Quantitative isotope-labeled crosslinker proteomics reveals developmental variation in protein interactions and posttranslational modifications in *Diaphorina citri*, the citrus greening insect vector

John S. Ramsey^{1,2}, Xuefei Zhong³, Surya Saha^{2,^}, Juan D. Chavez³, Richard Johnson³, Jaclyn E. Mahoney², Andrew Keller³, Kathy Moulton⁴, Lukas A. Mueller², David G. Hall⁴, Michael J. MacCoss³, James E. Bruce³, Michelle Heck^{1,2,5*}

¹Emerging Pests and Pathogens Research Unit, United States Department of Agriculture, Ithaca, NY, USA

²Boyce Thompson Institute, Ithaca, NY, USA

³Department of Genome Sciences, University of Washington, Seattle, WA, USA

⁴Subtropical Insects and Horticulture Research Unit, United States Department of Agriculture, Ft. Pierce, FL, USA

⁵Plant Pathology and Plant-Microbe Biology Section, School of Integrated Plant Sciences, Cornell University, Ithaca, NY, USA

[^]Current affiliation: School of Animal and Comparative Biomedical Sciences, 1117 E. Lowell Street, Tucson AZ 85721 USA

*Corresponding Author; email address: mlc68@cornell.edu

Running title: Protein interactions and PTMs during *Diaphorina citri* development

Abstract

Acquisition of the citrus greening bacterial pathogen, ‘*Candidatus Liberibacter asiaticus*’ (CLas) by Asian citrus psyllid (*Diaphorina citri*) nymphs is required efficient tree-to-tree transmission during the adult stage. Quantitative isotope-labeled protein interaction reporter (PIR) cross-linkers were used in parallel with protein quantification using spectral counting to quantify protein interactions within microbe-enriched cellular fractions of nymph and adult *D. citri*. Over 100 unique crosslinks were found between five insect histone proteins, and over 30% of these were more abundant in nymph compared to adult insects. Strikingly, some cross-links detected in *D. citri* proteins are conserved in cross-linking studies on human cells, suggesting these protein interaction topologies were present in the common ancestor (~750MYA) or are subject to convergent evolution. Analysis of posttranslational modifications of crosslinked histones revealed the presence of acetylated and methylated lysine residues, which may impact psyllid chromatin structure and gene expression. Histone H3 peptides acetylated in the N terminal tail region were found to be more abundant in nymph compared to adult insects in two orthogonal proteomics methods. The insect life stage-specific histone posttranslational modifications and protein interactions represent physical evidence that metamorphosis is associated with changes in chromatin structure that regulate genome-wide transcriptional reprogramming.

Introduction

Insect transmission of bacterial and viral pathogens is responsible for the spread of numerous agricultural pathogens that cause diseases of economic significance (1, 2). The Asian citrus psyllid (ACP, *Diaphorina citri*), is one of two psyllid species responsible for the worldwide spread of ‘*Candidatus Liberibacter asiaticus*’ (CLas), the gram negative, fastidious bacteria responsible for huanglongbing (HLB), also referred to as citrus greening disease. CLas

is a gram-negative α -proteobacteria which is not amenable to laboratory culture by standard methods, but which is capable of replicating in the disparate environments of citrus phloem sap and tissues of the insect vector (3, 4). Evidence on the mode of CLas transmission is consistent with a circulative, propagative transmission pathway in *D. citri*, whereby bacteria acquired into the gut of the insect vector from feeding on an infected tree must circulate and replicate within the insect body and enter the salivary glands in order to be transmitted to a healthy tree as a component of the insect's saliva (5). While both nymph and adult insects are capable of acquiring CLas into their guts, the bacterium can only be transmitted by adult insects that have acquired the pathogen during their nymphal stage (6). The weaker nymphal immune response may enable the bacteria to readily cross the gut transmission barrier in nymphs, which is fortified and difficult to penetrate in adults (7).

Analysis of the role of genes, transcripts, proteins, and metabolites in CLas transmission by *D. citri* may reveal novel approaches to disease control (8, 9). High-throughput genome and transcriptome analysis has provided details of gene sequences and expression levels in a wide range of insect vectors (10-12). Proteomics analyses rely on databases of proteins predicted from genomic and transcriptomic data, and application of mass spectrometry workflows to the study of vector biology has enabled the identification of proteins and metabolites that are present at high abundance in specific transmission barrier tissues or in response to experimental treatments such as host plant infection (13-16). Mass spectrometry-based workflows designed to identify protein-protein interactions provide complementary data to quantitative proteomics studies, revealing details of protein intramolecular structure. The conformation of protein complexes which may enable the development of targeted strategies to block pathogen transmission by insect pests (17, 18).

Crosslinking mass spectrometry (XL-MS) using protein interaction reporter (PIR) technology has been used for *in vivo* identification of specific regions of interaction between protein domains and within protein complexes in cell and tissue samples (17, 19, 20). A prior study combining PIR and quantitative proteomics identified protein interactions between adult *D. citri* proteins and bacterial proteins. Quantitative proteome analysis of nymph and adult insects from CLas(-) and CLas(+) colonies revealed a much greater impact of infection status on adult as compared to nymphal-stage insects (18). The less robust immune response to CLas observed in nymphs compared to adults may be related to the fact that CLas acquisition by nymphs is critical to enabling subsequent transmission by adults. The number of differentially abundant proteins identified between nymph and adult insects was several-fold greater than between CLas(-) and CLas(+) colonies within either developmental stage (18). More than 2000 proteins (nearly half of all identified proteins) were identified by spectral counting as differentially abundant between nymph and adult insects, of which several hundred were undetected at one developmental stage or the other (18). This is consistent with prior studies documenting genome-wide transcriptional changes regulated by changes in chromatin accessibility and remodeling associated with insect metamorphosis (21). The PIR results from that study yielded only very few cross-linked peptides between *D. citri* and CLas proteins (18). Peptides from CLas proteins predicted to catalyze the first two steps in the conversion of pantothenate to coenzyme A were found to be crosslinked to the *D. citri* proteins myosin and hemocyanin 1, the latter of which has extensively documented functions in arthropod metabolism and immunity (22-24).

In the current study, we attempted to increase the identification of microbe involved cross-linked peptides by using a percoll gradient centrifugation strategy to produce cellular fractions from CLas(+) nymph and adult *D. citri*. These fractions were treated with isotopic

variant PIR crosslinkers developed for quantitative PIR analysis, identical except for eight deuterium/hydrogen substitutions, were used to separately crosslink adult and nymph samples that were subsequently combined into a single sample for combined mass spectrometry analysis (25). Quantitative analysis of differences in crosslink abundance between developmental stages was performed in biological and technical replicate samples. In parallel, quantitative proteomics workflows were used to analyze the same Percoll fractions from CLas(+) and CLas(-) adult and nymph insects, and statistical analysis of spectral counting data revealed proteins whose abundance varied with developmental stage and/or infection status. Together, these data types allow us to distinguish between the developmental regulation of protein interaction formation as a result of changes in protein abundance or because of a change in protein interaction topology and protein complex formation during insect development.

Experimental Section

Insect Rearing

A colony of *D. citri* was continuously reared on CLas-infected *Citrus medica* at the USDA Agricultural Research Service Horticultural Research Laboratory in Fort Pierce, Florida, for several generations. Monthly quantitative PCR (qPCR) tests were performed on to determine the percent infection and average CLas Cq values as described (5). CLas(+) adult and nymph psyllids were collected from colonies in which the percent infection was 80%, with an average Cq value of 27.5. QPCR was performed on triplicate Percoll gradient fractionations to assess the distribution of CLas titer in the fractionation scheme (**Table 1**). Analysis of the *D. citri* wingless gene was performed as described (26) to assess the distribution of *D. citri* cells in the fractions (**Table 1**).

Experiment Design and Statistical Rationale

A total of 8 insect samples [4 biological replicates each adult CLas(+) and nymph CLas(+) samples] were fractionated using Percoll density centrifugation and the CLas-enriched fractions from each replicate were pooled and analyzed using the Quantitative Protein Interaction Reporter (qPIR) workflow. Each biological replicate comprises a population of insects (500 adults per sample, 750 nymphs per sample) to capture the biological variation within insect populations. Four biological replicates for each sample type (nymph/adult) were used to account for variability in sample preparation (**Figure 1**). Triplicate injection replicates were performed to account for analytical variability in sampling PIR labeled peptides. For each peptide crosslink identified in pooled nymph/adult samples, the 95% confidence interval was calculated for the mean log₂ ratio of nymph/adult crosslink abundance. A crosslink was identified as more abundant in nymph or adult samples when the {[absolute value of crosslink log₂ ratio] – [1/2 (95% CI)]} ≥ 1.

A total of 12 insect samples [3 biological replicates each adult CLas(+), adult CLas(-), nymph CLas(+), nymph CLas(-)] were taken through the Percoll fractionation and the same fractions as above were analyzed using data dependent acquisition and quantitative proteomics using spectral counting (Figure 1). The Fisher's exact test using peptide spectral counting was used to identify differentially abundant proteins in pairwise analyses between the four sample classes. Benjamini-Hochberg p-value adjustment for multiple testing was applied to determine threshold for differential abundance.

132

Figure 1. Experimental design of *Diaphorina citri* Quantitative Protein Interaction Reporter (PIR) and Quantitative Proteomics (Spectral Counting) studies. In the PIR study, percoll gradient cellular samples prepared from nymph and adult *D. citri* were treated with light or heavy crosslinkers (Light-XL, Heavy-XL) which generate unique mass profiles. Nymph and adult samples (four each) treated with Light-XL and Heavy-XL were pooled prior to peptide sample preparation and mass spectrometry analysis, and differences in crosslink abundance between nymph and adult samples were identified in replicate reciprocally labeled samples. In the Spectral Counting study, three replicates each of CLas(+)/CLas(-) nymph and adult percoll fraction samples were analyzed by mass spectrometry, and differentially abundant proteins were identified by pairwise statistical analysis between sample classes.

133

134 *Quantitative Protein Interaction Reporter (PIR) sample preparation and analysis*

135 Adult (~500 insects per sample) and nymph (~ 750 insects per sample) *D. citri* collected
 136 from CLas infected colonies were homogenized and fractionated by Percoll density
 137 centrifugation (14). Fractions 3, 4, 5 and 6 from each sample were pooled and suspended in a
 138 buffered solution consisting of 50 mM Na₂HPO₄ and 0.85 % (w/v) NaCl pH 7.4 for subsequent
 139 crosslinking reaction. Four adult samples and four nymph samples were crosslinked with 10 mM
 140 of light or heavy PIR cross-linker (Biotin Aspartate Proline-N-hydroxyphthalamide, BDP-NHP)

at room temperature for 1 hour (25). Within each developmental group, two samples were crosslinked with light PIR crosslinker (d_0 -BDP-NHP) and two samples were crosslinked with heavy PIR crosslinker (d_8 -BDP-NHP). After the reaction, cells in the reaction mixture were lysed by incubation with equal volume of 8 M urea in 0.1 M Tris buffer overnight and ultrasonication. Protein concentration was determined by Bradford assay. Equimolar nymph light-XL and adult heavy-XL (and reciprocally labeled) protein samples were combined for a single injection into the mass spectrometer, enabling relative quantitation of crosslinks in different sample types within a single instrument run. The pooled protein lysates were reduced with 5 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 minutes and alkylated with 10 mM iodoacetamide (IAA) for 45 minutes. Samples were diluted with 0.1 M ammonium bicarbonate buffer (pH=8) to urea concentration below 1 M, followed by overnight trypsin (Promega, WI) digestion at 37°C with 1:200 (enzyme: total protein) ratio. The digested mixtures were desalted with 1cc Sep-Pak C18 solid phase extraction (SPE) cartridges (Waters, MD), dried in vacuum centrifuge, and fractionated by strong cation exchange (SCX) HPLC (250 × 10 mm, Luna 5 µm diameter, 100 Å pore size particles, Phenomenex, Torrance, CA). The peptides mixture were eluted with a binary solvent gradient (Solvent A, 7 mM KH_2PO_4 (pH 2.6)/ 30% ACN, Solvent B, 7 mM KH_2PO_4 (pH 2.6)/350 mM KCl) at 1.5 mL/min mobile phase flow rate (0–2.5 min, 0% B; 2.5–7.5 min, 0–5% B; 7.5–47.5 min, 5–60% B; 47.5–67.5 min, 60–100% B; 67.5–77.5 min, 100% B). Five fractions of eluates (42.5–52.5min, 52.5–57.5 min, 57.5–62.5 min, 62.5–67.5min, 67.5–87.5 min) were collected, and further dried by vacuum centrifuge, resuspended in 5 mL 0.1 M NH_4HCO_3 , and pH adjusted to ~7.4. To enrich the crosslinked peptides, each collected fractions was incubated with ~200 µL UltraLink monomeric avidin beads slurry (Thermo Pierce, Rockford, IL) for 30 min. The crosslinked peptide pairs captured by the beads were eluted with 0.5 mL of

70% ACN with 0.5% formic acid. The samples were dried and resuspended in 0.1% formic acid prior to nano LC–MS analysis.

