Bacterial and Fungal Next Generation Sequencing Datasets and Metadata from Citrus Infected with ‘Candidatus Liberibacter asiaticus’

Nichole A. Ginnan,¹ Tyler Dang,¹ Sohrab Bodaghi,¹ Paul M. Ruegger,¹ Beth B. Peacock,¹ Greg McCollum,² Gary England,³ Georgios Vidalakis,¹,† Caroline Roper,¹,† Philippe Rolshausen,4,† and James Borneman,¹,†

¹ Department of Microbiology and Plant Pathology, University of California, Riverside 92521
² U.S. Department of Agriculture, Agricultural Research Service, U.S. Horticultural Research Laboratory, Fort Pierce, FL 34945
³ Hastings Agricultural Extension Center, University of Florida, Hastings 32145
⁴ Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

Abstract

Citrus production throughout the world is being severely threatened by Huanglongbing (HLB), which is a disease associated with the bacteria ‘Candidatus Liberibacter asiaticus’ (CLas), africanus, and americanus. This Resource Announcement provides amplicon-based next generation sequencing (NGS) datasets of the bacterial and fungal rRNA internal transcribed spacer (ITS) region from CLas-infected citrus budwood, leaves, and roots from five orchards located in different geographical regions in Florida (USA). To our knowledge, this is the first amplicon-based NGS study (i) that describes the fungal taxa associated with citrus and (ii) that provides comparative analyses of the bacterial and fungal taxa associated with budwood, leaves, and roots from the same citrus trees. This report also provides the sample metadata linked to these sequence datasets including HLB severity rating, tissue type, citrus rootstock, citrus scion, geographical region, and year trees were planted. When analyzed with other similar datasets, we anticipate that researchers will be able to obtain a greater understanding of the factors that shape the citrus microbiome as well as identify individual microorganisms or consortia of microorganisms that play a role in HLB suppression or exacerbation.

Background and Context for the Resource

Citrus is one of the highest valued fruit crops internationally, and it plays a vital role in the human diet and medicine. Citrus by-products also have considerable value because they possess antimicrobial and insecticidal properties (Nannapaneni et al. 2008; Oikeh et al. 2016; Rafiq et al. 2016; Siskos et al. 2008; Talon and Gmitter 2008).

First and second authors contributed equally to the work.

¹Corresponding authors: G. Vidalakis; E-mail: georgios.vidalakis@ucr.edu, C. Roper; E-mail: caroline.roper@ucr.edu, P. Rolshausen; E-mail: philippe.rolshausen@ucr.edu, and J. Borneman; E-mail: borneman@ucr.edu

Author Contributions: Author contributions were as follows: (i) C.R., G.M., G.V., J.B., and P.R. conceived of the work; (ii) G.E., G.M., N.A.G., P.R., S.B., and T.D. identified citrus orchards, collected samples, and/or processed samples; (iii) J.B., N.A.G., S.B., and T.D. performed experiments; (iv) J.B. and P.M.R. performed data analysis; (v) B.B.P., C.R., J.B., P.R., N.A.G., and T.D. prepared the manuscript; and (vi) B.B.P., C.R., G.M., G.V., J.B., N.A.G., P.M.R., P.R., S.B., and T.D. edited the manuscript.

*The e-Xtra logo stands for “electronic extra” and indicates that five supplementary figures and seven supplementary tables are published online.

Accepted for publication 23 March 2018.
Currently, citrus production throughout the world is severely threatened by Huanglongbing (HLB), which is a disease associated with three gram-negative, phloem-limited, alphaproteobacteria: ‘Candidatus Liberibacter asiaticus’ (CLas), africanus, and americanus (Wang et al. 2017). However, in the United States, only several different strains of CLas have been reported (Chen et al. 2010; Kunta et al. 2017; Zheng et al. 2017).

HLB is characterized by compromised nutrient transport, resulting in several distinct symptoms including yellow shoots, branch dieback, fruit remaining green, fruit size reduction, and ultimately tree death (Bové 2006; da Graça et al. 2016). In areas under high HLB pressure, tree decline can be swift with dramatic yield losses (Bové 2006; Wang and Trivedi 2013). HLB affects most citrus cultivars, and there currently are no effective prophylactic or therapeutic treatments (Bové 2006). HLB has resulted in the loss of both billions of dollars and thousands of jobs in Florida (USA) alone, and there is concern that its spread through other citrus growing regions could cause similar losses (da Graça et al. 2016; Hodges and Spreen 2012).

