foitzik, S., & Herbers, J. M. (2001). Colony structure of a slavemaking ant. I. Intracolony relatedness,
472 worker reproduction, and polydomy. *Evolution; International Journal of Organic Evolution*, *55*(2),
473 307–315. https://doi.org/10.1111/j.0014-3820.2001.tb01295.x

474 Franks, N. R., & Scovell, E. (1983). Dominance and reproductive success among slave-making worker
475 ants. *Nature*, *304*(5928), 724–725. https://doi.org/10.1038/304724a0

476 Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new
477 bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research*, *38*(Web Server issue),
478 W695-699. https://doi.org/10.1093/nar/gkq313

479 Gospocic, J., Shields, EJ., Glastad, KM., Lin, Y., Penick, CA., Yan, H., … Bonasio, R. (2017). The
480 neuropeptide corazonin controls social behavior and caste identity in ants. *Cell*, 170, 748–759. e12.
481 pmid:28802044. https://doi.org/10.1016/j.cell.2017.07.014

482 Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L.,
483 Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F.,
484 Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Trinity: Reconstructing a full-length
485 transcriptome without a genome from RNA-Seq data. *Nature biotechnology*, *29*(7), 644–652.
486 https://doi.org/10.1038/nbt.1883

487 Gstöttl, C., Stoldt, M., Jongepier, E., Bornberg-Bauer, E., Feldmeyer, B., Heinze, J., & Foitzik, S. (2020).
488 Comparative analyses of caste, sex, and developmental stage-specific transcriptomes in two
489 *Temnothorax* ants. *Ecology and Evolution*, *10*(10), 4193–4203. https://doi.org/10.1002/ece3.6187

490 Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*,
491 *7*(1), 1–16. https://doi.org/10.1016/0022-5193(64)90038-4

492 Hamilton, W. D. (1987). Kinship, recognition, disease, and intelligence: Constraints of social
493 evolution. In Y. Ito, J. Brown, & J. Kikkawa (Hrsg.), *Animal Societies: Theories and Facts* (S. 81–102).
494 Japan Scientific Societies Press.

495 Hazelrigg, T., Watkins, W. S., Marcey, D., Tu, C., Karow, M., & Lin, X. R. (1990). The exuperantia gene
496 is required for *Drosophila* spermatogenesis as well as anteroposterior polarity of the developing
497 oocyte, and encodes overlapping sex-specific transcripts. *Genetics*, *126*(3), 607–617.

498 Heinze, J. (1996a). Reproductive Hierarchies Among Workers of the Slave-making Ant, *Chalepoxenus*
499 *muellerianus*. *Ethology*, *102*(1), 117–127. https://doi.org/10.1111/j.1439-0310.1996.tb01109.x

500 Heinze, J. (1996b). The reproductive potential of workers in slave-making ants. *Insectes Sociaux*,
501 *43*(3), 319–328. https://doi.org/10.1007/BF01242933

502 Hölldobler, B., & Wilson, E. O. (1990). *The Ants*. Harvard University Press.

Hölldobler, B., & Wilson, E. O. (2009). *The superorganism: The beauty, elegance, and strangeness of insect societies*. WW Norton & Company.

Huerta-Cepas, J., Serra, F., & Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular Biology and Evolution*, *33*(6), 1635–1638. https://doi.org/10.1093/molbev/msw046

Johnson, C. A. (2008). Slave-Maker Ant Competition for a Shared Host and the Effect on Coevolutionary Dynamics. *Ecological Monographs*, *78*(3), 445–460. https://doi.org/10.1890/07-1515.1

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Kanehisa, M., & Sato, Y. (2020). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science*, *29*(1), 28–35. https://doi.org/10.1002/pro.3711

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology*, *428*(4), 726–731. https://doi.org/10.1016/j.jmb.2015.11.006

Kaur, R., Stoldt, M., Jongepier, E., Feldmeyer, B., Menzel, F., Bornberg-Bauer, E., & Foitzik, S. (2019). Ant behaviour and brain gene expression of defending hosts depend on the ecological success of the intruding social parasite. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1769), 20180192. https://doi.org/10.1098/rstb.2018.0192

Keller, L., & Genoud, M. (1997). Extraordinary lifespans in ants: A test of evolutionary theories of ageing. *Nature*, *389*(6654), 958–960. https://doi.org/10.1038/40130

