

RESEARCH ARTICLE

Deep Characterization of the Microbiomes of *Calophya* spp. (Hemiptera: Calophyidae) Gall-Inducing Psyllids Reveals the Absence of Plant Pathogenic Bacteria and Three Dominant Endosymbionts

Will A. Overholt¹✉, Rodrigo Diaz²✉, Erin Roskopf³✉, Stefan J. Green⁴✉, William A. Overholt⁵✉*

1 School of Biology, Georgia Institute of Technology, Atlanta, Georgia, United States of America, **2** Department of Entomology, Louisiana State University, Baton Rouge, Louisiana, United States of America, **3** United States Horticultural Research Laboratory, United States Department of Agriculture, Fort Pierce, Florida, United States of America, **4** Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois, United States of America, **5** Biological Control Research and Containment Laboratory, University of Florida, Fort Pierce, Florida, United States of America

✉ These authors contributed equally to this work.

* billlover@ufl.edu



CrossMark
click for updates

OPEN ACCESS

Citation: Overholt WA, Diaz R, Roskopf E, Green SJ, Overholt WA (2015) Deep Characterization of the Microbiomes of *Calophya* spp. (Hemiptera: Calophyidae) Gall-Inducing Psyllids Reveals the Absence of Plant Pathogenic Bacteria and Three Dominant Endosymbionts. PLoS ONE 10(7): e0132248. doi:10.1371/journal.pone.0132248

Editor: Matthias Horn, University of Vienna, AUSTRIA

Received: January 21, 2015

Accepted: June 11, 2015

Published: July 10, 2015

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: Raw sequence data (FASTQ files) were deposited in the NCBI Sequence Read Archive under project SRP049390. Other relevant data are within the paper and its Supporting Information files

Funding: This research was funded in part by grants from the Florida Fish and Wildlife Conservation Commission, the Florida Department of Agriculture and Consumer Services, and a National Science Foundation Graduate Research Fellowship, <https://www.nsfgrfp.org/>, (to WAO1) under Grant No.

Abstract

Bacteria associated with sap-feeding insect herbivores include not only symbionts that may increase their hosts' fitness but also harmful plant pathogens. *Calophya* spp. gall-inducing psyllids (Hemiptera: Calophyidae) are being investigated for their potential as biological control agents of the noxious weed, Brazilian peppertree (*Schinus terebinthifolia*), in Florida. Although there are no examples of plant pathogen transmission by members of the family Calophyidae, several insects in the superfamily Psylloidea are known to transmit pathogenic bacteria in the genera *Candidatus Liberibacter* and *Candidatus Phytoplasma*. To determine whether *Calophya* spp. harbor potentially harmful plant pathogenic bacteria, we sequenced small subunit (SSU) ribosomal RNA (rRNA) gene amplicons generated from individuals from four *Calophya* spp. populations. All microbial SSU gene sequences fell into the bacterial domain, with 98-99% belonging to the Proteobacteria. The *Calophya* microbiomes contained a relatively simple community, with 49-79 operational taxonomic units (OTUs; 97%) detected, and only 5-8 OTUs with greater than 1% abundance. *Candidatus Carsonella* showed the highest relative abundance, with OTUs from this candidate genus representing between 51 – 65% of all recovered sequences. The next most abundant clade observed was an unclassified Enterobacteriaceae group closely related to bacteria from the genera *Buchnera* and *Blochmannia* that ranged from 20-31% in relative abundance. *Wolbachia* populations were the third most abundant group and represented 7-27% of the diversity in microbial OTUs. No SSU rRNA gene sequences from putative pathogenic bacteria from the genera *Ca. Liberibacter* or *Ca. Phytoplasma* were detected in the microbiomes of the four *Calophya* populations. The probability that infected psyllids were present in our

2013172310. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

colonies, but were not sampled, was extremely low (1.39×10^{-10}). As far as we are aware, our study is the first to characterize the microbiome of a candidate biological control agent, and coupled with previous work demonstrating a high degree of host specificity and absence of plant viruses, suggests that releasing *Calophya* spp. in United States poses minimal risk to non-target plants.

Introduction

After habitat destruction, competition with invasive species is considered the greatest threat to imperiled native flora and fauna in the USA [1]. Within the continental USA, Florida is most vulnerable to invasion due to its island-like biogeography, high level of disturbance, and the incessant introduction of exotic species [2]. Of the 4,304 plants species that reproduce outside of cultivation in Florida, 1,440 (33%) are exotic [3] and 157 of these are considered invasive by the Florida Exotic Pest Plant Council [4]. Among the invasive plants, Brazilian peppertree (*Schinus terebinthifolia*) stands out as one of the most problematic, having invaded an estimated 283,000 ha of the state [5] where it severely reduces native community diversity [6]. Herbicides and mechanical control are the standard methods for management of Brazilian peppertree [7], but these tactics are costly [5], can harm non-target vegetation [8], and repeated treatments are required to prevent regrowth [5].

Classical biological control of Brazilian peppertree has been pursued for several years, but no natural enemies have yet been approved for release in the continental United States [9]. Recent foreign exploration in the coastal regions of Brazil has led to the discovery of leaf galling psyllids in the genus *Calophya* (Hemiptera: Calophyidae) attacking Brazilian peppertree [10,11]. Based on field collections, the *Calophya* spp. found on Brazilian peppertree have narrow host associations and in some cases appear to be monophagous [12,13]. *Calophya* spp. collected from Brazilian peppertree at four locations in Brazil were imported into Florida quarantine for further examination, including two populations of *Calophya latiforceps* Burckhardt from Bahia and Espirito Santo, one undescribed *Calophya* sp. from Espirito Santo, and *Calophya terebinthifolii* Burckhardt and Basset from Santa Catarina. *Calophya latiforceps* completes one generation in 39 d with 40% survival. Gall initiation and growth resulted in yellowing, deformation and abscission of Brazilian peppertree leaves [14]. Host specificity tests demonstrated that *C. latiforceps* nymphs can induce galls and complete development only on Brazilian peppertree and those adults exposed to non-target plants only had greatly reduced survival [15].

In spite of the promising biological control potential of *Calophya* spp., many species in the superfamily Psylloidea are known vectors of plant pathogenic bacteria including Gram-negative *Candidatus Liberibacter* [16] and Gram-positive *Candidatus Phytoplasma* [17]. The transmission of members of *Ca. Liberibacter* by the Asian citrus psyllid (*Diaphorina citri* Kumayama, Psyllidae), carrot psyllid (*Triozia apicalis* Förster, Triozidae), and potato psyllid (*Bactericera cockerelli* Sulc., Triozidae) has resulted in significant economic losses in citrus [18], carrots, and potatoes [19,20], respectively. Species of *Ca. Phytoplasma* cause more than 700 diseases and have been reported in the Psyllidae genera *Cacopsylla* and *Bactericera* [17,21]. Insect vectored pathogenic bacteria may also have positive effects. *Ca. Liberibacter europaeus*, vectored by the psyllid, *Arytainilla spartiophila* (Förster) (Psyllidae), may help suppress populations of the noxious weed, Scotch broom (Fabales: Fabaceae: *Cytisus scoparius*) in Europe and New Zealand [22].

The microbiomes of agricultural pest psyllids have been well characterized, initially with Sanger sequencing of cloned amplicons of the small subunit (SSU) rRNA gene, and more recently using next generation sequencing approaches [23–27]. The microbiomes of phloem feeding insects are typically characterized by a very low diversity of <10 bacterial operational taxonomic units, and for psyllids, are composed of the primary endosymbiont *Ca. Carsonella* and several secondary endosymbionts [25]. The vast majority of these studies focused on the microbiome of the potato psyllid (*Bactericera cockerelli*) and the Asian citrus psyllid (*Diaphorina citri*) [23,24,27]. Only one study has investigated bacteria associated with *Calophya* (*Calophya schini*), and this study only generated two SSU rRNA gene sequences from *Calophya schini*, one for *Ca. Carsonella* and a secondary endosymbiont associated with the *Enterobacteriaceae* [28,29].