Samples from each SCX fraction were analyzed in triplicate by a Waters NanoAcquity UPLC coupled to a Thermo Velos Fourier transform ion cyclotron resonance mass spectrometer implemented with Real-time Analysis for Cross-linked peptide Technology (ReACT) (27). The cross-linked peptide pairs were separated by a in-housed packed 40 cm × 75 µm C8 column (5 µm diameter, 100 Å pore size MichromMagic beads) with a 4 hour 10–30% B gradient (solvent A: 0.1% formic acid in water, solvent B: 0.1% formic acid in acetonitrile, flow rate 300 nL/min). High resolution full MS survey scan and CID MS/MS scans for ions bearing more than 3 charges were acquired by ICR cells at 50 k resolving power at m/z 400, to examine cross-linked peptide pair candidates in real-time. The presence of reporter ion mass (m/z 752.4129) in the CID MS/MS spectra, along with fragment ion masses satisfying the PIR mass relationship (precursor mass = reporter ion mass + peptide α mass + peptide β mass) within 20 ppm of mass tolerance among the top 500 fragment ions in the MS/MS spectra suggests the precursor ion as a potential candidate of cross-linked peptide. Real time Analysis for Cross-linked peptide Technology (ReACT) triggers subsequent low resolution CID MS/MS/MS scans of two linear peptides released from the cross-linked peptide pair candidate ion ‘on-the-fly’ (27). For each released peptide ion from one cross-linked peptide pair precursor, two MS/MS/MS scans were acquired by the linear ion trap with normalized collision energy of 35.

PIR Data analysis

A database consisting of 4439 *Diaphorina citri* protein sequences (ftp://ftp.citrusgreening.org/manuscripts/Ramsey2020/PIR_database.fasta) identified from DDA-MS of CLas infected adult and nymph insects, 1012 *Candidatus Liberibacter asiaticus*

sequences, 1032 *Wolbachia* sequences, 352 *Profftella* sequence, and 196 *Carsonella* sequences was used for identification of linear peptide sequences released from cross-linked peptide pairs. MS³ spectra from ReACT analysis were searched against the database containing both 7031 forward and 7031 reverse protein sequences by Comet (version 2017.01 rev. 4). The linear peptide sequences were then assembled to cross-linked peptide pairs according to the mass relationships recorded during the ReACT data acquisition. Comet search was conducted with the following parameters: precursor peptide mass tolerance 20 ppm, allowing for -1, 0, + 1, + 2, or +3 ¹³C offsets; fragment ion mass tolerance 1.005 Da with 0.4 Da offset; static modification, carbamidomethylation of Cys (57.0215 Da); variable modifications including Met oxidation (15.9949 Da), BDP stump modification (197.0324 Da for d₀-BDP-NHP, 201.0597 Da for d₈-BDP-NHP) of Lys and protein N-termini, where ‘stump’ refers to the residual group from the crosslinker attached to lysine side chain or peptide N-terminus after CID cleavage, Lys and Arg methylation (14.0156 Da), Lys dimethylation (28.0313 Da), Lys trimethylation (42.0470 Da), Lys acetylation (42.0106 Da); up to three missed tryptic cleavages allowed. All the target and decoy MS³ peptide-spectrum-matches containing an internal Lys modified by BDP stump were mapped back to the cross-linked relationships with a 20 ppm tolerance. XLinkProphet (28) was applied to model crosslink specific discriminating information, and the results were filtered at 1% FDR of unique crosslinks corresponding to minimum composite probability 0.63.

MassChroQ (29) was used for MS¹ peak area integration of all crosslinked peptide pairs identified. LC-MS/MS runs acquired from the same SCX fractions were grouped for retention time alignment using the Obiwrap method. Accurate precursor masses of the most abundant three isotopes, charge state, and MS² scan number of each peptide pair-spectra-match (PSM) of crosslinked peptide pair were used as input for MassChroQ. Peak areas under curves from

extracted ion chromatograms were integrated for all input PSMs with 20 ppm mass tolerance of precursor ion using the Zivy peak detection algorithm (29). $\text{Log}_2(\text{adult/nymph})$ ratio for each isotope of each PSM was calculated from peak area output by MassChroQ (29), and normalized by the median ratio within each biological replicates. All ratios belonging to one non-redundant crosslinked pair were grouped, the average ratio after outlier rejection was assigned as the ratio of the specific crosslinked peptide pair. The 95% confidence interval was computed for each non-redundant crosslink.

Quantitative proteomics sample preparation and analysis

Three biological replicate samples of adult and nymph *D. citri* (~ 200 insects per sample) collected from CLas-infected colonies were homogenized and fractionated by Percoll density centrifugation (14). The same four fractions collected for PIR analysis were used for protein extraction and peptide sample preparation, with fractions 3/4 and fractions 5/6 prepared processed separately for each biological sample. Protein precipitation solvent (10% trichloroacetic acid in acetone with 2% beta-mercaptoethanol) was made fresh and kept on ice until use. For Percoll gradient fraction protein extraction, one mL precipitation solvent was added to purified cells and the sample was sonicated to rupture cells. Several rounds of sonication (30 seconds, 15% amplitude, Branson digital probe sonicator, Danbury, CT) were needed to fully disrupt pellet, with samples chilled on ice between rounds of sonication to avoid heating. Protein samples were precipitated overnight in the -20°C freezer. Precipitated protein pellets were washed three times with one mL ice-cold acetone, and after decanting the supernatant from the final wash the pellet was dried to completion. Pellets were resuspended in 200 μL protein reconstitution solvent [8 M urea, 50 mM triethylammonium bicarbonate (TEAB) in water]. Pellets were reconstituted overnight with shaking. Samples were centrifuged at full

233 speed to pellet insoluble material, and the supernatant was collected for protein analysis. Quick
234 Start Bradford Protein Assay (Bio-Rad, Hercules, CA) was used to determine concentration of
235 protein samples. Based on Bradford assay data, 3.5 µg of protein from Percoll gradient fraction
236 samples was used for peptide sample preparation. For reduction, a 100 µL solution of each
237 protein sample was prepared in 10 mM dithiotreitol, 50 mM TEAB, and samples were vortexed,
238 centrifuged briefly, and incubated for one hour at 30°C. After cooling samples to room
239 temperature, methyl methanethiosulfonate (MMTS) in 50 mM TEAB was added to a final
240 concentration of 30 mM MMTS, samples were vortexed, centrifuged, and incubated for one hour
241 at room temperature in the dark. Samples were diluted with TEAB if necessary to reduce
242 concentration of urea below 1 M. Sequencing grade, modified trypsin (Promega, Madison, WI)
243 was added to protein samples (target ratio of trypsin: protein is between 1:20-1:100 by weight).
244 Samples were vortexed, centrifuged, and incubated at 30°C overnight.

245 Oasis MCX solid phase extraction cartridges (Waters, Milford, MA), 30 µm particle size,
246 were used to remove contaminants from the digested peptides. Trypsin-digested peptide samples
247 were dried in a vacuum concentrator and reconstituted in 0.1% formic acid in water.
248 Concentrated formic acid (5 µL) was added to peptide samples, which were spotted on litmus
249 paper to confirm pH>3. Cartridges were conditioned with 1) 1 mL 100% methanol, 2) 1 mL 3%
250 ammonium hydroxide in water, 3) 2 mL 100% methanol, 4) 3 mL 0.1% formic acid in water.
251 Samples were added to the column, salts were washed with 1 mL 0.1% formic acid in water,
252 neutrals were washed with 1 mL 0.1% formic acid in methanol, and peptides were eluted in 1
253 mL 3% ammonium hydroxide in methanol and dried down in a speed-vac.

254 Mass spectrometry was performed on a Fusion (Thermo Fisher Scientific) mass
255 spectrometer. Judging from the mass spectral response and total protein estimates, about one

microgram of each sample digest was loaded from the autosampler onto a 150- μ m Kasil fritted trap packed with Reprosil-Pur C18-AQ (3- μ m bead diameter, Dr. Maisch) to a bed length of 2 cm at a flow rate of 2 μ l/min. After loading and desalting using a total volume of 10 μ l of 0.1% formic acid plus 2% acetonitrile, the trap was brought on-line with a pulled fused-silica capillary tip (75- μ m i.d.) packed with the same Reprosil C18-AQ that was mounted in an in-house constructed microspray source and placed in line with a Waters Nanoacquity binary UPLC pump plus autosampler. Peptides were eluted off the column using a gradient of 2-25% acetonitrile in 0.1% formic acid over 100 minutes, followed by 25-60% acetonitrile over 40 minutes at a flow rate of 250 nl/min.

The mass spectrometer was operated using electrospray ionization (2 kV) with the heated transfer tube at 275 C using data dependent acquisition (DDA) in the so-called “Top Speed” mode, whereby one orbitrap mass spectrum (m/z 400-1600 with quadrupole isolation) was acquired with multiple linear ion trap tandem mass spectra every three seconds or less. The resolution for MS in the orbitrap was 120,000 at m/z 200, and for MS/MS the linear ion trap provided unit resolution. The automatic gain control targets for MS in the orbitrap was $2e5$, whereas for MS/MS it was $1e4$, and the maximum fill times were 20 and 35 msec, respectively. The MS/MS spectra were acquired using quadrupole isolation with an isolation width of 1.6 m/z and HCD collision energy (NCE) of 30%. The precursor ion threshold intensity was set to $1e4$ in order to trigger an MS/MS acquisition. Furthermore, MS/MS acquisitions were allowed for precursor charge states of 2-7. Dynamic exclusion (including all isotope peaks) was set for 5 seconds using monoisotopic precursor selection with a mass error of 15 ppm. The fragment ions were analyzed in the linear trap using the “rapid” scan rate.