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

Kleeberg, I., & Foitzik, S. (2016). The placid slavemaker: Avoiding detection and conflict as an alternative, peaceful raiding strategy. *Behavioral Ecology and Sociobiology*, *70*(1), 27–39. https://doi.org/10.1007/s00265-015-2018-6

Korb, J., & Heinze, J. (2016). Major Hurdles for the Evolution of Sociality. *Annual Review of Entomology*, *61*(1), 297–316. https://doi.org/10.1146/annurev-ento-010715-023711

Korb, J., Meusemann, K., Aumer, D., Bernadou, A., Elsner, D., Feldmeyer, B., Foitzik, S., Heinze, J., Libbrecht, R., Lin, S., Majoe, M., Monroy Kuhn, J. M., Nehring, V., Negroni, M. A., Paxton, R. J., Séguret, A. C., Stoldt, M., Flatt, T., & null, null. (2021). Comparative transcriptomic analysis of the mechanisms underpinning ageing and fecundity in social insects. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1823), 20190728. https://doi.org/10.1098/rstb.2019.0728

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Leonhardt, S. D., Menzel, F., Nehring, V., & Schmitt, T. (2016). Ecology and Evolution of Communication in Social Insects. *Cell*, *164*(6), 1277–1287. https://doi.org/10.1016/j.cell.2016.01.035

Lin, S., Werle, J., & Korb, J. (2021). Transcriptomic analyses of the termite, *Cryptotermes secundus,* reveal a gene network underlying a long lifespan and high fecundity. *Communications biology*, *4*(1), 1–12.

Love, M. I., Huber, W., & Anders, S. (2014a). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 550. https://doi.org/10.1186/s13059-014-0550-8

Love, M. I., Huber, W., & Anders, S. (2014b). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Luo, W., & Brouwer, C. (2013). Pathview: An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*, *29*(14), 1830–1831. https://doi.org/10.1093/bioinformatics/btt285

Macdonald, P. M., Luk, S. K., & Kilpatrick, M. (1991). Protein encoded by the exuperantia gene is concentrated at sites of bicoid mRNA accumulation in *Drosophila* nurse cells but not in oocytes or embryos. *Genes & Development*, *5*(12b), 2455–2466. https://doi.org/10.1101/gad.5.12b.2455

Morandin, C., Brendel, V. P., Sundström, L., Helanterä, H., & Mikheyev, A. S. (2019). Changes in gene DNA methylation and expression networks accompany caste specialization and age-related physiological changes in a social insect. *Molecular Ecology*, *28*(8), 1975–1993. https://doi.org/10.1111/mec.15062

Morandin, C., Hietala, A., & Helanterä, H. (2019). Vitellogenin and vitellogenin-like gene expression patterns in relation to caste and task in the ant *Formica fusca*. *Insectes Sociaux*, *66*(4), 519–531. https://doi.org/10.1007/s00040-019-00725-9

Morandin, C., Tin, M. M. Y., Abril, S., Gómez, C., Pontieri, L., Schiøtt, M., Sundström, L., Tsuji, K., Pedersen, J. S., Helanterä, H., & Mikheyev, A. S. (2016). Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. *Genome Biology*, *17*(1), 43. https://doi.org/10.1186/s13059-016-0902-7

Mori, A., Grasso, D.A., Visicchio, R., Le Moli, F. (2001). Comparison of reproductive strategies and raiding behaviour in facultative and obligatory slave-making ants: the case of *Formica sanguinea* and *Polyergus rufescens*. *Insectes Sociaux, 48*, 302-314. https://doi.org/10.1007/PL00001782

Negroni, M. A., Macit, M. N., Stoldt, M., Feldmeyer, B., & Foitzik, S. (2021). Molecular regulation of lifespan extension in fertile ant workers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1823), 20190736. https://doi.org/10.1098/rstb.2019.0736

Prebus, M. (2017). Insights into the evolution, biogeography and natural history of the acorn ants, genus *Temnothorax* Mayr (hymenoptera: Formicidae). *BMC Evolutionary Biology*, *17*(1), 250. https://doi.org/10.1186/s12862-017-1095-8

Rabeling, C. (2020). Social Parasitism. In C. K. Starr (Hrsg.), *Encyclopedia of Social Insects* (S. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-90306-4_175-1

