Phytopathology® • 2022 • 112:2273-2287 • https://doi.org/10.1094/PHYTO-02-22-0067-R

Bacteriology

Strain Tracking of '*Candidatus* Liberibacter asiaticus', the Citrus Greening Pathogen, by High-Resolution Microbiome Analysis of Asian Citrus Psyllids

Steven A. Higgins,[1] Marina Mann,[2] and Michelle Heck[1,2,†]

[1] Emerging Pests and Pathogens Research Unit, United States Department of Agriculture–Agricultural Research Service (USDA–ARS), Ithaca, NY 14853

[2] Plant Pathology and Plant Microbe Biology Department, Cornell University, Ithaca, NY 14853

Accepted for publication 6 June 2022.

Abstract

The Asian citrus psyllid, *Diaphorina citri,* is an invasive insect and a vector of '*Candidatus* Liberibacter asiaticus' (*C*Las), a bacterium whose growth in *Citrus* species results in huanglongbing (HLB), also known as citrus greening disease. Methods to enrich and sequence *C*Las from *D. citri* often rely on biased genome amplification and nevertheless contain significant quantities of host DNA. To overcome these hurdles, we developed a simple pretreatment DNase and filtration (PDF) protocol to remove host DNA and directly sequence *C*Las and the complete, primarily uncultivable microbiome from *D. citri* adults. The PDF protocol yielded *C*Las abundances upward of 60% and facilitated direct measurement of *C*Las and endosymbiont replication rates in psyllids. The PDF protocol confirmed our lab strains derived from a progenitor Florida *C*Las strain and accumulated 156 genetic variants, underscoring the utility of this method for bacterial strain tracking. *C*Las genetic polymorphisms arising in lab-reared psyllid populations included prophage-encoding regions with key functions in *C*Las pathogenesis, putative antibiotic resistance loci, and a single secreted effector. These variants suggest that laboratory propagation of *C*Las could result in different phenotypic trajectories among laboratories and could confound *C*Las physiology or therapeutic design and evaluation if these differences remain undocumented. Finally, we obtained genetic signatures affiliated with *Citrus* nuclear and organellar genomes, entomopathogenic fungal mitochondria, and commensal bacteria from laboratory-reared and field-collected *D. citri* adults. Hence, the PDF protocol can directly inform agricultural management strategies related to bacterial strain tracking, insect microbiome surveillance, and antibiotic resistance screening.

Keywords: Asian citrus psyllid, '*Candidatus* Liberibacter asiaticus', citrus greening, *Diaphorina citri*, high-throughput sequencing, insect microbiome

The abundance of host DNA in nucleic acid extracts is a significant barrier to cost-effective sequence analysis of host-associated microbiomes (Marotz et al. 2018). Microbial populations are also nonuniform in abundance and spatial organization, and sequencing effort is therefore directed toward host DNA, which limits the cost-effectiveness and utility of shotgun sequence data in microbiome analysis (Pereira-Marques et al. 2019). To help overcome these issues, a combination of pre- and post-DNA extraction methods have been devised to limit host DNA in nucleic acid extracts (Marotz et al. 2018). These methods include differential lysis of host tissues followed by exonuclease treatment (Marotz et al. 2018), host DNA depletion via methylation status followed by multiple displacement amplification (Liu et al. 2020; Wu et al. 2015; Zheng et al. 2014), use of targeted hybridization probes for selective enrichment of microbial DNA (Cai et al. 2019; Lefoulon et al. 2019), and cell enrichment by cultivation in alternative hosts (Li et al. 2021). While useful, these methods are relatively costly, are laborious, and can restrict analysis to a single taxon (e.g., probe hybridization) or introduce biases (e.g., using multiple displacement amplification) that limit the types of downstream analyses afforded by shotgun sequence data (Marine et al. 2014).

One host-microbe system that illustrates the impediment of host DNA in microbiome analyses is that of the citrus greening disease pathosystem (Duan et al. 2009; Tyler et al. 2009). Citrus trees worldwide are threatened by the rapidly spreading citrus greening disease, which, over a period of years, significantly reduces citrus fruit yield and marketability and eventually results in tree decline and death (Ghosh et al. 2018; Johnson et al. 2014; Wang and Trivedi 2013). Citrus greening disease is estimated to cost some U.S. citrus growers $1 billion annually and has reduced Florida citrus production by over 70% since the early 2000s (Dala-Paula et al. 2019; S. Li et al. 2020). In the United States, the disease is attributed to the growth of the obligately biotrophic, gram-negative bacterium '*Candidatus* Liberibacter asiaticus' (*C*Las) within citrus trees, which induces phloem-plugging and results in the characteristic citrus greening pathology of aborted ripening, blotchy leaf mottle, and tree root decline (Dala-Paula et al. 2019; Johnson et al. 2014; Wang and Trivedi 2013). *C*Las is proposed to be transmitted by its insect vector, the Asian citrus psyllid (*Diaphorina citri* Kuwayama, Hemiptera: Liviidae), in a circulative, propagative manner that involves extensive interaction with and replication within the vector's tissues (Ammar et al. 2016; Inoue et al. 2009). *D. citri* acquire *C*Las during the nymphal stage and become competent to transmit as adults when *C*Las titers peak in the insect's salivary glands (Ammar et al. 2016). Like the majority of phloem-feeding hemipteran insects, *D. citri* harbor bacterial symbionts whose functions in their biology have

†Corresponding author: M. Heck; michelle.cilia@usda.gov

Author contributions: S.A.H. conceived the idea, S.A.H. and M.H. developed the idea and planned experiments, M.M. performed field collection of *D. citri* adults, S.A.H. performed experiments and data analysis, and S.A.H., M.M., and M.H. wrote and revised the manuscript.

Funding: Support was provided by the California Citrus Research Board (grant 5300-202), the U.S. Department of Agriculture–Agricultural Research Service (CRIS projects 0500-00093-0001-00-D and 8062-22410-007-000-D), and of a U.S. Department of Agriculture-National Institute of Food and Agriculture predoctoral fellowship to M. Mann.

***e*-Xtra:** Supplementary material is available online.

The author(s) declare no conflict of interest.

been inferred from genome sequencing: *Wolbachia pipientis* strain wDi (Saha et al. 2012), '*Candidatus* Profftella armatura', a symbiont with predicted nutritional and defensive functions (Nakabachi et al. 2013, 2020), and '*Candidatus* Carsonella rudii', a nutritional symbiont (Nakabachi et al. 2006; Saha et al. 2012). These latter two symbionts are essential for *D. citri* survival, but their role in *C*Las transmission is speculative (Ukuda-Hosokawa et al. 2015). *D. citri* have spread *C*Las throughout citrus-growing regions in the United States since its initial discovery in Florida in 2005 (Wang and Trivedi 2013). Although we report on the impacts of citrus greening to the U.S. citrus industry for simplicity, we would like to emphasize that this disease presents a significant threat to citrus production worldwide (Ghosh et al. 2018; Wang 2019). Particularly alarming is that a handful of countries globally produce a majority of the world's citrus crops (Ghosh et al. 2018; Wang 2019) and that some at-risk countries possess large swaths of habitat favoring the spread of citrus greening disease (Ajene et al. 2020; Wang 2019).

Historically, the cultivation of microbial plant and insect pathogens has facilitated their experimentation, comparative sequence analysis, and design of therapeutics to control their spread, whereas currently uncultivable *Ca*. Liberibacter pathogens (such as *C*Las) are difficult to study due to their obligately biotrophic lifestyle (Haapalainen 2014). Because attempts to isolate and continuously culture *C*Las have been unsuccessful (Merfa et al. 2019), high-throughput sequencing has become an invaluable tool for hypothesis generation and performing metabolic, comparative genomics, and evolutionary analyses of this pathogenic bacterium (Coyle et al. 2019; Thapa et al. 2020; Wang and Trivedi 2013). Nevertheless, many hurdles to efficient and direct sequencing of *C*Las populations from both citrus and *D. citri* exist (Duan et al. 2009; Tyler et al. 2009). For example, due to the overabundance of host DNA in nucleic acid extracts from citrus and psyllid samples, multiple enrichment strategies that include deep sequencing, multiple displacement amplification of host-depleted DNA, and antibody- or RNA-probe capture hybridization are employed to facilitate sequence analysis of *C*Las and other *Ca.* Liberibacter populations (Cai et al. 2019; Duan et al. 2009; Katoh et al. 2014; Lin et al. 2011; Merfa et al. 2019).

Robust and facile sequencing of *C*Las and other pathogenic, host-associated microbial populations is necessary to keep pace with the rapid development and application of therapeutics to prevent their spread (Blaustein et al. 2018). Although a variety of antimicrobials have been applied to prevent the spread of citrus greening disease (Blaustein et al. 2018; S. Li et al. 2020), we are currently unable to perform cost-effective and rapid sequencing of *C*Las populations to screen for genetic resistance to therapeutics if and when it arises. Furthermore, methods to identify and compare *C*Las populations have been limited to a few dozen samples that have been sequenced worldwide (Thapa et al. 2020). This small sample size limits the identification and development of targeted therapeutics to control *C*Las spread and hampers our ability to contrast standing genetic variation detected in laboratory-propagated *C*Las populations with those actively spreading in agricultural settings. Thus, we developed a simple and efficient pretreatment DNase and filtration (PDF) protocol to remove host DNA and sequence *C*Las and other bacterial populations directly from individual members of their insect host, *D. citri*. We compare and contrast our method with existing sequencing methods used to sequence *C*Las populations and present results of our genomic analyses enabled by the PDF protocol on both lab-reared and field-collected *D. citri* sampled from a minimally managed citrus grove in Florida. Specifically, we report on replication rate estimates and genetic polymorphisms detected among microbial populations within individual *D. citri* adults and discuss their implications for citrus greening disease management and insect microbiome research. The PDF protocol enables simultaneous analysis of all microbial community members in a single *D. citri* adult, from symbionts to commensals, that will greatly inform diverse research fields, such as therapeutics development and application (Sundin and Wang 2018), holobiont theory (Simon et al. 2019), and ecological and evolutionary investigations of *D. citri* and other agriculturally relevant pests (Badial et al. 2018; Burdon and Thrall 2008).

Materials and Methods

Rearing and collection of laboratory-reared *D. citri* adults

Colonies of *D. citri* were reared in controlled growth chambers on *C*Las-infected *Citrus medica* (citron) plants under a 14/10 h light/dark cycle at 28°C. Citrus plants were regularly monitored by lab technicians and maintained by pruning and watering on a weekly basis or as needed. Individual *D. citri* adults from *C*Las-infected citron colonies were bulk-collected using a sterile 25-ml serological pipette under vacuum into a sterile 50-ml falcon tube. The falcon tube containing *D. citri* adults was then transferred onto ice to immobilize the insects prior to handling and sorting. Immobilized *D. citri* were then transferred into a plastic weigh boat on ice, and individual *D. citri* were placed into sterile 2-ml microcentrifuge tubes using a 70% (vol/vol) ethanol-wiped, fine-tipped paintbrush and kept on ice until processing for DNA extraction. A total of 33 and nine laboratory-reared *D. citri* adults were processed using the PDF and Direct protocols, respectively.

Field collection of *D. citri* adults

With permission from a local citrus grower, *D. citri* adults were collected from *Citrus sinensis* (Valencia orange) trees in a minimally managed citrus grove located in St. Lucie County, FL, in early October of 2018. This grove did not utilize pesticides, and natural enemies of *D. citri* were frequently observed (e.g., spiders and parasitoid wasps) (Kondo et al. 2015). Individual *D. citri* were collected from citrus leaf surfaces using mouth aspiration into sterile borosilicate glass collection vials from a few dozen citrus trees. The individual *D. citri* were then transferred into 95% ethanol (vol/vol) and kept stored on ice. In the lab, the ethanol was decanted and the *D. citri* adults were air dried on Kimwipes until storage at −80°C prior to DNA extraction.