Previously, we were unable to detect bacteria in *Calophya latiforceps* from the species *Ca. Liberibacter solanacearum*, *Ca. L. asiaticus*, *Ca. L. americanus* or *Ca. L. africanus* in amplification reactions employing species-specific primer sets [30]. We also found no evidence of plant viruses in *C. latiforceps* using both molecular and inoculation techniques [30]. Although targeted polymerase chain reaction (PCR) amplification reactions are highly sensitive to low levels of target deoxyribonucleic acid (DNA), the specificity of the assays can miss closely related taxa of interest. Furthermore, extensive evidence of associations between psyllids and plant pathogenic bacteria prompted us to employ a broader approach to detect potentially pathogenic bacteria. Therefore, the goal of this study was to characterize the entire microbial community present in several *Calophya* species. To achieve this, we performed ultra-deep high-throughput sequencing of SSU rRNA gene amplicons generated from genomic DNA (gDNA) extracted from *Calophya*. These results are useful in determining not only the potential safety of these candidate biological control agents but also provide new insight into the diversity of bacterial symbionts associated with psyllids. As far as we are aware, this is the first example of complete characterization of the microbiome of a candidate weed biological control agent.

Materials and Methods

Insect collection

Adults of *Calophya* spp. were obtained from colonies maintained at the biological control quarantine facility of the University of Florida in Fort Pierce, Florida. These colonies were initiated using adults collected from galled leaves at locations along coastal Brazil from 2012 to 2014 under permit no. 12BR008156/DF issued to W. A. Overholt from the Brazilian Institute of Environment and Renewable Natural Resources. The colonies included two populations of *Calophya latiforceps*, one population of *Calophya terebinthifolii* and one population of an undescribed *Calophya* species.

The designations used for the four *Calophya* populations used in this study, along with their origins are included in Table 1. Upon arrival to the quarantine facility, newly emerged adults from each location were reared on Brazilian peppertree saplings that were grown out-of-doors. In May 2014, three groups of 20 adults ranging in age from newly emerged to approximately 15 days old (maximum longevity [14]) were haphazardly collected from each population and placed at -20°C until processing.

DNA extraction and amplification

Psyllid DNA was extracted from three sub-samples of each population, using the DNeasy Blood & Tissue Kit (Qiagen, Venlo Netherlands) as reported for the extraction of DNA from the Asian citrus psyllid, *Diaphorina citri* [31]. DNA from twenty psyllids (a group) was

Table 1. Colony designations, species, origins and Florida State Collection of Arthropods voucher numbers of *Calophya* spp. examined for microbiome diversity.

Designation	Species	Origin in Brazil	Voucher no. Florida State Collection of Arthropods
Salvador	<i>Calophya latiforceps</i>	12.908°S, 38.336°W Salvador, Bahia State	E2013-3192-1
<i>C. tere</i>	<i>Calophya terebinthifolii</i>	26.921°S, 48.640°W Balneario Camboriu, Santa Catarina State	E2014-5749-1
Carapina	<i>Calophya latiforceps</i>	20.213°S, 40.229°W, Carapina, Espiritu Santo State	E2013-6744-1
Ubu	undescribed <i>Calophya</i> sp.	20.786°S, 40.579°W Ubu, Espiritu Santo State	E2013-6743-1

doi:10.1371/journal.pone.0132248.t001

extracted using the manufacturer’s instructions. Extracted DNA was stored at -20°C until used for PCR amplification.

Characterization of the psyllid microbiome

Genomic DNA was used as input for PCR amplification of the V4 variable region of microbial SSU rRNA genes using indexed primers as described previously [32]. Primers were assessed *in silico* using the Probe Match program associated with RDP and the program Pimer Prospector against our database [33,34]. These primers hit 99% of all high quality *Ca. Phytoplasma* sequences in the RDP database when allowing 1 mismatch (92% perfectly), and 100% of *Ca. Phytoplasma* sequences in our database.

After equalizing PCR yields, PCR products were pooled and cleaned. Amplification and sequencing were performed at the RTSF Genomics Core at Michigan State University. Pooled products were loaded on a standard MiSeq flow cell (v2) and sequenced in a paired end 2x250bp format. The resulting sequences were processed in QIIME v.1.8 [35]. Briefly, overlapping reads were merged using fastq-join [36], and quality filtering was performed using USEARCH v7.0 where reads with an expected error >0.5 were excluded from further analysis [37]. Only completely assembled reads of approximately 250 bases (after primer sequence removal) were utilized for further analysis. Operational taxonomic units (OTUs) were defined at 97% sequence similarity using USEARCH v.7.0 [37] and chimeras were removed *de novo* following the UPARSE pipeline and by comparing against representative sequences from the Greengenes database v.13_8 [38] clustered at 91% sequence identity to minimize computational time. OTUs were also defined at 99% sequence similarity using USEARCH v.7.0 as described above to determine if dominant OTUs from the populations originating from Salvador, Carapina, and Ubu could be distinguished.

A representative set of OTU sequences were classified using the Ribosomal Database Project classifying algorithm [33] that had been re-trained with representative sequences from the Greengenes database v.13_8. Sequences derived from plastids (chloroplast or mitochondria) were discarded from further analysis. Sequences were aligned using the PyNAST aligner [39] and the Greengenes v.13_8 database. Any sequence that failed to align was discarded from further analysis. Raw sequence data (FASTQ files) were deposited in the NCBI Sequence Read Archive under project SRP049390.

Statistical analysis

The QIIME generated OTU table was imported into R v.3.1.1 [40] and sampling effort was adjusted to a common scale by calculating size scaling factors using the R package DESeq2 [41]. Other normalization and transformation approaches were tested including using cumulative-sum scaling with the R package ‘metagenomeSeq’ [42], and subsampling to 100,000

sequences per library (S1 Fig). Shannon diversity for each sample was calculated from the DESeq2 adjusted OTU table using the package ‘Vegan’ [43]. An ANOVA was performed to test for differences in the means of the Shannon values, followed by Student-Newman-Keuls post-hoc means separation procedure ($P < 0.05$).

Rarefaction curves were generated using the ‘rarefy’ function in ‘vegan’ by repeatedly sub-sampling the OTU table at steps of 1000. A mean number of OTUs and the standard deviation for each sample per step was calculated using the R package ‘plyr’ and plotted using the R package ‘ggplot2’ [44,45]. Total richness was estimated using the R package ‘preseqR’ [46] by fitting a zero-truncated negative binomial model and extending to 100 million sequences per sample.

Two dominant classes for all microbiomes were identified, and the top OTUs belonging to these dominant classes were determined. OTU sequences were placed onto a pre-compiled SSU rRNA gene tree maintained by The Greengenes Database Consortium containing 408,000 full length Bacterial and Archaeal gene sequences using ARB [47]. Similar reference sequences within the database were selected and the phylogenetic tree was exported and visualized using iTOL [48].

To compare microbial community structures, all OTU counts were transformed using a variance stabilizing approach implemented in DESeq2 and recommended by McMurdie and Holmes [49]. A Bray-Curtis similarity matrix was calculated using the package ‘vegan’ in R [43] and similarities were visualized using a nonmetric multidimensional scaling plot (NMDS) generated using the R package ‘ggplot2’ [44]. Confidence regions (95%) for each group were calculated using ‘vegan’s’ oridellipse function and overlaid on the NMDS plot to estimate the area within the ordination space that contains the true value. Insect microbiome groupings (replicates) were statistically verified using an adonis test (R package ‘vegan’) and a SIMPROF test within the software package Primer-E v.6 [43,50,51].

Screening for potential plant pathogens

Initially, the RDP assigned taxonomy was screened for OTUs assigned to *Ca. Liberibacter* or *Ca. Phytoplasma* genera. After verifying that no OTU was annotated to either of these groups, a reference set of full length SSU rRNA gene sequences from each group was generated from sequences submitted to GenBank. The initial dataset included 34 *Ca. Liberibacter* and 589 *Ca. Phytoplasma* sequences of greater than 1400 bases. Reference sequences were screened phylogenetically to remove mis-annotated sequences (e.g. EU093079.1 –*Ca. Phytoplasma* RYL-GD, S2 Fig). An additional 3 of the sequences failed to align using the PyNAST aligner and were also excluded, yielding a total of 573 *Ca. Phytoplasma* sequences. Using nucleotide BLAST v.2.2.28+, all quality controlled amplicon sequences generated in this study were compared against this reference database (this database is provided in the supporting information). A positive detection for either group was considered >94% similarity across 80% of the sequence read.

