bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Transcriptomic basis and evolution of the ant social interactome

Michael R. Warner[1], Alexander S. Mikheyev[2,3], Timothy A. Linksvayer[1]

[1]University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2]Okinawa Institute of Science and Technology, Okinawa, Japan

[3]Research School of Biology, Australian National University, Canberra, Australia

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Social interactions are a fundamental feature of life for most organisms. Organismal development is often strongly regulated by interactions among close relatives, yet little is known about how genes expressed in social partners indirectly affect developmental trajectories and trait expression. In eusocial insects, social interactions between caregiving worker nurses and larvae strictly regulate larval development and resultant adult phenotypes. Here, we study the social interactome regulating larval development by collecting and sequencing interacting nurses and larvae across a time course of larval development. First, we find that the majority of nurse and larval transcriptomes exhibit parallel expression dynamics across larval development, providing a strong transcriptomic signature of the social interactome regulating larval development. Next, we leverage this widespread nurse-larva gene co-expression to infer social gene regulatory networks acting between nurses and larvae. We find that genes with the strongest social effects tend to be peripheral elements of within-tissue regulatory networks and are often known to encode secreted proteins. For example, our data suggest that the gene *giant-lens* expressed by nurses may inhibit larval epidermal growth factor signaling, which is known to affect various aspects of insect development, including caste in honey bees. Finally, we find that genes recruited for social regulatory processes tend to be relatively evolutionarily young and tend to experience relaxed selective constraint. Overall, our study provides a first glimpse into the molecular details and evolutionary features of the social mechanisms that regulate all aspects of life in eusocial insects.

Author Summary

From single-celled bacteria to complex plants and animals, most organisms experience and are strongly impacted by social interactions. Despite their obvious importance, little is known about the genetic basis of social traits and interactions. In eusocial insects, the development of larvae is strictly regulated by caregiving nurse workers through a myriad of social interactions occurring across larval development. In this paper, we study ant nurse-larva social interactions and the effect of interactions on gene expression

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

by simultaneously collecting and sequencing interacting nurses and larvae across larval development. We find that broad-scale gene expression patterns reflect ongoing social interactions. Based on nurse-larva gene co-expression, we identify genes expressed in nurses that exhibit strong putative effects on the expression of genes in larvae. Interestingly, such genes with strong social roles tend to be less important for pathways operating within individuals, and such highly social genes tend to experience relatively weaker natural selection in comparison to less social genes. This suggests that genes underlying social traits are characterized by distinct evolutionary and network features. Overall, this study represents a significant advance in our understanding of the intersection of genetics, behavior, and evolution.

Introduction

Social interactions play a prominent role in the lives of nearly all organisms [1] and strongly affect trait expression as well as fitness [2–4]. Social interactions in the context of development (e.g. parental care) often strongly regulate developmental trajectories and resultant adult phenotypes, for example via generationally-transferred compounds such as milk in mammals [5,6], milk-like secretions in arthropods [7,8], and other forms of nutritional provisioning [9,10]. In many taxa including certain birds, mammals, and insects, care for offspring and the regulation of offspring development has shifted at least in part from parents to adult siblings, who perform alloparental care [11]. In eusocial insect societies, sterile nurse workers strictly regulate the development of their larval siblings by modulating the quantity and quality of nourishment larvae receive [12–14], as well as through the direct transfer of growth-regulating hormones and proteins [15,16]. At the same time, larvae manipulate nurse provisioning behavior via pheromones [17–20] and begging behavior [21,22].

In general, traits such as caregiving behavior that are defined or influenced by social interactions are the property of the genomes of multiple interacting social partners [2,14]. This has implications for both the mechanistic (e.g., molecular) underpinnings of development and trait expression as well as the genetic basis of trait variation at the population level -- i.e. how allelic variation in the genomes of interacting social partners affects trait variation [2,14]. Furthermore, because social traits are expressed in

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

one individual but impact the fitness of other individuals, social behavior and socially-influenced traits experience distinct forms of selection, including kin selection and social selection [23,24]. Altogether, these distinct genetic features and patterns of selection are often thought to lead to distinct evolutionary features, such as rapid evolutionary dynamics in comparison to other traits [25–27]. In eusocial insects, previous studies show that variation in larval developmental trajectories and ultimate adult phenotypes (including reproductive caste, body size, etc.) depends on the combination of larval and nurse genotypes [28–34]. However, the identity of specific genes and molecular pathways that are functionally involved in the expression of social interactions (e.g., genes underlying nurse and larval traits affecting nurse-larva interactions), and patterns of molecular evolution for these genes, have remained less well studied [15,16,35,36].

Transcriptomic studies are often used to identify sets of genes underlying the expression of particular traits by performing RNA-sequencing on individuals that vary in the expression of such traits. For example, in social insects, recent studies have compared the transcriptomes of workers that perform nursing versus foraging tasks [37–39], or nurses feeding larvae of different stages or castes [35,40]. However, given the phenotypic co-regulation known to occur between interacting social partners (here, nurses and larvae), it is likely that genes expressed in one social partner affect the expression of genes in the other social partner, and vise-versa, such that interacting social partners are connected by "social" gene regulatory networks [14,32,41,42]. Thus, identifying the genes important for *social interactions* such as nurse-larva interactions is only possible by studying the transcriptomic dynamics of both interacting social partners across a time series of interactions. Similarly, host-symbiont interactions result in gene regulation between interacting partners, and recent studies have sought to reconstruct gene regulatory networks acting between hosts and symbionts by collecting and profiling the transcriptomes of each social partner across a time series of interactions [43–47].

Here, we use analogous methodology to study transcriptomic signatures of nurse-larva interactions in the pharaoh ant, *Monomorium pharaonis*. We sample a developmental time series of larvae as well as the nurses that feed each larval stage in this series, collecting individuals at the moment of

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

interaction in order to identify genes involved in the expression of nurse-larva interactions, as well as genes affected by these interactions (i.e. the full "social interactome" [14]). Given that pharaoh ant nurses tend to specialize on feeding young versus old larvae and nurses feeding young versus old larvae show different transcriptomic profiles [40], and larval transcriptomic profiles also change over development [48,49], we predicted that if nurse-larva interactions have functional implications as described above, we would observe concerted changes in broad-scale gene expression in larvae and their nurses across larval development (Fig 1). Based on our dual RNA-seq data, we infer social gene regulatory networks acting between nurses and larvae to identify genes predicted to have important social regulatory effects. Finally, we combine our measures of social regulatory effects with available population genomic data [48] to characterize the patterns of molecular evolution of genes underlying nurse-larva interactions.

Results

Transcriptome-wide signatures of nurse-larva co-expression across larval development

To elucidate transcriptomic signatures of nurse-larva interactions, we performed RNA-sequencing on worker-destined larvae across five developmental stages and nurses that fed larvae of each developmental stage (termed "stage-specific" nurses; see Fig S1 for sampling scheme, Table S1 for list of samples), building upon a previously published dataset focused on caste development in *M. pharaonis* [48]. We hypothesized that if genes expressed in larvae regulate the expression of genes in nurse and vise versa, we would observe correlated expression profiles across development in larvae and nurses (Fig 1). As a biological control, we also collected "random nurses" that we observed feeding any stage of larvae in the colony, and hence would not be expected to show correlated expression dynamics with larvae across the five larval developmental stages. We also concurrently collected reproductive-destined larvae across the last four developmental stages (reproductive-destined larvae can only be identified after the first larval stage [50]), but unless clearly stated otherwise, all analyses were conducted only on worker-destined larvae. To minimize experimental variance, we pooled ten individuals per sample. We sequenced whole bodies of larvae but separated nurse heads and abdomens prior to sequencing.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We grouped genes into co-expression profiles or "modules" using an algorithm designed to characterize gene co-expression dynamics across a short time series [51]. Each module represents a standardized pre-defined expression profile, consisting of five values that each represent the $\log_2$ fold-change between the given developmental stage and the initial stage (see Fig S2; this results in a total of 81 possible modules). We sorted genes into the module that most closely represented their expression profile by Pearson correlation. We identified modules containing a greater than expected number of genes, where we formed null expectations using permutation tests across developmental stage [51]. We identified such significantly-enriched modules separately for larvae, stage-specific nurse heads, stage-specific nurse abdomens, random nurse heads, and random nurse abdomens. We focused on both parallel (i.e. positive regulation or activation) and anti-parallel (i.e. inhibitory) correlated expression patterns by identifying significantly-enriched modules that were shared in both larvae and nurses (parallel), as well as significantly-enriched modules for which the inverse of the module was identified as significantly-enriched in the social partner (anti-parallel).

Larvae and stage-specific nurses shared many significantly-enriched modules (Table S2). These modules contained the majority of genes expressed in nurses (stage-specific nurse head: 65%; abdomen: 76%) and a substantial proportion of the larval transcriptome (shared with stage-specific nurse head: 22%; abdomen: 60%), such that there was a general widespread signature of correlated transcriptional patterns between stage-specific nurses and larvae across larval development (Fig 2A-D). These coordinated dynamics were dominated by parallel associations in nurse abdomens (possibly reflecting shared metabolic pathways) but anti-parallel associations in nurse heads (possibly reflecting the social regulation of larval growth). In contrast to stage-specific nurses, random nurses (our biological control) shared few significantly-enriched modules with larvae (Table S2), and modules shared between random nurses and larvae contained significantly fewer genes than modules shared between stage-specific nurses and larvae (Fig 2E).

Identification of genes putatively involved in social interactions

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Given that we observed transcriptome-wide patterns consistent with nurse-larva transcriptional co-regulation across larval development, we next identified the genes that might be driving these patterns (see Fig S3). We performed differential expression analysis to identify genes that varied in larval expression according to larval developmental stage, as well as genes that varied in nurse expression according to the developmental stage of larvae they fed. We identified 8125 differentially expressed genes (DEGs) in larvae (78% of 10446 total genes). We identified 2057 and 1408 DEGs in stage-specific nurse heads and abdomens, respectively, compared to 599 and 520 DEGs in random nurse heads and abdomens, respectively. We removed genes differentially expressed in both stage-specific and random nurses (N = 272 DEGs in heads, N = 140 DEGs in abdomens), which might differ among our colony replicates due to random colony-specific effects that were not consistently associated with social regulation of larval development. After this removal, we retained the top 1000 DEGs, sorted by P-value, for each sample type other than random nurses (larvae, stage-specific nurse heads, stage-specific nurse abdomens) for social gene regulatory network reconstruction.