PDF protocol and sequencing

To each tube containing an individual *D. citri* adult, we added 100 µl of master mix containing 8 U DNase I (TURBO DNA-free kit, catalog number AM1907, Thermo Fisher Scientific, Waltham, MA, U.S.A.), 1× DNase I buffer, 2× ProteoGuard EDTA-free protease inhibitor cocktail (catalog number 635673, Takara Bio USA, San Jose, CA, U.S.A.), and DNase and RNase-free RPI molecular grade water (catalog number 248700-500.0, Neta Scientific, Hainesport, NJ, U.S.A.). Psyllids were then manually homogenized within each tube for 30 s using an autoclaved Bel-Art polypropylene microcentrifuge pestle (catalog number 199230001, SP Industries, South Wayne, NJ, U.S.A.). Following homogenization, samples were briefly centrifuged and gently flicked with a finger to mix the tube contents, followed by incubation in a water bath at 32°C for 2 h. Tubes were gently mixed by hand agitation during incubation at 30-min intervals. Following this incubation, samples were placed on ice, and 100 µl of sterile 1× TE buffer (10 mM Tris HCl, 1 mM EDTA, pH 8.0) was added. The total mixture was then transferred to autoclaved, 1.5-ml plastic microcentrifuge tubes containing a 5 µm pore size PVDF filter cartridge (catalog number UFC30SV00, MilliporeSigma, Burlington, MA, U.S.A.), followed by centrifugation at 8,000 × *g* for 3 min. The supernatant was carefully removed by pipetting, and the resulting cell pellet was placed on ice. DNA was then extracted from the cell pellet by addition of 600 µl of buffer RLT (catalog number 79216, Qiagen Sciences, Germantown, MD, U.S.A.), followed by gentle pipetting to reconstitute the cell pellet. The mixture was then transferred to a sterile 2-ml microcentrifuge tube containing pre-autoclaved and air-dried 0.1-mm zirconia-silicate beads (catalog number 11079101Z, Bio Spec Products, Bartlesville, OK, U.S.A.) and homogenized using a

mixer mill 400 (Retsch GmbH, Haan, Germany) at 25 Hz for 1 min. A 450-µl aliquot of the homogenate was then transferred to a sterile 2-ml EconoSpin mini spin column (catalog number 1910-250, Epoch Life Science, Sugar Land, TX, U.S.A.) containing 450 µl of 70% (vol/vol) molecular grade ethanol (catalog number 3916EA, Decon Laboratories, King of Prussia, PA, U.S.A.) and gently pipetted up and down to mix. The samples were centrifuged at 8,000 rpm for 1 min, the flow-through discarded, and an additional 700 µl of 75% (vol/vol) molecular grade ethanol was added to the column followed by centrifugation at 8,000 rpm for 30 s and the flow-through discarded. This process was repeated an additional two times, followed by a 2-min centrifugation of the empty column at 12,000 rpm to remove residual ethanol. The DNA was eluted off the column using 25 µl of 0.1× TE buffer and DNA concentration and quality evaluated using a NanoDrop 2000 instrument (Thermo Fisher Scientific). DNA samples were provided to the Microbial Genome Sequencing Center (Pittsburgh, PA, U.S.A.; renamed SeqCenter as of June 6, 2022) for 2 × 151 bp sequencing on an Illumina NextSeq 2000 instrument at 200 Mbp or 1 Gbp depth offerings using a modification of the standard Illumina Nextera tagmentation protocol (Baym et al. 2015). Of note, the range of DNA concentrations in these samples was very low (~0.1 to 0.6 ng/µl), and the sequencing center used 10 µl for library preparation. In earlier iterations of this workflow, preparation of sequencing libraries was unsuccessful using DNA samples not stored or shipped on ice. Any DNA derived from this method should therefore be kept on ice when handled for in-house sequencing or shipped on ice or dry ice to sequencing facilities to ensure stability of the minute DNA quantities present in the processed samples. In addition to the PDF protocol, *D. citri* adults were placed into a sterile 2-ml microcentrifuge tube containing two sterile 3.2-mm steel beads and were snap frozen in liquid nitrogen. Next, samples were cryoground on the Retsch mixer mill 400, their DNA extracted, and samples submitted for sequencing as described above. We refer to this last step as the "Direct protocol" and compare and contrast results from both the PDF and Direct protocols in the results section (Fig. 1).

Quantitative polymerase chain reaction (qPCR) analysis

Prior to sequencing, DNA samples were analyzed by qPCR using gene-specific PCR primers to monitor *C*Las enrichment relative to DNA from its *D. citri* host. The PCR primers targeted either the *C*Las 16S rRNA gene (5′-TCGAGCGCGTATGCGAATAC-3′ and 5′-GCGTTATCCCGTAGAAAAAGGTAG-3′ forward and reverse primers, respectively) or the *D. citri* wingless (*wg*) gene (5′-GCTCTCAAAGATCGGTTTGACGG-3′ and 5′-GCTGCCA CGAACGTTACCTTC-3′ forward and reverse primers, respectively). PCR amplification was performed using 2X PowerSYBR green master mix (catalog number 4367695, Thermo Fisher Scientific). Briefly, the qPCR reaction was composed of 5 µl of 2X PowerSYBR green master mix, 0.25 µl each of the forward and reverse primers (10 µM stock), 0.125 µl of molecular grade bovine serum albumin (20,000 ng/µl stock, catalog number B9000S, New England Biolabs, Ipswich, MA, U.S.A.), 1 µl of DNA (~0.1 to 0.6 ng/µl), and 3.375 µl of molecular grade water. The qPCR

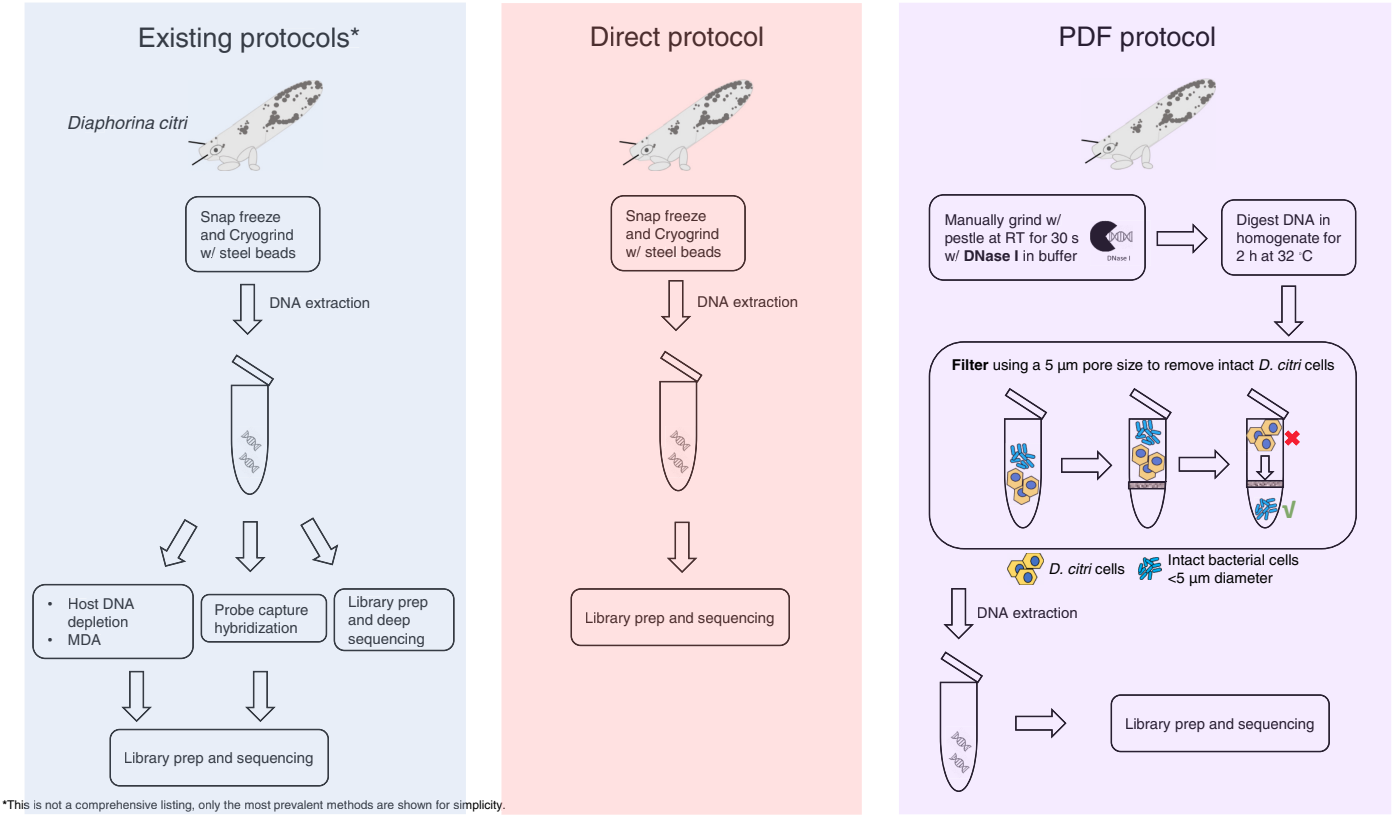

Fig. 1. Overview of the protocols to enrich and sequence '*Candidatus* Liberibacter asiaticus' (*C*Las) from its insect host, the Asian citrus psyllid, *Diaphorina citri*. The existing protocols to enrich and sequence *C*Las in the literature include post DNA extraction treatments, including host DNA depletion by methylation status (host DNA depletion, above) followed by multiple displacement amplification (MDA), use of complementary RNA probes targeting the *C*Las genome via probe capture hybridization, or deep sequencing of *Citrus* spp. or *D. citri* tissues with high *C*Las titers. The Direct protocol includes directly extracting DNA from *C*Las-infected *D. citri* and submitting for library preparation and DNA sequencing. The Direct protocol is used for comparison with the pretreatment DNase and filtration (PDF) protocol in the main text. Finally, the PDF protocol includes manually grinding *D. citri* tissues using a microcentrifuge pestle in a buffer containing DNase I enzyme, digestion of DNA for 2 h, following by filtration of the cell lysate through a five micron pore size filter by centrifugation. Then, DNA extraction is performed on the resulting cell pellet, followed by library preparation and sequencing. The asterisk next to the "Existing protocols" in the figure indicates that this is not a comprehensive listing of all methods to enrich and sequence *C*Las from its hosts, only that the most prevalent approaches have been presented for simplicity.

amplification was performed on an Applied Biosystems QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific). Thermocycler conditions consisted of a 2-min incubation at 50°C followed by a 10-min incubation at 95°C and 40 cycles at 95°C for 15 s and 55°C for 1 min followed by denaturation at 95°C for 15 s and melt curve analysis from 60°C (1 min) to 95°C (15 s) at a rate of 0.05°C/s. The qPCR Cq values were converted to *C*Las 16S rDNA gene copies using duplicate standard curves comprised of eight 10-fold serial dilutions of a known concentration of pUC57 plasmid (2,710 bp length) containing a *C*Las 16S rDNA gene fragment (427 bp length) prior to comparison with *D. citri* wingless genes. A similar methodology was used to calculate *D. citri wg* gene copies, except that genomic DNA was extracted as previously described (Igwe et al. 2021) from a pool of *D. citri* adults ($n = 10$) followed by six 10-fold serial dilutions to construct a *D. citri* genomic DNA standard curve. An estimated *D. citri* genome size of 485,705,802 bp was used to calculate *wg* gene copies for a given quantity of DNA in the dilution series. The range of *C*Las 16S rDNA gene copies in the standard curve dilution series were 8.4×10^2 to 8.4×10^9, and the assay possessed a slope of −3.13, y-intercept of 39.70, R^2 of 0.986, and efficiency of 108.7%. The range of *wg* gene copies in the standard curve dilution series were 2.5×10^{-1} to 2.5×10^5, and the *wg* qPCR assay possessed a slope of −2.98, y-intercept of 37.96, adjusted R^2 of 0.95, and efficiency of 118.37%. All gene copies are reported as gene copies per microliter of DNA and are hereafter referred to as "gene copies" for simplicity.

Data acquisition and analysis

For purposes of comparison, raw read data sets of *C*Las genomes derived from citrus tissues were accessed from the National Center for Biotechnology and Information (NCBI) Sequence Read Archive (SRA) (Leinonen et al. 2010) using accession numbers SRR016816, SRR7050882, SRR8177702, SRR9641236, and SRR9673120. These Illumina reads represent *C*Las genomic data from an unknown *C*Las strain identified in *Citrus sinensis* in Florida, U.S.A.; strain TX1712 collected from *C. sinensis* in Texas, U.S.A. (Cai et al. 2018); strain YNJS7C identified in *C. sinensis* from Yunnan province, China (Chen et al. 2019); strain AHCA17 identified in *C. maxima* (pummelo) in California, U.S.A. (Cai et al. 2020); and strain TaiYZ2 collected from *C. maxima* (pummelo) in Songkhla province, Thailand (T. Li et al. 2020), respectively.

The latest *D. citri* (version 3), *C. medica* (version 1, HZAU), and *C. sinensis* genome (version 2, HZAU) assemblies were downloaded from https://citrusgreening.org/ and https://citrusgenomedb.org/, respectively. All raw fastq sequence files were processed with FaQCs (v2.10) (Lo and Chain 2014) to trim (phred score cutoff ≥30) and remove low-quality and short (<50 bp length) Illumina reads. Individual samples and a total combined sample set were assembled or co-assembled, respectively, using the SPAdes assembler (v3.14.0) (Bankevich et al. 2012) with k-mer sizes of 33, 55, and 77 for individual samples and 77, 99, and 127 for the co-assembly (all raw reads from individual samples pooled prior to assembly). Trimmed Illumina reads were aligned to assembled scaffolds using bowtie2 (v2.3.4) (Langmead and Salzberg 2012) and coverage calculated using samtools (v1.9) (Li et al. 2009), mpileup, and coverage modules with default settings. All scaffolds were aligned to the nt database using BLAST+ (v2.11.0) (Camacho et al. 2009) and taxonomic assignments inferred using the BASTA (v 1.4.1) (Kahlke and Ralph 2019) "sequence" module with settings -i 80 -b True -p 90. Blobtools (v1.1.1) (Laetsch et al. 2020) was used to visualize GC%, coverage, and taxonomic assignments of assembled scaffolds (data not shown) and inspired the plots used in portions of the results and Supplementary Materials. Automatic binning of assembled scaffolds into coherent operational taxonomic units was performed using LAST (version 1145) (Kielbasa et al. 2011) alignments and BASTA taxonomic assignments, followed by bin consistency analysis using the CheckM (v1.1.3) tool's "lineage_wf" module to assess the bin completeness and consistency metrics (Parks et al. 2015). QUAST (v5.1.0rc1, settings -m 100 -f –rna-finding -b) (Mikheenko et al. 2016) was used to compare assembled genome bins of *D. citri* symbionts with high-quality, complete reference symbiont genomes and report assembly quality statistics.