Results

Characterization of the *Calophya* microbiome

The microbiomes of the four populations of *Calophya* spp., in triplicate, were determined by deep sequencing of microbial SSU rRNA gene amplicons generated from total insect DNA extracts (10 individuals per replicate sample). In total, over 2 million sequences were generated from 12 samples. All sequences passing quality control filters were derived from organisms within the domain Bacteria, with greater than 98% of the sequences derived from organisms within the phylum Proteobacteria. The microbiomes of all tested organisms were highly similar, had low overall species richness with 49–79 detected OTUs, and housed few dominant taxa (Fig 1). Only 5 OTUs had >1% abundance in populations from Salvador, Ubu, and Carapina,

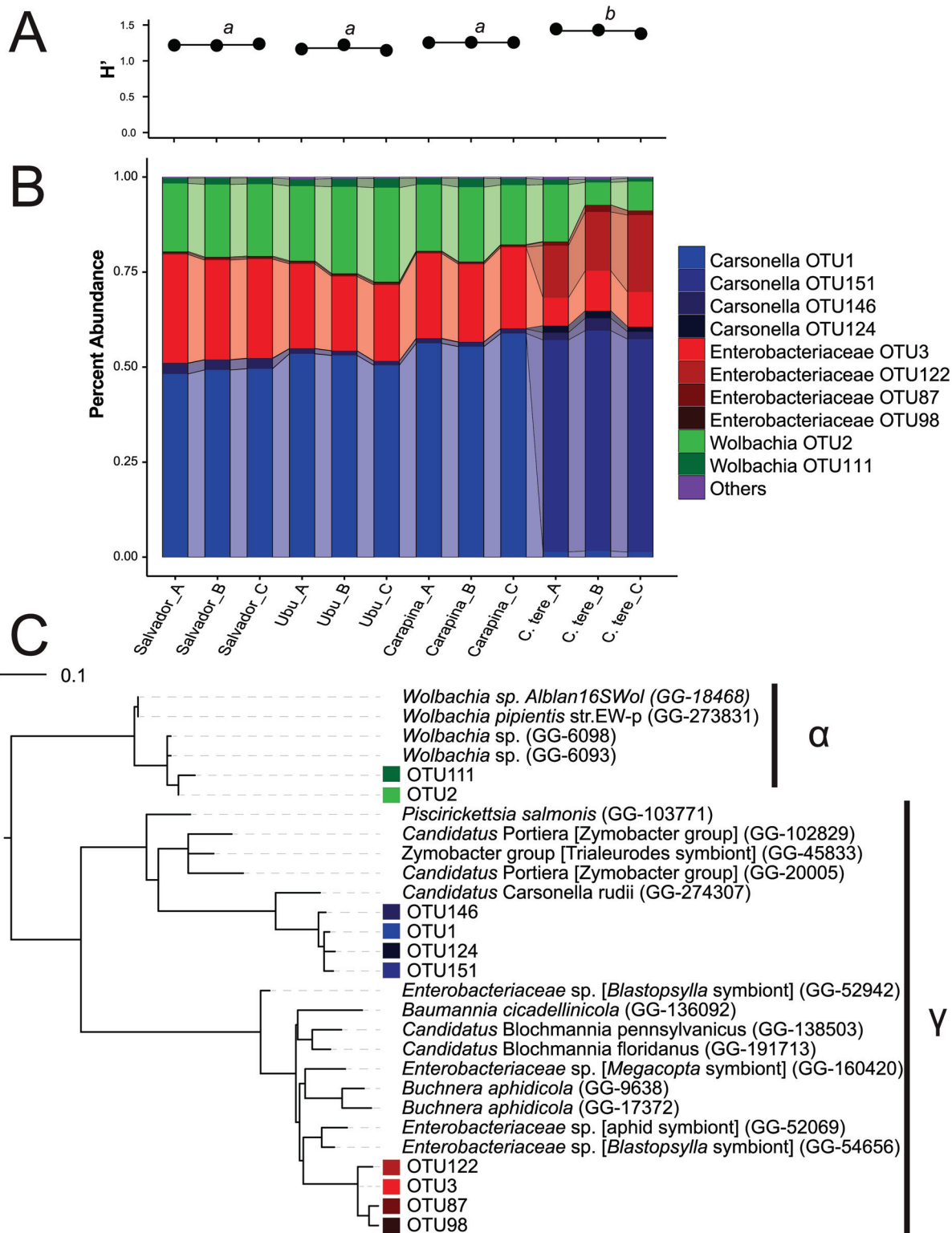


Fig 1. The diversity, composition, and phylogeny of dominant members of the *Calophya* spp. microbiome. (A) Shannon diversity indices for each microbiome sampled. The average value for each *Calophya* sp. (run in triplicate) is indicated with a line. Superscripts indicate groups that are statistically different by pairwise t-tests. (B) Microbial community composition for each microbiome analyzed. OTU abundances were estimated by converting libraries to a common scale using DESeq2, and normalized using total counts. OTUs from each dominant clade are colored with the same hue, and in all cases these dominant clades represent >98% of the total community. OTU bars are connected with semi-transparent edges to make it easier to distinguish between

different OTUs within the same group. (C) Dominant OTUs were inserted into a pre-compiled phylogenetic tree using ARB. OTU are labeled using the same color scheme as in (B).

doi:10.1371/journal.pone.0132248.g001

while 8 OTUs >1% were present in *C. terebinthifolii* (Table 2). Approximately 80% of the total richness was sampled as estimated using a fitted zero-truncated negative binomial model at 10 million sequences per sample. Rarefaction curves indicated that the majority of these OTUs were extremely low in abundance (Fig 2). Shannon diversity indices (base e) ranged from 1.16 to 1.44, and the microbial community sampled from *C. terebinthifolii* was significantly more diverse than the other three populations (Fig 1). The Salvador population was more diverse than Carapina, but Salvador was not different from Ubu and Ubu was not more diverse than Carapina ($P < 0.05$ for all comparisons).

The microbial community structures for all *Calophya* spp. were compared using the Bray-Curtis diversity index, which uses the microbial populations present and their relative abundances. A relatively extreme transformation, variance stabilization approach using the R package 'DESeq2', was applied, similar to what was observed with a log (X+1) transformation (S1 Fig). Transformation choice played a relatively important role in distinguishing between the populations from Salvador, Ubu, and Carapina. In all cases, a nonparametric ANOVA (adonis) test revealed sample groups were significant ($R^2 > 0.90$, $P < 0.01$), but these tests do not have enough power to resolve post-hoc pairwise testing. However, *C. terebinthifolii* was significantly different from the other three populations based on non-overlapping 95% confidence ellipsoids and a SIMPROF test with 9999 permutations, independent of the normalization and transformations that were applied (Fig 3A and 3B, S1 Fig). Conversely, *Calophya latiforceps*-Salvador could only be distinguished from Ubu and Carapina as an internal cluster (with non-overlapping 95% confidence ellipsoids) after applying a strong transformation. Using the R package 'metagenomeSeq' to normalize libraries to a common scale (using cumulative-sum scaling to the 75th percentile) [42], we can distinguish between *C. latiforceps*-Salvador and the other two populations.

The microbiome of *C. terebinthifolii* was highly distinct from that of all other *Calophya* species tested. Nonetheless, the *C. terebinthifolii* microbiome contained dominant taxa that were closely related to the dominant taxa in other species. Bacteria from the candidate genus *Ca. Carsonella* were the most abundant in all *Calophya* tested, with four OTUs representing between 51–65% of rRNA gene amplicon sequences from each sample. Two dominant OTUs within the *Ca. Carsonella* clade were detected. *Ca. Carsonella* OTU_1 was present in all samples but was highly enriched in *C. latiforceps* (Salvador and Carapina) and *Calophya* sp. (Ubu) relative to *C. terebinthifolii*. The *C. terebinthifolii* microbiome was dominated by a distinct *Ca. Carsonella* population, OTU_151, which was not detected in the other three insect populations (Fig 1B). The next most abundant clade, consisting of four OTUs, was affiliated with an unclassified Enterobacteriaceae group (Gammaproteobacteria) similar to *Buchnera* and *Blochmannia* (90–94% sequence identity) that ranged from 20–31% abundance (Fig 1B and 1C). Enterobacteriaceae OTU_87 was more abundant in *C. terebinthifolii*, while OTU_3 was enriched in the other three insect microbiomes. *Wolbachia* (Alphaproteobacteria) taxa were detected in all microbiomes ranging in abundance from 7–27%, and were primarily represented by a single OTU (OTU2). Overall the *C. terebinthifolii* microbiome had a lower abundance of *Wolbachia* sequences than the microbiome of other *Calophya* species, although there was substantial variation between the *C. terebinthifolii* replicates (Figs 1B and 3A).