Reconstruction of social gene regulatory networks

To infer gene-by-gene social regulatory relationships between nurses and larvae, we reconstructed gene regulatory networks acting within and between nurses and larvae (Fig S3). To identify the most highly connected (i.e. centrally located, upstream) genes of regulatory networks, we calculated within-tissue connectivity and social connectivity by averaging the strength of connections across each connection a gene made, differentiating between within-tissue (nurse-nurse or larva-larva) and social connections (nurse-larva) (Fig 1B). On average, within-tissue connectivity was higher than social connectivity (Wilcoxon rank-sum test; $P < 0.001$ in all tissues), and within-tissue connectivity was negatively correlated with social connectivity in each tissue (Fig S4). The top enriched gene ontology terms based on social connectivity in nurses were entirely dominated by metabolism (Tables S3,S4).

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Secreted proteins and social gene regulation

While based on our data it is not possible to distinguish between genes that code for protein products that are actually exchanged between nurses and larvae versus genes that affect behavior or physiology within organisms (Fig 1A), proteins that are known to be cellularly secreted represent promising candidates for the social regulation of larval development (40). Genes coding for proteins that are cellularly secreted in *Drosophila melanogaster* had higher social connectivity than genes coding for non-secreted orthologs in nurse heads (Fig 3A; Wilcoxon rank-sum test; $P = 0.011$), though not for nurse abdomens ($P = 0.094$). While many of the genes with the highest social connectivity have unknown function, one of the genes with the highest social connectivity within nurse heads was the protein *giant-lens* (Table S5). *Giant-lens* (also known as *argos*) is a secreted protein that is known to inhibit EGFR signalling in *D. melanogaster* [52]. *Giant-lens* expression in nurse heads was negatively correlated with the expression of the homolog of *eps8*, human EGFR substrate 8 (Fig 3B; note however that this pattern was not seen for all genes in the EGFR pathway). Furthermore, *giant-lens* expression in larvae steadily drops throughout development (Fig S5), suggesting that *giant-lens* expressed in nurses and passed to larvae may inhibit EGFR signaling in larvae, even as EGFR inhibition within larvae via *giant-lens* decreases over the course of development. Interestingly, *eps8* does not exhibit a similar peak and drop in expression level in reproductive-destined larvae in comparison to worker-destined larvae (Fig S6), suggesting that this pathway could be involved in the regulation of caste development.

Molecular evolution of social gene regulatory networks

To investigate the selective pressures shaping social regulatory networks, we used population genomic data from 22 resequenced *M. pharaonis* workers, using one *M. chinense* worker as an outgroup [48]. Using polymorphism and divergence data, we estimated gene-specific values of selective constraint, which represents the intensity of purifying selection that genes experience [53]. To identify genes disproportionately recruited to the core of social regulatory networks, we calculated "sociality index" as the difference between social connectivity and within-tissue connectivity for each gene. Sociality index

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

was negatively correlated to selective constraint due to a positive correlation between within-tissue connectivity and constraint and a negative correlation between social connectivity and constraint (Fig 4A-C). Additionally, genes differed in sociality index according to their estimated evolutionary age, with ancient genes exhibiting lower sociality indices than genes in younger age categories (Fig 4D). Finally, while evolutionary age and evolutionary rate appear to be somewhat empirically confounded [54], selective constraint and evolutionary age were each independently associated with sociality index, based on a model including both variables as well as tissue type (GLM; LRT; evolutionary age: $\chi^2 = 21.536$, $P < 0.001$; selective constraint: $\chi^2 = 22.191$, $P < 0.001$).

Discussion

In organisms with extended offspring care, developmental programs are controlled in part by socially-acting gene regulatory networks that operate between caregivers and developing offspring [14,41]. In this study, we sequenced the transcriptomes of ant nurses and larvae as they interacted across larval development to assess the effects of social interactions on gene expression dynamics. We found that large sets of genes (i.e. modules) expressed in ant larvae and their caregiving adult nurses show correlated changes in expression across development (Fig 2). The majority of nurse and larval transcriptomes was represented in these correlated modules, indicating that the tight phenotypic co-regulation characterizing nurse-larva interactions over the course of larval development is strongly reflected at the molecular level.

To characterize the overall network and evolutionary patterns of genes involved in nurse-larva interactions, we reverse engineered nurse-larva gene regulatory networks and calculated the "social connectivity" for each gene, defined as the sum of inferred social regulatory effects on all genes expressed in social partners. We found that genes with high social connectivity tended to have low within-individual connectivity (Fig S4; where within-individual connectivity is defined as the sum of inferred regulatory effects acting within a given tissue). Nurse-expressed genes with higher sociality indices (i.e disproportionately higher social connectivity than within-individual connectivity) tended to be

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

232 evolutionarily young and rapidly evolving due to relaxed selective constraint (Fig 4). Genes with high
233 social connectivity were enriched for a number of Gene Ontology (GO) categories associated with
234 metabolism (Tables S3,S4), consistent with the idea that molecular pathways associated with metabolism
235 are involved in the expression of social behavior [55,56]. Finally, we found that genes encoding for
236 orthologs of cellularly-secreted proteins important for intercellular signaling in *Drosophila melanogaster*
237 tended to exhibit higher levels of social connectivity than their non-secreted counterparts (Fig 3A).

238 One gene that stands out in terms of being cellularly secreted and exhibiting a relatively high
239 social connectivity is the gene *giant-lens*, which inhibits epidermal growth factor receptor (EGFR)
240 signaling [52]. EGFR signaling affects eye and wing development [57] and body size in *D. melanogaster*
241 [58], caste development in the honey bee *Apis mellifera* [58,59] via the transfer of royalactin from nurses
242 to larvae [58], and worker body size variation in the ant *Camponotus floridanus* [60]. While further
243 experimental work is necessary to ascertain whether *giant-lens* is actually orally secreted by nurses and
244 transferred to larvae, gene expression dynamics suggest that *giant-lens* transferred from nurses to larvae
245 may inhibit EGFR signaling at the end of larval development in worker-destined larvae (Fig 3B), but not
246 in reproductive-destined larvae (Fig S6). While caste in *M. pharaonis* is socially regulated in the first
247 larval stage [49], social inhibition of EGFR signaling could play a role in the regulation of worker body
248 size [60] or secondary caste phenotypes such as wings [61,62].

249 In terms of broad evolutionary patterns, our study complements previous results suggesting genes
250 with worker-biased expression (in comparison to genes with queen-biased expression) tend to be rapidly
251 evolving, evolutionarily young, and loosely connected in regulatory networks [37,48,49,63–65]. Because
252 workers are obligately sterile, their traits are shaped indirectly by kin selection, based on how they affect
253 the reproductive success of fertile relatives (i.e. queens and males) [23,66]. As a result, all-else-equal,
254 genes associated with worker traits are expected to evolve under relaxed selection relative to genes
255 associated with queen traits [67,68]. In general, the suite of genic characteristics commonly associated
256 with worker-biased genes (rapidly evolving, evolutionarily young, loosely connected) are all consistent
257 with relaxed selection acting on genes associated with workers [49]. Here, we show that within the

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

worker caste, genes that appear to be functionally involved in the expression of social behavior (i.e. nursing) experience relaxed selective constraint relative to genes important for within-worker processes. Therefore, the combination of kin selection as well rapid evolution thought to be characteristic of social traits [25] likely act in concert to shape the labile evolutionary patterns commonly associated with worker-biased genes.

In this study, we reconstructed regulatory networks acting between nurses and larvae, which of course assumes that nurse gene expression changes as a function of larval stage fed. This is much more likely to be the case when nurses are specialized on feeding particular larval stages. About half of all feeding events were performed by specialist nurses in a previous study in *M. pharaonis* [40], so our pooled nurse samples likely contained a mix of generalist and specialist nurses. This would serve to effectively dilute the signal of nurse-larva interactions. Clearly, future studies would ideally focus entirely on specialists, as well as on tissues such as brains and the specific exocrine glands [69] known to be important for social behavior and communication. Despite these limitations, we were still able to observe strong transcriptomic signatures of social interactions.

Conclusions

In this study, we uncovered strong transcriptomic signatures of social interactions (i.e. co-regulation of nurse and larval genes and modules) and identified distinct evolutionary features of genes that were inferred to have strong social regulatory roles. Because we simultaneously collected nurses and larvae over a time series of interactions, we were able to elucidate the molecular basis of social interactions. This is a promising approach that could be readily extended to study the molecular underpinnings of all forms of social regulation in social insect colonies, including regulation of foraging, regulation of reproduction, etc.. Furthermore, by adapting the methodology presented here (i.e. simultaneous collection over the course of interactions followed by sequencing), the molecular mechanisms and evolutionary features of genes underlying a diverse array of social interactions, including courtship behavior, dominance hierarchy formation, and regulation of biofilm production could all be investigated. Overall, this study provides

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

284 empirical, theoretical, and methodological foundations on which future research can build upon to realize
285 the full power of the genomics revolution for understanding the genetic underpinnings and evolution of
286 interacting phenotypes.
287
288 ## Methods
289 This study builds on previous work investigating genomic signatures of kin selection in which we
290 characterized transcriptomic profiles from adult queens and workers, as well as queen- and worker-
291 destined larvae [48]. While stage-specific nurses were used in the previous analysis, the knowledge of the
292 developmental stage of larvae they fed was not, as they were simply treated as adult workers. This study
293 complements the past dataset with new data from random nurses, which were collected concurrently with
294 previous samples.
295
296 *Study Design*
297 We created thirty total replicate experimental colonies of approximately equal sizes (~300-400 workers,
298 ~300-400 larvae) from a mixture of many genetically heterogeneous colonies of the ant *Monomorium*
299 *pharaonis*. We removed queens from ½ the study colonies to promote the production of reproductive-
300 destined larvae. In this study, we pooled samples across colonies with and without queens, as queen
301 presence appeared to have a negligible effect on gene expression (0 DEGs at FDR < 0.1 for each sample
302 type). We pre-assigned colonies to one of five larval developmental stages (labeled L1-L5, where L1 and
303 L2 refer to 1st-instar and 2nd-instar larvae and L3, L4, and L5 refer to small, medium, and large 3rd-
304 instar larvae [50]). We sampled individuals (larvae as well as nurses) across larval development time (3-4
305 days between each time point based on the developmental stage of the youngest larvae in colonies without
306 queens). From each colony, we sampled stage-specific nurses and worker-destined larvae, as well as
307 random nurses from colonies with queens and reproductive-destined larvae from colonies without queens
308 (starting at the L2 stage, because at L1 caste cannot be distinguished [50,70]. See Table S1 for full sample
309 list.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

For each time point in each assigned colony, we collected stage-specific nurses, nurses feeding larvae of the specified developmental stage (L1, L2, etc). Concurrently, we haphazardly collected random nurses, nurses we observed feeding a larva of any developmental stage. After collecting nurses, we anaesthetized the colony using carbon dioxide and collected larvae of the specified developmental stage. We separated heads and abdomens of nurses to sequence separately and sequenced whole bodies of larvae. We collected ten samples of each type to pool into single samples (separated by sample type and tissue) for each colony.