The iRep tool (Brown et al. 2016) with default settings was used to estimate the index of replication (average, instantaneous population replication rate) against representative reference genomes for *C*Las, *Wolbachia pipientis* strain wDi, '*Candidatus* Profftella armatura' (hereafter Prof), and '*Candidatus* Carsonella ruddii' (hereafter Car) populations detected within individual *D. citri* adults processed in the present study. Snippy (v4.6.0) (Seeman 2015) was used with parameters –mincov 10 –minfrac 9 and –minqual 100 to detect, enumerate, and compare genetic polymorphisms among *D. citri* symbionts and reference symbiont genomes. Gubbins (v2.4.1) (Croucher et al. 2015) was used with settings –tree_builder raxml –raxml_model GTRGAMMA and additional default settings to flag and generate a recombination-free single nucleotide polymorphism (SNP) alignment among whole *C*Las genomes assembled in the present study and publicly available reference genomes, which was supplied to IQ-TREE (v1.6.1) (Nguyen et al. 2015). IQ-TREE was run with settings -m K3Pu + F -bb 1000 -bnni to generate a SNP phylogeny. Some *C*Las genomes (both publicly available and in the present study) were highly fragmented and/or missing >25% of the genome and were not included in the phylogenetic analysis of SNPs. Visualization of SNP data and phylogeny was performed using the R packages BioCircos v0.3.4 (Chen and Cui 2019) and ggtree (Yu et al. 2017).

Statistical analysis

All statistics were performed in R 4.1.2 (R Core Team 2018). The function t.test() was used to perform a two-sample Welch's *t* test to examine whether coverage of select *D. citri* bacterial populations between protocols were equal. Significant deviation from the null expectation was accepted at $\alpha < 0.05$.

Data availability

Raw Illumina sequence data were deposited to NCBI Sequence Read Archive under BioProject PRJNA779156. The assembled data, their taxonomic assignments, SNP alignment and phylogeny, and additional raw data presented in the manuscript are accessible in the figshare repository (https://doi.org/10.6084/m9.figshare.c.5810090.v1).

Code availability

All scripts used to process data and generate figures or analyses can be found in the figshare repository (https://doi.org/10.6084/m9.figshare.c.5810090.v1).

Results

Comparison of *C*Las enrichment and sequencing protocols using lab-reared *D. citri*

A variety of methods, including deep sequencing, multiple displacement amplification, methylation-based DNA depletion, and probe capture hybridization, have been published to enrich and sequence *C*Las from both citrus and *D. citri* samples (hereafter referred to as "existing protocols", Supplementary Table S1). Because access to raw sequence data (i.e., Illumina reads) for most published *C*Las genomes produced using existing protocols were unavailable on public sequence repositories, such as the SRA (Leinonen et al. 2010), we performed comparisons using reported total read counts or the number of read alignments from the literature (Supplementary Table S1; Fig. 2A). Therefore, we compared sequence data produced using the PDF protocol, which uses pre-extraction DNase treatment and filtration to remove host DNA, with data produced by direct DNA extraction from *C*Las-infected *D. citri* adults (Direct protocol, Fig. 1). We reasoned that robust comparisons of

*C*Las-infected *D. citri* adults with variable *C*Las titers could be made using the PDF and Direct protocols, whereas the same analysis was not feasible with other published methods because publicly available raw data sets were lacking. Regardless, improvements to host DNA removal facilitated by the PDF protocol were made possible by comparison with the Direct protocol.

To determine the effect of the PDF protocol on *C*Las titers relative to the insect host *D. citri* DNA, we quantified *C*Las 16S rRNA and *D. citri* wingless (*wg*) gene copies in DNA samples generated using the PDF protocol and compared these values with the Direct protocol (Fig. 1). The values of *C*Las 16S rRNA gene copies for the PDF and Direct protocols ranged from 1.1×10^5 to 3.2×10^7 and 1.4×10^3 to 1.6×10^7, respectively, whereas *D. citri wg* gene copies ranged between 8.7×10^2 to 2.4×10^3 and 1.3×10^5 to 1.2×10^6, respectively. Thus, *D. citri* genomic DNA (which encodes the *wg* gene) was reduced by up to three orders of magnitude when the PDF protocol was utilized (Supplementary Fig. S1).

To assess the impact of host DNA removal on sequence coverage of *C*Las populations, we sequenced DNA samples prepared using the PDF and Direct protocols that possessed a range of *C*Las 16S rRNA gene copies for Illumina shotgun sequencing. The reduction in *D. citri wg* gene copies we detected also corresponded to a significant reduction in the average coverage of the *D. citri* genome (mean coverage of 0.3 and 0.9 for the PDF and Direct protocols,

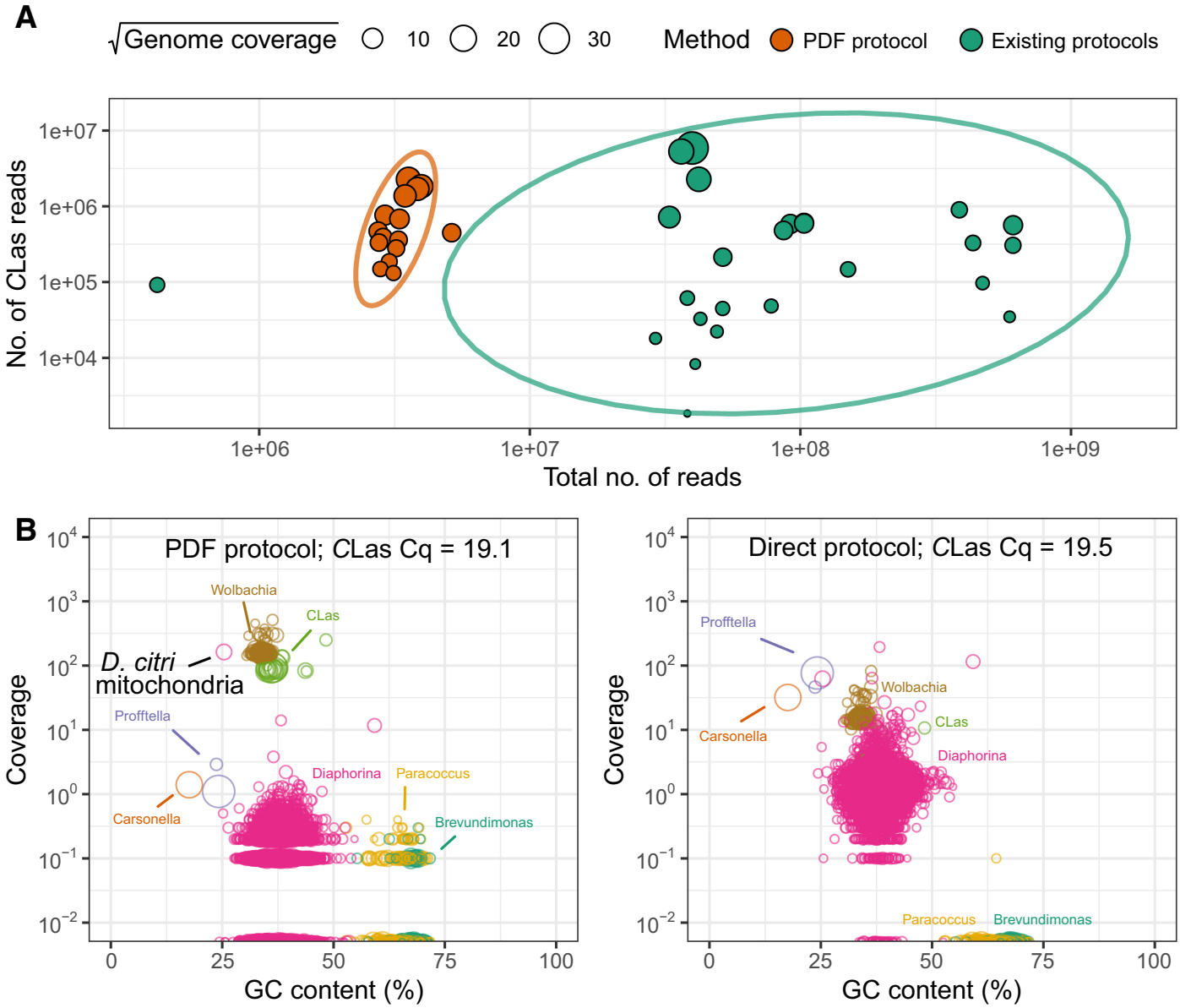

Fig. 2. Comparison of '*Candidatus* Liberibacter asiaticus' (*C*Las) and additional microbial community member abundances **A,** between the pretreatment DNase and filtration (PDF) protocol and existing *C*Las sequencing protocols in the literature or **B,** between the PDF and Direct protocols using laboratory-reared *Diaphorina citri* adults with equivalent *C*Las titers. In the top panel, the PDF protocol (orange circles) generates equivalent or greater amounts of *C*Las data with orders of magnitude fewer sequencing data than existing protocols (green circles), reducing sequencing costs and increasing useable microbiome data. B, In addition to increasing *C*Las abundance relative to the host, the PDF protocol also improves detection and analysis of *Wolbachia pipientis* strain wDi and other commensal bacteria (*Paracoccus* sp. and *Brevundimonas* sp.). For *D. citri* individuals with comparable *C*Las titers (assessed via qPCR; text inside bottom panels), the relative abundance of *C*Las and *Wolbachia* in these samples is an order of magnitude greater with the PDF protocol compared to directly sequencing DNA extracted from an individual *D. citri* adult ("Direct protocol" in bottom right panel). The size and color of points in B represent the length and taxonomic affiliations of contigs generated from the coassembly of sequence reads from all lab-reared *D. citri* individuals processed in the present study. Hence, taxa with fewer, larger circles represent larger, more contiguous assemblies. Taxa with the most abundant and numerous contigs with consistent taxonomic affiliations are shown for clarity. Taxonomic annotations of all contigs can be accessed via the figshare repository link in the data availability section. Plots in B were inspired by BlobTools (Laetsch et al. 2020).

respectively; $t_{welch} = 8.6$, df = 33, $P = 6.7 \times 10^{-10}$) and a concomitant increase in sequencing effort afforded to *C*Las populations (Supplementary Fig. S1). In fact, when we compared the PDF protocol with existing protocols for sequencing *C*Las genomes available in the literature (Supplementary Table S1), the PDF protocol provided comparable coverage of *C*Las populations at orders of magnitude lower sequencing depth (Fig. 2A). Even for samples with high *C*Las titers (i.e., low Cq values observed for *C*Las 16S rRNA gene qPCR assays), average genome coverages of *C*Las populations in DNA samples generated using the PDF protocol were on average 35-fold greater (mean coverage of 43.2 and 1.24 for PDF and Direct protocols, respectively; $t_{welch} = -3.37$, df = 31, $P = 0.002$) than directly sequencing DNA extracted from an individual *D. citri* adult ("Direct protocol", Fig. 2B). In addition to *C*Las, the PDF protocol significantly increased coverage of *Wolbachia* symbiont populations within *D. citri* adults by more than sevenfold compared with the Direct protocol ($t_{welch} = -11.8$, df = 33, $P = 2.7 \times 10^{-13}$, Fig. 2B). Whereas average coverage of *Wolbachia* symbiont populations approached 20× in the Direct protocol, they increased to 140× using the PDF protocol, greatly enhancing sequencing effort toward additional symbionts present within *D. citri* adults. Conversely, average coverage of the *D. citri* symbionts Car and Prof were significantly reduced by six- and threefold, respectively, when comparing the PDF and Direct protocols, respectively. Nevertheless, average coverage of Car and Prof populations were 5× and 39×, respectively, using the PDF protocol, which is sufficient for downstream genome assembly or comparative genomic analyses.

Aside from genome coverage estimation, we performed de novo genome assembly of individual and pooled samples (i.e., a coassembly) and examined the taxonomic identity of each contig by alignment to high-quality reference genome sequences or NCBI's nucleotide database in concert with the BASTA tool for taxonomic annotation (Kahlke and Ralph 2019). Assembly statistics of de novo genome assemblies for all *D. citri* symbionts analyzed in the present study can be found in Supplemental Materials (Supplementary Table S2). Of note, we assembled draft bacterial genomes from laboratory-reared *D. citri* adults assigned to the genus *Paracoccus* and *Brevundimonas* at 85 and 96% completeness, respectively, and with less than 5% contamination as assessed via the CheckM tool (Parks et al. 2015). Due to the uneven abundances of these taxa among *D. citri* individuals and their low coverage (Fig. 2B), their genomes were highly fragmented, with low N50 values of 3,480 and 8,650 nt for *Paracoccus* and *Brevundimonas*, respectively, which are the sequence lengths at which sequences this long or longer compose at least 50% of the genome length. All assembled microbiome sequence data reported for laboratory-reared and field-collected *D. citri* adults (see below) can be accessed in the data availability section.