Table 2. Sequencing statistics for *Calophya* spp. microbiomes.

SampleID ^a	Sequences Retrieved ^b	Number of OTUs ^c	Estimated Richness ^d	Shannon (H')	Num OTU > 1% Abundance
Salvador A	185424	50 (45)	65 [77%]	1.25 (1.25)	5 (5)
Salvador B	190544	54 (46)	66.46 [81%]	1.26 (1.26)	5 (5)
Salvador C	200392	49 (43)	61.52 [80%]	1.26 (1.25)	5 (5)
Ubu A	180116	69 (59)	85.82 [80%]	1.22 (1.22)	5 (5)
Ubu B	167006	64 (53)	77.08 [83%]	1.21 (1.22)	5 (5)
Ubu C	155472	62 (56)	81.51 [76%]	1.24 (1.24)	5 (5)
Carapina A	172604	58 (55)	79.69 [73%]	1.16 (1.17)	5 (5)
Carapina B	158362	61 (60)	87.45 [70%]	1.22 (1.22)	5 (5)
Carapina C	151683	68 (63)	92.16 [74%]	1.15 (1.15)	5 (5)
<i>C. tere</i> _A	184772	71 (63)	91.68 [77%]	1.44 (1.45)	8 (8)
<i>C. tere</i> _B	179729	79 (69)	101.33 [78%]	1.43 (1.44)	8 (8)
<i>C. tere</i> _C	107640	56 (56)	81.29 [69%]	1.38 (1.38)	8 (8)

^aSalvador and Carapina belong to *C. latiforceps*, Ubu refers to an undescribed *Calophya* species, and *C. tere* refers to *C. terebinthifolii* (see Table 1). Values in parentheses were calculated using rarefied libraries to 100,000 sequences per sample, values before parentheses were calculated using DESeq2 normalized OTU counts.

^bNumber of sequences after quality control.

^cOTU defined at 97% sequence similarity.

^dPredicted using a fitted zero truncated negative binomial model to 100 million sequences per sample for rarefied data only. Values in brackets are % coverage using DESeq2 normalized OTU counts.

doi:10.1371/journal.pone.0132248.t002

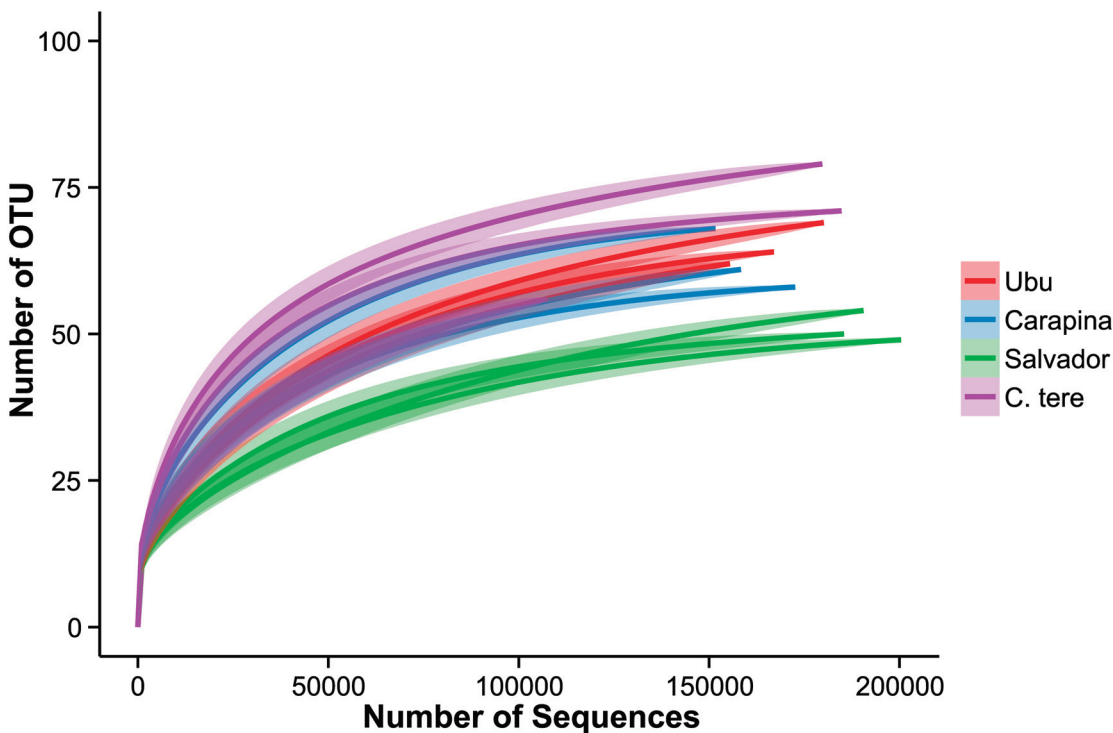


Fig 2. The microbial community richness sampled from *Calophya* spp. Rarefaction curves generated in R using the package ‘vegan’. Standard deviations for each step are indicated with the semi-transparent shading.

