DYNAMIC VARIATION IN A SOCIALLY EXCHANGED FLUID INDICATES A ROLE IN THE REGULATION OF ANT COLONY DEVELOPMENT AND MATURATION

Hakala SM[1], Meurville M-P[1], Stumpe M[2], LeBoeuf AC[1]✉

[1] Department of Biology, University of Fribourg, Chemin du Musée 10, 1700, Fribourg, Switzerland, [2] Metabolomics and Proteomics Platform, Department of Biology, University of Fribourg, Chemin du Musée 10, 1700, Fribourg, Switzerland ✉email: adrialeboeuf@gmail.com

Abstract

Socially exchanged fluids, like seminal fluid and milk, present a direct and effective means through which an individual can influence conspecifics. As organisms heavily adapted to social life, social insects have evolved a suite of behavioral, morphological and molecular mechanisms to ensure cooperation and inclusive fitness benefits for all group members, just as multicellular organisms have. As a part of their social physiology, some species have developed a social circulatory system, where exogenous food and endogenously produced material, such as hormones, proteins and small molecules, are transferred mouth-to-mouth from the foregut of one individual to another. This fluid transfer – stomodeal trophallaxis – ensures that both food and endogenously produced components are distributed throughout the social network.

To understand how the endogenous materials in trophallactic fluid relate to colony life, we investigated trophallactic fluid of the carpenter ant *Camponotus floridanus*, monitoring this fluid within colonies under different biotic and abiotic conditions and in different individuals within colonies. Using quantitative proteomic analyses of over 100 colony and single-individual trophallactic fluid samples, we established a set of core trophallactic fluid proteins. By combining frequentist, empirical Bayesian and machine-learning classification tools, we identified sets of proteins that are significantly induced in trophallactic fluid under different conditions: proteins that differ between young and mature colonies in the field, proteins that differ between colonies in the field and in the lab, and proteins that differ between the trophallactic fluid of nurses and foragers.

Results reveal that different stages of the colony life cycle utilize classic metabolic processes in different ways, with young colonies prioritizing direct carbohydrate processing, while mature colonies prioritizing transmission of stored resources over the trophallactic network. Further, proteins from pathways that govern the fecundity-longevity trade-off and that have been previously implicated in social insect caste determination are being transferred between individuals within colonies, potentially acting as superorganismal hormones. Nurses in particular show higher abundances in proteins and pathways involved in defense against aging. Thus, we show that the protein composition of ant trophallactic fluid varies with social and developmental conditions both at the colony and at the individual level, suggesting that the trophallactic fluid proteome plays a key role in the social physiology of colonies.

Key words

Social behavior, aging, development, Hymenoptera: Formicidae, individual variation, life history evolution, machine learning, network, population proteomics, reproduction, senescence, social evolution, social physiology, stomodeal trophallaxis, tradeoff, life cycle

Introduction

In the course of social evolution, related organisms have formed cooperative entities such as multicellular organisms or groups of social animals (Hamilton, 1963; Queller, 1992; Szathmáry, 2015). In groups of social animals, collective decisions on movement, reproduction and even development are needed for survival (Couzin, 2009; Miller et al., 2013). Countering this drive toward shared interests and collective good, social conflicts arising from conflicting interests complicate cooperation, but fundamentally also drive the evolution of increasingly coordinated behavior – social systems rely on mechanisms to control their development and cohesion, and to suppress social conflicts (Bourke, 2011, Strassmann et al., 2011; Veit, 2019). Some social groups have taken this coordination to a very high level: social insect societies develop and function as a single unit instead of as competing individuals, as 'superorganisms' paralleling the development of multicellular organisms as a single unit rather than as a set of competing cells (Boomsma & Gawne, 2018; Johnson & Linksvayer, 2010).

In superorganismal societies, reproductive queens and males function as the germline, and workers as the soma. Similarly to different tissues in multicellular organisms, workers can be further specialized and exhibit division of labor across different behavioral and morphological castes (Bourke, 2011). While morphological castes (e.g. workers of different sizes) are determined during development, the behavioral caste of an individual typically changes during its lifetime. At the beginning of their adult life, workers specialize inside the nest as nurses focusing on brood care, and as they age, they switch to foraging outside of the nest (Huang & Robinson, 1996). Colonies also go through life stages. Young social insect colonies have an initial growth phase where they solely produce one type of workers, and only later in their life cycle do they produce more specialized worker castes and finally, males and queens (Wilson, 1971). The switch to reproductive phase is a major life-history transition at the colony level, and tightly connected to female caste determination. In social Hymenoptera, determination of whether a female larva develops into a queen or a worker, and what kind of worker exactly, is controlled by intricate differences of gene expression of the same female genome, guided primarily by environmental factors, in particular nutrition and social cues, but sometimes also influenced by genetics (Anderson et al., 2008; Rajakumar et al., 2018; Schwander et al., 2010; Wheeler, 1986).

Coordinated function of tightly integrated groups such as social insect colonies, and subgroups such as their different castes, has been described as social physiology (Friedman et al., 2020), consisting of various behavioral, morphological and molecular mechanisms that ensure cooperation and inclusive fitness benefits for all group members. As a part of their social physiology, some social insect societies have developed a form of social circulatory system, where nutrition and endogenously produced functional molecules, such as hormones, are transferred mouth-to-mouth from the foregut of one individual to another (LeBoeuf et al., 2016, 2018). This social fluid transfer is called stomodeal trophallaxis (Meurville & LeBoeuf, 2021). It ensures not only that food is distributed to all adults and larvae within the colony, but also that all individuals of the colony are interconnected through shared bodily fluids. Trophallactic fluid of ants and bees typically contains endogenous proteins involved in digestion, immune defense and developmental regulation (LeBoeuf et al., 2016), indicating that this fluid transmits more than food.

Molecular signals are important in controlling the colony life histories and guiding caste determination both at the colony level and at the individual level. Queen pheromones are central signaling molecules acting across individuals (Kocher & Grozinger, 2011; Nijhout & Wheeler, 1982; Pamminger et al., 2016). Juvenile hormone and vitellogenin are central signaling molecules in classical insect development that may also play across-individual roles in some social insects (Harwood et al., 2019; LeBoeuf et al., 2016, 2018; Scharf et al., 2007). Together with fundamental

nutrient-response signaling pathways (insulin, TOR), these molecules establish the developmental trajectories of individuals (Chandra et al., 2018; Libbrecht et al., 2013; Morandin et al., 2014). In solitary organisms, such molecules are produced and function solely within the organism's own body. In contrast, in social Hymenoptera even the molecules traditionally functioning within-individuals can be secreted to the crop and distributed among the society members through trophallaxis and the social circulatory system (LeBoeuf et al., 2016).

Molecular components transmitted through trophallaxis, namely juvenile hormone and juvenile hormone esterase-like proteins, have been shown to influence the development of ant larvae (LeBoeuf et al., 2016, 2018). Thus it is possible that molecules in trophallactic fluid may influence caste determination, similarly to honeybee workers feeding larvae with royal jelly to direct their development toward a queen fate (Buttstedt et al., 2014, 2016; Kamakura, 2011; Kucharski et al., 2015; LeBoeuf et al., 2016). The molecular functions of trophallactic fluid are still largely unstudied, but it is known, for example, that social isolation changes its composition (LeBoeuf et al., 2016), suggesting that this fluid correlates with social environment. Further, we propose that trophallactic fluid could both reflect and affect the social environments of the colony, thus providing important environmental cues for collective decision making. However, it is still largely unknown how much and what kind of qualitative and quantitative variation is present in trophallactic fluid and what causes it.

To understand how endogenous materials exchanged over the social circulatory system relate to colony life, we investigate the variation of the protein content in the trophallactic fluid of the carpenter ant *Camponotus floridanus* at different scales, monitoring this fluid between colonies under different biotic and abiotic conditions and in different individuals within colonies. We hypothesize that the target of some worker-produced trophallactic fluid proteins are developing larvae, and thus expect that colonies at different stages of their life cycle circulate different sets of proteins in the fluid. We hypothesize that young and mature colonies will have trophallactic fluid that differs in proteins that regulate nutritional usage and caste determination. Similarly, we predict that the trophallactic fluid of brood-caring individuals (nurses) is likely to contain proteins targeting larval development while the trophallactic fluid of individuals specialized in foraging (foragers) is expected to have more proteins involved in the first steps of digestion. We also expect the content of trophallactic fluid to reflect differences of the environmental context between field-collected and laboratory-reared colonies. Further, we aim to pinpoint the core trophallactic fluid proteins from the noise inherently present in this fluid that is a mix of endogenous and exogenous material.

Materials and methods

Study species

Camponotus floridanus is a common species of carpenter ant in the south-eastern USA, and has already been the focus of previous trophallactic fluid analyses (LeBoeuf et al., 2016, 2018). They live in dead wood or in man-made structures, often in urban habitats, and forage for honeydew, floral nectar, extra-floral nectar, and arthropod prey. Each colony has a single, singly mated queen (Gadau et al., 1996), and polydomous nest-structures where queenless satellite nests are common. Colonies grow to tens of thousands of workers and produce sexual brood only after multiple years of initial growth. Large established colonies have two morphologically differentiated worker castes, with variably sized small-headed minors focusing on brood care when young and

foraging when old, and big-headed majors that engage in nest defense, foraging and food storage (Deyrup, 2017).

Colony and sample identification

The species was identified based on worker and queen morphology (Deyrup, 2003, 2017; Moreau et al., 2014). In line with previous studies, we use the name *C. floridanus* with the knowledge that the taxonomy and nomenclature of the C. *atriceps* complex (to which it belongs) is not fully resolved (Deyrup, 2017).

We collected full young colonies (0-80 workers) and mature colony extracts (30-200 workers) on several Florida Keys islands (SI Figure 1 and SI Table 1) in winter 2019 and 2020. A colony was deemed "young" if the worker population was <100, primarily minors, and the queen was found (meaning both that the species could be clearly identified and that the nest was not a queenless satellite of an established colony), and "mature" if the opened nest contained many large aggressive majors. Young colonies lack majors (Gibson, 1989) and individuals are generally less aggressive. We only collected mature colony samples when we also found larval brood in the opened nest. In our study area, we observed that young colonies are typically found nesting in different material than are mature colonies. Young colonies are often found under stones or in lumps of clay-like mud associated with crab burrows a short distance from the water, whereas the mature colonies were found nesting in large pieces of damp rotting wood.