We performed RNA-sequencing on all samples concurrently using Illumina HiSeq 2000 at Okinawa Institute of Science and Technology Sequencing Center. Reads were mapped to the NCBI version 2.0 *M. pharaonis* assembly [37], and we used RSEM [71] to estimate counts per locus and fragments per kilobase mapped (FPKM) for each locus. For further details on RNA extraction and library preparation, see [48].

Transcriptome-wide signatures of nurse-larva co-expression across larval development

We used an algorithm that categorizes genes based on their expression dynamics over time into a number of modules represented by pre-defined expression profiles [51]; see Fig S2 for workflow). To create modules, we started at 0 and either doubled, halved, or kept the expression level the same at each subsequent stage, resulting in 81 possible modules (3*3*3*3 = 81; four stages after L1). To generate gene-specific expression profiles based on real results, we calculated the average $\log_2$ fold change in expression (FPKM) of the gene at each developmental stage compared to the initial expression level at stage L1. We then assigned each gene to the closest module by Pearson correlation between gene expression profile and module expression profile [51]. To identify significantly-enriched modules, we generated null distributions of the number of genes present in each module (based on permutation of expression over time), and retained modules with a significantly greater than expected number of genes based on these null distributions (FDR < 0.05 after Bonferroni multiple correction [51]).

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Identification of genes putatively involved in social interactions

We used the package EdgeR [72] to construct models including larval developmental stage and replicate and performed differential expression analysis for each sample type separately. We retained genes differentially expressed according to a nominal P-value of less than 0.05 (i.e. no false discovery correction), as the purpose of this step was simply to identify genes that could be involved in interactions that shape larval development (rather than spurious interactions arising from replicate-specific effects).

Social regulatory network reconstruction

We normalized expression for each gene using the inverse hyperbolic sine transformation of FPKM. As input to the algorithm, we constructed "meta-samples" by combining expression data within the same replicate and time point from nurses and larvae and labeling genes according to the tissue they were expressed in, along the lines of previous host-symbiont studies [46,47]. We utilized the program GENIE3 [73], the top-performing individual program in a wide survey of GRN reconstruction programs [74], to construct two types of networks: those acting between larvae and nurse heads, and those acting between larvae and nurse abdomens. We repeated the reconstruction process 1000 times and averaged pairwise connection strengths across runs. To capture the total effect of each gene on the transcriptome dynamics within tissues, we averaged the regulatory effects each gene had on all other 999 genes expressed in the same tissue ("within-individual connectivity"). Similarly, to capture the effect each gene had on the transcriptome of social partners, we averaged regulatory effects each gene had on the 1000 genes expressed in social partners ("social connectivity").

Estimation of selective constraint, and evolutionary rate

Previously, we performed whole-genome resequencing on 22 diploid *M. pharaonis* workers as well as one diploid *M. chinense* worker to serve as an outgroup [48]. We estimated selective constraint using MKtest2.0 [75], assuming an equal value of alpha (an estimate of the proportion of nonsynonymous substitutions fixed by positive selection) across all genes. Selective constraint is the estimate of the

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

proportion of nonsynonymous mutations that are strongly deleterious and thereby do not contribute to polymorphism or divergence [75]. Selective constraint is estimated using polymorphism data, so it represents the strength of purifying selection genes experience within the study population [53].

Phylostratigraphic Analysis

Phylostrata are hierarchical taxonomic categories, reflecting the most inclusive taxonomic grouping for which an ortholog of the given gene can be found [76–79]. We focused on distinguishing between genes that were evolutionarily "ancient", present in non-insect animals, versus genes present in only insects, hymenopterans, or ants [49]. We constructed a database containing 48 hymenopteran genomes, 10 insect non-hymenopteran genomes, and 10 non-insect animal genomes, and estimated evolutionary age based on the most evolutionarily distant identified BLASTp hit (E-value 10^{-10}).

Gene Set Enrichment Analysis

We performed gene set enrichment analysis based on social connectivity for each gene in each tissue separately using the R package topGO [80]. We identified enriched gene ontology terms using Kolmogorov-Smirnov tests ($P < 0.05$).

General Analyses

We performed all statistical analyses and generated all plots using R version in R version 3.4.0 [81], aided by the packages "reshape2" [82], "plyr" [83], and "ggplot2" [84].

Acknowledgments

We would like to thank the following: Mandy Tin for constructing RNA-sequencing libraries and performing RNA-sequencing, Luigi Pontieri for images of pharaoh ants, Chao Tong for compiling hymenopteran genomes for use in phylostratigraphy, and Rohini Singh for comments on the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

1. Frank SA. All of life is social. Curr Biol. 2007 Aug 21;17(16):R648–50.

2. Moore AJ, Brodie ED III, Wolf JB. Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. Evolution. 1997;51(5):1352–62.

3. Bleakley BH, Wolf JB, Moore AJ. The quantitative genetics of social behaviour. In: Székely T, Moore AJ, Komdeur J, editors. Social behaviour: genes, ecology and evolution. Cambridge, MA: Cambridge University Press; 2010. pp. 29–54.

4. Wolf JB, Brodie ED III, Cheverud JM, Moore AJ, Wade MJ. Evolutionary consequences of indirect genetic effects. Trends Ecol Evol. 1998 Feb 1;13(2):64–9.

5. Bijma P. Estimating maternal genetic effects in livestock. J Anim Sci. 2006 Apr;84(4):800–6.

6. Bouwman AC, Bergsma R, Duijvesteijn N, Bijma P. Maternal and social genetic effects on average daily gain of piglets from birth until weaning. J Anim Sci. 2010;88(9):2883–92.

7. Chen Z, Corlett RT, Jiao X, Liu S-J, Charles-Dominique T, Zhang S, et al. Prolonged milk provisioning in a jumping spider. Science. 2018 Nov 30;362(6418):1052–5.

8. Stay B, Coop AC. "Milk" secretion for embryogenesis in a viviparous cockroach. Tissue and Cell. 1974;6(4):669–93.

9. Moczek AP. Horn polyphenism in the beetle *Onthophagus taurus*: larval diet quality and plasticity in parental investment determine adult body size and male horn morphology. Behav Ecol. 1998 Jan 1;9(6):636–41.

10. Lindström J. Early development and fitness in birds and mammals. Trends Ecol Evol. 1999 Sep;14(9):343–8.

11. Wilson Edward O. Sociobiology: the new synthesis. Cambridge, MA: Belknap; 1975.

12. Linksvayer TA, Kaftanoglu O, Akyol E, Blatch S, Amdam GV, Page RE. Larval and nurse worker control of developmental plasticity and the evolution of honey bee queen–worker dimorphism. J Evol Biol. 2011 Sep 1;24(9):1939–48.

13. Wheeler DE. Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. Am Nat. 1986;128(1):13–34.

14. Linksvayer TA. The molecular and evolutionary genetic implications of being truly social for the social insects. In: Zayed A, Kent CF, editors. Advances in insect physiology. Academic Press; 2015. pp. 271–92.

15. LeBoeuf AC, Waridel P, Brent CS, Gonçalves AN, Menin L, Ortiz D, et al. Oral transfer of chemical cues, growth proteins and hormones in social insects. eLife Sciences. 2016 Nov 29;5:e20375.

16. LeBoeuf AC, Cohanim AB, Stoffel C, Brent CS. Molecular evolution of juvenile hormone esterase-like proteins in a socially exchanged fluid. Scientific Reports 2018 Dec 13;17830.

17. Penick CA, Liebig J. A larval "princess pheromone" identifies future ant queens based on their

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

424 juvenile hormone content. Anim Behav. 2017 June;128:33-40.

425 18. Brian MV. Larval recognition by workers of the ant *Myrmica*. Anim Behav. 1975
426 Nov;23(4):745–56.

427 19. Slessor KN, Winston ML, Le Conte Y. Pheromone communication in the honeybee (*Apis*
428 *mellifera L.*). J Chem Ecol. 2005 Nov;31(11):2731–45.

429 20. Le Conte Y, Sreng L, Poitout SH. Brood pheromone can modulate the feeding behavior of *Apis*
430 *mellifera* workers (Hymenoptera: Apidae). J Econ Entomol. 1995 Aug 1;88(4):798–804.

431 21. Kaptein N, Billen J, Gobin B. Larval begging for food enhances reproductive options in the
432 ponerine ant *Gnamptogenys striatula*. Anim Behav. 2005;69(2):293–9.

433 22. Creemers B, Billen J, Gobin B. Larval begging behaviour in the ant *Myrmica rubra*. Ethol Ecol
434 Evol. 2003 Jul 1;15(3):261–72.

435 23. Hamilton WD. The genetical evolution of social behaviour. I. J Theor Biol. 1964 Jul;7(1):1–16.

436 24. Wolf JB, Brodie ED III, Moore AJ. Interacting phenotypes and the evolutionary process. II.
437 Selection resulting from social interactions. Am Nat. 1999;153(3):254–66.

438 25. West-Eberhard MJ. Sexual selection, social competition, and speciation. Q Rev Biol.
439 1983;58(2):155–83.

440 26. McGlothlin JW, Moore AJ, Wolf JB, Brodie ED III. Interacting phenotypes and the evolutionary
441 process. III. Social evolution. Evolution. 2010 Sep;64(9):2558–74.

442 27. Bailey NW, Marie-Orleach L, Moore AJ, Simmons L. Indirect genetic effects in behavioral
443 ecology: does behavior play a special role in evolution? Behav Ecol. 2018 Jan 13;29(1):1–11.

444 28. Osborne KE, Oldroyd BP. Possible causes of reproductive dominance during emergency queen
445 rearing by honeybees. Anim Behav. 1999 Aug;58(2):267–72.

446 29. Beekman M, Oldroyd BP. Effects of cross-feeding anarchistic and wild type honey bees:
447 anarchistic workers are not queen-like. Naturwissenschaften. 2003 Apr;90(4):189–92.