The plant microbiome is involved in various aspects of plant health and disease such as growth rate, vigor, immunity, infection, and disease protection (Klein et al. 2013; Schlaeppi and Bulgarelli 2015; Vogel et al. 2016). Understanding how the microbiome influences and interacts with the plant will require implementation of a variety of experimental approaches, including meta-analyses of large microbial datasets in relation to important variables associated with plant health, defense and disease.

The objective of this study was to provide unique next generation sequencing (NGS) datasets of the bacteria and fungi associated with CLas-infected citrus budwood, leaves and roots from five orchards located in different geographical regions in Florida. This report also provides the sample metadata linked to these datasets including HLB severity rating, tissue type, citrus rootstock, citrus scion, geographical region, and year trees were planted. In this paper, we define tissue type as either leaves, budwood, or roots.

Methods

Field sampling experimental design and methods. Ten citrus trees with varying levels of HLB symptoms distributed throughout each orchard were selected from five different locations (50 individual trees in total). The five orchards were located in two different geographic regions (northern and southern) in the state of Florida; see Supplementary Fig. S1. The northern Florida sampling locations included orchards in Weirsdale, Umatilla, and Howey-in-the-Hills, while the southern Florida sampling locations included Indiantown and Fort Pierce. Citrus budwood, leaves and roots were collected from all trees in March 2016. All trees in the study were determined to be CLas-positive by illumina bacterial rRNA ITS sequencing. Sample metadata including HLB severity rating, citrus rootstock, citrus scion, geographical region, and year trees were planted can be found in Supplementary Table S1. An explanation of the HLB severity rating scheme can be found in Supplementary Fig. S2.

Each tree was divided into four quadrants (north, south, east, and west). Two approximately 1-year-old budwood sticks ~20 to 25 cm long, with at least 10 mature and fully expanded leaves attached, were collected from each of the quadrants and pooled in large sealed plastic bags. Feeder roots were sampled from two sides of the tree; topsoil approximately 0.5 meter from the base of the trunk on two sides of the tree near the irrigation drip was removed and feeder roots were pulled from the ground, shaken to remove soil, and sealed in separate plastic bags. Between each sampled tree, the clippers and trowels were cleaned by spraying a 20% household bleach solution on them, and then wiping them clean with paper towels. Researchers also changed their gloves between each tree that was sampled. All samples were immediately placed in a cooler with ice, and then in the same day they were put in a laboratory refrigerator and stored at 4°C until the tissue was prepared for lyophilization, which was performed within 24 h of sample collection.

Samples were not surface-sterilized; thus, our microbial NGS datasets and analyses included both epiphytes and endophytes. Budwood and leaf samples were cut into 1 to 2 cm long pieces and approximately 6 g of leaves and 10 g of budwood were placed into 50-ml conical tubes and stored at ~80°C; root samples were rinsed with autoclaved purified water (Barnstead Mega-Pure System MP-6a, Thermo Fisher Scientific, Waltham, MA) and approximately 5 g of rinsed root tissue was placed into 50-ml conical tubes and stored at ~80°C (Supplementary Fig. S3 provides representative pictures of samples at this stage of the processing). Tissue samples were then lyophilized with a bench-top freeze dryer (Labconco FreeZone 4.5L, Kansas City, MO) for 16
to 20 h. The lyophilized tissues were stored at −80°C until they were shipped to UC Riverside (USDA permit P526P-16-00352) on dry ice for DNA isolation and microbiome analyses.