Laboratory rearing

Young colonies were brought to the lab and maintained in fluon-coated plastic boxes with a mesh-ventilated lid, at 25 °C with 60 % relative humidity and a 12 h light/dark cycle. Each colony was provided with one or more glass tube for nesting, 10% sugar water, and a Bhatkar & Whitcomb diet (1970) and some *Drosophila melanogaster*. One week prior to proteomic sampling, we substituted the honey-based food with maple syrup-based food to avoid contamination with honeybee proteins (as in LeBoeuf et al 2016).

Trophallactic fluid collection

Field samples of trophallactic fluid were collected within eight hours of ant collection. Of the 20 young colonies and 23 mature colonies, workers collected from two of the mature colonies (L and N) were subdivided into six fragments to assess variation within a single colony (two samples from major workers, two samples from brood-associated workers, and two samples from the remaining minor workers). For all other analyses only one of these for each colony (referred to as minors1) was used to avoid pseudoreplication. In the laboratory, the trophallactic fluid samples underlying the *Field vs. Lab* comparison were sampled after six months in the lab. The four colonies used for the single individual analyses had been in the lab for 18 months at the time of trophallactic fluid collection.

Trophallactic fluid was obtained from CO_2- or cold-anesthetized workers whose abdomens were gently squeezed to force them to regurgitate the contents of their crops (as in LeBoeuf et al., 2016). For each colony, at least 30 individuals were sampled to obtain at least 10 µl of raw trophallactic fluid. For many young colonies only smaller samples were possible, because of the low number of workers (SI Table 1). Young colony samples were only used for further analysis if at least 2.5 µl of trophallactic fluid were collected. For single individual samples, workers with visibly full abdomens were chosen and the obtained sample volumes ranged from 0.7 µl to 2.2 µl. An individual was classified as forager, when it was seen outside the nest tube in the feeding area of an undisturbed laboratory nest box, and a nurse, when it remained in the nest tube even after the

tube was removed from the original laboratory nest and placed into a new one. For colonies from which individual samples were collected, a pooled sample was also taken from less obviously full individuals that remained after individual sampling. Samples were collected with glass capillaries into 5 µl of 1 x Sigmafast Protease Inhibitor Cocktail (Sigma-Aldrich) with 50 mM Tris pH 9 in LoBind eppendorf tubes and were stored -80 C until further analysis. The total proteomics sample number is 73 colony samples of following types: 23 mature colonies with two of them sampled 6 times, 20 young colonies in the field, 16 young colonies in the lab, four laboratory colonies used for single individual sampling; and 40 individual samples: 20 nurses and 20 foragers.

Protein mass spectrometry sample preparation and analysis

Samples were mixed with Laemmli sample buffer and pH was adjusted with 1 M Tris-Cl, pH 7. After reduction with 1 mM DTT for 10 min at 75°C and alkylation using 5.5 mM iodoacetamide for 10 min at room temperature protein samples were separated on 4-12% gradient gels (ExpressPlus, GeneScript). Each gel lane was cut into small pieces, proteins were in-gel digested with trypsin (Promega) and the resulting peptide mixtures were processed on STAGE tips (Rappsilber et al., 2007; Shevchenko et al., 2006).

LC-MS/MS measurements were performed on a QExactive plus mass spectrometer (Thermo Scientific) coupled to an EasyLC 1000 nanoflow-HPLC. HPLC-column tips (fused silica) with 75 µm inner diameter were self-packed with Reprosil-Pur 120 C18-AQ, 1.9 µm (Dr. Maisch GmbH) to a length of 20 cm. A gradient of A (0.1% formic acid in water) and B (0.1% formic acid in 80% acetonitrile in water) with increasing organic proportion was used for peptide separation (loading of sample with 0% B; separation ramp: from 5-30% B within 85 min). The flow rate was 250 nl/min and for sample application 650 nl/min. The mass spectrometer was operated in the data-dependent mode and switched automatically between MS (max. of $1x10^6$ ions) and MS/MS. Each MS scan was followed by a maximum of ten MS/MS scans using normalized collision energy of 25% and a target value of 1000. Parent ions with a charge state form z = 1 and unassigned charge states were excluded from fragmentation. The mass range for MS was m/z = 370-1750. The resolution for MS was set to 70,000 and for MS/MS to 17,500. MS parameters were as follows: spray voltage 2.3 kV; no sheath and auxiliary gas flow; ion-transfer tube temperature 250°C.

The MS raw data files were uploaded into MaxQuant software (Tyanova, Temu, & Cox, 2016), version 1.6.2.10, for peak detection, generation of peak lists of mass error corrected peptides, and for database searches. MaxQuant was set up to search both the UniProt (RRID:SCR_002380, www.uniprot.org) and NCBI (RRID:SCR_003496, www.ncbi.nlm.nih.gov) databases restricted to *C. floridanus* (UniProt, February 2020 version; NCBI RefSeq, version 7.5), along with common contaminants, such as keratins and enzymes used for digestion. Carbamidomethylcysteine was set as fixed modification and protein amino-terminal acetylation and oxidation of methionine were set as variable modifications. Three missed cleavages were allowed, enzyme specificity was trypsin/P, and the MS/MS tolerance was set to 20 ppm. The average mass precision of identified peptides was in general less than 1 ppm after recalibration. Peptide lists were further used by MaxQuant to identify and relatively quantify proteins using the following parameters: peptide and protein false discovery rates, based on a forward-reverse database, were set to 0.01, minimum peptide length was set to 7, minimum number of peptides for identification and quantitation of proteins was set to one which must be unique. The 'match-between-run' option (0.7 min) was used. All proteins labelled as contaminants, reverse or only identified by site were excluded, as well as proteins with scores less than 70. After the filtering the dataset contained 519 proteins. Quantitative analysis was performed using iBAQ values. Intensity-based absolute quantification (iBAQ) is the quotient of sum of all identified peptides and the number of theoretically observable peptides of a protein (Schwanhäusser et al., 2011).

Statistical analyses

Analyses of dataset characteristics were performed in Perseus v1.6.15.0 (Tyanova, Temu, Sinitcyn, et al., 2016), R 3.6.1 (R Development Core Team, 2013) and Matlab R2020b (Figures 1 and 2). From the raw iBAQ data we created the final proteome dataset by filtering out contaminants, reverse proteins, proteins identified by cite, and proteins with scores lower than 70. Hierarchical clustering was performed on Pearson correlation values between samples on median-subtracted raw iBAQ data in Matlab (corrcoef; clustergram using Euclidean distance and knnimpute to cope with missing values). Principal component analysis was run in Perseus on raw iBAQ values, separately on colony and single individual datasets. Differences in protein numbers among the sample types were analyzed with a negative binomial model, using the function nb.glm from the package MASS 7.3-53 (Venables & Ripley, 2002). Proteome variability per sample type, as measured by the coefficient of variation of the iBAQ abundance of each protein, was analysed with a generalized linear model with gamma distribution and log-link with the R-package LME4 (1.1-26) (Bates et al., 2015). The package multcomp 1.4-15 was used for post-hoc testing for both of these models.

The samples were subdivided according to three main comparisons: *Young vs. Mature* colonies from the field, young colonies in the *Field vs. Lab*, and individual *Nurses vs. Foragers* in the lab. In addition, the extent of spatial effects was analysed for the field-collected *Young vs. Mature* dataset by dividing the sampling locations to two areas (*East vs. West*, SI Figure 1). Quantitative proteomic comparisons between sample types were performed independently with three different approaches: 1) Frequentist, 2) Bayesian, and 3) Machine-learning.

Frequentist analysis

Within each dataset only proteins present in over 70% of the samples were analysed. Out of an original 519 proteins, the final datasets for each comparison contained the following number of proteins: *Young vs. Mature*, 172; *Field vs. Lab*, 137; and *Nurse vs. Forager*, 136. All data were log2 transformed and median-centered, and missing data were imputed by random sampling from normal distribution with 2SD downward shift and 0.3 width for each sample. For colony datasets, we used the permutation-based FDR of 0.05, and for the single individual dataset that contained more borderline-significant proteins, we used a more stable Benjamini-Hochberg FDR with a stricter threshold of 0.01. S0 parameter was set to 2 for all analyses. All comparisons were run as two-sample t-tests, with the *Field vs. Lab* as paired.

For the individual dataset, the combined effects of colony identity and behavioral role (*Nurse vs. Forager*) and their interaction, were analysed with two-way ANOVA, with Benjamini-Hochberg FDR corrections, in R with the base R 3.6.1 command 'p.adjust'. Both factors were also analysed separately with multiple- and two-sample t-tests (for colony identity and behavioral role, respectively). To allow comparison to the other statistical methods, only the simple *Nurse vs. Forager* analysis without the interaction was used for combining the lists of significantly different protein abundances. Our balanced sampling guarantees the results of this simpler model are robust enough to find the most descriptive proteins for nurse and forager trophallactic fluid.

Empirical Bayesian analysis

We implemented LIMMA (Linear Models for Microarray Data), an empirical Bayes method for two group comparison, on our raw iBAQ proteomic datasets with the LIMMA-pipeline-proteomics pipeline 3.0.0 (http://doi.org/10.5281/zenodo.4050581) developed for R 4.0.2. Full data was used and no outliers were removed. Data were median-normalized before comparison and all comparisons were run with a log2 (fold change) cutoff of 2.

Random Forest analysis

We used random forest models (sklearn.ensemble.RandomForestClassifier version 0.22.1, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011) to classify samples into one of two groups for each comparison. These analyses were performed in Python 3.7.6 in a Jupyter notebook. For each comparison, ten analyses were performed, each with a different seed. For each seed, the dataset was split into 80% training set and 20% test set, and a model was fit, tested and accuracy computed. If accuracy was below 85%, hyper-parameter tuning was performed with GridSearchCV (sklearn 0.22.1), and the model re-fit. A seed and its corresponding model were not retained for further analysis if accuracy could not be improved above 75%. The typical parameters: max_depth, 3 or 5; max_features, 'auto'; min_samples_leaf, 3; min_samples_split, 8 or 12; n_estimators, 100 or 500. Samples were classified with out-of-box scores (SI Table 4). The average accuracies of classification for comparisons were: *Young vs. Mature*, 87%; *Nurse vs. Forager*, 93%; *Field vs. Lab*, 91%; *East vs. West*, 58%.

