Examination of the quantitative relationship between proteome- and transcriptome-level responses in larval guts of *Frankliniella occidentalis* infected with an orthotospovirus.

Jinlong Han[1] and Dorith Rotenberg[1*]

[1]North Carolina State University, Department of Entomology and Plant Pathology, Raleigh, North Carolina, 27695, U.S.A

Current address of J. Han: Department of Agricultural Biology, Colorado State University, Fort Collins, CO 80523, USA

[*] Correspondence: Dorith Rotenberg, Email: drotenb@ncsu.edu

Abstract. The western flower thrips, *Frankliniella occidentalis*, is the principal thrips vector of *Orthotospovirus tomatomaculae* (order *Bunyavirales*, family *Tospoviridae*), a devastating plant-pathogenic virus commonly referred to as tomato spotted wilt virus (TSWV). The larval gut is the gateway for virus transmission by *F. occidentalis* adults to plants. In a previous report, gut expression at the transcriptome-level was subtle but significant in response to TSWV in L1s. Since it has been well documented that the relationship between the expression of mRNA and associated protein products in eukaryotic cells is often discordant, we performed identical, replicated experiments to identify and quantify virus-responsive larval gut proteins (427) to expand our understanding of insect host response to TSWV. While we documented statistically-significant, positive correlations between mRNA and protein abundance for genes (4,192) expressed in first and second instar guts, there was virtually no alignment of individual genes identified to be differentially modulated by virus infection at the transcriptome and proteome levels. Predicted protein-protein interaction networks associated with clusters of co-expressed

proteins revealed wide variation in correlation strength between protein and cognate transcript abundance, which appeared to be associated with the type of cellular processes, cellular compartments, and network connectivity represented by the proteins. In total, our findings indicate distinct and dynamic regulatory mechanisms of transcript and protein abundance (expression, modifications, and/or turnover) in gut tissues regardless of virus-infection. This study provides molecular candidates for future functional analysis of thrips vector competence, and underscores the necessity of examining complex virus-vector interactions at a systems level.

Keywords: western flower thrips; tomato spotted wilt virus; insect gut tissue; virus-vector interactions; weighted gene co-expression network analysis

1. Introduction

Quantitative measurements of molecular species, such as RNA and protein, are essential for dissecting molecular mechanisms in disease biology. mRNA abundance in cells has been commonly used as a proxy for protein expression (Vogel and Marcotte. 2012), however, the accumulating body of evidence generated from animal systems with numerous genomics data sets (eg. human) clearly shows that the correlation of mRNA and protein levels varies significantly across cell types (Du et al., 2019) and transcript abundance insufficiently predicts protein abundance (Liu et al., 2016; Payne, 2015). This context-dependent and discordant association between mRNA and protein levels has been ascribed to a variety of complex regulatory mechanisms, including post-transcriptional and post-translational silencing, mRNA decay and protein degradation (Gygi et al., 1999; Vogel & Marcotte, 2012). As such, it is imperative to investigate both transcriptional and translational events and/or outcomes in order to

gain more precise views of host cellular, tissue system and physiological responses associated with pathogen invasion.

Orthotospovirus tomatomaculae, in the genus *Orthotospovirus* (family *Tospoviridae*, order Bunyavirales) and commonly known as tomato spotted wilt virus (TSWV), causes severe damage on many economically important food crops. It is transmitted by the western flower thrips, *Frankliniella occidentalis* (Pergande), in a circulative-propagative manner. The first (L1) and early second instar (L2) larval thrips are critical to the virus transmission cycle because adults - the most significant plant inoculators of TSWV during probing and feeding in the landscape - are only vector competent if they acquire TSWV as young larvae (Moritz et al., 2004; van de Wetering et al., 1996). Upon entry of the L1 midgut, the virus replicates and spreads throughout surrounding visceral gut muscles and disseminates to the salivary glands during the late L2 stage. Virus is retained, albeit at lower abundance during the pupal stages (Schneweis et al., 2017), and when adults eclose, virus increases in the adult principal salivary glands (Montero-Astúa et al., 2016). The virus is inoculated into the plant through thrips feeding (Montero-Astúa et al., 2016; Ullman, 1992; Ullman et al., 1993, 1995). Virus infection and replication of this circulative virus in thrips offers opportunities for unraveling complex tissue response to infection. Considering the critical role of larval guts for successful establishment of infection, unveiling molecular responses to virus invasion of this tissue system is a significant step toward devising novel crop disease control strategies that disrupt the gut-virus interaction.

Numerous studies have utilized next-generation sequencing to investigate the molecular aspects of interactions between insect vectors and their associated plant viruses (Dietzgen et al., 2016; Ghosh & Ghanim, 2021;; Heck, 2018; Rotenberg et al., 2015; Martin et al., 2017). However, only a handful of studies have examined the effect of plant virus infection on the gut

of vector species (Brault et al., 2010; Geng et al., 2018; Han & Rotenberg, 2021; Zhao et al., 2016). Previously, we reported the first thrips gut transcriptome of *F. occidentalis* larvae in response to TSWV infection (Han & Rotenberg, 2021), and provided several candidate biomarkers of infection at a stage when TSWV is known to enter and invade the epithelium of the anterior midgut of L1 (*i.e.*, limited to a few cells) (Ullman et al., 2012; Badillo-Vargas et al., 2019), and when virus titer is at its lowest during the transmission cycle (Badillo-Vargas et al., 2012, Rotenberg and Whitfield, 2018; Han and Rotenberg, 2021). It was the early L1 stage - three hours after a 24-hour exposure period on TSWV-infected plant tissue – that exhibited a relatively larger transcriptome-wide response to infection compared to the L2 stage (48 hours post exposure), when virus titers in the gut had increased 3-fold (Han and Rotenberg, 2021), and virus has disseminated and invaded other regions of the midgut (Montero-Astúa et al., 2016a).

This study attempts to gain a deeper understanding of the larval gut-TSWV interaction by i) quantifying modulation of the gut proteome by the virus; ii) identifying protein co-expression networks and functional enrichments represented by network members; and iii) examining the association between the gut proteome (this study) and transcriptome of TSWV-exposed and non-exposed larval thrips, as well as within network protein-transcript abundance correlations, by capitalizing on previously published gut transcriptome data. Collectively, our analyses revealed distinct responses of thrips gut tissues to virus infection at transcriptional and translational levels, stressing the importance of investigating virus-vector interactions using complementary approaches.

3. Results and Discussion

3.1. TSWV proteins detected in larval guts. Comparisons were made between first instar larvae, three hours after a 24-hour exposure time to TSWV-infected plant tissue (3-L1-V) and

second instar larvae ensuing from the 3-L1 cohort, 48 hours after removal from the infected plant tissue (48-L2-V). Structural and nonstructural proteins encoded by the five open reading frames of the TSWV genome were identified in the gut proteome. We documented a significantly greater ($P < 0.05$) normalized abundance of nucleocapsid protein (N), a virion structural protein, and two non-structural viral proteins (NSs and NSm), in 48-L2, compared to 3-L1 of virus-exposed insects (**Figure 1**). Virus accumulation in thrips guts after removal of the source of inoculum coupled with an increased abundance of non-structural proteins is indicative of virus replication (Ullman et al., 1993; Onishi et al., de Asis Fihlo et al., 2007). Increases in viral RNA (TSWV-N) from L1 to L2 bodies (Schneweis et al., 2017; Badillo-Vargas et al., 2012) or guts (Han and Rotenberg, 2021) of *F. occidentalis* during growth and development have been reported previously, corroborating our findings at the protein level.

While all attempts were made to avoid cross-contamination of virus-exposed (V) and non-exposed (NV) samples during larvae exposure, tissue dissection, and sample processing, there were peptides that identified TSWV proteins in the NV-larval gut samples, albeit at lower numbers (47% - 88% fewer unique peptides that identified a viral protein) and lower normalized abundance per protein (23% to 99% lower) as compared to V-guts (**Supplementary Table S1**). We attribute the occurrence of these TSWV peptides to some unidentifiable source or means of contamination as has been suggested previously in tissue proteomics (Rajarapu at al., 2022) and transcriptomics (Stafford-Banks et al., 2014) studies with this virus-vector system.

3.2. Influence of larval developmental stage and virus on normalized protein abundance in the gut proteome. Our proteomic study detected a total of 47,934 unique peptides across all gut samples, which identified 4,192 different *F. occidentalis* proteins. Principal component analysis

(PCA) on a correlation matrix of normalized thrips protein abundance values by larval stage-treatment groups (NV and V) revealed that larval stages (3-L1 vs 48-L2), regardless of treatment, clustered separately along PC1, which explained the majority of the variation in the data set (31.8%), and that 3-L1 clustered by virus infection status (NV vs V) along PC2 (18.7% of the variation) (data not shown). The finding that virus exposure and/or infection had a subtle effect on the gut proteome compared to the physiological status of the thrips vector is consistent with a report that *F. occidentalis* sex (male vs female adults) had a larger influence on the salivary gland proteome than TSWV infection of these adult glands (Rajarapu et al., 2022).

3.3. Differential abundance of larval gut proteins in response to virus. The larval gut response to TSWV significantly modulated the abundance of 384 and 73 proteins expressed in 3-L1 and 48-L2 guts, respectively (**Figure 2A**; $P < 0.05$ and at least 2-fold change compared to NV guts), referred herein as virus-modulated, differentially abundant proteins (vDAPs, **Supplementary Table S2**). Comparatively, the number and magnitude of change of vDAPs in the *F. occidentalis* larval guts (**Figure 2A,** mean fold change = 18.9 (3-L1) and 8.6 (48-L2)) indicated a more robust gut response to TSWV exposure and infection in 3-L1 compared to 48-L2. This finding is consistent with our previous gut transcriptome study (Han and Rotenberg, 2021), where twice as many transcript sequences were differentially-abundant in guts of 3-L1 (105) compared to 48-L2 (49). Notably, virus exposure led to a greater perturbation of gut proteins compared to transcripts, highlighting a more pronounced impact of virus at the protein level than at the transcriptional level. The temporal proximity of exposure to and acquisition from the TSWV-infected plant host for the 3-L1 cohorts (3 hours post-removal of inoculum source) may, in part, explain the relatively larger response of L1s to the virus treatment. Within

the time-frame of the present study, gut infection of young L1s is localized to epithelial cells of the anterior midgut, and within the L2 stage, TSWV has spread to other regions of the midgut and surrounding visceral muscles. This well documented phenomenon of tissue tropism (reviewed in Montero-Astua et al., 2016b), and the relative abundance of TSWV proteins in L1 compared to L2 guts (**Figure 1**) did not appear to explain the more robust effect of TSWV on L1s. We hypothesize that ingestion of infected plant tissue contributes to the larger gut response in L1s compared to L2s, which were sampled 48 hours after removal of the source of inoculum (~24 hours post molting from L1 to L2). Examination of indirect (infected plant) vs. direct effects of TSWV on gut physiology is warranted to fully understand the nature of the virus-vector interaction.

The majority of vDAPs in larval guts were predominantly down-regulated (81% in 3-L1, 71% in 48-L2) (**Figure 2B**), as was reported for a whole-body proteome study of *F. occidentalis* L1s infected with TSWV (Badillo-Vargas et al., 2012). Of the 30 vDAPs determined to be shared between the two larval stages in the present study (**Figure 2B**), alpha- and beta-tubulins showed the most reduced abundance (**Figure 2C**) in 3-L1 guts. Alpha- and beta-tubulins are structural, cytoskeletal proteins that form microtubules (MT), adaptable tube-like structures that can rapidly reorganize, extend and depolymerize under different physiological states and external stimuli (Desai and Mitchison, 1997; Gasic et al., 2019). Depolymerization of microtubules (reviewed in Ohi et al., 2021) can lead to tubulin mRNA instability (Gasic et al., 2019) and subsequent reduced tubulin synthesis (Ben-Ze'ev et al., 1979 ref), referred to as tubulin autoregulation. Autoregulation may have contributed to the reduced abundance of tubulins in guts of TSWV-infected L1s. Tubulin cytoskeletal networks in animal cells have been implicated in virus attachment to cell surfaces, intracellular trafficking, and virus replication (Naghavi and

Walsh, 2017). More proximal to the present study, rice stripe virus (RSV), a circulative-propagated, plant-pathogenic RNA virus, was shown to upregulate and utilize alpha-tubulin for accumulation and dissemination along the route from gut cells to salivary glands of the small brown planthopper (Li et al., 2020). Negative regulation of tubulins in TSWV-infected L1 whole body proteomes (Badillo Vargas et al, 2012) and the transcriptome of pro-pupae (P1) of *F. occidentalis* (Schneweis et al., 2017) is consistent with the present findings of larval gut tubulins.