448 30. Linksvayer TA. Direct, maternal, and sibsocial genetic effects on individual and colony traits in
449 an ant. Evolution. 2006 Dec;60(12):2552–61.

450 31. Linksvayer TA. Ant species differences determined by epistasis between brood and worker
451 genomes. PLoS One. 2007 Oct 3;2(10):e994.

452 32. Linksvayer TA, Fondrk MK, Page RE Jr. Honeybee social regulatory networks are shaped by
453 colony-level selection. Am Nat. 2009 Mar;173(3):E99–107.

454 33. Teseo S, Châline N, Jaisson P, Kronauer DJC. Epistasis between adults and larvae underlies
455 caste fate and fitness in a clonal ant. Nat Commun. 2014;5:3363.

456 34. Villalta I, Blight O, Angulo E, Cerdá X, Boulay R. Early developmental processes limit socially
457 mediated phenotypic plasticity in an ant. Behav Ecol Sociobiol. 2016 Feb 1;70(2):285–91.

458 35. Vojvodic S, Johnson BR, Harpur BA, Kent CF, Zayed A, Anderson KE, et al. The transcriptomic

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

459 and evolutionary signature of social interactions regulating honey bee caste development. Ecol Evol.
460 2015;5(21):4795–807.

461 36. Benowitz KM, McKinney EC, Cunningham CB, Moore AJ. Predictable gene expression related
462 to behavioral variation in parenting. Behav Ecol;ary179;doi:10.1093/beheco/ary179.

463 37. Mikheyev AS, Linksvayer TA. Genes associated with ant social behavior show distinct
464 transcriptional and evolutionary patterns. eLife. 2015 Jan 26;4:e04775.

465 38. Kohlmeier P, Alleman AR, Libbrecht R, Foitzik S, Feldmeyer B. Gene expression is more
466 strongly associated with behavioural specialisation than with age or fertility in ant workers. Mol
467 Ecol; 2018 Dec 7;doi:10.1111/mec.14971.

468 39. Manfredini F, Lucas C, Nicolas M, Keller L, Shoemaker D, Grozinger CM. Molecular and social
469 regulation of worker division of labour in fire ants. Mol Ecol. 2014 Feb;23(3):660–72.

470 40. Walsh JT, Warner MR, Kase A, Cushing BJ, Linksvayer TA. Ant nurse workers exhibit
471 behavioural and transcriptomic signatures of specialization on larval stage. Anim Behav. 2018 Jul
472 1;141:161–9.

473 41. Linksvayer TA, Fewell JH, Gadau J, Laubichler MD. Developmental evolution in social insects:
474 regulatory networks from genes to societies. J Exp Zool B Mol Dev Evol. 2012 May;318(3):159–69.

475 42. Bloch G, Grozinger CM. Social molecular pathways and the evolution of bee societies. Philos
476 Trans R Soc Lond B Biol Sci. 2011 Jul 27;366(1574):2155–70.

477 43. Burns JA, Zhang H, Hill E, Kim E, Kerney R. Transcriptome analysis illuminates the nature of
478 the intracellular interaction in a vertebrate-algal symbiosis. eLife 2017;6:e22054.

479 44. Westermann AJ, Förstner KU, Amman F, Barquist L, Chao Y, Schulte LN, et al. Dual RNA-seq
480 unveils noncoding RNA functions in host–pathogen interactions. Nature. 2016 Jan
481 20;529(7587):496–501.

482 45. Westermann AJ, Gorski SA, Vogel J. Dual RNA-seq of pathogen and host. Nat Rev Microbiol.
483 2012 Sep;10(9):618–30.

484 46. Schulze S, Henkel SG, Driesch D, Guthke R, Linde J. Computational prediction of molecular
485 pathogen-host interactions based on dual transcriptome data. Front Microbiol. 2015 Feb 6;6:65.

486 47. Tierney L, Linde J, Müller S, Brunke S, Molina JC, Hube B, et al. An interspecies regulatory
487 network inferred from simultaneous RNA-seq of *Candida albicans* invading innate immune cells.
488 Front Microbiol. 2012 Mar 12;3:85.

489 48. Warner MR, Mikheyev AS, Linksvayer TA. Genomic signature of kin selection in an ant with
490 obligately sterile workers. Mol Biol Evol. 2017 Jul 1;34(7):1780–7.

491 49. Warner MR, Qiu L, Holmes MJ, Mikheyev AS. The convergent evolution of caste in ants and
492 honey bees is based on a shared core of ancient reproductive genes and many plastic genes. Preprint.
493 Available from: https://www.biorxiv.org/content/early/2018/10/26/454645.abstract

494 50. Warner MR, Kovaka K, Linksvayer TA. Late-instar ant worker larvae play a prominent role in
495 colony-level caste regulation. Insectes Soc. 2016 Nov 1;63(4):575–83.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

51. Ernst J, Nau GJ, Bar-Joseph Z. Clustering short time series gene expression data. Bioinformatics. 2005 Jun;21 Suppl 1:i159–68.

52. Schweitzer R, Howes R, Smith R, Shilo BZ, Freeman M. Inhibition of *Drosophila* EGF receptor activation by the secreted protein *Argos*. Nature. 1995 Aug 24;376(6542):699–702.

53. Kreitman M, Hudson RR. Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics. 1991 Mar;127(3):565–82.

54. Moyers BA, Zhang J. Evaluating phylostratigraphic evidence for widespread de novo gene birth in genome evolution. Mol Biol Evol. 2016 May;33(5):1245–56.

55. Toth AL, Robinson GE. Evo-devo and the evolution of social behavior. Trends Genet. 2007 Jul;23(7):334–41.

56. Toth AL, Varala K, Henshaw MT, Rodriguez-Zas SL, Hudson ME, Robinson GE. Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. Proc Biol Sci. 2010 Jul 22;277(1691):2139–48.

57. Shilo BZ. Signaling by the *Drosophila* epidermal growth factor receptor pathway during development. Exp Cell Res. 2003 Mar 10;284(1):140–9.

58. Kamakura M. Royalactin induces queen differentiation in honeybees. Nature. 2011 May 26;473(7348):478–83.

59. Mutti NS, Dolezal AG, Wolschin F, Mutti JS, Gill KS, Amdam GV. IRS and TOR nutrient-signaling pathways act via juvenile hormone to influence honey bee caste fate. J Exp Biol. 2011 Dec 1;214(23):3977–84.

60. Alvarado S, Rajakumar R, Abouheif E, Szyf M. Epigenetic variation in the *Egfr* gene generates quantitative variation in a complex trait in ants. Nat Commun. 2015 Mar 11;6:6513.

61. Abouheif E, Wray GA. Evolution of the gene network underlying wing polyphenism in ants. Science. 2002 Jul 12;297(5579):249–52.

62. Trible W, Kronauer DJC. Caste development and evolution in ants: it's all about size. J Exp Biol. 2017 Jan 1;220(Pt 1):53–62.

63. Morandin C, Tin MMY, Abril S, Gómez C, Pontieri L, Schiøtt M, et al. Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. Genome Biol. 2016 Mar 7;17:43.

64. Harpur BA, Kent CF, Molodtsova D, Lebon JMD, Alqarni AS, Owayss AA, et al. Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. Proc Natl Acad Sci U S A. 2014 Feb 18;111(7):2614–9.

65. Johnson BR, Tsutsui ND. Taxonomically restricted genes are associated with the evolution of sociality in the honey bee. BMC Genomics. 2011 Mar 29;12:164.

66. Bourke AFG, Franks NR. Social evolution in ants. Princeton, NJ: Princeton University Press; 1995.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

533 67. Linksvayer TA, Wade MJ. Genes with social effects are expected to harbor more sequence
534 variation within and between species. Evolution. 2009 Jul;63(7):1685–96.

535 68. Linksvayer TA, Wade MJ. Theoretical predictions for sociogenomic data: the effects of kin
536 selection and sex-limited expression on the evolution of social insect genomes. Front Ecol Evol;
537 2016 June;4:65.

538 69. Eelen D, Børgesen L, Billen J. Functional morphology of the postpharyngeal gland of queens
539 and workers of the ant *Monomorium pharaonis (L.)*. Acta Zool. 2006;87(2):101–11.

540 70. Warner MR, Lipponen J, Linksvayer TA. Pharaoh ant colonies dynamically regulate
541 reproductive allocation based on colony demography. Behav Ecol Sociobiol. 2018 Mar 1;72(3):31.

542 71. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without
543 a reference genome. BMC Bioinformatics. 2011 Aug 4;12:323.

544 72. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential
545 expression analysis of digital gene expression data. Bioinformatics. 2010 Jan 1;26(1):139–40.

546 73. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from
547 expression data using tree-based methods. PLoS One 2010 Sep;doi:10.1371/journal.pone.0012776.

548 74. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, et al. Wisdom of crowds
549 for robust gene network inference. Nat Methods. 2012 Jul 15;9(8):796–804.

550 75. Welch JJ. Estimating the genomewide rate of adaptive protein evolution in *Drosophila*. Genetics.
551 2006 Jun;173(2):821–37.

552 76. Quint M, Drost H-G, Gabel A, Ullrich KK, Bönn M, Grosse I. A transcriptomic hourglass in
553 plant embryogenesis. Nature. 2012 Oct 4;490(7418):98–101.

554 77. Drost H-G, Gabel A, Grosse I, Quint M. Evidence for active maintenance of phylotranscriptomic
555 hourglass patterns in animal and plant embryogenesis. Mol Biol Evol. 2015 May;32(5):1221–31.

556 78. Domazet-Lošo T, Tautz D. A phylogenetically based transcriptome age index mirrors
557 ontogenetic divergence patterns. Nature. 2010 Dec 9;468(7325):815–8.

558 79. Domazet-Lošo T, Brajković J, Tautz D. A phylostratigraphy approach to uncover the genomic
559 history of major adaptations in metazoan lineages. Trends Genet. 2007;23(11):533–9.

560 80. Alexa A, Rahnenfuhrer J. topGO: enrichment analysis for gene ontology. R package version
561 2.28.0.

562 81. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R
563 Foundation for Statistical Computing; 2017.

564 82. Wickham H. Reshaping data with the reshape package. J Stat Soft. 2007;21(12):1-20.

565 83. Wickham H. The split-apply-combine strategy for data analysis. J Stat Soft. 2011;40(1):1-29.