Estimation of replication rates in *D. citri* endosymbionts

The depth of genome coverage of *C*Las and the other psyllid endosymbionts allowed us to estimate genome replication rates for all four symbiont populations detected within individual *D. citri*. The genomic region predicted to contain the origin of replication (*oriC*) in bacteria can be distinguished by both GC skew (Fig. 3A, green line), a peak in average genome coverage (Fig. 3A, gray line), or conserved genetic elements (usually *dnaA*) associated with the *oriC*, but whose genomic location differs in the alphaproteobacterial symbionts investigated herein and other genes are localized there instead (e.g., *hemE*, Fig. 3A) (Brown et al. 2016; Ioannidis et al. 2007). Using the iRep tool, which performs regression analysis of ordered sequence read coverage over complete or draft microbial genome assemblies to estimate the index of replication (iRep), we were able to calculate an instantaneous measure of the average population replication rate (Brown et al. 2016) for the *D. citri* endosymbionts (Fig. 3B). The iRep values for the four symbionts were ordered Car > Prof > *C*Las > *Wolbachia* (Fig. 3C), but they differed between the PDF and Direct protocols. These differences were largely attributed to differential coverage between symbiont populations when comparing both protocols (Fig. 3D). For example, low sequence coverage of *C*Las and *Wolbachia* populations tended to result in higher iRep estimates in the Direct protocol data, whereas symbionts Car and Prof possessed higher coverage and lower iRep estimates using the same protocol. When comparing samples with coverage >20×, iRep values for Car and Prof populations in the PDF protocol data (1.52 ± 0.03 and 1.53 ± 0.11, respectively) approached an asymptote comparable to iRep estimates of Car and Prof populations in the Direct protocol data (1.67 ± 0.17 and 1.36 ± 0.09, respectively), but were not identical (Fig. 3C and D). In contrast, *Wolbachia* populations achieved sufficient coverage (∼20×) for genome analysis in Direct protocol data, but their iRep estimates between the PDF and Direct protocol differed (1.33 ± 0.05 and 1.11 ± 0.02, respectively). To include samples from the Direct protocol, we chose to present unfiltered iRep values because coverage of some taxa, especially for the Direct protocol samples, failed to satisfy cutoff criteria required by the iRep tool to report an iRep estimate (Brown et al. 2016). We also calculated iRep values for the handful of publicly available raw sequencing data available for *C*Las strains sequenced from *Citrus sinensis* and *Citrus maxima* DNA extracts, which ranged from 1.32 to 1.64 (Supplementary Table S3). However, average genome coverage of these *C*Las strains (see Methods for details) ranged from only 5 to 18×, which could inflate iRep estimates compared with *C*Las strains herein whose iRep estimates plateaued with increasing average genome coverage (Fig. 3D; PDF protocol panels). Overall, we observed that sequence data generated by the PDF protocol provides ample sequence coverage required to estimate stable iRep values for *C*Las (and other *D. citri* symbionts), especially when *C*Las titers in DNA extracts are sufficiently high (*C*Las 16S rRNA gene copies $\geq 1 \times 10^5$) (Supplementary Fig. S1). All iRep estimates and associated metrics for symbionts in lab-reared *D. citri* and *C*Las genomes isolated from *Citrus* spp. can be found in Supplementary Table S3.

Genetic polymorphisms in microbial endosymbionts from lab-reared *D. citri*

In addition to genetic diversity and replication rate estimation, another aim was to investigate the relationships of and genetic variation between *C*Las strains sequenced with the PDF protocol and publicly available *C*Las strains. Phylogenetic analysis of putatively non-recombinogenic SNPs identified with Gubbins (Croucher et al. 2015) (see Methods for details) revealed low, yet detectable, genetic variation among the *C*Las genomes from our *D. citri* colonies (Fig. 4). These strains were observed to be descended from *C*Las strains (psy62 and JRPAMB1) derived from *D. citri* colonies maintained at USDA facilities in Florida (Duan et al. 2009) (unpublished strain JRPAMB1 location accessed with NCBI BioSample accession SAMN11842353; Fig. 4). In the 7 years since arriving in New York from Florida rearing facilities, our *C*Las strains accumulated as many as 156 (78 ± 56, mean ± s.d.) genetic polymorphisms compared to reference *C*Las strain psy62 from FL (Fig. 5). In contrast, *C*Las sequence data produced using the Direct protocol did not reveal detectable (minimum SNP coverage = 10) genetic polymorphisms in *C*Las, even in samples with high *C*Las titers (Fig. 2B, right panel). Of note, the PDF protocol enabled detection of genetic variants in two *C*Las genes predicted to be involved in antibiotic resistance by the ProGenomes database (Mende et al. 2017), a putative cysteine desulfurase and N-acetyltransferase (Fig. 5, purple text). *C*Las strains 20E and 22E possessed a missense mutation in the cysteine desulfurase *sufS* gene, at position 1,066,382, resulting in a base change of A to G and an associated amino acid sequence change of glutamine to arginine (amino acid 37 of *C*Las strain psy62 protein accession WP_015824969.1). We also detected a deletion mutation in the gene encoding a putative N-acetyltransferase (WP_102134467.1), in which a CA to C mutation at position 1,158,964 resulted in a frameshift mutation after amino acid position 36 in the encoded amino acid sequence.

Interestingly, the gene (CLIBASIA_05345) encoding the putative N-acetyltransferase is located within a recently predicted prophage region of the *C*Las genome (Dominguez-Mirazo et al. 2019) and could constitute an important mobile element transferable among *C*Las strains if the prophage becomes lytic or is excised from the *C*Las genome (Fleites et al. 2014).

In addition to potential antibiotic resistance, *C*Las encodes secreted protein effectors, which could enable *C*Las to establish or persist during infection of its citrus or *D. citri* hosts (Jain et al. 2015; Thapa et al. 2020). No variants were detected in loci encoding the secreted peroxidase SC2_gp095 (CLIBASIA_00485), the putative glutathione-dependent peroxidase SC2_gp100

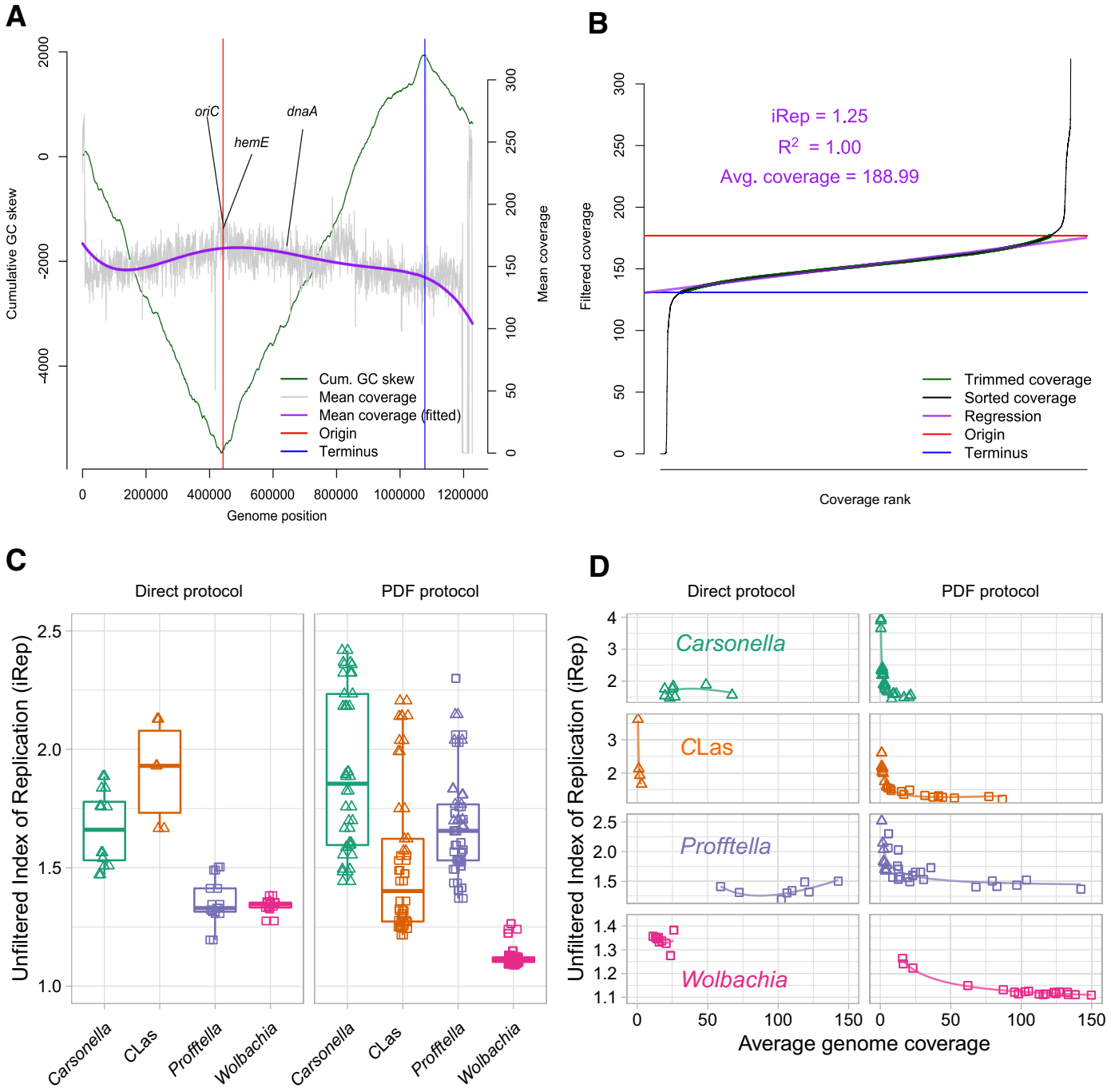

Fig. 3. Index of replication (iRep) estimates for the four *Diaphorina citri* symbionts detected in sequence data sets generated from the pretreatment DNase and filtration (PDF) and Direct protocols. **A,** The mean coverage (5 kb length windows, 10 nt sliding window) and GC skew across the '*Candidatus* Liberibacter asiaticus' (*C*Las) genome provide information about the replication origin (red line) and terminus (blue line). **B,** The iRep tool (Brown et al. 2016) filters and sorts the coverage by rank along the *C*Las genome to derive a metric consistent with an average instantaneous replication rate within a bacterial population. **C,** The iRep estimates were highest for '*Candidatus* Carsonella armatura' in both the PDF and Direct protocols followed by '*Candidatus* Profftella armatura', *C*Las, and *Wolbachia pipientis* strain wDi but **D,** were shown to be highly dependent upon genome coverage. Squares and triangles in C and D indicate whether iRep could robustly calculate the filtered index of replication or not, respectively. Lines in D are second order polynomial regressions fit by formula iRep ~ poly(log(coverage), 2) in R 4.1.2. Unfiltered index or replication values are shown to directly compare estimates from the PDF and Direct protocols. *oriC* = predicted origin or replication determined by the DoriC database (Luo and Gao 2019) for *C*Las strain psy62; *hemE* = gene encoding uroporphyrinogen decarboxylase involved in heme biosynthesis and often detected adjacent to *oriC* in members of the *Alphaproteobacteria* (Ioannidis et al. 2007); *dnaA* = encodes DnaA protein that binds to *dnaA* boxes to promote strand separation at the *oriC* region. All iRep data displayed in panels C and D can be found in Supplementary Table S3.

(CLIBASIA_05595), or the Sec-translocon dependent effector SDE1 (CLIBASIA_05315). However, we did detect an insertion mutation in CLIBASIA_04530, encoding a putative Sec-dependent effector, which was shown to be predominantly expressed by *C*Las in citrus tissues (Thapa et al. 2020) (Supplementary Table S4). Although the exact functional consequences of the genetic polymorphisms detected between *C*Las strains in our lab and ancestral *C*Las strains in FL are unknown, the PDF protocol enabled robust detection and documentation of these mutations between *C*Las strains.

As previously noted, *C*Las strains also harbor distinct lytic and non-lytic prophage-encoding regions that confer a variety of functions related to *C*Las pathogenicity and countering host immune responses (Dominguez-Mirazo et al. 2019; Fleites et al. 2014; Jain et al. 2015; Merfa et al. 2019). In the three predicted prophage elements, FP1, FP2, and a recently discovered Type IV prophage element (Dominguez-Mirazo et al. 2019), we detected a total of 69, 11, and one polymorphisms in predicted coding regions, respectively, and only five polymorphisms in noncoding elements of the FP1 prophage region and none in the FP2 and the Type IV prophage element (Fig. 5). This was in contrast to non-prophage portions of the *C*Las genome, where only 33 polymorphisms were detected in coding regions compared with 52 in non-coding elements. The prophage variants occurred in phage primases and terminases, and four of the *C*Las strains we sequenced possessed an 8-bp insertion in the bacteriophage repressor protein c1, resulting in a frameshift mutation that could alter repression of the prophage lytic state within the strains this mutation was observed (Supplementary Table S4) (Miller and Kokjohn 1987).