Other stage-shared proteins in the present study infer roles in DNA-binding/replication, intracellular transport, proteolysis and other enzymatic processes (ligation, hydrolysis, redox), transcription of non-coding RNAs, lipid and carbohydrate metabolism, and structural constituents of ribosomes (**Figure 2C**, **Supplementary Table S3**). For the most part, these shared proteins and their associated normalized abundances under the infection state indicate a negative regulation of gene expression and/or enhanced protein turnover. If and how TSWV inhibits or interferes with these cellular processes in thrips vector cells remains to be determined, and the current lack of a robust, immortalized thrips cell line for research beckons us to innovate ways to address these basic questions.

3.4. Gene ontologies (GO) and functional enrichment of larval gut proteins in response to TSWV. Of the 427 non-redundant vDAPs identified across the two larval stages, 90% were assigned GO terms (**Supplementary Figure S2, Supplementary Table S2**). The most predominant GO annotations were ‘gene expression’ and ‘hydrolase activity’ for both stages, and the most common cellular compartments assigned to the vDAPs were ‘nucleus’ and ‘integral component of membrane’ for 3-L1 and 48-L2 guts, respectively. Using STRING, it was determined that 88% and 76% of the thrips vDAP sequences for the 3-L1 and 48-L2 guts,

respectively, identified orthologs (394 in total) in the annotated *D. melanogaster* proteome, and enrichment analysis of these orthologs revealed more granular, functional categories encompassed by the mapped vDAPs (**Figure 3**). Not surprisingly, categories for the collection of vDAPs mirrored those of the larval stage overlap (replication, transcription, and intracellular protein transport) (**Figure 2C**), with an emphasis on the multifunctional role of alpha-tubulin (TubA85E ortholog).

Other enriched categories were revealed, namely immune systems and post-translational protein modification. We found that a vast majority of immune-associated vDAPs was down-regulated in larval gut tissues after virus infection. The 3-L1 guts exhibited a more pronounced immune response than 48-L2 guts (34 vs 9 vDAPs). The most perturbed gut protein associated with immunity was Mdr50 (fold change (FC), V/NV) = -288.33), a subunit of ABC transporter complex involved in xenobiotic transporter activity (Gerrard, et al. 1993). Among a few up-regulated vDAPs, the protein pellino, which acts as a positive regulator of the Toll signaling pathway (Susanna Valanne, et al. 2011), showed a 4.61-fold increase in abundance in 3-L1 guts following virus infection.

There were 34 vDAPs in 3-L1 guts classified in the post-translational protein modification category, and all but five exhibited reduced abundance under infection (**Figure 3, Supplementary Table S4**). This category represented proteins with roles in ubiquitination or deubiquitination (proteolysis), Ras-related functions (vesicle-mediated transport, signal transduction), autophagy and phagocytosis, nuclear import of transcription factors (antibacterial, humoral immunity) and chromatin/histone binding (methylation), Golgi function, proteasome structure and peptidase function, regulation of alternative mRNA splicing (spliceosome), protein trafficking, and chitin synthesis and protein glycosylation (UDP-N-acetylglucosamine

diphosphorylase). vDAPs at the extremes in magnitude in change in response to TSWV for this functional category (**Figure 3**) were Rab3 (FC = -130.26), a Ras-related protein known for its role in exocytosis by regulating the final step in synaptic vesicle fusion, and F-box protein 11 isoform A (FC = 17.44), a subunit of a SCF E3-ubiquitin ligase complex involved in degradation of diverse cellular proteins, and has been implicated in silencing by si- and miRNAs at the step of RISC-loading in *Drosophila melanogaster* (Lee et al., 2009). More recently, it was reported that silencing of USB7 (a E3-ubiquitin ligase) in *F. occidentalis* adults resulted in reduced virus transmission to the host, but had no effect on acquisition (i.e., larval gut) (Shi et al., 2023). In total, these vDAP identified orthologs with known, conserved functions with vital roles in cell biology and virology.

3.5. Larval gut co-expression protein networks. Using the collection of all *F. occidentalis* larval gut proteins identified in this study (4,192), WGCNA revealed a total of 10 distinct clusters (i.e., modules) of co-expressed proteins grouped into four branches of the eigengene dendrogram, referred to as meta-modules (Langfelder & Horvath. 2008) (**Figure 4A**). vDAPs were distributed across eight modules with 64% located in the 'turquoise' module; all down-regulated in response to TSWV (**Supplementary Table S2**). Modules that were positively correlated (**Figure 4B**, $P < 0.01$) with the TSWV condition were the 'black' (10% of vDAPs) and 'red' (3.5% of vDAPs) modules, and all of the vDAPs in these modules were more abundant (upregulated) in response to virus (3-L1), with the exception of the sole five associated with the 48-L2 stage (**Supplementary Table S2**). The 'black' and 'red' modules of proteins resided in two different meta-modules, suggesting different biological roles (Langfelder and Horvath, 2007) with regards to perturbation by the virus condition.

Network membership analysis of the 'black' (210 proteins in total) and 'red' (261 proteins in total) modules identified 10 and 18 highly interconnected hub proteins, respectively, each with a module membership value (kME) no less than 0.90 and the greatest number of intramodular connections (kIN > 32) (**Supplementary Table S5**; **Figure 5**). Five of the 'black' hub proteins were vDAPs (**Figure 5B)**, while none of the 'red' hubs were responsive to TSWV under our experimental conditions (**Figure 5D**). Given that these hub proteins share the most connections with other co-expressed proteins (including other hubs), they likely play a critical role in regulating or coordinating activities of 'neighborhood' proteins in response to viral activities in larval gut cells.

Systems-level analyses of disease-associated omics datasets have led to the identification of key molecular pathways and potential new targets and biomarkers of diseases (Altaf-Ul-Amin et al., 2014). Functional analysis of the hub proteins and their connecting proteins with regards to thrips biology and thrips vector competence warrants further study. For example, in the present study, F9188, an uncharacterized protein with unknown function, was identified as a hub protein in the 'black' module. Notably, it was also shown as the most connected central hub gene in a co-expression network that was enriched with TSWV-responsive transcripts in *F. occidentalis* larval guts (Han & Rotenberg, 2021). While its protein abundance was not significantly disturbed by virus infection, the abundance of its corresponding mRNA was shown to be ~3-fold higher in TSWV-infected 3-L1 guts compared to uninfected controls.

3.6. Comparative analysis of the proteomes and transcriptomes of the larval gut of TSWV-exposed and non-exposed thrips. Identifiers (FOCC gene IDs) of all nonredundant larval gut proteins identified (4,192) were used to recover their associated transcripts and normalized

abundance values from a previously published RNA-Seq expression data set for *F. occidentalis* larval transcriptomes under V and NV conditions (Han and Rotenberg, 2021) (**Supplementary Table S6**). The search recovered transcript expression data for 99.3% of the protein data set. Statistical analysis of the paired transcript-protein expression data ($\log_2$ normalized abundance) revealed significant, positive correlations (Pearson, r) for guts of the 3-L1 (r = 0.59, $P < 0.0001$) and 48-L2 (r = 0.567, $P < 0.0001$) larval stages. Correlations of this strength and significance are remarkable for heterogeneous gut tissue (diverse cell types) that represent different regions of the insect gut (foregut, midgut and hindgut). Regression analyses on (x,y) pairs of TSWV-modulated proteins for each virus condition (NV/V) revealed a more significant, yet weak association between protein and transcript abundance in V-3-L1 guts ($R^2 = 0.09$, $P < 0.0001$, n = 374) compared to NV-3-L1 ($R_2 = 0.0$, $P = 0.52$, n = 374) guts (**Figure 6B vs 6A**). Despite significant positive correlations in larval guts in general, use of transcript abundance to explain the variation in protein abundance had limited value for TSWV-modulated proteins.

The two ‘omics expression datasets were cross-queried (by OGS IDs) to identify overlapping sequences that were significantly ($P < 0.05$) perturbed by TSWV as compared to NV counterparts. Only three sequences were shared between the data sets. Two were responsive in 3-L1 guts, and were discordant in direction of change [FOCC011070: protein (FC ~ 4), transcript (FC ~ -4); and FOCC004976: protein (FC ~ -8), transcript (FC ~ 3)] while the third was down-regulated at the transcript and protein levels in 48-L2 guts [FOCC001083: protein (FC ~ -17), transcript (FC ~ -9)]. The annotations for these three sequences are dopamine N-acetyltransferase-like isoform X4, lysozyme C-3-like, and cuticle protein 6-like, respectively, providing little added insight about the effect of virus infection on gut response. The gut-level

response of lysozyme C-3-like downregulation is consistent with whole-body proteomic responses to TSWV infection (Badillo-Vargas et al., 2012).

Among the 427 non-redundant vDAPs, 99% displayed unaltered mRNA abundances after virus exposure, implying undisturbed transcription while post-transcriptional processes influenced translation efficiency (Sunnerhagen P. 2007; Zhao, et al. 2016). This gave rise to variable translation, with some proteins being enhanced and others diminished. Similarly, among the 147 non-redundant, differentially-expressed transcripts from our previous gut transcriptome study (Han and Rotenberg, 2021), 98% maintained stable protein levels despite potential fluctuations in mRNA levels. This protein stability may arise from adjustments in translation rates, changes in mRNA stability, and post-translational modifications like phosphorylation, ubiquitination, or acetylation, ensuring protein constancy even in the face of shifting mRNA abundances (Mata et al. 2005; McCarthy, 1998; Pradet-Balade et al., 2001; Deribe et al., 2010). Collectively our data indicate distinct transcriptional and translational responses of the larval gut to TSWV infection, given that consistent experimental conditions, parameters, and timing were followed intentionally for the two ‘omics studies. These distinct responses likely reflect temporally-dynamic and direct physical and biochemical interactions between the bunyavirus and host cell machinery, molecular building blocks (nucleotides and amino acids), and myriad unknown host factors available in the cytoplasm to complete its ‘life-cycle’ [transcription, translation, replication, virion assembly] (reviewed in Elliott and Schmaljohn 2013). TSWV-induced, post-translational modifications of host proteins (Gong et al., 2023) may also contribute to the discordance between transcriptomic and proteomic response. Mass spectrometry-based, discovery proteomics methodologies, such as the one used in the present study, are not designed to handle identification of PTMs (Kim et al., 2015), especially with the use of a reference

proteome database composed of single protein models per gene model. Therefore, use of existing RNA-Seq datasets for *F. occidentalis* in public repositories to identify transcript variants and isoforms for gene models and incorporation of PTM detection techniques for new proteome datasets may provide a more complete view of host response to virus or other persistent perturbations.

3.7. **Intramodular analysis of correlations between protein and transcript abundance, and functional enrichment.** In an attempt to approach the basis of correlations between transcriptome and proteome expression patterns in larval guts, rank correlation coefficients (Spearman's rank coefficient, ρ) were calculated for protein-transcript pairs for the 10 sets of module members for each larval stage-treatment combination. The analyses revealed significant, positive associations (ρ) for all of the 10 modules. However, the strength of the association varied considerably among the 10 modules (**Figure 7**), with an apparent trend of tighter associations in 3-L1 compared to 48-L2 guts for each module, as was the case for the entire protein-transcript abundance data set (**Figure 6**).

