AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems®

RESEARCH ARTICLE

Interactions of the Intracellular Bacterium *Cardinium* with Its Host, the House Dust Mite *Dermatophagoides farinae*, Based on Gene Expression Data

Jan Hubert,[a,b] Marta Nesvorna,[b] Pavel B. Klimov,[c,d] Tomas Erban,[b] Bruno Sopko,[b] Scot E. Dowd,[e] Erin D. Scully[f]

[a]Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czechia
[b]Crop Research Institute, Prague, Czechia
[c]School of Natural Sciences, Bangor University, Bangor, United Kingdom
[d]Faculty of Biology, Tyumen State University, Tyumen, Russia
[e]MR DNA (Molecular Research LP), Shallowater, Texas, USA
[f]USDA-ARS Center for Grain and Animal Health Research, Stored Product Insect and Engineering Research Unit, Manhattan, Kansas, USA

ABSTRACT *Dermatophagoides farinae* is inhabited by an intracellular bacterium, *Cardinium*. Using correlations between host and symbiont gene expression profiles, we identified several important molecular pathways that potentially regulate/facilitate their interactions. The expression of *Cardinium* genes collectively explained 95% of the variation in the expression of mite genes assigned to pathways for phagocytosis, apoptosis, the MAPK signaling cascade, endocytosis, the tumor necrosis factor (TNF) pathway, the transforming growth factor beta (TGF-β) pathway, lysozyme, and the Toll/Imd pathway. In addition, expression of mite genes explained 76% of the variability in *Cardinium* gene expression. In particular, the expression of the *Cardinium* genes encoding the signaling molecules BamD, LepA, SymE, and VirD4 was either positively or negatively correlated with the expression levels of mite genes involved in endocytosis, phagocytosis, and apoptosis. We also found that *Cardinium* possesses a complete biosynthetic pathway for lipoic acid and may provide lipoate, but not biotin, to mites. *Cardinium* gene expression collectively explained 84% of the variation in expression related to several core mite metabolic pathways, and, most notably, a negative correlation was observed between bacterial gene expression and expression of mite genes assigned to the glycolysis and citric acid cycle pathways. Furthermore, we showed that *Cardinium* gene expression is correlated with expression levels of genes associated with terpenoid backbone biosynthesis. This pathway is important for the synthesis of pheromones, thus providing an opportunity for *Cardinium* to influence mite reproductive behavior to facilitate transmission of the bacterium. Overall, our study provided correlational gene expression data that can be useful for future research on mite-*Cardinium* interactions.

IMPORTANCE The molecular mechanisms of mite-symbiont interactions and their impacts on human health are largely unknown. Astigmatid mites, such as house dust and stored-product mites, are among the most significant allergen sources worldwide. Although mites themselves are the main allergen sources, recent studies have indicated that mite-associated microbiomes may have implications for allergen production and human health. The major medically important house dust mite, *D. farinae*, is known to harbor a highly abundant intracellular bacterium belonging to the genus *Cardinium*. Expression analysis of the mite and symbiont genes can identify key mite molecular pathways that facilitate interactions with this endosymbiont and possibly shed light on how this bacterium affects mite allergen production and physiology in general.

KEYWORDS transcriptome, *Cardinium*, symbiont, house dust mite, interactions, allergens, allergens, endosymbiont, host-pathogen interactions

Citation Hubert J, Nesvorna M, Klimov PB, Erban T, Sopko B, Dowd SE, Scully ED. 2021. Interactions of the intracellular bacterium *Cardinium* with its host, the house dust mite *Dermatophagoides farinae*, based on gene expression data. mSystems 6:e00916-21. https://doi.org/10.1128/mSystems.00916-21.

Editor Seth Bordenstein, Vanderbilt University

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Jan Hubert, carpoglyphus@gmail.com.

Received 13 July 2021
Accepted 7 October 2021
Published 2 November 2021

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

Microbes associated with arthropods can have significant impacts on host life history, ecology, digestion, and reproductive physiology and can shape their evolution. Interactions between microbes and their hosts can be broadly categorized as mutualistic, commensal, or parasitic, and these associations can be obligate or facultative. As examples of mutualistic associations, some arthropod species rely on the metabolic capabilities of symbiotic bacteria to colonize otherwise unavailable, nutrient-poor ecological niches (1–4). *Cardinium* is a Gram-negative endosymbiotic bacterium associated with insects, arachnids, and nematodes and belongs to the phylum *Cytophaga-Flavobacterium-Bacteroides*. *Cardinium* is very common among arthropods, with an estimated 13% of extant arthropod species potentially having associations with this bacterium (5–9). This bacterium can cause cytoplasmic incompatibility and feminization (10) and affect vector competence, sex ratios, and biotin provisioning in its hosts (11). The functional role of *Cardinium* associated with mites is unknown thus far.

Here, we investigated the molecular mechanisms underlying the interactions of *Cardinium* and its acarine host, the American house dust mite, *Dermatophagoides farinae* Hughes, 1961 (Pyroglyphidae). This is one of the most medically important species of house dust mites (12–14). House dust mites, including *D. farinae*, produce almost 40 different allergens causing allergic diseases that affect 15 to 30% of the world's population (15). Mite allergens and biologically active molecules of mite-associated bacteria have diverse molecular functions, affecting both mites and bacteria, which has implications for human health. For example, some mite immunogenic proteins may control bacterial populations based on their hydrolytic function (e.g., cysteine protease [Der f 1], trypsin [Der f 3], chymotrypsin [Der f 6]) and other antibacterial proteins (e.g., a bactericidal permeability-increasing protein-like protein [Der f 7], a bacteriolytic enzyme [Der f 38]) (15). In turn, mite-associated bacteria can also play a role in modulating mite allergen production (16–19) or induce IgE sensitization to microbial antigens, causing further impacts on human health (20).

Cardinium associated with *D. farinae* is a common hyperabundant bacterial species in U.S. and European populations (21–23). Phylogenetically, it is part of *Cardinium* group A (24) and a subgroup that is nearly exclusively associated with stored-product mites and predatory cheyletid mites (25). In *Brevipalpus* mites, *Cardinium* is localized in mesenchymal tissues, the reproductive tract, and appendages (26). A previous study (27) using Cy5-labeled universal bacterial probes indicated the presence of bacteria in the reproductive tract, mesodermal tissues, appendages, and hemolymph of *D. farinae*, although the specific locations of *Cardinium* in mite tissues still need to be confirmed. PCR-based analysis with symbiont-specific primers indicated that the bacterium was also associated with eggs and that it was more prevalent in females than in males, representing 92 and 67% of all bacterial taxa in females and males, respectively (25). A whole-genome sequence of *Cardinium* associated with *D. farinae* has been assembled. The total assembly length is 1.48229 Mb (GC content, 37.3%), and the genome codes for 1,247 proteins (25). Relative to other *Cardinium* genomes, *Cardinium* associated with *D. farinae* had a larger genome size and coded for larger numbers of protein-coding genes but was comparable in size to the genome of "*Candidatus* Amoebophilus asiaticus," an intracellular bacterium associated with *Acanthamoeba*. Due to its smaller genome size relative to free-living bacteria and its intracellular localization, the current hypothesis is that the endosymbiont of *D. farinae* show hallmarks of obligate endosymbiosis. Mite-*Cardinium* symbiosis may have an important medical consequence, since the presence of this bacterium is correlated with the presence of bacterial endotoxins, which can interfere with industrial mite-based vaccine production. For example, *D. farinae* usually has abundant endotoxins derived from Gram-negative bacteria compared to another house dust species, *Dermatophagoides pteronyssinus*, which typically lacks *Cardinium* (18, 28–30). Some *D. farinae* populations from Asia, however, may host different Gram-negative bacteria, i.e., *Bartonella* (16, 17, 31).

Previous attempts to eliminate *Cardinium* from *D. farinae* by applying tetracycline at a rate of 1 mg/g of diet to study the impact of these intracellular bacteria on mite

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

physiology were unsuccessful (22). Therefore, we employed correlation analyses among global gene expression profiles of *D. farinae* and *Cardinium* to identify potential molecular interactions between hosts and their symbionts (32). Aspects related to cytoplasmic incompatibility, nutrient exchange, and host cell interactions have already been addressed in the literature using the same methodology in nonmite hosts (33, 34). We therefore focused on other aspects important for mite-*Cardinium* symbiosis, particularly mite genes involved in the immune and metabolic pathways.

RESULTS

Metatranscriptome of *D. farinae*. We obtained 6.01×10^5 quality-trimmed transcriptomic reads that mapped to a previously published *Cardinium* assembly (25), and 161×10^6 reads mapped to the *de novo* assembled *D. farinae* transcriptome (see Data Set S1, sheets 1 to 7, in the supplemental material). Overall, transcripts derived from the following taxa were identified: *Cardinium* (1,358 predicted proteins), Alveolata (2,269), Bacteria excluding *Cardinium* (264), Microsporidia (620), and *D. farinae* (8,301). We further analyzed the mite (3×10^4 to 1.9×10^5 reads/sample) and *Cardinium* (0.5×10^3 to 2.8×10^3 reads/sample) gene expression data. A rarefaction analysis of normalized data sets showed that younger mite cultures generally expressed lower numbers of unique transcripts than old cultures (Fig. S1). The numbers of *Cardinium* 16S rRNA gene copies per individual mite did not differ between young and old cultures by the Mann-Whitney test (Fig. S2), as follows. For pooled mite samples, the values were z-score$_{(1,10)}$ = 0.081, $P = 0.94$, median (25th to 75th percentile) for young, 4.14 (4.09 to 4.21) $\times 10^4$ copies/individual, and median for old, 4.13 (4.07 to 4.24) $\times 10^4$ copies/individual. For individual mites, the values were z-score$_{(1,18)}$ = 1.85, $P = 0.0621$, median (25th to 75th percentile) for young, 4.58 (3.76 to 5.01) $\times 10^4$ copies/individual, and median for old, 3.16 (2.75 to 4.33) $\times 10^4$ copies/individual.

Effect of culture age on gene expression. In our experiments, old and young mite cultures differed in population structure, as older cultures had a higher proportion of adult mites than younger cultures. Distance-based redundancy analysis (dbRDA) correlation models indicated that the variable "culture age" explained 22% of the variability in predicted *Cardinium* KEGG gene expression and 48% variability in predicted mite gene KEGG gene expression (Table 1). When the "culture age" variable was entered in dbRDA models using either *Cardinium* or mite expression as the factor, it was eliminated by the forward selection procedure. Furthermore, the age of the mite culture was correlated with the expression of 79 predicted protein-encoding *D. farinae* genes. The expression of only 4 *Cardinium* predicted proteins prevailed in young culture, while the rest of the predicted proteins (e.g., BRU, PORCN, ZFPM1, and B9D) were not influenced by culture age (Fig. S3 and Data Set S1, sheet 8). However, variation in mite and *Cardinium* expression levels was observed among the biological replications of both the old and young samples, allowing this data set to be analyzed to identify correlations between mite and *Cardinium* genes that could be important for mediating interactions.