**DNA isolation.** Freeze-dried leaves and roots were crushed into small pieces (<0.5 cm) with sterile stainless-steel spatulas, and 100 mg of freeze-dried tissue was transferred to 2-ml microcentrifuge tubes (Eppendorf Safe-Lock tubes, Hamburg, Germany) containing a single 4 mm stainless steel grinding ball (SPEX SamplePrep, Metuchen, NJ). Approximately 200 mg of whole freeze-dried budwood was transferred to 5-ml frosted polyethylene vials containing a single 9.5-mm stainless steel grinding ball (SPEX SamplePrep). Samples were chilled at −80°C for 15 min and then pulverized to a powder using a 2010 Geno/Grinder (SPEX SamplePrep) at 1,680 rpm for 20 to 30 s, twice. One milliliter of 4 M guanidine thiocyanate was added to the pulverized leaf and root samples. Pulverized budwood was transferred from the 5-ml frosted polyethylene vials to 5 ml microcentrifuge tubes (Eppendorf Safe-Lock tubes) and then 2 ml of 4 M guanidine thiocyanate was added. All samples were incubated at 4°C for 15 min and subsequently centrifuged for 1 h at 17,500 × g. DNA was isolated using the MagMAX-96 DNA Multi-Sample Kit (Thermo Fisher Scientific) with the protocol “4413021ForPlants” on a MagMAX Express-96 Deep Well Magnetic Particle Processor; each reaction contained 300 μl of supernatant, 250 μl of 100% isopropanol, 300 μl of Multi-Sample DNA Lysis Buffer, and 20 μl of DNA Binding Bead mix. Samples were subjected to two successive washes with 150 μl of Wash Solution 1 followed by two successive washes with 150 μl of Wash Solution 2. The final DNA was eluted in 100 μl of DNA Elution Buffer. The concentration of the eluted DNA was assessed with an Infinite M1000 Pro (Tecan, Männedorf, Switzerland). The DNA was aliquoted into four 1.5-ml Eppendorf DNA LoBind Tubes with 25 μl in each tube and dried using a SpeedVac Concentrator (Thermo Fisher Scientific). The dried DNA was stored at −20°C until further use for bacterial and fungal Illumina library construction.

**NGS of the bacterial rRNA ITS region.** Illumina bacterial rRNA ITS libraries were constructed as follows. PCRs were performed in an MJ Research PTC-200 thermal cycler (Bio-Rad Inc., Hercules, CA) as 25-μl reactions containing: 50 mM Tris (pH 8.3), bovine serum albumin (BSA) at 500 μg/ml, 2.5 mM MgCl2, 250 μM of each deoxynucleotide triphosphate (dNTP), 400 nM of the forward PCR primer, 200 nM of each reverse PCR primer, 2.5 μl of DNA template, and 0.625 units JumpStart Taq DNA polymerase (Sigma-Aldrich, St. Louis, MO). PCR primers targeted a portion of the small-subunit (ITS-1507F, GGTGAAGTCGTAACAAGGTA) and large-subunit (ITS-23SR, GGTTBCCCCATTCRG) rRNA genes and the hypervariable ITS region (Ruegger et al. 2014), with the reverse primers including a 12-bp barcode and both primers including the Illumina sequences needed for cluster formation; primer binding sites are the reverse and complement of the commonly used small-subunit rRNA gene primer 1492R (Frank et al. 2008) and the large-subunit rRNA gene primer 129F (Hunt et al. 2006). PCR primers were only frozen and thawed once. Thermal cycling parameters were 94°C for 5 min; 35 cycles of 94°C for 20 s, 56°C for 20 s, and 72°C for 40 s; followed by 72°C for 10 min. PCR products were purified using a Qiagen QIAquick PCR Purification Kit (Qiagen, Valencia, CA) according to the manufacturer’s instructions.

DNA sequencing (single-end 150 base) was performed using an Illumina MiSeq (Illumina, Inc., San Diego, CA). Clusters were created using template concentrations 2.5 pM and PhiX at 65 K/mm². We used the UPARSE pipeline for de-multiplexing, length trimming, quality filtering, and operational taxonomic unit (OTU) picking using default parameters or recommended guidelines that were initially described in (Edgar 2013) and which have been updated at https://www.drive5.com/usearch/manual/uparse_pipeline.html. Briefly, after demultiplexing and using the recommended 1.0 expected error threshold, sequences were trimmed to a uniform length of 149 bp, which kept nearly 90% of reads. Sequences were then dereplicated and clustered using the UPARSE-OTU algorithm, which also detects and removes chimeric sequences. An OTU table was then generated using the otutab command. OTUs having nonbacterial DNA were identified and enumerated by performing a local BLAST search (Altschul et al. 1990) of their seed sequences against the nucleotide database. OTUs were removed if any of their highest scoring BLAST hits contained taxonomic IDs within the citrus family, fungal kingdom, or PhiX. Taxonomic assignments to bacterial OTUs were made by finding the lowest common taxonomic level of the highest BLAST hits excluding unclassified designations. Data were normalized within each sample by dividing the number of reads in each OTU by the total number of reads in that sample. The bacterial sequence mapping file with sample metadata and the OTU table can be found in Supplementary Tables S1 and S2, respectively.