Aside from *C*Las, the Prof symbiont was observed to possess the next greatest number of genetic polymorphisms when compared with the latest Prof reference genome from https://citrusgreening.org/ when using either protocol (Fig. 5). The location of genetic polymorphisms detected in Prof using either protocol generally

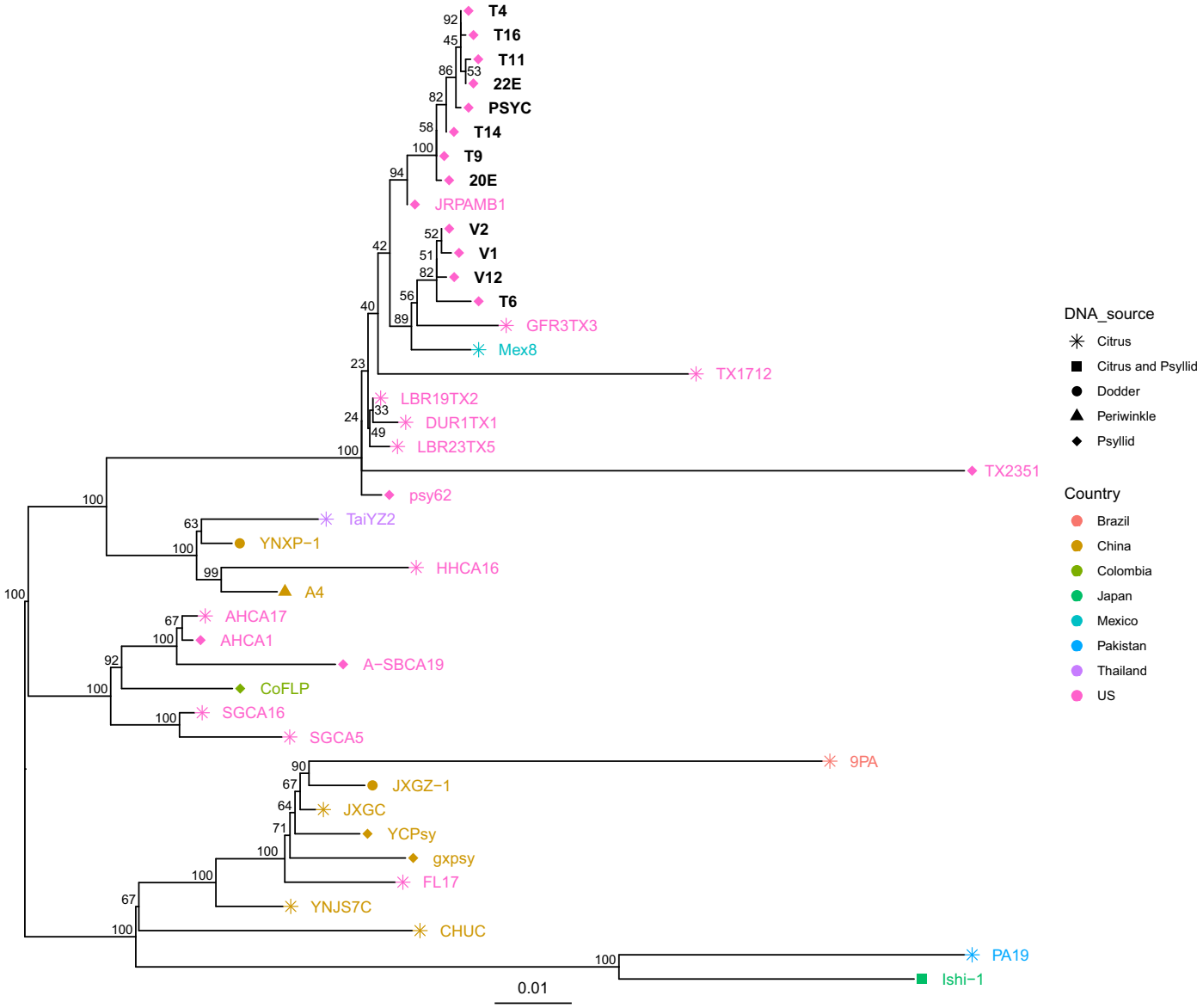

Fig. 4. A midpoint-rooted, maximum-likelihood, whole genome SNP phylogeny of the '*Candidatus* Liberibacter asiaticus' (*C*Las) strains derived from laboratory-reared *Diaphorina citri* adults using the pretreatment DNase and filtration (PDF) protocol (black tip labels) or *Citrus* spp. and *D. citri* tissues from public databases (colored tip labels). Shapes next to the tree tips indicate the DNA tissue source and colors of the shapes and tip labels represent the country of origin of the *C*Las genome (except black tip labels, which are U.S. samples from this study). Values at nodes represent ultrafast bootstrap support reported by IQ-TREE (Nguyen et al. 2015) estimated using 1,000 bootstrap replicates. The scale bar represents the number of nucleotide substitutions per site. Strains with genome assemblies possessing median contig sizes of <30 kb for more than half the contigs in the assembly were excluded, except for strain 9PA, which was included due to its unique geographic context in Brazil (Silva et al. 2021).

agreed, although coverage was usually higher with the Direct protocol in Car and Prof populations, and only a few unique polymorphisms were detected with each protocol (Fig. 5). Notably, a majority of the variants detected were frameshift or non-synonymous polymorphisms within genetic machinery involved in polyketide synthesis (Supplementary Table S4), which could indicate alteration of the defensive compounds (i.e., diaphorin) provided to *D. citri* by the Prof symbiont (Nakabachi et al. 2020) and could be attributed to continuous laboratory propagation of *D. citri*. Genetic variants detected within Car, the nutritional symbiont of *D. citri*

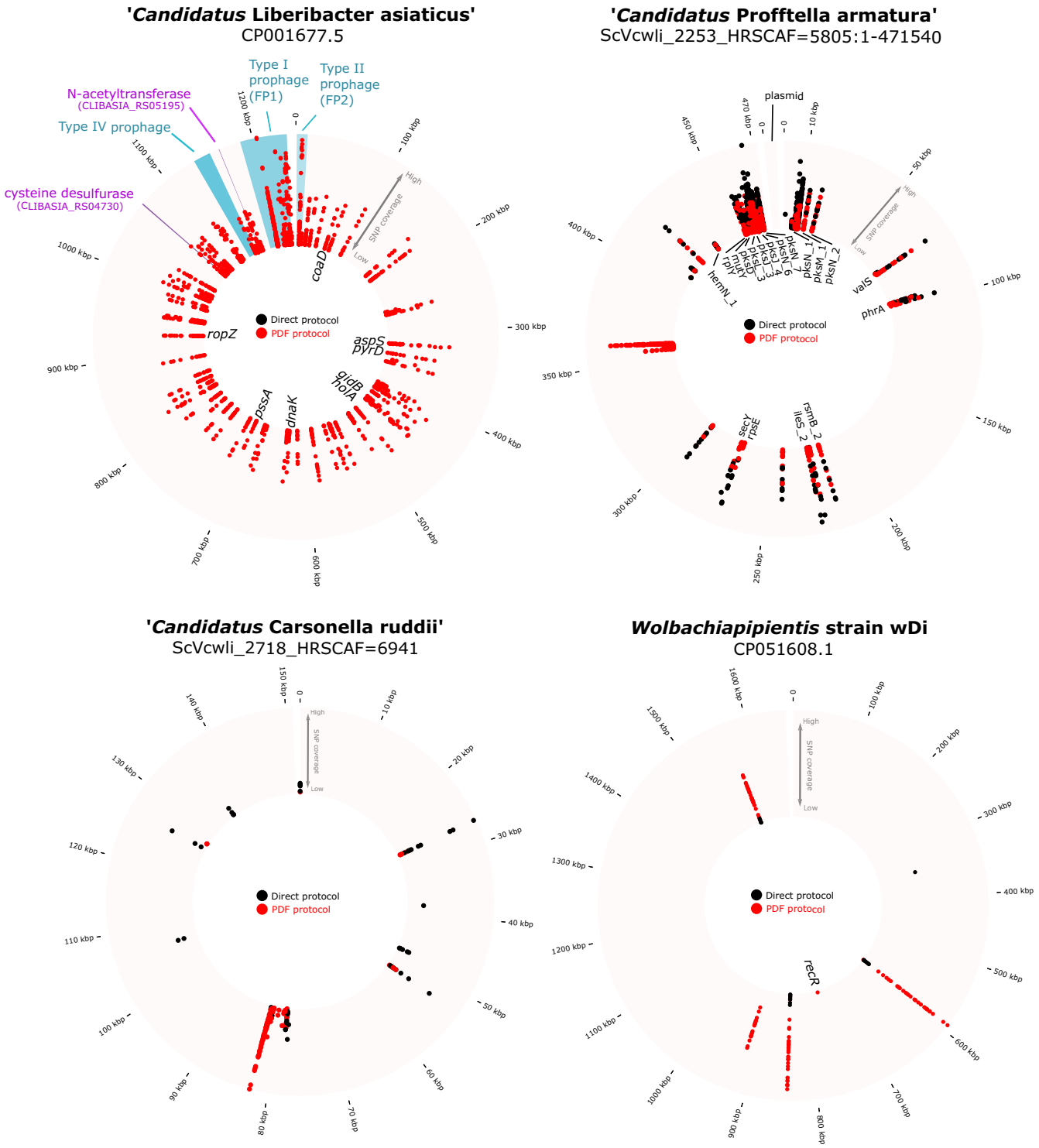

Fig. 5. Coverage profiles of genetic polymorphisms detected within genomes of bacterial pathogens and symbionts of *Diaphorina citri*, the Asian citrus psyllid. Genetic polymorphisms identified in sequence reads generated using the pretreatment DNase and filtration (PDF) protocol (red points) versus the Direct protocol (black points) are compared. Teal-colored regions in the top left panel are predicted prophage-encoding regions of the '*Candidatus* Liberibacter asiaticus' (*C*Las) genome (Dominguez-Mirazo et al. 2019). Regions and labels in purple in the top left are two genes predicted to encode antibiotic resistance genes and observed to possess genetic polymorphisms compared to *C*Las strain psy62. Four letter codes inside each ring are gene abbreviations associated with annotated loci in which polymorphisms were detected. If a four-letter abbreviation is absent, the polymorphism was either located within a gene lacking an annotation or was located in a noncoding region of the genome. All polymorphisms detected by snippy (Seeman 2015), including their four letter annotations, can be found in Supplementary Table S4.

(Nakabachi et al. 2020, 2006), which were localized to genes encoding hypothetical proteins, RNA polymerase subunit beta, DNA polymerase III subunit alpha, and the tRNA gene for phenylalanine (Supplementary Table S4). The genomes of *Wolbachia pipentis* strain wDi possessed the lowest number of polymorphisms (n = 6) compared with the *W. pipientis* strain wDi reference genome, with the highest coverage of polymorphisms detected in samples generated using the PDF protocol (Fig. 5). However, all variants detected were frameshift or missense mutations in key genes putatively involved cell wall remodeling, conjugation, and cell regulation (Supplementary Table S4). We did not detect genetic variants in *Wolbachia* loci encoding proteins with a demonstrated interaction with *C*Las (Jain et al. 2017). All genetic variants detected in *D. citri* symbionts can be accessed in Supplementary Table S4.

Sequence analysis of field-collected *D. citri* microbiomes

Next, we utilized the PDF protocol to sequence, assemble, and examine the microbiomes of field-collected *D. citri* adults (n = 6) collected from a citrus grove in Florida, U.S.A., in November 2018. Unfortunately, *C*Las titers in these samples were quite low (*C*Las 16S rRNA gene Cq values >30), and consequently, any assembled *C*Las sequences detected in these samples were short in length and possessed low coverage (mean coverage ± s.d. = 0.74 ± 1.18). However, we did observe large contigs assigned to the *D. citri* symbionts Car and Prof in the coassembly of the six field-collected *D. citri* adults (Fig. 6A). Rather unexpectedly, we also detected 10,830 contigs (>2,500 nt length) in the coassembly that were assigned to the *C. sinensis* genome, and, in particular, large, high-coverage contigs aligning to chloroplast and mitochondrial sequence elements of this same species (Fig. 6B). Unlike our laboratory-reared specimens, in which we observed a significant sequence signal associated with the *D. citri* genome, only six contigs (>2,500 nt length) were assigned to the *D. citri* genome in field-collected adults, the longest of which was an ∼15 kb contig affiliated with the *D. citri* mitochondrial genome (Fig. 6A). We also detected large numbers of low-coverage contigs assigned to commensal bacterial populations including members of the genera *Escherichia*, *Flavobacterium*, *Pseudomonas*, and *Xanthomonas*. Of note, only contigs assigned to the genus *Pseudomonas* achieved appreciably high coverage (∼20×) in one of the field-collected *D. citri* adults we sequenced (panel SPK8; Fig. 6A). A shotgun sequence analysis of field-collected *D. citri* in California also reported variable levels of sequences assigned to the genus *Pseudomonas* (Huang et al. 2021) and reinforces the utility of the PDF protocol for psyllid microbiome analysis in general. Finally, we observed high-coverage contigs, some of which were large (∼15 to 30 kb), that were assigned to fungal mitochondrial sequences from well-known entomopathogenic fungal genera *Entomophthora*, *Capillidium*, and *Neoconidiobolus* (Fig. 6A). These sequences achieved appreciable coverage in only two of the six field-collected *D. citri* adults we sequenced (panels SPK1 and SPK4; Fig. 6A), suggesting that the PDF protocol could enable identification of in situ populations of entomopathogenic fungi potentially infecting *D. citri* in citrus groves.

