



Article

GreeningDB: A Database of Host–Pathogen Protein–Protein Interactions and Annotation Features of the Bacteria Causing Huanglongbing HLB Disease

Cristian D. Loaiza ¹, Naveen Duhan ¹ and Rakesh Kaundal ^{1,2,3,*}

¹ Department of Plants, Soils and Climate, Utah State University, Logan, UT 84322, USA; cdloaiza@aggiemail.usu.edu (C.D.L.); naveen.duhan@usu.edu (N.D.)

² Bioinformatics Facility, Center for Integrated BioSystems, Utah State University, Logan, UT 84322, USA

³ Department of Computer Science, Utah State University, Logan, UT 84322, USA

* Correspondence: rkaundal@usu.edu; Tel.: +1-(435)-797-4117; Fax: +1-(435)-797-2766

Citation: Loaiza, C.D.; Duhan, N.; Kaundal, R. GreeningDB: A Database of Host–Pathogen Protein–Protein Interactions and Annotation Features of the Bacteria Causing Huanglongbing HLB Disease. *Int. J. Mol. Sci.* **2021**, *22*, 10897. <https://doi.org/10.3390/ijms221910897>

Academic Editors: Prabhakaran Soundararajan and Abinaya Manivannan

Received: 26 August 2021

Accepted: 27 September 2021

Published: 8 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: The *Citrus* genus comprises some of the most important and commonly cultivated fruit plants. Within the last decade, citrus greening disease (also known as huanglongbing or HLB) has emerged as the biggest threat for the citrus industry. This disease does not have a cure yet and, thus, many efforts have been made to find a solution to this devastating condition. There are challenges in the generation of high-yield resistant cultivars, in part due to the limited and sparse knowledge about the mechanisms that are used by the *Liberibacter* bacteria to proliferate the infection in *Citrus* plants. Here, we present GreeningDB, a database implemented to provide the annotation of *Liberibacter* proteomes, as well as the host–pathogen comparatome tool, a novel platform to compare the predicted interactomes of two HLB host–pathogen systems. GreeningDB is built to deliver a user-friendly interface, including network visualization and links to other resources. We hope that by providing these characteristics, GreeningDB can become a central resource to retrieve HLB-related protein annotations, and thus, aid the community that is pursuing the development of molecular-based strategies to mitigate this disease's impact. The database is freely available at <http://bio-info.usu.edu/GreeningDB/>.

Keywords: Citrus greening disease; huanglongbing; citrus; database; protein annotations; host–pathogen interactions; protein–protein interactions

1. Introduction

Citrus greening disease (or huanglongbing) is the most devastating condition affecting citrus fields around the globe. Within the last few years, the citrus industry has been heavily affected due to the rapid spread of huanglongbing (HLB). In the US alone, the disease has caused a 21% decrease in the fresh citrus fruit market and around 71% in the production of oranges [1]. This infectious condition triggers when a Gram-negative bacterium *Candidatus Liberibacter* gets established into the phloem of a healthy citrus tree. *Ca. Liberibacter* bacteria are transmitted into citrus trees through the insect vector *Diaphorina citri* (also known as the Asian citrus psyllid), which carry the bacteria and release them into the trees during the feeding process. Infected trees typically have asymmetrical blotchy mottle, yellow shoots, and partially green, lopsided fruits [2]. Three species of *Liberibacter* genus have been identified as citrus pathogens, named according to the continent from where they were originally discovered: *Ca. Liberibacter asiaticus*, *Ca. Liberibacter americanus*, and *Ca. Liberibacter africanus*. Within this genus, there is another major pathogenic species (*Ca. Liberibacter solanacearum*) that although does not seem to be affecting citrus, is a causative agent of zebra chip disease [3]. Except for *L. crescens*, which is

a non-pathogenic strain [4], no other *Ca. Liberibacter* strains have been successfully cultured in artificial media, making the possibility to apply traditional molecular and genetic analysis difficult.

Notwithstanding, research groups have been able to elucidate some of the characteristics behind HLB infection. Now, it is well known that *Ca. Liberibacter* is an obligate parasite that lacks some of the housekeeping genes of regular bacteria, meaning that this bacterium is only able to multiply if it is inside a eukaryotic host, either a plant or a psyllid [5]. Furthermore, some molecular mechanisms used by HLB bacteria to overcome plant immune systems have been elucidated, which includes increasing the levels of salicylic acid in plants [6], *Liberibacter* prophages assisting in the suppressing of *Citrus* plant defenses [7], and the induction of the expression of immune response genes, such as those from the mitogen-activated protein kinase signaling pathway [8]. These processes often involve the prompt of protein–protein interactions (PPIs). PPIs constitute an essential part of all biological processes within the cell for all living beings. PPIs can be classified as intraspecies, within an organism; or interspecies, involving at least two organisms. Interspecies interactions are critical to understanding the molecular basis of pathogenesis between host plants and their pathogens [9,10].

Identifying the citrus proteins targeted by HLB bacteria is fundamental to advance into better management of this disease. There are few resources available to retrieve PPI data for many plants in terms of host–pathogen systems [11,12]; unfortunately for citrus, except for the study on the citrus targets of an SDE1 effector from *Ca. Liberibacter asiaticus* [13], there is no publicly accessible information of inter-species PPIs. Implementation of databases incorporating host–pathogen interaction features have been used to bolster knowledge about the disease, thus accelerating the development of novel drugs and therapeutics [14,15]. Furthermore, the development of genomic databases is crucial to the generation of advanced molecular approaches. For citrus genomics, there is a Citrus Genome Database (CGD) [16] available, a comprehensive resource with various modules that is helpful in enhancing studies on citrus breeding and genomics. Databases alike are also implemented in virulent bacteria [17,18]; however, no such tool exists for the *Liberibacter* genus.

To deliver a platform to facilitate the study of HLB bacteria, we have developed GreeningDB, a database implemented to store and offer the annotation of *Liberibacter* strains proteomes, and a novel host–pathogen ‘comparatomics’ tool that allows users to compare predicted host–pathogen interactomes of *Citrus* and HLB bacteria. GreeningDB is built to deliver a user-friendly interface, including network visualization and links to external annotation resources. This database can serve as a central resource for scientists, and the community in general, who are pursuing development of molecular-based strategies to mitigate the impact of citrus greening disease. GreeningDB is freely accessible through <http://bioinfo.usu.edu/GreeningDB>.