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. *PLOS ONE*, *6*(9), e22594. https://doi.org/10.1371/journal.pone.0022594

Seifert, B., Buschinger, A., Aldawood, A., Antonova, V., Bharti, H., Borowiec, L., Dekoninck, W., Dubovikoff, D., Espadaler, X., Flegr, J., Georgiadis, C., Heinze, J., Neumeyer, R., Ødegaard, F., Oettler,

583 J., Radchenko, A., Schultz, R., Sharaf, M., Trager, J., … Zettel, H. (2016). Banning paraphylies and
584 executing Linnaean taxonomy is discordant and reduces the evolutionary and semantic information
585 content of biological nomenclature. *Insectes Sociaux*, *63*(2), 237–242.
586 https://doi.org/10.1007/s00040-016-0467-1

587 Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large
588 phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

589 Stoldt, M., & Foitzik, S. (2021). Slave-Making in Ants (Dulosis). In C. K. Starr (Hrsg.), *Encyclopedia of*
590 *Social Insects* (S. 807–814). Springer International Publishing. https://doi.org/10.1007/978-3-030-
591 28102-1_105

592 Stoldt, M., Klein, L., Beros, S., Butter, F., Jongepier, E., Feldmeyer, B., & Foitzik, S. (2021). Parasite
593 Presence Induces Gene Expression Changes in an Ant Host Related to Immunity and Longevity.
594 *Genes*, *12*(1), 95. https://doi.org/10.3390/genes12010095

595 Suefuji, M., & Heinze, J. (2014). The genetic population structure of two socially parasitic ants: "The
596 active slave-maker" *Myrmoxenus ravouxi* and the "degenerate slave-maker" *M. kraussei*.
597 *Conservation Genetics*, *15*(1), 201–211. https://doi.org/10.1007/s10592-013-0531-6

598 Touchard, A., Aili, S. R., Fox, E. G. P., Escoubas, P., Orivel, J., Nicholson, G. M., & Dejean, A. (2016).
599 The Biochemical Toxin Arsenal from Ant Venoms. *Toxins*, *8*(1), 30.
600 https://doi.org/10.3390/toxins8010030

601 Visicchio, R., Mori, A., Grasso, D.A., Castracani, C., Le Moli, F. (2001). Glandular sources of
602 recruitment, trail, and propaganda semiochemicals in the slave-making ant *Polyergus rufescens*.
603 *Ethology Ecology and Evolution 13*, 361-372. https://doi.org/10.1080/08927014.2001.9522767

604 Ward, P. S., Brady, S. G., Fisher, B. L., & Schultz, T. R. (2015). The evolution of myrmicine ants:
605 Phylogeny and biogeography of a hyperdiverse ant clade (Hymenoptera: Formicidae). *Systematic*
606 *Entomology*, *40*(1), 61–81. https://doi.org/10.1111/syen.12090

607 Wilson, E. O. (1971). *The Insect Societies*. Belknap Press of Harvard University Press.

608 Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: An R Package for Comparing Biological
609 Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*, *16*(5), 284–287.
610 https://doi.org/10.1089/omi.2011.0118

611

612 **Data Accessibility**

613 All raw sequence data underlying this study have been deposited in the National Centre for
614 Biotechnological Information (NCBI) Sequence Read Archive (SRA) and will be accessible upon
615 publication of this manuscript (BioProject accession number XY will follow in the next version of the
616 manuscript).

617

Author Contributions

The study was conceived by JH, EBB and SF, and was designed by EJ, BF, JH, EBB and SF. DG provided ant samples. BF, CS, JW and EJ conducted the analyses. All authors contributed to writing the paper.

Authors declare no conflict of interest.

622 **Table 1**: Number of differentially expressed genes with a fecundity or longevity functionality based
623 on a text mining approach using the UniProt database as a reference. Results of Fisher's exact test
624 are given indicating whether the number of genes with either fecundity or longevity association in
625 caste and lifestyle are more than expected by chance.
626

Treatment	**Fecundity**		**Longevity**	
	N of genes	P-Value	N of genes	P-value
Caste	**1076**	**2.2e-16**	**412**	**0.0003**
Queen	630		262	
Worker	447		150	
Lifestyle	**31**	**6.619e-09**	13	0.39
Host	9		6	
Slave-maker	23		7	

627

628

Table 2: Shared enriched GO terms between queen-worker caste associated modules from this study, a study on 16 ant species from multiple genera (Morandin et al. 2016), and the termite *Cryptotermes secundus* (Lin et al. 2021).