Mite regulatory pathways and *Cardinium* gene expression. Correlations between *Cardinium* and *D. farinae* gene expression were identified using dbRDA. We detected a total of 232 KEGG-annotated *D. farinae* genes belonging to signaling and regulatory pathways associated with mite immune responses, such as phagocytosis, apoptosis, the mitogen-activated protein kinase (MAPK) signaling cascade, endocytosis, the Toll pathway, the tumor necrosis factor (TNF) pathway, the transforming growth factor beta (TGF-β) pathway, lysozyme, and the Toll/Imd pathway (Data Set S1, sheet 9). In addition, four *D. farinae* genes (DF_06356, DF_09960, DF_14487, and DF_17932) that were not annotated by KEGG were included (Data Set S1, sheet 9) based on their high protein sequence similarity to Toll-like receptors. When the expression levels of these 236 mite genes were used as independent variables, they explained the same amount of variation ($R^2 = 0.756$ and 0.750) in *Cardinium* gene expression as the full data set (3,189 mite genes; $R^2 = 0.750$) (Table 1). *Cardinium* gene expression data entered as independent variables explained the higher variation ($R^2 = 0.952$) of expression levels in mite regulatory/immune pathways

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

TABLE 1 Correlation models based on dbRDA of expression of mite and *Cardinium* genes as dependent and independent variables[a]

ID	Dependent variable	Independent variable	df	*F*	*P*	R^2
1	*Cardinium*#	Culture_age	1, 10	2.801	0.13	0.219
2	Mite_genomic#	Culture_age	1, 10	9.144	0.002	0.478
3	*Cardinium*#	Mite_genomic	5, 6	3.6	0.001	0.75
4	Mite_genomic#	*Cardinium*	5, 6	7.283	0.001	0.859
5	Mite_regulatory pathways	Culture_age	1, 10	12.401	0.002	0.55
6	Mite_regulatory pathways	*Cardinium*	6, 5	16.623	0.001	0.952
7	*Cardinium*	Mite_regulatory pathways	5, 6	3.712	0.001	0.756
8	*Cardinium*	Lysosome	4, 7	3.933	0.001	0.692
9	*Cardinium*	Apoptosis	4, 7	3.824	0.001	0.686
10	*Cardinium*	Phagocytosis	4, 7	3.7	0.001	0.679
11	*Cardinium*	Endocytosis	4, 7	3.496	0.001	0.664
12	*Cardinium*	MAPK	3, 8	3.78	0.001	0.586
13	*Cardinium*	TOLL	2, 9	3.119	0.001	0.408
14	*Cardinium*	TOLL_IMD	3, 8	2.407	0.002	0.475
15	*Cardinium*	TNF	2, 9	3.138	0.002	0.411
16	*Cardinium*	TGF	2, 9	2.722	0.001	0.377
17	Phagocytosis	*Cardinium*	8, 3	40.176	0.001	0.991
18	Apoptosis	*Cardinium*	7, 4	33.034	0.001	0.983
19	MAPK	*Cardinium*	7, 4	27.883	0.001	0.98
20	Endocytosis	*Cardinium*	6, 5	22.181	0.001	0.967
21	TNF	*Cardinium*	7, 4	13.088	0.001	0.952
22	TGF	*Cardinium*	6, 5	9.416	0.001	0.919
23	Lysozome	*Cardinium*	6, 5	8.8	0.001	0.913
24	TOLL	*Cardinium*	7, 4	11.778	0.001	0.953
25	TOLL_IMD	*Cardinium*	6, 5	6.8373	0.001	0.891
26	Mite_metabolic pathways	Culture_age	1, 10	2.791	0.004	0.218
27	Mite_metabolic pathways	*Cardinium*	6, 5	4.408	0.001	0.841
28	*Cardinium*	Mite_metabolic pathways	3, 5	3.387	0.001	0.56

[a]These models are based on gene expression estimates for predicted KEGG genes for *Cardinium* and select mite molecular pathways (see Tables S7 and S9 for details). The models were either full models (marked by #) or partial models inferred by the forward selection procedure. R^2 values describe the explained variability by the models. The significance of the models was tested by a permutation test ($n = 999$ permutations). For each model, degrees of freedom (df), *F*, and *P* values are provided.

than the inverse ($R^2 = 0.756$). Partial dbRDA models were constructed using the expression levels of *Cardinium* genes as the dependent variables (all predicted KEGG proteins) and the expression of mite genes belonging to selected pathways as independent factors (i.e., phagocytosis, 31 predicted KEGG-annotated proteins; apoptosis, 30; MAPK signaling cascade, 48; endocytosis, 78; Toll pathway, 21; TNF pathway, 17; TGF-β pathway, 26; lysozyme, 49; Toll/Imd pathway, 23) (Data Set S1, sheet 9; Table 1). The explained variability in these dbRDA models decreased in the following order: lysosome, apoptosis, phagocytosis, endocytosis, MAPK signaling cascade, Toll/Imd pathway, Toll pathway, TNF pathway, and TGF-β pathway (Table 1). These models indicate a correlation (R^2 ranging from 0.377 to 0.692) between the expression of predicted *Cardinium* genes as the dependent variable and the expression of genes in these pathways as the independent variable. The expression of predicted *Cardinium* genes as an independent variable (R^2 ranging from 0.891 to 0.991) was likewise strongly correlated with the expression of predicted mite genes in immune regulatory pathways, decreasing in the following order: phagocytosis, apoptosis, MAPK signaling cascade, endocytosis, TNF pathway, TGF-β pathway, lysosome, Toll pathway, and Toll/Imd pathway.

For the matrix of the mite KEGG-annotated genes in these 9 pathways, including 236 predicted mite genes and the expression data of 448 predicted *Cardinium* genes, we calculated Spearman rank correlations (Data Set S1, sheets 10 and 11) and obtained 6,684 significant correlations ($P < 0.05$). These data were then visualized as a heat map (Fig. 1) and network graph (Fig. S4). There were a total of 105,020 combinations, and 3.5% of the combinations showed significant positive pairwise correlations while 2.9% showed significant negative pairwise correlations (Fig. 1). A cluster analysis of these

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

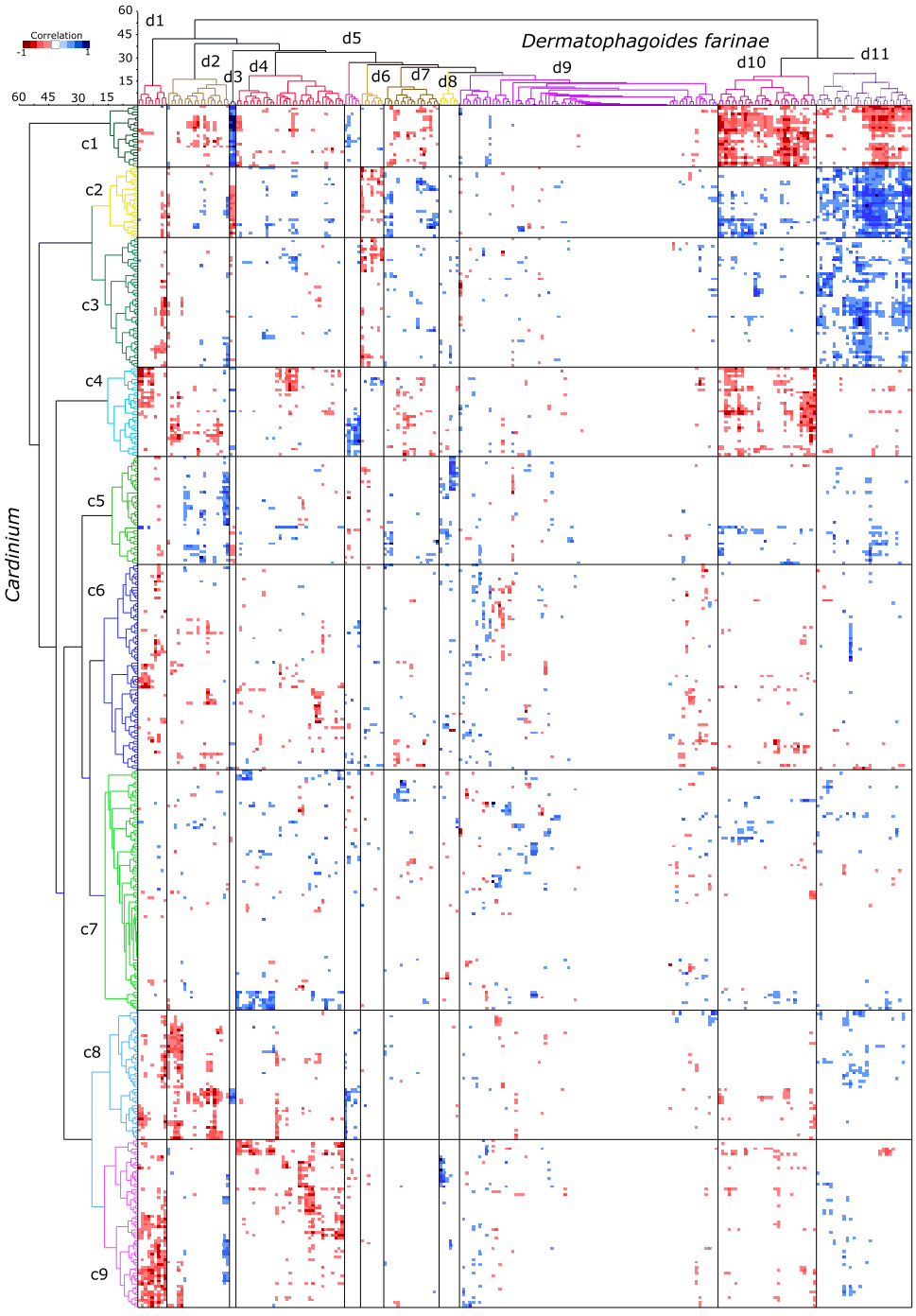


FIG 1 Spearman correlations of *Cardinium* expression and mite expression for select regulatory *D. farinae* pathways. Only significant correlations are shown ($P < 0.05$). The data were clustered by Ward's method; clusters and predicted KEGG-annotated proteins are described in Data Set S1, sheets 10 and 11, in the supplemental material.

pairwise expression levels revealed 11 clusters of mite transcripts (d clusters = 11) with similar expression profiles and nine clusters of *Cardinium* expression patterns (c clusters = 9). The heat map showed the following features. Peaks of positive correlations were observed for the following cluster combinations: c1 and d3 (98% of the genes in this cluster were significantly and positively correlated with *Cardinium* expression), c2 and d11 (65%), c3 and d11 (34%), c4 and d5 (27%), and c5 and d8 (23%) (Fig. 1). Peaks of negative correlations were observed for the following combinations

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

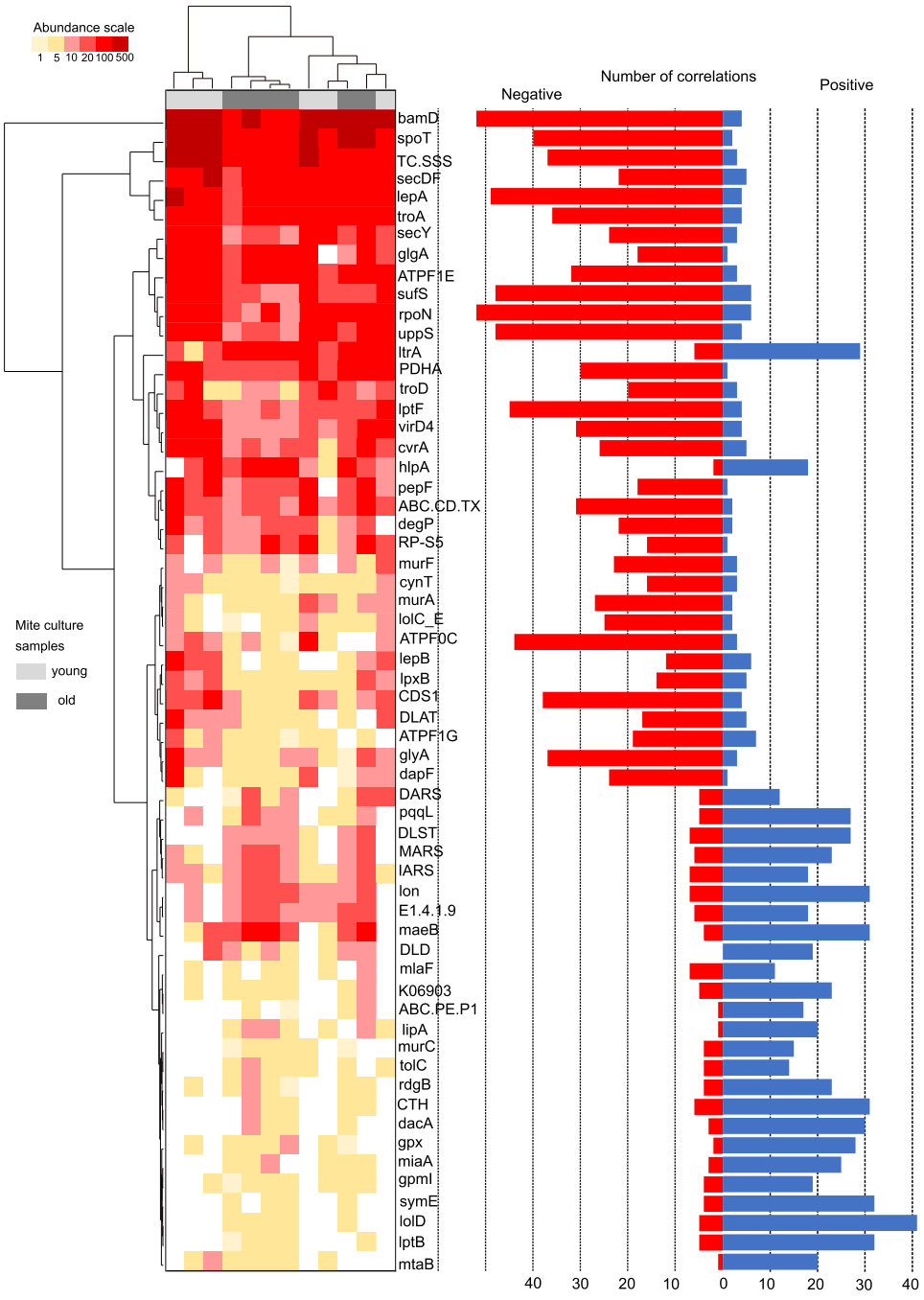


FIG 2 Expression of genes predicted to code for select *Cardinium* proteins and their correlations with the expression of genes belonging to mite regulatory and immune pathways. Standardized abundance data were used to generate the heat map; Ward's method was used for clustering. The *Cardinium* predicted proteins are described in Data Set S1, sheet 10.

of clusters: c2 and d3 (98%), c2 and d10 (65%), c2 and d11 (38%), c9 and d1 (36%), c2 and d6 (27%), c8 and d2 (22%), and c4 and d1 (20%) (Fig. 1). Details regarding these pairwise mite-symbiont expression correlations are given in Data Set S1, sheet 10. There was a marked pattern of the levels of highly expressed *Cardinium* predicted proteins (e.g., BamD, LepA, LptF, and VirD4 but not LtrA) having negative correlations with the expression of mite genes belonging to immune/regulatory pathways (Fig. 2), while the reverse was also true, i.e., the levels of weakly expressed *Cardinium* predicted

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

proteins (e.g., PltB, SymE, MurC) were mostly positively correlated with expression of genes from mite immune/regulatory pathways.