566 84. Wickham H. 2016. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag;
567 2016.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

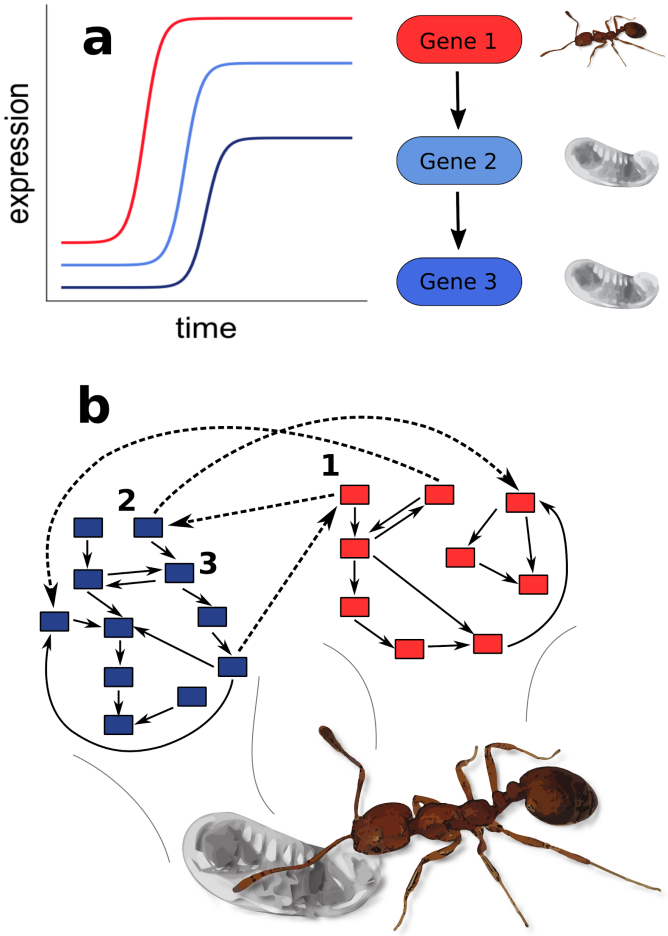

Fig 1. Social regulation of gene expression between ant nurses and larvae.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

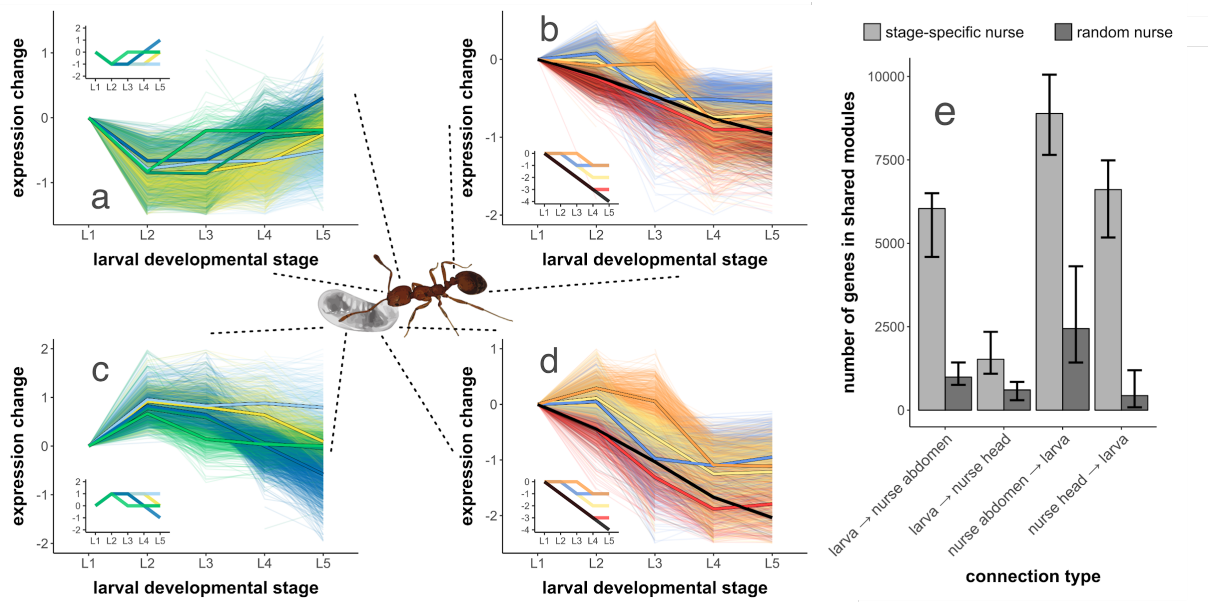

Fig 2. Nurse and larval transcriptomes show strong signatures of gene co-expression across larval development.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

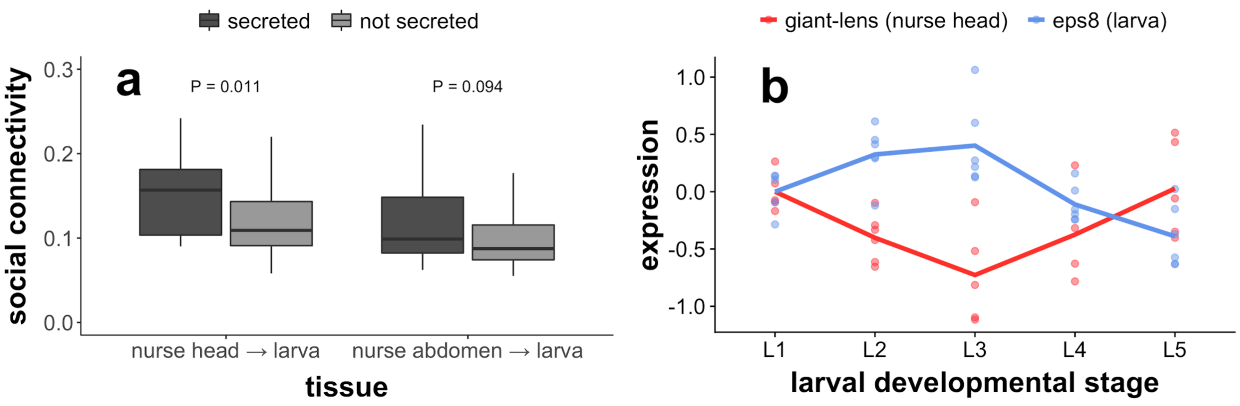

Fig 3. Genes encoding secreted proteins such as *giant-lens* are important for social gene regulation.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

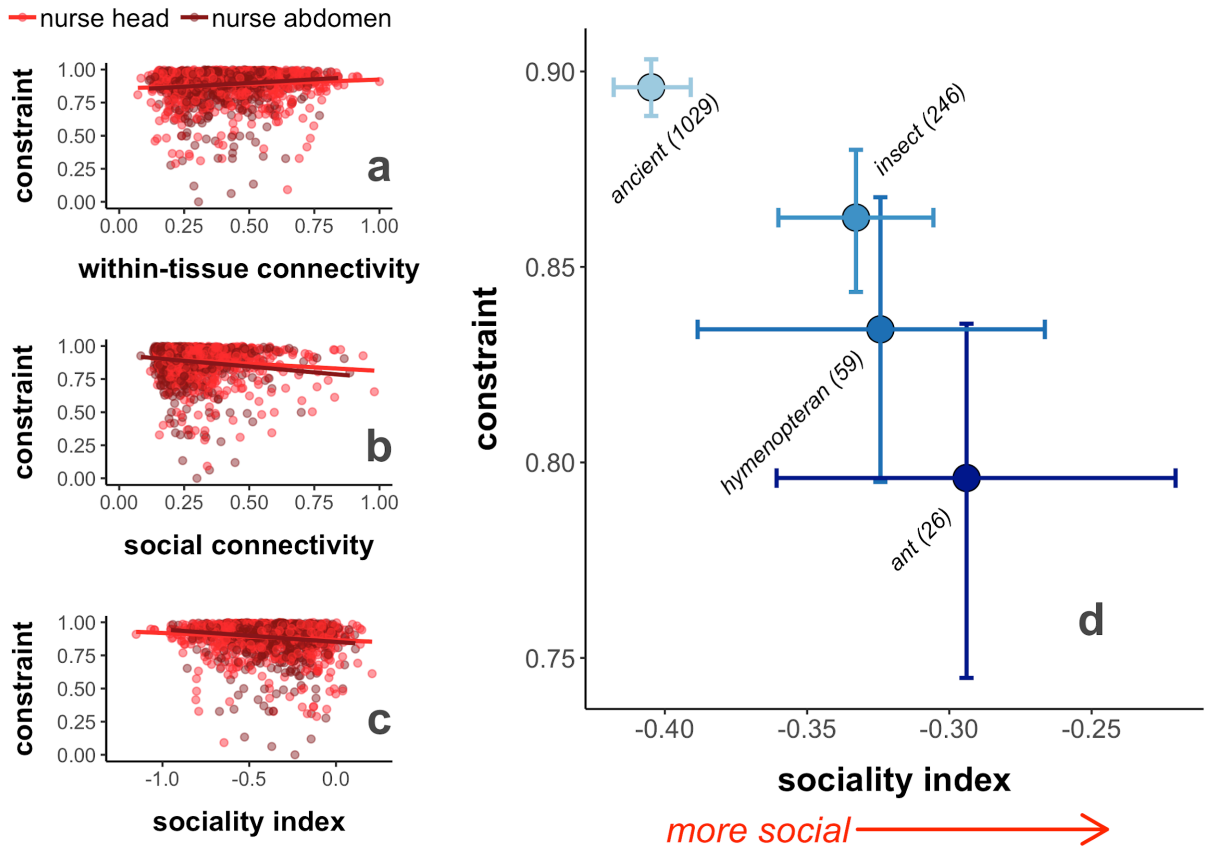

Fig 4. Highly social genes tend to be less evolutionarily constrained.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