Organism(s)	Shared enriched GO-terms with this study
Multiple ants	cellular protein modification process
Multiple ants	protein modification process
Multiple ants	monovalent inorganic cation transport
C. secundus	cyclic nucleotide biosynthetic process
C. secundus	signal transduction
C. secundus	transmembrane transport
C. secundus	DNA replication initiation
C. secundus	transcription
C. secundus	G protein-coupled receptor signaling pathway
C. secundus	intracellular protein transport
C. secundus	proteolysis

Figure 1: Sample dendrogram and heatmap depicting a strong clustering according to caste and a weaker effect of lifestyle. The heatmap is based on the top 50 highest expressed genes. Slave-making species are highlighted in bold.

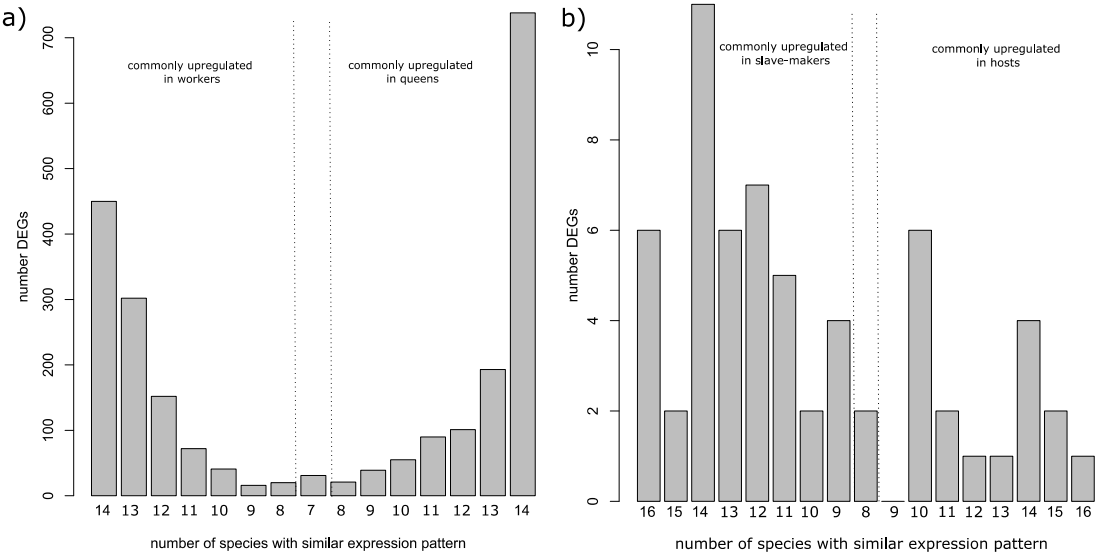

Figure 2: Number of differentially expressed genes (DEGs) which showed shared expression patterns across a majority of study species during the pairwise comparisons of queens vs workers or slave-makers vs hosts. a) Number of DEGs upregulated in workers (left) or upregulated in queens (right) across more than half the study species. b) Number of DEGs upregulated in slave-makers (left) or upregulated in hosts (right) across more than half the study species. The vertical lines delimit the start of a majority of species in either direction, with the bar between the lines showing DEGs that showed a common expression pattern in one direction in half the species, and in the other direction in half the species - e.g. in graph a), DEGs which were upregulated in queens in 7 species, and upregulated in workers in the other 7 species. Please note that the host-slave-maker pairs add up to 16 comparisons instead of 14, since *Harpagoxenus sublaevis* parasitizes two host species: *L. acervorum* and *L. muscorum.*