***Cardinium* and metabolic pathways of *D. farinae*.** For *D. farinae*, 372 transcripts were assigned to KEGG pathways of interest (Data Set S1, sheets 12 to 15). These predicted mite genes were assigned to 63 pathways, forming three clusters with low, intermediate, and strong expression levels (Fig. 3). The pathway with the highest expression levels included transcripts associated with carbohydrate metabolism, purine metabolism, and glycan biosynthesis (Fig. 3). The expression levels of mite transcripts in metabolic pathways entered as independent variables showed a correlation with the expression of predicted *Cardinium* KEGG genes as dependent variables in the dbRDA model ($R^2 = 0.56$) (Table 1). The dbRDA model using *Cardinium* KEGG gene expression as an independent variable and mite metabolic pathway expression as a dependent variable explained higher variability ($R^2 = 0.841$) than the previous model. The "culture age" variable explained substantially lower amounts of variation ($R^2 = 0.21$) than the *Cardinium* expression levels.

Additionally, we calculated Spearman correlations for expression in each predicted KEGG *Cardinium* pathway versus the metabolic pathways (Data Set S1, sheets 14 and 15), which resulted in 28,098 pairwise comparisons, of which 2% (565) were significant ($P < 0.05$) and positive and 6% (1,637) were significant and negative (Fig. 4). A two-way cluster analysis of expression patterns identified 8 *Cardinium* clusters (c1 to c8) and 7 clusters for mite metabolite pathways (d1 to d7) (Fig. 4). Significant negative correlations were identified for the following cluster pairs: c1/d1 (41% and 0% of pathways with negative and positive correlations, respectively; $n = 429$), c2/d1 (58% and 0%; $n = 396$), c3/d1 (24% and 0%; $n = 374$), c4/d2 (63% and 0%; $n = 270$), c5/d7 (26% and 0%; $n = 700$), and c6/d1 (21% and 0%; $n = 360$). Significant positive correlations were found for the following cluster pairs: c1/d7 (8% positive and 0% negative, respectively; $n = 564$), c2/d7 (6% and 0%; $n = 504$), and c4/d3 (57% and 0%; $n = 90$). Thus, most of the correlations were negative (Fig. 5). Furthermore, *Cardinium* genes belonging to clusters c1, c2, and c4 (showing both positive and negative correlations) could potentially both up- and downregulate mite gene expression; in contrast, cluster c6 could likely only downregulate mite gene expression (Fig. S9, correlation network analysis). In contrast to the patterns detected for immune/regulatory pathways (Fig. 2), several highly expressed *Cardinium* genes (e.g., a gene coding for VirD4) had more positive correlations with mite metabolic pathway gene expression than genes with lower expression levels (e.g., genes coding for RfbB, DcuB, and LolD) (Fig. 5).

DISCUSSION

Mite metagenomics. The results of metatranscriptomic analysis again confirmed that cultured *D. farinae* mites are associated with *Cardinium* (25). We identified *Cardinium* as the most abundant mite symbiont (1,358 predicted proteins), followed by other bacteria (260 predicted proteins). Our quantitative PCR (qPCR) abundance estimates suggest that, on average, each mite had 103 to 104 *Cardinium* cells, a result consistent with previous studies (25). We also identified two unknown and possibly novel eukaryotic mite symbionts: Alveolata (1,145 to 5,500 reads per sample) and Microsporidia (203 to 392 reads per sample). The presence of the following alveolates associated with *D. farinae* has also been previously shown using light microscopy: *Gregarina* (Apicomplexa) (35, 36) and unidentified flagellated protozoa (36, 37). However, to our knowledge, Microsporidia have not yet been detected in *D. farinae*, although they are known to occur in stored-food mites, such as *Tyrophagus putrescentiae* (22, 27, 38). These microsporidians should therefore be identified morphologically in the future.

Effect of culture age. The age of mite cultures can affect the levels of allergens and other proteins (39, 40), as well as the microbiome composition (22) in mites. *D. farinae* cultures show specific growth patterns during cultivations. There are three population growth phases: (i) a latency phase (0 to 10th week) that is characterized by a slow increase in the mite population size; (ii) an exponential phase (10th to 20th week) that

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

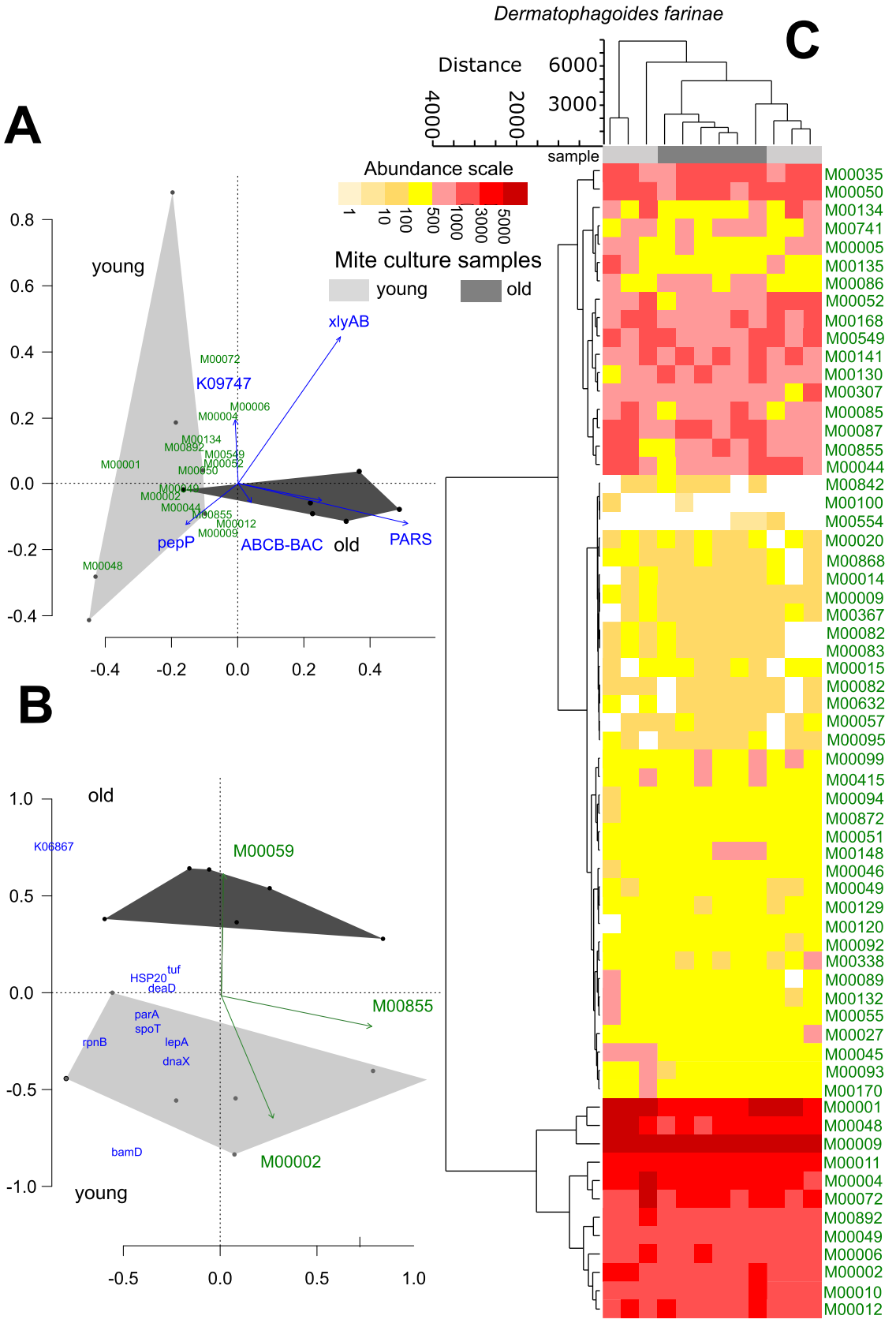


FIG 3 Interaction of expression of *Cardinium* genes with the expression of genes from select metabolic pathways of *D. farinae*. (A and B) Distance-based redundancy analysis (dbRDA) models describing the distribution of samples from young and old mite cultures. (A) Model using the expression of mite pathways (green) as the dependent variable and *Cardinium* predicted proteins as the

(Continued on next page)

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

is characterized by exponential growth of the mite population; and (iii) a "decline" phase (20th week) that is characterized by a large decrease in the mite population size (41, 42). Here, we showed that samples collected in the exponential (young) and decline (old) phases largely did not differ in *Cardinium* abundance or gene expression, with only 4 *Cardinium* genes having significant expression differences. This indicates that *Cardinium* probably does not contribute to mite culture collapse during the decline phase. Remarkably, our rarefaction analysis showed some differences between young and old cultures in the number of expressed transcripts, which was 2-fold higher in the mites and 0.5-fold higher in *Cardinium* in old cultures than in young cultures. Different demographic structures (i.e., a higher number of juveniles in the young cultures) are probably responsible for this effect on mites.

Effect of *Cardinium* on mite immune/regulatory pathways and metabolism. It is commonly accepted that *Wolbachia* interacts with the host immune system, influencing gene expression to ensure its persistence within the host (43). The host immune system expresses effector molecules that promote the proliferation of these intracellular symbionts inside the host (44, 45). The proliferation of *Cardinium* may be affected by a similar mechanism. In particular, the expression of 31% of *Cardinium* genes was correlated with the expression of mite genes belonging to mite immune and/or regulatory pathways.

Several molecular pathways and receptors influenced by *Cardinium* were identified. The TGF-β receptor pathway regulates parasites in the host gut (46, 47). The Toll pathway is involved in mediating innate immune responses in arthropods and is activated by Toll-like receptors (TLRs) (48). The TNF pathway stimulates inflammation and triggers apoptosis through caspase (49). Two predicted proteins from our *de novo* transcriptome assembly were candidates for TNF receptors (DF_02523, DF_22218). However, their expression levels were very low. The NF-kappa-B receptor belonging to the Toll pathway was not detected here, although a candidate NF-kappa-B receptor was previously found in *D. pteronyssinus* (50). We identified four possible Toll-like receptors in the *D. farinae* transcriptome that were not assigned to any KEGG pathway (DF_06356, DF_14487, DF_09960, DF_17932). Their transcripts showed both positive and negative correlations to transcripts from the *Cardinium* endosymbiont (Data Set S1, sheet 10). We hypothesize that *Cardinium* can activate mite TLRs with its surface patterns (see below for membrane-associated proteins), as is known for other invertebrate hosts (51–54).

The mitogen-activated protein kinase (MAPK) signaling pathway is involved in the regulation of multiple cellular processes, such as inflammation, immunity, cell differentiation, proliferation, and death. Caspases involved in MAPK induce apoptosis when overexpressed in tissue culture cells, and bacteria have evolved molecules that downregulate caspases (55). Here, we identified mite transcripts encoding CASP3 and CASP7. CASP3 expression was correlated with the expression of 24 *Cardinium* genes (e.g., negative correlations for genes coding for metabolism-associated proteins PqqL and RdgB and signaling molecule protein MlaF, and positive correlations for genes coding for metabolism-associated proteins DLAT, MurF, and SufS). CASP7 expression was negatively correlated with that of 4 *Cardinium* genes (e.g., coding for membrane transport protein LptF and environmental information processing protein RpoN) and positively correlated with 14 genes (e.g., coding for signaling and cellular process protein CTH and uncharacterized protein SufD and for metabolism protein DacA). By these mechanisms, intracellular bacterial symbionts can manipulate host cells, for example, to avoid apoptosis of nurse cells, which provides nutrients to growing symbiont-infested oocytes (56–58).