To understand which proteins contributed to the classification, we used SHAP (SHapley Additive exPlanations, shap package v0.37.0 for Python 3), a game theory tool that explains the output of machine learning models (Lundberg & Lee, 2017). To analyse the importance of each protein in a given comparison (feature importance), we averaged the absolute value of the Shapley values per protein across the data. Then for each protein, we averaged the feature importances for each of the ten seeded models. Proteins that have no impact on the model classification receive a feature importance value of 0. When ranked according to average feature importance, the data had an approximate Pareto distribution with an inflection point typically at feature importance of ~0.15. Thus, because there is no established cutoff for significance in this form of analysis, we chose to include as 'significant' in further analyses all proteins with a feature importance of > 0.15 (SI Table 4).

For random forest predictions, models trained on the classification between young and mature colonies were used to classify the same young colonies after six months in the laboratory. Out-of-box scores were averaged over five seeded models.

Orthology, gene ontology, and protein network analyses

Because little functional work has been done in ants, we analysed gene ontology terms for the *Drosophila* orthologs to our genes of interest. Orthologs to *C. floridanus* trophallactic fluid proteins were determined with OMA ("Orthologous MAtrix" Jan 2020 release (Altenhoff et al., 2021)). If no ortholog was found within OMA for a given gene, the protein sequence was BLASTed against *Drosophila melanogaster*. In some cases, no ortholog could be found.

GO analysis was performed using both Flybase (Larkin et al., 2021) and STRING v11 (Szklarczyk et al., 2019). STRING was also used for protein-protein interaction and pathway analyses, including KEGG and Reactome (SI Table 5). The protein-protein interaction enrichment analysis in STRING used a hypergeometric test with Benjamini-Hochberg corrected FDR. Only 43 out of the 60 most abundant proteins had sufficient annotation for use by STRING while 52 of the 55 differentially abundant proteins had sufficient annotation.

Results

We sought to determine whether variation in protein presence and abundance in trophallactic fluid correlate with important aspects of colony life. To assess this, we analyzed the trophallactic fluid proteomes of colonies at different stages in the colony life-cycle (*Young vs. Mature*), of colonies when kept in the field or in the lab (*Field vs. Lab*), and between colonies found on different

nearby islands (*East vs. West*). Because trophallactic fluid proteins may be signals transmitted between individuals within a colony, we also analysed protein presence and abundance in trophallactic fluid of single individuals in different colony 'tissues' – in-nest workers taking care of brood and out-of-nest workers (*Nurse vs. Forager*).

Overall proteome variation

Over 73 colony samples and 40 single individual trophallactic fluid samples analysed, a total of 519 proteins were identified (Figure 1). Trophallactic fluid samples contained a set of 27 'core' trophallactic fluid proteins that were present in all samples regardless of life-cycle, life-stage or environmental conditions. We also observed sets of inducible proteins that were routinely observed in different types of samples, and we observed a large number of proteins that were more variably present. The majority of the 519 proteins we observed (298, 57%) were present in less than half of the samples. Even though the most common proteins displayed higher average abundance, across the entire dataset, protein abundance did not correlate with the proportion of samples containing the protein – even proteins present in only a small proportion of the samples in some cases exhibited high abundance (SI Figure 2). The overall protein abundance was higher in colony samples relative to single individual samples, somewhat reflective of the trophallactic fluid volume collected. The number of proteins identified for a given sample correlated with collected trophallactic fluid sample volume (Pearson correlation test $p < 0.05$, $r = 0.24$ for colony samples and $p < 0.05$, $r = -0.40$ for single individual samples).

Field-collected samples exhibited more variable proteomes than did lab-collected samples (Figure 1, gamma GLM posthoc p-values < 0.001, SI Table 2). Further, colonies that had been in the lab for more than one year showed less variable proteomes than did colonies that had been in the lab for only six months (Figure 1, gamma GLM $z=-4.46$, $SE=0.04$, $p <0.001$). The trophallactic proteome variability of young and mature colonies did not differ significantly, nor did nurses' and foragers' (Figure 1). However, foragers had fewer identified proteins in their trophallactic fluid than did nurses (Figure 1, negative binomial $z=3.72$ $SE=0.08$ $P=0.005$), while there is no significant difference among the full colony samples (SI Table 3).

When hierarchically clustered based on correlation across protein abundances, the trophallactic fluid samples tended broadly to align with others of the same type though clusters were not fully distinct (Figure 2). For the colony trophallactic fluid proteomes (Figure 2A and C), mature samples clustered primarily with other mature samples, and young field and laboratory samples grouped together in mixed clusters. In trophallactic fluid collected from single individuals, samples correlated both with other samples of the same colony and with individuals of other colonies having the same behavioral role (nurse or forager; Figure 2B and D).

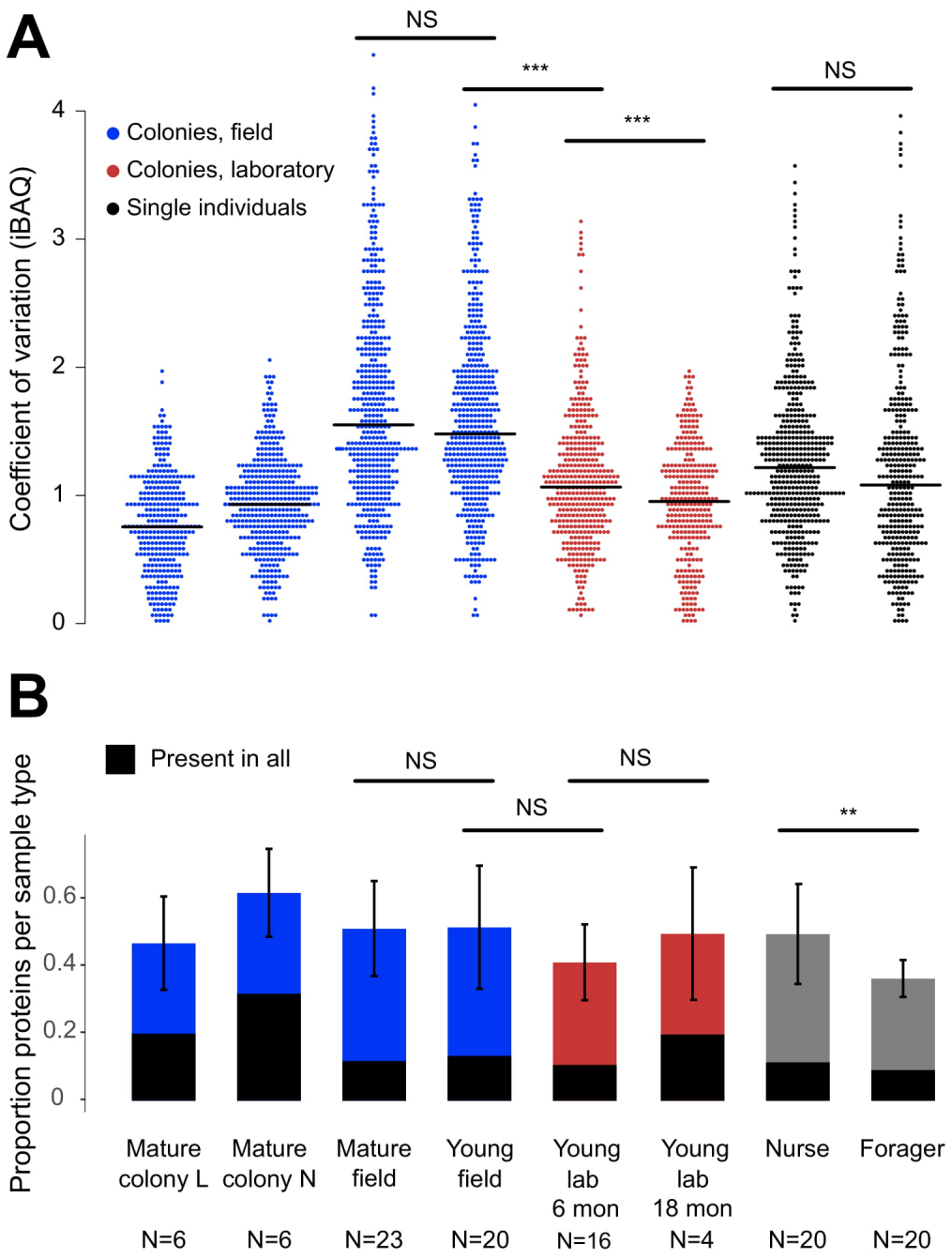

*Figure 1: Protein presence in trophallactic fluid varies with biotic and abiotic factors. A. Coefficient of variation (standard deviation/mean) for the iBAQ values of the 519 proteins identified by sample type. B. Mean ± SD of the proportion of proteins present in samples of a given type. Proportion of proteins present in all samples of a given type are highlighted in black. Sample sizes per type are given under their names. Mature L and Mature N are mature colonies that were sampled six times to assess within-colony variation in colony samples. Significance of comparisons based on gamma GLM (A) or negative binomial GLM (B): NS indicated when p > 0.05 significant, ** p <0.01, *** p < 0.001 (full results in SI Table 2).*