doi:10.1371/journal.pone.0132248.g002

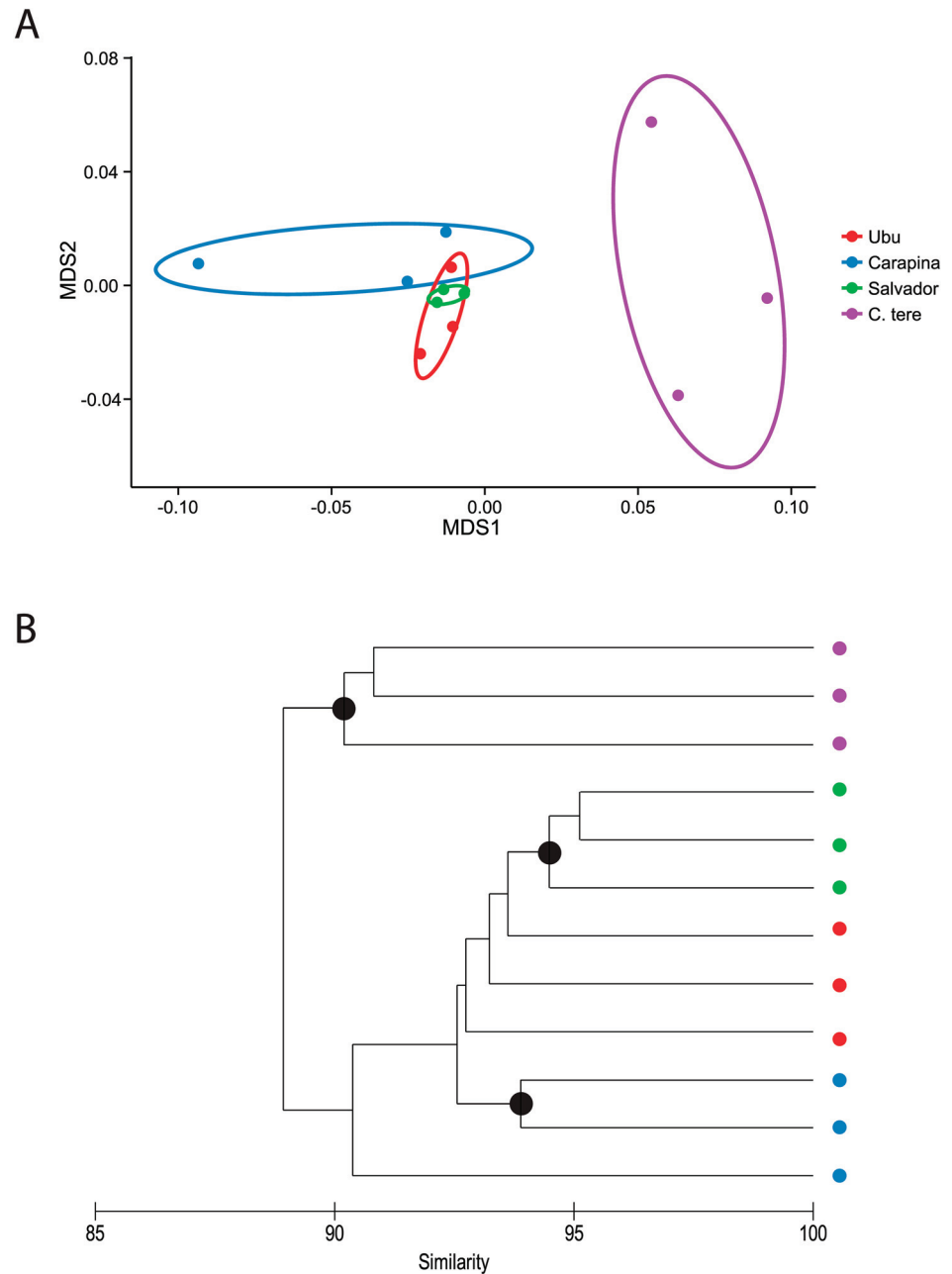


Fig 3. Comparing microbial communities associated with *Calophya* spp. (A) *Calophya* microbiomes were compared by computing Bray-Curtis similarities using variance stabilized OTU counts, and visualized using a non-metric multidimensional scaling plot (stress = 0.09). 95% confidence ellipses are shown, and *C. terebinthifolii* confidence ellipses clearly do not overlap with the other *Calophya* spp. (B) A SIMPROF test was performed and visualized using Primer-6. Black circles indicate groups that cannot be significantly differentiated. Again, *C. terebinthifolii* forms a cohesive group distinguishable from the other three *Calophya* spp. which cannot be separated into distinct groups.

doi:10.1371/journal.pone.0132248.g003

Absence of pathogenic bacteria

A database of reference sequences of disease causative agents in the genera *Ca. Liberibacter* and *Ca. Phytoplasma* was generated that contained 603 high quality, phylogenetically verified, >1400bp sequences (Supporting Information). All 2 million quality controlled SSU rRNA gene amplicon reads were aligned to the reference database using blastn [52], retaining any hits >94% sequence similarity for at least 200 bases (~80% of the sequence). No sequences met these criteria, out of a total of 2,077,290 sequences. The closest sequences were 93% similar to *Ca. Liberibacter crescens* (Genbank accession number NR_102476); however, they were 94% similar to *Rhizobium leguminosarum* (AY509900), and 99% similar to an unclassified *Rhizobium* (JQ977368). These *Rhizobium* and *Ca. Liberibacter* strains were approximately 92% similar to each other across nearly full length SSU rRNA gene sequences. This analysis demonstrated that there were no detectable populations closely related to *Ca. Liberibacter* or *Ca. Phytoplasma* in the *Calophya* microbiomes examined in this study.

Discussion

Phylogeny of bacteria and *Calophya*

The objectives of this study were to determine if *Calophya* species harbor plant pathogenic bacteria, and more broadly to assess the whole-body microbiome of these promising biocontrol agents. Despite an extremely deep SSU rRNA gene amplicon sequencing effort, no pathogens were detected, and we conclude that releasing *Calophya* species as biocontrol agents of Brazilian peppertree poses no risk of introducing plant pathogenic bacteria.

Overall, the *Calophya* microbiome of multiple species was found to be extraordinarily limited, consistent with prior studies of psyllid microbiomes [25]. Specifically, all four populations of *Calophya* spp. contained the primary obligate endosymbiont *Ca. Carsonella* [53]. *Ca. Carsonella* is considered a primary endosymbiont of Psylloidea, with a role in synthesizing essential nutrients required by their sap-feeding hosts [28,54]. Recent work has shown that the highly degraded genome of *Ca. Carsonella* is complemented in amino acid synthesis by genes from secondary endosymbionts [55] and host genes, some of which may have been gained through horizontal transfer from *Ca. Carsonella* [56].

As expected, our results are consistent with the assumption of co-diversification of *Ca. Carsonella* and the host *Calophya* species. A phylogenetic reconstruction of *Calophya* using the mitochondria cytochrome c oxidase I (MT-COI) gene revealed that *C. terebinthifolii* was 13–14% divergent from the other three *Calophya* populations [15]. This corresponded to a 3% sequence divergence in the V4 region of the SSU rRNA gene for the *Ca. Carsonella* populations. Higher similarities between the SSU rRNA genes relative to the MT-COI genes were expected due to the high degree of sequence conservation in the SSU rRNA gene [28,57]. MT-COI analysis further revealed a 6% sequence divergence between *C. latiforceps*, and *C. sp. Ubu*, which along with morphological differences, was sufficient to warrant further investigation into describing a new species [15]. However, the dominant *Ca. Carsonella* populations retrieved from *C. sp. Ubu* were >99% identical to the ones found in *C. latiforceps*. We believe it is unlikely that *Ca. Carsonella* has been shared across matrilineages [58], and this pattern is due to the lower resolving power in short SSU rRNA gene fragments [28]. Finally, the *Ca. Carsonella* population from *C. latiforceps* was indistinguishable from the one retrieved from Carapina, which was unsurprising since the host populations were also indistinguishable using MT-COI [15].

The majority of psyllids studied to date have contained one or many co-primary symbionts [59,60]. *Ca. Carsonella* species have a highly reduced genome that does not contain sufficient amino acid biosynthesis pathways to meet the needs of the host, and previous studies have

shown these functions can be complemented by an obligate secondary endosymbiont [61], or possibly by incorporation of symbiont genes into the host genome [55,61,62]. In our study, all populations of *Calophya* contained sequences derived from bacteria of the broadly distributed family Enterobacteriaceae. Closely related and described strains include the primary symbionts *Blochmannia* from ants, and *Buchnera* from aphids [28,63,64]. Both *Blochmannia* and *Buchnera* have demonstrated roles in nutrient provision [65–68]. Sequences derived from Enterobacteriaceae in this study were highly similar (94–98% identity) to a sequence isolated from *Calophya schini* Tuthill [28], an insect which feeds on a congener of Brazilian peppertree, *Schinus molle* L.

While a definitive phylogenetic tree for the genus *Calophya* has not been conclusively generated, based on morphology, *C. schini* is an outgroup to the *Calophya* spp. used in this study [10,11]. This taxonomy is consistent with the divergent microbiome of *C. terebinthifolii*. This divergent microbiome included a distinct dominant *Ca. Carsonella* OTU and a distinct Enterobacteriaceae OTU (>3%) that was absent from the other *Calophya* populations. Furthermore, the three *Calophya* populations with indistinguishable *Ca. Carsonella* populations also had indistinguishable dominant Enterobacteriaceae populations (>99% similar). Interestingly, *C. terebinthifolii* maintained the distinct Enterobacteriaceae group and the shared Enterobacteriaceae group. We hypothesize that this was due to horizontal re-acquisition of the Enterobacteriaceae strain present in the other three populations, as horizontal transfers of secondary symbionts have been shown to occur [28]. However, we cannot verify that both Enterobacteriaceae populations were present in all *C. terebinthifolii* individuals, as each replicate was composed of a mixture of 10 insects. Hypothetically, this same pattern would emerge if approximately 30% of *C. terebinthifolii* insects contained only the more typical Enterobacteriaceae population, and ca. 70% contained only the divergent strain.

The final dominant microbial population detected was affiliated with the Alphaproteobacterial genus, *Wolbachia*. *Wolbachia* are a well-studied and diverse group of intracellular bacteria that infect arthropods and nematodes. *Wolbachia* infect more than 65% of insect species, where they are generally considered reproductive parasites because they manipulate host reproduction to favor their spread. This is accomplished through several mechanisms including cytoplasmic incompatibility, induction of parthenogenesis, feminization and male killing (reviewed in [69]). In addition, *Wolbachia* has been shown to play a mutualistic role in *Drosophila melanogaster* by increasing resistance to RNA viruses [70]. The role of *Wolbachia* in *Calophya* spp. is unknown, but parthenogenesis, feminization and male killing can probably be ruled out, as there has been no evidence for these effects in the established colonies at the University of Florida. Bacteria from the genus *Wolbachia* typically do not share phylogenetic congruencies with their host, and indeed in this study we observed one dominant *Wolbachia* population in all *Calophya* species. This is similar to other studies of psyllid endosymbionts [25,28,71]

Using the relationships between the entire *Calophya* microbiomes, we also observe a strong congruency to the phylogeny of *Calophya*. The 13–14% divergence in MT-COI genes between *C. terebinthifolii* corresponds to an average Bray-Curtis dissimilarity value of 11% using a variance stabilizing transformation (78% using untransformed data). This was expected since *Ca. Carsonella* represented ~50% of the total microbiome for each *Calophya* sp. Hence, utilizing both the similarities between microbial community structures and individual endosymbiont populations, we were still incapable of distinguishing between *Calophya latiforceps* and *Calophya* sp. Ubu.