Functional enrichment analyses of protein-protein interactions (PPI) in four module networks were performed, two of which represented extreme strengths in protein-transcript associations in the data set [pink (mean ρ = 0.817) and turquoise (mean ρ = 0.528)], and the two that were significantly (positive) associated with TSWV (**Figure 4B**), which also varied in strength of association [red (mean ρ = 0.731) and black (mean ρ = 0.556)] (**Figure 7**). A summary of the functional enrichments and associated metrics for these four networks is presented in **Supplementary Table S7.** The pink network (135 nodes) was enriched ($P < 1x10^{-16}$) in ribosomal proteins and other highly conserved translational machinery and translation

factors (e.g. eukaryotic initiation factors). The red network (222 nodes) was enriched ($P < 1x10^{-16}$) in proteins associated with mitochondrial structure and processes, including cristae formation, mitochondrial ribosomes, electron transport, fatty acid biosynthesis (Acyl-CoA synthetase, val/leu/isoleu degradation), metabolism of structural and signaling lipids, and ER transmembrane proteins. The black network (161 nodes) had relatively fewer member connections (no. of edges in relation to no. of hubs) compared to the other modules of comparable neighborhood size, and functional enrichment was statistically supported, albeit at a lower significance ($P < 5x10^{-4}$). It was enriched for functions associated with mRNA splicing and alternative splicing, the tricarboxylic acid cycle (TCA), regulation of endocytosis, and ER integral proteins. The turquoise network (1060 nodes) was enriched ($P < 1x10^{-16}$) in a diverse set of functional roles, not unexpected given the number of proteins in the network, including those known to be associated with DNA replication initiation and elongation (e.g., Minichromosome (Mcm) helicase complex), ribosome biogenesis (Noc), nerve cells (axons), pre-mRNA processing (spliceosome complex), post-embryonic digestive tract morphogenesis, vesicle-mediated transport (clathrin adapter), immune system (interferon alpha/beta signaling) and response to xenobiotics.

WGCNA-derived modules represent different sets of interconnected genes, and by virtue of these connections, indicate coordinated interactions that culminate into distinguishable cellular processes. Based on the intramodular correlations between protein and transcript abundance and functional enrichments determined by our analyses, we posit that the strength of the transcript-protein association may be explained, in part, by the type of cellular processes and cellular compartmentalization represented by the module (Du et al., 2019, Tasaki et al., 2022), and possibly connectivity (number of connections) of a protein with other proteins within the PPI

network, like conserved translational and post-translational machinery (e.g., ribosomal protein complexes, proteasome) (Tasaki et al., 2022). Indeed, enrichment of the pink and red networks with ribosomal complexes and mitochondrial compartmentalized proteins, respectively, is consistent with this hypothesis, as is the black network with fewer PPI edges and enrichment in proteins that play roles in pre-mRNA processing. One study addressed the discordance of transcript to protein abundance by training models on cancer tumor 'omics datasets and validating the predictive power of non-cognate transcript abundance on proteins in large and small multiprotein complexes associated with diverse biological processes (Srivasta et al, 2022). Determination of transcript variants and post-translational modifications in tissues of *F. occidentalis* and/or other thrips vectors of orthotospoviruses may provide a pathway towards identification of reliable biomarkers of vector competence.

3. Experimental Procedures

3.1 Biological source for larval gut protein samples. The source of *F. occidentalis* (lab colony on green bean pods), isolate of TSWV (*Emilia sonchifolia* lab culture), and the TSWV acquisition protocol - age-synchronized L1 cohorts subjected to a 24-hour acquisition access period (AAP) on infected (V) or non-infected (NV) plant tissue, then moved to green beans for gut clearing and development - were prepared and used precisely as described for our larval gut transcriptome study (Han & Rotenberg, 2021). The gut proteome experiment consisted of the two treatments (V and NV) and two sampling time-points post AAP [3-hrs (L1 stage) and 48-hrs (L2 stage), and the experiment was conducted four independent times (i.e., biological replicates) (**Supplementary Figure S1**). Infection rates of larval cohorts (V and NV) were determined by arbitrarily subsampling 10 individuals (L1) and subjecting each larval body to RNA extraction and cDNA synthesis followed by real-time PCR detection of TSWV nucleocapsid (N) RNA and

F. occidentalis actin RNA (internal reference) as described previously (Han and Rotenberg, 2021). Based on this strategy, it was confirmed that each V cohort (all four bio-replicates) exhibited a 100% infection rate, and TSWV N RNA was not detected in the NV cohorts.

3.2. Thrips gut dissection. Larval dissections were performed to carefully remove gut tissue, which included the foregut, midgut, and partial hindgut as described previously (Han & Rotenberg, 2021), with adaptations for downstream protein analyses. A detailed experimental design is illustrated in Figure S1 in the current study. Each thrips larva was placed on a sterile microscope slide, immersed in ice-cold phosphate-buffered saline (PBS) buffer (~ 80 ul), then decapitated using a sterile, polymer-coated scalpel blade (#15, Southmedic Inc., Barrie, Ontario, Canada). Undesired tissues, including principal and tubular salivary glands, fat body, and other residual body parts were carefully removed from the gut tissue using an ultrafine single deer hair tool (Ted Pella, Redding, CA). After dissection, an additional ice-cold PBS buffer (~100 ul) was added to the slide to further wash away the unwanted residual tissue and body contents from the gut tissue. Each gut was immediately pulled through the surface of the buffer with the single deer hair tool and transferred into 200 ul of fresh ice-cold PBS buffer in a sterile 1.5-ml microfuge tube. This process was repeated to pool a total of 100 guts in the same tube, which was placed on ice throughout the entire process. Each replicate of the experiment was performed under the same laboratory conditions, during the same time period of the day, and by alternating between 20 NV-larvae and 20 V-larvae until obtaining 100 guts for both samples (using separate tools for each treatment). Sample tubes containing gut tissues were centrifuged at 8,000 g for 5 minutes and excess PBS buffer was carefully removed to achieve 20 ul in each tube. Gut samples were immediately flash-frozen in liquid nitrogen and stored at -80□ until use. A total of 16 gut samples were prepared (two treatments, two sampling time points, four bio-replicates). To ensure

that 100 guts would provide sufficient protein concentrations for downstream processing, we performed the bicinchoninic acid assay (BCA) (Thermo Fisher Scientific, Waltham, MA) to estimate total protein concentration on an entirely separate pool of 100 L1 guts.

3.3. Tissue-processing and ultra-performance liquid chromatography-tandem mass spectrometry analysis. The 16 frozen gut samples were submitted to the Duke Proteomics and Metabolomics Core Facility for sample processing and ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) analysis. Briefly, each sample was spiked with undigested bovine casein as an internal quality control standard, which were then supplemented with 100 ul of 8M urea/50 mM ammonium bicarbonate. Proteins were solubilized by probe sonication (30% power, 3X). Protein content was estimated based on the protein concentration determined by BCA assay, and 4.5 ug of protein extract was used for enzymatic digestion with 20 ng/ul sequencing grade trypsin (Promega, Madison, WI) for 1 hour at 47°C, and eluted using 50 mM TEAB, followed by 0.2% FA, and lastly using 50% ACN/0.2% FA. All samples were then lyophilized to dryness and resuspended in 12 ul of 1% trifluoroacetic acid/2% acetonitrile with 12.5 fmol/ul of yeast alcohol dehydrogenase (ADH). A study pool quality control (SPQC) was created by combining equal volumes of each sample.

Quantitative LC/MS/MS was performed with 4 ul of each sample using a nanoAcquity UPLC system (Waters Corp) coupled to a Thermo Orbitrap Fusion Lumos high resolution accurate mass tandem mass spectrometer (Thermo) equipped with a FAIMSPro device via a nanoelectrospray ionization source. Briefly, the sample was first trapped on a Symmetry C18 20 mm×180 um trapping column (5 ul/min at 99.9/0.1 v/v water/acetonitrile), after which the analytical separation was performed using a 1.8 um Acquity HSS T3 C18 75 um×250 mm column (Waters Corp.) with a 90-min linear gradient of 5 to 30% acetonitrile with 0.1% formic

acid at a flow rate of 400 nl/min with a column temperature of 55□. Data collection on the Fusion Lumos mass spectrometer was performed for three different compensation voltages (CV) (-40v, -60v, -80v). Within each CV, a data-dependent acquisition (DDA) mode of acquisition with a r=120,000 (@ m/z 200) full MS scan from m/z 375-1500 with a target AGC value of 4e5 ions was performed. MS/MS scans were acquired in the linear ion trap in "rapid" mode with a target AGC value of 1e4 and max fill time of 35 milliseconds. The total cycle time for each CV was 0.66 seconds, with total cycle times of 2 seconds between full MS scans. A 20 seconds dynamic exclusion was employed to increase depth of coverage. The total analysis cycle time for each sample injection was approximately 2 hours.

3.4. Mass spectrometry data acquisition and analysis. Following 20 total UPLC-MS/MS analyses (including four replicative SPQC samples), data were imported into Proteome Discoverer 2.4 (Thermo Scientific Inc.) and individual UPLC-MS/MS data files were aligned based on the accurate mass and retention time of detected precursor ions ("features") using Minora Feature Detector algorithm in Proteome Discoverer. Relative peptide abundance was measured based on peak intensities of selected ion chromatograms of the aligned features across all runs. The MS/MS data was searched against the *Frankliniella occidentalis* database, tomato spotted wilt virus protein sequences [AAF80981.1 (G_N/G_C glycoprotein precursor); AAL55403.1 (RNA-dependent RNA polymerase); and in house clones for NSs, NSm, and N] and a common contaminant/spiked protein database (bovine albumin, bovine casein, yeast ADH, etc.), and an equal number of reversed-sequence "decoys" for false discovery rate (FDR) determination. Mascot Distiller and Mascot Server (v 2.5, Matrix Sciences) were utilized to produce fragment ion spectra and to perform database searches. Database search parameters included fixed

modification on Cys (carbamidomethyl) and variable modification on Met (oxidation). Search tolerances were 2ppm precursor and 0.8Da product ion with full trypsin enzyme rules. Peptide Validator and Protein FDR Validator nodes in Proteome Discoverer were used to annotate the data at a maximum 1% protein FDR based on q-value calculations. A peptide matched to multiple different proteins was exclusively assigned to the protein that had more identified peptides. Protein homology was addressed by grouping proteins that had the same set of peptides to account for their identification. A master protein within a group was assigned based on percent coverage. **Supplementary Table S8** provides raw peptide sequences, metrics and raw abundance values for the 20 total UPLC-MS/MS analyzed samples.

Prior to imputation, a peptide filter was deployed which required a peptide to have a measurable quantitative value in >50% of the samples within a treatment group. After filtering, any missing values were imputed using the following rules: 1) if only a single signal was missing within the group of three, an average of the other two values was used, or 2) if two out of three signals were missing within the group of three, a randomized intensity within the bottom 2% of the detectable signals was used. To summarize the protein abundance, all peptides belonging to the same protein were summed into a single intensity. The protein abundance data was then normalized using a robust mean normalization, in which the highest and lowest 10% of the signals were excluded and the average of the remaining signals were normalized to be the same across each individual channel.

In order to assess technical reproducibility, % coefficient of variation (%CV) was calculated for ADH spiked into each sample prior to any type of data normalization. The variation was 23.4%, which is well within expected for non-normalized spiked control data. Then, %CV was calculated for each protein across four injections of the SPQC pool that were

interspersed throughout the study. The mean %CV of the SPQC pools was 9.2% which is well within expected analytical tolerances for this type of analysis. To assess biological plus technical variability, %CVs were calculate for each protein across the four stage-treatment sample groups, which were 33.6%, 21.9%, 22.3% and 22.9% for NV-3-L1, V-3-L1, NV-48-L2, and V-48-L2, respectively. Overall, these variations are well within desired variations and indicate a consistent protein extraction, processing and data acquisition.

3.5. Identification and annotation of virus-mediated, differentially abundant proteins (vDAPs). A two-tailed heteroscedastic *t*-test was employed to identify vDAPs in V guts compared to NV guts based on normalized protein abundance and using the criteria of p-value < 0.05 and fold change ≥ 2. Amino acid sequences of all vDAPs, i.e., non-redundant vDAPs identified across both larval developmental stages, were subjected to Blastp to query the NCBI non-redundant protein database followed by mapping and annotation with Gene Ontology (GO) terms ($E < 1.0E-5$) for three provisional functional GO categories (biological process, molecular function, and cellular component) using OmicsBox v2.0 (Götz et al., 2008). To further discriminate more specific putative roles, the vDAP amino acid sequence (one file containing multiple fasta-formatted sequences) were subjected to STRING analysis (https://string-db.org/; Szklarczyk, et al. 2023) to map vDAPs to the program-supplied, annotated *Drosophila melanogaster* proteome in order to predict functional and physical protein networks and to identify significantly enriched functions represented by the vDAPs.