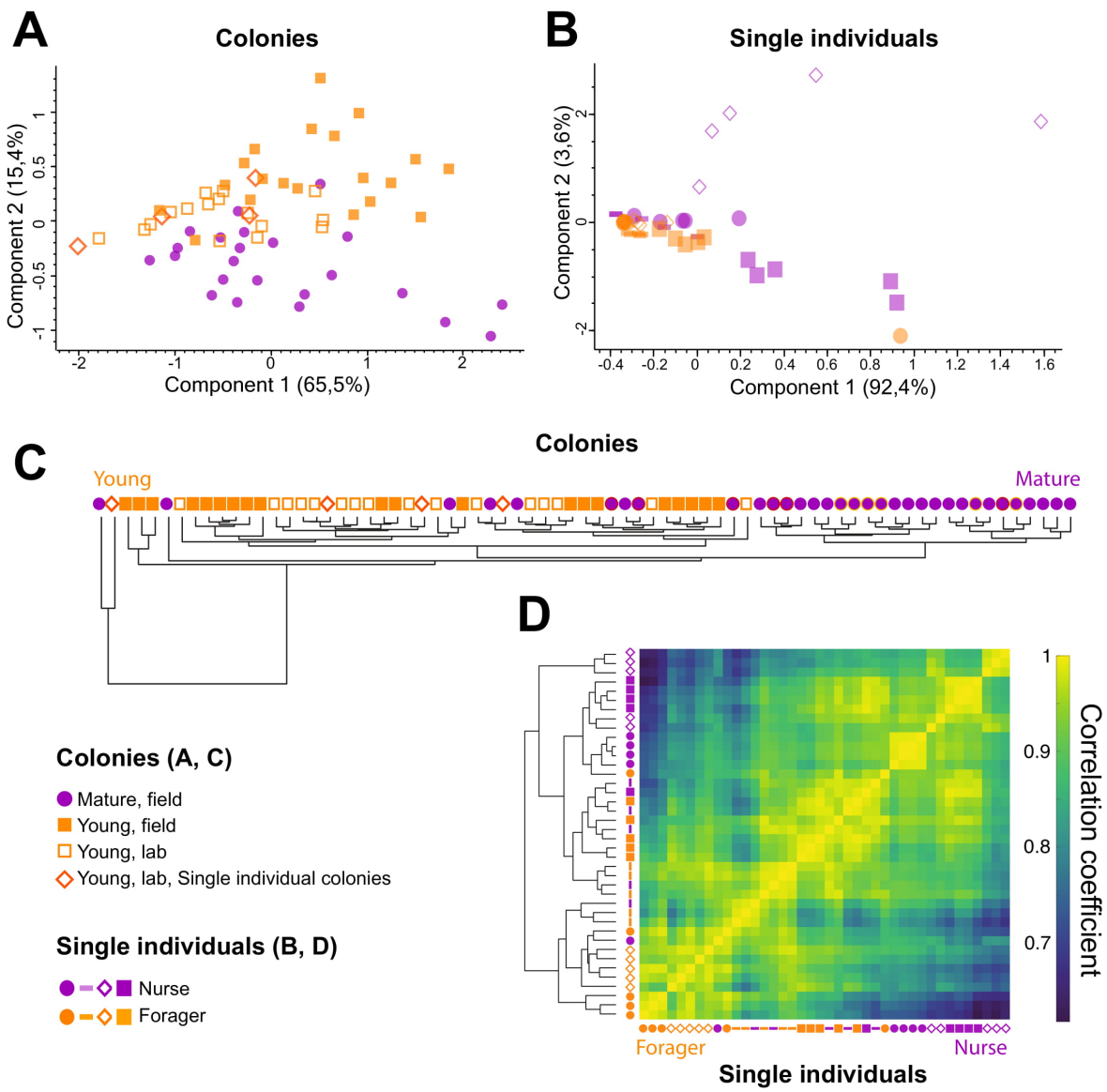

Figure 2: Similarity across trophallactic fluid proteome samples of colonies and single individuals. Principal component analysis for all 519 proteins for (A) colony samples and (B) single individual samples from the four colonies marked with red diamonds in 2A. Hierarchical clustering using Euclidean distance of the colony samples (C) and for single individual samples (D) based on correlation coefficients across all protein abundances. Heatmap of correlation coefficients is shown in (D). Symbols correspond across (A) and (C) for colony samples, and (B) and (D) for single individual samples. In (C) the six replicate samples of the two mature colonies are highlighted with red (L) or orange (N) circles.

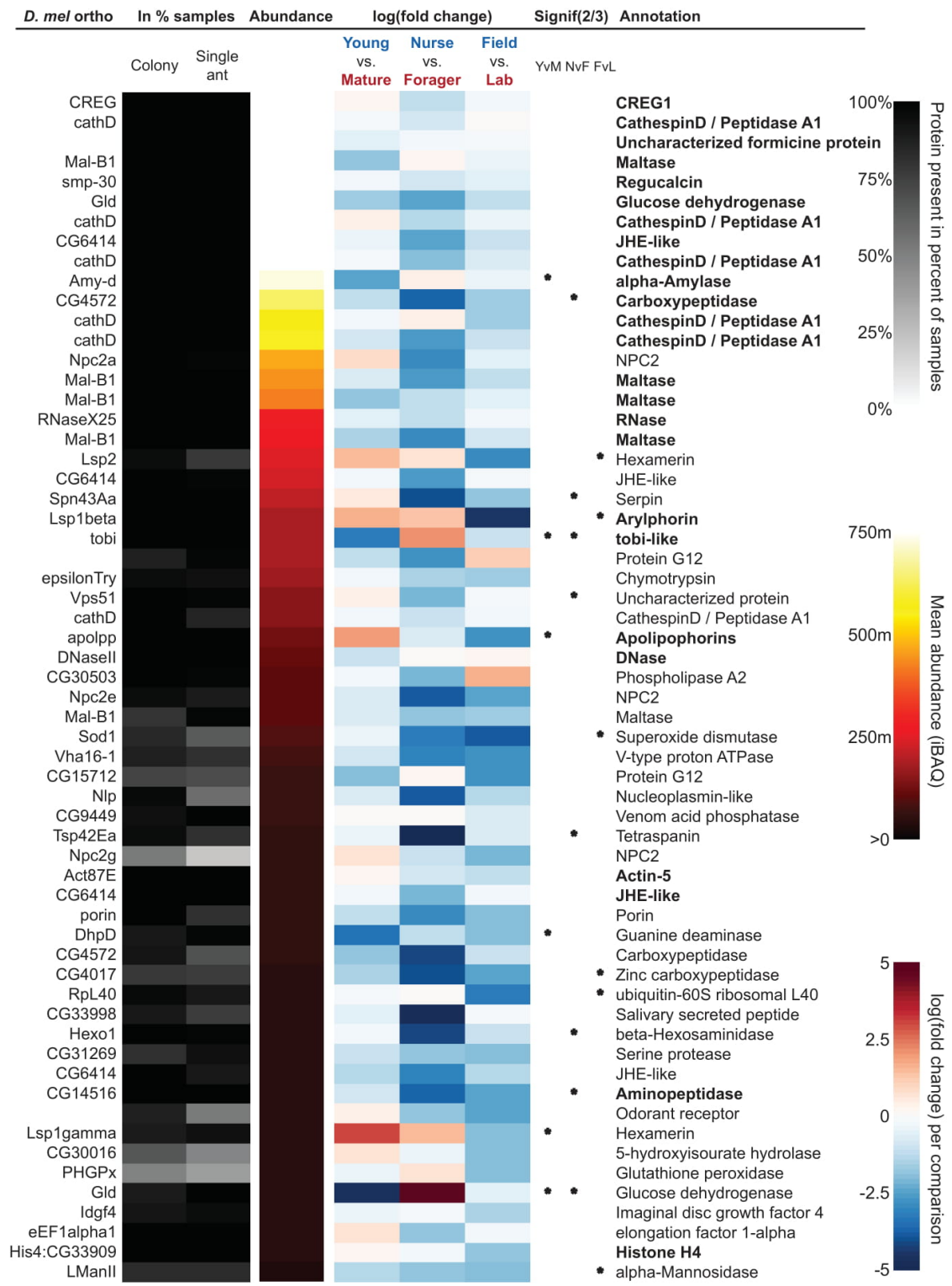

Figure 3: The sixty most abundant proteins in trophallactic fluid over 73 colony and 40 single individual samples. From left to right, Drosophila melanogaster ortholog gene names, proportion of samples in which the protein was identified in colony samples and single individual samples, average iBAQ abundance across all samples, log2 of the fold change in abundance between types for a given comparison, the comparisons for which the protein was significant in two out of three methods, annotation terms (full results in SI Table 4). Annotation terms are in bold for these 25 out of 27 core trophallactic fluid proteins.

Comparisons of trophallactic fluid across conditions

In addition to characterizing the most abundant and core protein of the trophallactic fluid (Figure 3), we wanted to robustly confirm proteins that differ significantly in our comparisons despite the noise inherently present in samples originating from the foregut. To accomplish this, we chose to overlay three distinct statistical approaches: classical frequentist, empirical Bayesian and machine-learning in the form of random forest classification. In our main comparisons, *Young vs. Mature* colonies from the field, young colonies in the *Field vs. Lab*, and individual *Nurses vs. Foragers* in the lab, we found significant differences between groups with all three analysis methods (full results in SI Figure 3 and SI Table 4).

For the *Young vs. Mature* comparisons, there were 10, 17 and 40 differentially abundant proteins according to frequentist t-test, empirical Bayesian LIMMA and the random forest approach, respectively. Similarly, for the *Nurse vs. Forager* comparison there were 21, 103 and 26 differentially abundant proteins, and when young colonies were brought to the laboratory and resampled after six months, 17, 52 or 29 proteins had significantly different abundance. The average accuracies of classification for comparisons with the random forest approach were: *Young vs. Mature*, 87%; *Nurse vs. Forager*, 93%; and *Field vs. Lab*, 91%.

To leverage the unique benefits of the different forms of analysis we focused on proteins significant in two out of the three forms of analysis. Here, young and mature colonies differed by 11 proteins, and nurses and foragers differed by 25 proteins (Figure 4). When young colonies were brought to the laboratory and resampled six months later, the proteomes differed significantly by 23 proteins. We found no clear signature of spatial structure (eastern versus western islands) in the trophallactic fluid proteomes. The frequentist analysis between different sampling areas found no significantly different proteins, and the random forest model did not reach high enough accuracy for this dataset to be informative (58% classification accuracy). Only the empirical Bayesian model found eight proteins that significantly differed in that dataset (SI Table 4).

The single individual comparison between behavioral castes showed clear interaction between the colony and behavioral role of the individuals, with 35 proteins significant only through the interaction (significantly separating the nurses and the foragers but only in some colonies (SI Table 4).