Mass spectrometry data files were searched against a combined database of predicted *D. citri*, CLas (GCF_000023765.2), and endosymbiont [*Candidatus* Carsonella ruddii] DC (GCF_000441575.1), '*Candidatus* Proffotella armatura' (GCF_000441555.1), and *Wolbachia pipientis* (strain *Wolbachia* wDi, GCF_000331595.1], proteins using Mascot Daemon 2.5.1 (Matrix Science, Boston, MA). *D. citri* proteins from NCBI GNOMON annotation 100 of the diaci version 1.1 genome and MCOT transcriptome (30) were clustered with cd-hit (31) to remove redundancy (parameters: -c 1.00 -d 0 -n 5). The resultant *D. citri* protein set was merged with Carsonella, Proffotella, *Wolbachia*, and CLas to create the combined database (http://ftp.citrusgreening.org/manuscripts/Ramsey2020/Gnomon_MCOTahrd_clustered_endosymbionts.faa). MS/MS search parameters included fixed modifications (cysteine: methylthio), variable modifications (lysine: methylation, dimethylation, acetylation; arginine: acetylation; asparagine, glutamine: deamidation; methionine: oxidation), and maximum one missed cleavage. Thermo *.raw files were converted into Mascot Generic Format using MSConvert in ProteoWizard (32). Files with the *.dat extension were exported from Mascot and loaded into Scaffold Q+ 4.8.9 (Proteome Software, Portland, OR) and used to calculate normalized spectral count for each protein from each sample. Scaffold protein and peptide thresholds were set at 95%, with a minimum peptide number of two per protein – this resulted in a protein false discovery rate (FDR) of 1.8% and a peptide FDR of less than 0.1% (33, 34). Statistical analysis of weighted spectral count data for identification of proteins differentially abundant between sample categories was performed in Scaffold using Fisher's exact test (significance level $p < 0.05$ using the Benjamini-Hochberg multiple test correction). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020412 (35, 36).

Protein structure modeling

ICM (Molsoft LLC, San Diego, CA) and Rasmol (37) software were used to visualize .pdb files representing predicted protein three dimensional structure.

Results and Discussion

Protein Interaction Reporter crosslink quantification

A *D. citri* protein interaction network was established using isotope-labeled Protein Interaction Reporter (PIR) mass spectrometry of microbe-enriched cellular fractions isolated from adult and nymph *D. citri* collected from CLas-infected citrus trees. QPCR analysis on three Percoll gradient fractionations showed CLas found in every fraction and slightly enriched in fractions 3 and 4 (**Table 1**). This crosslinking network is available at XLinkDB 3.0 <http://xlinkdb.gs.washington.edu/xlinkdb/index.php> (network name: ACP_Percoll_Fraction_XL_Bruce), which provides visualization of the quantitative site interaction network and details of identified crosslinks (38, 39). A total of 1571 non-redundant crosslinked peptide pairs were identified (**Table S1**), including 326 crosslinks between peptides derived from 93 unique protein pairs (**Table S2**). Out of the total 1571 crosslinks identified, 665 were quantified in two or more biological replicates from adult and nymph samples – for these, log₂ ratios [with 95% confidence interval (CI)] of crosslink abundance between nymph and adult samples were calculated. Crosslinks were determined to be more abundant in one sample type (nymph or adult) when {[absolute value of crosslink log₂ ratio] – [1/2 (95% CI)]} ≥ 1. Using these criteria, 78 crosslinks involving peptides from a total of 24 proteins were determined to be more abundant in adult *D. citri* compared to nymph *D. citri*, and 162 crosslinks involving peptides from a total of 82 proteins were determined to be more abundant in nymph *D. citri* compared to adult *D. citri* (**Table S3**).

Table 1. Average CLas and wingless gene quantification by qPCR in Percoll gradient fractions

Fraction	Average CLas Cq ²	Average Wg Cq ²
1, 2	26.18 (0.17)	23.36 (0.14)
3,4	24.38 (0.07)	24.43 (0.06)
5,6	20.65 (0.04)	21.33 (0.03)
7,8	23.78 (0.08)	24.46 (0.06)
9,10	25.10 (0.06)	26.80 (0.09)
11,12	29.52 (0.07)	29.69 (0.23)
Whole ¹	19.29 (0.16)	23.57 (0.07)

¹ Average Cq values for pooled whole insects used prior to Percoll gradient fractionation.

² Standard deviation of the mean in parenthesis.

Functional characterization of differentially abundant crosslinks

These differentially abundant crosslinks were assigned to functional categories according to the annotation of the proteins associated with the crosslinked peptides (**Figure 2**). Crosslinks within and between a small number of myosin proteins comprise a large fraction of the total identified crosslinks, and the protein functional category most represented in the crosslinks found to be more abundant in both nymph and adult samples is cytoskeleton/muscle (**Figure 2**). This indicates that unique conformations or complexes of cytoskeleton/muscle proteins are present in nymph and adult insects, which is likely related to the dramatic differences in morphology and behavior between developmental stages, including the presence of wings and the concomitant significance of flight muscle in adults. Histone proteins are the second most represented category of crosslinks found to be more abundant in nymph *D. citri*, while no crosslinks containing histone proteins were more abundant in adult *D. citri* (**Figure 2**). A total of 109 unique crosslinks containing peptides mapping to histone proteins were identified, and 35 of these histone crosslinks were found to be more abundant in the nymph samples (**Table S1, S3**). Out of these 35 differentially abundant crosslinks, 12 are intermolecular crosslinks representing interactions between five pairs of histones: Histone 3/4; Histone 2b/4; Histone 3/2b; Histone 3/2a; Histone 2a/2b (**Table 2**).

346

Table 2. Histone cross-linked peptides detected as higher in relative abundance in CLas (+) nymphs vs. adult insects

Peptide A ¹	Peptide B ¹	Protein A accession and description	Protein B accession and description	Crosslink log2ratio adult/nymph
TVTAMDVVYALK ¹¹⁸ R	YQK ⁵⁷ STELLIR	MCOT08963.0.MT_Histone H4	MCOT05881.0.CC_Histone H3	-4.83
DAVTYTEHAK ¹⁰⁴ R	EIAQDFK ⁸⁰ TDLR	MCOT08963.0.MT_Histone H4	MCOT05881.0.CC_Histone H3	-2.57
GVLK ⁸⁶ VFLENVIR	LLPGELAK ¹⁰⁹ H AVSEGTK	MCOT08963.0.MT_Histone H4	A0A1S3DIC0_Histone H2B	-2.39
DAVTYTEHAK ¹⁰⁴ R	LLPGELAK ¹⁰⁹ H AVSEGTK	MCOT08963.0.MT_Histone H4	A0A1S3DIC0_Histone H2B	-2.07
DAVTYTEHAK ¹⁰⁴ R	K ³⁵ ESYAIYIK	MCOT08963.0.MT_Histone H4	A0A1S3DIC0_Histone H2B	-3.49
YQK ⁵⁷ STELLIR	K ¹³⁹ SETAK	MCOT05881.0.CC_Histone H3	MCOT13323.0.CO_Histone H2A	-2.11
VTIMPK ¹²³ DIQLAR	LLSGVTIAQGG VLPNIQAVLLP K ¹⁴⁴ KTEK	MCOT05881.0.CC_Histone H3	MCOT08626.0.MT_Histone H2A	-2.77
YQK ⁵⁷ STELLIR	LLSGVTIAQGG VLPNIQAVLLP K ¹⁴⁴ KTEK	MCOT05881.0.CC_Histone H3	MCOT08626.0.MT_Histone H2A	-3.93
K ⁶⁵ LPFQR	K ³⁵ ESYAIYIK	MCOT05881.0.CC_Histone H3	A0A1S3DIC0_Histone H2B	-2.85
EIAQDFK ⁸⁰ TDLR	LLPGELAK ¹⁰⁹ H AVSEGTK	MCOT05881.0.CC_Histone H3	A0A1S3DIC0_Histone H2B	-2.74
K ⁶² GNYAER	K ³⁵ ESYAIYIK	MCOT08626.0.MT_Histone H2A	A0A1S3DIC0_Histone H2B	-3.12
K ⁶² GNYAER	ESYAIYIK ⁴⁴ VLK	MCOT08626.0.MT_Histone H2A	A0A1S3DIC0_Histone H2B	-3.7

¹ Crosslinked residues are indicated in bold with crosslinked residue number in superscript.

350

Figure 2. Functional categories assigned to proteins involved in crosslinks found to be more abundant in **a)** adult *D. citri* samples, and **b)** nymph *D. citri* samples. For each sample type, the number of crosslinks associated with the most represented category is given. Detailed information on differentially abundant crosslinks is given in **Tables S5**. The category entitled Regulation included proteins involved in transcription, translation, and posttranslational modification.

In the remaining 23 differentially abundant histone crosslinks, both peptides map to the same protein (**Table 3**). Eight of these 23 crosslinks included peptides from histone H3 with posttranslational modifications (PTMs, one dimethylation and seven acetylations). Two of these histone H3 crosslinks include a peptide acetylated at K¹⁰ and a third includes a peptide dimethylated at the same position. In contrast to the convention in our predicted protein database, the histone numbering convention used in the literature counts the amino acid after the initiator methionine as the first residue – on account of this, K¹⁰ in our histone H3 protein is universally described in the literature as H3K9. Both acetylation and methylation of H3K9 are known to increase the hydrophobicity of the histone H3 tail, with the consequence of rendering this domain more available for binding of proteins functioning in chromatin remodeling, including chromodomain helicase DNA binding protein 3 (CHD3) (40). One of the histone crosslinks found to be more abundant in nymphs is the histone H3 homodimer crosslinked at H3K4 (TK⁵QTAR, **Table 3**). The intranucleosomal distances between the crosslinked residues of histone H3 protein pairs were estimated from analysis of the .pdb structure (PDB 1KX5) of the published X-ray crystal structure of the nucleosome(41). In cases where the predicted intranucleosomal distance between crosslinked homodimer residues was greater than the length of the crosslinker (30 angstroms), these crosslinks likely represent internucleosomal interactions. The lysine residues involved in the histone H3 crosslinks from **Table 3** are depicted on structural representations of the two copies of histone H3 within the nucleosome octamer (**Figure 3**).

Table 3. Crosslinks where both peptides are derived from the same histone protein which are more abundant in nymph compared to adult *D. citri*.