Discussion

Methods that enable or enhance sequencing of insect-associated microbiomes can inform solutions to longstanding challenges in diverse fields such as phytopathology, insect biology, and holobiont theory (Guyomar et al. 2018; Simon et al. 2019). For example, sequence-based analyses of microbiome composition and function can generate hypotheses related to the microbiome's influence on insect fitness or the number and types of ecosystem services an insect and its microbiome (i.e., the holobiont) can provide (Simon et al. 2019). Furthermore, sequenced-based investigations of genetic modules underlying microbially mediated pesticide resistance can facilitate new strategies to circumvent pesticide resistance and uncover novel resistance phenotypes of use to beneficial insects (Gressel 2018; Simon et al. 2019). However, the high quantities of host DNA detected within insect microbiomes can hamper detection sensitivity of both microbial diversity and genetic polymorphisms present therein (Pereira-Marques et al. 2019). Furthermore, many insect-associated microorganisms of great importance are obligate or facultative symbionts that are as yet uncultivable (Hosokawa et al. 2016; Moran et al. 2008), and therefore, cultivation-independent methods to predict functions of uncultivable taxa are vital research tools (Coyle et al. 2019; Guyomar et al. 2018). Thus, new approaches are needed to circumvent the limitations inherent to sequence-based analyses of insect microbiomes. The PDF protocol described herein enables cost-effective, efficient, and direct access to insect-associated microorganisms at orders of magnitude lower cost and total sequencing effort (Fig. 2A; Supplementary Table S1). Hence, the PDF protocol generates more microbiome data and facilitates investigations of microbe-microbe interactions or evolution at the molecular level.

Aside from the operational benefits afforded by the PDF protocol, it also enabled a comprehensive assessment of the microbiota present within and on *D. citri* adults. The PDF protocol not only enabled genetic analysis of both obligate (Car and Prof) and facultative (*C*Las and *Wolbachia*) symbionts, but it also captured a variety of commensal bacterial and eukaryotic sequences in the *D. citri* adults examined (Figs. 2 and 6). The eukaryotic sequences we detected in field-collected *D. citri* are particularly intriguing because they represent genetic signatures of *Citrus* organelles retained due to *D. citri*'s phloem-feeding lifestyle. Thus, the shotgun sequence data provided by the PDF protocol could contain genetic signatures consistent with an insect's diet, suggesting that the PDF method could complement or potentially supplant existing PCR-based approaches for gut content analysis in psyllids and other insects (Cooper et al. 2019). The latter is particularly promising because PCR amplicons have limited value in taxonomic assignment due to their short length, whereas citrus chloroplast and mitochondrial contigs detected in PDF protocol data are tens to hundreds of kilobases in length, which will greatly assist with taxonomic identification of citrus or other plant species fed on by *D. citri*. Of note, stark differences in the composition of genetic signatures detected between the laboratory-reared and field-collected *D. citri* were observed. For example, laboratory-reared *D. citri* possessed much higher quantities of the *D. citri* genome than field-collected *D. citri* (Fig. 2B; except for a high-coverage mitochondrial contig in field-collected *D. citri*, Fig. 6A), whereas the field-collected *D. citri* data mostly contained contigs assigned to the *Citrus sinensis* genome (Fig. 6B). The exact reasons for these differences are unknown, but they could be due to methodological differences in the capture and storage of laboratory-reared versus field-collected *D. citri*. For example, field-collected *D. citri* were preserved in 95% (vol/vol) ethanol when field collected and then subsequently frozen, whereas laboratory-reared *D. citri* were collected and frozen immediately. Thus, the cells of field-collected *D. citri* stored in ethanol prior to freezing could facilitate cell lysis, releasing more *D. citri* DNA that can be degraded during the DNase I treatment in the PDF protocol. The rigid wall of plant cells found in field-collected *D. citri* might have been less affected by ethanol storage, and their cells were probably intact during DNase I treatment. This could explain why genetic signatures of citrus were higher in field-collected rather than laboratory-reared *D. citri*, especially if host cell lysis was less effective without ethanol storage or a freeze-thaw step. Nonetheless, these storage differences need to be explored in future iterations of the PDF protocol and could enhance digestion of *D. citri* or another insect host's DNA by DNase treatment.

In addition to genetic signatures indicative of the *D. citri* diet, we also detected mitochondrial signatures assigned to putatively entomopathogenic fungi, which suggests that some members of the field-collected *D. citri* we studied were actively infected at the time of sampling (Fig. 6). Interestingly, the mitochondrial

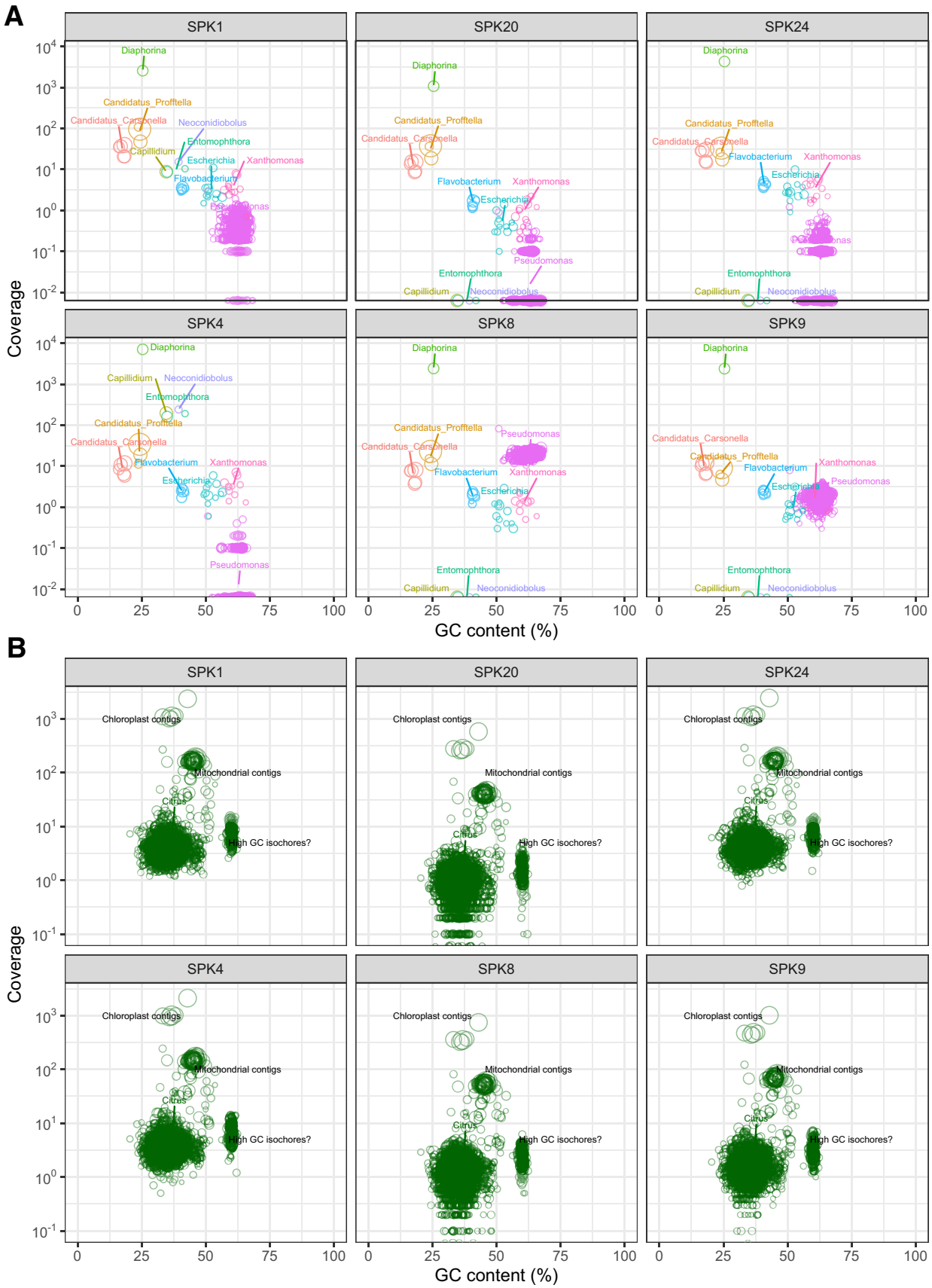

Fig. 6. Taxonomic composition, coverage, size, and percent GC content of assembled contigs (>2,500 nt) from six individual field-collected *Diaphorina citri* adults processed using the pretreatment DNase and filtration (PDF) protocol. The size and color of points in all panels represent the length and taxonomic affiliation of contigs generated from a coassembly of sequence reads from all individuals. The assembled contigs were taxonomically assigned to **A,** *D. citri* (mitochondrial contigs), entomopathogenic fungi (*Capillidium*, *Neoconidiobolus*, *Entomophthora*), and *D. citri* symbionts or commensal bacterial taxa or to **B,** the genomes of *Citrus sinensis* or *C. maxima*. Note the variation in coverage of contigs assigned to putative entomopathogenic fungi or the bacterial genus *Pseudomonas* in A and the consistently high coverage of citrus chloroplast and mitochondrial contigs in B. Taxa with the most abundant and numerous contigs with consistent taxonomic affiliations are shown for clarity. Taxonomic annotations of all contigs can be accessed by the figshare repository link located in the data availability section.

signatures assigned to putatively entomopathogenic fungi we detected are distinct from fungal genera previously reported to infect *D. citri* (Grafton-Cardwell et al. 2013) and suggest that additional, unexplored fungal taxa could be utilized as biocontrol agents for *D. citri* populations in the field. Fungi previously reported to infect *D. citri*, including members of the genera *Isaria*, *Hirsutellum*, *Lecanicillium*, and *Beauveria*, are all classified within the order Hypocreales, whereas the mitochondrial signatures we detected in the field-collected psyllids were all classified within the order Entomophthorales, members of which have been reported to infect *D. citri* in Mexico (Berlanga-Padilla et al. 2018; Guizar-Guzman and Sanchez-Peña 2013) but whose potential as biocontrol agents have not been explored. Despite the modest number of field-collected psyllids processed using the PDF protocol ($n = 6$), these abundant fungal signatures were also detectable, albeit at low coverage, within an additional 8 out of 17 (47%) *D. citri* adults that were collected from the same citrus grove but were processed and sequenced using the Direct protocol (see Data availability to access raw data for the $n = 17$ *D. citri* individuals reported). These findings suggest an epizootic infection was occurring in the *D. citri* populations sampled from the citrus grove in Florida at the time of collection. Therefore, the PDF protocol can facilitate the identification of native or potentially novel fungal pathogens of invasive insects, such as *D. citri*, and could assist the development of biocontrol strategies to prevent their spread (Clifton et al. 2019).

The sequence data afforded by our PDF protocol also enabled robust estimation of the replication rate of each symbiont present within the *D. citri* microbiome (Fig. 3). The same analysis is unlikely to be compatible with alternative bacterial enrichment protocols such as multiple displacement amplification or probe capture hybridization methods because the former results in biased genome amplification and the latter presents variable capture efficiencies and loss of relative abundance among sequence elements captured by the probes. However, probe hybridization is capable of enriching and sequencing *C*Las DNA from citrus tissues (Cai et al. 2019), which is currently unavailable with the PDF protocol described herein but is under active development. These methods are the most widely used to sequence *C*Las genomes from *D. citri* adults or citrus samples, but both require greater time and financial investment and do not provide a comprehensive analysis of the *D. citri* microbiome. Nevertheless, the PDF protocol also has limitations, including reduced sequence coverage of obligate symbionts (Car and Prof) within *D. citri* compared with direct extraction and sequencing of DNA from *D. citri* adults using the Direct protocol. Despite this limitation, coverage of Car and Prof populations within *D. citri* adults enabled robust iRep estimates on par with those calculated from the Direct protocol (Fig. 3D). Importantly, the PDF and Direct protocols are complementary approaches, and a small portion of the initial psyllid homogenate can be saved for DNA extraction via the Direct protocol and sequenced if additional coverage of the host or obligate bacterial endosymbionts is needed.

The estimation of *C*Las replication rates in *D. citri* adults is of interest to citrus greening disease research because actively replicating *C*Las populations are more likely to be spread by their *D. citri* vector (Ammar et al. 2018). However, whether *C*Las actively replicates or only accumulates within its insect vector via feeding has been a matter of debate (Ammar et al. 2016; Pelz-Stelinski et al. 2010). According to iRep values recorded in the present study, approximately 20 to 30% of the *C*Las population with >25× coverage in the *D. citri* adults examined were actively replicating their genomes (Fig. 3). These values were much lower than Car and Prof populations, which at equivalent coverage (25× or greater) suggest that 50 to 75% of these symbiont populations are actively replicating their genomes, consistent with the key nutritional and defensive roles of these bacteria within *D. citri* (Nakabachi et al. 2006, 2013, 2020). Considering the high prevalence of the *Wolbachia* symbiont within *D. citri* (Kruse et al. 2017; Ren et al. 2018), we were surprised by the low iRep values estimated for this symbiont. Our data suggest that only 10% of the *Wolbachia* populations examined were actively replicating (Fig. 3). It is possible that the unique mode of transovarial transmission by *Wolbachia*, competition with other symbionts, or inhibition by *D. citri*'s immune system might limit its replication within *D. citri* (Kruse et al. 2017; Ren et al. 2018), but the exact factors underpinning the limited replication detected within *Wolbachia* populations in the lab-reared *D. citri* we examined should be explored.