Reduced bacterial diversity in phloem-feeding insects

Our results provide support for the hypothesis that sap-feeding Hemipterans have low bacterial diversity compared to other insects [25]. Using SSU rRNA gene amplicon sequencing, we

detected between 5 and 8 OTUs with an abundance of >1% and an average Shannon index (H' , log e) of 1.27 (range = 1.16–1.44) per *Calophya* sample. In contrast, non-sap-feeding Hemipterans have larger numbers of bacterial OTUs per sample including Hymenoptera (11 OTUs), Lepidoptera (32 OTUs), and Isoptera (90 OTUs) [72]. Our OTU numbers and Shannon indices are consistent with those reported for the potato psyllid, *B. cockerelli* (Triozidae), and the Asian citrus psyllid, *D. citri* (Psyllidae), which ranged between 3 and 7 OTUs [25,27,71] and Shannon indices (log e) of 0.58 and 1.18, respectively [25]. Similarly, whiteflies (*Bemisia tabaci* Gennadius, [25]) and aphids (*Acyrtosiphon pisum* Harris, [25]; *Aphis glycines* Matsumura, [73]) have low bacterial diversity. Insects feeding on more complex diets appear to require a more diverse microbiota for digestive purposes. For example, the number of OTUs per sample for xylophagous insects of decaying wood and for detritivores was 103 and 53, respectively [72]. The reason for the low diversity of bacteria found in phloem-feeding Hemipterans is unknown, but could be due to the low variation in physical and chemical composition of sap [74], which might not require an association with a large number of bacteria for nutritional purposes. Furthermore, plant galls are nutrient sinks [75] which provide an enriched food source for gall-inducing Hemipterans such as *Calophya* spp. This high quality diet may require less microbial processing than typical phloem sap. Future studies could test this hypothesis by comparing the microbiome diversity in non-galling and gall-inducing *Calophya*.

Absence of plant pathogens in *Calophya* spp.

The psyllid (Psylloidea) families Psyllidae and Triozidae include several members that transmit plant pathogenic bacteria [76], including species in the genera *Ca. Phytoplasma* and *Ca. Liberibacter* [17]. Examples include Asian citrus psyllid (*Diaphorina citri* Kumayama (Psyllidae) transmission of species of *Ca. Liberibacter* that cause Huanglongbing in *Citrus* spp., and carrot yellows decline transmitted by *Triozia apicalis* Förster and *Bactericera trigonica* Hodkinson (Triozidae) [19,20]. *Ca. Phytoplasma* diseases include European stone fruit yellows (*P. prunorum*), transmitted by *Cacopsylla pruni* Scopoli [77] and pear decline (*P. pyri*) transmitted by *Cacopsylla pyricola* (Förster) [78]. Although we have not been able to find any records of bacterial plant pathogen transmission by species in the family Calophyidae, we examined *Calophya* to determine whether any plant pathogenic bacteria could be detected. In this study, extraordinarily deep SSU rRNA gene amplicon sequencing (>100,000 sequences per sample) did not reveal any known pathogens. This result confirms prior assays in which specific primers to detect the presence of four species of the genus *Ca. Liberibacter* (*i.e.*, *Ca. L. solanacearum*, *Ca. L. asiaticus*, *Ca. L. americanus* and *Ca. L. africanus*) were negative [15]. The probability that psyllids infected with plant pathogenic bacteria were present in our colonies, but not sampled, can be calculated if we assume a background level of infection. The reported levels of infection by *Ca. Liberibacter asiaticus*, the causative agent of citrus greening disease, in field collected *D. citri* in Florida ranged from 8.75% (Munjanuth et al. 2008) to 81.0% (Ammar et al. 2011). If we conservatively assume the lower level of infection, by assaying 240 psyllids out of an estimated colony size of 4000, the probability that plant pathogenic bacteria were present but not detected was extremely low (1.39×10^{-10} , hypergeometric test).

Virus transmission by *C. latiforceps* is also unlikely as only one member of the superfamily Psylloidea is known to vector a virus [79]. Regardless, we previously tested *C. latiforceps* for Tomato torrado virus and Tomato chocolate spot virus using specific primers, and also inoculated highly susceptible indicator plants with homogenized insects, and found no evidence of viruses [30]. Additional evidence of lack of pathogen transmission would be available if we could rear the insects on infected plants to examine acquisition, but host range testing on 89 plant species demonstrated that *C. calophya* will only feed on Brazilian peppertree [30]. Thus,

we have not been able to determine whether transmission is theoretically possible. Nonetheless, due to the extremely restricted diet and the largely obligate and simple microbiome associated with *Calophya* species, we believe that these organisms are safe to release in the United States.

Supporting Information

S1 Fig. Comparing the effect of OTU table transformations and normalizations on *Calophya* spp. microbiome beta diversity. (A) OTU counts have been normalized using a variance stabilized approach as described in the main methods section. (B) A $\log_{10}(X+1)$ transformation on the DESeq2 common scaled OTU table, Love et al. (2014) stated that the variance stabilized normalization was similar to a log normalization [41]. (C) OTU counts have been normalized to a common scale using DESeq2. (D) OTU counts have been normalized using the cumulative-sum scaling method applied with the R package 'metagenomeSeq' [42]. (E) OTU counts have been normalized to a common-scale using DESeq2 and square root transformed. (F) OTU counts have been normalized using cumulative-sum scaling and square root transformed. (G) OTU counts have been randomly subsampled without replacement to 100,000 sequences per sample. (H) Raw, non-transformed and non-normalized, OTU counts. (I) Subsampled OTU counts have been square root transformed. (J) Subsampled OTU counts have been $\log_{10}(X+1)$ transformed.

(PDF)

S2 Fig. Validating the taxonomic annotation of GenBank sequences associated with *Phytoplasma* used in the reference database. As described in the methods section, the GenBank nucleotide database was screened for any sequence affiliated with the SSU rRNA gene from *Phytoplasma* spp. All sequences were aligned using PyNAST and a maximum likelihood tree was generated using MEGA v6. Misannotated sequences were excluded from the database.

(PDF)

S1 Dataset. Reference database of full length or nearly full length *Liberibacter* and *Phytoplasma* sequences.

(FASTA)

Acknowledgments

Carolina Arguijo, Liz Calise and Amy Dubois from the Weed Biological Control Program at BCRCCL assisted with the maintenance of insect and plant colonies.

Author Contributions

Conceived and designed the experiments: Will A. Overholt RD ER SJG William A. Overholt. Performed the experiments: ER SJG RD. Analyzed the data: Will A. Overholt. Contributed reagents/materials/analysis tools: William A. Overholt ER SJG. Wrote the paper: Will A. Overholt RD SJG ER William A. Overholt.