Endocytosis is a mechanism of entry of symbionts into the host cell. For example,

FIG 3 Legend (Continued)
factor (indicated by blue arrows). (B) Model using expression of predicted *Cardinium* proteins as the dependent variable (blue) and mite metabolic pathways as the factor (green arrows). Both models had selected factors by forward selection. The orthogonal hulls indicate sample distribution (young and old cultures) in the model. (C) Heat map of the expression of *D. farinae* predicted proteins in metabolic pathways based on standardized data set; Ward's method was used for clustering.

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

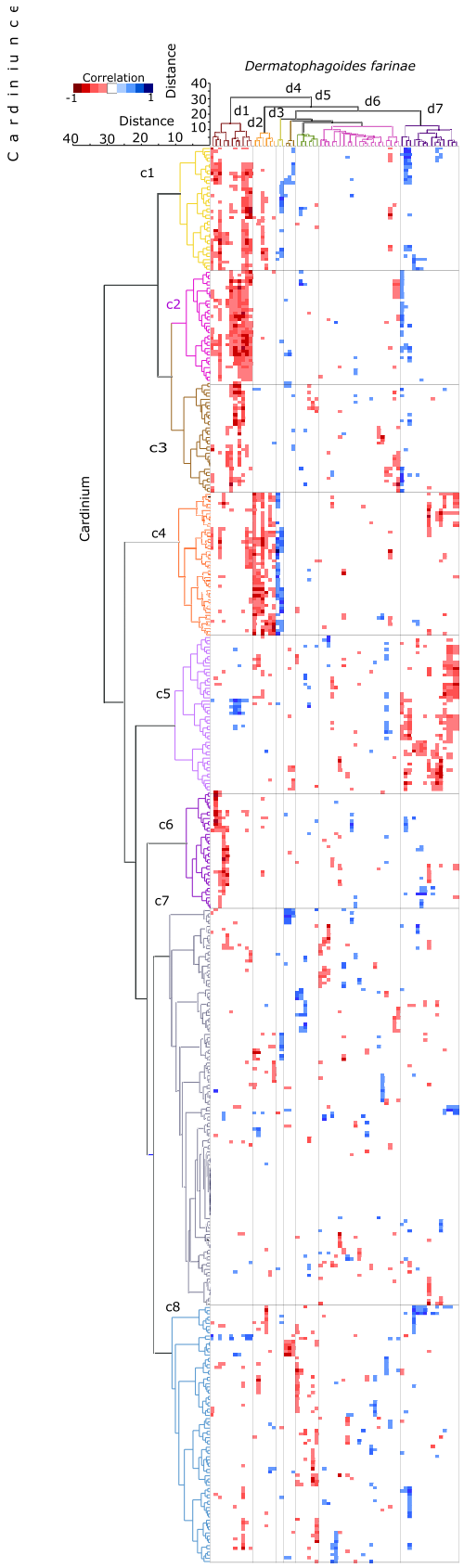


FIG 4 Spearman correlations of gene expression of metabolic pathways of *D. farinae* and *Cardinium* expression. Only significant correlations ($P < 0.05$) are shown. The data were clustered by Ward's method; clusters are described in Data Set S1, sheets 14 and 15.

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

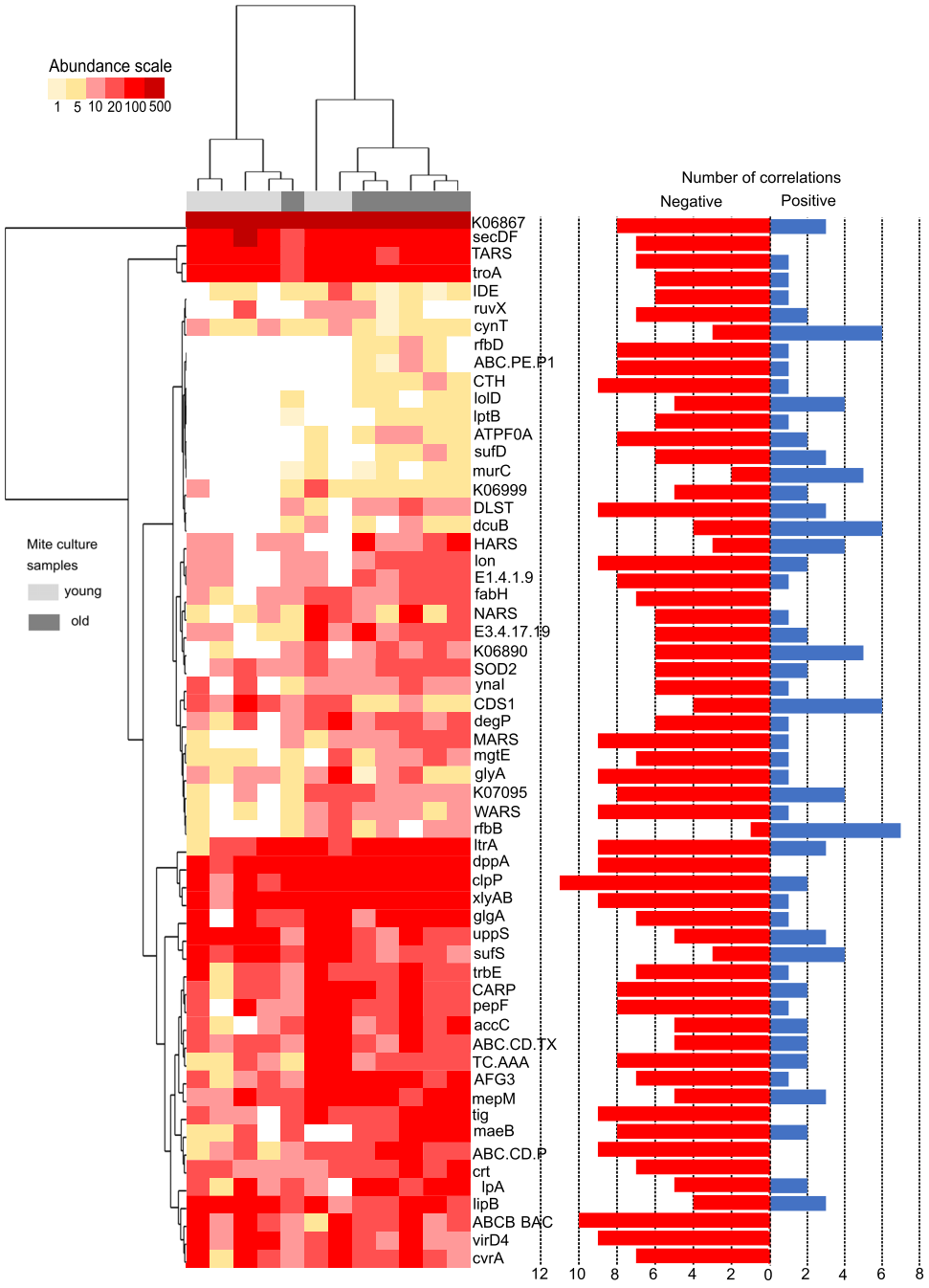


FIG 5 Expression of select *Cardinium* genes and their correlations with the expression of mite genes belonging to metabolic pathways. Standardized abundance data were used to generate the heat map; Ward's method was used for clustering. The *Cardinium* predicted proteins are described in Data Set S1, sheet 15.

Buchnera aphidicola) facilitates TOR pathway-mediated endocytosis to gain entry into host cells (*Acyrthosiphon pisum*) in the aphid blastula (59, 60). *Wolbachia* utilizes host cell phagocytic and clathrin/dynamin-dependent endocytic machinery for horizontal cell-to-cell transfer (61). Additionally, *Wolbachia* resides in phagosomes, or vacuoles derived from the host endoplasmic reticulum (ER), which help the bacterium replicate and escape host immune surveillance (62). A similar mechanism for cytoplasmic entry and intracellular localization is expected for *Cardinium*.

Previous studies involving fluorescent *in situ* hybridization with Cy5-labeled universal bacterial probes identified bacteria in the reproductive tract, appendages, and

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

mesenchymal tissues and inside hemocytes of *D. farinae* (27). The majority of the bacterial reads from *D. farinae* were identified as *Cardinium* in this and previous studies (25). For this reason, it is currently hypothesized that *Cardinium* occurs inside hemocytes, indicating the hemocyte-based control of *Cardinium* through phagocytosis (63). In this study, we found that expression of two mite cathepsins (CTSL and CTSF) was negatively correlated with *Cardinium* protein expression (Data Set S1, sheet 10). Inside phagocytes, *Cardinium* may be degraded by the proteolytic action of cathepsins, indicating a potential regulatory function of mite cathepsins for *Cardinium*.

Predicted *Cardinium* proteins and mite molecular pathways. Several *Cardinium* genes were predicted to up- or downregulate genes from mite immune pathways via interaction with pathway receptors. The Toll receptor is an example of such a mite receptor that may recognize several *Cardinium* proteins, such as DegP (serine protease Do [EC 3.4.21.107]), MlaD (phospholipid/cholesterol/gamma-HCH transport system substrate-binding protein), and BamD (outer membrane protein assembly factor). DegP is a bifunctional bacterial protein acting both as a protease and a chaperone regulating cell growth (64). In the host cell, DegP of "*Candidatus* Liberibacter asiaticus" can presumably digest host proteins using zinc as a cofactor; therefore, DegP is considered a virulence factor (65). The Mla A-F system is universally conserved in Gram-negative bacteria; this system plays an important role in phospholipid transport to ensure the translocation of phospholipids between the inner membrane and periplasmic proteins (66). BamD is an outer membrane lipoprotein of Gram-negative bacteria that is essential for viability. The protein is translocated across the inner membrane and maintained in an assembly-competent state by periplasmic chaperones (67). In our system, the expression of genes coding for the metabolism-associated protein DegP, membrane transport protein MlaD, and signaling protein BamD showed negative (AP1G1, SEC61A, CANX, TUBA, CDC42, HSPA1s, ARFGAP1, RAB11A, RAB35, SRC, and STK3) or positive (e.g., MRT43, PLD1_2, EMC, and LMNB) correlations with expressed genes from all analyzed mite immune/regulatory pathways. The lipopolysaccharide transport system including LptF (lipopolysaccharide export system permease protein) in Gram-negative bacteria is responsible for transporting lipopolysaccharide from the cytoplasmic surface of the inner membrane (68). In our system, LptF expression showed a positive correlation with mite endocytosis, indicating that an increase in *Cardinium* lipopolysaccharide production may trigger an endocytic response from the mite host to control *Cardinium*.

Cardinium assembles Fe-S clusters, used as cofactors for many biological processes, by the sulfur formation system Suf (69). Interestingly, in our system, *Cardinium* SufS (cysteine desulfurase/selenocysteine lyase; EC:2.8.1.7 and EC:4.4.1.16) had strong correlations with several mite genes, mainly with genes coding for proteins involved in MAKK ($n = 16$) and endocytosis ($n = 21$). The prevailing correlations were negative, which may be explained by the need of the bacterium to compete for iron with its host.

The function of the GTP-binding protein LepA in bacterial cells is known in *Legionella*, in which LepA and LepB were shown to promote the nonlytic release of bacteria from their protozoan host by commandeering the exocytic pathway to release type IV secretion system (T4SS) effectors (70–72). A similar function is suggested for this protein in *Cardinium*; however, we did not identify proteins associated with the T4SS, except for VirD4. We found positive correlations of LepA with clathrin-mediated endocytosis.

The protein RpoN (RNA polymerase sigma-54 factor) is a regulator of genes involved in nitrogen metabolism and nitrogen assimilation under nitrogen-limited conditions (73); *rpoN* has been found in many free-living and symbiotic bacteria, e.g., "*Ca.* Amoebophilus asiaticus" (74), "*Ca.* Walczuchella monophlebidarum" (75), *Blattabacterium* (76), and *Sodalis* (77). The increase of RpoN expression in *Cardinium* is correlated with increased expression of genes coding for Toll receptors of mite immune/regulatory pathways, both an increase and decrease of expression of genes involved in the endocytosis pathway, a reduced expression of genes encoding caspases (apoptosis), a reduced expression of

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

genes coding for polyamine biosynthesis and purine metabolism, and a reduced expression of genes coding for proteins involved in glycosaminoglycan and pyrimidine metabolism in mites. These results highlight the importance of *rpoN* in *Cardinium*-host interactions.

We also identified the following proteins in *Cardinium* that had not been previously detected (11, 78, 79): DdpD (ATP-binding cassette domain-containing protein; GenBank no. WP_186292434), VirD4 (T4SS protein VirD4 [EC 7.4.2.8]), and the toxin protein SymE (Fig. S5 and S6). Our alignment of SymE indicated a high similarity to the UniProtKB A0A368ZU39 protein of *Schleiferia thermophila* (Fig. S5). Our finding of genes encoding components of the T4SS (VirD4) is remarkable because this secretion system has not yet been found in other known species of *Cardinium* (11), and this finding is consistent with the presence of LepA (see above), which is also known to interact with this system. The protein encoded by *virD4* is involved in the production of a transport membrane protein of the T4SS (80). We observed that the expression of the gene predicted to code for VirD4 was correlated with regulatory and immunological pathways in *D. farinae* (Fig. 5). However, other components of the T4SS were not found in our *Cardinium* transcriptome.