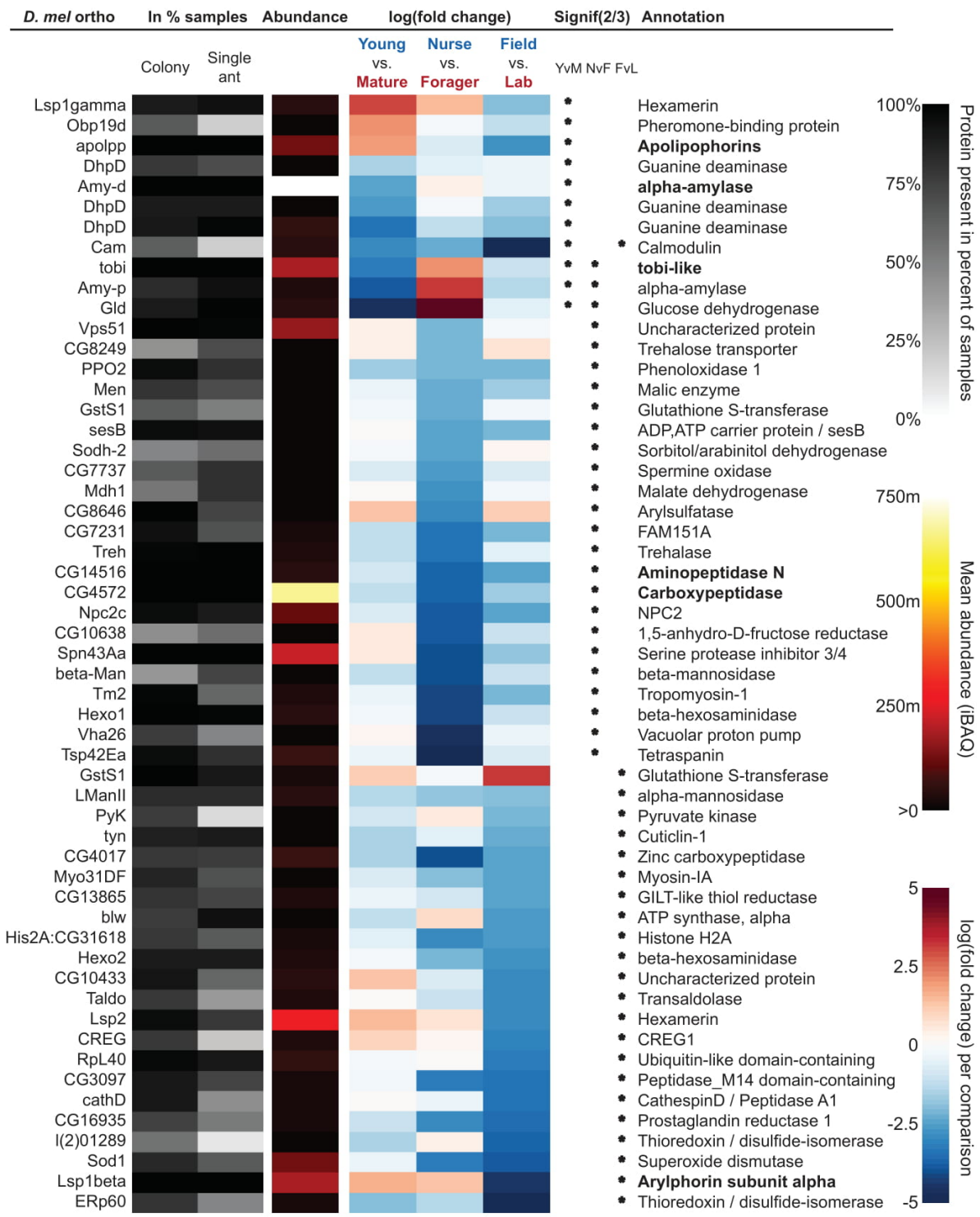

Figure 4: All proteins that significantly differ in two out of three of the analysis methods (frequentist, empirical Bayesian and random forest classification with SHAP values to assess importance). From left to right, Drosophila melanogaster ortholog gene names, proportion of samples in which the protein was identified in colony samples and single individual samples, average iBAQ abundance across all samples, log2 of the fold change in abundance between types for a given comparison, the comparisons for which the protein was significant in two out of three methods, annotation terms (full results in SI Table 4). Annotation terms are in bold for these core trophallactic fluid proteins.

Abundant and significantly differing proteins in trophallactic fluid

To investigate the individual proteins found in trophallactic fluid and to compare the significant proteins of our analyses, we analysed the gene ontology terms and protein-protein interaction (PPI) networks of the trophallactic fluid proteins' *Drosophila melanogaster* orthologs. The sixty most abundant proteins of trophallactic fluid (Figure 3) are predominantly involved in the biological processes of carbohydrate metabolism, lipid and sterol transport (FDR corrected enrichment $p<0.00038$, $p<0.0013$ and 0.0087 respectively). The larval serum protein complex was represented by 3 out of 4 members in both the most abundant proteins and in the significant proteins (hexamerins/arylphorins: Lsp1beta, Lsp1gamma and Lsp2). A strong representation of the innate immune system (Reactome pathway $p < 6.57e-5$) was clear as were lysosomal processes (KEGG pathway, $p < 3.21e-9$). Beyond the sixty most abundant proteins many others are of interest as well; for example, vitellogenin is the 93rd most abundant protein in trophallactic fluid, present in 77% and 88% of colony and single individual samples, respectively.

The PPI network of our differentially abundant protein set (55 proteins, Figure 4) was similar to that of the abundant proteins (Figure 3) but with increased interaction in the networks of the proteins themselves beyond what would be expected by chance (PPI enrichment p-value < 1.22e-15 in differentially abundant proteins relative to p-value < 1.59e-9 in abundant proteins) and with stronger enrichment in carbohydrate metabolic processes ($p < 1.06e-6$) and additional enrichment in oxidation-reduction processes ($p < 0.0032$).

There was a notable overlap of 18 proteins between the most abundant proteins and the proteins significant in two out of three of our statistical methods. To better understand the functions of the proteins important in each comparison, we analysed the GO terms and PPI networks of proteins significant in two out of three statistical methods separately for each comparison. Of our three primary comparisons, the *Nurse vs. Forager* comparison yielded a network of proteins with significantly more interaction than would have been predicted by chance (PPI enrichment p-value < 9.76e-06) as well as a higher degree of PPI enrichment than the other two comparisons (Young v Mature $p<0.0003$ and Field v Lab $p<0.006$). The orthologs of differentially abundant proteins found in the behavioral caste comparison involved not only carbohydrate processing, but also oxidation-reduction and malate metabolic processes (FDR corrected enrichment $p < 0.0042$ and $p < 0.033$, respectively). These pathways have been implicated in the determination of lifespan (Wiley & Campisi, 2016). Indeed, three of the 55 differentially abundant proteins over all comparisons have *D. melanogaster* orthologs with the gene ontology term 'determination of adult lifespan' (Men, sesB, Sod1). The *C. floridanus* tetraspanin is a one-to-many ortholog to the family of Tsp42E genes, one of which has also been implicated in determination of adult lifespan. Twenty-four of the orthologs of proteins significant in two out of three of our analyses are annotated to be expressed in the adult head.

Many of the trophallactic fluid proteins, abundant or significant, were represented in trophallactic fluid by multiple genes from the same protein family, in some cases part of tandem repeats in the genome, indicative of relatively recent evolution. Multiple proteins of the same family were found in the most abundant trophallactic fluid proteins for a family of cathepsinD-like proteins (six in the top 60; (C. Hamilton et al., 2010; LeBoeuf et al., 2016) and a family of Maltase-B1-like proteins (five in the top 60). In the significant protein (Figure 4), we observed fewer members of these families and instead saw three guanine deaminase proteins, all of which were significant in the *Young vs. Mature* comparison. Other families that showed duplications, some of which were significant in one or more analyses were glucose dehydrogenases, CREG1 and tobi-like proteins (target-of-brain-insulin). Three of the most abundant trophallactic fluid proteins had no similarity to *Drosophila* genes, but none of these were significant in more than one analysis for a given comparison.

While trophallactic fluid samples of young and mature colonies were relatively well clustered (Figure 2) and classifiable by random forest algorithms, we wanted to see if our trained classifier could assess the maturity of our young colonies after they had spent six months in the laboratory. Our random forest classifier assigned an average out-of-box maturity score to our 16 laboratory samples of 42% mature, reflecting the intermediate position of the laboratory colony samples in Figure 2.

Discussion

When an ant colony matures, the protein content of trophallactic fluid changes, suggesting that these proteins that circulate amongst individuals play fundamental roles in age-related colony physiology. In ant trophallactic fluid, there are core proteins that are always present but can change in abundance, and inducible proteins whose presence varies based on colony life cycle and environmental context. At the individual level certain proteins also correspond to individuals' behavioral role, in accordance to individual task requirements and age. Certain overarching patterns across our comparisons can be indicative of function. Proteins abundant in mature colonies and nurses are likely factors required to build large workers and sexuals. Proteins abundant in foragers and young colonies are likely to be molecular machinery for fast nutrient uptake. The trophallactic fluid proteomes reveal a rich network of proteins involved in the principal functions of ant colonies and their life cycle.

The among-colony variation in protein abundance in trophallactic fluid declines over time when colonies are brought from the field to the laboratory, showing that the fluid reflects environmental complexity – typical of traits that have evolved to deal with specific environmental cues and stressors (e.g. immunity, Lazzaro & Little, 2009). This is also reflected in laboratory colony development in this species, where lab colonies rarely grow to the size or maturity of field colonies. In our comparisons between young and mature colonies, and between the different behavioral castes of individuals within colonies, the overall level of variation stays the same. However, many specific proteins differ in abundance in these comparisons, reflecting adaptation and the functional role of these proteins in the social physiology of ant colonies.

Considering the protein machinery transmitted over the social circulatory system over three timescales reveals fundamental priorities of social insect colonies: controlling the development and life cycle of the whole colony, controlling how individuals work with resources, and finally how evolution has acted as this social fluid has evolved.

Priorities for resource use as colonies mature

Maximizing resource usage is fundamental to reproduction, aging and life history evolution (Rodrigues & Flatt, 2016; Stearns, 2000; Wiley & Campisi, 2016). Through significantly differing genes involved in these processes between young and mature ant colonies, our results link trophallactic fluid to the control of the colony life cycle. The main axis separating young and mature colonies is the tradeoff between maintenance and reproduction (Gadgil & Bossert, 1970): young colonies focus on rapid growth whereas mature colonies consolidate resources into more resource-expensive individuals, sexuals. Many of the proteins that significantly differ in trophallactic fluid between young and mature colonies correspond to those priorities – resource transmission and consolidation through hexamerins and apolipophorins in mature colonies (Burmester, 2002; Burmester & Scheller, 1999; Thorsten Burmester, 1999), and rapid usage of and response to nutrients through tobi, glucose dehydrogenase, and amylases in young colonies (Buch et al., 2008; White et al., 2021). While the observed difference could reflect different environmental

conditions and resources used by young and mature colonies, these findings suggest that trophallaxis and the social circulatory system are tools to share information and make collective decisions on whether to produce more workers or start producing sexuals, tilting the colony from one life-stage to another.