Peptide A	Peptide B ^{1,2}	PTM (Peptide A or B)	Protein accession and description	Crosslink log2ratio adult/nymph
TK ⁵ QTAR	K ¹⁰ _{Me2} STGGK ¹⁵ APR	Dimethylation (B)	MCOT05881.0.CC_Histone H3	-2.12
K ¹⁰ _{Ac} STGGK ¹⁵ APR	K ¹⁹ QLATK ²⁴ _{Ac} AAR	Acetylation (A and B)	MCOT05881.0.CC_Histone H3	-3.5
K ¹⁰ STGGK _{Ac} APR	K ¹⁹ QLATK ²⁴ _{Ac} AAR	Acetylation (A and B)	MCOT05881.0.CC_Histone H3	-3.14
K ¹⁹ QLATK ²⁴ _{Ac} A	YQK ⁵⁷ STELLIR	Acetylation	MCOT05881.0.CC_Histone H3	-3.14
TK ⁵ QTAR	K ¹⁰ STGGK ¹⁵ _{Ac} APR	Acetylation	MCOT05881.0.CC_Histone H3	-3.03
TK ⁵ QTAR	K ¹⁹ QLATK ²⁴ _{Ac} AAR	Acetylation	MCOT05881.0.CC_Histone H3	-3.07
TK ⁵ QTAR	K ¹⁰ _{Ac} STGGK ¹⁵ APR	Acetylation	MCOT05881.0.CC_Histone H3	-2.73
TK ⁵ QTAR	K ¹⁹ _{Ac} QLATK ²⁴ AAR	Acetylation	MCOT05881.0.CC_Histone H3	-2.79
K ⁶⁵ LFPQR	YQK ⁵⁷ STELLIR	N/A	MCOT05881.0.CC_Histone H3	-3.75
K ¹⁰ STGGK	TK ⁵ QTAR	N/A	MCOT05881.0.CC_Histone H3	-3.15
TK ⁵ QTAR	YQK ⁵⁷ STELLIR	N/A	MCOT05881.0.CC_Histone H3	-2.9
TK ⁵ QTAR	STGGK ¹⁵ APR	N/A	MCOT05881.0.CC_Histone H3	-2.95
TK ⁵ QTAR	TK ⁵ QTAR	N/A	MCOT05881.0.CC_Histone H3	-2.45
K ⁸⁸ YLVSAAVAGK	VDAQK ⁸¹ LAPFIR	N/A	MCOT07573.0.MM_Histone H1.1	-2.73
TEK ¹⁴⁸ KA	LLSGVTIAQGGVLPNIQAVLLPK ¹⁴⁴ K	N/A	MCOT08626.0.MT_Histone H2A	-3.48
LAHYNK ⁸⁶ R	K ³⁵ ESYAIYIK	N/A	A0A1S3DIC0_Histone	-3.4
RK ³⁵ ESYAIYIY	VLK ⁴⁷ QVHPDTGVSSK	N/A	A0A1S3DIC0_Histone	-3.29
AQK ¹⁹ NIAK	K ³⁵ ESYAIYIK	N/A	A0A1S3DIC0_Histone	-3.15
AVTK ¹²¹ YTSSK	ESYAIYIK ⁴⁴ VLK	N/A	A0A1S3DIC0_Histone	-2.87
K ³⁵ ESYAIYIYK	VLK ⁴⁷ QVHPDTGVSSK	N/A	A0A1S3DIC0_Histone	-2.75
AVTK ¹²¹ YTSSK	LLPGELAK ¹⁰⁹ HAVSEGTK	N/A	A0A1S3DIC0_Histone	-2.21
AVTK ¹²¹ YTSSK	VLK ⁴⁷ QVHPDTGVSSK	N/A	A0A1S3DIC0_Histone	-2.1

¹ Crosslinked residue is indicated in bold with residue number in superscript following crosslinked lysine.

² Peptides containing PTMs are indicated with PTM type in subscript following modified lysine (Ac=acetylation; Me2=dimethylation).

Four of the histone H3 crosslinks found to be more abundant in nymph compared to adult insects are predicted to be internucleosomal crosslinks, because the distance between crosslinked residues either within one H3 protein, or between the two H3 proteins within a nucleosome, is greater than the 30 Å crosslinker distance constraint (**Figure 3**).

Figure 3. Structural representation of the two copies of histone H3 within a nucleosome highlighting residues involved in H3-H3 crosslinks found to be more abundant in nymph compared to adult insects. Distance between crosslinked residues less than 30 Å are represented by black lines, and distances greater than 30 Å are represented by red lines. The crosslinker size is 30 Å, and so the crosslinks represented by red lines are predicted to be intermolecular crosslinks between nucleosomes. Crosslinked lysine residues highlighted in black were found to have posttranslational modifications: acetylated at K¹⁰, K¹⁵, K¹⁹, K²⁴, and dimethylated at K¹⁰.

Increased internucleosomal interactions in nymphs may be due to differences in chromatin configuration between nymph and adult insects. Alternative chromatin states have previously been identified between insect castes, wherein genetically identical organisms exhibit polyphenisms associated with distinct social identities within the colony (42, 43). Sequencing of chromatin immunoprecipitation (ChIP-seq) samples from worker and queen honeybee castes, and from major and minor worker carpenter castes, reveals caste-specific chromatin signatures associated with distinctive transcription profiles (44, 45). Nutritional differences between

worker and queen bees have been associated with variation in histone PTMs, with a histone deacetylase inhibitor identified as a prominent component of royal jelly, the honeybee secretion which is consumed in greater abundance by queens compared to workers (46). Chromatin remodeling represents a effective method for an organism to effect the extensive and coordinated changes in gene and protein expression which characterize different insect developmental states. In previous studies of CLas(+) *D. citri*, significantly more nuclear fragmentation was observed in adult compared to nymph midgut cells, and protein biomarkers of apoptosis were more abundant in midguts of CLas(+) compared to CLas(-) adult insects (7, 47). Differences in histone crosslinks observed between CLas(+) nymph and adult *D. citri* in the present study may be related to the fact that while nuclear fragmentation and apoptosis were observed in adult midguts during CLas acquisition, nymph midgut nuclear morphology was largely unaffected by insect exposure to CLas.

Twenty-eight crosslinks involving metabolic proteins were upregulated in nymph compared to adult *D. citri* (**Table S3**). These include four intramolecular or dimeric crosslinks involving two aminopeptidase N proteins, two crosslinks between peptides from cathepsin proteases, and two crosslinks between peptides from carboxypeptidases. In related insect vectors, orthologs of these proteases have been found to play critical roles in pathogen transmission: a pea aphid (*Acyrtosiphon pisum*) aminopeptidase N was found to serve as the receptor mediating virus uptake into the gut (48), and a green peach aphid (*Myzus persicae*) cathepsin B was associated with reduced virus transmission (49). These highly abundant proteins are expressed in the gut epithelial cells of hemipterans and serve as sites of microbe-vector interaction at the molecular level.

A crosslink between hemocyanin 1 and hemocyanin 2 was identified as more abundant in nymph compared to adult *D. citri*. The hemocyanin 1 protein has previously been identified as a protein involved in the response of *D. citri* to CLas, and hemocyanin 2 has been found to be significantly more abundant in nymph compared to adult *D. citri* (18). Hemocyanins are closely related to the larval storage proteins known as hexamerins, which have been characterized as playing crucial roles in insect development (22, 50). A related crosslink found to be more abundant in nymph *D. citri* was identified between two hexamerin proteins. Proteins with predicted roles in energy metabolism (*e.g.* ATP metabolism, electron transport in the mitochondria) were found to be part of crosslinks upregulated in both adult and nymph samples. In contrast, crosslinks between proteins involved in metabolism were identified as more abundant only in nymph samples.

A crosslink between two proteins annotated as Prohibitin-1 and Prohibitin-2 was found to be more abundant in *D. citri* nymph samples (**Table S3**). Prohibitins (PHBs) are pleiotropic proteins localized to the inner mitochondrial membrane (51), and insect PHBs have been found to function in larval metabolism and pupal development in *Drosophila melanogaster* (52) and as a dengue virus receptor in *Aedes aegypti* (53). While no high-resolution crystal structural models for PHBs exist, a coiled-coil domain comprises the C terminal region of members of this protein family, which has previously been identified by PIR as a site of PHB homo- and hetero-dimerization in human cells (20, 54). There is substantial conservation between human and *D. citri* sequences, and strikingly, a *D. citri* PHB-1 homodimer was identified at the same residue (K202) as was found using PIR in human cells (20, 54). The PHB heterodimer which was found at higher levels in nymph compared to adults was represented by a crosslink between K202 of PHB-1 (the homodimer site), and K237 in the coiled-coil domain of PHB-2. A PHB heterodimer

between these regions had also previously been identified in human cells. These data suggest a conserved function of the PHB interactions across the animal kingdom (20, 54).

Posttranslational modifications of crosslinked peptides

Out of the 1571 unique peptide crosslinks identified in this study, 59 were found to contain at least one of the following PTMs: lysine acetylation (32), methylation (13), and dimethylation (14); and arginine acetylation (1) (**Table S4**). PTM residues are numbered in superscript and annotated in subscript (Ac=acetylation, Me=methylation, Me2=dimethylation, Me3=trimethylation). In four of these crosslinks, all mapping to the N terminal tail region of histone H3, PTMs were identified on both peptides in one crosslink. In three cases both peptides are acetylated, and in the fourth case one peptide is dimethylated ($K^{10}_{Me2}STGGK^{15}APR$), and the other is acetylated ($K^{19}QLATK^{24}_{Ac}AAR$) (**Table S4**). The dimethylated lysine on histone H3 aligns to the residue known as H3K9, and the acetylated lysine aligns to the residue known as H3K23. While H3K9 methylation (mono, di, and tri) has been associated with heterochromatin structures and decreased transcription (55, 56), H3K23 acetylation has been associated with transcriptional activation (57, 58). In 25 out of the 30 crosslinks containing acetylated lysines, both peptides mapped to one of four core histone proteins (histones H3, H4, H2A/2B). The five other acetylated crosslinked peptides mapped to a myosin heavy chain and an uncharacterized protein (**Table S4**). Three crosslinks were identified where both peptides map to the N terminal tail region of histone H3, with the same peptide found to be either acetylated and dimethylated at H3K9 ($K^{10}_{Ac/Me2}STGGK^{15}APR$) in all three cases. An additional 24 crosslinks between peptides mapped to myosin or actin proteins were found to contain peptides with methylation or dimethylation PTMs.

Protein homodimer crosslinks

Out of the 1245 crosslinks where each peptide in a pair is from the same protein, 65 crosslinks are considered to represent unambiguous homodimers due to the fact that the crosslinker-modified amino acid residue of each peptide is precisely the same (**Table S5**). Eight unique homodimer crosslinks were identified between peptides derived from the core histone proteins H4, H3.3, H2a, and H2B (**Table 4**).

Table 4. Crosslinked peptides representing unambiguous homodimers of histone proteins.

Protein ID	Protein Description	Distance (Å) ¹	Peptide 1 ²	Peptide 2 ²
MCOT08626.0.MT	Histone H2A	14.2	K ⁶² GNYAER	K ⁶² GNYAER
A0A1S3DIC0	Histone H2B	56.6	AV T K ¹²¹ YTSSK	AV T K ¹²¹ YTSSK
MCOT05881.0.CC	Histone H3.3	116	T K ⁵ QTAR	T K ⁵ QTAR
MCOT05881.0.CC	Histone H3.3	99.7	K ¹⁹ QLAT K ²⁴ _{Ac} AAR	K ¹⁹ QLAT K ²⁴ _{Ac} AAR
MCOT05881.0.CC	Histone H3.3	15.6	VTIMP K ¹²³ DIQLAR	VTIMP K ¹²³ DIQLAR
MCOT05881.0.CC	Histone H3.3	49.4	Y Q K ⁵⁷ STELLIR	Y Q K ⁵⁷ STELLIR
MCOT08963.0.MT	Histone H4	13.5	TVTAMDVVYAL K ¹¹⁸ R	TVTAMDVVYAL K ¹¹⁸ R
MCOT08963.0.MT	Histone H4	13.5	TVTAMDVVYAL K ¹¹⁸ R	KTVTAMDVVYAL K ¹¹⁸ R

¹ Distances in angstroms between homodimer residues of each protein pair within the nucleosome core particle predicted from the modeled X-ray structure. Distances greater than the 30 Å crosslinker length are predicted to represent interactions between nucleosomes, while shorter distances may represent interactions between protein pairs within a single histone octamer.

² Crosslinked lysine in bold and numbered in superscript, acetylated lysines indicated with K_{Ac} and numbered in superscript.

The distance between two crosslinked **K**¹²³ residues within a nucleosome is predicted to be 16 Å (**Table 4, Figure 4A**) and therefore could represent an intranucleosomal homodimer. This crosslinked residue is in the histone H3 C terminal region, which has been reported to be the H3 dimerization interface (59). In contrast, the predicted distance between two H3 N terminal residues (K19) within a nucleosome is 100 Å, suggesting that this crosslink represents an internucleosomal interaction (**Figure 4A**). Internucleosomal interactions have significant impacts on chromatin structure and gene transcription, are frequently mediated by posttranslational modifications of histone N terminal tail regions (60), and are indicative of

480 chromatin in its closed (heterochromatin) rather than open (euchromatin) configuration (**Figure**
481 **4B**).