**NGS of the fungal rRNA ITS region.** Illumina fungal rRNA ITS libraries were constructed as follows. PCRs were performed in an MJ Research PTC-200 thermal cycler (Bio-Rad Inc.) as
25-µL reactions containing: Phusion High-Fidelity DNA Polymerase Mix (New England Biolabs, Ipswich, MA) supplemented with BSA at 500 µg/ml, 1 mM MgCl₂, 250 µM of each deoxy-nucleotide triphosphate (dNTP), 400 nM of each primer, and 2.5 µl of DNA template. The PCR primers gITS7 (GTGARTCATCGARTCTTTG) and ITS4-1 (TCCTCCGCTTATTGATATGC) targeted the ITS2 region of the ribosomal rRNA gene operon (Ihrmark et al. 2012; White et al. 1990), with the reverse primers including 12-base barcodes, and both primers including the Illumina sequences needed for cluster formation. Thermal cycling parameters were 94°C for 5 min; 35 cycles of 94°C for 20 s, 56°C for 20 s, and 72°C for 30 s; followed by 72°C for 10 min. PCR products were purified using a Qiagen QIAquick PCR Purification Kit according to the manufacturer’s instructions.

DNA sequencing (single-end 300 base) was performed using an Illumina MiSeq. Clusters were created using template concentrations 2.5 pM and PhiX at 65 K/mm². We used the UPARSE pipeline for de-multiplexing, length trimming, quality filtering, and OTU picking using default parameters or recommended guidelines that were initially described in (Edgar 2013) and which have been updated at https://www.drive5.com/usearch/manual/uparse_pipeline.html. Briefly, after demultiplexing and using the recommended 1.0 expected error threshold, sequences were trimmed to a uniform length of 254 bp, which kept approximately 80% of reads. Sequences were then dereplicated and clustered into zero-radius OTUs using the UNOISE3 algorithm (Edgar 2016), which detects and removes chimeric sequences. An OTU table was then generated using the otutab command. Taxonomic assignments to the fungal OTUs were performed using the RDP Classifier version 2.12 (Wang et al. 2007), trained on the ver7_99_s_10.10.2017 release of the UNITE database (Kõljalg et al. 2013), and OTUs having nonfungal assignments were removed. Data were normalized within each sample by dividing the number of reads in each OTU by the total number of reads in that sample. The fungal sequence mapping file with sample metadata and the OTU table can be found in Supplementary Tables S1 and S3, respectively.

**Microbiome data analyses.** Taxonomy plots and data tables were generated using QIIME 1.9.1 (Caporaso et al. 2010) and analyzed using Prism (GraphPad, San Diego, CA). Hellinger beta diversity analyses were performed and statistically assessed by Adonis tests using QIIME.

**Results**

This report describes the bacteria and fungi associated with CLas-infected citrus from five orchards located in different geographical regions in Florida. To our knowledge, this is the first amplicon-based NGS study that (i) describes the fungal taxa associated with citrus and (ii) provides comparative analyses of the bacterial and fungal taxa associated with budwood, leaves, and roots from the same citrus trees. Because the citrus samples were not surface sterilized, our datasets and microbial analyses included both epiphytes and endophytes.

**Bacterial rRNA ITS analysis.** Besides the higher taxonomic resolution (Ruegger et al. 2014), another advantage of using this bacterial rRNA ITS analysis over bacterial 16S rRNA analyses is the relatively low number of mitochondrial and plastid sequences obtained when examining plant samples. For example, in this study, after low quality sequences had been removed but before nonbacterial sequences were removed, the average total number of sequences per sample was 87,506. Within these sequences, the average number of mitochondria, chloroplast, and other plastid sequences per sample (and their percentage of the average total number of sequences per sample) were 6.88 (0.0079%), 0.18 (0.0002%), and 14.42 (0.0165%), respectively. In addition, the range of the number of sequences per sample for mitochondria, chloroplasts and other plastids were 0 to 169, 0 to 9, and 0 to 1,209, respectively.

**Bacterial community analyses.** Beta diversity analyses were performed on the bacterial and fungal sequence data obtained from budwood, leaf and root samples from CLas-infected citrus trees (Supplementary Fig. S4). The most distinct differences in both the bacterial and fungal communities among the three tissue types were between the roots and the other two tissues (budwood or leaves) (Adonis tests, $P = 0.001$). Comparisons of the microbial communities associated with the budwood and leaves showed that the bacteria were different (Adonis test, $P = 0.001$) between these two tissues but that the fungi were not different (Adonis test, $P = 0.059$).