In connection to the colony life cycle, we found that trophallactic fluid contains numerous factors linked to aging and coping with oxidative stress, including two of the three most well-known antioxidant enzymes: superoxide dismutase and glutathione peroxidase (Monaghan et al., 2009). Considering these and other proteins and pathways that have been implicated in aging, like redox pathways and malate metabolism (Koch et al., 2021; Wiley & Campisi, 2016), nurses show the strongest relative abundance in anti-aging proteins. This links trophallactic fluid to one of the main topics of evolutionary ecology: the longevity-fecundity tradeoff between the metabolic needs for reproduction and coping with consequent oxidative stress (Edward & Chapman, 2011; Flatt, 2011; Monaghan et al., 2009). Social insects seemingly escape this tradeoff with long-lived and highly reproductive queens and short-lived, non-reproductive workers (Edward & Chapman, 2011; Heinze & Schrempf, 2008; Monaghan et al., 2009). Our results show that both metabolism and oxidative stress related factors can be shared over the social circulatory system, which may play a central role in explaining why social insects seem to subvert this tradeoff. If molecules dealing with oxidative stress can be spread into the circulatory system, certain individuals may bear the costs that others in the network incur. This could potential account for some of the puzzling results of previous studies on senescence in social insects (Lucas & Keller, 2020), and provides a new perspective to analyze the regulatory changes of social insect reproductive castes with regard to aging (Corona et al., 2005; Elsner et al., 2018; Gstöttl et al., 2020; Negroni et al., 2019).

While both the altered longevity-fecundity trade-off and caste determination in social insects have long been recognized to implicate juvenile hormone, vitellogenin, and insulin signaling (Corona et al., 2016; Rodrigues & Flatt, 2016; Von Wyschetzki et al., 2015), beyond the somewhat controversial link between MRJP1/royalactin and EGFR (Buttstedt et al., 2016; Kamakura, 2011), no mechanism has been proposed to enact the robust regulation of colony maturation observed across social insect species. Here we show that the exchange of materials over the social circulatory system can provide that mechanism. This socially transmitted fluid contains nutrient storage proteins, hormones, major regulators of metabolism and major players in pathways involved in aging. While most previous work has focused almost exclusively on gene expression, this research shows that for species that engage in trophallaxis, expression studies are necessary but insufficient to understand where in the colony the relevant genes are having their effects. It is conceivable that proteins produced in certain individuals could be sequestered by other individuals, allowing transmission of critical materials across superorganismal tissues (Friedman et al., 2020). In this view, proteins that exert effects in different individuals than the individuals that have produced them can be considered superorganismal hormones.

Modulation and transmission of resources within colony

Ant trophallactic fluid includes proteins dealing with metabolism and nutrient usage and storage, notably transmission of large storage proteins (Lsps, apolpp) and rapid usage of sugar resources (Gld, tobi, Amy, Mal). Many of these metabolism-related proteins are core trophallactic fluid proteins present in high abundance, but many also differ significantly among the colony and individual life stages in our study, showing that optimizing metabolism over levels of organization and life stages can happen through trophallaxis. Similar shifts can be seen in the tissues of multicellular organisms as life-stage priorities change, for example in the midgut of drosophila females after mating, where significant changes in expression are observed in many of the same genes we note

here (White et al., 2021). Some targets of these proteins in trophallactic fluid may be adult physiology and others larval physiology. The *Drosophila* larval hemolymph proteome changes as development unfolds, in particular around metamorphosis where nutrients need to be mobilized and repurposed (Handke et al., 2013). Many of these significantly differing proteins over larval development (Handke et al., 2013) also appear important in our comparisons, suggesting that this regulation may occur over the social network in some ants. Metamorphosis involves substantial autophagy and autophagic flux, repurposing nutrients through lysosomal pathways (Rusten et al., 2004) that are well represented in our abundant and significantly differing protein lists as well.

The protein composition of trophallactic fluid, including many of the above-mentioned metabolism and redox proteins, differs between the nurses and foragers, suggesting that different behavioral castes have different roles at the molecular level as well. We confirm that trophallactic fluid indeed plays a role in the social physiology of the colonies, with different functions distributed among the individuals, similarly to multicellular organisms distributing functions among organs (Friedman et al., 2020; Linksvayer, 2015). Overall, most of the significantly different proteins have higher abundances in nurses, who also have overall more proteins than foragers, likely reflecting their role in larval care and development. An earlier study found that in honey bees, foragers have greater expression of immune response, sugar processing and detoxification genes especially in glands in and around the head, tissues dealing with nectar transfer and trophallaxis (Vannette et al., 2015). Many of the genes differentially expressed in foragers in their study were significantly more abundant in trophallactic fluid of foragers in our study (namely sugar processing enzymes and glutathione-S-transferases), and our analysis of Reactome pathways showed significant enrichment in innate immunity genes.

Evolution of social fluids

Trophallactic fluid is not the only type of social fluid in the animal kingdom - also milk and seminal fluid are similar examples of direct transfers of biological material among individuals. Such socially exchanged materials often contain molecules that target receivers' physiology beyond the fundamental reason for the transfer (Bromfield et al., 2014; Savino et al., 2013). Some of the proteins we find to be important in our comparisons have previously been implicated in these other social transfers. For example, one of our protein hits is orthologous to Drosophila's CG10433, a seminal fluid protein (Findlay et al., 2008) that impacts juvenile hormone associated hatch-rate post-mating (Liu et al., 2014). An interesting aspect of ant societies is that the males are generally not present in them, except as genes functioning in workers. Thus it seems possible that some genes traditionally considered to be only male-expressed (such as those coding for seminal fluid proteins like CG10433, Hexo2, regucalcin, (Findlay et al., 2008)) may have evolved to have new, paternally-imprinted functions in workers, as a way for the male genome to impact the society life (Drewell et al., 2012; Galbraith et al., 2016).

As another interesting parallel to a phylogenetically distant social fluid, trophallactic fluid's most abundant protein is a CREG1, a secreted glycoprotein abundant in mammalian milk (Zhang et al., 2015); another CREG1 protein is significantly more abundant in our field colonies. Additionally, the connection to the previously mentioned gene expression changes in *Drosophila* midgut after mating (White et al., 2021) suggests that genes involved in a physiological response to one social fluid (sperm) may have evolved to be important also in connection to another (trophallactic fluid). Beyond these social fluids, there are also exceptional parallels to venom, a fluid evolved for interspecific interaction (Mariano et al., 2019; Surm & Moran, 2021). These kinds of similarities between social fluids may indicate that certain molecular machinery is required to build functional transmitted fluids. Finding molecular parallels in distinct behavioral processes begs for

further investigation on the role of such exchanges in the evolution of social physiology. If individuals affect each other's behavioral trajectories and development in ways that have fitness relevance, this substantially complicates the pathways of inheritance, and allow social effects to directly influence the evolutionary process in a form of indirect genetic effect (Linksvayer, 2015; McGlothlin et al., 2010; Wolf et al., 1998).

While many of the proteins in trophallactic fluid are secreted extracellular proteins whose orthologs are also clearly secreted extracellular proteins (e.g. maltases, amylases, hexamerins, superoxide dismutase), many others are classically considered to function within cells either in cytosol, in membranes or specific organelles (e.g. cathepsins, pyruvate kinase, ATP synthase, porin, tetraspanin). Many of these not-classically-secreted proteins are critical players in the secretion machinery, especially in high secretion tissues (Martínez et al., 2020). In some cases these proteins have been shown to be secreted but in specific contexts; for example pyruvate kinase, when secreted, has been implicated in metastatic cancers (Wang et al., 2020). Consistent with roles in secretion we also find markers of exosomes, extracellular vesicles typically present in social fluids (e.g. tetraspanin, (de la Torre Gomez et al., 2018).

Lysosomal pathways are enriched in our set of significantly varying trophallactic fluid proteins according to the KEGG analysis. Lysosomes are clearly acidic and can be major players in secretion, autophagic flux and exocytosis (Csizmadia et al., 2018; Maruzs et al., 2019; Tancini et al., 2020). The lysosomal signatures we see in trophallactic fluid may be traces of the mechanism of secretion, or perhaps be evidence of how this fluid has evolved. As trophallactic fluid has become acidified in formicine ants (Tragust et al., 2020), lysosomal genes may have been duplicated and neofunctionalized to a new role in this acidic fluid. This would be a similar repurposing to what was observed with the multiple juvenile-hormone-esterase-like proteins in trophallactic fluid (LeBoeuf et al., 2018). The fact that many abundant trophallactic fluid proteins represent clusters of related proteins from a few families (cathepsins, guanine deaminases, maltases) suggest there has been recent evolution in the proteins arriving in this fluid. This hypothesis can be tested by analyzing trophallactic fluid over the ant phylogeny, between species who engage in trophallaxis and those that don't, from different subfamilies including those without an acidified gut. It is an exciting future goal to assess how classical metabolism-related proteins have evolved to function in their new role in inter-individual transmission in trophallactic fluid.

A large proportion of the proteins in trophallactic fluid seem to be inducible, meaning that their expression and abundance is triggered by certain environmental, social or developmental contexts. The current study focuses on some of these contexts, but there are still many more that lie outside the scope of this work. For example, based on our limited data on the four colonies we sampled, individuals of different colonies harbor significantly different sets of proteins in their trophallactic fluid. Ant colonies typically have colony signatures in many traits, and when based on genetic differences, these can extend to larger scale spatial patterns due to the strong population genetic structures of ants (Hakala et al., 2019). However, this is not the case here: the trophallactic fluid proteomes do not differentiate between sampling locations although they do differ among colonies, indicating that colony signatures are not genetic, but rather plastic and functional. Previous screens on population proteomics in other study systems have found differences among populations in similar spatial scales (Gonzalez et al., 2010; Karr, 2008; Mathé-Hubert et al., 2019). As our individual-level samples were collected from laboratory colonies and thus lack many of the environmental reasons for differentiation (Madeira et al., 2020), we suggest the most likely explanations are health status and slight differences in the social context (for example brood composition). Pinpointing all contexts that induce changes in trophallactic fluid, along with the exact functions of the proteins, are important subjects for future work.

