**Systems analysis to understand parasitic manipulation of host behavior: a meta-analysis**

As mentioned above, over the past couple of years we have identified several genes of interest in *C. floridanus* ants: (1) *Rhythmic genes*: several thousands of genes that show 24h rhythms in the ant brain (REF), (2) *Plasticity genes*: genes that underlie behavioral plasticity in healthy ants (REF), and (3) *Manipulation genes*: ant genes that show differential expression during *Ophiocordyceps*-induced behavior manipulation (REF). To understand how *Ophiocordyceps* can induce changes to ant’s rhythmic behavior, we can gain by integrating the different transcriptomic datasets to form novel data-driven hypothesis. In this section, we illustrate the usefulness of using network tools and published datasets to better understand the mechanisms underlying parasitic manipulation of host behavior.

HOW AND WHY NETWORKS ARE USEFUL FOR STUDYING HOST-PARASITE INTERACTIONS?

DISCUSS THE USEFULNESS/POPULARITY OF WGCNA FOR PERFORMING NETWORK ANALYSIS.

The questions we set out to answer was: (1) Which regions of the host’s gene expression network is affected by the parasite during manipulation? (2) Do the manipulating parasite target regions in the host network that are under clock control or drive behavioral plasticity? The first step was to build the gene co-expression network (GCN) for *C. floridanus.* The expression of a gene fluctuates over a 24h day. Collecting gene expression at multiple time points throughout the day allows us to capture these daily fluctuations and better quantify the level of co-expression for gene pairs. The accuracy of the gene-gene similarity (correlation) is a necessary step for building a robust GCN and for downstream identification of modules or clusters in the network (REFS).

To construct the GCN of *C. floridanus*, we used time-course transcriptomics datasets of nurse and forager brains generated for a separate study in our lab (REF). In the study, brains of forager and nurse ants were sampled for RNA-Sequencing every 2h, over a 24h period, under 12h:12h light-dark conditions. First, we combined the twenty-four forager and nurse datasets to construct a generalized ant GCN and identified modules or clusters of highly co-expressed genes in the network using functions from the WGCNA package in R (REF) (Supp. File 1). Next, we annotated the ant GCN by identifying where our genes of interests are located in the network. To do so, we used Fisher’s exact test to identify the significant overlaps between genes in different modules and: 1) rhythmic genes, (2) plasticity genes, and (3) manipulation genes (Supp. File 1).

Of the 34 genes significantly higher expressed in forager brains (for-UP; Fig. 1), 20 were located in module-9 and the overlap was significant (odds-ratio = 67; p-value = 3e-23). For the 47 genes significantly higher expressed in nurse brains (or, lower expressed in foragers, for-DOWN; Fig. 1), 18 were located in module-6 and this overlap was also significant (odds-ratio = 34, p-value = 5e-18). No other modules in the ant GCN showed a significant overlap with the for-UP or for-DOWN genesets. The 209 highly co-expressed genes in module-9 is likely important for driving the physiological and behavioral state that characterize a forager ant. As such, we will refer to module-9 as the Forager-cluster, whereas module-6 will be referred to as the Nurse-cluster.

Next, using the same approach as above, we set out to identify the gene clusters that contain most of the manipulation genes. Will and colleagues found 232 genes to be significantly up-regulated in the forager head at manipulation as compared to healthy controls (ophio-UP; Fig. 1). Of these ophio-UP genes, 32 were located in the Forager-cluster and the overlap was found to be significant (odds-ratio = 9, p-value = 2e-15). As for the 574 genes down-regulated at manipulation (ophio-DOWN; Fig. 1), 42 were located in the Nurse-cluster and showed a significant overlap (odds-ratio = 19, p-value = 4e-13). In addition to the Nurse-cluster, the ophio-DOWN genes were also overrepresented in module-5. This makes sense when we focus on how the different modules are connected in the ant GCN; expression of the Nurse-cluster is positively correlated to the expression of module-5 (Fig. 1). Therefore, severe dysregulation of the Nurse-cluster can change expression of genes in the neighboring modules that are connected to it which includes cluster-5.