Effect of *Cardinium* on host metabolism. Previous studies have associated the presence of intracellular symbionts with increased levels of mitochondrial oxidative phosphorylation in the host, indicating that these bacteria stimulate host metabolism (81, 82). Although the precise mechanism by which these bacteria stimulate oxidative phosphorylation is not known, key nutrients provided by intracellular bacteria, such as vitamins, lipoate, pantothenic acid, biotin, and riboflavin, may play a role (32, 81, 83, 84). Our study indicated that *Cardinium* gene expression was negatively correlated with the expression of genes involved in several metabolic mite pathways (Data Set S1, sheet 10). Negative expression correlations were 3 times more frequent than positive correlations (Fig. 5). Mite populations heavily infested with *Cardinium* have significantly lower population growth rates (22) and can thus potentially affect mite energy metabolism. The genomes of many intracellular symbionts are highly reduced in comparison to *Escherichia coli*, e.g., whose genome is 4.64 Mb and has 4,609 genes (85), while the genome of the *Cardinium* associated with *D. farinae* is only 1.48 Mb (1,403 genes). The symbionts are largely dependent on their host cells for several nutrients (86, 87). Likewise, the *Cardinium* genome is reduced, making this bacterium highly dependent on its eukaryotic host for obtaining most key metabolites and energy in the form of ATP (11). Collectively, this could explain the interactions between *Cardinium* genes and mite genes coding for proteins involved in core metabolism.

Cardinium bacteria associated with *D. farinae* have a complete pathway for lipoic acid biosynthesis, as is known for *Cardinium hertigii* (11). Lipoate is a highly conserved sulfur-containing cofactor involved in oxidative reactions in a previously characterized *Cardinium* strain (11, 88). Lipoate produced by *Cardinium* could help the mite host overcome the negative effect on oxidative phosphorylation. Previous studies have shown that biotin may be a key nutrient provided by *Cardinium* to some insect hosts (11, 79), while this was not true for bacteria associated with a strain of *Bemisia tabaci* (89) or the nematode *Pratylenchus penetrans* (5). Interestingly, the mite-associated *Cardinium* also lacks genes encoding the enzymes BioA, BioB, BioD, and BioF but expresses biotin transporters (BioY, BioN, BioM).

Potential effect of *Cardinium* on mite pheromone production. Some bacteria are known to mediate the aggregation of their hosts (90) and manipulate their defense strategies (91). For example, a decrease in *Wolbachia* infection in *Thasus neocalifornicus* was associated with reduced production of defensive and alarm pheromones by the host (92). *D. farinae* has opisthosomal glands (93) that produce a variety of compounds, including hydrocarbons, terpenes, aromatics, and alkaloids, with suggested defensive, alarm, aggregation, and sex pheromone functions (94). The terpenoid pathways of *D. farinae* likely produce geranylgeranyl diphosphate, which is required to produce defense-related secretions in aphids and termites (95, 96). In our study, the expression of genes from the pathway responsible for the synthesis of terpenoids was positively

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

correlated with the expression of *Cardinium* genes from cluster c4 (Fig. 4). We hypothesize that *Cardinium* can stimulate the production of sexual or aggregation pheromones (97–99) to encourage mites to congregate and increase its own vertical transmission.

Conclusions. The molecular interactions between the mite *D. farinae* and *Cardinium* host included the correlation in the transcriptome of mite immune/regulatory and metabolism pathways (apoptosis, endocytosis, lysozome, MAPK, TGF, TNF, and TOLL/IMD) and *Cardinium* genes (e.g., BamD, LepA, SymE, and VirD4). *Cardinium* protein expression was negatively correlated with 3-fold-more predicted mite metabolic pathways (glycolysis and citrate cycles) than those of positive correlations. *Cardinium* possesses a complete pathway for lipoic acid biosynthesis but not biotin. The correlation analysis indicates the possible effect *Cardinium* on terpenoid pathways of *D. farinae* and production of sexual or aggregation pheromones.

MATERIALS AND METHODS

Mite populations. The laboratory cultures of *D. farinae* originated from rearing facilities of the Crop Research Institute, Prague, Czechia, and were maintained as described previously (22). Mites were mass reared in Iwaki flasks on house dust mite diet (HDMd) and then harvested. The detailed rearing and harvesting methodology was described previously (21).

Sampling. Mites were reared in 12 Iwaki chambers, and 1-month (young) and 3-month (old) cultures were collected for further analysis. Live mites were collected from the plugs and surfaces of the flasks with a brush and placed into sterile microcentrifuge tubes. We modified a previously published protocol to collect the mites for RNA extraction and perform surface sterilization to reduce contamination from environmental microbes that may be present on the exoskeleton (27), as follows: (i) 30- to 40-mg samples of mites were weighed (the fresh weight of a mite is approximately 6 μg, so each sample contained 5,000 to 6,700 mites); (ii) each mite sample was placed into 100% ethanol, vortexed for 5 s, and centrifuged (13,000 $\times$ *g* for 1 min); (iii) the supernatant was replaced with a 1:10 bleach solution containing 5% sodium hypochlorite, and then samples were mixed by vortexing for 5 s and centrifuged (13,000 $\times$ *g* for 2 min); and (iv) samples were washed twice with double-distilled water (ddH_2O) to remove residual bleach. Steps i to iv were performed on ice, and we collected six biological replicates for each 1- and 3-month culture for a total of 12 samples. The same methodology was used for mite sampling for DNA isolation.

RNA extraction. RNA isolation was performed using the NucleoSpin RNA kit (catalog no. 740984.50; Macherey-Nagel, Duren, Germany) according to the manufacturer's instructions, with the following modifications: (i) mites were homogenized for 30 s in a glass tissue grinder (Kavalierglass, Prague, Czechia) in 500 μl of lysis buffer; (ii) samples were centrifuged at 2,000 $\times$ *g* for 3 s; and (iii) DNA was degraded by DNase I at 37°C according to the manufacturer's protocol (Riboclear plus, catalog no. 313-50; GeneAll, Lisbon, Portugal). RNA quality was evaluated using a NanoDrop (NanoDrop One; Thermo Scientific, Waltham, MA, USA) and an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and samples were shipped on dry ice to the MrDNA laboratory (Shallowater, TX) for downstream processing and sequencing.

DNA extraction for quantification of *Cardinium* by qPCR. Mites were transferred to 2.0-ml screwcap MCT conical NS tubes (catalog no. CP5912; Alpha Laboratories, Eastleigh, Hampshire, UK) with 0.5 g of an equal mixture of 0.3-mm and 1.0-mm garnet sharp particles (catalog no. 11079103gar and 11079110gar; BioSpec) and 1 glass bead (3-mm diameter; catalog no. R155761; BioSpec). The samples were homogenized in a mini-beadbeater (BioSpec) for 5 min and then centrifuged (10,000 $\times$ *g* for 2 min). The supernatant was transferred to a new sterile microcentrifuge tube with 20 μl of proteinase K and incubated at 56°C overnight. The remainder of the DNA extraction procedure was performed using the QIAamp DNA micro kit (Qiagen, Hilden, Germany; catalog no. 56304) and by following the manufacturer's protocol for tissue samples. This method was used to extract DNA from pools for mites. In contrast, DNA from individual mites was isolated using 50 μl of a 5% solution of Chelex sodium (catalog no. C7901; Sigma-Aldrich, Saint Louis, MO, USA). Mites were crushed by pipetting, and then 3 μl of proteinase K was added. Samples were incubated at 56°C overnight, and the rest of the DNA extraction procedure was performed according to a previously described protocol (100). The abundance of *Cardinium* bacteria was also quantified using a previously described protocol (22). In brief, amplification of a 160-bp fragment of the 16S rRNA gene of *Cardinium* was performed on a StepOnePlus real-time PCR system (Life Technologies, Carlsbad, CA, USA) using 96-well plates with GoTaq qPCR master mix (catalog no. A6001; Promega) containing SYBR green as a double-stranded-DNA-binding dye. The primers Card_6QF3-TGCAAATCTCAAAAGCATGT-5 and Card_6QR 3-TCAAGCTCTACCAACTCCCA-5 (101) were used.

RNA processing and sequencing. The concentration of total RNA was determined using the Qubit RNA assay kit (Life Technologies). For rRNA depletion, first, 1,000 ng of total RNA was used to remove any residual DNA contamination using Baseline-ZERO DNase (Epicentre, Madison, WI, USA) by following the manufacturer's instructions, followed by purification using the RNA Clean & Concentrator kit (Zymo Research, Irvine, CA, USA). DNA-free RNA samples were used for rRNA removal with the Ribo-Zero Gold rRNA removal epidemiology kit (Illumina, San Diego, CA, USA), and final purification was performed using RNA Clean & Concentrator columns (Zymo Research). DNA-free and rRNA-depleted RNA samples were used for library preparation using Kapa mRNA hyperprep kits (Kapa Biosystems, Wilmington, MA,

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

USA) by following the manufacturer's instructions. In all the libraries, DNA concentration was measured using the Qubit dsDNA HS assay kit (Life Technologies), and the average library size was determined using an Agilent 2100 bioanalyzer (Agilent Technologies). The libraries were then pooled in equimolar ratios at 2 nM, and the library pool (10.0 pM) was clustered using cBot (Illumina) and sequenced using the Illumina HiSeq 2500 system as 250 × 250 bp paired-end (PE) ends (500 cycles).

Sequence processing. Read processing was performed using CLC Genomic Workbench v.20 (Qiagen, Venlo, Netherlands). PE reads were trimmed to remove residual adapter sequences and then filtered by length and quality score according to the default software parameters. The filtered reads were mapped to the previously described genome of *Cardinium* (25) (BioProject accession no. PRJNA555788; GCA_007559345.1), and unmapped reads were *de novo* assembled and processed as the *D. farinae* meta-transcriptome (25) using CLC Genomic Workbench software.

Contig processing. The reads were *de novo* assembled to generate 6×10^4 contigs with a contig N_{50} of 1,423 bp. In this data set, 3×10^4 coding DNA sequences (CDSs) were predicted and annotated from these contigs by using a set of known proteins as the trusted source (see below) using Prokka (102, 103) installed on the Galaxy server (104).

Expression analyses. Trimmed reads were mapped against the annotated transcriptome assembly generated as described above using CLC Workbench. The mapped reads ranged from 32×10^6 to 90×10^6 paired reads per sample for *D. farinae* and 0.16×10^6 to 1.2×10^6 paired reads per sample for *Cardinium*. The total numbers of reads mapped to each mite transcript and *Cardinium* locus were used to generate a count matrix for the expression analysis.

Protein identification. The *Cardinium* data set was annotated using the *Cardinium* annotated genome (GenBank GCA_007559345.1; see Erban et al. [25]). For improved identification of mite proteins, proteins from previously assembled and annotated genomes were used for comparison: arthropods (*D. pteronyssinus*, *Sarcoptes scabiei*, *Stegodyphus mimosarum*, *Tetranychus urticae*, and *Tribolium castaneum* [105–107]), alveolates/protozoans (*Besnoitia besnoiti*, *Toxoplasma gondii*, *Vitrella brassicaformis*, and *Gregarina niphandrodes*), nonmicrosporidian fungi (*Aspergillus fumigatus*, *Wallemia ichthyophaga* [108], *Candida albicans* [109], and *Saccharomyces cerevisiae*), bacteria (*Bacillus cereus*, *Kocuria rhizophila*, and *Bartonella* [110–114]), and microsporidia (*Nematocida displodere* and *Vittaforma corneae*).

PHMMER (115) bundled with the Galaxy server was used to match predicted proteins from the *de novo* transcriptome to annotated proteins in the reference library above. Matches with bit scores higher than 300 were retained and used for functional predictions. Some proteins were manually annotated with the EMBL-EBI website with reference proteomes and assembled genomes as databases (116). Proteins with the best matches to Archaea, Chordata, Cyanobacteria, plants, and nonmicrosporidian fungi were discarded as possible contaminants of the transcriptome from the mite diet. We also eliminated apicomplexan, microsporidian, and bacterial proteins (except *Cardinium*) from prior analyses (Data Set S1, sheets 4 and 7).

Selected proteins and metabolic pathway analyses. The KEGG mapper tool was used to infer cellular functions from protein sequences using the GhostKOALLA annotator against the genus prokaryotes + family eukaryotes + viruses databases (117). A comparison of predicted proteins of *D. farinae* (with or without expression) to *D. pteronyssinus* is given in Data Set S1, sheet 16.