Conclusions

We show that the protein composition of ant trophallactic fluid varies across different external contexts and internal conditions both at the colony and at the individual level, suggesting that the changing trophallactic fluid proteome has key functions in social physiology and life cycle of colonies. We have laid the groundwork for future studies on the functions of these proteins in controlling the colony life cycle, senescence and behavior.

Acknowledgements

We thank Dr. Joanne Reiter for help with field work. The work was funded by a Swiss National Science Foundation grant PR00P3_179776 to ACL, and a Swiss Government Excellence grant to SMH.

References

Altenhoff, A. M., Train, C. M., Gilbert, K. J., Mediratta, I., Mendes de Farias, T., Moi, D., Nevers, Y., Radoykova, H. S., Rossier, V., Warwick Vesztrocy, A., Glover, N. M., & Dessimoz, C. (2021). OMA orthology in 2021: website overhaul, conserved isoforms, ancestral gene order and more. *Nucleic Acids Research*, *49*(D1), D373–D379.

Anderson, K. E., Linksvayer, T. A, & Smith, C. R. (2008). The causes and consequences of genetic caste determination in ants (Hymenoptera : Formicidae). *Ecology*, *11*(August), 119–132.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bhatkar, A., & Whitcomb, W. H. (1970). Artificial diet for rearing various species of ants. *The Florida Entomologist*, *53*(4), 229.

Boomsma, J. J., & Gawne, R. (2018). Superorganismality and caste differentiation as points of no return: how the major evolutionary transitions were lost in translation. *Biological Reviews*, *93*(1), 28–54.

Bourke, A. F. (2011). *Principles of social evolution*. Oxford University Press.

Bromfield, J. J., Schjenken, J. E., Chin, P. Y., Care, A. S., Jasper, M. J., & Robertson, S. A. (2014). Maternal tract factors contribute to paternal seminal fluid impact on metabolic phenotype in offspring. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2200–2205.

Buch, S., Melcher, C., Bauer, M., Katzenberger, J., & Pankratz, M. J. (2008). Opposing effects of dietary protein and sugar regulate a transcriptional target of drosophila insulin-like peptide signaling. *Cell Metabolism*, *7*(4), 321–332.

Burmester, T. (2002). Origin and evolution of arthropod hemocyanins and related proteins. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, *172*(2), 95–107.

Burmester, T., & Scheller, K. (1999). Ligands and receptors: Common theme in insect storage protein transport. *Naturwissenschaften*, *86*(10), 468–474.

Burmester, Thorsten. (1999). Evolution and function of the insect hexamerins. *European Journal of Entomology, 96(3)* 213–225.

Buttstedt, A., Ihling, C. H., Pietzsch, M., & Moritz, R. F. A. (2016). Royalactin is not a royal making of a queen. *Nature*, *537*(7621), E10–E12.

Buttstedt, A., Moritz, R. F. A., & Erler, S. (2014). Origin and function of the major royal jelly proteins of the honeybee (*Apis mellifera*) as members of the yellow gene family. *Biological Reviews*, *89*(2), 255–269.

Chandra, V., Fetter-Pruneda, I., Oxley, P. R., Ritger, A. L., McKenzie, S. K., Libbrecht, R., & Kronauer, D. J. C. (2018). Social regulation of insulin signaling and the evolution of

eusociality in ants. *Science*, *361*(6400), 398–402.
Corona, M., Hughes, K. A., Weaver, D. B., & Robinson, G. E. (2005). Gene expression patterns associated with queen honey bee longevity. *Mechanisms of Ageing and Development*, *126*(11), 1230–1238.
Corona, M., Libbrecht, R., & Wheeler, D. E. (2016). Molecular mechanisms of phenotypic plasticity in social insects. *Current Opinion in Insect Science*, *13*, 55–60.
Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, *13*(1), 36–43.
Csizmadia, T., Lorincz, P., Hegedus, K., Széplaki, S., Low, P., & Juhász, G. (2018). Molecular mechanisms of developmentally programmed crinophagy in *Drosophila*. *Journal of Cell Biology*, *217*(1), 361–374.
de la Torre Gomez, C., Goreham, R. V., Bech Serra, J. J., Nann, T., & Kussmann, M. (2018). "Exosomics" - A review of biophysics, biology and biochemistry of exosomes with a focus on human breast milk. *Frontiers in Genetics*, *9*(MAR), 1–11.
Deyrup, M. (2003). An updated list of florida ants (Hymenoptera: Formicidae). *The Florida Entomologist*, *86*(1), 43–48.
Deyrup, M. (2017). *Ants of Florida - Identification and natural history*.Taylor & Francis Group.
Drewell, R. a, Lo, N., Oxley, P. R., & Oldroyd, B. P. (2012). Kin conflict in insect societies: a new epigenetic perspective. *Trends in Ecology & Evolution*, *27*(7), 367–373.
Edward, D., & Chapman, T. (2011). Mechanisms underlying reproductive trade-offs: costs of reproduction. In F. T & H. A (Eds.), *Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs* (pp. 137–152). Oxford University Press.
Elsner, D., Meusemann, K., & Korb, J. (2018). Longevity and transposon defense, the case of termite reproductives. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(21), 5504–5509.
Findlay, G. D., Yi, X., MacCoss, M. J., & Swanson, W. J. (2008). Proteomics reveals novel *Drosophila* seminal fluid proteins transferred at mating. *PLoS Biology*, *6*(7), 1417–1426.
Flatt, T. (2011). Survival costs of reproduction in *Drosophila*. *Experimental Gerontology*, *46*(5), 369–375.
Friedman, D. A., Johnson, B. R., & Linksvayer, T. A. (2020). Distributed physiology and the molecular basis of social life in eusocial insects. *Hormones and Behavior*, *122*(March), 104757.
Gadau, J., Heinze, J., Hölldobler, B., & Schmid, M. (1996). Population and colony structure of the carpenter ant *Camponotus floridanus*. *Molecular Ecology*, *5*, 785–792.
Galbraith, D. A., Yi, S. V., & Grozinger, C. M. (2016). Evaluation of possible proximate mechanisms underlying the kinship theory of intragenomic conflict in social insects. *Integrative and Comparative Biology*, *56*(6), 1206–1214.
Gibson, R. L. (1989). Soldier production in *Camponotus novaeboracensis* during colony growth. *Insectes Sociaux*, *36*(1), 28–41.
Gonzalez, E. G., Krey, G., Espiñeira, M., Diez, A., Puyet, A., & Bautista, J. M. (2010). Population proteomics of the European hake (*Merluccius merluccius*). *Journal of Proteome Research*, *9*(12), 6392–6404.
Gstöttl, C., Stoldt, M., Jongepier, E., Bornberg-Bauer, E., Feldmeyer, B., Heinze, J., & Foitzik, S. (2020). Comparative analyses of caste, sex, and developmental stage-specific transcriptomes in two Temnothorax ants. *Ecology and Evolution*, *10*(10), 4193–4203.
Hamilton, C., Lejeune, B. T., & Rosengaus, R. B. (2010). Trophallaxis and prophylaxis: Social immunity in the carpenter ant *Camponotus pennsylvanicus*. *Biology Letters*, *7*(1), 89–92.
Hamilton, W. D. (1963). The evolution of altruistic behavior. *The American Naturalist*, *97*(896), 354–356.
Handke, B., Poernbacher, I., Goetze, S., Ahrens, C. H., Omasits, U., Marty, F., Simigdala, N., Meyer, I., Wollscheid, B., Brunner, E., Hafen, E., & Lehner, C. F. (2013). The hemolymph proteome of fed and starved *drosophila* larvae. *PLoS ONE*, *8*(6), 1–10.
Harwood, G., Amdam, G., & Freitak, D. (2019). The role of Vitellogenin in the transfer of immune elicitors from gut to hypopharyngeal glands in honey bees (*Apis mellifera*). *Journal of Insect Physiology*, *112*(June 2018), 90–100.

Heinze, J., & Schrempf, A. (2008). Aging and reproduction in social insects - A mini-review. *Gerontology*, *54*(3), 160–167.
Huang, Z. Y., & Robinson, G. E. (1996). Regulation of honey bee division of labor by colony age demography. *Behavioral Ecology and Sociobiology*, *39*(3), 147–158.
Johnson, B. R., & Linksvayer, T. A. (2010). Deconstructing the superorganism: social physiology, groundplans, and sociogenomics. *The Quarterly Review of Biology*, *85*(1), 57–79.
Kamakura, M. (2011). Royalactin induces queen differentiation in honeybees. *Nature*, *473*(7348), 478–483.
Karr, T. L. (2008). Application of proteomics to ecology and population biology. *Heredity*, *100*(2), 200–206.
Koch, R., Buchanan, K., Casagrande, S., Crino, O., Dowling, D., Hill, G., Hood, W., McKenzie, M., Mariette, M., Noble, D., Pavlova, A., Seebacher, F., Sunnucks, P., Udino, E., White, C., Salin, K., & Stier, A. (2021). Integrating Mitochondrial Aerobic Metabolism into Ecology and Evolution. *Trends in Ecology & Evolution*, 1–12.
Kocher, S. D., & Grozinger, C. M. (2011). Cooperation, conflict, and the evolution of queen pheromones. *Journal of Chemical Ecology*, *37*(11), 1263–1275.
Kucharski, R., Foret, S., & Maleszka, R. (2015). EGFR gene methylation is not involved in Royalactin controlled phenotypic polymorphism in honey bees. *Scientific Reports*, *5*(August), 11–16.
Larkin, A., Marygold, S. J., Antonazzo, G., Attrill, H., Dos Santos, G., Garapati, P. V., Goodman, J. L., Gramates, L. S., Millburn, G., Strelets, V. B., Tabone, C. J., & Thurmond, J. (2021). FlyBase: updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Research*, *49*(D1), D899–D907.
Lazzaro, B. P., & Little, T. J. (2009). Immunity in a variable world. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1513), 15–26.
LeBoeuf, A. C., Cohanim, A. B., Stoffel, C., Brent, C. S., Waridel, P., Privman, E., Keller, L., & Benton, R. (2018). Molecular evolution of juvenile hormone esterase-like proteins in a socially exchanged fluid. *Scientific Reports*, *8*(1), 1–10.
LeBoeuf, A. C., Waridel, P., Brent, C. S., Gonçalves, A. N., Menin, L., Ortiz, D., Riba-Grognuz, O., Koto, A., Soares, Z. G., Privman, E., Miska, E. A., Benton, R., & Keller, L. (2016). Oral transfer of chemical cues, growth proteins and hormones in social insects. *ELife*, *5*(NOVEMBER2016), 1–27.
Libbrecht, R., Corona, M., Wende, F., Azevedo, D. O., Serrão, J. E., & Keller, L. (2013). Interplay between insulin signaling, juvenile hormone, and vitellogenin regulates maternal effects on polyphenism in ants. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 11050–11055.
Linksvayer, T. A. (2015). The molecular and evolutionary genetic implications of being truly social for the social insects. *Advances in Insect Physiology*, *48*, 271–292.
Liu, C., Wang, J. L., Zheng, Y., Xiong, E. J., Li, J. J., Yuan, L. L., Yu, X. Q., & Wang, Y. F. (2014). Wolbachia-induced paternal defect in Drosophila is likely by interaction with the juvenile hormone pathway. *Insect Biochemistry and Molecular Biology*, *49*(1), 49–58.
Lucas, E. R., & Keller, L. (2020). The co-evolution of longevity and social life. *Functional Ecology*, *34*(1), 76–87.
Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *ArXiv*, *Section 2*, 1–10.
Madeira, D., Araújo, J. E., Madeira, C., Mendonça, V., Vitorino, R., Vinagre, C., & Diniz, M. S. (2020). Seasonal proteome variation in intertidal shrimps under a natural setting: Connecting molecular networks with environmental fluctuations. *Science of the Total Environment*, *703*, 134957.
Mariano, D. O. C., de Oliveira, Ú. C., Zaharenko, A. J., Pimenta, D. C., Rádis-Baptista, G., & Prieto-Da-Silva, Á. R. de B. (2019). Bottom-up proteomic analysis of polypeptide venom components of the giant ant *Dinoponera quadriceps*. *Toxins*, *11*(8).
Martínez, J., Marmisolle, I., Tarallo, D., & Quijano, C. (2020). Mitochondrial bioenergetics and dynamics in secretion processes. *Frontiers in Endocrinology*, *11*(May), 1–18.