2. Results

2.1. GreeningDB Overview

The GreeningDB database comprises 10 citrus species and 25 *Ca. Liberibacter* strains. The database contains five tabs: Home, Tools, Features, Datasets, and Help. The “Home” page gives an overview of the database and the host–pathogen species that are present. The “Tools” page has three tools: interactome comparison, a sophisticated search module, and BLAST search. The “Features” tab connects to nine distinct features pages: host–pathogen interactions, protein annotations, secreted effectors, subcellular localization, gene ontology, functional domain mapping, Citrus genes altered by HLB infection, virulence effectors, and predicted effectors. The “Datasets” page contains genomic information for citrus species and strains of *Ca. Liberibacter*.

2.2. GreeningDB Tools

2.2.1. Novel Comparactomics Tool

The main feature of *GreeningDB* is the comparison of host–pathogen interactomes. The comparison tool functions in two ways: the user may compare a citrus host to two *Liberibacter* strains, or they can compare two citrus hosts to one *Liberibacter* strain. The interolog method implemented in this comparison tool is designed in such a manner that it can anticipate whole-genome protein–protein interactions in a matter of seconds. It also allows users to tailor their search by selecting protein–protein interaction databases and a configurable BLAST alignment search criteria. When a user starts a comparison, a unique job ID is generated to allow the user to track the progress. Users can also receive an e-mail upon completion of the job by providing an e-mail address in the given box. When the task is completed, the results are shown in a tabular format, allowing the user to sort the table by column or filter the data using keywords. The network visualization page includes links to the NCBI database for chosen node proteins, as well as descriptions and degrees. Users may also use the force atlas functionality to alter the network layout and output network as an SVG image or JSON format. The search results for each unique job ID are kept on the server for 30 days.

2.2.2. BLAST Search Tool

The standalone version of the NCBI BLAST is implemented on *GreeningDB* to provide the homology search functionality. All host–pathogen proteomes used to build *GreeningDB* are also available as BLAST databases with all *Citrus* susceptible to HLB, all *Citrus* tolerant to HLB, all *Ca. Liberibacter* infecting *Citrus*, or all *Ca. Liberibacter* infecting *Solanacearum*. A nucleotide or amino acid sequence may be uploaded by users and the system automatically determines the BLAST variant to use from BLASTp or BLASTx. The BLAST results page provides options to visualize alignments in a tabular or standard format. A more detailed version of the results is also implemented using BlasterJS [19].

2.2.3. Advanced Search Module

GreeningDB's advanced search module provides an interface via which users may search for information filtered by various keywords and other parameters, such as protein length, genomic range, and subcellular localization in a certain species, and download the results as a tab-delimited file. This search module is extensive, allowing searching for scholarly papers, protein annotations, GO keywords, experiment descriptions, and a variety of additional data that matches the searched keyword. Furthermore, a more basic option for doing a rapid search of a protein accession is accessible throughout *GreeningDB*, and both the “advanced” and “basic” searches will reveal all of the information that may be gathered from our database records.

2.3. Features in GreeningDB

Data acquired from the literature or results from the different annotation pipelines are given in various search modules; *GreeningDB* comprises a total of nine search modules. Several categories of *Liberibacter* protein annotations (protein annotations, secreted effectors, subcellular localization annotation, gene ontology (GO) term annotation, functional domain mappings (InterProScan), virulence factors, and predicted effectors (EffectiveDB)) can be retrieved from seven of those modules. In addition, two additional modules (citrus genes regulated by HLB infection, and host–pathogen interactions) are provided to collect annotations on citrus proteins discovered to be linked to HLB by literature mining. The module data will be shown in accordance with the dataset chosen. The dataset options within the *Liberibacter* annotation modules can also be set to specify which specific strain of *Liberibacter* (Table 1) to display, whereas in the other two modules, there is an option to specify from which host species to display annotation (*Citrus sinensis* or *Citrus clementina*)

and the two host species from which the majority of the manual annotations were retrieved.

Table 1. List of the *Liberibacter* proteomes that are integrated into GreeningDB. Five species within *Liberibacter* genus are included: *Ca. Liberibacter asiaticus*, *Ca. Liberibacter africanus*, *Ca. Liberibacter americanus*, *Ca. Liberibacter solanacearum*, and *L. crescens*. Sources of unpublished proteomes are provided in the form of NCBI BioProject identifiers.

Organism	Strain	Protein Sequences	Genome Size (M)	GC %	HLB Pathogen	Reference
<i>Candidatus Liberibacter asiaticus</i>	gxpsy	1075	1.27	36.6	Yes	[20]
	A4	1027	1.23	36.4	Yes	[21]
	psy62	1021	1.22	36.5	Yes	[22]
	JXGC	1040	1.22	36.4	Yes	BioProject: 376787
	Ishi-1	997	1.19	36.3	Yes	[23]
	AHCA1	1010	1.23	36.6	Yes	BioProject: 470611
	FL17	1029	1.28	36.5	Yes	BioProject: 269509
	YNJS7C	1073	1.26	36.6	Yes	BioProject: 488522
	YCPsy	1042	1.23	36.5	Yes	[24]
	SGCA5	1005	1.20	36.4	Yes	[25]
	TX2351	1026	1.25	36.5	Yes	BioProject: 361117
	HHCA	827	1.15	36.5	Yes	[26]
	SGCA1	499	0.23	36.3	Yes	BioProject: 470611
<i>Candidatus Liberibacter africanus</i>	PTSAPSY	1017	1.19	31.1	Yes	[27]
<i>Candidatus Liberibacter americanus</i>	Sao Paulo	945	1.18	31.1	Yes	[28]
	PWSP	925	1.20	31.1	Yes	[29]
<i>Candidatus Liberibacter solanacearum</i>	Clso-ZC1	1151	1.26	35.2	No	[30]
	LsoNZ1	1199	1.31	35.3	No	[31]
	FIN114	1124	1.25	35.2	No	[32]
	HenneA	1104	1.21	34.9	No	[31]
	RSTM	1182	1.29	35.1	No	[33]
	FIN111	1087	1.20	34.9	No	[32]
	ISR100	1210	1.30	35.0	No	BioProject: 427973
	R1	1112	1.20	35.3	No	[34]
<i>Liberibacter crescens</i>	BT-1	1235	1.50	35.4	No	[35]