Despite its uncultivable nature, *C*Las is capable of propagation in laboratory settings using young citrus trees and laboratory-reared *D. citri* (Hilf and Lewis 2016; Hosseinzadeh et al. 2021; Skelley and Hoy 2004). Although this method of *C*Las cultivation has generated valuable insights related to therapeutics development and the response of *D. citri*'s immune system to *C*Las infection (Arp et al. 2016; Kruse et al. 2018; Mann et al. 2018), continual laboratory propagation of bacteria can lead to bottleneck effects and introduce genetic variation not represented in founding or wild bacterial populations (Elena and Lenski 2003; Marks et al. 2010). This fact is immediately apparent when comparing SNPs between reference *C*Las strain psy62 and the *C*Las strains sequenced in our laboratory (Figs. 4 and 5). The great number of SNPs detected between *C*Las strain psy62 (a FL *C*Las strain) and our *C*Las strains in New York are concerning for several reasons. Because our *C*Las strains derive from those maintained in Florida greenhouses and *D. citri*-rearing facilities, and the number of differences suggest that our *C*Las strains have diverged from Florida strains (average of ~22 mutations/year), which could result in inconsistent responses by therapeutics or other treatments being explored among facilities. To our knowledge, only a dozen *C*Las strains have been sequenced from commercial citrus groves or field-collected *D. citri* populations (Dai et al. 2019; Kunta et al. 2017; Liu et al. 2020; Thapa et al. 2020; Zheng et al. 2021). Rather, genome sequences of most *C*Las strains are derived from laboratory-reared *D. citri* infected with *C*Las or infected citrus scions grafted onto healthy citrus rootstock (Supplementary Table S1). This suggests that our understanding of the existing genomic diversity of *C*Las strains in commercial groves or urban settings is poor. Research conducted using institute-specific *C*Las strains may obfuscate or stymie investigations of *C*Las physiology or therapeutics development without dedicated methods to monitor and account for genetic variation among lab-propagated *C*Las strains. Our results and others (Pitino et al. 2014) also support the idea that future studies consider a role for *C*Las strain diversity and propensity for mutation in fundamental studies on plant infection and psyllid transmission (Ammar et al. 2018).

The biodiversity of insects is vastly underexplored, and we have only begun to scratch the surface in our understanding of the relationship between insects and their microbial partners (Gurung et al. 2019). Furthermore, few methods are available to explore the host-microbe interface at the molecular level in insects and these methods are desperately needed for basic and applied entomological studies of insect vectors with medical and agricultural importance (Chevrette et al. 2019; French et al. 2021). We hope the PDF protocol and any future iterations thereof fulfills these needs and can enable advances in insect biology and microbiology to the benefit of all.

Acknowledgments

S. A. Higgins is grateful to Ashley M. Frank for assistance with revising the manuscript. S. A. Higgins and M. Heck thank Hannah J. MacLeod and Stephanie Preising for their helpful scientific discussions regarding the development of the PDF protocol. All authors thank Stacy DeBlasio, David Igwe, Luke Thompson, and Julie Blaha for their tireless efforts maintaining and provisioning the citrus and *C*Las-infected *D. citri* adults utilized in this work.

Literature Cited

Ajene, I. J., Khamis, F., van Asch, B., Pietersen, G., Rasowo, B. A., Ekesi, S., and Mohammed, S. 2020. Habitat suitability and distribution potential of

Liberibacter species ('*Candidatus* Liberibacter asiaticus' and '*Candidatus* Liberibacter africanus') associated with citrus greening disease. Divers. Distrib. 26:575-588.

Ammar, E. D., Hall, D. G., Hosseinzadeh, S., and Heck, M. 2018. The quest for a non-vector psyllid: Natural variation in acquisition and transmission of the huanglongbing pathogen '*Candidatus* Liberibacter asiaticus' by Asian citrus psyllid isofemale lines. PLoS One 13:1-23.

Ammar, E. D., Ramos, J. E., Hall, D. G., Dawson, W. O., and Shatters, R. G. 2016. 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 11: 1-18.

Arp, A. P., Hunter, W. B., and Pelz-Stelinski, K. S. 2016. Annotation of the Asian citrus psyllid genome reveals a reduced innate immune system. Front. Physiol. 7:1-18.

Badial, A. B., Sherman, D., Stone, A., Gopakumar, A., Wilson, V., Schneider, W., and King, J. 2018. Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. Plant Dis. 102:1648-1652.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., MA, A., and Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455-477.

Baym, M., Kryazhimskiy, S., Lieberman, T. D., Chung, H., Desai, M. M., and Kishony, R. K. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One 10:1-15.

Berlanga-Padilla, A. M., Gallou, A., Ayala-Zermeño, M. A., Serna-Domínguez, M. G., Montesinos-Matías, R., Rodríguez-Rodríguez, J. C., and Arredondo-Bernal, H. C. 2018. Entomopathogenic fungi associated to *Diaphorina citri* (Hemiptera: Liviidae) in Colima. Mexico. Rev. Mex. Biodivers. 89:986-1001.

Blaustein, R. A., Lorca, G. L., and Teplitski, M. 2018. Challenges for managing '*Candidatus* Liberibacter spp.' (Huanglongbing disease pathogen): Current control measures and future directions. Phytopathology 108:424-435.

Brown, C. T., Olm, M. R., Thomas, B. C., and Banfield, J. F. 2016. Measurement of bacterial replication rates in microbial communities. Nat. Biotechnol. 34:1256-1263.

Burdon, J. J., and Thrall, P. H. 2008. Pathogen evolution across the agro-ecological interface: Implications for disease management. Evol. Appl. 1: 57-65.

Cai, W., Nunziata, S., Costanzo, S., Kumagai, L., Rascoe, J., and Stulberg, M. J. 2020. Genome resource for the huanglongbing causal agent '*Candidatus* Liberibacter asiaticus' strain AHCA17 from citrus root tissue in California, USA. Plant Dis. 104:627-629.

Cai, W., Nunziata, S., Rascoe, J., and Stulberg, M. J. 2019. SureSelect targeted enrichment, a new cost effective method for the whole genome sequencing of '*Candidatus* Liberibacter asiaticus'. Sci. Rep. 9:1-8.

Cai, W., Yan, Z., Rascoe, J., and Stulberg, M. J. 2018. Draft whole-genome sequence of '*Candidatus* Liberibacter asiaticus' strain TX1712 from citrus in Texas. Genome Announc. 6:1-2.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10:1-9.

Chen, X., and Cui, Y. 2019. BioCircos: Interactive circular visualization of genomic data using "htmlwidgets" and "BioCircos.js." https://cran.r-project.org/package=BioCircos

Chen, Y., Li, T., Zheng, Z., Xu, M., and Deng, X. 2019. Draft whole-genome sequence of a '*Candidatus* Liberibacter asiaticus' strain from Yunnan, China. Microbiol. Resour. Announc. 8:e01413-18.

Chevrette, M. G., Carlson, C. M., Ortega, H. E., Thomas, C., Ananiev, G. E., Barns, K. J., Book, A. J., Cagnazzo, J., Carlos, C., Flanigan, W., Grubbs, K. J., Horn, H. A., Hoffmann, F. M., Klassen, J. L., Knack, J. J., Lewin, G. R., McDonald, B. R., Muller, L., Melo, W. G. P., Pinto-Tomás, A. A., Schmitz, A., Wendt-Pienkowski, E., Wildman, S., Zhao, M., Zhang, F., Bugni, T. S., Andes, D. R., Pupo, M. T., and Currie, C. R. 2019. The antimicrobial potential of *Streptomyces* from insect microbiomes. Nat. Commun. 10:1-11.

Clifton, E. H., Castrillo, L. A., Gryganskyi, A., and Hajek, A. E. 2019. A pair of native fungal pathogens drives decline of a new invasive herbivore. Proc. Natl. Acad. Sci. U.S.A. 116:9178-9180.

Cooper, W. R., Horton, D. R., Wildung, M. R., Jensen, A. S., Thinakaran, J., Rendon, D., Nottingham, L. B., Beers, E. H., Wohleb, C. H., Hall, D. G., and Stelinski, L. L. 2019. Host and non-host "whistle stops" for psyllids: Molecular gut content analysis by high-throughput sequencing reveals landscape-level movements of Psylloidea (Hemiptera). Environ. Entomol. 48:554-566.

Coyle, J. F., Lorca, G. L., and Gonzalez, C. F. 2019. Understanding the physiology of '*Candidatus* Liberibacter asiaticus': An overview of the demonstrated molecular mechanisms. J. Mol. Microbiol. Biotechnol. 28:116-127.

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., Parkhill, J., and Harris, S R. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 43:e15.

Dai, Z., Wu, F., Zheng, Z., Yokomi, R., Kumagai, L., Cai, W., Rascoe, J., Polek, M., Chen, J., and Deng, X. 2019. Prophage diversity of '*Candidatus* Liberibacter asiaticus' strains in California. Phytopathology 109: 551-559.

Dala-Paula, B. M., Plotto, A., Bai, J., Manthey, J. A., Baldwin, E. A., Ferrarezi, R. S., and Gloria, M. B. A. 2019. Effect of huanglongbing or greening disease on orange juice quality, a review. Front. Plant Sci. 9:1-19.

Dominguez-Mirazo, M., Jin, R., and Weitz, J. S. 2019. Functional and comparative genomic analysis of integrated prophage-like sequences in '*Candidatus* Liberibacter asiaticus'. mSphere 4:1-13.

Duan, Y., Zhou, L., Hall, D. G., Li, W., Doddapaneni, H., Lin, H., Liu, L., Vahling, C. M., Gabriel, D. W., Williams, K. P., Dickerman, A., Sun, Y., and Gottwald, T. 2009. Complete genome sequence of citrus huanglongbing bacterium, '*Candidatus* Liberibacter asiaticus' obtained through metagenomics. Mol. Plant-Microbe Interact. 22:1011-1020.

Elena, S. F., and Lenski, R. E. 2003. Microbial genetics: Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. Nat. Rev. Genet. 4:457.

Fleites, L. A., Jain, M., Zhang, S., and Gabriel, D. W. 2014. '*Candidatus* Liberibacter asiaticus' prophage late genes may limit host range and culturability. Appl. Environ. Microbiol. 80:6023-6030.

French, E., Kaplan, I., Iyer-Pascuzzi, A., Nakatsu, C. H., and Enders, L. 2021. Emerging strategies for precision microbiome management in diverse agroecosystems. Nat. Plants 7:256-267.

Ghosh, D. K., Motghare, M., and Gowda, S. 2018. Citrus greening: Overview of the most severe disease of citrus. Adv. Agric. Res. Technol. J. 2:83-100.

Grafton-Cardwell, E. E., Stelinski, L. L., and Stansly, P. A. 2013. Biology and management of Asian citrus psyllid, vector of the huanglongbing pathogens. Annu. Rev. Entomol. 58:413-432.

Gressel, J. 2018. Microbiome facilitated pest resistance: Potential problems and uses. Pest Manag. Sci. 74:511-515.

Guizar-Guzman, L., and Sanchez-Peña, S. R. 2013. Infection by *Entomophthora sensu stricto* (Entomophthoromycota: Entomophthorales) in *Diaphorina citri* (Hemiptera: Liviidae) in Veracruz. Mexico. Florida Entomol. 96: 624-627.

Gurung, K., Wertheim, B., and Salles, F. J. 2019. The microbiome of pest insects: It is not just bacteria. Entomol. Exp. Appl. 167:156-170.

Guyomar, C., Legeai, F., Jousselin, E., Mougel, C., Lemaitre, C., and Simon, J. C. 2018. Multi-scale characterization of symbiont diversity in the pea aphid complex through metagenomic approaches. Microbiome 6:1-21.

Haapalainen, M. 2014. Biology and epidemics of *Candidatus* Liberibacter species, psyllid-transmitted plant-pathogenic bacteria. Ann. Appl. Biol. 165:172-198.

Hilf, M. E., and Lewis, R. S. 2016. Transmission and propagation of '*Candidatus* Liberibacter asiaticus' by grafting with individual citrus leaves. Phytopathology 106:452-458.

Hosokawa, T., Ishii, Y., Nikoh, N., Fujie, M., Satoh, N., and Fukatsu, T. 2016. Obligate bacterial mutualists evolving from environmental bacteria in natural insect populations. Nat. Microbiol. 1:1-7.

Hosseinzadeh, S., Higgins, S. A., Ramsey, J., Howe, K., Griggs, M., Castrillo, L., and Heck, M. 2021. Proteomic polyphenism in color morphotypes of *Diaphorina citri*, insect vector of citrus greening disease. J. Proteome Res. 20:2851-2866.

Huang, J., Dai, Z., Zheng, Z., da Silvia, P. A., Kumagai, L., Xiang, Q., Chen, J., and Deng, X. 2021. Bacteriomic analyses of Asian citrus psyllid and citrus samples infected with '*Candidatus* Liberibacter asiaticus' in southern California and huanglongbing management implications. Front. Microbiol. 12:1-11.

Igwe, D. O., Higgins, S. A., and Cilia, M. 2021. An excised leaf assay to measure acquisition of '*Candidatus* Liberibacter asiaticus' by psyllids associated with citrus Huanglongbing disease. Phytopathology 112:69-75.

Inoue, H., Ohnishi, J., Ito, T., Tomimura, K., Miyata, S., Iwanami, T., and Ashihara, W. 2009. Enhanced proliferation and efficient transmission of '*Candidatus* Liberibacter asiaticus' by adult *Diaphorina citri* after acquisition feeding in the nymphal stage. Ann. Appl. Biol. 155:29-36.