References

1. Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E. Quantifying threats to imperiled species in the United States. *Bioscience*. 1998; 48: 607–15.
2. Simberloff D, Schmitz DC, Brown TC. *Strangers in paradise: impact and management of nonindigenous species in Florida*. Washington DC: Island Press; 1997.
3. Wunderlin RP, Hansen BF. *Atlas of Florida Vascular Plants*. Inst Syst Bot Univ south Florida, Tampa [Internet]. University of South Florida; 2008. Available: <http://www.florida.plantatlas.usf.edu/Family.aspx?id=42>

4. FLEPPC [Internet]. 2013. Available: <http://www.fleppc.org/>
5. Rodgers L, Bodle M, Black D, Laroche F, Pratt P, Boughton A, et al. Status of nonindigenous species. South Florida Environ Rep 2012 (Chapter 7). 2012; 1–35. Available: http://www.sfwmd.gov/portal/page/portal/pg_grp_sfwmd_sfer/portlet_prevreport/2012_sfer/v1/chapters/v1_ch7.pdf
6. Workman R. *Schinus*. Technical Proceedings of Techniques for Control of *Schinus* in South Florida: A Workshop for Natural Area Managers Sanibel-Captiva Conservation Foundation, Sanibel, FL. 1979.
7. Gioeli K, Langeland K. Brazilian pepper-tree control. University of Florida, Cooperative Extension Service. Institute of Food and Agricultural Sciences. SS-AGR-17; 1997.
8. Doren RF, Jones DT. Plant management in Everglades National Park. In: Simberloff D, Schmitz DC, Brown TC, editors. Strangers in Paradise: Impact and Management of Nonindigenous Species in Florida. Washington DC: Island Press; 1997. pp. 275–86.
9. Manrique V, Cuda JP, Overholt WA. Brazilian peppertree: a poster child for invasive plants in Florida. J Florida Stud. 2013; 1: 1–14.
10. Burckhardt D, Basset Y. The jumping plant-lice (Hemiptera, Psylloidea) associated with *Schinus* (Anacardiaceae): systematics, biogeography and host plant relationships. J Nat Hist. 2000; 34: 57–155.
11. Burckhardt D, Cuda JP, Manrique V, Diaz R, Overholt WA, Williams DA., et al. *Calophya latiforceps*, a new species of jumping plant lice (Hemiptera: Calophyidae) associated with *Schinus terebinthifolius* (Anacardiaceae) in Brazil. Florida Entomol. 2011; 94: 489–499.
12. Hardy NB, Cook LG. Gall-induction in insects: evolutionary dead-end or speciation driver? BMC Evol Biol. 2010; 10: 257. doi: [10.1186/1471-2148-10-257](https://doi.org/10.1186/1471-2148-10-257) PMID: [20735853](https://pubmed.ncbi.nlm.nih.gov/20735853/)
13. Burckhardt D, Ouvrard D, Queiroz D, Percy D. Psyllid host-plants (Hemiptera: Psylloidea): resolving a semantic problem. Florida Entomol. 2014; 97: 242–6.
14. Diaz R, Moscoso D, Manrique V, Williams D, Overholt WA. Native range density, host utilisation and life history of *Calophya latiforceps* (Hemiptera: Calophyidae): an herbivore of Brazilian Peppertree (*Schinus terebinthifolia*). Biocontrol Sci Technol. 2014; 24: 536–53.
15. Diaz R, Dickey AM, Shatters RGJ, Manrique V, Overholt WA. New species diversity revealed from molecular and morphological characterization of gall-inducing *Calophya* spp. (Hemiptera: Calophyidae) from Brazilian peppertree. Florida Entomol. 2015; 98: 776–779.
16. Haapalainen M. Biology and epidemics of *Candidatus Liberibacter* species, psyllid-transmitted plant-pathogenic bacteria. Ann Appl Biol. 2014; 165: 172–98.
17. Weintraub PG, Beanland L. Insect vectors of phytoplasmas. Annu Rev Entomol. 2006; 51: 91–111. PMID: [16332205](https://pubmed.ncbi.nlm.nih.gov/16332205/)
18. Hall DG, Richardson ML, Ammar E, Halbert SE. Asian citrus psyllid, *Diaphorina citri*, vector of citrus huanglongbing disease. Entomol Exp Appl. 2013; 146: 207–23.
19. Munyaneza JE. Psyllids as vectors of emerging bacterial diseases of annual crops. Southwest Entomol. 2010; 35: 471–7.
20. Munyaneza JE. Zebra chip disease of potato: biology, epidemiology, and management. Am J Potato Res. 2012; 89: 329–50.
21. Tedeschi R, Lauterer P, Brusetti L, Tota F, Alma A. Composition, abundance and phytoplasma infection in the hawthorn psyllid fauna of northwestern Italy. Eur J plant Pathol. 2009; 123: 301–10.
22. Thompson S, Fletcher JD, Ziebell H, Beard S, Panda P, Jorgensen N, et al. First report of '*Candidatus Liberibacter europaeus*' associated with psyllid infested Scotch broom. New Dis Rep. 2013; 27: 6.
23. Arp AP, Chapman R, Crosslin JM, Bextine B. Low-Level Detection of *Candidatus Liberibacter Solanacearum* in *Bactericera cockerelli* (Hemiptera: Trioziidae) by 16s rRNA Pyrosequencing. 2013; 42: 868–73.
24. Hail D, Dowd SE, Bextine B. Identification and location of symbionts associated with potato psyllid (*Bactericera cockerelli*) lifestyles. Environ Entomol. 2012; 41: 98–107. doi: [10.1603/EN11198](https://doi.org/10.1603/EN11198) PMID: [22525064](https://pubmed.ncbi.nlm.nih.gov/22525064/)
25. Jing X, Wong AC-N, Chaston JM, Colvin J, McKenzie CL, Douglas AE. The bacterial communities in plant phloem-sap-feeding insects. Mol Ecol. 2014; 23: 1433–44. doi: [10.1111/mec.12637](https://doi.org/10.1111/mec.12637) PMID: [24350573](https://pubmed.ncbi.nlm.nih.gov/24350573/)
26. Subandiyah S, Nikoh N, Tsuyumu S. Complex endosymbiotic microbiota of the citrus psyllid *Diaphorina citri* (Homoptera: Psylloidea). Zoolog Sci. 2000; 17: 983–9.
27. Fagen JR, Giongo A, Brown CT, Davis-Richardson AG, Gano KA, Triplett EW. Characterization of the relative abundance of the citrus pathogen *Ca. Liberibacter asiaticus* in the microbiome of its insect vector, *Diaphorina citri*, using high throughput 16S rRNA sequencing. Open Microbiol J. 2012; 6: 29. doi: [10.2174/1874285801206010029](https://doi.org/10.2174/1874285801206010029) PMID: [22529882](https://pubmed.ncbi.nlm.nih.gov/22529882/)
28. Thao ML, Clark M a, Baumann L, Brennan EB, Moran N a, Baumann P. Secondary endosymbionts of psyllids have been acquired multiple times. Curr Microbiol. 2000; 41: 300–4. PMID: [10977900](https://pubmed.ncbi.nlm.nih.gov/10977900/)

29. Thao ML, Moran NA, Abbot P, Brennan EB, Burckhardt DH, Baumann P. Cospeciation of psyllids and their primary prokaryotic endosymbionts. *Appl Environ Microbiol*. 2000; 66: 2898–905. PMID: [10877784](#)
30. Diaz R, Manrique V, Munyaneza J, Sengoda V, Adkins S, Hendricks K, et al. Host specificity testing and examination for plant pathogens reveals that the gall-inducing psyllid, *Calophya latiforceps* (Hemiptera: Calophyidae), is safe to release for biological control of Brazilian peppertree. *Entomol Exp Appl*. 2014; 154: 1–14.
31. Matos LA, Hilf ME, Chen J, Folimonova SY. Validation of 'variable number of tandem repeat'—based approach for examination of '*Candidatus* Liberibacter asiaticus' diversity and its applications for the analysis of the pathogen populations in the areas of recent introduction. *PLoS One*. 2013; 8: e78994. doi: [10.1371/journal.pone.0078994](#) PMID: [24223873](#)
32. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012; 6: 1621–4. doi: [10.1038/ismej.2012.8](#) PMID: [22402401](#)
33. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*. 2007; 73: 5261–7. PMID: [17586664](#)
34. Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer N, Knight R. PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics*. 2011; 27: 1159–61. doi: [10.1093/bioinformatics/btr087](#) PMID: [21349862](#)
35. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010; 7: 335–6. doi: [10.1038/nmeth.f.303](#) PMID: [20383131](#)
36. Aronesty E. Command-line tools for processing biological sequencing data [Internet]. 2011. Available: <http://code.google.com/p/ea-utils>
37. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods*. 2013; 10: 996–8. doi: [10.1038/nmeth.2604](#) PMID: [23955772](#)
38. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006; 72: 5069–72. PMID: [16820507](#)
39. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics*. 2010; 26: 266–7. doi: [10.1093/bioinformatics/btp636](#) PMID: [19914921](#)
40. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2014. Available: <http://www.r-project.org/>
41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol*. 2014; 15: 550. PMID: [25516281](#)
42. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013; 10: 1200–2. doi: [10.1038/nmeth.2658](#) PMID: [24076764](#)
43. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package. 2013.
44. Wickham H. ggplot2: elegant graphics for data analysis [Internet]. Springer New York; 2009. Available: <http://had.co.nz/ggplot2/book>
45. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. *J Stat Softw* [Internet]. 2011; 40: 1–29. Available: <http://www.jstatsoft.org/v40/i01/>
46. Deng C, Daley T, Smith AD. preseqR: Predicting Species Accumulation [Internet]. 2014. Available: <http://cran.r-project.org/package=preseqR>
47. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, et al. ARB: a software environment for sequence data. *Nucleic Acids Res*. 2004; 32: 1363–71. PMID: [14985472](#)
48. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*. 2007; 23: 127–8. PMID: [17050570](#)
49. McMurdie PJ, Holmes S. Waste Not, Want Not: Why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. 2014; 10: e1003531. doi: [10.1371/journal.pcbi.1003531](#) PMID: [24699258](#)
50. Clarke K, Warwick R. Changes in marine communities: an approach to statistical analysis and interpretation. Plymouth: PRIMER-E; 2001.
51. Clarke K. Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol*. 1993; 18: 117–43.
52. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990; 215: 403–10. PMID: [2231712](#)