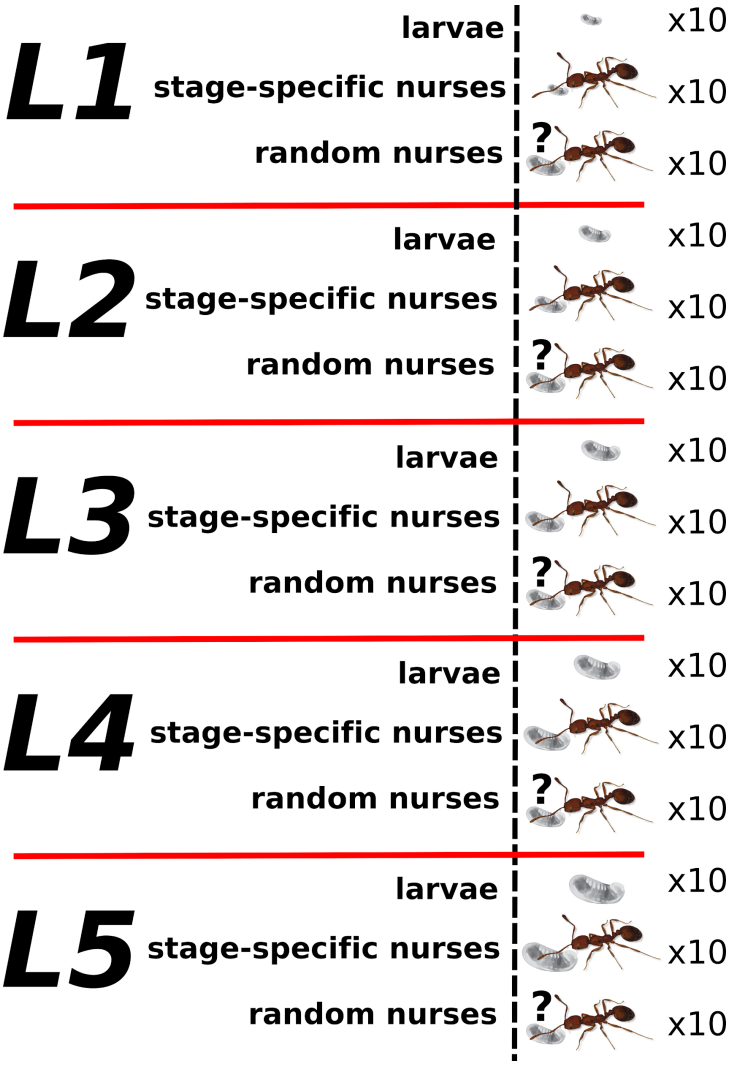

Fig S1. Diagram of sampling scheme.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

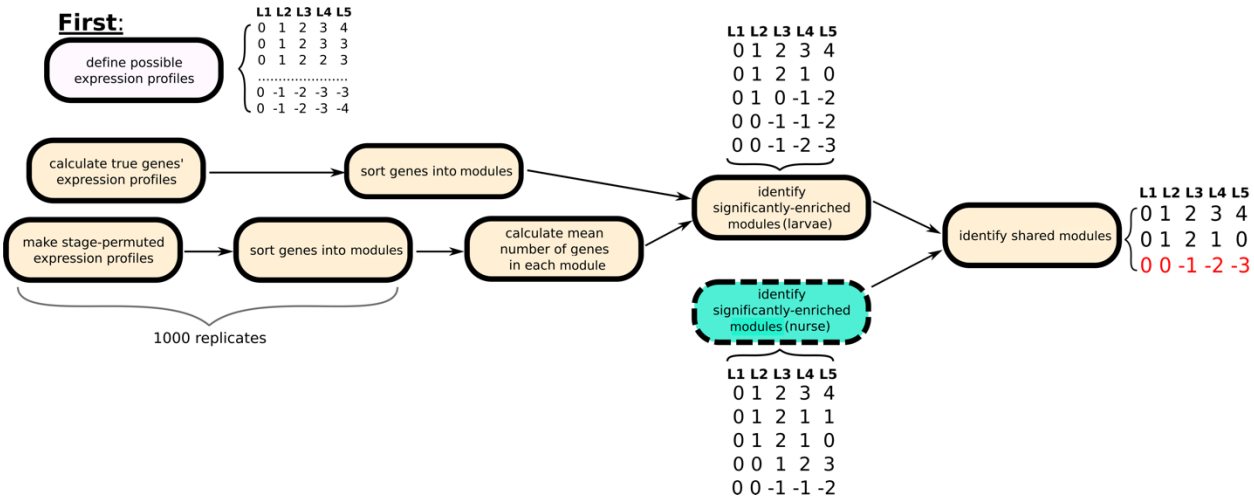

Fig S2. Identification of significantly-enriched modules shared between larvae and nurses.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

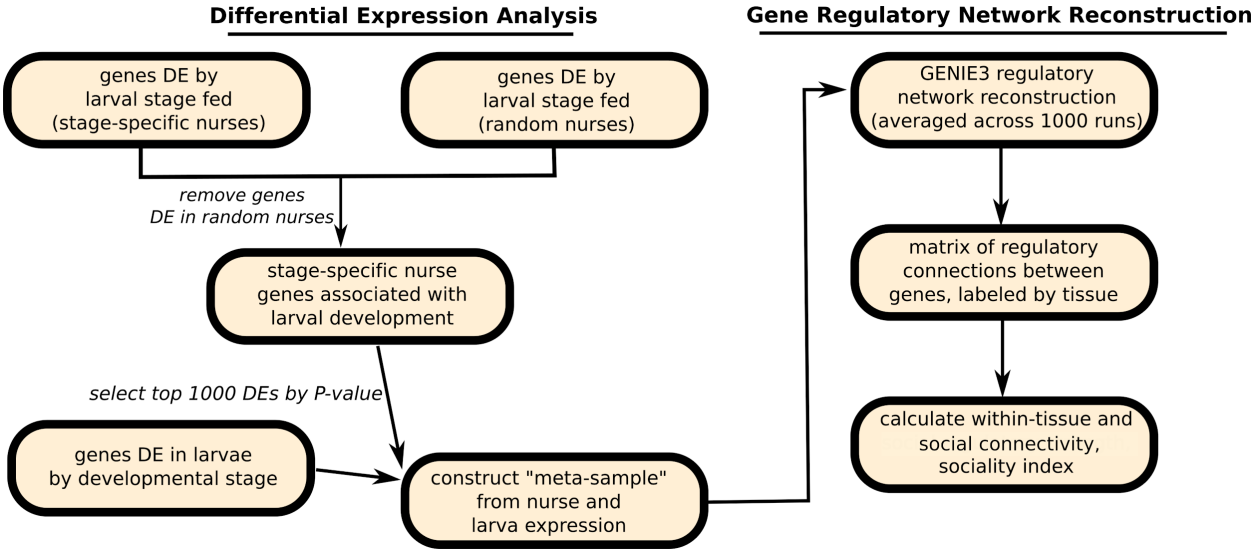

Figure S3. Workflow of preliminary differential expression analysis and gene regulatory network reconstruction.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

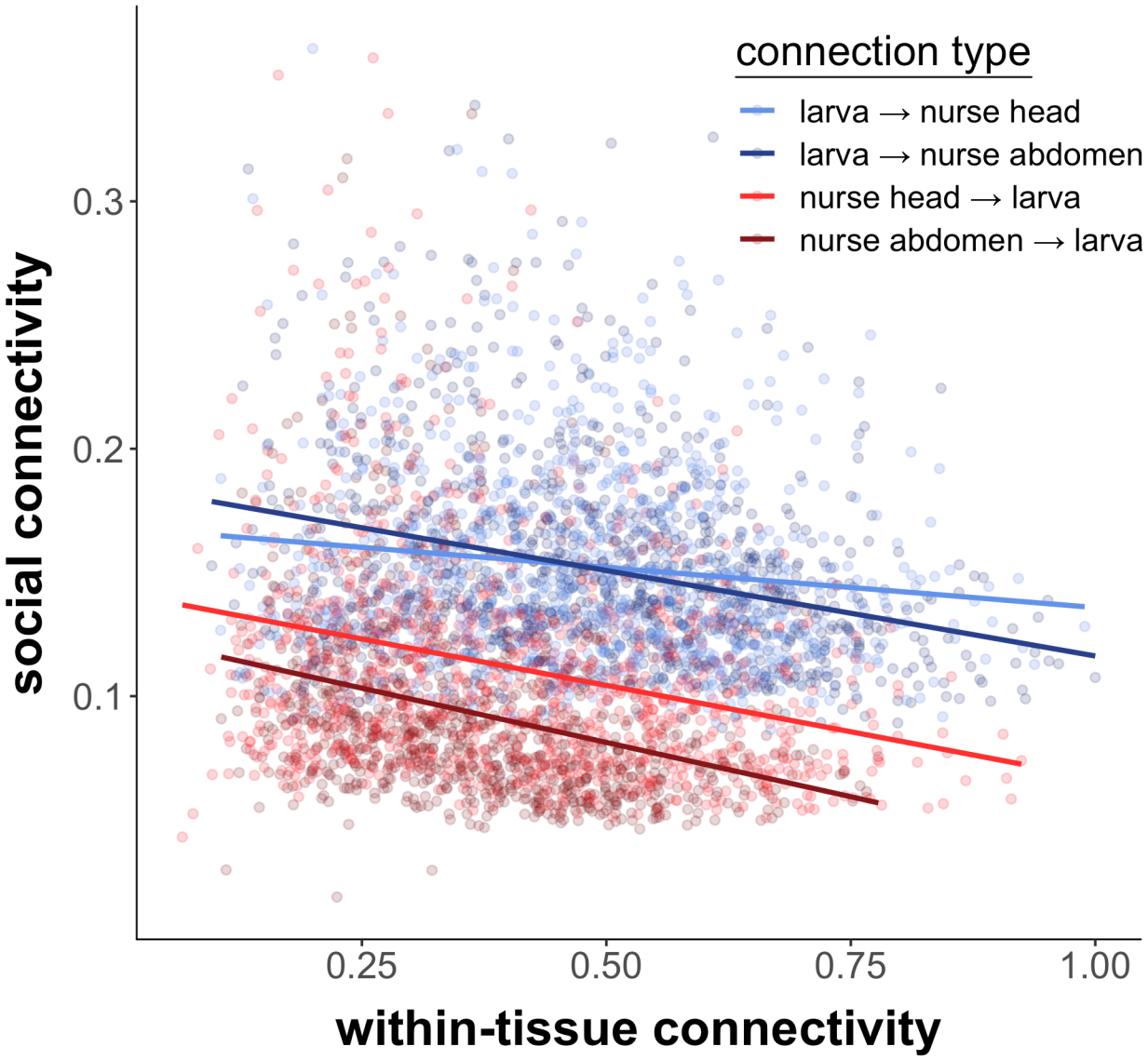

Figure S4. Genes highly connected in within-tissue networks tended to have low values of social connectivity.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