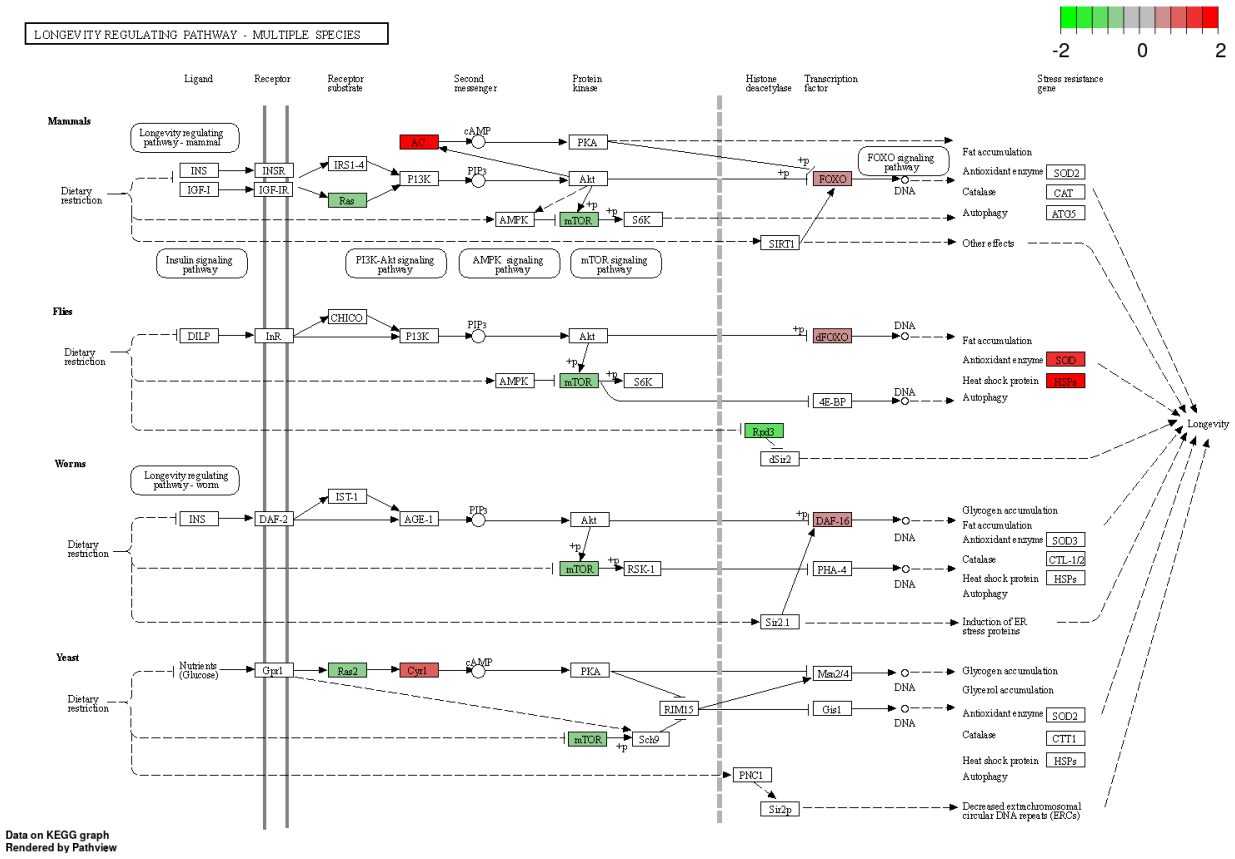

Figure 3: KEGG map depicting the "longevity regulating pathway" with up-regulated genes in queens (negative logFC; green) versus up-regulated in workers (positive logFC; red).