3.6. Weighted gene co-expression network analysis (WGCNA). A matrix of $\log_2$-transformed, normalized abundance values for all detected proteins (4,229) across the 16 gut samples were subjected to WGCNA for identifying modules (clusters) of co-expressed proteins and their statistical associations with an external sample trait, in this case, TSWV treatment (Langfelder &

Horvath, 2008). The adjacencies between proteins were calculated using a soft-thresholding power of 10 under the 'signed' network. The adjacency matrix was transformed into a topological overlap matrix for calculating the corresponding dissimilarity. Modules of co-expressed proteins were identified by applying the 'dynamic tree cut' method with minimum module size set as 30. By virtue of including all gut proteins identified in the analysis, modules have the potential to include both vDAPs (TSWV-responsive) and non-vDAPs (not responsive to TSWV as compared to NV treatment). Similar modules were merged when their eigengenes were highly correlated (80%). The module-trait relationships were analyzed by correlating module eigengenes and the infection status (NV or V) of gut samples. Modules that were significantly associated with TSWV treatment were further analyzed for identifying intramodular hub proteins. Proteins that ranked in the top 10% of both module membership (kME, correlation of a protein to its module eigengene) and intramodular connectivity (kIN, sum of connection weights between a protein with all its connected neighbors) among all module proteins were considered as hub proteins. The intramodular interaction networks were visualized with VisANT software v5.53 (Hu et al., 2004). In addition, as described above (section 3.5) for all DAPs, STRING analyses were performed on module sequence sets that were determined to be significantly ($P < 0.05$) associated with the trait of TSWV treatment in order to identify enriched functional categories within modules.

3.7. Cross-datasets comparisons of transcriptomes and proteomes of *F. occidentalis* larval guts. Sequence identifiers (OGSv1.1, FOCC IDs, Rotenberg et al., 2019, 2020) of all non-redundant proteins identified in L1 and L2 guts were used to query a *F. occidentalis* larval (L1 and L2) gut transcriptome sequence database assembled previously for a gut RNAseq experiment (Han & Rotenberg, 2021) that utilized the same insect and virus source, larval developmental

stage-time of sample, experimental set-up, conditions and personnel as implemented in the present study. A data matrix consisting of sequence identifiers, sample factors [larval stage (3-L1, 48-L2), treatment (V/NV)], average (n=4) $\log_2$(normalized transcript abundance) (x) and average (n=4) $\log_2$(normalized protein abundance) (y) was prepared. Statistical correlation coefficients (Pearson, r) were calculated and linear regression analyses were performed on all x, y data (all proteins identified with a transcript match) and vDAPs only for each larval stage – treatment combination to determine the strength and significance of the relationship between protein and transcript abundance. Spearman rank correlation coefficients (□) were calculated for the sets of x, y data associated for each WGCNA module of co-expressed proteins to examine the possible association between enriched functions ascribed to a module and the strength of the association between protein-transcript abundance. In addition, the list of TSWV-responsive transcripts identified in larval guts previously (105 for 3-L1 and 49 for 48-L2) and gut DAPs identified here were examined for overlap. All correlation and regression analyses were performed with JMP Pro v.15.

Supplementary Materials:

Supplementary Figures.

Figure S1: Experimental design for preparing *Frankliniella occidentalis* larval gut samples for the proteomics study (.pdf)

Figure S2: Distribution of TSWV-modulated, differentially-abundant proteins (vDAPs, up- and down-regulated) into gene ontologies (GO) assigned for up- and down-regulated proteins identified in guts of 3-L1 and 48-L2 stages (.pdf)

Supplementary File 1: Workbook of large supporting data tables (Table S1 - S8) (.xlsx)

Author Contributions: Conceptualization, D.R and J.H..; methodology, D.R. and J.H.; formal analysis, J.H. and D.R; investigation, J.H.; resources, D.R.; data curation, J.H.; writing-original draft preparation and editing, J.H. and D.R.; visualization, J.H.; supervision, D.R.; project administration, D.R.; funding acquisition, D.R.. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the NC Agricultural Research Service and by USDA National Institute of Food and Agriculture grant no. 2018-67013-28495.

Data Availability Statement: The raw mass spectrometry datasets generated in this study have been deposited in the MassIVE repository (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp) under project ID: MSV000089637 (available upon publication)

Acknowledgements. We thank the Duke University School of Medicine for the use of the Proteomics and Metabolomics Core Facility, which provided consultation, sample processing, protein identification and normalization, and statistical analysis of differential protein expression services. We thank Swapna Priya Rajarapu for her assistance with virus acquisition assays and discussions about best practices for thrips dissection for proteomics. We also thank Marlonni Maurastoni and Anna Whitfield for internal review of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Figure legends.

Figure 1. Accumulation of tomato spotted wilt virus (TSWV) proteins in *Frankliniella occidentalis* larval guts. One-way analysis of variance was performed for each protein to determine the effect of larval stage on $\log_2$-transformed, normalized abundance of TSWV proteins detected in the quantitative proteomics dataset: RNA-dependent RNA polymerase (RdRp), two non-structural proteins (NSm and NSs), glycoprotein precursor (GN_GC), and nucleocapsid protein (N). Error bar = standard error of the mean; ns = no significant difference; *P-values* are displayed above bars. 3-L1 = first instar larvae, three hours after a 24-hour exposure time on TSWV-infected plant tissue; 48-L2 = second instar larvae ensuing from the 3-L1 cohort, 48 hours after removal from infected plant tissue.

Figure 2. Differentially abundant proteins in *Frankliniella occidentalis* larval guts after exposure and infection by tomato spotted wilt virus (vDAPs). (A) Volcano plot showcasing the distribution of gut proteins by magnitude of change in normalized abundance of protein between virus (V) and no-virus (NV) conditions and level of significance of that change; proteins with a p-value < 0.05 and fold change ≥ 2 were identified as vDAPs. 3-L1 = first instar larvae, three hours after a 24-hour exposure time to plant tissue of the two conditions (V, NV); 48-L2 = second instar larvae ensuing from the 3-L1 cohort, 48 hours after removal from plant tissue. (B) Venn diagrams showing the numbers of up- and down-regulated vDAPs that are either shared or unique between the two larval stages. The values in parentheses indicate the total number of up- or down-regulated vDAPs in each larval stage. (C) Magnitude and direction of fold changes for vDAPs in common between larval stages.

Figure 3. Fold change and STRING functional enrichment analysis of differentially abundant proteins in *Frankliniella occidentalis* larval guts after tomato spotted wilt virus exposure (V) and infection compared to the non-exposed (NV) condition (ie., vDAPs = 427 non-redundant

proteins across the two larval stages); 3-L1 = first instar larvae, three hours after a 24-hour exposure time to plant tissue of the two conditions (V, NV); 48-L2 = second instar larvae ensuing from the 3-L1 cohort, 48 hours after removal from plant tissue. Eighty-four percent of the vDAP dataset mapped to the *D. melanogaster* (Dm) proteome provided by the program. Reactome categories were selected because of broad coverage and finer resolution of cellular activities enriched in the vDAP annotations; vDAPs with extreme fold change values within each category are labeled with Dm ortholog notations: Mcm2 = DNA replication licensing factor Mcm2 (part of minichromosome maintenance complex); Src64B = Src oncogene at 64b, isoform d (Tyrosine-protein kinase); TUBA85E = Tubulin alpha-2 chain; TfIIFbeta = General transcription factor IIF subunit 2; cdk8 = Cyclin-dependent kinase 8; Fbxo11 = F-box protein 11, isoform a (subunit of a SCF-like E3 ubiquitin ligase); Mdr50 = Multi drug resistance 50 (Atp-binding cassette, subfamily b, member 1; Sra1 = Specifically Rac1-associated protein 1 (cytoplasmic FMR1-interacting protein); CG6142 = uncharacterized protein (belongs to the GMC oxidoreductase family); sqd = RNA-binding protein squid (belongs to hnRNPA family); rab3 = Ras-related protein rab-3c.

Figure 4. Weighted gene co-expression network analysis (WGCNA) identifies clusters of co-expressed larval gut proteins of *Frankliniella occidentalis.* A) The dendrogram illustrates intermodular relationships, with similar modules grouped together under the same meta-modules (MM); shown as branches on the eigengene dendrogram. Asterisks highlight modules that include sets of differentially abundant proteins in response to tomato spotted wilt virus exposure and infection (vDAPs). B) Module-trait relationships demonstrate significant correlations (p-value < 0.05) between modules (clusters of co-expressed proteins, represented as colored blocks on the left side of the column) and the external trait of tomato spotted wilt virus (TSWV)

infection. WGCNA analysis was carried out using log_2-transformed, normalized abundance values of all gut-expressed proteins, using samples from both virus-exposed (V) and non-exposed (NV) conditions with four biological replicates. Asterisks denote modules that exhibit a significant association with the external trait of TSWV infection.

Figure 5. Intramodular network analysis of protein co-expression modules significantly correlated with the TSWV condition (see Figure 4B, red and black modules). Module membership (kME, correlation of a node to a module eigengene) and intramodular connectivity (kIN, sum of connection weights between a node and all its network neighbors) are plotted for all co-expressed proteins in the (A) black and (C) red modules. Pearson's correlation coefficients (r) were calculated on the matrix of kME and kIN values for all proteins in each module. Proteins that ranked within the top 10% in both kME and kIN were considered candidate hub proteins (labeled data points). Intramodular relationships are visualized between co-expressed proteins in the (B) black and (D) red module networks. Proteins enclosed by the rounded square boxes represent tomato spotted wilt virus-modulated differentially abundant proteins (vDAPs) *Frankliniella occidentalis* larval guts. Hub proteins in each network are depicted as red rhombuses. Proteins shared between the 3-L1 and 48-L2 developmental stages, as well as those unique to each stage, are indicated by triangle, circle, and square shapes, respectively. For improved network visualization, a weight cutoff of ≥ 0.25 and ≥ 0.20 was applied in the black and red modules, respectively. The thickness of connecting lines reflects the strength of the connections between proteins. 3-L1 = first instar larvae, three hours after a 24-hour exposure time to plant tissue; 48-L2 = second instar larvae ensuing from the 3-L1 cohort, 48 hours after removal from plant tissue.

Figure 6. Correlations between the abundance of *Frankliniella occidentalis* larval gut proteins and their cognate transcripts for (A) first instar larvae, 3 hours after a 24-hour exposure time on healthy plant tissue (NV 3-L1); (B) first instar larvae, 3 hours after a 24-hour exposure time on TSWV-infected plant tissue (V 3-L1); (C) second instar larvae ensuing from the NV 3-L1 cohort, 48 hours after removal of plant tissue (NV 48-L2); and (D) second instar larvae ensuing from the V 3-L1 cohort, 48 hours after removal of infected plant tissue (V 48-L2). Pearson's correlation and linear regression analyses were performed on $\log_2$-transformed normalized abundances values for proteins and their cognate transcript. Blue- and green-colored data points indicate TSWV-modulated, differentially-abundant proteins (vDAPs). Gray and blue/green lines depict best-fit regression lines for the indicated subsets of proteins.

Figure 7. Associations between larval gut proteins and their cognate transcripts in *Frankliniella occidentalis* for the 10 network-modules of co-expressed proteins. Spearman's rank correlation coefficients (ρ) were calculated on $\log_2$-transformed, normalized abundance of proteins and transcripts each larval stage - virus condition combination for each of the network modules. The number within parentheses () indicates the number of proteins within each module. V/NV = virus-exposed (and infected)/non-exposed; 3-L1 = first instar larvae, three hours after a 24-hour exposure to the condition (V or NV); 48-L2 = second instar larvae ensuing from the 3-L1 cohort, 48 hours post exposure to the condition.