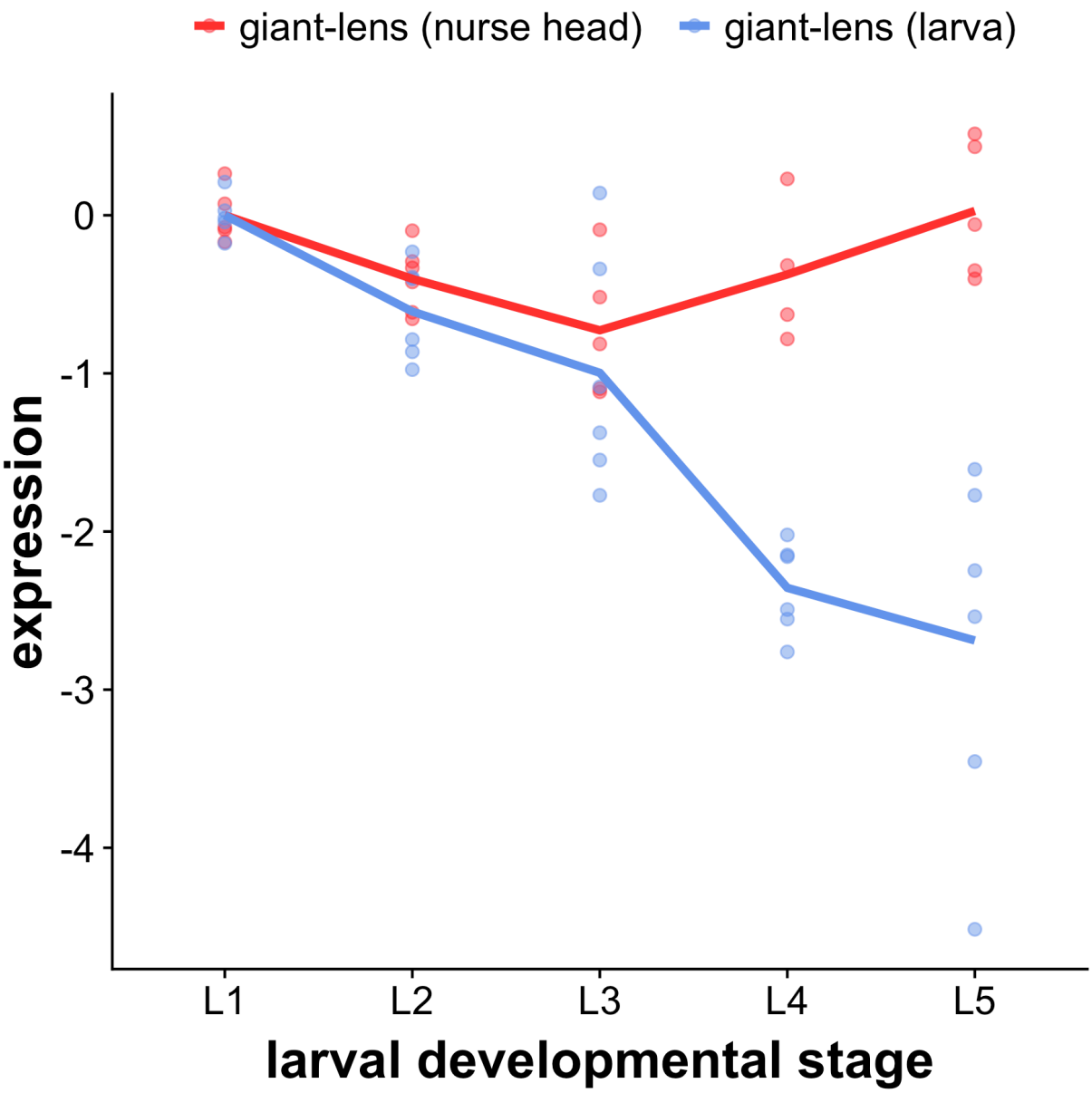

Fig S5. Expression of *giant-lens* in nurse heads and worker-destined larvae.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

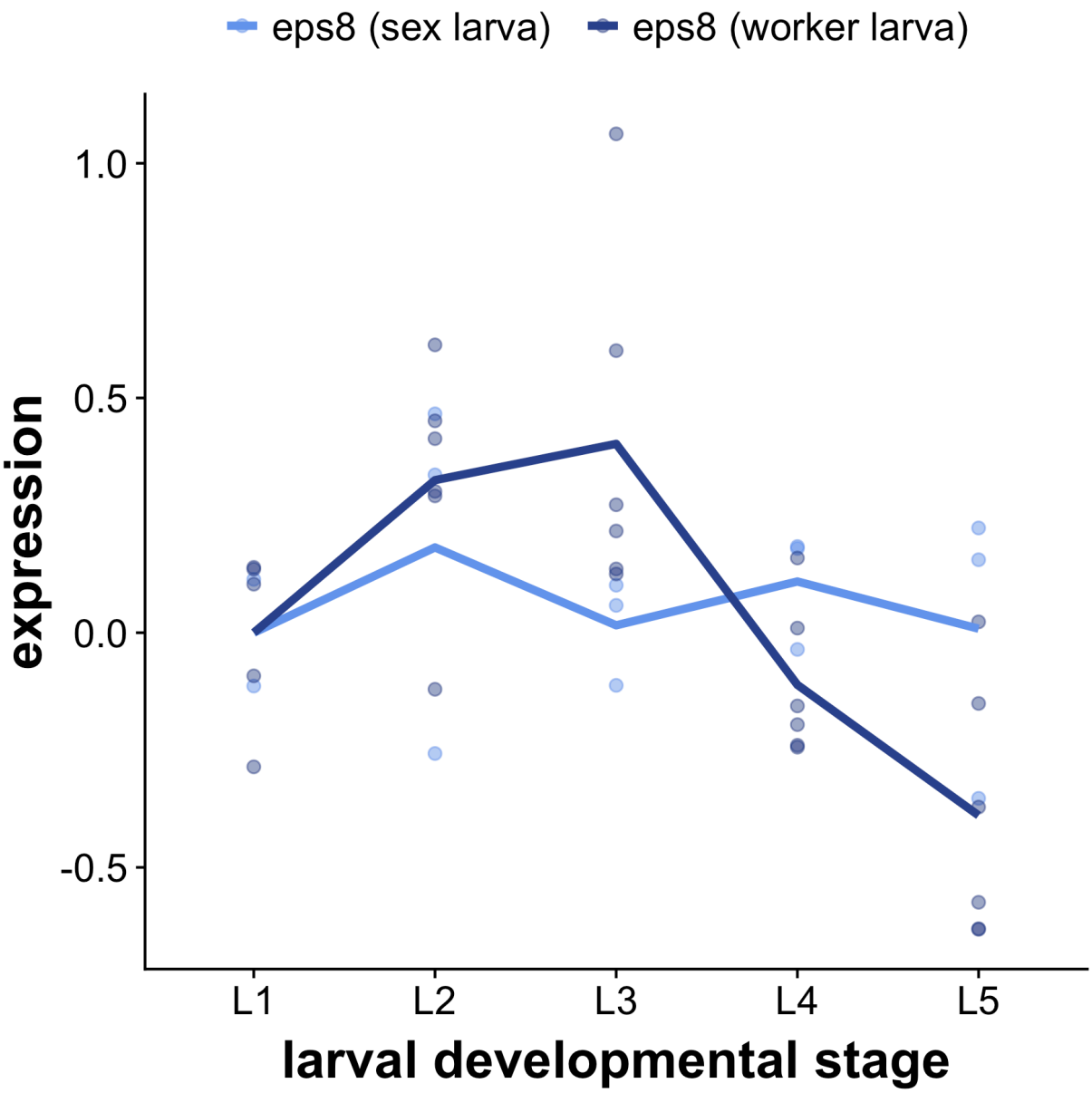

Fig S6. Expression of *eps8* (epidermal growth factor receptor substrate 8) in worker-destined and reproductive-destined larvae.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

development stage	sample type	number of samples
L1	larva (W/R)	5
	stage-specific nurse head	5
	stage-specific nurse abdomen	6
	random nurse head	3
	random nurse abdomen	2
L2	larva (W)	6
	larva (R)	3
	stage-specific nurse head	6
	stage-specific nurse abdomen	5
	random nurse head	3
	random nurse abdomen	3
L3	larva (W)	6
	larva (R)	3
	stage-specific nurse head	5
	stage-specific nurse abdomen	6
	random nurse head	3
	random nurse abdomen	3
L4	larva (W)	6
	larva (R)	3
	stage-specific nurse head	4
	stage-specific nurse abdomen	5
	random nurse head	2
	random nurse abdomen	2
L5	larva (W)	6
	larva (R)	3
	stage-specific nurse head	5
	stage-specific nurse abdomen	5
	random nurse head	3
	random nurse abdomen	3

672

673 **Table S1. Description of samples included in study.**
674
675

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

	number of significantly-enriched modules	modules positively shared with larvae	modules negatively shared with larvae	number of genes in shared modules
stage-specific nurse head	9	0	5	6838
random nurse head	9	0	2	209
stage-specific nurse abdomen	21	13	4	7943
random nurse abdomen	10	0	1	1400

Table S2. Number of nurse significantly-enriched modules shared with larvae.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0044710	single-organism metabolic process	81	14	7.82	0.00011
GO:0055114	oxidation-reduction process	45	11	4.35	0.00230
GO:0019637	organophosphate metabolic process	18	1	1.74	0.00506
GO:0044711	single-organism biosynthetic process	18	0	1.74	0.00506
GO:0006629	lipid metabolic process	12	2	1.16	0.00565
GO:0009117	nucleotide metabolic process	14	0	1.35	0.00755
GO:0006812	cation transport	18	2	1.74	0.00781
GO:0015672	monovalent inorganic cation transport	11	1	1.06	0.00898
GO:0090407	organophosphate biosynthetic process	11	0	1.06	0.01033
GO:0055086	nucleobase-containing small molecule metabolic process	16	0	1.55	0.01156

Table S3. Nurse head social connectivity GO terms based on GSEA of social connectivity.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

GO.ID	Term	Annotated	Significant	Expected	P-value
GO:0055114	oxidation-reduction process	52	7	4.96	0.022
GO:0008152	metabolic process	246	25	23.46	0.033

Table S4. Nurse abdomen GO terms based on GSEA of social connectivity.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

nurse head	nurse abdomen
Basement membrane-specific heparan sulfate proteoglycan core protein	Procollagen-lysine 2-oxoglutarate 5-dioxygenase 3
Collagen alpha-1(IV) chain	Collagen alpha-1(IV) chain
Spondin-1	Tubulointerstitial nephritis antigen-like
Serine proteinase stubble	Papilin
Angiotensin-converting enzyme (Fragment)	Semaphorin-2A
Thrombospondin-4	Transferrin
Protein giant-lens	Basement membrane-specific heparan sulfate proteoglycan core protein
Protein lev-9	Protein NPC2 homolog
Papilin	Testican-2
Glypican-6	Lysozyme

Table S5. Swiss-Prot annotations for the top genes top genes coding for secreted proteins, sorted by social connectivity.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure Captions

Fig 1. Social regulation of gene expression between ant nurses and larvae.