Statistical analyses. We tested whether the numbers of *Cardinium* 16S rRNA gene copies in young and old mite cultures were significantly different using a Mann-Whitney test in PAST (118) on $\log_{10}$-transformed data. Expression analyses of mite, *Cardinium*, microsporidian, and alveolate genes were performed separately and then subsampled by recalculation to the lowest number of reads per sample (119). The standardized and unstandardized data were compared by nonmetric multidimensional scaling (NMDS) (Fig. S8), and the outliers of unstandardized data are shown in Fig. S9. Rarefaction analyses were performed in PAST (118) to standardize reads. Expression of the genes assigned to KEGG pathways of interest was used for further analyses. Comparisons between the effects of the cultures (old and young) were performed using dbRDA on the Bray-Curtis dissimilarity distance matrix in the R package vegan (120). To identify genes with different expression levels between old and young mite cultures, we compared the expression levels of genes encoding predicted KEGG-annotated proteins by parametric analysis of variance (ANOVA) using the Benjamini-Hochberg correction and visualized the results as volcano plots in XLSTAT (121).

Potential interactions among KEGG-annotated gene expression of *Cardinium* and mite genes were analyzed by two types of correlation analyses, dbRDA and Spearman correlation analysis. In dbRDA, mite expression levels were set as "species," *Cardinium* gene expression levels were set as "environmental variables" (122), and the Bray-Curtis dissimilarity distance matrix was used. To identify variables with the strongest influence on the dbRDA model, we applied a variable selection procedure ("both") using the ordistep function (123) to reduce a full data set to a subset of predictor environmental variables. Then, a new model based on a reduced set of "environmental" variables was constructed and analyzed as described above. dbRDA models were generated separately for mite whole-transcriptome data sets and metabolic pathways and for regulatory and immune pathways associated with the host reaction to infection or symbiotic bacteria (apoptosis, endocytosis, lysozyme, MAPK signaling cascade, TGF-β pathway, TNF pathway, and Toll/Imd pathway). Models with the best predictive power were selected based on their explained variability (R^2). Spearman correlations using bootstrap permutation were calculated in PAST. Correlation values were filtered based on their significance level ($P \leq 0.05$) in R and were visualized in Cytoscape v3.8.0 (124) using a correlation-based network in Metscape 3 (125). Only edges with an absolute R^2 of $\geq$0.75 were included in the initial analysis. To further reduce the numbers of edges, only edges with "strength correlations" of $\geq$8 ($R^2 \geq 0.86$) were considered. The final network was organized in the Compound Spring Embedder (CoSE) layout (126).

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

Heat maps were constructed from the standardized abundance of predicted proteins or Spearman correlation coefficients. Data sets were clustered by Ward's method using the Euclidean distance in PAST.

Data availability. Raw reads were deposited into NCBI's Sequence Read Archive (SRA) database under accession numbers SRR11058239 to SRR11058250, associated with BioProject PRJNA605718.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

DATA SET S1, XLSX file, 19 MB.

FIG S1, PDF file, 0.7 MB.

FIG S2, PDF file, 0.6 MB.

FIG S3, PDF file, 0.7 MB.

FIG S4, PDF file, 0.7 MB.

FIG S5, PDF file, 0.5 MB.

FIG S6, PDF file, 0.7 MB.

FIG S7, PDF file, 0.8 MB.

FIG S8, PDF file, 0.6 MB.

FIG S9, PDF file, 0.5 MB.

ACKNOWLEDGMENTS

We are obligated to anonymous referees for suitable comments on a previous version of this article and to Martin Markovic for providing excellent assistance.

J.H., M.N., B.S., and T.E. were supported by a project from the Czech Science Foundation (GACR no. 19-09998S) and by the Ministry of Agriculture of the Czech Republic (project no. RO0418); P.B.K. was supported by a grant from the Russian Science Foundation (project no. 19-14-00004).

M.N. performed sampling of mites, RNA extraction, and cleanup, S.E.D. performed RNA sequencing, J.H. performed bioinformatics analysis, E.D.S., P.B.K., S.E.D., B.S., T.E., and J.H. contributed to scientific writing, and J.H. wrote the main manuscript.

REFERENCES

1. Ju J-F, Bing X-L, Zhao D-S, Guo Y, Xi Z, Hoffmann AA, Zhang K-J, Huang H-J, Gong J-T, Xu Zhang X, Hong X-Y. 2020. *Wolbachia* supplement biotin and riboflavin to enhance reproduction in planthoppers. ISME J 14:676–687. https://doi.org/10.1038/s41396-019-0559-9.
2. Ren F-R, Sun X, Wang T-Y, Yao Y-L, Huang Y-Z, Zhang X, Luan J-B. 2020. Biotin provisioning by horizontally transferred genes from bacteria confers animal fitness benefits. ISME J 14:2542–2553. https://doi.org/10.1038/s41396-020-0704-5.
3. Husnik F, Hypsa V, Darby A. 2020. Insect-symbiont gene expression in the midgut bacteriocytes of a blood-sucking parasite. Genome Biol Evol 12:429–442. https://doi.org/10.1093/gbe/evaa032.
4. Zhu Y-X, Song Z-R, Song Y-L, Hong X-Y. 2020. Double infection of *Wolbachia* and *Spiroplasma* alters induced plant defense and spider mite fecundity. Pest Manag Sci 76:3273–3281. https://doi.org/10.1002/ps.5886.
5. Brown AMV, Wasala SK, Howe DK, Peetz AB, Zasada IA, Denver DR. 2018. Comparative genomics of *Wolbachia-Cardinium* dual endosymbiosis in a plant-parasitic nematode. Front Microbiol 9:2482. https://doi.org/10.3389/fmicb.2018.02482.
6. Siozios S, Pilgrim J, Darby AC, Baylis M, Hurst GDD. 2019. The draft genome of strain *c*Cpun from biting midges confirms insect *Cardinium* are not a monophyletic group and reveals a novel gene family expansion in a symbiont. PeerJ 7:e6448. https://doi.org/10.7717/peerj.6448.
7. Morag N, Klement E, Saroya Y, Lensky I, Gottlieb Y. 2012. Prevalence of the symbiont *Cardinium* in Culicoides (Diptera: Ceratopogonidae) vector species is associated with land surface temperature. FASEB J 26:4025–4034. https://doi.org/10.1096/fj.12-210419.
8. Chang J, Masters A, Avery A, Werren JH. 2010. A divergent *Cardinium* found in daddy long-legs (Arachnida: Opiliones). J Invertebr Pathol 105:220–227. https://doi.org/10.1016/j.jip.2010.05.017.
9. Duron O, Hurst GDD, Hornett EA, Josling JA, Engelstadter J. 2008. High incidence of the maternally inherited bacterium *Cardinium* in spiders. Mol Ecol 17:1427–1437. https://doi.org/10.1111/j.1365-294X.2008.03689.x.
10. Chigira A, Miura K. 2005. Detection of '*Candidatus* Cardinium' bacteria from the haploid host *Brevipalpus californicus* (Acari: Tenuipalpidae) and effect on the host. Exp Appl Acarol 37:107–116. https://doi.org/10.1007/s10493-005-0592-4.
11. Penz T, Schmitz-Esser S, Kelly SE, Cass BN, Muller A, Woyke T, Malfatti SA, Hunter MS, Horn M. 2012. Comparative genomics suggests an independent origin of cytoplasmic incompatibility in *Cardinium hertigii*. PLoS Genet 8:e1003012. https://doi.org/10.1371/journal.pgen.1003012.
12. Colloff MJ. 2009. Dust mites. CSIRO Publishing, Collingwood, VIC, Australia.
13. Spieksma FTM. 1997. Domestic mites from an acarologic perspective. Allergy 52:360–368. https://doi.org/10.1111/j.1398-9995.1997.tb01012.x.
14. OConnor BM. 1979. Evolutionary origins of astigmatid mites inhabiting stored products, p 273–278. *In* Rodriguez GJ (ed), Recent advances in acarology, vol 1. Academic Press, New York, NY. https://doi.org/10.1016/b978-0-12-592201-2.50038-5.
15. WHO/IUIS Allergen Nomenclature Sub-Committee. 2020. Allergen nomenclature. Astigmata. WHO/IUIS Allergen Nomenclature Sub-Committee. http://www.allergen.org/search.php?TaxOrder=Astigmata.
16. Kim JY, Yi M-H, Hwang Y, Lee JY, Lee I-Y, Yong D, Yong T-S. 2018. 16S rRNA profiling of the *Dermatophagoides farinae* core microbiome: *Enterococcus* and *Bartonella*. Clin Exp Allergy 48:607–610. https://doi.org/10.1111/cea.13104.
17. Lee J, Kim JY, Yi M-h, Hwang Y, Lee I-Y, Nam S-H, Yong D, Yong T-S. 2019. Comparative microbiome analysis of *Dermatophagoides farinae*, *Dermatophagoides pteronyssinus*, and *Tyrophagus putrescentiae*. J Allergy Clin Immunol 143:1620–1623. https://doi.org/10.1016/j.jaci.2018.10.062.
18. Valerio CR, Murray P, Arlian LG, Slater JE. 2005. Bacterial 16S ribosomal DNA in house dust mite cultures. J Allergy Clin Immunol 116:1296–1300. https://doi.org/10.1016/j.jaci.2005.09.046.
19. Hubert J, Nesvorna M, Klimov P, Dowd SE, Sopko B, Erban T. 2019. Differential allergen expression in three *Tyrophagus putrescentiae* strains inhabited by distinct microbiome. Allergy 74:2502–2507. https://doi.org/10.1111/all.13921.
20. Dzoro S, Mittermann I, Resch-Marat Y, Vrtala S, Nehr M, Hirschl AM, Wikberg G, Lundeberg L, Johansson C, Scheynius A, Valenta R. 2018.