REFERENCES

Altaf-Ul-Amin, M., Afendi, F. M., Kiboi, S. K., & Kanaya, S. (2014). Systems biology in the context of big data and networks. *BioMed Research International*, *2014*. https://doi.org/10.1155/2014/428570

Badillo-Vargas, I. E., Rotenberg, D., Schneweis, D. J., Hiromasa, Y., Tomich, J. M., & Whitfield, A. E. (2012). Proteomic analysis of *Frankliniella occidentalis* and differentially expressed proteins in response to Tomato spotted wilt virus infection. *J. Virology, 86*(16), 8793–8809. https://doi.org/10.1128/jvi.00285-12

Badillo-Vargas, I.E., Chen, Y., Martin, K.M., Rotenberg, D., & Whitfield, A.E. (2019). Discovery of novel thrips vector proteins that bind to the viral attachment protein of the plant bunyavirus, tomato spotted wilt virus. *J. Virology. 93*(21): e00699-19. https://doi: 10.1128/JVI.00699-19

Ben-Ze'ev, A., Farmer, S.R., and Penman, S. (1979). Mechanisms of regulating tubulin synthesis in cultured mammalian cells. *Cell 17*, 319–325. doi: 10.1016/0092-8674(79)90157-0.

Brault, V., Tanguy, S., Reinbold, C., le Trionnaire, G., Arneodo, J., Jaubert-Possamai, S., Guernec, G., & Tagu, D. (2010). Transcriptomic analysis of intestinal genes following acquisition of pea enation mosaic virus by the pea aphid *Acyrthosiphon pisum. J. Gen.Virology, 91*(3), 802–808. https://doi.org/10.1099/vir.0.012856-0

de Assis Filho, F. M., Naidu, R. A., Deom, C. M., and Sherwood, J. L. (2002). Dynamics of tomato spotted wilt virus replication in the alimentary canal of two thrips species. *Phytopathology 92*(7), 729-733. https://doi.org/10.1094/PHYTO.2002.92.7.729

Deribe, Y., Pawson, T. & Dikic, I. (2010). Post-translational modifications in signal integration. Nat Struct Mol Biol 17, 666–672. https://doi.org/10.1038/nsmb.1842.

Desai, A., & Mitchison, T .J. 1997. Microtubule polymerization dynamics. *Annu Rev Cell Dev Biol. 13*: 83–117. https://doi.org/10.1146/annurev.cellbio.13.1.83

Dietzgen, R. G., Mann, K. S., & Johnson, K. N. (2016). Plant virus-insect vector interactions: Current and potential future research directions. *Viruses*, *8*(11), 1–21. https://doi.org/10.3390/v8110303

Du, Y., Clair, G. C., Al Alam, D., Danopoulos, S., Schnell, D., Kitzmiller, J. A., Misra, R. S., Bhattacharya, S., Warburton, D., Mariani, T. J., Pryhuber, G. S., Whitsett, J. A., Ansong, C., & Xu, Y. (2019). Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung. *Am J Physiol Lung Cell Mol Physiol*, *317*(3) :L347-L360. doi: 10.1152/ajplung.00475.2018.

R.M. Elliott, R. M., & Schmaljohn, C.S. (2013). Bunyaviridae, Fields Virology (edn 6), D.M. Knipe, P.M. Howley (Eds.), Wolters Kluwer, pp. 1244-1282

Gasic I, Boswell SA, Mitchison TJ (2019). Tubulin mRNA stability is sensitive to change in microtubule dynamics caused by multiple physiological and toxic cues. *PLoS Biol 17*(4): e3000225. https://doi.org/10.1371/journal.pbio.3000225.

Geng, L., Qian, L. X., Shao, R. X., Liu, Y. Q., Liu, S. S., & Wang, X. W. (2018). Transcriptome profiling of whitefly guts in response to *Tomato yellow leaf curl virus* infection. *Virology Journal*, *15*(1), 1–12. https://doi.org/10.1186/s12985-018-0926-6

Gerrard, B., Stewart, C., Dean, M. (1993). Analysis of Mdr50: a *Drosophila* P-glycoprotein/multidrug resistance gene homolog. *Genomics*,*17*(1):83-88. doi: 10.1006/geno.1993.1286.

Ghosh, S., & Ghanim, M. (2021). Factors determining transmission of persistent viruses by *Bemisia tabaci* and emergence of new virus-vector relationships. In *Viruses* (Vol. 13, Issue

9, p. 1808). Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/v13091808

Gong, Y., Li, Y., Liu, D., Jiang, L., Liang, H., Wu, Y., Wang, F., Yang, J. (2023). Analysis of lysine acetylation in tomato spotted wilt virus infection in *Nicotiana benthamiana*. Front Microbiol. 14:1046163. doi: 10.3389/fmicb.2023.1046163.

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435. https://doi.org/10.1093/NAR/GKN176

Gray, S. M., & Banerjee, N. (1999). Mechanisms of Arthropod Transmission of Plant and Animal Viruses. *Microbiology and Molecular Biology Reviews*, *63*(1), 128–148. https://doi.org/10.1128/MMBR.63.1.128-148.1999

Gygi, S. P., Rochon, Y., Franza, B. R., & Aebersold, R. (1999). Correlation between Protein and mRNA Abundance in Yeast. *Molecular and Cellular Biology*, *19*(3), 1720–1730. https://doi.org/10.1128/MCB.19.3.1720/ASSET/DBC7033A-DCFA-438A-BA94-F79C524792FD/ASSETS/GRAPHIC/MB0391587007.JPEG

Han, J., & Rotenberg, D. (2021). Integration of transcriptomics and network analysis reveals co-expressed genes in *Frankliniella occidentalis* larval guts that respond to tomato spotted wilt virus infection. *BMC Genomics*, *22*(1), 1–17. https://doi.org/10.1186/s12864-021-08100-4

Heck, M. (2018). Insect transmission of plant pathogens: a systems biology perspective. *MSystems*, *3*(2). https://doi.org/10.1128/msystems.00168-17

Hu, Z., Mellor, J., Wu, J., & DeLisi, C. (2004). VisANT: An online visualization and analysis tool for biological interaction data. *BMC Bioinformatics*, *5*, 17. https://doi.org/10.1186/1471-2105-5-17

Kim, M.S., Zhong, J., & Pandey, A. (2016) Common errors in mass spectrometry-based analysis of post-translational modifications. Proteomics, 16(5):700-14. doi: 10.1002/pmic.201500355.

Langfelder P, Horvath S. (2007). Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol. 1*:54. doi: 10.1186/1752-0509-1-54.

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

Li, Y., Chen, D., Hu, J., Zhang, K., Kang, L., Chen, Y., et al. (2020). The α-tubulin of Laodelphax striatellus mediates the passage of rice stripe virus (RSV) and enhances horizontal transmission. PLoS Pathog 16(8): e1008710. https://doi.org/10.1371/journal.ppat.1008710

Lee, Y., Pressman, S., Andress, A. et al. (2009). Silencing by small RNAs is linked to endosomal trafficking. *Nat Cell Biol 11*, 1150–1156. https://doi.org/10.1038/ncb1930

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, *165*(3), 535–550. https://doi.org/10.1016/J.CELL.2016.03.014

Martin, K.M., Barandoc-Alviar, K., Schneweis, D.J., Stewart, C.L., Rotenberg.,, D, Whitfield, A.E. (2017). Transcriptomic response of the insect vector, *Peregrinus maidis*, to maize

mosaic rhabdovirus and identification of conserved responses to propagative viruses in hopper vectors. Virology, 509, 71-81. doi: 10.1016/j.virol.2017.05.019.

Mata, J., Marguerat, S., & Bähler, J. (2005) Post-transcriptional control of gene expression: a genome-wide perspective. Trends in Biochem Sci, 30 (9): 506-514. doi: 10.1016/j.tibs.2005.07.005.

McCarthy, J. E. G. (1998) Posttranscriptional control of gene expression in yeast. Microbiol. Mol. Bio. Rev. 62(4):1492-1553. doi: 10.1128/MMBR.62.4.1492-1553.

Montero-Astúa, M., Ullman, D. E., & Whitfield, A. E. (2016a). Salivary gland morphology, tissue tropism and the progression of tospovirus infection in *Frankliniella occidentalis*. Virology, 493, 39–51. https://doi.org/10.1016/j.virol.2016.03.003

Montero-Astúa, M., Stafford, C., Badillo-Vargas, I., Rotenberg, D., Ullman, D.E., and Whitfield, A. E. (2016b). Tospovirus-thrips biology. In Vector-Mediated Transmission of Plant Pathogens. J.K. Brown, ed. APS Press, pp. 289-309.

Moritz, G., Kumm, S., & Mound, L. (2004). Tospovirus transmission depends on thrips ontogeny. *Virus Research*, *100*(1), 143–149. https://doi.org/10.1016/j.virusres.2003.12.022

Naghavi, M. H., & Walsh, D. 2017. Microtubule regulation and function during virus infection. *J Virol 91*:e00538-17. https://doi.org/10.1128/JVI.00538-17.

Ohi R, Strothman C, Zanic M. (2021). Impact of the 'tubulin economy' on the formation and function of the microtubule cytoskeleton. *Curr Opin Cell Biol. 68*:81-89. doi: 10.1016/j.ceb.2020.09.005.

Ohnishi, J., Knight, L. M., Hosokawa, D., Fujisawa, I., and Tsuda, S. 2001. Replication of Tomato spotted wilt virus after ingestion by adult *Thrips setosus* is restricted to midgut epithelial cells. *Phytopathology*, *91*:1149-1155. doi: 10.1094/PHYTO.2001.91.12.1149.

Payne, S. H. (2015). The utility of protein and mRNA correlation. *Trends in Biochemical Sciences*, *40*(1), 1–3. https://doi.org/10.1016/J.TIBS.2014.10.010

Pradet-Balade, B., Boulmé, F., Beug, H., Müllner, E. W., Garcia-Sanz, J. A. (2001). Translation control: bridging the gap between genomics and proteomics? Trends Biochem Sci., 26(4):225-9. doi: 10.1016/s0968-0004(00)01776-x.

Rotenberg, D., Jacobson, A. L., Schneweis, D. J., & Whitfield, A. E. (2015). Thrips transmission of tospoviruses. *Current Opinion in Virology*, *15*, 80–89. https://doi.org/10.1016/j.coviro.2015.08.003

Rotenberg, D. & Whifield (2018). Molecular interactions between tospoviruses and thrips vectors, *Current Opinion in Virology*, *33*:191-197. doi.org/10.1016/j.coviro.2018.11.007.

Rotenberg, D., Robertson, H. M., Oliver, J. E., Benoit, J. B., Snoeck, S., Baumann, A. A., Mahmoud, S. ben, Veenstra, J. A., Jacobs, C. G. C., Park, Y., Panfilio, K. A., Widana Gamage, S. M. K., Ahn, S.-J., Vargas Jentzsch, I. M., Jennings, E. C., Szuter, E. M., Christiaens, O., Scanlan, J., Taning, C. N. T., … Warren, J. H. (2019). Frankliniella occidentalis *Official Gene Set OGSv1.0. Ag Data Commons (Database)*. https://doi.org/10.15482/USDA.ADC/1504029

Rotenberg, D., Baumann, A.A., Ben-Mahmoud, S., Christiaens, O., Dermauw, W., Ioannidis, P., Jacobs, C.G.C., Vargas Jentzsch, I.M., Oliver, J.E., Poelchau, M.F., Rajarapu, S.P., Schneweis, D.J., Snoeck, S., Taning, C.N.T., Wei, D., Widana-Gamage, S.M.K., Hughes, D.S.T., Murali, S.C., Bailey, S., Bejerman, N.E., Holmes, C.J., Jennings, E.C., Rosendale, A.J., Rosselot, A., Hervey, K., Schneweis, B.A., Cheng, S., Childers, C., Simão, F.A., Dietzgen, R.G., Chao, H., Dinh, H., Doddapaneni, H., Dugan, S., Han, Y., Lee, S.L., Muzny, D.M., Qu, J., Worley, K.C., Benoit, J.B., Friedrich, M., Jones, J.W., Panfilio, K.A., Park, Y., Robertson, H.M., Smagghe, G., Ullman, D.E., Van Der Zee, M., Van Leeuwen, T., Veenstra, J.A., Waterhouse, R.M., Weirauch, M.T., Werren, J.H., Whitfield, A.E., Zdobnov, E.M., Gibbs, R.A. And Richards, S. (2020). Genome-enabled insights into the biology of thrips as crop pests. BMC Biology 18: 142. https://doi.org/10.1186/s12915-020-00862-9.