3. Discussion

With a resource such as GreeningDB available, numerous activities may now be accomplished that were hard to complete previously. In the GreeningDB user-friendly interface, users can easily retrieve annotation of a *Ca. Liberibacter asiaticus* protein (Figure 1a); *C. sinensis* genes found regulated under HLB disease conditions (Figure 1b); secreted effectors of a *Ca. Liberibacter africanus* strain; obtaining proteins from a particular *Ca. Liberibacter asiaticus* strain that have “resistance” functional domains (Figure 1c); and extracellular proteins of a certain *Ca. Liberibacter asiaticus* strain, etc. By offering these functionalities inside GreeningDB, this database can become a primary resource for retrieving (and submitting) HLB-related protein annotations.

WP_012778388.1

Strain:
Ca. Liberibacter asiaticus psy62

Description:
NAD(P)-binding protein

Replicon:
chr

Replicon Accession:
[NC_012985.3](#)

Position in replicon (bp):
49722–50864

Gene name:
–

Locus:
CLIBASIA_RS00230

Length:
380

Old Locus:
CLIBASIA_00255

(a)

Subcellular localization:
Cytoplasmic

Interpro annotation:
[IPR036188](#): FAD/NAD(P)-binding domain superfamily
[IPR002938](#): FAD-binding domain

Functional domains:
PF01494: FAD binding domain
PR00420: Aromatic-ring hydroxylase (flavoprotein monooxygenase) signature

Gene ontology:

Salicylic acid (SA) hydroxylase. Salicylic acid (SA) and its derivatives play a central role in plant defense by mediating responses against pathogens in many plant species, and they are important for activating PTI, ETI, and SAR. ([28488467](#))

Ca. Liberibacter asiaticus TX2351

Show 25 entries

Accession	Localization	Score
WP_012778609.1	Extracellular	9.45
WP_015452479.1	Extracellular	9.64
WP_015452627.1	Extracellular	10
WP_015452633.1	Extracellular	9.71
WP_015452869.1	Extracellular	9.45
WP_015452985.1	Extracellular	9.65

Showing 1 to 6 of 6 entries (filtered from 1,001 total entries)

(b)

Search: extracellular

Ca. Liberibacter asiaticus gxpsy

Show 25 entries

Accession	Source	Domain	Description
WP_015452566.1	Pfam	PF02673	Bacitracin resistance protein BacA
WP_015452825.1	Pfam	PF12344	Ultra-violet resistance protein B
WP_015824951.1	TIGRFAM	TIGR00710	efflux_Bcr_CfIA: drug resistance transporter, Bcr/CfIA subfamily

Showing 1 to 3 of 3 entries (filtered from 7,803 total entries)

(c)

Search: resistance

Figure 1. Examples of the annotation features of the GreeningDB search modules. (a) Description of WP_012778388.1, a ‘*Ca. Liberibacter asiaticus*’ str. psy62 protein; (b) Proteins that are located extracellularly in ‘*Ca. Liberibacter asiaticus*’ str. TX2351; (c) Proteins found with a “resistance” functional domain in ‘*Ca. Liberibacter asiaticus*’ str. gxpsy.

3.1. Host–Pathogen Interactome Comparison: In the Direction of a Better Knowledge of HLB Infection Mechanisms

Computational approaches for predicting protein–protein interactions have grown in popularity in recent years because conventional methods such as yeast two-hybrid [36] and co-immunoprecipitation [37] are time-consuming and expensive when used on a large scale. Furthermore, the use of computational techniques to predict PPIs has been shown to occasionally be more accurate than traditional strategies [38]. It is critical to account for variations in host–pathogen interactomes because strain-specific PPI patterns may play a critical role in the creation of strain-optimized treatments.

Similarly, there are variations in HLB development among citrus species [39], distinctions that may be significant and worth identifying because they may point to genes that should be targeted from an HLB-tolerant variety to increase resistance in a susceptible one. Our innovative comparatomics module has been deployed to assist researchers in identifying PPI patterns that may be unique or common among the various HLB systems. To the best of our knowledge, this is the first report of a tool capable of performing this HPI comparison study, which we refer to as *comparatomics*.

This novel *comparatomics* tool offers two types of comparisons. A user can compare two hosts (*Citrus*) to a pathogen (*Liberibacter*) or two pathogens (*Liberibacter*) to a single host (*Citrus*). This tool supports the host and pathogen datasets specified in the “Data collection” section. In this tool, the user may also choose which protein–protein interaction database to utilize as a template in the prediction process, as well as configure BLASTp alignment filters to find homolog proteins. GreeningDB also includes a network visualization platform built using SigmaJS; this plugin was specifically chosen for its efficiency in showing huge networks. A user may visualize a collection of attributes for each node (species, description, degree) from the host–pathogen network visualization of a typical comparison result; moreover, a user can quickly locate hub nodes (nodes with a higher number of edges) as shown in Figure 2.