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

House dust mites as potential carriers for IgE sensitization to bacterial antigens. Allergy 73:115–124. https://doi.org/10.1111/all.13260.
21. Erban T, Hubert J. 2008. Digestive function of lysozyme in synanthropic acaridid mites enables utilization of bacteria as a food source. Exp Appl Acarol 44:199–212. https://doi.org/10.1007/s10493-008-9138-x.
22. Hubert J, Nesvorna M, Kopecky J, Erban T, Klimov P. 2019. Population and culture age influence the microbiome profiles of house dust mites. Microb Ecol 77:1048–1066. https://doi.org/10.1007/s00248-018-1294-x.
23. Klimov P, Molva V, Nesvorna M, Pekar S, Shcherbachenko E, Erban T, Hubert J. 2019. Dynamics of the microbial community during growth of the house dust mite *Dermatophagoides farinae* in culture. FEMS Microbiol Ecol 95:fiz153. https://doi.org/10.1093/femsec/fiz153.
24. Nakamura Y, Kawai S, Yukuhiro F, Ito S, Gotoh T, Kisimoto R, Yanase T, Matsumoto Y, Kageyama D, Noda H. 2009. Prevalence of *Cardinium* bacteria in planthoppers and spider mites and taxonomic revision of "*Candidatus* Cardinium hertigii" based on detection of a new *Cardinium* group from biting midges. Appl Environ Microbiol 75:6757–6763. https://doi.org/10.1128/AEM.01583-09.
25. Erban T, Klimov P, Molva V, Hubert J. 2020. Whole genomic sequencing and sex-dependent abundance estimation of *Cardinium* sp., a common and hyperabundant bacterial endosymbiont of the American house dust mite, *Dermatophagoides farinae*. Exp Appl Acarol 80:363–380. https://doi.org/10.1007/s10493-020-00475-5.
26. Kitajima EW, Groot TVM, Novelli VM, Freitas-Astua J, Alberti G, de Moraes GJ. 2007. *In situ* observation of the *Cardinium* symbionts of *Brevipalpus* (Acari: Tenuipalpidae) by electron microscopy. Exp Appl Acarol 42: 263–271. https://doi.org/10.1007/s10493-007-9090-1.
27. Hubert J, Kopecky J, Perotti MA, Nesvorna M, Braig HR, Sagova-Mareckova M, Macovei L, Zurek L. 2012. Detection and identification of species-specific bacteria associated with synanthropic mites. Microb Ecol 63:919–928. https://doi.org/10.1007/s00248-011-9969-6.
28. Trivedi B, Valerio C, Slater JE. 2003. Endotoxin content of standardized allergen vaccines. J Allergy Clin Immunol 111:777–783. https://doi.org/10.1067/mai.2003.1338.
29. Valerio C, Murray P, Arlian LG, Slater JE. 2004. Endotoxin in dust mite allergen extracts. J Allergy Clin Immunol 113:S136–S137. https://doi.org/10.1016/j.jaci.2003.12.489.
30. Valerio C, Arlian LG, Slater JE. 2009. Bacterial DNA sequences isolated from standardized dust mite extracts and wild mites. J Allergy Clin Immunol 123:S216–S216. https://doi.org/10.1016/j.jaci.2008.12.826.
31. Kim JY, Yi M, Lee J, Lee I, Yong T. 2019. Potential immunomodulatory effect of the house dust mite microbiome. Allergy 74:LBTP1825.
32. Smith TE, Moran NA. 2020. Coordination of host and symbiont gene expression reveals a metabolic tug-of-war between aphids and *Buchnera*. Proc Natl Acad Sci U S A 117:2113–2121. https://doi.org/10.1073/pnas.1916748117.
33. Mann E, Stouthamer CM, Kelly SE, Dzieciol M, Hunter MS, Schmitz-Esser S. 2017. Transcriptome sequencing reveals novel candidate genes for *Cardinium hertigii*-caused cytoplasmic incompatibility and host-cell interaction. mSystems 2:e00141-17. https://doi.org/10.1128/mSystems.00141-17.
34. Li C, He M, Yun Y, Peng Y. 2020. Co-infection with *Wolbachia* and *Cardinium* may promote the synthesis of fat and free amino acids in a small spider, *Hylyphantes graminicola*. J Invertebr Pathol 169:107307. https://doi.org/10.1016/j.jip.2019.107307.
35. Martinez-Giron R, Doganci L, Iraola V. 2009. Gregarines in *Dermatophagoides* spp. (Acari: Pyroglyphidae): light microscopy observation. J Med Entomol 46:367–368. https://doi.org/10.1603/033.046.0223.
36. Martinez-Giron R. 2018. Flagellated protozoa detected in *Dermatophagoides* by light microscopy. Allergol Immunopathol (Madr) 46:304–306. https://doi.org/10.1016/j.aller.2017.05.008.
37. Ribas A, Martinez-Giron R. 2006. Protozoal forms in house-dust mites and respiratory allergy. Allergy Asthma Proc 27:347–349. https://doi.org/10.2500/aap.2006.27.2878.
38. Larsson JIR, Steiner MY, Bjornson S. 1997. *Intexta acarivora* gen. et sp. n. (Microspora: Chytridiopsidae)—ultrastructural study and description of a new microsporidian parasite of the forage mite *Tyrophagus putrescentiae* (Acari: Acaridae). Acta Protozool 36:295–304.
39. Andersen A. 1991. Nutritional value of yeast for *Dermatophagoides pteronyssinus* (Acari: Epidermoptidae) and the antigenic and allergenic composition of extracts during extended culturing. J Med Entomol 28: 487–491. https://doi.org/10.1093/jmedent/28.4.487.
40. Eraso E, Martinez J, Garcia-Ortega P, Martinez A, Palacios R, Cisterna R, Guisantes JA. 1998. Influence of mite growth culture phases on the biological standardization of allergenic extracts. J Invest Allergol Clin Immunol 8:201–206.
41. Eraso E, Martinez J, Martinez A, Palacios R, Guisantes JA. 1997. Quality parameters for the production of mite extracts. Allergol Immunopathol (Madr) 25:113–117.
42. Eraso E, Guisantes JA, Martinez J, Saenz-de-Santamaria M, Martinez A, Palacios R, Cisterna R. 1997. Kinetics of allergen expression in cultures of house dust mites, *Dermatophagoides pteronyssinus* and *D. farinae* (Acari: Pyroglyphidae). J Med Entomol 34:684–689. https://doi.org/10.1093/jmedent/34.6.684.
43. Kremer N, Charif D, Henri H, Gavory F, Wincker P, Mavingui P, Vavre F. 2012. Influence of *Wolbachia* on host gene expression in an obligatory symbiosis. BMC Microbiol 12:S7. https://doi.org/10.1186/1471-2180-12-S1-S7.
44. Herbert RI, McGraw EA. 2018. The nature of the immune response in novel *Wolbachia*-host associations. Symbiosis 74:225–236. https://doi.org/10.1007/s13199-017-0503-6.
45. Pan X, Pike A, Joshi D, Bian G, McFadden MJ, Lu P, Liang X, Zhang F, Raikhel AS, Xi Z. 2018. The bacterium *Wolbachia* exploits host innate immunity to establish a symbiotic relationship with the dengue vector mosquito *Aedes aegypti*. ISME J 12:277–288. https://doi.org/10.1038/ismej.2017.174.
46. Surachetpong W, Singh N, Cheung KW, Luckhart S. 2009. MAPK ERK signaling regulates the TGF-beta1-dependent mosquito response to *Plasmodium falciparum*. PLoS Pathog 5:e1000366. https://doi.org/10.1371/journal.ppat.1000366.
47. Patrnogic J, Heryanto C, Eleftherianos I. 2018. Transcriptional up-regulation of the TGF-beta intracellular signaling transducer Mad of *Drosophila* larvae in response to parasitic nematode infection. Innate Immun 24: 349–356. https://doi.org/10.1177/1753425918790663.
48. Kawasaki T, Kawai T. 2014. Toll-like receptor signaling pathways. Front Immunol 5:461. https://doi.org/10.3389/fimmu.2014.00461.
49. Varfolomeev EE, Ashkenazi A. 2004. Tumor necrosis factor: an apoptosis JuNKie? Cell 116:491–497. https://doi.org/10.1016/S0092-8674(04)00166-7.
50. Waldron R, McGowan J, Gordon N, McCarthy C, Mitchell EB, Doyle S, Fitzpatrick DA. 2017. Draft genome sequence of *Dermatophagoides pteronyssinus*, the European house dust mite. Genome Announc 5:e00789-17. https://doi.org/10.1128/genomeA.00789-17.
51. Salzet M. 2001. Vertebrate innate immunity resembles a mosaic of invertebrate immune responses. Trends Immunol 22:285–288. https://doi.org/10.1016/S1471-4906(01)01895-6.
52. Imler J-L, Zheng L. 2004. Biology of Toll receptors: lessons from insects and mammals. J Leukoc Biol 75:18–26. https://doi.org/10.1189/jlb.0403160.
53. Zheng L, Zhang L, Lin H, McIntosh MT, Malacrida A. 2005. Toll-like receptors in invertebrate innate immunity. Invertebr Surviv J 2:105–113.
54. Coscia MR, Giacomelli S, Oreste U. 2011. Toll-like receptors: an overview from invertebrates to vertebrates. Invertebr Surviv J 8:210–226.
55. Zychlinsky A, Sansonetti P. 1997. Perspectives series: host/pathogen interactions. Apoptosis in bacterial pathogenesis. J Clin Invest 100: 493–495. https://doi.org/10.1172/JCI119557.
56. Zhukova MV, Kiseleva E. 2012. The virulent *Wolbachia* strain wMelPop increases the frequency of apoptosis in the female germline cells of *Drosophila melanogaster*. BMC Microbiol 12:S15. https://doi.org/10.1186/1471-2180-12-S1-S15.
57. Guo Y, Hoffmann AA, Xu X-Q, Zhang X, Huang H-J, Ju J-F, Gong J-T, Hong X-Y. 2018. *Wolbachia*-induced apoptosis associated with increased fecundity in *Laodelphax striatellus* (Hemiptera: Delphacidae). Insect Mol Biol 27:796–807. https://doi.org/10.1111/imb.12518.
58. Almeida F, Suesdek L. 2017. Effects of *Wolbachia* on ovarian apoptosis in *Culex quinquefasciatus* (Say, 1823) during the previtellogenic and vitellogenic periods. Parasit Vectors 10:398. https://doi.org/10.1186/s13071-017-2332-0.
59. Lu H-I, Chang C-c, Wilson ACC. 2016. Amino acid transporters implicated in endocytosis of *Buchnera* during symbiont transmission in the pea aphid. Evodevo 7:24. https://doi.org/10.1186/s13227-016-0061-7.
60. Koga R, Meng X-Y, Tsuchida T, Fukatsu T. 2012. Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. Proc Natl Acad Sci U S A 109:E1230–E1237. https://doi.org/10.1073/pnas.1119212109.
61. White PM, Pietri JE, Debec A, Russell S, Patel B, Sullivan W. 2017. Mechanisms of horizontal cell-to-cell transfer of *Wolbachia* spp. in *Drosophila melanogaster*. Appl Environ Microbiol 83:e03425-16. https://doi.org/10.1128/AEM.03425-16.
62. Fattouh N, Cazevieille C, Landmann F. 2019. *Wolbachia* endosymbionts subvert the endoplasmic reticulum to acquire host membranes without