Schneweis, D. J., Whitfield, A. E., & Rotenberg, D. (2017). Thrips developmental stage-specific transcriptome response to tomato spotted wilt virus during the virus infection cycle in *Frankliniella occidentalis*, the primary vector. *Virology*, *500*, 226–237. https://doi.org/10.1016/j.virol.2016.10.009

Shi, J., Zhou, J., Jiang, F., Li, Z., Zhu, S. (2023). The effects of the E3 ubiquitin-protein ligase UBR7 of Frankliniella occidentalis on the ability of insects to acquire and transmit TSWV. *PeerJ 11*:e15385. doi: 10.7717/peerj.15385.

Shrestha, A., Champagne, D. E., Culbreath, A. K., Rotenberg, D., Whitfield, A. E., & Srinivasan, R. (2017). Transcriptome changes associated with *Tomato spotted wilt virus* infection in various life stages of its thrips vector, *Frankliniella fusca* (Hinds). *Journal of General Virology*, *98*(8), 2156–2170. https://doi.org/10.1099/jgv.0.000874

Srivastava, H., Lippincott, M. J., Currie, J., Canfield, R., Lam. M. P. Y., & Lau, E. (2022). Protein prediction models support widespread post-transcriptional regulation of protein abundance by interacting partners. *PLoS Comput Biol*, *18*(11): e1010702. https://doi.org/10.1371/journal.pcbi.1010702

Sunnerhagen, P. (2007). Cytoplasmic post-transcriptional regulation and intracellular signalling. Mol Genet Genomics 277, 341–355. doi: 10.1007/s00438-007-0221-5.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., von Mering, C. (2023) The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucl. Acids Res.*, *51*(D1):D638–D646. https://doi.org/10.1093/nar/gkac1000

Tasaki, S., Xu, J., Avey, D.R. et al. (2022). Inferring protein expression changes from mRNA in Alzheimer's dementia using deep neural networks. Nat Commun 13:655. https://doi.org/10.1038/s41467-022-28280-1

Ullman, D. E. (1992). A midgut barrier to tomato spotted wilt virus acquisition by adult western flower thrips. *Phytopathology*, *82*(11), 1333. https://doi.org/10.1094/phyto-82-1333

Ullman, D. E., German, T. L., Sherwood, J. L., Westcot, D. M., & Cantone, F. A. (1993). Tospovirus replication in insect vector cells: immunocytochemical evidence that the nonstructural protein encoded by the S RNA of tomato spotted wilt tospovirus is present in thrips vector cells. *Phytopathology*, *83*(4), 456. https://doi.org/10.1094/phyto-83-456

Ullman, D. E., Westcot, D. M., Chenault, K. D., Sherwood, J. L., German, T. L., Bandla, M. D., Cantone, F. A., & Duer, H. L. (1995). Compartmentalization, intracellular transport, and

autophagy of tomato spotted wilt tospovirus proteins in infected thrips cells. *Phytopathology*, *85*(6), 644–654. https://doi.org/10.1094/Phyto-85-644

Valanne, S., Wang, J.-H., Rämet, M. (2011). The Drosophila Toll signaling pathway. *J Immunol* ,*186* (2): 649–656. https://doi.org/10.4049/jimmunol.1002302

van de Wetering, F., Goldbach, R., & Peters, D. (1996). Tomato spotted wilt tospovirus ingestion by first instar larvae of *Frankliniella occidentalis* is a prerequisite for transmission. *Phytopathology*, *86*(9), 900–905. https://doi.org/10.1094/phyto-86-900

Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics 2012 13:4*, *13*(4), 227–232. https://doi.org/10.1038/nrg3185

Zhao, W., Lu, L., Yang, P., Cui, N., Kang, L., & Cui, F. (2016). Organ-specific transcriptome response of the small brown planthopper toward rice stripe virus. *Insect Biochemistry and Molecular Biology*, *70*, 60–72. https://doi.org/10.1016/j.ibmb.2015.11.009

Zhao, B., Roundtree, I. & He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol 18, 31–42 . doi: 10.1038/nrm.2016.132.

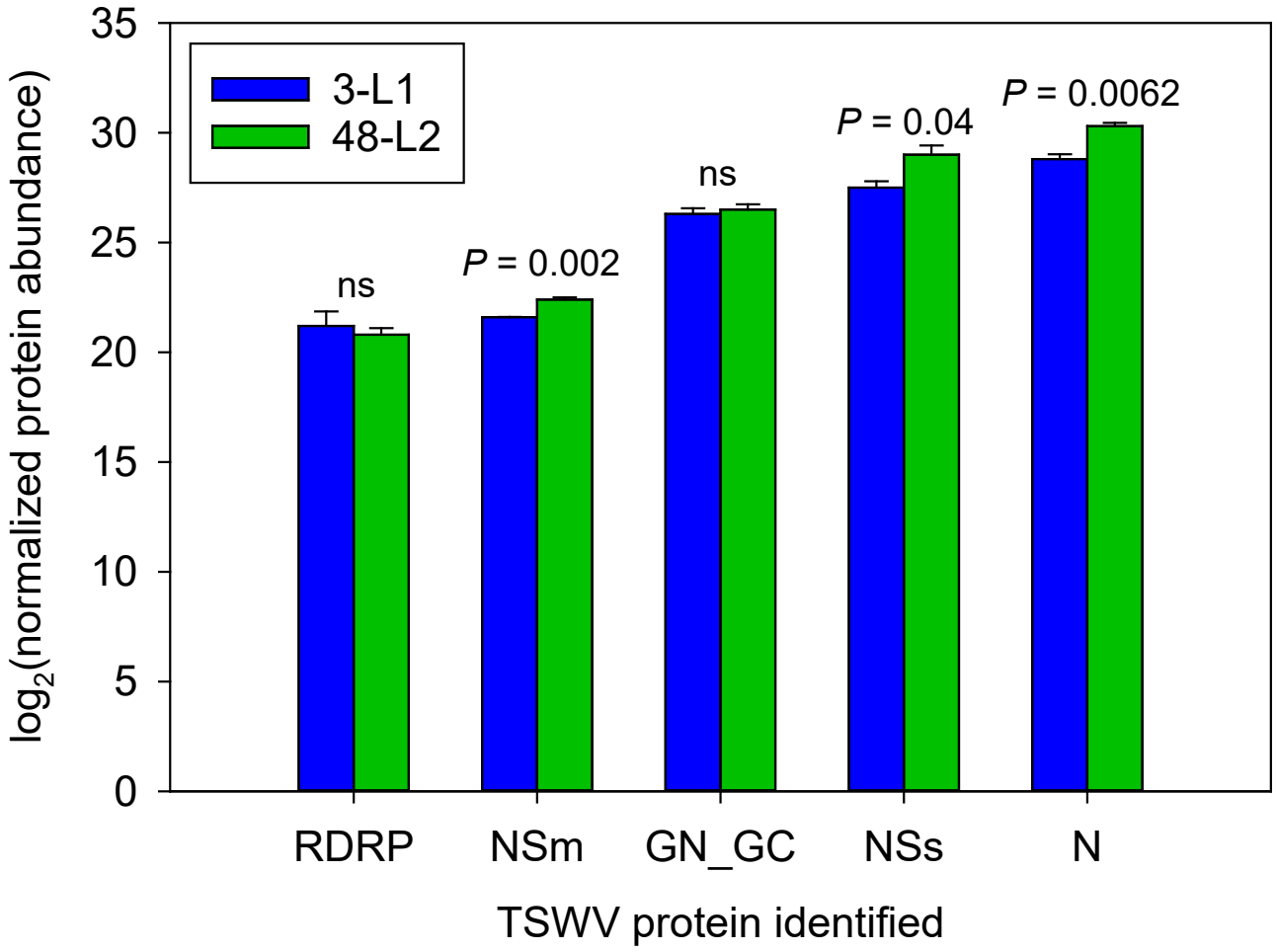
35
30
25
20
15
10
5
0
log2(normalized protein abundance)
3-L1
48-L2
ns
P = 0.002
ns
P = 0.04
P = 0.0062
RDRP
NSm
GN_GC
NSs
N
TSWV protein identified

Figure 1

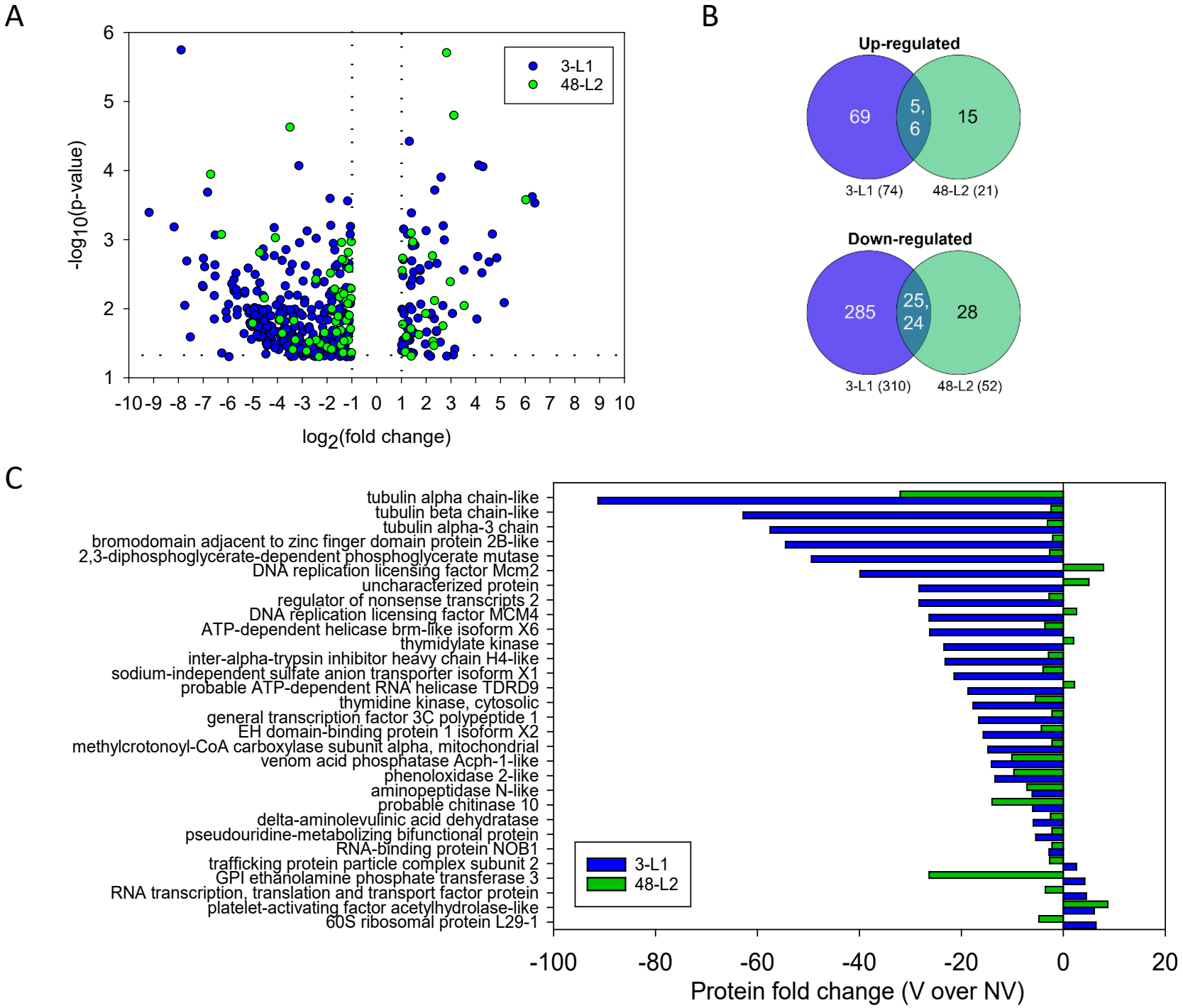
A
3-L1
48-L2
-log10(p-value)
log2(fold change)
B
Up-regulated
69
5, 6
15
3-L1 (74)
48-L2 (21)
Down-regulated
285
25, 24
28
3-L1 (310)
48-L2 (52)
C
tubulin alpha chain-like
tubulin beta chain-like
tubulin alpha-3 chain
bromodomain adjacent to zinc finger domain protein 2B-like
2,3-diphosphoglycerate-dependent phosphoglycerate mutase
DNA replication licensing factor Mcm2
uncharacterized protein
regulator of nonsense transcripts 2
DNA replication licensing factor MCM4
ATP-dependent helicase brm-like isoform X6
thymidylate kinase
inter-alpha-trypsin inhibitor heavy chain H4-like
sodium-independent sulfate anion transporter isoform X1
probable ATP-dependent RNA helicase TDRD9
thymidine kinase, cytosolic
general transcription factor 3C polypeptide 1
EH domain-binding protein 1 isoform X2
methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial
venom acid phosphatase Acph-1-like
phenoloxidase 2-like
aminopeptidase N-like
probable chitinase 10
delta-aminolevulinic acid dehydratase
pseudouridine-metabolizing bifunctional protein
RNA-binding protein NOB1
trafficking protein particle complex subunit 2
GPI ethanolamine phosphate transferase 3
RNA transcription, translation and transport factor protein
platelet-activating factor acetylhydrolase-like
60S ribosomal protein L29-1
3-L1
48-L2
Protein fold change (V over NV)