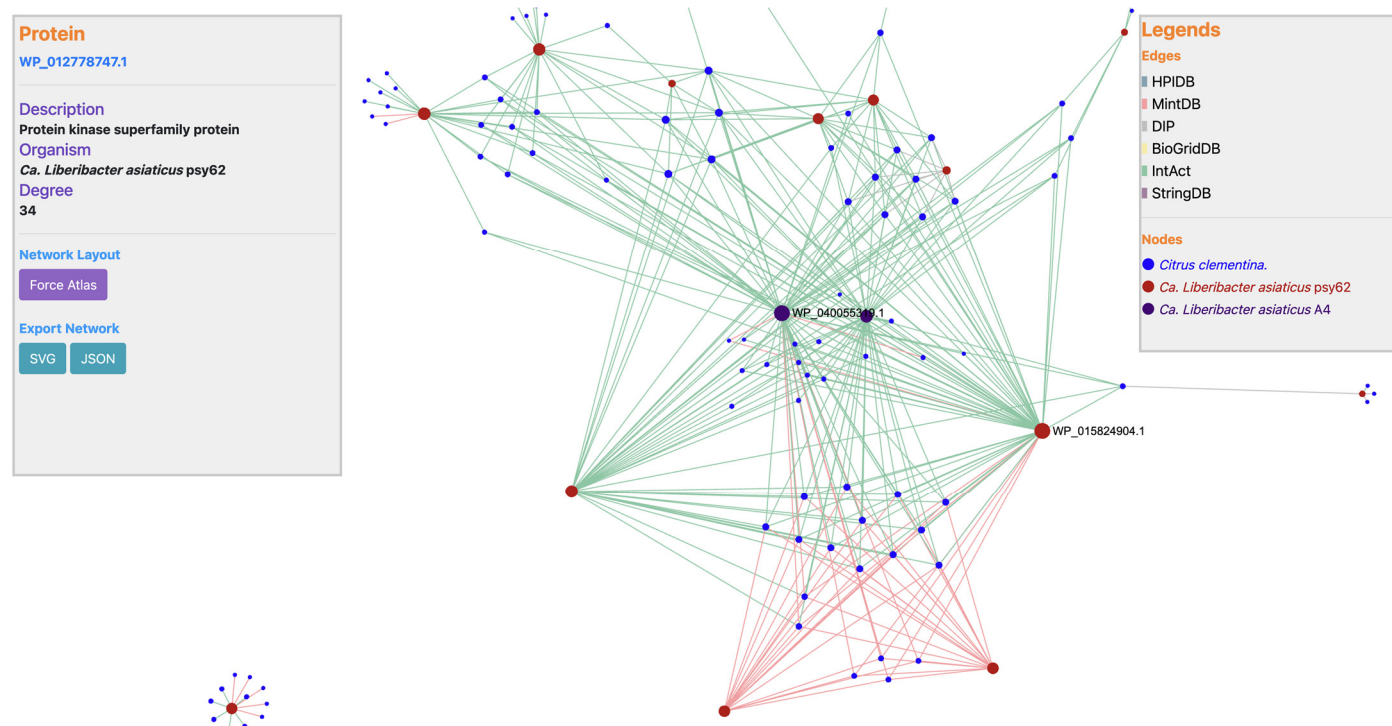


Figure 2. Host–pathogen interaction network visualization obtained through GreeningDB’s *comparatomics* module. *Citrus clementina* vs. ‘*Ca. Liberibacter asiaticus*’ str. psy62 interactome is compared against *Citrus clementina* vs. ‘*Ca. Liberibacter asiaticus*’ str. A4 interactome.

This is beneficial because hub nodes have been discovered to be critical in understanding several infectious disease pathways [40]. The network analysis offered by our database is not restricted to the user; the resulting network files may be downloaded and viewed in any third-party network analyzer program that can handle JSON or tabular network files.

In biological networks, it is typical to identify proteins that interact with several proteins at the same time, and knowing their role can help to obtain a better understanding of the pathogenicity processes employed by the *Liberibacter* bacteria to resist citrus defenses. The variability in HPI patterns among a set of *Liberibacter* strains can cause differences in virulence and infection processes; thus, we strongly believe that providing proteomes from commonly occurring strains, as well as strains that do not appear to infect Citrus (e.g., *L. crescens* and *Ca. Liberibacter solanacearum*), will improve our understanding of this disease.

3.2. GreeningDB's Future Development and Limitations

Because of existing limitations in genome assembly and annotation availability, some characteristics are not available to a few species in the database. Because studies on HLB molecular processes are often conducted using more refined genome assemblies, such as *C. clementina* and *C. sinensis*, there is little information available on Citrus genes regulated by HLB in other species. In terms of the pool of genomes available in *GreeningDB*, we expect to add more *L. crescens* strains, proteomes, and resistant Citrus datasets in the future, in addition to *P. trifoliata* and *C. excavata*. We were additionally constrained because we did not wish to include capabilities that were already available in other citrus genome resources, for example, the visualization of genomic variants or high-throughput sequencing experiments via JBrowse or citrus metabolic pathways using BioCyC. We will most likely collaborate with these citrus databases in the future or update the present version of those genomic tools in our database.

One of the most difficult challenges for a community-run database such as *GreeningDB* is its long-term viability, particularly in terms of content updates. To ensure that *GreeningDB* remains relevant to the community, we will update our database once a year, which includes upgrading the backend proteome files (both host and pathogen species) and then updating the backend HPI databases with the latest version (e.g., new version of HPIDB, STRINGdb, MINT, BioGRID, and other databases). Similarly, we will undertake an annual curation of HLB publication material in order to incorporate that data into our database records, as well as implement new features, such as HLB-related QTL and genetic markers in the second version of *GreeningDB*.

Depending on the influx of users and their specific requirements, the next version of the *GreeningDB* database might also include an API to access the database contents programmatically.

4. Materials and Methods

4.1. Data Collection

Citrus spp. and *Liberibacter* spp. proteomes were acquired from various sources. The *Citrus sinensis* Annotation Project (CAP) [41] was used to download the proteomes of *Fortunella hindsii* [42], *Citrus reticulata*, *Citrus grandis*, *Citrus ichangensis*, *Citrus medica*, and *Atalantia buxifolia* [43]. Protein sequences of *Citrus clementina* and *Citrus sinensis* [44] were obtained from Phytozome [45].

Citrus ecotypes exhibiting resistance to HLB infection include *Poncirus trifoliata* and *Clausena excavata* [46]; however, proteomes for these species were not available at the time this database was created. Nonetheless, we added protein sequences from both species to create a more comprehensive *GreeningDB*. TransDecoder [47] was used to transform the transcriptome of *Clausena excavata* [48] and the unigenes of *Poncirus trifoliata* [16] into protein sequences.