**Taxonomic analyses.** Analyses were also performed to describe and compare the most abundant bacterial and fungal phyla and genera associated with budwood, leaves, and roots from CLas-infected citrus trees (Supplementary Fig. S5).
The predominant bacterial phyla associated with all three tissues were the Proteobacteria, Actinobacteria, Bacteroidetes, and Firmicutes. However, the distributions were different among the three tissues with budwood and leaves having more Proteobacteria and Bacteroidetes than the roots (Supplementary Table S4). Conversely, the roots had considerably more Actinobacteria than either the budwood or leaves, which is consistent with the large amount of Streptomyces that was associated with the roots (described below).

For the fungi, the predominant phylum associated with all three tissues was the Ascomycota. Budwood and leaves also contained relatively large amounts of Basidiomycota compared with the roots (Supplementary Table S5). Conversely, roots contained a larger amount of the Glomeromycota than budwood or leaves, which is consistent with this phylum being comprised of members of the root-associated mycorrhizal fungi.

At the genus level, the budwood and leaves contained more of the HLB-associated Liberibacter than the roots (Supplementary Table S6). Other genera that were more abundant in or on the budwood and leaves than the roots were Methylbacterium, Hymenobacter, Pantoea, Curtobacterium, and Spirosoma. All of these genera except for Spirosoma have been previously identified from leaves or roots from CLas-infected or uninfected citrus (Blaustein et al. 2017; Trivedi et al. 2010, 2011; Zhang et al. 2013a, b). Spirosoma has been previously identified as an endophyte of the fruit of Palmer navel oranges (Gomba et al. 2017). Conversely, the roots contained a much larger amount of Streptomyces than budwood or leaves. Two other genera that were more abundant in or on the roots than budwood or leaves were Bacillus and Lactobacillus, and both these genera along with Streptomyces have been previously identified from roots or leaves from CLas-infected or uninfected citrus (Trivedi et al. 2010, 2011, 2012; Zhang et al. 2013b).

For the fungi, budwood and leaves had higher levels of the genera Cladosporium, Sporobolomyces, Symmetrospora, Camptophora, and Hannaella compared with the roots (Supplementary Table S7). In prior research, Cladosporium has been described as a leaf endophyte of citrus (Araújo et al. 2001), while Sporobolomyces has been associated with citrus leaves (Furuya et al. 2012). To our knowledge, this is the first report of Symmetrospora, Camptophora, and Hannaella inhabiting citrus. Conversely, the genera Exophiala, Fusarium, Glomus, Rhizophagus, and Acrocalymma were more abundant in or on the roots than budwood or leaves. The two mycorrhizal genera, Rhizophagus and Glomus, have been previously associated with citrus roots (Song et al. 2015; Wu et al. 2016). Fusarium has also been previously identified with citrus roots (Tatum et al. 1987). However, to our knowledge, this is the first report of Exophiala and Acrocalymma inhabiting citrus.

**Future Directions**

This Resource Announcement provides both unique amplicon-based NGS datasets of the bacteria and fungi associated with CLas-infected citrus budwood, leaves and roots along with the sample metadata including HLB severity rating, tissue type, citrus rootstock, citrus scion, geographical region, and year trees were planted. We anticipate that our datasets combined with a previously described amplicon-based NGS dataset of leaf and root bacteria from citrus possessing varying HLB symptoms (Blaustein et al. 2017) and a metagenomic NGS dataset from CLas-infected and uninfected citrus roots (Zhang et al. 2017) will provide the seeds for an important HLB resource and/or database that can be utilized in future HLB research. Such studies could involve obtaining a greater understanding of the factors that shape the citrus microbiome as well as identifying individual microorganisms or consortia of microorganisms that play a role in HLB suppression or exacerbation by performing meta-analyses of multiple NGS datasets that are linked to key metadata. We also anticipate that these datasets will aid in directing targeted experiments that aim to understand the in-planta microbe–microbe or microbiome–host interactions associated with HLB.

**Availability of Data and Materials**

The bacterial and fungal NGS datasets have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under SRA Identifier Number SRP127690 (accession numbers SRFX3520308 to SRFX3520607), Supplementary Tables S1, S2, and S3 contain the bacterial and fungal sequence mapping file with sample metadata, the bacterial OTU table, and the fungal OTU table, respectively. Supplementary Figures S1 and S3
contain a sampling map and a representative picture of budwood, leaf, and root samples, respectively.

Acknowledgments
We thank K. Kelly and N. Maharaj for assistance with sample collection and processing. Material presented here will be used in part to satisfy the requirements for the Ph.D. degree of T. Dang and N. A. Ginnan at the University of California, Riverside.

Literature Cited