Figure 2

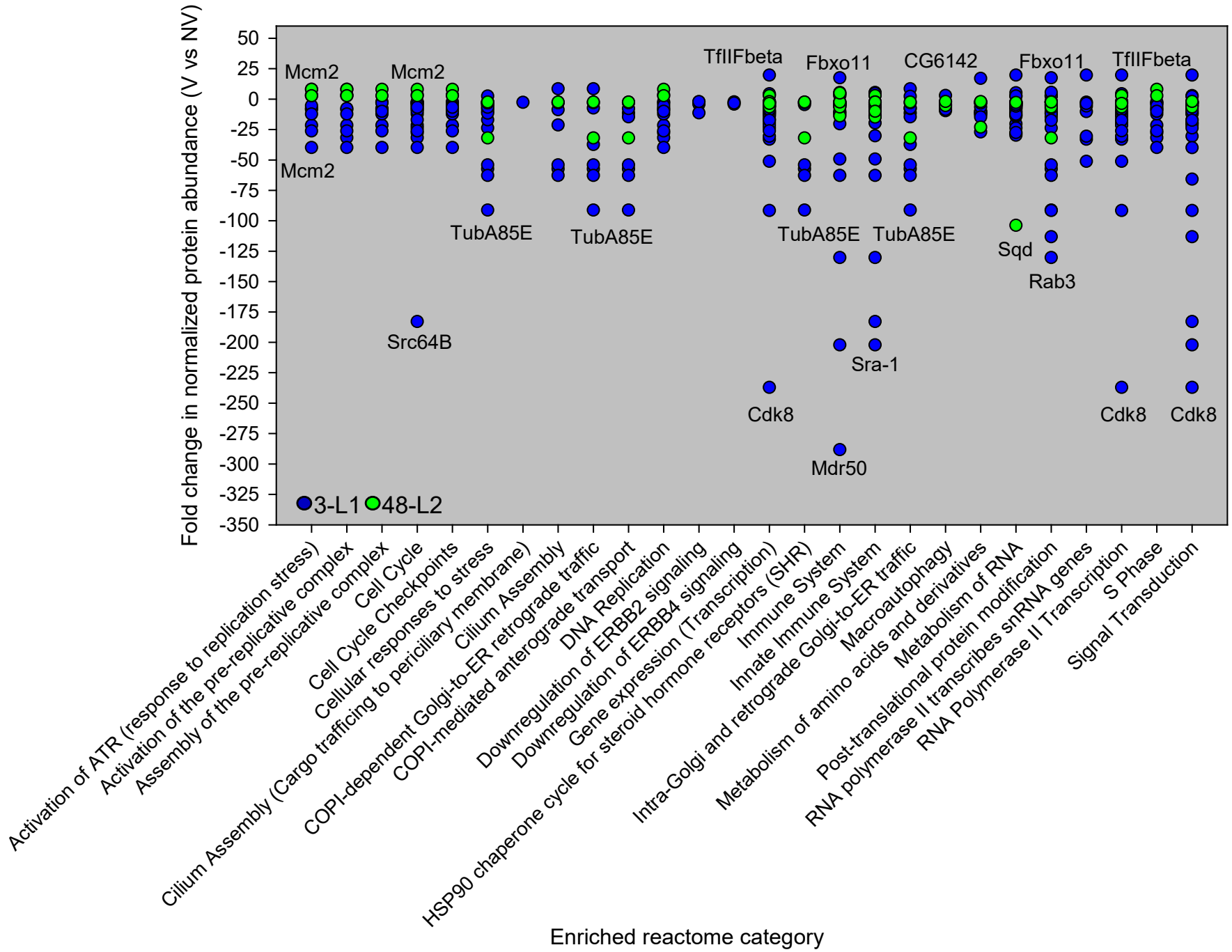
Fold change in normalized protein abundance (V vs NV)
50
25
0
-25
-50
-75
-100
-125
-150
-175
-200
-225
-250
-275
-300
-325
-350
Mcm2
Mcm2
Mcm2
TfIIFbeta
Fbxo11
CG6142
Fbxo11
TfIIFbeta
TubA85E
TubA85E
TubA85E
TubA85E
Sqd
Rab3
Src64B
Sra-1
Cdk8
Cdk8
Cdk8
Mdr50
3-L1
48-L2
Activation of ATR (response to replication stress)
Activation of the pre-replicative complex
Assembly of the pre-replicative complex
Cell Cycle
Cell Cycle Checkpoints
Cellular responses to stress
Cilium Assembly (Cargo trafficing to periciliary membrane)
Cilium Assembly
COPI-dependent Golgi-to-ER retrograde traffic
COPI-mediated anterograde transport
DNA Replication
Downregulation of ERBB2 signaling
Downregulation of ERBB4 signaling
Gene expression (Transcription)
HSP90 chaperone cycle for steroid hormone receptors (SHR)
Immune System
Innate Immune System
Intra-Golgi and retrograde Golgi-to-ER traffic
Macroautophagy
Metabolism of amino acids and derivatives
Metabolism of RNA
Post-translational protein modification
RNA polymerase II transcribes snRNA genes
RNA Polymerase II Transcription
S Phase
Signal Transduction
Enriched reactome category

Figure 3

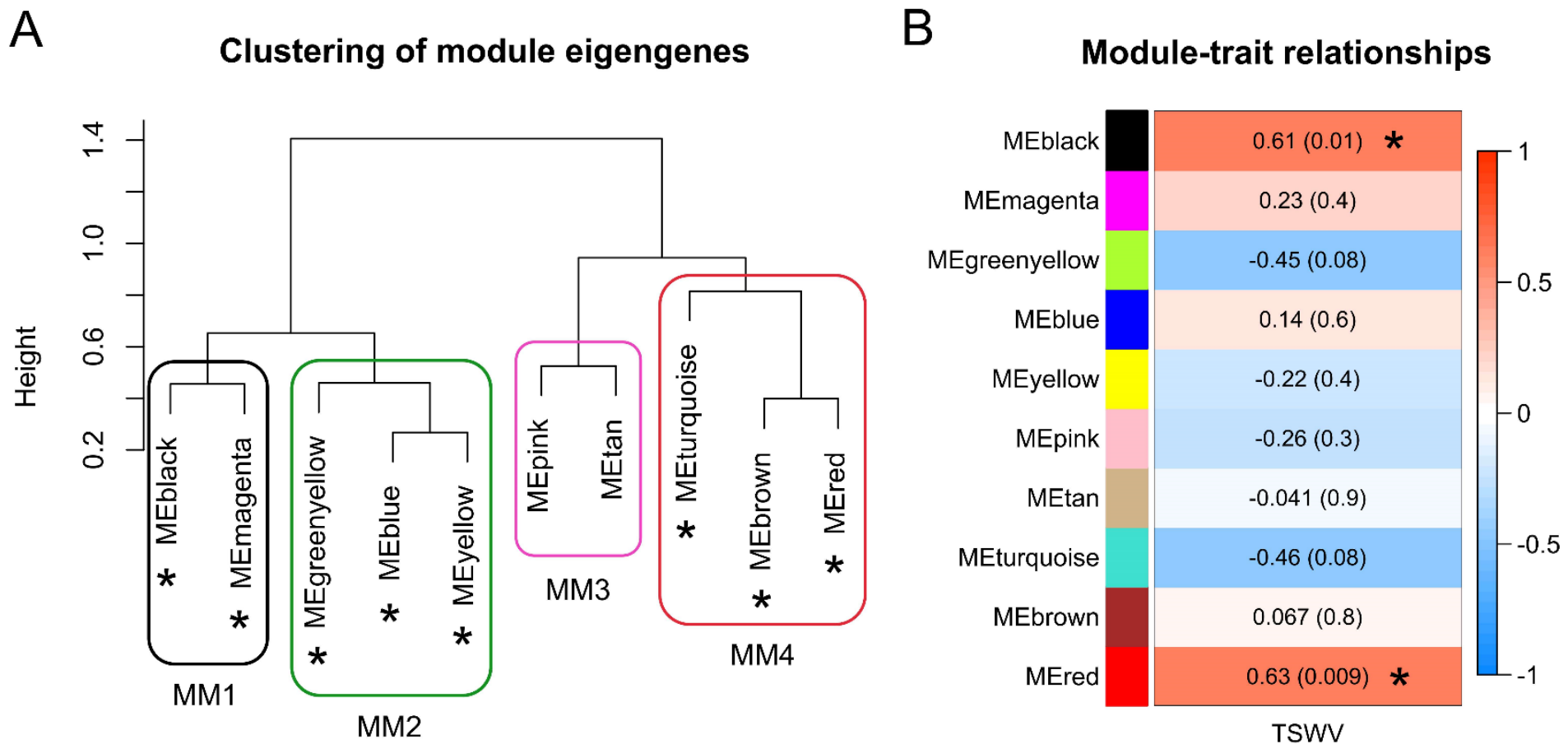
A
Clustering of module eigengenes
Height
0.2
0.6
1.0
1.4
MEblack
MEmagenta
MEgreenyellow
MEblue
MEyellow
MEpink
MEtan
MEturquoise
MEbrown
MEred
MM1
MM2
MM3
MM4
B
Module-trait relationships
MEblack 0.61 (0.01) *
MEmagenta 0.23 (0.4)
MEgreenyellow -0.45 (0.08)
MEblue 0.14 (0.6)
MEyellow -0.22 (0.4)
MEpink -0.26 (0.3)
MEtan -0.041 (0.9)
MEturquoise -0.46 (0.08)
MEbrown 0.067 (0.8)
MEred 0.63 (0.009) *
TSWV
1
0.5
0
-0.5
-1