The proteomes of *Ca. Liberibacter* were acquired from the NCBI (<https://www.ncbi.nlm.nih.gov/>, accessed on November 11, 2018) (Table 1). This collection contains 13 *Ca. Liberibacter asiaticus* strains, 1 *Ca. Liberibacter africanus* strain, and 2 *Ca. Liberibacter americanus* strains. As a negative control, 8 *Ca. Liberibacter solanacearum* strains, as well as the *L. crescens* BT-1 proteome, are included in this resource. The proteome of the insect vector *Diaphorina citri* was also incorporated into *GreeningDB*, and psyllid protein sequences (*D. citri* OGS v2.0) were obtained from the Citrus Greening Solutions database [49].

4.2. Protein–Protein Interaction Database Collection

We utilized several PPI databases from The International Molecular Exchange Consortium (IMEx) [50] as templates to construct the interolog method in the comparatomics module. In the host–pathogen comparatomics module, *GreeningDB* uses the five major IMEX datasets: HPIDB v3.0 (2019) [12], an inter-species interaction database including four intra-species interaction datasets, IntAct (November 20, 2018) [51], DIP (November 20, 2018) [52], MINT (November 20, 2018) [53], and BioGRID v.3.5.169 [54]. STRINGdb [55] was also added to improve sensitivity in wide-range tests.

4.3. Comparatomics: A Host–Pathogen Interactome Comparison Tool

We developed a new module called the *comparatomics* tool to compare interactions between two distinct sets of HPIs. An interolog method was used within *GreeningDB* to achieve this on the backend. The interolog technique is based on the idea of transferring protein–protein interactions across comparable systems. For example, if A and A' are orthologs and B and B' are orthologs, then the interactions between A and B (in one system) and A' and B' (in another system) are interologs [56]. The ortholog proteins produced because of this approach are then used as a query to search the PPI databases. If a match in the PPI database matches to a host and pathogen protein, that protein pair is expected to interact. *GreeningDB's* backend uses this approach to anticipate host–pathogen PPIs between *Citrus* and *Liberibacter* proteins. In fact, each proteome was matched to each of the six PPI databases independently, yielding six alignment result files per proteome. BLASTp was run using the default settings to align the sequences. Following alignment, further steps for interolog prediction are performed using our in-house R scripts that call the SQL functions; all of this occurs when a user submits a task. SQL tables representing PPI databases were indexed for interactor A and interactor B columns to speed up the ortholog match searching process.

4.4. Pathogen Protein Annotation

The full protein descriptions of *Liberibacter* proteins were obtained from the NCBI Genome Assembly and Annotation Report's "Protein Details." Conserved domain areas and Gene Ontology (GO) terms were predicted using InterProScan [57] with the "iprscan" option in all bacteria proteomes for functional annotation. Using PSORTdb [58], we were also able to determine the subcellular localization of all of the bacterium proteins. In addition, we ran EffectiveDB [59] to predict effectors for all *Liberibacter* strains.

4.5. Secreted Effectors and HLB-Related Proteins from Literature

To refine the annotation data received from the online sites and tools, we examined the literature for information on proteins or genes associated with citrus greening disease. Much of the data gathered comes from a limited number of articles in particular. For example, in a thorough research to predict *Liberibacter* Sec-dependent extra cytoplasmic proteins [60], the authors predicted signal peptides from diverse *Liberibacter* strains by integrating the findings of four signal peptide prediction methods. Another paper that used the data employed a bioinformatics process to predict 28 potential effector proteins from

the *Ca. Liberibacter asiaticus* genome [61]. In addition to computational prediction analysis, we obtained annotation data from a transcriptome profile study that contrasted tolerant and susceptible plants; *Citrus clementina* genes that were differently expressed via treatments and associated with illness were mined [62]. We were able to retrieve more than a dozen articles over the previous five years because of this manual literature mining, demonstrating the relevance of HLB research.

4.6. GreeningDB Implementation and Architecture

The database *GreeningDB* is hosted at USU's Bioinformatics Facility using a Linux virtual machine within a high-performance computational cluster. The backend of the database was written in PHP, an open-source server-side scripting language, and implemented through an Apache server. Front-end visualization was written using HTML5, Bootstrap 4, and JavaScript.

For the backend, all the PPI databases listed in section “Protein–Protein interaction database collection” were downloaded and installed locally as separate MySQL tables. Similarly, each of the features described in the section “Features in *GreeningDB*” were implemented locally as independent MySQL tables. The interolog prediction backend was implemented by combining in-house PHP and R scripts with SQL databases. PPI networks were visualized using SigmaJS library dedicated to graph drawing (<http://sigmajs.org>, accessed on December 20, 2019). BLAST results were visualized using the BlasterJS plugin. The overall architecture of the database is depicted in Figure 3.