In summary, what we found was unexpected; the ant genes affected during *Ophiocordyceps*-induced behavioral manipulation are primarily located in the same two modules that contain most of the genes underlying ant behavioral plasticity (Fig. 1). In other words, the manipulating parasite seems to be targeting the same genes and processes that allow ants to display extreme behavioral plasticity. To get a glimpse of the functions of these genes, we identified the overrepresented GO terms in the Forager- and Nurse-cluster (Fig. 1, Supp. File 1). Genes in the Forager-cluster seem to be largely involved in energy production from ATP (ATP binding and ATP hydrolysis), whereas ones in the Nurse-cluster seem to be involved in metabolism (carbohydrate metabolic process) and catalytic activity (oxidoreductase activity, protein kinase activity) (Fig. 1).

Thus far, we have discussed the importance of host plasticity genes in parasite-induced behavioral manipulation. Now, we focus our attention on the links between the ant clock and behavioral plasticity, and how *Ophiocordyceps* might be affecting the host’s rhythmic gene expression to induce behavioral changes. Although the Forager- and Nurse-clusters in the ant GCN are not enriched for diurnal (24h) genes, they are directly connected to rhythmic modules (module-4 and module-12; Fig. 1). To alter or disrupt the processes that are under clock control, *Ophiocordyceps* could indirectly target these connected rhythmic modules by altering the expression of either Forager- or Nurse-cluster.

When we visualized the daily expression pattern of the Forager- and Nurse-cluster in the forager brain, a more direct link appeared. The daily expression of Forager- and Nurse-cluster show ultradian rhythms but only in forager brains; no such rhythms were present in nurse brains (Supp. Fig. 1). Using Fisher’s exact test, we confirmed that the Forager-cluster was indeed overrepresented in 12h-rhythmic genes (p-value = 2e-04) and the Nurse-cluster in 8h-rhythmic genes (p-value = 8e-17) but only in forager brains. In nurse brains, the clusters did not show enrichment for any rhythmic gene set. In conclusion, the presence or absence of daily rhythms in the Forager- and Nurse-cluster seems to be associated with two distinct behavioral states in ants. A hypothesis arises: distinct changes in the rhythmic properties – amplitude, phase or periodicity – of these two clusters can induce specific behavioral changes, even different behavioral states. To test the hypothesis, further studies to quantify the changes in rhythmic properties of the Forager- and Nurse-cluster during *Ophiocordyceps*-infectionwill be necessary.