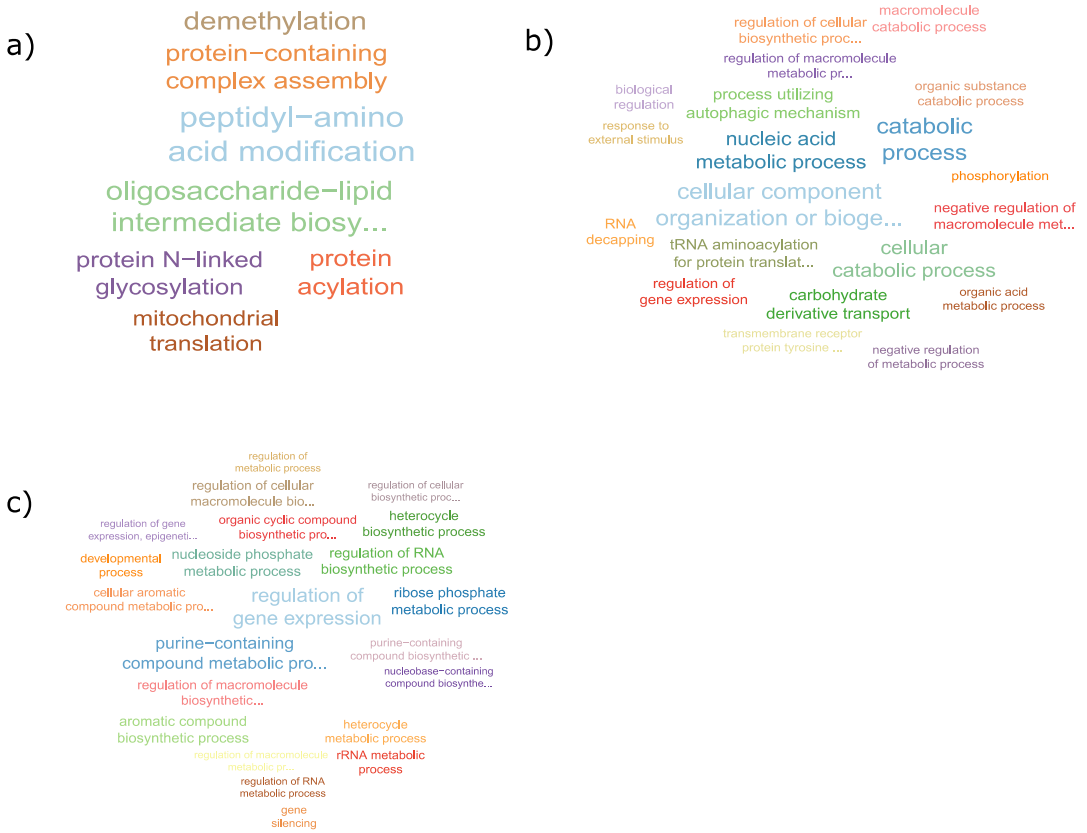

Figure 4: Word clouds of enriched GO functions based on genes with signature of selection in a) slave-makers, b) hosts, c) non-hosts.

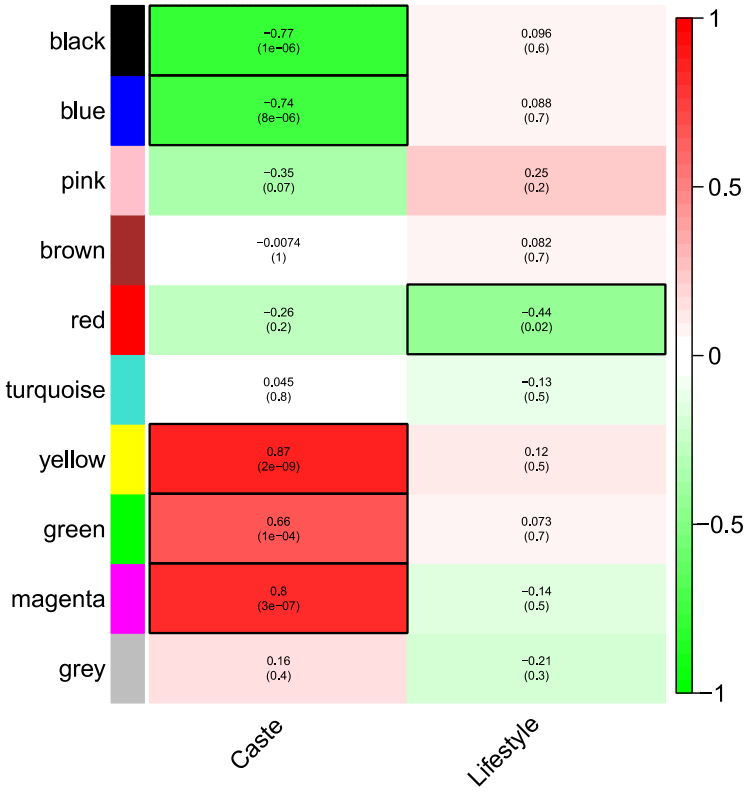

Figure 5: Module trait relationship of the 10 gene co-expression clusters with caste and lifestyle. The strength of the correlation is given in the upper numbers and significance levels in parentheses below. Significant modules are highlighted by black boxes. Modules are labelled by colours according to standard WGCNA output. Green indicates modules of genes overexpressed in workers / slave-makers, red indicates those with genes more expressed in queens / hosts.