Ioannidis, P., Hotopp, J. C. D., Sapountzis, P., Siozios, S., Tsiamis, G., Bordenstein, S. R., Baldo, L., Werren, J. H., and Bourtzis, K. 2007. New criteria for selecting the origin of DNA replication in *Wolbachia* and closely related bacteria. BMC Genomics 8:1-15.

Jain, M., Fleites, L. A., and Gabriel, D. W. 2015. Prophage-encoded peroxidase in '*Candidatus* Liberibacter asiaticus' is a secreted effector that suppresses plant defenses. Mol. Plant-Microbe Interact. 28:1330-1337.

Jain, M., Fleites, L. A., and Gabriel, D. W. 2017. A small *Wolbachia* protein directly represses phage lytic cycle genes in '*Candidatus* Liberibacter asiaticus' within psyllids. mSphere 2:1-12.

Johnson, E. G., Wu, J., Bright, D. B., and Graham, J. H. 2014. Association of '*Candidatus* Liberibacter asiaticus' root infection, but not phloem plugging

with root loss on huanglongbing-affected trees prior to appearance of foliar symptoms. Plant Pathol. 63:290-298.
Kahlke, T., and Ralph, P. J. 2019. BASTA – Taxonomic classification of sequences and sequence bins using last common ancestor estimations. Methods Ecol. Evol. 10:100-103.
Katoh, H., Miyata, S. I., Inoue, H., and Iwanami, T. 2014. Unique features of a Japanese '*Candidatus* Liberibacter asiaticus' strain revealed by whole genome sequencing. PLoS One 9:1-11.
Kielbasa, S. M., Wan, R., Sato, K., Horton, P., and Frith, M. C. 2011. Adaptive seeds tame genomic sequence comparison. Genome Res. 21:487-493.
Kondo, T., Guillermo, G. F., Tauber, C., Sarmiento, Y. C. G., Mondragon, A. F. V., and Forero, D. 2015. A checklist of natural enemies of *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) in the department of Valle del Cauca, Colombia and the world. Insecta Mundi. 0457:1-14.
Kruse, A., Fattah-Hosseini, S., Saha, S., Johnson, R., Warwick, E. R., Sturgeon, K., Mueller, L., MacCoss, M. J., Shatters, R. G. Jr, and Heck, M. C. 2017. Combining 'omics and microscopy to visualize interactions between the Asian citrus psyllid vector and the Huanglongbing pathogen '*Candidatus* Liberibacter asiaticus' in the insect gut. PLoS One 12:1-28.
Kruse, A., Ramsey, J. S., Johnson, R., Hall, D. G., MacCoss, M. J., and Heck, M. 2018. '*Candidatus* Liberibacter asiaticus' minimally alters expression of immunity and metabolism proteins in hemolymph of *Diaphorina citri*, the insect vector of huanglongbing. J. Proteome Res. 17:2995-3011.
Kunta, M., Zheng, Z., Wu, F., da Graca, J. V, Park, J.-W., Deng, X., and Chen, J. 2017. Draft whole-genome sequence of '*Candidatus* Liberibacter asiaticus' strain TX2351 isolated from Asian citrus psyllids in Texas, USA. Genome Announc. 5:e00170-17.
Laetsch, D. R., Blaxter, M. L., and Leggett, R. M. 2020. BlobTools: Interrogation of genome assemblies [version 1; peer review: 2 approved with reservations]. F1000Research 6:1-18.
Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9:357-359.
Lefoulon, E., Vaisman, N., Frydman, H. M., Sun, L., Foster, J. M., and Slatko, B. E. 2019. Large enriched fragment targeted sequencing (LEFT-SEQ) applied to capture of *Wolbachia* genomes. Sci. Rep. 9:1-10.
Leinonen, R., Sugawara, H., and Shumway, M. 2010. The sequence read archive. Nucleic Acids Res. 39:1-3.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and samtools. Bioinformatics 25:2078-2079.
Li, S., Wu, F., Duan, Y., Singerman, A., and Guan, Z. 2020. Citrus greening: Management strategies and their economic impact. HortScience 55: 604-612.
Li, T., Thaochan, N., Huang, J., Chen, J., Deng, X., and Zheng, Z. 2020. Genome sequence resource of '*Candidatus* Liberibacter asiaticus' from Thailand. Plant Dis. 104:624-626.
Li, T., Zhang, L., Deng, Y., Deng, X., and Zheng, Z. 2021. Establishment of a *Cuscuta campestris*-mediated enrichment system for genomic and transcriptomic analyses of '*Candidatus* Liberibacter asiaticus'. Microbiol. Biotechnol. 14:737-751.
Lin, H., Lou, B., Glynn, J. M., Doddapaneni, H., Civerolo, E. L., Chen, C., Duan, Y., Zhou, L., and Vahling, C. M. 2011. The complete genome sequence of '*Candidatus* Liberibacter solanacearum', the bacterium associated with potato Zebra Chip disease. PLoS One 6:1-13.
Liu, K., Atta, S., Cui, X., Zeng, C., Chen, J., Zhou, C., and Wang, X. 2020. Genome sequence resources of two '*Candidatus* Liberibacter asiaticus' strains from Pakistan. Plant Dis. 104:2048-2050.
Lo, C., and Chain, P. S. G. 2014. Rapid evaluation and quality control of next generation sequencing data with FaQCs. BMC Bioinformatics 15:1-8.
Luo, H., and Gao, F. 2019. DoriC 10.0: An updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. Nucleic Acids Res. 47:D74-D77.
Mann, M., Fattah-Hosseini, S., Ammar, E.-D., Stange, R., Warrick, E., Sturgeon, K., Shatters, R., and Heck, M. 2018. *Diaphorina citri* nymphs are resistant to morphological changes induced by '*Candidatus* Liberibacter asiaticus' in midgut epithelial cells. Infect. Immun. 86:e00889-17.
Marine, R., McCarren, C., Vorrasane, V., Nasko, D., Crowgey, E., Polson, S. W., and Wommack, K. E. 2014. Caught in the middle with multiple displacement amplification: The myth of pooling for avoiding multiple displacement amplification bias in a metagenome. Microbiome 2:1-8.
Marks, M. E., Castro-Rojas, C. M., Teiling, C., Du, L., Kapatral, V., Walunas, T. L., and Crosson, S. 2010. The genetic basis of laboratory adaptation in *Caulobacter crescentus*. J. Bacteriol. 192:3678-3688.
Marotz, C. A., Sanders, J. G., Zuniga, C., Zaramela, L. S., Knight, R., and Zengler, K. 2018. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome 6:1-9.
Mende, D. R., Letunic, I., Huerta-Cepas, J., Li, S. S., Forslund, K., Sunagawa, S., and Bork, P. 2017. ProGenomes: A resource for consistent functional and taxonomic annotations of prokaryotic genomes. Nucleic Acids Res. 45:D529-D534.
Merfa, M. V., Pérez-López, E., Naranjo, E., Jain, M., Gabriel, D. W., and De La Fuente, L. 2019. Progress and obstacles in culturing '*Candidatus* Liberibacter asiaticus', the bacterium associated with huanglongbing. Phytopathology 109:1092-1101.
Mikheenko, A., Saveliev, V., and Gurevich, A. 2016. MetaQUAST: Evaluation of metagenome assemblies. Bioinformatics 32:1088-1090.
Miller, R. V., and Kokjohn, T. A. 1987. Cloning and characterization of the c1 repressor of *Pseudomonas aeruginosa* bacteriophage D3: A functional analog of phage lambda cI protein. J. Bacteriol. 169:1847-1852.
Moran, N. A., McCutcheon, J. P., and Nakabachi, A. 2008. Genomics and evolution of heritable bacterial symbionts. Annu. Rev. Genet. 42:165-190.
Nakabachi, A., Piel, J., Malenovsky, I., and Hirose, Y. 2020. Comparative genomics underlines multiple roles of *Profftella*, an obligate symbiont of psyllids: Providing toxins, vitamins, and carotenoids. Genome Biol. Evol. 12:1975-1987.
Nakabachi, A., Ueoka, R., Oshima, K., Teta, R., Mangoni, A., Gurgui, M., Oldham, N. J., van Echten-Deckert, G., Okamura, K., Yamamoto, K., Inoue, H., Ohkuma, M., Hongoh, Y., Miyagishima, S. Y., Hattori, M., Piel, J., and Fukatsu, T. 2013. Defensive bacteriome symbiont with a drastically reduced genome. Curr. Biol. 23:1478-1484.
Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., and Hattori, M. 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267.
Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268-274.
Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25:1043-1055.
Pelz-Stelinski, K. S., Brlansky, R. H., Ebert, T. A., and Rogers, M. E. 2010. Transmission parameters for '*Candidatus* Liberibacter asiaticus' by Asian citrus psyllid (Hemiptera: Psyllidae). J. Econ. Entomol. 103: 1531-1541.
Pereira-Marques, J., Hout, A., Ferreira, R. M., Weber, M., Pinto-Ribeiro, I., Van Doorn, L. J., Knetsch, C. W., and Figueiredo, C. 2019. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. Front. Microbiol. 10:1-9.
Pitino, M., Hoffman, M. T., Zhou, L., Hall, D. G., Stocks, I. C., and Duan, Y. 2014. The phloem-sap feeding mealybug (*Ferrisia virgata*) carries '*Candidatus* Liberibacter asiaticus' populations that do not cause disease in host plants. PLoS One 9:1-7.
R Core Team. 2018. R: A language and environment for statistical computing. https://www.r-project.org/
Ren, S. L., Li, Y. H., Ou, D., Guo, Y. J., Qureshi, J. A., Stansly, P. A., and Qiu, B L. 2018. Localization and dynamics of *Wolbachia* infection in Asian citrus psyllid *Diaphorina citri*, the insect vector of the causal pathogens of huanglongbing. Microbiol. Open. 7:1-11.
Saha, S., Hunter, W. B., Reese, J., Morgan, J. K., Marutani-Hert, M., Huang, H., and Lindeberg, M. 2012. Survey of endosymbionts in the *Diaphorina citri* metagenome and assembly of a *Wolbachia* wDi draft genome. PLoS One 7:1-12.
Seeman, T. 2015. snippy: Fast bacterial variant calling from NGS reads. https://github.com/tseemann/snippy
Silva, P. A., Huang, J., Wulff, N. A., Zheng, Z., Krugner, R., and Chen, J. 2021. Genome sequence resource of '*Candidatus* Liberibacter asiaticus' strain 9PA from Brazil. Plant Dis. 105:199-201.
Simon, J. C., Marchesi, J. R., Mougel, C., and Selosse, M. A. 2019. Host-microbiota interactions: From holobiont theory to analysis. Microbiome 7:1-5.
Skelley, L. H., and Hoy, M. A. 2004. A synchronous rearing method for the Asian citrus psyllid and its parasitoids in quarantine. Biol. Control. 29:14-23.
Sundin, G. W., and Wang, N. 2018. Antibiotic resistance in plant-pathogenic bacteria. Annu. Rev. Phytopathol. 56:161-180.
Thapa, S. P., De Francesco, A., Trinh, J., Gurung, F. B., Pang, Z., Vidalakis, G., Wang, N., Ancona, V., Ma, W., and Coaker, G. 2020. Genome-wide analyses of *Liberibacter* species provides insights into evolution, phylogenetic relationships, and virulence factors. Mol. Plant Pathol. 21: 716-731.
Tyler, H. L., Roesch, L. F. W., Gowda, S., Dawson, W. O., and Triplett, E. W. 2009. Confirmation of the sequence of '*Candidatus* Liberibacter asiaticus' and assessment of microbial diversity in huanglongbing-infected citrus phloem using a metagenomic approach. Mol. Plant-Microbe Interact. 22:1624-1634.
Ukuda-Hosokawa, R., Sadoyama, Y., Kishaba, M., Kuriwada, T., Anbutsu, H., and Fukatsuf, T. 2015. Infection density dynamics of the citrus greening bacterium '*Candidatus* Liberibacter asiaticus' in field populations of the psyllid

Diaphorina citri and its relevance to the efficiency of pathogen transmission to citrus plants. Appl. Environ. Microbiol. 81:3728-3736.
Wang, N. 2019. The citrus Huanglongbing crisis and potential solutions. Mol. Plant. 12:607-609.
Wang, N., and Trivedi, P. 2013. Citrus huanglongbing: A newly relevant disease presents unprecedented challenges. Phytopathology 103:652-665.
Wu, F., Zheng, Z., Deng, X., Cen, Y., Liang, G., and Chen, J. 2015. Draft genome sequence of '*Candidatus* Liberibacter asiaticus' from *Diaphorina citri* in Guangdong, China. Genome Announc. 3:1-2.
Yu, G., Smith, D., Zhu, H., Guan, Y., and Lam, T. T.-Y. 2017. ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8: 28-36.
Zheng, Y., Guo, J., Deng, X., and Zheng, Z. 2021. Genome sequence resource of '*Candidatus* Liberibacter asiaticus' strain myan16 from Myanmar. Plant Dis. 105:1162-1164.
Zheng, Z., Deng, X., and Chen, J. 2014. Whole-genome sequence of '*Candidatus* Liberibacter asiaticus' from Guangdong, China. Genome Announc. 2:1-2.