Fig 4

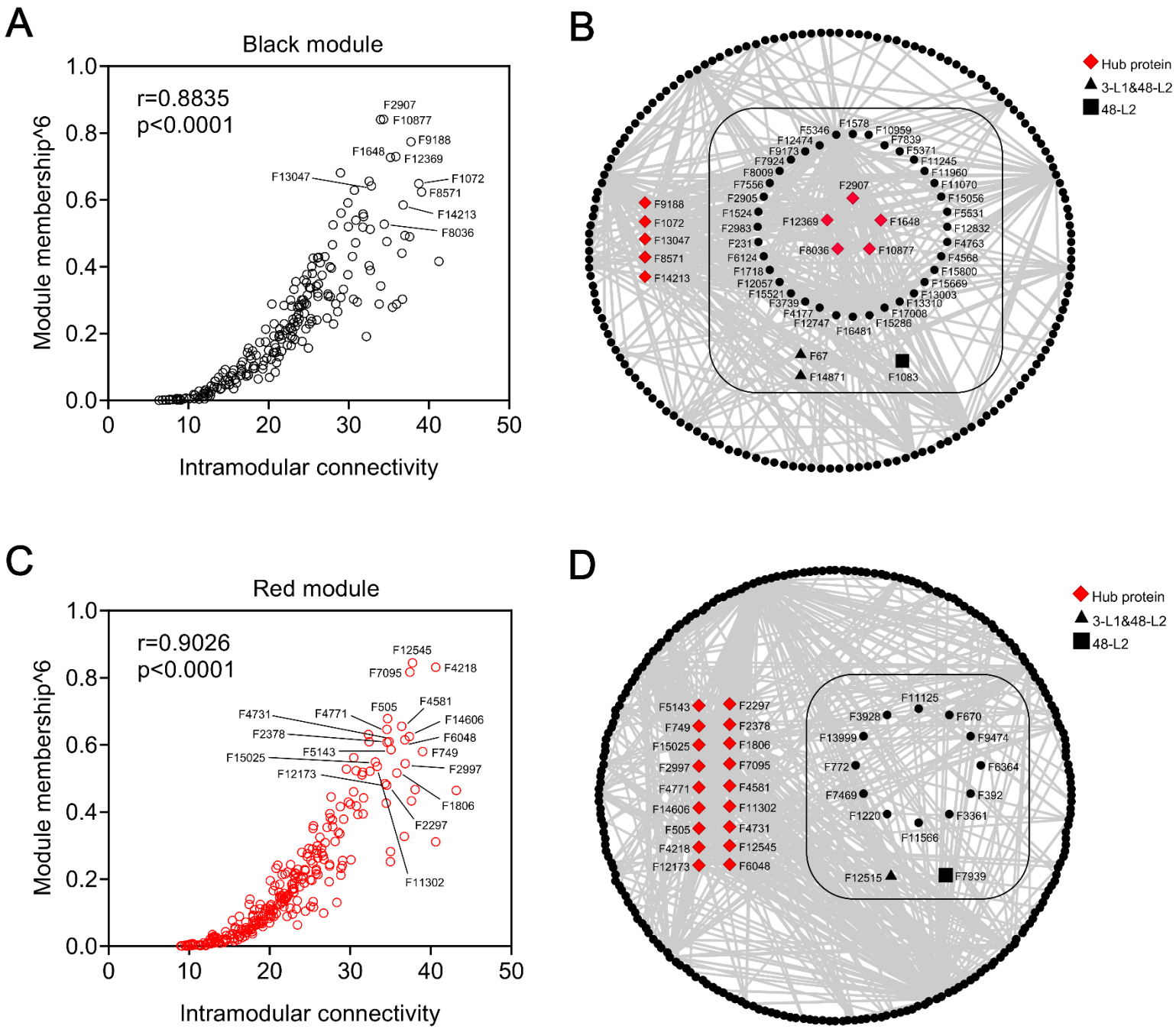
A
Black module
r=0.8835
p<0.0001
Module membership^6
Intramodular connectivity
F2907
F10877
F9188
F1648
F12369
F13047
F1072
F8571
F14213
F8036
B
Hub protein
3-L1&48-L2
48-L2
F5346
F1578
F10959
F12474
F7839
F9173
F5371
F7924
F11245
F8009
F11960
F7556
F2907
F11070
F2905
F15056
F1524
F5531
F2983
F12369
F1648
F12832
F231
F4763
F8036
F10877
F6124
F4568
F1718
F15800
F12057
F15669
F15521
F13003
F3739
F13310
F4177
F17008
F12747
F15286
F16481
F67
F14871
F1083
C
Red module
r=0.9026
p<0.0001
Module membership^6
Intramodular connectivity
F12545
F7095
F4218
F4581
F505
F4731
F4771
F14606
F2378
F6048
F5143
F749
F15025
F2997
F12173
F1806
F2297
F11302
D
Hub protein
3-L1&48-L2
48-L2
F5143
F2297
F749
F2378
F15025
F1806
F2997
F7095
F4771
F4581
F14606
F11302
F505
F4731
F4218
F12545
F12173
F6048
F11125
F3928
F670
F13999
F9474
F772
F6364
F7469
F392
F1220
F3361
F11566
F12515
F7939

Fig 5

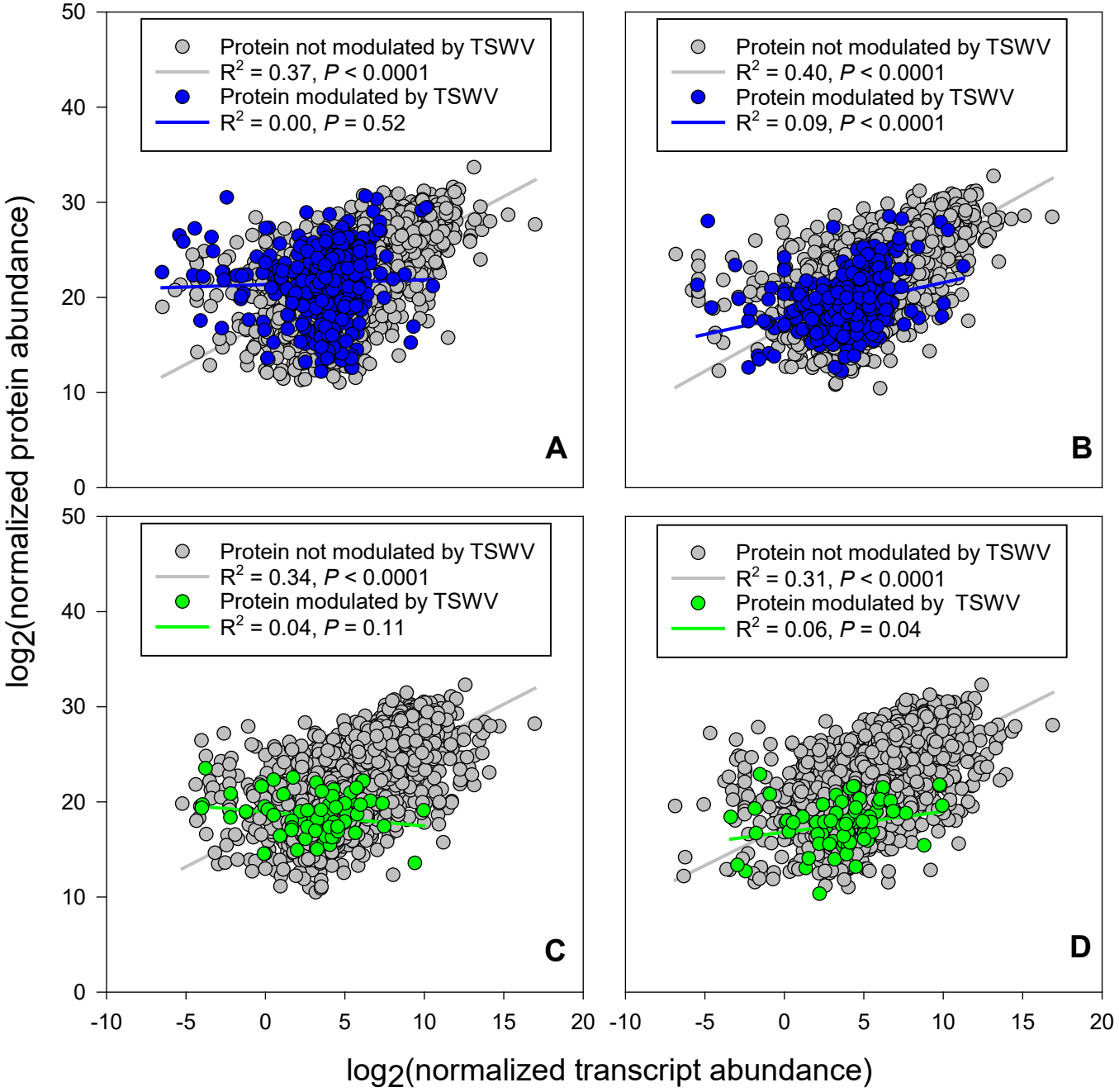
Protein not modulated by TSWV
R2 = 0.37, P < 0.0001
Protein modulated by TSWV
R2 = 0.00, P = 0.52
A
Protein not modulated by TSWV
R2 = 0.40, P < 0.0001
Protein modulated by TSWV
R2 = 0.09, P < 0.0001
B
Protein not modulated by TSWV
R2 = 0.34, P < 0.0001
Protein modulated by TSWV
R2 = 0.04, P = 0.11
C
Protein not modulated by TSWV
R2 = 0.31, P < 0.0001
Protein modulated by TSWV
R2 = 0.06, P = 0.04
D
log2(normalized protein abundance)
log2(normalized transcript abundance)
0
10
20
30
40
50
-10
-5
0
5
10
15
20

Figure 6

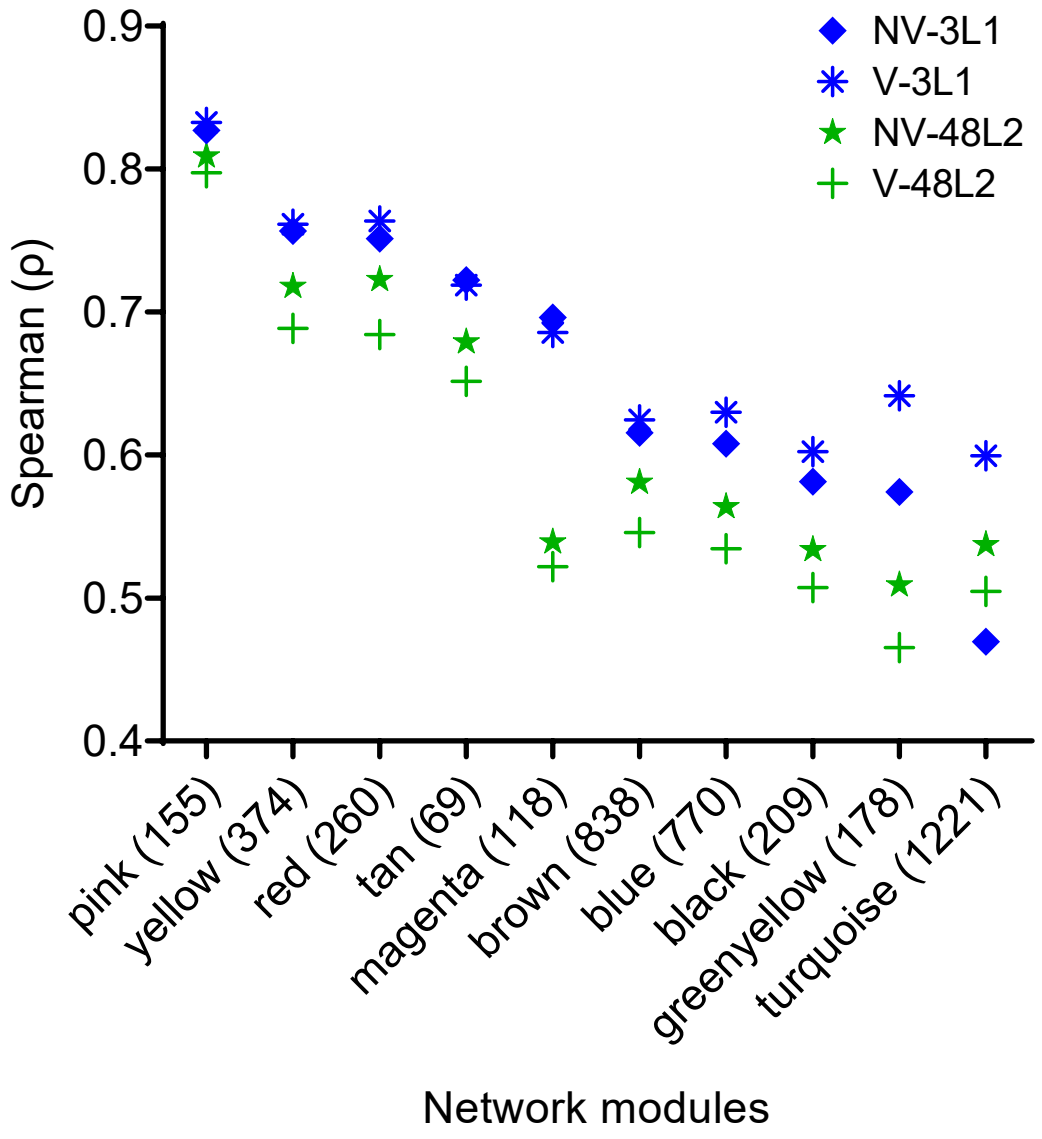
0.9
0.8
0.7
0.6
0.5
0.4
Spearman (ρ)
NV-3L1
V-3L1
NV-48L2
V-48L2
pink (155)
yellow (374)
red (260)
tan (69)
magenta (118)
brown (838)
blue (770)
black (209)
greenyellow (178)
turquoise (1221)
Network modules

Figure 7