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

triggering ER stress. PLoS Negl Trop Dis 13:e0007218. https://doi.org/10.1371/journal.pntd.0007218.
63. Gerardo NM, Hoang KL, Stoy KS. 2020. Evolution of animal immunity in the light of beneficial symbioses. Philos Trans R Soc Lond B Biol Sci 375:20190601. https://doi.org/10.1098/rstb.2019.0601.
64. Chang Z. 2016. The function of the DegP (HtrA) protein: protease versus chaperone. IUBMB Life 68:904–907. https://doi.org/10.1002/iub.1561.
65. Kuykendall LD, Shao JY, Hartung JS. 2012. 'Ca. Liberibacter asiaticus' proteins orthologous with *pSymA*-encoded proteins of *Sinorhizobium meliloti*: hypothetical roles in plant host interaction. PLoS One 7:e38725. https://doi.org/10.1371/journal.pone.0038725.
66. Chi X, Fan Q, Zhang Y, Liang K, Wan L, Zhou Q, Li Y. 2020. Structural mechanism of phospholipids translocation by MlaFEDB complex. Cell Res 30:1127–1135. https://doi.org/10.1038/s41422-020-00404-6.
67. Doyle MT, Bernstein HD. 2019. Bacterial outer membrane proteins assemble via asymmetric interactions with the BamA β-barrel. Nat Commun 10:3358. https://doi.org/10.1038/s41467-019-11230-9.
68. Bowyer A, Baardsnes J, Ajamian E, Zhang L, Cygler M. 2011. Characterization of interactions between LPS transport proteins of the Lpt system. Biochem Biophys Res Commun 404:1093–1098. https://doi.org/10.1016/j.bbrc.2010.12.121.
69. Ayala-Castro C, Saini A, Outten FW. 2008. Fe-S cluster assembly pathways in bacteria. Microbiol Mol Biol Rev 72:110–125. https://doi.org/10.1128/MMBR.00034-07.
70. Backert S, Meyer TF. 2006. Type IV secretion systems and their effectors in bacterial pathogenesis. Curr Opin Microbiol 9:207–217. https://doi.org/10.1016/j.mib.2006.02.008.
71. Best AM, Abu Kwaik Y. 2019. Evasion of phagotrophic predation by protist hosts and innate immunity of metazoan hosts by *Legionella pneumophila*. Cell Microbiol 21:e12971. https://doi.org/10.1111/cmi.12971.
72. Flieger A, Frischknecht F, Hacker G, Hornef MW, Pradel G. 2018. Pathways of host cell exit by intracellular pathogens. Microb Cell 5:525–544. https://doi.org/10.15698/mic2018.12.659.
73. Hunt TP, Magasanik B. 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc Natl Acad Sci U S A 82:8453–8457. https://doi.org/10.1073/pnas.82.24.8453.
74. Schmitz-Esser S, Tischler P, Arnold R, Montanaro J, Wagner M, Rattei T, Horn M. 2010. The genome of the amoeba symbiont "*Candidatus* Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. J Bacteriol 192:1045–1057. https://doi.org/10.1128/JB.01379-09.
75. Rosas-Perez T, Rosenblueth M, Rincon-Rosales R, Mora J, Martinez-Romero E. 2014. Genome sequence of "*Candidatus* Walczuchella monophlebidarum" the flavobacterial endosymbiont of *Llaveia axin axin* (Hemiptera: Coccoidea: Monophlebidae). Genome Biol Evol 6:714–726. https://doi.org/10.1093/gbe/evu049.
76. Kambhampati S, Alleman A, Park Y. 2013. Complete genome sequence of the endosymbiont *Blattabacterium* from the cockroach *Nauphoeta cinerea* (Blattodea: Blaberidae.) Genomics 102:479–483. https://doi.org/10.1016/j.ygeno.2013.09.003.
77. Toh H, Weiss BL, Perkin SAH, Yamashita A, Oshima K, Hattori M, Aksoy S. 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Res 16:149–156. https://doi.org/10.1101/gr.4106106.
78. Showmaker KC, Walden KKO, Fields CJ, Lambert KN, Hudson ME. 2018. Genome sequence of the soybean cyst nematode (*Heterodera glycines*) endosymbiont "*Candidatus* Cardinium hertigii" strain cHgTN10. Genome Announc 6:e00624-18. https://doi.org/10.1128/genomeA.00624-18.
79. Zeng Z, Fu Y, Guo D, Wu Y, Ajayi OE, Wu Q. 2018. Bacterial endosymbiont *Cardinium* cSfur genome sequence provides insights for understanding the symbiotic relationship in *Sogatella furcifera* host. BMC Genomics 19:688. https://doi.org/10.1186/s12864-018-5078-y.
80. Fronzes R, Christie PJ, Waksman G. 2009. The structural biology of type IV secretion systems. Nat Rev Microbiol 7:703–714. https://doi.org/10.1038/nrmicro2218.
81. Heddi A, Grenier A-M, Khatchadourian C, Charles H, Nardon P. 1999. Four intracellular genomes direct weevil biology: nuclear, mitochondrial, principal endosymbiont, and *Wolbachia*. Proc Natl Acad Sci U S A 96:6814–6819. https://doi.org/10.1073/pnas.96.12.6814.
82. Heddi A, Lefebvre F, Nardon P. 1993. Effect of endocytobiotic bacteria on mitochondrial enzymatic activities in the weevil *Sitophilus oryzae* (Coleoptera: Curculionidae). Insect Biochem Mol Biol 23:403–411. https://doi.org/10.1016/0965-1748(93)90024-M.
83. Douglas AE. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. Annu Rev Entomol 43:17–37. https://doi.org/10.1146/annurev.ento.43.1.17.
84. Dale C, Moran NA. 2006. Molecular interactions between bacterial symbionts and their hosts. Cell 126:453–465. https://doi.org/10.1016/j.cell.2006.07.014.
85. Blattner FR, Plunkett G, III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1462. https://doi.org/10.1126/science.277.5331.1453.
86. Ankrah NYD, Chouaia B, Douglas AE. 2018. The cost of metabolic interactions in symbioses between insects and bacteria with reduced genomes. mBio 9:e01433-18. https://doi.org/10.1128/mBio.01433-18.
87. McCutcheon JP, Moran NA. 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. Genome Biol Evol 2:708–718. https://doi.org/10.1093/gbe/evq055.
88. Gupta A, Nair S. 2020. Dynamics of insect-microbiome interaction influence host and microbial symbiont. Front Microbiol 11:1357. https://doi.org/10.3389/fmicb.2020.01357.
89. Santos-Garcia D, Rollat-Farnier P-A, Beitia F, Zchori-Fein E, Vavre F, Mouton L, Moya A, Latorre A, Silva FJ. 2014. The genome of *Cardinium* cBtQ1 provides insights into genome reduction, symbiont motility, and its settlement in *Bemisia tabaci*. Genome Biol Evol 6:1013–1030. https://doi.org/10.1093/gbe/evu077.
90. Wada-Katsumata A, Zurek L, Nalyanya G, Roelofs WL, Zhang A, Schal C. 2015. Gut bacteria mediate aggregation in the German cockroach. Proc Natl Acad Sci U S A 112:15678–15683. https://doi.org/10.1073/pnas.1504031112.
91. Van Arnam EB, Currie CR, Clardy J. 2018. Defense contracts: molecular protection in insect-microbe symbioses. Chem Soc Rev 47:1638–1651. https://doi.org/10.1039/c7cs00340d.
92. Becerra JX, Venable GX, Saeidi V. 2015. *Wolbachia*-free heteropterans do not produce defensive chemicals or alarm pheromones. J Chem Ecol 41:593–601. https://doi.org/10.1007/s10886-015-0596-4.
93. Brody AR, Wharton GW. 1970. *Dermatophagoides farinae*: ultrastructure of lateral opisthosomal dermal glands. Trans Am Microsc Soc 89:499–513. https://doi.org/10.2307/3224560.
94. Raspotnig G. 2010. Oil gland secretions in Oribatida (Acari), p 235–239. *In* Sabelis M, Bruin J (ed), Trends in acarology. Springer, Dordrecht, Netherlands. https://doi.org/10.1007/978-90-481-9837-5_38.
95. Vandermoten S, Mescher MC, Francis F, Haubruge E, Verheggen FJ. 2012. Aphid alarm pheromone: an overview of current knowledge on biosynthesis and functions. Insect Biochem Mol Biol 42:155–163. https://doi.org/10.1016/j.ibmb.2011.11.008.
96. Hojo M, Matsumoto T, Miura T. 2007. Cloning and expression of a geranylgeranyl diphosphate synthase gene: insights into the synthesis of termite defence secretion. Insect Mol Biol 16:121–131. https://doi.org/10.1111/j.1365-2583.2007.00709.x.
97. Steidle JLM, Barcari E, Hradecky M, Trefz S, Tolasch T, Gantert C, Schulz S. 2014. Pheromonal communication in the European house dust mite, *Dermatophagoides pteronyssinus*. Insects 5:639–650. https://doi.org/10.3390/insects5030639.
98. Tatami K, Mori N, Nishida R, Kuwahara Y. 2001. 2-Hydroxy-6-methylbenzaldehyde: the female sex pheromone of the house dust mite *Dermatophagoides farinae* (Astigmata: Pyroglyphidae). Med Entomol Zool 52:279–286. https://doi.org/10.7601/mez.52.279.
99. Kuwahara Y, Leal WS, Suzuki T. 1990. Pheromone study on astigmatid mites XXVI. Comparison of volatile components between *Dermatophagoides farinae* and *D. pteronyssinus* (Astigmata, Pyroglyphidae). Med Entomol Zool 41:23–28. https://doi.org/10.7601/mez.41.23.
100. Konakandla B, Park Y, Margolies D. 2006. Whole genome amplification of Chelex-extracted DNA from a single mite: a method for studying genetics of the predatory mite *Phytoseiulus persimilis*. Exp Appl Acarol 40:241–247. https://doi.org/10.1007/s10493-006-9042-1.
101. Hubert J, Erban T, Kopecky J, Sopko B, Nesvorna M, Lichovnikova M, Schicht S, Strube C, Sparagano O. 2017. Comparison of microbiomes between red poultry mite populations (*Dermanyssus gallinae*): predominance of *Bartonella*-like bacteria. Microb Ecol 74:947–960. https://doi.org/10.1007/s00248-017-0993-z.
102. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
103. Cuccuru G, Orsini M, Pinna A, Sbardellati A, Soranzo N, Travaglione A, Uva P, Zanetti G, Fotia G. 2014. Orione, a web-based framework for NGS

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.

analysis in microbiology. Bioinformatics 30:1928–1929. https://doi.org/10.1093/bioinformatics/btu135.
104. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Gruning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544. https://doi.org/10.1093/nar/gky379.
105. Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, Gibbs R, Beeman RW, Brown SJ, Bucher G, Friedrich M, Grimmelikhuijzen CJ, Klingler M, Lorenzen M, Richards S, Roth S, Schroder R, Tautz D, Zdobnov EM, Muzny D, Gibbs RA, Weinstock GM, Attaway T, Bell S, Buhay CJ, Chandrabose MN, Chavez D, Clerk-Blankenburg KP, Cree A, Dao M, Davis C, Chacko J, Dinh H, Dugan-Rocha S, Fowler G, Garner TT, Garnes J, Gnirke A, Hawes A, Hernandez J, Hines S, Holder M, Hume J, Jhangiani SN, Joshi V, Khan ZM, Jackson L, Kovar C, Kowis A, Lee S, Tribolium Genome Sequencing Consortium, et al. 2008. The genome of the model beetle and pest *Tribolium castaneum*. Nature 452:949–955. https://doi.org/10.1038/nature06784.
106. Rider SD, Jr, Morgan MS, Arlian LG. 2015. Draft genome of the scabies mite. Parasit Vectors 8:585. https://doi.org/10.1186/s13071-015-1198-2.
107. Grbic M, Van Leeuwen T, Clark RM, Rombauts S, Rouze P, Grbic V, Osborne EJ, Dermauw W, Ngoc PC, Ortego F, Hernandez-Crespo P, Diaz I, Martinez M, Navajas M, Sucena E, Magalhaes S, Nagy L, Pace RM, Djuranovic S, Smagghe G, Iga M, Christiaens O, Veenstra JA, Ewer J, Villalobos RM, Hutter JL, Hudson SD, Velez M, Yi SV, Zeng J, Pires-daSilva A, Roch F, Cazaux M, Navarro M, Zhurov V, Acevedo G, Bjelica A, Fawcett JA, Bonnet E, Martens C, Baele G, Wissler L, Sanchez-Rodriguez A, Tirry L, Blais C, Demeestere K, Henz SR, Gregory TR, Mathieu J, Verdon L, Farinelli L, Schmutz J, Lindquist E, et al. 2011. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature 479:487–492. https://doi.org/10.1038/nature10640.
108. Zajc J, Liu Y, Dai W, Yang Z, Hu J, Gostincar C, Gunde-Cimerman N. 2013. Genome and transcriptome sequencing of the halophilic fungus *Wallemia ichthyophaga*: haloadaptations present and absent. BMC Genomics 14:617. https://doi.org/10.1186/1471-2164-14-617.
109. Muzzey D, Schwartz K, Weissman JS, Sherlock G. 2013. Assembly of a phased diploid *Candida albicans* genome facilitates allele-specific measurements and provides a simple model for repeat and indel structure. Genome Biol 14:R97. https://doi.org/10.1186/gb-2013-14-9-r97.
110. Stromsten NJ, Benson SD, Burnett RM, Bamford DH, Bamford JKH. 2003. The *Bacillus thuringiensis* linear double-stranded DNA phage Bam35, which is highly similar to the *Bacillus cereus* linear plasmid pBClin15, has a prophage state. J Bacteriol 185:6985–6989. https://doi.org/10.1128/JB.185.23.6985-6989.2003.
111. Segers FHID, Kesnerova L, Kosoy M, Engel P. 2017. Genomic changes associated with the evolutionary transition of an insect gut symbiont into a blood-borne pathogen. ISME J 11:1232–1244. https://doi.org/10.1038/ismej.2016.201.
112. Kosoy M, Morway C, Sheff KW, Bai Y, Colborn J, Chalcraft L, Dowell SF, Peruski LF, Maloney SA, Baggett H, Sutthirattana S, Sidhirat A, Maruyama S, Kabeya H, Chomel BB, Kasten R, Popov V, Robinson J, Kruglov A, Petersen LR. 2008. *Bartonella tamiae* sp. nov., a newly recognized pathogen isolated from three human patients from Thailand. J Clin Microbiol 46:772–775. https://doi.org/10.1128/JCM.02120-07.
113. Tokuda G, Elbourne LDH, Kinjo Y, Saitoh S, Sabree Z, Hojo M, Yamada A, Hayashi Y, Shigenobu S, Bandi C, Paulsen IT, Watanabe H, Lo N. 2013. Maintenance of essential amino acid synthesis pathways in the *Blattabacterium cuenoti* symbiont of a wood-feeding cockroach. Biol Lett 9:20121153. https://doi.org/10.1098/rsbl.2012.1153.
114. Koga R, Moran NA. 2014. Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. ISME J 8:1237–1246. https://doi.org/10.1038/ismej.2013.235.
115. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37. https://doi.org/10.1093/nar/gkr367.
116. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. Nucleic Acids Res 46:W200–W204. https://doi.org/10.1093/nar/gky448.
117. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27.
118. Hammer O, Harper DAT, Ryan PD. 2001. PAST: paleontological statistics software package for education and data analysis. Palaeontol Electron 4:4. https://palaeo-electronica.org/2001_1/past/issue1_01.htm.
119. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79:5112–5120. https://doi.org/10.1128/AEM.01043-13.
120. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MH, Szoecs E, Wagner H. 2019. Package 'vegan': community ecology package, version 2.5–6. CRAN—The Comprehensive R Archive Network. https://cran.r-project.org/web/packages/vegan/vegan.pdf.
121. Addinsoft. 2020. XLSTAT. Addinsoft, New York, NY. https://www.xlstat.com.
122. jokergoo. 2020. Visualize big correlation matrix. A Bioinformagician. http://web.archive.org/web/20200424071358/http://jokergoo.github.io/blog/html/large_matrix_circular.html.
123. Zeleny D. 2017. Analysis of community ecology data in R: constrained ordination. https://www.davidzeleny.net/anadat-r/doku.php/en:forward_sel.
124. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303.
125. Basu S, Duren W, Evans CR, Burant CF, Michailidis G, Karnovsky A. 2017. Sparse network modeling and Metscape-based visualization methods for the analysis of large-scale metabolomics data. Bioinformatics 33:1545–1553. https://doi.org/10.1093/bioinformatics/btx012.
126. Genc B, Dogrusoz U. 2016. An algorithm for automated layout of process description maps drawn in SBGN. Bioinformatics 32:77–84. https://doi.org/10.1093/bioinformatics/btv516.

Downloaded from https://journals.asm.org/journal/msystems on 09 November 2021 by 94.38.94.90.