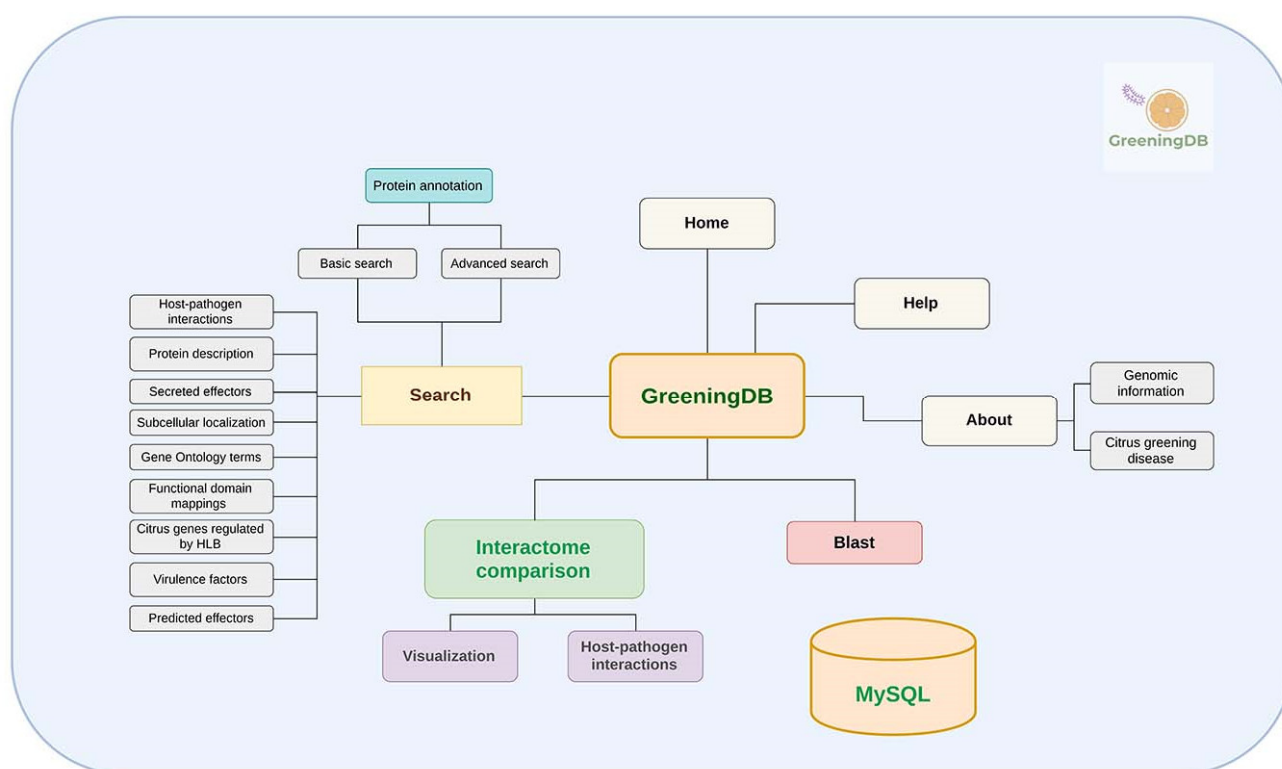


Figure 3. An overall view of the *GreeningDB* database architecture.

5. Conclusions

GreeningDB is a comprehensive database that offers the scientific community functional annotation and a complete collection of HPI characteristics for the majority of the sequenced *Liberibacter* proteomes. We have developed a unique host–pathogen “comparactomics” tool as part of this resource, which is a prediction platform that allows us to compare two HLB interactomes at the same time. We anticipate that *GreeningDB* will be a

valuable resource for the citrus breeding and experimental biology communities, bolstering and accelerating molecular research aimed at mitigating the Huanglongbing disease.

Author Contributions: R.K. formulated and designed the research. C.D.L. analyzed the data. C.D.L. and N.D. designed and constructed the web database. Writing—original draft preparation, C.D.L.; writing—review and editing, N.D. and R.K.; visualization, C.D.L., N.D., and R.K.; supervision, R.K.; project administration, R.K.; funding acquisition, R.K. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the support to this study from faculty start-up funds to R.K. from the Center for Integrated BioSystems / Department of Plants, Soils and Climate, USU. Partial support to C.L. from the USDA NIFA grant #2016-70016-24781 funded to R.K. is duly acknowledged. This research was also supported by the Utah Agricultural Experiment Station (UAES), USU, and approved as journal paper number 9523.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: The data can be accessed at <http://bioinfo.usu.edu/GreeningDB/>.

Acknowledgments: The authors sincerely thank Shelby McCowan (Linux Systems Administrator) who helped with the designing of the database and implementing the database successfully on an HPC cluster.

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

References

1. Dala-Paula, B.M.; Plotto, A.; Bai, J.; Manthey, J.A.; Baldwin, E.A.; Ferrarezi, R.S.; Gloria, M.B.A. Effect of Huanglongbing or Greening Disease on Orange Juice Quality, a Review. *Front. Plant. Sci.* **2019**, *9*, 1976, doi:10.3389/fpls.2018.01976.
2. Bové, J.M. Huanglongbing: A destructive, newly-emerging, century-old disease of citrus. *J. Plant. Pathol.* **2006**, *88*, 7–37.
3. Secor, G.A.; Rivera, V.V.; Abad, J.A.; Lee, I.-M.; Clover, G.R.G.; Liefting, L.W.; Li, X.; De Boer, S.H. Association of “*Candidatus Liberibacter solanacearum*” with Zebra Chip Disease of Potato Established by Graft and Psyllid Transmission, Electron Microscopy, and PCR. *Plant. Dis.* **2009**, *93*, 574–583, doi:10.1094/PDIS-93-6-0574.
4. Fagen, J.R.; Leonard, M.T.; Coyle, J.F.; McCullough, C.M.; Davis-Richardson, A.G.; Davis, M.J.; Triplett, E.W. *Liberibacter crescens* gen. nov., sp. nov., the first cultured member of the genus *Liberibacter*. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 2461–2466, doi:10.1099/ijs.0.063255-0.
5. Haapalainen, M. Biology and epidemics of *Candidatus Liberibacter* species, psyllid-transmitted plant-pathogenic bacteria. *Ann. Appl. Biol.* **2014**, *165*, 172–198, doi:10.1111/aab.12149.
6. Wang, N.; Pierson, E.A.; Setubal, J.C.; Xu, J.; Levy, J.G.; Zhang, Y.; Li, J.; Rangel, L.T.; Martins, J. The *Candidatus Liberibacter*–Host Interface: Insights into Pathogenesis Mechanisms and Disease Control. *Annu. Rev. Phytopathol.* **2017**, *55*, 451–482, doi:10.1146/annurev-phyto-080516-035513.
7. Jain, M.; Fleites, L.A.; Gabriel, D.W. Prophage-Encoded Peroxidase in ‘*Candidatus Liberibacter asiaticus*’ Is a Secreted Effector That Suppresses Plant Defenses’. *Mol. Plant.-Microbe Interact.* **2015**, *28*, 1330–1337, doi:10.1094/MPMI-07-15-0145-R.
8. Wang, N. The Citrus Huanglongbing Crisis and Potential Solutions. *Mol. Plant.* **2019**, *12*, 607–609, doi:10.1016/j.molp.2019.03.008.
9. Sahu, S.S.; Weirick, T.; Kaundal, R. Predicting genome-scale *Arabidopsis-Pseudomonas syringae* interactome using domain and interolog-based approaches. *BMC Bioinform.* **2014**, *15*, 1–8, doi:10.1186/1471-2105-15-S11-S13.
10. Ma, S.; Song, Q.; Tao, H.; Harrison, A.; Wang, S.; Liu, W.; Lin, S.; Zhang, Z.; Ai, Y.; He, H. Prediction of protein–protein interactions between fungus (*Magnaporthe grisea*) and rice (*Oryza sativa* L.). *Brief. Bioinform.* **2019**, *20*, 448–456, doi:10.1093/bib/bbx132.
11. Wattam, A.R.; Abraham, D.; Dalay, O.; Disz, T.L.; Driscoll, T.; Gabbard, J.L.; Gillespie, J.J.; Gough, R.; Hix, D.; Kenyon, R.; et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* **2014**, *42*, D581–D591, doi:10.1093/nar/gkt1099.
12. Ammari, M.G.; Gresham, C.R.; McCarthy, F.M.; Nanduri, B. HPIDB 2.0: A curated database for host-pathogen interactions. *Database* **2016**, *2016*, baw103, doi:10.1093/database/baw103.
13. Clark, K.; Franco, J.Y.; Schwizer, S.; Pang, Z.; Hawara, E.; Liebrand, T.W.H.; Pagliaccia, D.; Zeng, L.; Gurung, F.B.; Wang, P.; et al. An effector from the Huanglongbing-associated pathogen targets citrus proteases. *Nat. Commun.* **2018**, *9*, 1–11, doi:10.1038/s41467-018-04140-9.
14. Memišević, V.; Kumar, K.; Zavaljevski, N.; DeShazer, D.; Wallqvist, A.; Reifman, J. DBSecSys 2.0: A database of *Burkholderia mallei* and *Burkholderia pseudomallei* secretion systems. *BMC Bioinform.* **2016**, *17*, 387, doi:10.1186/s12859-016-1242-z.
15. Yue, J.; Zhang, D.; Ban, R.; Ma, X.; Chen, D.; Li, G.; Liu, J.; Wisniewski, M.; Droby, S.; Liu, Y. PCPPI: A comprehensive database for the prediction of *Penicillium*–crop protein–protein interactions. *Database* **2017**, *2017*, doi:10.1093/database/baw170.