DISCUSS/HIGHLIGHT SOME GENES

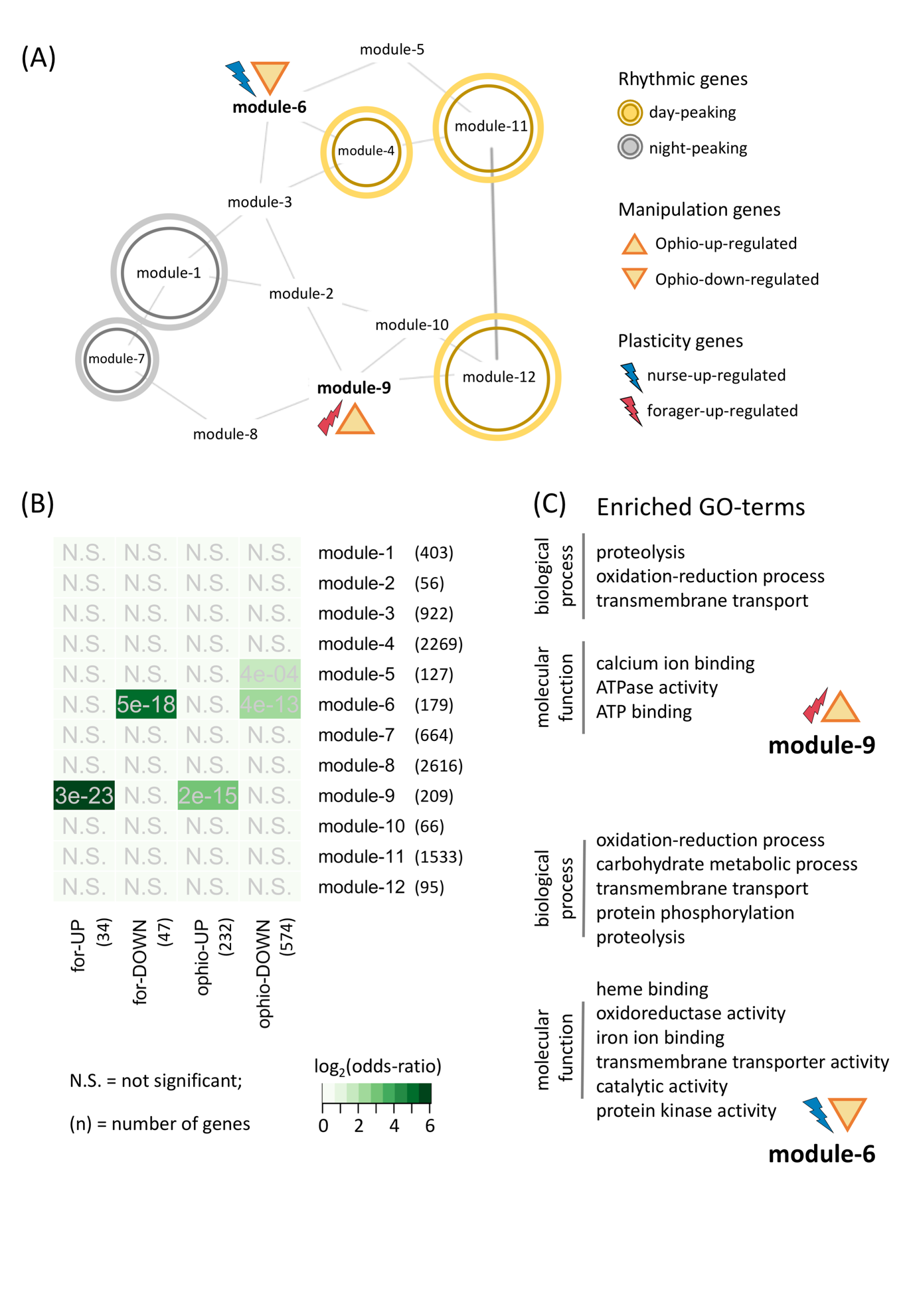
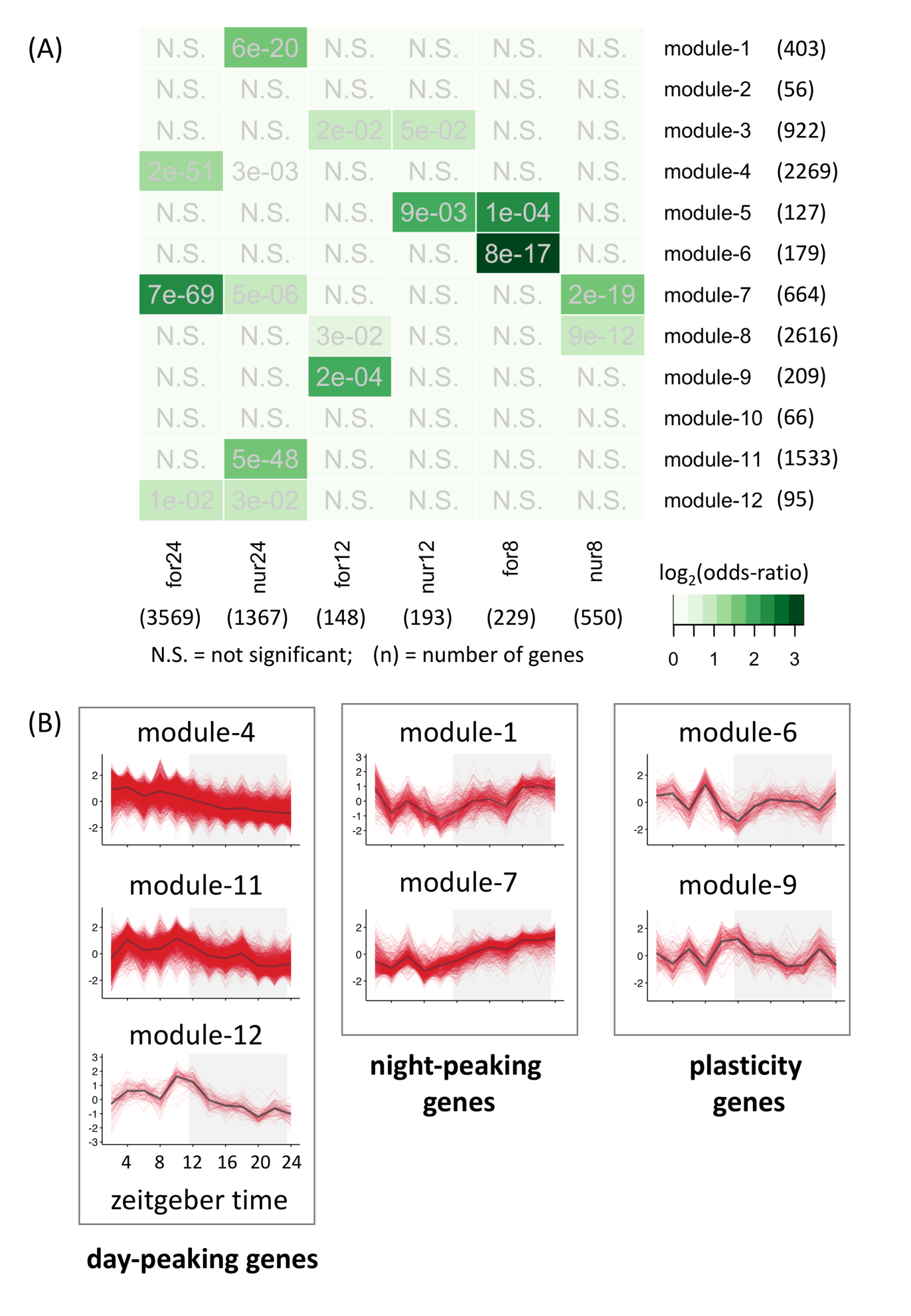