Figure 4. a) Structure of the nucleosome core particle as predicted from the X-ray diffraction pattern (PDB 1KX5). DNA is wound around the histone octamer, with histone N terminal tails extending from the nucleosome. The distances between the homodimer crosslink residues K123/K123 (16 Å) and K19/K19 (100 Å) in the two copies of histone H3 in the nucleosome is given. The distance between the two K19 residues is greater than the 30 Å crosslinker distance constraint, suggesting that the K19 homodimer represents an internucleosomal interaction between the H3 N terminal tails. **b)** Schematic of chromatin in its closed configuration (heterochromatin) and open configuration (euchromatin). DNA (green strand) is wrapped around nucleosome particles, with histone N terminal tails depicted in blue.

482 The H3K18 homodimer crosslink residue is a well-characterized site for functionally
483 significant histone acetylation – when acetylation at this site and others in histone H3 of a
484 pathogenic fungus is blocked by a wheat microbiome bacteria, fungal growth and virulence is
485 inhibited. Both peptides comprising this homodimer were acetylated at the lysine known as
486 H3K23 (57, 58) ($\text{K}^{19}\text{QLATK}^{24}_{\text{Ac}}\text{AAR}$), which is a histone acetylation associated with highly
487 expressed genes in plants and insects (57, 58, 61) (**Table 4**). The tetramer comprising two H3/H4
488 dimers is predicted to form via interactions between C terminal regions of Histone H3 (62),

which is represented in our data by the H3K123 homodimer (**Table 3, Figure 4A**). In addition, two unique peptide pairs were found to be crosslinked at the same lysine residue on the C terminus of Histone H4 (**Table 4**). The crosslinked residue corresponds to the Histone H4 C terminal lysine known as H4K91, which is found at the interface of the H3/H4 tetramer and H2A/H2B dimers and has been found to be acetylated and ubiquitinated with consequences on chromatin structure (63-65).

Other homodimers of biological significance which were identified in *D. citri* samples include three homodimer crosslinks between peptides in the C terminal domain of hemocyanin 1. Hemocyanin is known to form oligomeric structures, and has a predicted immunity function in numerous arthropod species (22, 50). The hemocyanin peptide LNHK⁷⁹⁴SFNYR, which was found in this study to form a homodimer, was previously identified as a component of a crosslink between hemocyanin and a CLas bacterial protein involved in coenzyme A biosynthesis (18).

Bacterial endosymbiont protein crosslinks

Surprisingly, although the Percoll fractions were apparently enriched for CLas using qPCR detection, no crosslinks involving CLas proteins were identified. Eleven crosslinks involving peptides of bacterial endosymbionts of *D. citri* were identified – in one crosslink one peptide was derived from a bacterial protein and the other peptide was derived from an insect protein, while in the remainder both crosslinked peptides are bacterial (**Table 5**). The insect-bacterial crosslink involves a polyketide synthase (PedF) involved in production of diaphorin, an abundant polyketide produced by the *D. citri* endosymbiont '*Candidatus Profftella armatura*', and an insect protein with a predicted secretion signal but no homology to known protein domains. Five of the bacterial crosslinks include Profftella proteins, including two crosslinks with both peptides mapping to a second polyketide synthase, PedI. The other five bacterial

crosslinks involve the *D. citri* endosymbiont '*Candidatus Carsonella rudii*', including a crosslink between the alpha and beta subunits of an ATP synthase, which are known interacting partners (66).

Quantitative mass spectrometry proteomics of D. citri percoll fraction samples

Adult and nymph *Diaphorina citri* insects were collected from healthy (CLas-) and infected (CLas+) citrus plants. Percoll gradient fractionation of insect homogenate was used to generate a concentrated sample of insect and microbe cells. The same fractions from the percoll gradient as were used for PIR analysis - those selected as having the highest abundance of CLas based on qPCR analysis for CLas titer - were used as starting material for quantitative

Table 5. Crosslinks involving microbial proteins identified in *Diaphorina citri* samples.

Organisms	Protein A ID; Xlink residue	Protein A Description	Protein B ID; Xlink residue	Protein B Description
Profftella- <i>D.citri</i>	WP_020915462.1; 8131	Type I polyketide synthase, PedF-like protein	MCOT00458.1.CT; 244	unknown function
Profftella	S5RPP0; 3702	Type I polyketide synthase, PedI-like protein	S5RPP0; 3688	Type I polyketide synthase, PedI-like protein
Profftella	S5RPP0; 3317	Type I polyketide synthase, PedI-like protein	S5RPP0; 3306	Type I polyketide synthase, PedI-like protein
Profftella	S5R3E1; 100	50S ribosomal protein L7/L12	S5R3E1; 114	50S ribosomal protein L7/L12
Profftella	S5R113; 361	Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex	S5R113; 361	Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex
Profftella	S5R0V0; 273	Acetyl-coenzyme A carboxylase carboxyl transferase beta	S5R0V0; 257	Acetyl-coenzyme A carboxylase carboxyl transferase beta
Carsonella	WP_045118136.1;891	unknown function	WP_045118136.1; 895	unknown function
Carsonella	WP_045118136.1; 601	unknown function	WP_045118136.1; 963	unknown function
Carsonella	WP_020915898.1; 331	elongation factor Tu	WP_020915898.1;346	elongation factor Tu
Carsonella	WP_020915723.1; 100	F0F1 ATP synthase alpha	WP_020915723.1;30	F0F1 ATP synthase alpha
Carsonella	WP_020915723.1; 358	F0F1 ATP synthase alpha	WP_020915725.1; 396	F0F1 ATP synthase beta

proteomics. Proteins were extracted from percoll fraction samples, and following sample

preparation, trypsin-digested peptides were analyzed by mass spectrometry. A total of 1780 insect and microbe proteins were identified (**Table S6**). More than 98% of these proteins are insect proteins, and 37 microbe proteins derived from the three primary bacterial endosymbionts of *D. citri* [*Candidatus Proffittella armatura* (22), *Wolbachia pipientis* (11), and *Candidatus Carsonella rudii* (4)] were identified. Peptides were also assigned to two CLas ATPase proteins (WP_015452542.1, WP_015452544.1) and a CLas DnaK molecular chaperone (WP_015824904.1). However, these peptide spectra were identified in both CLas(-) and CLas(+) samples, and the protein region to which the peptides map is highly conserved between CLas and the three endosymbionts, suggesting that these proteins are likely derived from one of the endosymbionts, which are present in all sample types.

Statistical analysis of quantitative mass spectrometry data

Statistical analysis of spectral counts was performed to identify proteins which were differentially abundant between sample categories (healthy nymph vs healthy adult, infected nymph vs infected adult, healthy nymph vs infected nymph, healthy adult vs infected adult, **Table 6, Table S7**). The fewest number of differentially abundant proteins (50) were identified between CLas(-)/CLas(+) nymphs, whereas nearly three times as many differentially abundant proteins (132) were identified between CLas(-)/CLas(+) adults (**Table 6**). These results are consistent with data from a previous quantitative proteomics analysis of whole insect nymph and adult samples – in this prior study, four times as many proteins in the adult insect were regulated by CLas infection as in the nymph insect (18). Many more proteins (>600) were found at significantly different abundance between nymph and adult percoll fraction samples than between CLas(+) and CLas(-) samples (**Table 6, Table S7**). Vitellogenins are proteins with documented function in lipid transport, reproduction, and immunity in insects. While some

vitellogenins which function in reproduction are expressed predominantly in female insects, a vitellogenin in the potato psyllid (*Bactericera cockerelii*) was found to be expressed at similar levels in male and female insects and may have an immune function (67). Several vitellogenin proteins, which were found at higher levels in this study in CLas(+) adult compared to CLas(-) adult samples, were either undetected or found at the limit of detection in nymphs of both infection status (**Table S7**). These include Vg-1 vitellogenins (XP_008476565, MCOT18962.0.CT), which were previously identified in the hemolymph of adult *D. citri* both as among the most highly abundant proteins in this tissue, and as more abundant in hemolymph from CLas(+) compared to CLas(-) insects (13).

Table 6. Summary of statistical analysis of quantitative proteomics data.

Sample Category Comparison (A vs B)	Number upregulated proteins ¹	Number downregulated proteins ¹
CLas(-) adult vs CLas(-) nymph	336	352
CLas(+) adult vs CLas(+) nymph	294	303
CLas(+) adult vs CLas(-) adult	65	67
CLas(+) nymph vs CLas(-) nymph	25	25

¹ For each sample category comparison, the number of proteins upregulated and downregulated in category A vs B are given.

Subsequent to performing this work, data from our lab has revealed substantial variability between the different color morphs which are found within a population of *D. citri* (68). Because blue and non-blue *D. citri* were pooled into the same sample in this study, the abundance of proteins which may be significantly different between color morphs will not be precisely quantified. Follow-up work to separate insects from each developmental stage into separate color morph classes for analysis will provide insight into the impact of this significant polyphenism on insect development and immunity.

Integration of PIR and Spectral Count data

For the peptide crosslinks identified at significantly higher levels in nymph or adult *D. citri*, the increased abundance of these crosslinks may be attributed to the higher abundance of the parent proteins in one developmental stage relative to the other. However, in several instances a crosslink was found at significantly higher levels in nymphs compared to adults, but the proteins from which the peptides comprising that crosslink were derived were found at significantly lower levels in nymphs compared to adults. The PHB1-PHB2 heterodimer crosslink was found at four-fold higher abundance in nymphs compared to adults, but the spectral count data indicate that PHB1 and PHB2 are found at nine-fold and five-fold higher levels, respectively, in adults compared to nymphs. This indicates that protein complex formation is not simply a function of the concentration of the constituent proteins, and suggests that differences in protein complex and/or protein interaction topology is responsible for differences in crosslink abundance. Additional proteins involved in promoting PHB1-PHB2 complex formation may be more abundant in nymphs relative to nymphs, leading to increased abundance of this crosslink in nymphs despite the decreased abundance of the component proteins. Similarly, a crosslink between two hemocyanin 1 peptides was found by PIR to be more abundant in CLas(+) nymphs compared to adults, while spectral counting data indicate that hemocyanin 1 is one of the most abundant proteins in CLas(+) adults, and is over ten times as abundant in these samples compared to CLas(+) nymphs (**Tables S3, S7**).

Statistical analysis of spectral counting data suggests that histone protein expression is regulated by insect developmental state and CLas infection status. Histone H2B (two isoforms) and histone H4 were both found at higher abundance in adult compared to nymph insects, regardless of infection status (**Table 7**). In contrast, multiple crosslinks involving these core histones were found to be more abundant in nymphs compared to adults (**Tables 2, 3**). CLas

infection is associated with reduced histone H2B expression: both histone H2B isoforms were found at higher abundance in CLas(-) compared to CLas(+) nymph and adult insects. Histone H4 was also more abundant in CLas(-) compared to CLas(+) adult insects, but no statistical difference between spectral counts was found between CLas(-) and CLas(+) nymph samples. The core histone variant H3.3 was found at higher levels in CLas(-) nymphs compared to CLas(-) adults, but in CLas(+) samples the protein was more abundant in adults compared to nymphs (Table 7).

Table 7. Histone proteins found to be differentially abundant between insect sample classes based on development stage and CLas infection status.