Maruzs, T., Simon-Vecsei, Z., Kiss, V., Csizmadia, T., & Juhász, G. (2019). On the fly: Recent progress on autophagy and aging in *Drosophila*. *Frontiers in Cell and Developmental Biology*, *7*(JULY), 1–15.

Mathé-Hubert, H., Kremmer, L., Colinet, D., Gatti, J.-L., Van Baaren, J., Delava, É., & Poirié, M. (2019). Variation in the venom of parasitic wasps, drift, or selection? Insights from a multivariate QST analysis. *Frontiers in Ecology and Evolution*, *7*(May).

McGlothlin, J. W., Moore, A. J., Wolf, J. B., & Brodie, E. D. (2010). Interacting phenotypes and the evolutionary process. III. Social evolution. *Evolution*, *64*(9), 2558–2574.

Meurville, M.-P., & LeBoeuf, A. (2021). Trophallaxis: the functions and evolution of social fluid exchange in ant colonies (Hymenoptera: Formicidae). *Myrmecological News*, *10*(September), 4136.

Miller, N., Garnier, S., Hartnett, A. T., & Couzin, I. D. (2013). Both information and social cohesion determine collective decisions in animal groups. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(13), 5263–5268.

Monaghan, P., Metcalfe, N. B., & Torres, R. (2009). Oxidative stress as a mediator of life history trade-offs: Mechanisms, measurements and interpretation. *Ecology Letters*, *12*(1), 75–92.

Morandin, C., Havukainen, H., Kulmuni, J., Dhaygude, K., Trontti, K., & Helanterä, H. (2014). Not only for egg yolk—functional and evolutionary insights from expression, selection, and structural analyses of *Formica* ant vitellogenins. *Molecular Biology and Evolution*, *31*(8), 2181–2193.

Moreau, C. S., Deyrup, M. A., Davis, L. R., & Oliveira, P. (2014). Ants of the Florida Keys: Species accounts, biogeography, and conservation (Hymenoptera: Formicidae). *Journal of Insect Science*, *14*(1), 1–8.

Negroni, M. A., Foitzik, S., & Feldmeyer, B. (2019). Long-lived *Temnothorax* ant queens switch from investment in immunity to antioxidant production with age. *Scientific Reports*, *9*(1), 1–10.

Nijhout, H. F., & Wheeler, D. E. (1982). Juvenile hormone and the physiological basis of insect polymorphisms. *The Quarterly Review of Biology*, *57*(2), 109–133.

Pamminger, T., Buttstedt, A., Norman, V., Schierhorn, A., Botías, C., Jones, J. C., Basley, K., & Hughes, W. O. H. (2016). The effects of juvenile hormone on Lasius niger reproduction. *Journal of Insect Physiology*, *95*, 1–7.

Queller, D. C. (1992). A general model for kin selection. *Evolution*, *46*(2), 376.

R Development Core Team. (2013). A language and environment for statistical computing. In *Vienna Austria R Foundation for Statistical Computing* (Vol. 1, Issue 09/18/2009, p. ISBN 3-900051-07-0). R Foundation for Statistical Computing.

Rajakumar, R., Koch, S., Couture, M., Favé, M. J., Lillico-Ouachour, A., Chen, T., De Blasis, G., Rajakumar, A., Ouellette, D., & Abouheif, E. (2018). Social regulation of a rudimentary organ generates complex worker-caste systems in ants. *Nature*, *562*(7728), 574–577.

Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, *2*(8), 1896–1906.

Rodrigues, M. A., & Flatt, T. (2016). Endocrine uncoupling of the trade-off between reproduction and somatic maintenance in eusocial insects. *Current Opinion in Insect Science*, *16*(1995), 1–8.

Rusten, T. E., Lindmo, K., Juhász, G., Sass, M., Seglen, P. O., Brech, A., & Stenmark, H. (2004). Programmed autophagy in the *Drosophila* fat body is induced by ecdysone through regulation of the PI3K Pathway. *Developmental Cell*, *7*(2), 179–192.

Savino, F., Benetti, S., Liguori, S. A., Sorrenti, M., & Cordero Di Montezemolo, L. (2013). Advances on human milk hormones and protection against obesity. *Cellular and Molecular Biology*, *59*(1), 89–98.

Scharf, M. E., Buckspan, C. E., Grzymala, T. L., & Zhou, X. (2007). Regulation of polyphenic caste differentiation in the termite *Reticulitermes flavipes* by interaction of intrinsic and extrinsic factors. *Journal of Experimental Biology*, *210*(24), 4390–4398.

Schwander, T., Lo, N., Beekman, M., Oldroyd, B. P., & Keller, L. (2010). Nature versus nurture

in social insect caste differentiation. *Trends in Ecology & Evolution*, *25*(5), 275–282.
Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*(7347), 337–342.
Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V., & Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, *1*(6), 2856–2860.
Stearns, S. C. (2000). Life history evolution: Successes, limitations, and prospects. *Naturwissenschaften*, *87*(11), 476–486.
Strassmann, J. E., Queller, D. C., Avise, J. C., & Ayala, F. J. (2011). In the light of evolution V: Cooperation and conflict. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(SUPPL. 2), 10787–10791.
Surm, J. M., & Moran, Y. (2021). Insights into how development and life-history dynamics shape the evolution of venom. *EvoDevo*, *12*(1), 1–18.
Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), 10104–10111.
Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Von Mering, C. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613.
Tancini, B., Buratta, S., Delo, F., Sagini, K., Chiaradia, E., Pellegrino, R. M., Emiliani, C., & Urbanelli, L. (2020). Lysosomal exocytosis: The extracellular role of an intracellular organelle. *Membranes*, *10*(12), 1–21.
Tragust, S., Herrmann, C., Häfner, J., Braasch, R., Tilgen, C., Hoock, M., Milidakis, M. A., Gross, R., & Feldhaar, H. (2020). Formicine ants swallow their highly acidic poison for gut microbial selection and control. *ELife*, *9*, 1–25.
Tyanova, S., Temu, T., & Cox, J. (2016). The MaxQuant computational platform for mass spectrometry – based shotgun proteomics. *Nature Protocols*, *11*(12), 2301–2319.
Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., & Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, *13*(9), 731–740.
Vannette, R. L., Mohamed, A., & Johnson, B. R. (2015). Forager bees (*Apis mellifera*) highly express immune and detoxification genes in tissues associated with nectar processing. *Scientific Reports*, *5*(November), 1–9.
Veit, W. (2019). Evolution of multicellularity: cheating done right. *Biology and Philosophy*, *34*(3), 1–22.
Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S. Fourth Edition*. Springer.
Von Wyschetzki, K., Rueppell, O., Oettler, J., & Heinze, J. (2015). Transcriptomic signatures mirror the lack of the fecundity/longevity trade-off in ant queens. *Molecular Biology and Evolution*, *32*(12), 3173–3185.
Wang, C., Zhang, S., Liu, J., Tian, Y., Ma, B., Xu, S., Fu, Y., & Luo, Y. (2020). Secreted pyruvate kinase M2 promotes lung cancer metastasis through activating the integrin Beta1/FAK signaling pathway. *Cell Reports*, *30*(6), 1780-1797.e6.
Wheeler, D. E. (1986). Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. *American Naturalist*, *128*(1), 13–34.
White, M. A., Bonfini, A., Wolfner, M. F., & Buchon, N. (2021). *Drosophila melanogaster* sex peptide regulates mated female midgut morphology and physiology. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(1).
Wiley, C. D., & Campisi, J. (2016). From ancient pathways to aging cells - connecting metabolism and cellular senescence. *Cell Metabolism*, *23*(6), 1013–1021.
Wilson, Edward, O. (1971). *The insect societies*. Harward University Press.
Wolf, J. B., Brodie, E. D., Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998). Evolutionary consequences of indirect genetic effects. *Trends in Ecology and Evolution*, *13*(2), 64–69.

Zhang, L., Boeren, S., Hageman, J. A., Van Hooijdonk, T., Vervoort, J., & Hettinga, K. (2015). Bovine milk proteome in the first 9 days: Protein interactions in maturation of the immune and digestive system of the newborn. *PLoS ONE*, *10*(2), 1–19.