Figure 1: (A) The gene co-expression network (GCN) of ant brains; **Rhythmic genes** refer to genes oscillating every 24h in the ant brain with peaks during the day (lights-on) or night (lights-off), **Manipulation genes** are the ones that were differentially expressed in the ant head at *Ophiocordyceps* (Ophio) inducedmanipulated biting behavior, and **Plasticity genes** refer to the genes found to be differentially expressed between forager and nurse ant brains. (B) The heatmap shows the results from Fisher’s exact test performed to test if different sets of genes show significant overlap (modules in the GCN, manipulation genes, and plasticity genes). (C) The GO terms (biological processes and molecular functions) enriched in module-9 (Forager-cluster) and module-6 (Nurse-cluster).



Supplementary Figure 1: (A) Significant overlap between rhythmic genes and GCN modules. (B) Daily gene expression profiles of genes in different modules that are enriched for 24h-rhythmic genes (day-peaking and night-peaking) and plasticity genes (differentially expressed between forager and nurse ant brains).

Supplementary File 1: PDF/HTML file that contains all the code for creating and analyzing the ant GCN.

References

1. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.** *BMC Bioinformatics* 2008, **9:**559.

2. Langfelder P, Luo R, Oldham MC, Horvath S: **Is My Network Module Preserved and Reproducible?** *PLOS Computational Biology* 2011, **7:**e1001057.