Protein ID	Annotation	N ⁽⁻⁾	N ⁽⁺⁾	A ⁽⁻⁾	A ⁽⁺⁾	Significant Difference ¹
MCOT10597.0.OO	Histone H2B	80.49	15.35	129.08	43.66	Development: N ⁽⁻⁾ <A ⁽⁻⁾ ; N ⁽⁺⁾ <A ⁽⁺⁾ ; CLas infection: N ⁽⁺⁾ <N ⁽⁻⁾ ; A ⁽⁺⁾ <A ⁽⁻⁾
A0A1S3DIC0	Histone H2B	199.48	61.42	294.43	134.16	Development: N ⁽⁻⁾ <A ⁽⁻⁾ ; N ⁽⁺⁾ <A ⁽⁺⁾ ; CLas infection: N ⁽⁺⁾ <N ⁽⁻⁾ ; A ⁽⁺⁾ <A ⁽⁻⁾
MCOT08963.0.MT	Histone H4	63.33	44	78	63.33	Development: N ⁽⁻⁾ <A ⁽⁻⁾ ; N ⁽⁺⁾ <A ⁽⁺⁾ ; CLas infection: A ⁽⁺⁾ <A ⁽⁻⁾
MCOT05881.0.CC	Histone H3.3	17.99	6.22	13.16	11.56	Development: N ⁽⁻⁾ >A ⁽⁻⁾ ; N ⁽⁺⁾ <A ⁽⁺⁾
XP_008482743.1	Histone H1E	9.01	3.95	3.14	2.33	Development: N ⁽⁻⁾ >A ⁽⁻⁾ ; N ⁽⁺⁾ >A ⁽⁺⁾
XP_008480034.1	Histone H1.M6.2	36	12.66	4.66	3.33	Development: N ⁽⁻⁾ >A ⁽⁻⁾ ; N ⁽⁺⁾ >A ⁽⁺⁾
MCOT07573.0.MM	Histone H1-II-1	161.86	67.86	50.76	40.93	Development: N ⁽⁻⁾ >A ⁽⁻⁾ ; N ⁽⁺⁾ >A ⁽⁺⁾ ; CLas infection: N ⁽⁺⁾ <N ⁽⁻⁾ ; A ⁽⁺⁾ <A ⁽⁻⁾
MCOT12721.3.CO	Micronuclear Histone H1	11.9	1.65	0.66	0.32	Development: N ⁽⁻⁾ >A ⁽⁻⁾ ; N ⁽⁺⁾ >A ⁽⁺⁾ ; CLas infection: N ⁽⁺⁾ <N ⁽⁻⁾

¹ Average spectral counts for each sample class (N=3), and sample classes found to have significantly different spectral counts, are listed. N⁽⁻⁾=nymph CLas(-); N⁽⁺⁾=nymph CLas(+); A⁽⁻⁾=adult CLas(-); A⁽⁺⁾=adult CLas(+).

Whereas histone H2B, H4, and H3.3 are all components of the nucleosomal core particle around which DNA are wound, linker histone H1 proteins bind to the DNA entry and exit sites on the nucleosome, stabilizing chromatin structure and influencing gene expression (69, 70). In contrast to the higher expression of histone H2B and H4 in adults compared to nymphs, four histone H1 variants were found to be much more abundant in nymph compared to adult insects regardless of infection status. In addition, micronuclear histone H1 (nymphs) and histone H1-II-1 (adults and nymphs) were also more abundant in CLas(-) compared to CLas(+) insects (Table 7).

Two proteins which function in histone binding and modification were found at significantly different levels between sample classes: a histone lysine N-methyltransferase (MCOT17739.0.CT) was significantly more abundant in CLas(+) compared to CLas(-) adult samples, while the histone binding protein rbbp4 (MCOT00465.0.CT) was significantly more abundant in CLas(+) adult compared to CLas(+) nymph samples (**Table S7**).

Identification of posttranslational modifications in quantitative proteomics data

Lysine acetylation, methylation, and dimethylation were selected as variable modifications during Mascot searching. The data are summarized as the number of assigned modified spectra per protein (**Table S8**), and the modified peptide sequences are available in the peptide spectrum report (**Table S9**). Peptides with acetylated lysines were mapped to 112 proteins, peptides with methylated lysines were mapped to 255 proteins, and peptides with dimethylated lysines were mapped to 103 proteins (**Table S8**). Acetylated and methylated peptides were mapped to several of the histone proteins which were found at different abundance in adult compared to nymph and/or CLas(-) compared to CLas(+) samples (**Tables 7,8**). The acetylated peptide K⁸²_{Ac}YLVSAVAAGK is found in two histone H1 proteins found to be more abundant in nymph compared to adult samples. Based on our data it is not possible to definitively map this acetylated peptide to either of these proteins, or to know whether both of these proteins, or only one of them, are acetylated at this peptide. The acetylated residue is found at K82 on XP_008482743.1 (histone H1E) and K88 on MCOT07573.0.MM (histone H1-II-1). This acetylated peptide aligns to K85 of the human histone H1 (AAA63187.1), which is subject to post-translational acetylation significantly impacting chromatin structure (71). Both of these histone H1 proteins were more abundant in nymph compared to adult samples (**Table 7**).

While the peptide KYLVSAVAAGK is found in both adult and nymph samples, the modified

Table 8. Acetylated and methylated peptides mapping to the histone proteins identified by quantitative proteomics as differentially abundant between *D. citri* sample classes.

			# modified/unmodified peptide spectra ¹			
Protein ID	Protein Description	Modified Peptide ²	nymph CLas(-)	nymph CLas(+)	adult CLas(-)	adult CLas(+)
XP_008482743.1	histone H1E	K ⁸² _{Ac} YLVSAVAAGK*	12/93	6/39	0/36	0/39
XP_008480034.1	histone H1.M6.2	KPAQAQPAAGK ⁵⁹ _{Me}	24/0	1/0	6/0	5/0
A0A1S3DIC0	histone H2B	QVHPDTGVSSK ⁵⁸ _{Ac} *	14/88	6/36	0/66	4/68
A0A1S3DIC0	histone H2B	QVHPDTGVSSK ⁵⁸ _{Me} *	4/88	6/36	0/66	0/68
A0A1S3DIC0	histone H2B	VLK ⁴⁷ _{Ac} QVHPDTGVSSK*	0/62	0/32	2/40	0/28
A0A1S3DIC0	histone H2B	HAVSEGTK ¹¹⁷ _{Me}	10/657	5/119	10/1062	12/346
A0A1S3DIC0	histone H2B	LLLPGELAK ¹⁰⁹ _{Me}	20/64	6/26	42/67	47/63
MCOT05881.0.CC	Histone H3.3	K ²⁸ _{Ac} SAPSTGGVK	1/9	3/2	0/4	0/3
MCOT05881.0.CC	Histone H3.3	K ²⁸ _{Me2} SAPSTGGVK	20/9	27/2	17/4	16/3
MCOT05881.0.CC	Histone H3.3	STGGK ¹⁵ _{Ac} APR	44/0	12/0	22/0	10/0
MCOT05881.0.CC	Histone H3.3	QLATK ²⁴ _{Ac} AAR*	44/0	8/0	25/0	28/0
MCOT05881.0.CC	Histone H3.3	EIAQDFK ⁸⁰ _{Me2} TDLR*	4/42	4/22	22/36	14/40
MCOT05881.0.CC	Histone H3.3	EIAQDFK ⁸⁰ _{Me} TDLR*	34/42	14/22	68/36	66/40
MCOT08963.0.MT	Histone H4	DAVTYTEHAK ¹⁰⁴ _{Me}	2/31	1/8	0/30	0/33
* These peptides each map to 2 unique proteins (XP_008482743.1/MCOT07573.0.MM; XP_008481587.1/MCOT05138.0.CC; MCOT05881.0.CC/XP_008484413.1) and cannot be assigned to one or the other based on these data						

peptide with the first lysine acetylated is only found in nymph samples (**Table 8**).

¹ The total number of modified/unmodified spectra identified in each sample class (sum of three reps/class) are given.

² Modification type and residue number are indicated for each modified lysine.

Multiple methylated and acetylated peptides derived from histone H2B, H3.3, and H4 were identified (**Table 8**). Unique methylated and acetylated forms of a single histone H2B peptide were among five modified peptides mapping to histone H2B. Six modified peptides mapped to histone H3.3, including two acetylated peptides (STGGK¹⁵_{Ac}APR, QLATK²⁴_{Ac}AAR) for which no spectra representing unmodified peptides were identified (**Table 8**). These acetylated peptides, for which more spectra were found in nymph CLas(-) samples compared to

all other sample types, were identified as part of five separate crosslinks found by PIR to be more abundant in nymph *D. citri* (**Figure 5**). The acetylated peptides identified by PIR are identical except for the presence of an additional N terminal lysine, which is the crosslinked residue (**K**¹⁰STGGK¹⁵_{Ac}APR, **K**¹⁹QLATK²⁴_{Ac}AAR).

Quantitative Protein Interaction Reporter (PIR)			Quantitative Proteomics (Spectral Counting)				
Histone H3 Xlink log ₂ Ratio Adult CLas(+)/Nymph CLas(+)	Histone H3 Peptide A	Histone H3 Peptide B	Histone H3 Acetylated Peptide	nymph CLas(-)	nymph CLas(+)	adult CLas(-)	adult CLas(+)
-3.07	TK ⁹ QTAR	K ¹⁰ STGGK ¹⁵ _{Ac} APR	STGGK ¹⁵ _{Ac} APR	44	12	22	10
-3.14	K ¹⁹ QLATK ²⁴ _{Ac} AAR	K ¹⁰ STGGK ¹⁵ _{Ac} APR					
-3.03	TK ⁹ QTAR	K ¹⁹ QLATK ²⁴ _{Ac} AAR	QLATK ²⁴ _{Ac} AAR	44	8	25	28
-3.14	K ¹⁰ _{Ac} STGGK ¹⁵ APR	K ¹⁹ QLATK ²⁴ _{Ac} AAR					
-3.5	YQK ⁵⁷ STELLIR	K ¹⁹ QLATK ²⁴ _{Ac} AAR					

Figure 5. Impact of development and CLas infection on histone H3 acetylation: integration of PIR and spectral counting data. PIR crosslinks found to be more abundant in nymph compared to adult *D. citri* involve lysine-acetylated peptides (K15, 24) in the N terminal tail region histone H3. Crosslinked residues are in bold. Peptides containing these acetylated lysines were also identified in spectral counting data obtained from separate samples. Total spectral counts for the two modified peptides in the four sample classes (sum of three reps/class) are given.

The presence of an additional lysine on the peptides identified by PIR can be explained by the presence in the histone H3 protein sequence of an arginine residue on the N terminal side of this lysine. The trypsin enzyme used in peptide sample preparation for both PIR and spectral counting analysis cuts on the carboxyl side of either a lysine or arginine. In the PIR samples it is possible that the presence of the crosslinker on the lysine residue influences trypsin digestion such that the preferred cutting site is on the carboxyl side of the upstream arginine. In this way the N terminal lysine is retained in the PIR samples, while in the peptides prepared for quantitative proteomics the preferred trypsin digestion site is on the carboxyl side of this lysine.