(A) Cartoon depicting positive gene regulation (i.e. activation) between larvae and nurses, where gene 1 is expressed in nurses and genes 2 and 3 are expressed in larvae. After the expression of gene 1 increases, the expression of gene 2 increases as a result of the social interaction of nursing (depicted in [B]). This can occur if gene 1 itself codes for a protein passed to larvae, if the mRNA transcript is passed directly, or if gene 1 activates the expression of some other gene in nurses, which in turn is passed as mRNA (or codes for a protein that is passed) to larvae. Following the increase in expression of gene 2, the expression of gene 3, which is shown to be activated by gene 2, also increases. While we have depicted a time-lag in this social regulation of gene expression, the time lags are likely too short to observe in our data, as larvae were collected every 3-4 days across development. Therefore, correlated transcriptome dynamics over development (see Fig 2) would reflect mechanisms shown here. (B) Gene regulatory networks act between and within individuals engaged in social interactions. Blue boxes are genes expressed in larvae, and red boxes are genes expressed in nurses. Solid lines depict regulatory interactions within tissues (here, within larvae or within nurses), while dashed lines represent social connections (nurse-larva or vise versa).

Fig 2. Nurse and larval transcriptomes show strong signatures of gene co-expression across larval development.

Plots (A-D) depict the expression profiles of individual genes (light lines) as expressed in (A) nurse head, and (B) nurse abdomens, as well as larvae (C) larvae, shared with nurse heads, and (D) larvae, shared with nurse abdomens. Dark lines indicate the median expression values of all genes sorted into modules, with pre-defined expression profiles of modules depicted in plot insets. Only the five shared modules containing the most nurse-expressed genes are shown for clarity. Larval expression profiles are divided by the nurse tissue they are shared with, such that (C) depicts larval gene expression shared with nurse

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

803 heads (A), while (D) depicts larval gene expression shared with nurse abdomens (B). Note that nurse
804 heads and larvae shared inversely-related expression profiles. (E) Stage-specific nurses have more genes
805 than random nurses in modules shared with larvae than do random nurses, reflecting more broad-scale co-
806 expression across development. Error bars indicate 95% confidence intervals derived from systematic
807 drop-1 jackknifing of nurse samples. N = 10944 genes total.
808
809 **Fig 3. Genes encoding secreted proteins such as *giant-lens* are important for social gene regulation.**
810 (A) Genes encoding for proteins that are secreted in *Drosophila melanogaster* exhibit higher social
811 connectivity (i.e. more strongly socially regulate larval expression) in nurse heads than genes encoding
812 for non-secreted proteins (Wilcoxon rank-sum test). (B) The protein *giant-lens* is one of the genes coding
813 for secreted proteins with the highest social connectivity in nurse heads. Based on our data, *giant-lens*
814 expressed in stage-specific nurse heads (red) appears to inhibit the expression of the homolog of human
815 EGFR substrate 8 (*eps8*) expressed in worker-destined larvae (blue). Expression at stage *i* is equal to
816 $\log_2(\text{expression}_i/\text{expression}_1)$, i.e. the ratio of expression at the given stage to expression at the initial
817 stage.
818
819 **Fig 4. Highly social genes tend to be less evolutionarily constrained.**
820 Selective constraint, estimated from whole-genome polymorphism data, is (A) positively correlated with
821 within-tissue connectivity (Spearman correlation; head: $r = 0.122$, $P < 0.001$; abdomen: $r = 0.217$, $P <$
822 0.001), but negatively correlated with (B) social connectivity (head: $r = -0.090$, $P = 0.009$; abdomen: $r = -$
823 0.150, $P < 0.001$) and (C) sociality index (head: $r = -0.132$, $P < 0.001$; abdomen: $r = -0.223$, $P < 0.001$),
824 where sociality index is the difference between social and within-tissue connectivity per gene. Each point
825 in (a-c) indicates a single gene, as expressed in nurse heads or abdomens. Lines indicate trendline from
826 linear model. (D) Highly social genes also tend to be taxonomically restricted. Individual points depict
827 average values across nurse heads and abdomens for all genes within each estimated age class, indicated

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

828 by labels on points. Error bars depict 95% confidence intervals from bootstrapping. Numbers in
829 parentheses indicate number of genes in each age class.
830
831 **Fig S1. Diagram of sampling scheme.**
832 We collected ten worker-destined larvae, ten stage-specific nurses, and ten random nurses from each
833 colony (six colonies per time point, where time points represent larval developmental stages L1, L2, etc).
834 We collected stage-specific nurses when we observed them feeding larvae of the given developmental
835 stage. We collected random nurses when we observed them feeding any stage of larvae. We identified the
836 five separate stages of larvae as in [50].
837
838 **Fig S2. Identification of significantly-enriched modules shared between larvae and nurses.**
839 Inset tables depict expression profiles of sample modules genes can be assigned to. First, we construct
840 modules using all possible expression profiles (top left bubble). Expression profiles consist of five values,
841 starting at zero, that indicate the $\log_2$ fold-change in expression from the initial value (at stage L1). At
842 each subsequent stage, we either double, halve, or keep the expression level the same. This process is
843 repeated to produce 81 (four stages after L1; 3*3*3*3 = 81) total modules. Next, for each tissue
844 separately (here we depict workflow in larvae with yellow bubbles), we calculate individual gene
845 expression profiles as the $\log_2$ fold-change in expression from the initial value and assign genes to the
846 closest related module by Pearson correlation. Concurrently, we permute the developmental stage labels
847 for each gene and assign the stage-permuted genes to modules (repeated 1000 times). From these stage-
848 permuted results, we calculate the mean number of genes assigned to each module and treat this number
849 as a null expectation (as each expression profile is not equally likely to occur by chance). We then
850 identify significantly-enriched modules using a one-way binomial test (with the calculated mean as the
851 null), with a Bonferroni-corrected false discovery rate of 0.05. This entire process is repeated in a nurse
852 tissue and significantly-enriched modules are found (blue bubble). Finally, we compare significantly-
853 enriched modules between larvae and nurses and retain identical and inverse modules as shared profiles.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

854 An example of an inversely related profile is shown in red, where larvae exhibit the enriched module [0,
855 0, -1, -2, -3] and nurses exhibit the inverse module, [0, 0, 1, 2, 3].
856
857 **Figure S3. Workflow of preliminary differential expression analysis and gene regulatory network**
858 **reconstruction.**
859 On the left, we identify putatively socially-acting genes through differential expression analysis. First, for
860 nurse heads and abdomens separately, we perform differential expression analysis in stage-specific and
861 random nurses to identify genes differentially expressed according to larval stage fed, using a nominal P-
862 value of 0.05. We remove genes differentially expressed in random nurses, as these correspond to colony-
863 specific environmental effects unrelated to social regulation of larval development. Next, we select the
864 top 1000 differentially expressed genes by P-value in stage-specific nurses (after removing those DE in
865 random nurses) as well as the top 1000 differentially expressed genes in larvae. From these genes, we
866 create "meta-samples" by combining gene expression of larvae and stage-specific nurses collected from
867 the same colony (separately for heads and abdomens), and labeling genes by the tissue they are expressed
868 in. Using these meta-samples, we perform gene regulatory reconstruction (right) to identify genes
869 expressed in nurses that regulate larval gene expression, and vise-versa. We repeat gene regulatory
870 reconstruction 1000 times and average connection strength across runs, as the algorithm is non-
871 deterministic. The output of gene regulatory reconstruction is a matrix of regulatory connections acting
872 between genes. From this matrix, we calculate the average connectivity for each gene, separating within-
873 tissue (larva-larva or nurse head-nurse head) from social (nurse-larva) connections. Genes with high
874 connectivity are predicted to interact with many genes, i.e. are central to the network. Finally, we
875 calculate each genes' sociality index as the difference between social connectivity and within-tissue
876 connectivity.
877
878 **Figure S4. Genes highly connected in within-tissue networks tended to have low values of social**
879 **connectivity.**

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Connectivity is representative of the number and strength of regulatory connections each gene makes. Points indicate the average connectivity for a given gene, as measured within-tissue (x-axis; i.e. larva-larva or nurse-nurse) or socially (y-axis; i.e. larva-nurse). Points are colored by tissue the connectivity is measured in (e.g., dark blue indicates genes expressed in larvae, with connectivity measured in networks constructed with nurse abdomens). Spearman r = -0.166, -0.374, -0.276, -0.342 for the four tissues as ordered in legend; $P < 0.001$ in all cases.

Fig S5. Expression of *giant-lens* in nurse heads and worker-destined larvae. Expression at stage *i* is equal to $\log_2(\text{expression}_i/\text{expression}_1)$, i.e. the ratio of expression at the given stage to expression at the initial stage.

Fig S6. Expression of *eps8* (epidermal growth factor receptor substrate 8) in worker-destined and reproductive-destined larvae. Expression at stage *i* is equal to $\log_2(\text{expression}_i/\text{expression}_1)$, i.e. the ratio of expression at the given stage to expression at the initial stage.

Table S1. Description of samples included in study. Worker-destined larvae are indicated by larva (W), and reproductive-destined larvae are indicated by larva (R). Larval caste cannot be distinguished at the L1 stage, so L1 larvae are labeled larva (W/R). For network reconstruction, "meta" samples were used as input for network reconstruction, in which genes were labeled by sample type and grouped such that each gene contained a measurement of expression in worker-destined larvae, nurse heads, and nurse abdomens. Because sampling was uneven after removing low-quality samples, we used the minimum number of samples contained across tissues at a given stage for stage-specific nurse heads and abdomens, and randomly dropped excess samples. Overall, 24 "aggregate" samples were used as input for gene regulatory network reconstruction.

Table S2. Number of nurse significantly-enriched modules shared with larvae.

bioRxiv preprint doi: https://doi.org/10.1101/514356; this version posted January 8, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

905 Significantly-enriched modules are defined as modules with a statistically significant number of genes
906 assigned, as determined by a permutation test (FDR < 0.05). Left column is the total number of
907 significant modules for each tissue, while the second and third columns indicate number shared with
908 larvae (out of 24 larval significantly-enriched modules). The last column indicates the total number of
909 genes in these shared modules.