53. Sloan DB, Moran NA. Genome reduction and co-evolution between the primary and secondary bacterial symbionts of psyllids. *Mol Biol Evol.* 2012; 29:3781–92. doi: [10.1093/molbev/mss180](https://doi.org/10.1093/molbev/mss180) PMID: [22821013](https://pubmed.ncbi.nlm.nih.gov/22821013/)
54. Clark MA, Baumann L, Thao ML, Moran NA, Baumann P. Degenerative minimalism in the genome of a psyllid endosymbiont. *J Bacteriol.* 2001; 183: 1853–61. PMID: [11222582](https://pubmed.ncbi.nlm.nih.gov/11222582/)
55. Nakabachi A, Koshikawa S, Miura T, Miyagishima S. Genome size of *Pachypsylla venusta* (Hemiptera: Psyllidae) and the ploidy of its bacteriocyte, the symbiotic host cell that harbors intracellular mutualistic bacteria with the smallest cellular genome. *Bull Entomol Res.* 2010; 100: 27–33. doi: [10.1017/S0007485309006737](https://doi.org/10.1017/S0007485309006737) PMID: [19302725](https://pubmed.ncbi.nlm.nih.gov/19302725/)
56. Sloan DB, Nakabachi A, Richards S, Qu J, Murali SC, Gibbs RA, et al. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol Biol Evol.* 2014; 31: 857–71. doi: [10.1093/molbev/msu004](https://doi.org/10.1093/molbev/msu004) PMID: [24398322](https://pubmed.ncbi.nlm.nih.gov/24398322/)
57. Woese CR, Fox GE, Zablen L, Uchida T, Bonen L, Pechman K, et al. Conservation of primary structure in 16S ribosomal RNA. *Nature.* 1975; 254 (5495): 83–6. PMID: [1089909](https://pubmed.ncbi.nlm.nih.gov/1089909/)
58. Tamames J, Gil R, Latorre A, Peretó J, Silva FJ, Moya A. The frontier between cell and organelle: genome analysis of *Candidatus Carsonella ruddii*. *BMC Evol Biol.* 2007; 7: 181. PMID: [17908294](https://pubmed.ncbi.nlm.nih.gov/17908294/)
59. Buchner P. *Endosymbiosis of animals with plant microorganisms.* New York: John Wiley & Sons; 1965.
60. Moran NA, McCutcheon JP, Nakabachi A. Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet.* 2008; 42: 165–90. doi: [10.1146/annurev.genet.41.110306.130119](https://doi.org/10.1146/annurev.genet.41.110306.130119) PMID: [18983256](https://pubmed.ncbi.nlm.nih.gov/18983256/)
61. Nakabachi A, Ueoka R, Oshima K, Teta R, Mangoni A, Gurgui M, et al. Defensive bacteriome symbiont with a drastically reduced genome. *Curr Biol.* 2013; 23: 1478–84. doi: [10.1016/j.cub.2013.06.027](https://doi.org/10.1016/j.cub.2013.06.027) PMID: [23850282](https://pubmed.ncbi.nlm.nih.gov/23850282/)
62. Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, et al. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science.* 2006; 314 (5797): 267. PMID: [17038615](https://pubmed.ncbi.nlm.nih.gov/17038615/)
63. Baumann P. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu Rev Microbiol.* 2005; 59: 155–89. PMID: [16153167](https://pubmed.ncbi.nlm.nih.gov/16153167/)
64. Moran NA, Telang A, Smith M. Bacteriocyte of insects symbionts: a variety of insect groups harbor ancient endosymbionts prokaryotic. *Bioscience.* 1998; 48: 295–304.
65. Douglas AE. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annu Rev Entomol.* 1998; 43:17–37. PMID: [15012383](https://pubmed.ncbi.nlm.nih.gov/15012383/)
66. Feldhaar H, Straka J, Kruschke M, Berthold K, Stoll S, Mueller MJ, et al. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. *BMC Biol.* 2007; 5: 48. PMID: [17971224](https://pubmed.ncbi.nlm.nih.gov/17971224/)
67. Sandström J, Moran N. How nutritionally imbalanced is phloem sap for aphids? *Proceedings of the 10th International Symposium on Insect-Plant Relationships.* Springer; 1999. pp. 203–10.
68. Moran NA, Degnan PH. Functional genomics of *Buchnera* and the ecology of aphid hosts. *Mol Ecol.* 2006; 15: 1251–61. PMID: [16626452](https://pubmed.ncbi.nlm.nih.gov/16626452/)
69. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol.* 2008; 6: 741–51. doi: [10.1038/nrmicro1969](https://doi.org/10.1038/nrmicro1969) PMID: [18794912](https://pubmed.ncbi.nlm.nih.gov/18794912/)
70. Teixeira L, Ferreira Á, Ashburner M. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol.* 2008; 6: e1000002.
71. Nachappa P, Levy J, Pierson E, Tamborindeguy C. Diversity of endosymbionts in the potato psyllid, *Bactericera cockerelli* (Triozidae), vector of zebra chip disease of potato. *Curr Microbiol.* 2011; 62: 1510–20. doi: [10.1007/s00284-011-9885-5](https://doi.org/10.1007/s00284-011-9885-5) PMID: [21327558](https://pubmed.ncbi.nlm.nih.gov/21327558/)
72. Colman DR, Toolson EC, Takacs-Vesbach CD. Do diet and taxonomy influence insect gut bacterial communities? *Mol Ecol.* 2012; 21: 5124–37. doi: [10.1111/j.1365-294X.2012.05752.x](https://doi.org/10.1111/j.1365-294X.2012.05752.x) PMID: [22978555](https://pubmed.ncbi.nlm.nih.gov/22978555/)
73. Bansal R, Mian MAR, Michel AP. Microbiome diversity of *Aphis glycines* with extensive superinfection in native and invasive populations. *Environ Microbiol Rep.* 2014; 6: 57–69. doi: [10.1111/1758-2229.12108](https://doi.org/10.1111/1758-2229.12108) PMID: [24596263](https://pubmed.ncbi.nlm.nih.gov/24596263/)
74. Sandström J, Pettersson J. Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (*Acyrtosiphon pisum*) performance. *J Insect Physiol.* 1994; 40: 947–55.
75. Koyama Y, Yao I, Akimoto S. Aphid galls accumulate high concentrations of amino acids: a support for the nutrition hypothesis for gall formation. *Entomol Exp Appl.* 2004; 113: 35–44.
76. Hodkinson ID. Life cycle variation and adaptation in jumping plant lice (Insecta: Hemiptera: Psylloidea): a global synthesis. *J Nat Hist.* 2009; 43: 65–179.
77. Carraro L, Osler R, Loi N, Ermacora P, Refatti E. Transmission of European stone fruit yellows phytoplasma by *Cacopsylla pruni*. *J Plant Pathol.* 1998; 80: 233–39.
78. Jensen DD, Erwin WR, Gonzales CQ, Griggs WH, Gotan SM, Westigard PH, et al. The relation of pear psylla to pear decline. . . greenhouse tests. *Calif Agr.* 1963; 17: 1–2.