Conclusions

Complementary proteomics analyses were used to identify and quantify proteins and protein interactions in nymph and adult samples of *Diaphorina citri*, the insect vector of the

bacterial pathogen ('*Candidatus* Liberibacter asiaticus', CLas) associated with citrus greening disease. Quantitative isotope-labeled protein interaction reporter (PIR) technology identified more than 1500 unique crosslinks between pairs of insect peptides, between pairs of peptides derived from insect bacterial endosymbionts, and in one case between an insect and a microbe peptide. Crosslinker data analysis revealed that insect histone protein crosslinks were overrepresented among three classes of protein interactions: unambiguous homodimers, crosslinks more abundant in nymph compared to adult insects, and crosslinks containing methylated and acetylated lysines. Quantitative proteomics spectral counting analysis of nymph and adult insects collected from healthy and infected citrus trees [CLas(-) and CLas(+) samples] indicated that development and infection status influence histone protein levels and posttranslational modification of histone peptides. While histone H2B and H4 were both found at higher abundance in adult compared to nymph insects, multiple crosslinks involving these core histones were found to be more abundant in nymphs compared to adults. These histone crosslinks may represent PTM-mediated internucleosomal interactions reflecting developmental stage-specific chromatin configurations. Several crosslinked peptides identified in psyllids have been found in crosslinked peptides derived from human cells, showing selection for particular protein interaction topologies between orthologous proteins conserved across the animal kingdom. Variation in histone protein abundance, protein interactions, and postranslational modifications between nymph and adult insects may reflect changes in chromatin structure during development with potentially dramatic consequences on transcriptional activity and vector competence of this globally significant agricultural pest.

Data Availability

The mass spectrometry proteomics crosslinking data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD023382. The mass spectrometry proteomics spectral counting data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020412.

Supplemental Data

This article contains supplemental data.

References

1. Gray, S. M.; Banerjee, N., Mechanisms of arthropod transmission of plant and animal viruses. *Microbiology and Molecular Biology Reviews* **1999**, 63, (1), 128-+.
2. Perilla-Henao, L. M.; Casteel, C. L., Vector-Borne Bacterial Plant Pathogens: Interactions with Hemipteran Insects and Plants. *Frontiers in Plant Science* **2016**, 7.
3. Duan, Y. P.; Zhou, L. J.; Hall, D. G.; Li, W. B.; Doddapaneni, H.; Lin, H.; Liu, L.; Vahling, C. M.; Gabriel, D. W.; Williams, K. P.; Dickerman, A.; Sun, Y. J.; Gottwald, T., Complete Genome Sequence of Citrus Huanglongbing Bacterium, '*Candidatus Liberibacter asiaticus*' Obtained Through Metagenomics. *Molecular Plant-Microbe Interactions* **2009**, 22, (8), 1011-1020.
4. Zheng, Z.; Wu, F.; Kumagai, L. B.; Polek, M.; Deng, X.; Chen, J., Two '*Candidatus Liberibacter asiaticus*' Strains Recently Found in California Harbor Different Prophages. *Phytopathology* **2017**, 107, (6), 662-668.
5. Ammar, E. D.; Ramos, J. E.; Hall, D. G.; Dawson, W. O.; Shatters, R. G., Acquisition, Replication and Inoculation of '*Candidatus Liberibacter asiaticus*' following Various Acquisition Periods on Huanglongbing-Infected Citrus by Nymphs and Adults of the Asian Citrus Psyllid. *Plos One* **2016**, 11, (7).
6. Inoue, H.; Ohnishi, J.; Ito, T.; Tomimura, K.; Miyata, S.; Iwanami, T.; Ashihara, W., Enhanced proliferation and efficient transmission of '*Candidatus Liberibacter asiaticus*' by adult *Diaphorina citri* after acquisition feeding in the nymphal stage. *Annals of Applied Biology* **2009**, 155, (1), 29-36.
7. Mann, M.; Fattah-Hosseini, S.; Ammar, E.; Stange, R.; Warrick, E.; Sturgeon, K.; Shatters, R.; Heck, M., *Diaphorina citri* Nymphs Are Resistant to Morphological Changes

Induced by "*Candidatus* Liberibacter asiaticus" in Midgut Epithelial Cells. *Infection and Immunity* **2018**, 86, (4).

8. Mann, M.; Saha, S.; Cicero, J. M.; Pitino, M.; Moulton, K.; Cano, L.; Hunter, W. B.; Mueller, L. A.; Heck, M., Quantification of new and archived *Diaphorina citri* transcriptome data using a chromosomal length *D. citri* genome assembly reveals the vector's tissue-specific transcriptional response to citrus greening disease. *bioRxiv* **2021**, 2021.09.19.460957.

9. Heck, M.; Brault, V., Targeted disruption of aphid transmission: a vision for the management of crop diseases caused by Luteoviridae members. *Curr Opin Virol* **2018**, 33, 24-32.

10. Gallo-Franco, J. J.; Duque-Gamboa, D. N.; Toro-Perea, N., Bacterial communities of *Aphis gossypii* and *Myzus persicae* (Hemiptera: Aphididae) from pepper crops (*Capsicum* sp.). *Scientific Reports* **2019**, 9.

11. Kaur, N.; Hasegawa, D. K.; Ling, K. S.; Wintermantel, W. M., Application of Genomics for Understanding Plant Virus-Insect Vector Interactions and Insect Vector Control. *Phytopathology* **2016**, 106, (10), 1213-1222.

12. Yin, C. L.; Shen, G. Y.; Guo, D. H.; Wang, S. P.; Ma, X. Z.; Xiao, H. M.; Liu, J. D.; Zhang, Z.; Liu, Y.; Zhang, Y. Q.; Yu, K. X.; Huang, S. Q.; Li, F., InsectBase: a resource for insect genomes and transcriptomes. *Nucleic Acids Research* **2016**, 44, (D1), D801-D807.

13. Kruse, A.; Ramsey, J. S.; Johnson, R.; Hall, D. G.; MacCoss, M. J.; Heck, M., '*Candidatus* Liberibacter asiaticus' minimally alters expression of immunity and metabolism proteins in the hemolymph of *Diaphorina citri*, the insect vector of Huanglongbing. *J Proteome Research* **2018**, 17(9):2995-3011.

14. Ramsey, J. S.; Johnson, R. S.; Hoki, J. S.; Kruse, A.; Mahoney, J.; Hilf, M. E.; Hunter, W. B.; Hall, D. G.; Schroeder, F. C.; MacCoss, M. J.; Cilia, M., Metabolic Interplay between the Asian Citrus Psyllid and Its Proffittella Symbiont: An Achilles' Heel of the Citrus Greening Insect Vector. *Plos One* **2015**, 10, (11).

15. Wang, H.; Wu, K. K.; Liu, Y.; Wu, Y. F.; Wang, X. F., Integrative proteomics to understand the transmission mechanism of Barley yellow dwarf virus-GPV by its insect vector *Rhopalosiphum padi*. *Scientific Reports* **2015**, 5.

16. Zhao, J.; Chi, Y.; Zhang, X. J.; Lei, T.; Wang, X. W.; Liu, S. S., Comparative proteomic analysis provides new insight into differential transmission of two begomoviruses by a whitefly. *Virology Journal* **2019**, 16.

17. Chavez, J. D.; Cilia, M.; Weisbrod, C. R.; Ju, H. J.; Eng, J. K.; Gray, S. M.; Bruce, J. E., Cross-linking Measurements of the Potato leafroll virus Reveal Protein Interaction Topologies Required for Virion Stability, Aphid Transmission, and Virus-Plant Interactions. *Journal of Proteome Research* **2012**, 11, (5), 2968-2981.

18. Ramsey, J. S., Chavez, J.D., Johnson, R., Hosseinzadeh, S., Mahoney, J.E., Mohr, J.P., Robison, F., Zhong, X., Hall, D.G., MacCoss, M., Bruce, J., Cilia, M., Protein Interaction Networks at the host-microbe interface in *Diaphorina citri*, the insect vector of the citrus greening pathogen. *Royal Society Open Science* **2017**, 4, (160545).
19. Chavez, J. D.; Lee, C. F.; Caudal, A.; Keller, A.; Tian, R.; Bruce, J. E., Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. *Cell Systems* **2018**, 6, (1), 136.
20. Chavez, J. D.; Weisbrod, C. R.; Zheng, C. X.; Eng, J. K.; Bruce, J. E., Protein Interactions, Post-translational Modifications and Topologies in Human Cells. *Molecular & Cellular Proteomics* **2013**, 12, (5), 1451-1467.
21. van der Burg, K. R. L.; Lewis, J. J.; Martin, A.; Nijhout, H. F.; Danko, C. G.; Reed, R. D., Contrasting Roles of Transcription Factors Spineless and EcR in the Highly Dynamic Chromatin Landscape of Butterfly Wing Metamorphosis. *Cell Reports* **2019**, 27, (4), 1027.
22. Burmester, T., Expression and evolution of hexamerins from the tobacco hornworm, *Manduca sexta*, and other Lepidoptera. *Insect Biochemistry and Molecular Biology* **2015**, 62, 226-234.
23. Decker, H.; Jaenicke, E., Recent findings on phenoloxidase activity and antimicrobial activity of hemocyanins. *Developmental and Comparative Immunology* **2004**, 28, (7-8), 673-687.
24. Lee, S. Y.; Lee, B. L.; Soderhall, K., Processing of an antibacterial peptide from hemocyanin of the freshwater crayfish *Pacifastacus leniusculus*. *Journal of Biological Chemistry* **2003**, 278, (10), 7927-7933.
25. Zhong, X. F.; Navare, A. T.; Chavez, J. D.; Eng, J. K.; Schweppe, D. K.; Bruce, J. E., Large-Scale and Targeted Quantitative Cross-Linking MS Using Isotope-Labeled Protein Interaction Reporter (PIR) Cross-Linkers. *Journal of Proteome Research* **2017**, 16, (2), 720-727.
26. Chu, C. C.; Gill, T. A.; Hoffmann, M.; Pelz-Stelinski, K. S., Inter-Population Variability of Endosymbiont Densities in the Asian Citrus Psyllid (*Diaphorina citri* Kuwayama). *Microb Ecol* **2016**, 71, (4), 999-1007.
27. Weisbrod, C. R.; Chavez, J. D.; Eng, J. K.; Yang, L.; Zheng, C. X.; Bruce, J. E., In Vivo Protein Interaction Network Identified with a Novel Real-Time Cross-Linked Peptide Identification Strategy. *Journal of Proteome Research* **2013**, 12, (4), 1569-1579.
28. Keller, A.; Chavez, J. D.; Bruce, J. E., Increased sensitivity with automated validation of XL-MS cleavable peptide crosslinks. *Bioinformatics* **2019**, 35, (5), 895-897.
29. Valot, B.; Langella, O.; Nano, E.; Zivy, M., MassChroQ: A versatile tool for mass spectrometry quantification. *Proteomics* **2011**, 11, (17), 3572-3577.

30. Saha, S.; Hosmani, P. S.; Villalobos-Ayala, K.; Miller, S.; Shippy, T.; Flores, M.; Rosendale, A.; Cordola, C.; Bell, T.; Mann, H.; DeAvila, G.; DeAvila, D.; Moore, Z.; Buller, K.; Ciolkevich, K.; Nandyal, S.; Mahoney, R.; Van Voorhis, J.; Dunlevy, M.; Farrow, D.; Hunter, D.; Morgan, T.; Shore, K.; Guzman, V.; Izsak, A.; Dixon, D. E.; Cridge, A.; Cano, L.; Cao, X.; Jiang, H.; Leng, N.; Johnson, S.; Cantarel, B. L.; Richards, S.; English, A.; Shatters, R. G.; Childers, C.; Chen, M. J.; Hunter, W.; Cilia, M.; Mueller, L. A.; Munoz-Torres, M.; Nelson, D.; Poelchau, M. F.; Benoit, J. B.; Wiersma-Koch, H.; D'Elia, T.; Brown, S. J., Improved annotation of the insect vector of citrus greening disease: biocuration by a diverse genomics community. *Database (Oxford)* **2017**, 2017.
31. Fu, L. M.; Niu, B. F.; Zhu, Z. W.; Wu, S. T.; Li, W. Z., CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, 28, (23), 3150-3152.
32. Kessner, D.; Chambers, M.; Burke, R.; Agusand, D.; Mallick, P., ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **2008**, 24, (21), 2534-2536.
33. Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R., Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry* **2002**, 74, (20), 5383-5392.
34. Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R., A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry* **2003**, 75, (17), 4646-4658.
35. Vizcaino, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Research* **2016**, 44, (D1), D447-D456.
36. Vizcaino, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P. A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H. J.; Albar, J. P.; Martinez-Bartolome, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H., ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology* **2014**, 32, (3), 223-226.
37. Sayle, R. A.; Milnerwhite, E. J., Rasmol - Biomolecular Graphics for All. *Trends in Biochemical Sciences* **1995**, 20, (9), 374-376.
38. Keller, A.; Chavez, J. D.; Eng, J. K.; Thornton, Z.; Bruce, J. E., Tools for 3D Interactome Visualization. *Journal of Proteome Research* **2019**, 18, (2), 753-758.
39. Zheng, C. X.; Weisbrod, C. R.; Chavez, J. D.; Eng, J. K.; Sharma, V.; Wu, X.; Bruce, J. E., XLink-DB: Database and Software Tools for Storing and Visualizing Protein Interaction Topology Data. *Journal of Proteome Research* **2013**, 12, (4), 1989-1995.

- 855
- 856 40. Tencer, A. H.; Cox, K. L.; Di, L.; Bridgers, J. B.; Lyu, J.; Wang, X. D.; Sims, J. K.;
857 Weaver, T. M.; Allen, H. F.; Zhang, Y.; Gatchalian, J.; Darcy, M. A.; Gibson, M. D.; Ikebe, J.;
858 Li, W.; Wade, P. A.; Hayes, J. J.; Strahl, B. D.; Kono, H.; Poirier, M. G.; Musselman, C. A.;
859 Kutateladze, T. G., Covalent Modifications of Histone H3K9 Promote Binding of CHD3. *Cell*
860 *Reports* **2017**, 21, (2), 455-466.
- 861
- 862 41. Davey, C. A.; Sargent, D. F.; Luger, K.; Maeder, A. W.; Richmond, T. J., Solvent
863 mediated interactions in the structure of the nucleosome core particle at 1.9 angstrom resolution.
864 *Journal of Molecular Biology* **2002**, 319, (5), 1097-1113.
- 865
- 866 42. Wheeler, D. E., Developmental and Physiological Determinants of Caste in Social
867 Hymenoptera - Evolutionary Implications. *American Naturalist* **1986**, 128, (1), 13-34.
- 868
- 869 43. Wilson, E. O., The Origin and Evolution of Polymorphism in Ants. *Quarterly Review of*
870 *Biology* **1953**, 28, (2), 136-156.
- 871
- 872 44. Simola, D. F.; Ye, C. Y.; Mutti, N. S.; Dolezal, K.; Bonasio, R.; Liebig, J.; Reinberg, D.;
873 Berger, S. L., A chromatin link to caste identity in the carpenter ant *Camponotus floridanus*.
874 *Genome Research* **2013**, 23, (3), 486-496.
- 875
- 876 45. Wojciechowski, M.; Lowe, R.; Maleszka, J.; Conn, D.; Maleszka, R.; Hurd, P. J.,
877 Phenotypically distinct female castes in honey bees are defined by alternative chromatin states
878 during larval development. *Genome Research* **2018**, 28, (10), 1532-1542.
- 879
- 880 46. Dickman, M. J.; Kucharski, R.; Maleszka, R.; Hurd, P. J., Extensive histone post-
881 translational modification in honey bees. *Insect Biochemistry and Molecular Biology* **2013**, 43,
882 (2), 125-137.
- 883
- 884 47. Kruse, A.; Fattah-Hosseini, S.; Saha, S.; Johnson, R.; Warwick, E.; Sturgeon, K.;
885 Mueller, L.; MacCoss, M. J.; Shatters, R. G.; Heck, M. C., Combining 'omics and microscopy to
886 visualize interactions between the Asian citrus psyllid vector and the Huanglongbing pathogen
887 *Candidatus Liberibacter asiaticus* in the insect gut. *Plos One* **2017**, 12, (6).
- 888
- 889 48. Linz, L. B.; Liu, S.; Chougule, N. P.; Bonning, B. C., In Vitro Evidence Supports
890 Membrane Alanine Aminopeptidase N as a Receptor for a Plant Virus in the Pea Aphid Vector.
891 *Journal of Virology* **2015**, 89, (22), 11203-11212.
- 892
- 893 49. Pinheiro, P. V.; Ghanim, M.; Alexander, M.; Rebelo, A. R.; Santos, R. S.; Orsburn, B. C.;
894 Gray, S.; Cilia, M., Host Plants Indirectly Influence Plant Virus Transmission by Altering Gut
895 Cysteine Protease Activity of Aphid Vectors. *Molecular & Cellular Proteomics* **2017**, 16, (4),
896 S230-S243.
- 897
- 898 50. Burmester, T., Origin and evolution of arthropod hemocyanins and related proteins.
899 *Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology*
900 **2002**, 172, (2), 95-107.

51. Signorile, A.; Sgaramella, G.; Bellomo, F.; De Rasmio, D., Prohibitins: A Critical Role in Mitochondrial Functions and Implication in Diseases. *Cells* **2019**, 8, (1).
52. Eveleth, D. D.; Marsh, J. L., Sequence and Expression of the Cc Gene, a Member of the Dopa Decarboxylase Gene-Cluster of *Drosophila* - Possible Translational Regulation. *Nucleic Acids Research* **1986**, 14, (15), 6169-6183.
53. Londono-Renteria, B.; Troupin, A.; Conway, M. J.; Vesely, D.; Ledizet, M.; Roundy, C. M.; Cloherty, E.; Jameson, S.; Vanlandingham, D.; Higgs, S.; Fikrig, E.; Colpitts, T. M., Dengue Virus Infection of *Aedes aegypti* Requires a Putative Cysteine Rich Venom Protein. *Plos Pathogens* **2015**, 11, (10).
54. Chavez, J. D.; Keller, A.; Zhou, B.; Tian, R.; Bruce, J. E., Cellular Interactome Dynamics during Paclitaxel Treatment. *Cell Reports* **2019**, 29, (8), 2371.
55. Audergon, P. N.; Catania, S.; Kagansky, A.; Tong, P.; Shukla, M.; Pidoux, A. L.; Allshire, R. C., Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. *Science* **2015**, 348, (6230), 132-5.
56. Monaghan, L.; Massett, M. E.; Bunschoten, R. P.; Hoose, A.; Pirvan, P. A.; Liskamp, R. M. J.; Jorgensen, H. G.; Huang, X., The Emerging Role of H3K9me3 as a Potential Therapeutic Target in Acute Myeloid Leukemia. *Front Oncol* **2019**, 9, 705.
57. Lu, L.; Chen, X.; Sanders, D.; Qian, S.; Zhong, X., High-resolution mapping of H4K16 and H3K23 acetylation reveals conserved and unique distribution patterns in *Arabidopsis* and rice. *Epigenetics* **2015**, 10, (11), 1044-53.
58. Wang, L.; Zhang, F.; Rode, S.; Chin, K. K.; Ko, E. E.; Kim, J.; Iyer, V. R.; Qiao, H., Ethylene induces combinatorial effects of histone H3 acetylation in gene expression in *Arabidopsis*. *BMC Genomics* **2017**, 18, (1), 538.
59. Ramachandran, S.; Vogel, L.; Strahl, B. D.; Dokholyan, N. V., Thermodynamic Stability of Histone H3 Is a Necessary but not Sufficient Driving Force for its Evolutionary Conservation. *Plos Computational Biology* **2011**, 7, (1).
60. Luger, K.; Richmond, T. J., The histone tails of the nucleosome. *Current Opinion in Genetics & Development* **1998**, 8, (2), 140-146.
61. Li, K. L.; Zhang, L.; Yang, X. M.; Fang, Q.; Yin, X. F.; Wei, H. M.; Zhou, T.; Li, Y. B.; Chen, X. L.; Tang, F.; Li, Y. H.; Chang, J. F.; Li, W.; Sun, F., Histone acetyltransferase CBP-related H3K23 acetylation contributes to courtship learning in *Drosophila*. *Bmc Developmental Biology* **2018**, 18.

62. Chen, S. D.; Rufiange, A.; Huang, H. D.; Rajashankar, K. R.; Nourani, A.; Patel, D. J., Structure-function studies of histone H3/H4 tetramer maintenance during transcription by chaperone Spt2. *Genes & Development* **2015**, 29, (12), 1326-1340.
63. Hyland, E. A.; Cosgrove, M. S.; Molina, H.; Wang, D. X.; Pandey, A.; Cottee, R. J.; Boeke, J. D., Insights into the role of histone H3 and histone H4 core modifiable residues in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **2005**, 25, (22), 10060-10070.
64. Yan, Q. S.; Dutt, S.; Xu, R.; Graves, K.; Juszczynski, P.; Manis, J. P.; Shipp, M. A., BBAP Monoubiquitylates Histone H4 at Lysine 91 and Selectively Modulates the DNA Damage Response. *Molecular Cell* **2009**, 36, (1), 110-120.
65. Ye, J. X.; Ai, X.; Eugeni, E. E.; Zhang, L. W.; Carpenter, L. R.; Jelinek, M. A.; Freitas, M. A.; Parthun, M. R., Histone H4 lysine 91 acetylation: A core domain modification associated with chromatin assembly. *Molecular Cell* **2005**, 18, (1), 123-130.
66. Omote, H.; Park, M. Y.; Maeda, M.; Futai, M., The Alpha/Beta Subunit Interaction in H⁺-Atpase (Atp Synthase) - an Escherichia-Coli Alpha-Subunit Mutation (Arg-Alpha-296 -] Cys) Restores Coupling Efficiency to the Deleterious Beta-Subunit Mutant (Ser-Beta-174 -] Phe). *Journal of Biological Chemistry* **1994**, 269, (14), 10265-10269.
67. Ibanez, F.; Tang, X.; Tamborindeguy, C., *Bactericera cockerelli* vitellogenin-6 like, a vitellogenin without a direct reproductive function? *Insect Molecular Biology* **2018**, 27, (2), 166-176.
68. Hosseinzadeh, S.; Ramsey, J.; Mann, M.; Bennett, L.; Hunter, W. B.; Shams-Bakhsh, M.; Hall, D. G.; Heck, M., Color morphology of *Diaphorina citri* influences interactions with its bacterial endosymbionts and '*Candidatus* Liberibacter asiaticus'. *PLoS One* **2019**, 14, (5), e0216599.
69. Hergeth, S. P.; Schneider, R., The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *Embo Reports* **2015**, 16, (11), 1439-1453.
70. Kalashnikova, A. A.; Rogge, R. A.; Hansen, J. C., Linker histone H1 and protein-protein interactions. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms* **2016**, 1859, (3), 455-461.
71. Li, Y. L.; Li, Z. M.; Dong, L. P.; Tang, M.; Zhang, P.; Zhang, C. H.; Cao, Z. Y.; Zhu, Q.; Chen, Y. C.; Wang, H.; Wang, T. Z.; Lv, D. Y.; Wang, L. N.; Zhao, Y.; Yang, Y.; Wang, H. Y.; Zhang, H. Q.; Roeder, R. G.; Zhu, W. G., Histone H1 acetylation at lysine 85 regulates chromatin condensation and genome stability upon DNA damage. *Nucleic Acids Research* **2018**, 46, (15), 7716-7730.