16. Humann, J.; Frank, M.; Jung, S.; Cheng, C.-H.; Lee, T.; Hough, H.; Buble, K.; Zheng, P.; Scott, K.; Piaskowski, J.; et al. Citrus Genome Database Resources for HLB and Citrus Genomics, Genetics, and Breeding Research. In *Proceedings of the 6th International Research Conference on Huanglongbing*, March 10–15, 2019, Riverside, California, USA.
17. Winsor, G.L.; Khaira, B.; Van Rossum, T.; Lo, R.; Whiteside, M.D.; Brinkman, F.S.L. The Burkholderia Genome Database: Facilitating flexible queries and comparative analyses. *Bioinformatics* **2008**, *24*, 2803–2804, doi:10.1093/bioinformatics/btn524.
18. Winsor, G.L.; Lam, D.K.W.; Fleming, L.; Lo, R.; Whiteside, M.D.; Yu, N.Y.; Hancock, R.E.W.; Brinkman, F.S.L. Pseudomonas Genome Database: Improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Res.* **2011**, *39*, D596–D600, doi:10.1093/nar/gkq869.
19. Blanco-Míguez, A.; Fdez-Riverola, F.; Sánchez, B.; Lourenço, A. BlasterJS: A novel interactive JavaScript visualization component for BLAST alignment results. *PLoS ONE* **2018**, *13*, e0205286, doi:10.1371/journal.pone.0205286.
20. He, Y.; Wu, L.; Liu, X.; Zhang, X.; Jiang, P.; Ma, H. Yeast two-hybrid screening for proteins that interact with PFT in wheat. *Sci. Rep.* **2019**, *9*, 1–10, doi:10.1038/s41598-019-52030-x.
21. Free, R.B.; Hazelwood, L.A.; Sibley, D.R. Identifying Novel Protein-Protein Interactions Using Co-Immunoprecipitation and Mass Spectroscopy. *Curr. Protoc. Neurosci.* **2009**, *46*, 5.28.1–5.28.14, doi:10.1002/0471142301.NS0528S46.
22. Cong, Q.; Anishchenko, I.; Ovchinnikov, S.; Baker, D. Protein interaction networks revealed by proteome coevolution. *Science* **2019**, *365*, 185–189, doi:10.1126/science.aaw6718.
23. Folimonova, S.Y.; Robertson, C.J.; Garnsey, S.M.; Gowda, S.; Dawson, W.O. Examination of the Responses of Different Genotypes of Citrus to Huanglongbing (Citrus Greening) Under Different Conditions. *Phytopathology* **2009**, *99*, 1346–1354, doi:10.1094/PHYTO-99-12-1346.
24. Saha, S.; Sengupta, K.; Chatterjee, P.; Basu, S.; Nasipuri, M. Analysis of protein targets in pathogen–host interaction in infectious diseases: A case study on *Plasmodium falciparum* and *Homo sapiens* interaction network. *Brief. Funct. Genom.* **2017**, doi:10.1093/bfpg/elx024.
25. Wang, J.; Chen, D.; Lei, Y.; Chang, J.-W.; Hao, B.-H.; Xing, F.; Li, S.; Xu, Q.; Deng, X.-X.; Chen, L.-L. Citrus sinensis Annotation Project (CAP): A Comprehensive Database for Sweet Orange Genome. *PLoS ONE* **2014**, *9*, e87723, doi:10.1371/journal.pone.0087723.
26. Zhu, C.; Zheng, X.; Huang, Y.; Ye, J.; Chen, P.; Zhang, C.; Zhao, F.; Xie, Z.; Zhang, S.; Wang, N.; et al. Genome sequencing and CRISPR/Cas9 gene editing of an early flowering Mini-Citrus (*Fortunella hindsii*). *Plant. Biotechnol. J.* **2019**, pbi.13132, doi:10.1111/pbi.13132.
27. Wang, X.; Xu, Y.; Zhang, S.; Cao, L.; Huang, Y.; Cheng, J.; Wu, G.; Tian, S.; Chen, C.; Liu, Y.; et al. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat. Genet.* **2017**, *49*, 765–772, doi:10.1038/ng.3839.
28. Wu, G.A.; Prochnik, S.; Jenkins, J.; Salse, J.; Hellsten, U.; Murat, F.; Perrier, X.; Ruiz, M.; Scalabrin, S.; Terol, J.; et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat. Biotechnol.* **2014**, *32*, 656–662, doi:10.1038/nbt.2906.
29. Goodstein, D.M.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R.D.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.; et al. Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* **2012**, *40*, 1178–1186, doi:10.1093/nar/gkr944.
30. Ramadugu, C.; Keremane, M.L.; Halbert, S.E.; Duan, Y.P.; Roose, M.L.; Stover, E.; Lee, R.F. Long-Term Field Evaluation Reveals Huanglongbing Resistance in Citrus Relatives. *Plant. Dis.* **2016**, *100*, 1858–1869, doi:10.1094/PDIS-03-16-0271-RE.
31. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-Seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512, doi:10.1038/nprot.2013.084.
32. Bae, D.Y.; Eum, S.M.; Lee, S.W.; Paik, J.-H.; Kim, S.-Y.; Park, M.; Lee, C.; Tran, T.B.; Do, V.H.; Heo, J.-Y.; et al. Enrichment of genomic resources and identification of simple sequence repeats from medicinally important *Clausena excavata*. *3 Biotech.* **2018**, *8*, 133, doi:10.1007/s13205-018-1162-x.
33. Flores, M.; Saha, S.; Hosmani, P.S.; Brown, S.; Mueller, L.A. Community portal for the Citrus greening disease resources. *figshare* 2017; website online: https://figshare.com/articles/poster/Community_portal_for_the_Citrusgreening_disease_resources/5307391, accessed on March 23, 2020.
34. Orchard, S.; Kerrien, S.; Abbani, S.; Aranda, B.; Bhate, J.; Bidwell, S.; Bridge, A.; Briganti, L.; Brinkman, F.S.L.; Brinkman, F.; et al. Protein interaction data curation: The International Molecular Exchange (IMEx) consortium. *Nat. Methods* **2012**, *9*, 345–350, doi:10.1038/nmeth.1931.
35. Kerrien, S.; Aranda, B.; Breuza, L.; Bridge, A.; Broackes-Carter, F.; Chen, C.; Duesbury, M.; Dumousseau, M.; Feuermann, M.; Hinz, U.; et al. The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* **2012**, *40*, D841–D846, doi:10.1093/nar/gkr1088.
36. Xenarios, I.; Rice, D.W.; Salwinski, L.; Baron, M.K.; Marcotte, E.M.; Eisenberg, D. DIP: The Database of Interacting Proteins. *Nucleic Acids Res.* **2000**, *28*, 289–291, doi:10.1093/nar/28.1.289.
37. Licata, L.; Briganti, L.; Peluso, D.; Perfetto, L.; Iannuccelli, M.; Galeota, E.; Sacco, F.; Palma, A.; Nardozza, A.P.; Santonico, E.; et al. MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res.* **2012**, *40*, D857–D861, doi:10.1093/nar/gkr930.
38. Chatr-aryamontri, A.; Tyers, M.; Theesfeld, C.; Chang, C.; Rust, J.; Dolinski, K.; Oughtred, R.; Breitkreutz, B.-J.; Stark, C.; O'Donnell, L.; et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **2016**, *45*, D369–D379, doi:10.1093/nar/gkw1102.

39. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.; Roth, A.; Bork, P.; et al. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **2017**, *45*, D362–D368, doi:10.1093/nar/gkw937.
40. Matthews, L.R.; Vaglio, P.; Reboul, J.; Ge, H.; Davis, B.P.; Garrels, J.; Vincent, S.; Vidal, M. Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs”. *Genome Res.* **2001**, *11*, 2120–2126, doi:10.1101/gr.205301.
41. Mitchell, A.; Sangrador-Vegas, A.; Quinn, A.F.; McAnulla, C.; Binns, D.; Nuka, G.; McWilliam, H.; Chang, H.-Y.; Maslen, J.; Fraser, M.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240, doi:10.1093/bioinformatics/btu031.
42. Yu, N.Y.; Wagner, J.R.; Laird, M.R.; Melli, G.; Rey, S.; Lo, R.; Dao, P.; Sahinalp, S.C.; Ester, M.; Foster, L.J.; et al. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **2010**, *26*, 1608–1615, doi:10.1093/bioinformatics/btq249.
43. Eichinger, V.; Nussbaumer, T.; Platzer, A.; Jehl, M.-A.; Arnold, R.; Rattei, T. EffectiveDB—updates and novel features for a better annotation of bacterial secreted proteins and Type III, IV, VI secretion systems. *Nucleic Acids Res.* **2016**, *44*, D669–D674, doi:10.1093/nar/gkv1269.
44. Prasad, S.; Xu, J.; Zhang, Y.; Wang, N. SEC-translocon dependent extracytoplasmic proteins of *Candidatus Liberibacter asiaticus*. *Front. Microbiol.* **2016**, *7*, 1–9, doi:10.3389/fmicb.2016.01989.
45. Shi, Q.; Pitino, M.; Zhang, S.; Krystel, J.; Cano, L.M.; Shatters, R.G.; Hall, D.G.; Stover, E. Temporal and spatial detection of *Candidatus Liberibacter asiaticus* putative effector transcripts during interaction with Huanglongbing-susceptible, -tolerant, and -resistant citrus hosts. *BMC Plant. Biol.* **2019**, *19*, 1–12, doi:10.1186/s12870-019-1703-4.
46. Wang, Y.; Zhou, L.; Yu, X.; Stover, E.; Luo, F.; Duan, Y. Transcriptome profiling of Huanglongbing (HLB) tolerant and susceptible citrus plants reveals the role of basal resistance in HLB tolerance. *Front. Plant. Sci.* **2016**, *7*, 933.
47. Lin, H.; Han, C.S.; Liu, B.; Lou, B.; Bai, X.; Deng, C.; Civerolo, E.L.; Gupta, G. Complete Genome Sequence of a Chinese Strain of “*Candidatus Liberibacter asiaticus*.” *Genome Announc.* **2013**, *1*, doi:10.1128/genomeA.00184-13.
48. Zheng, Z.; Deng, X.; Chen, J. Whole-Genome Sequence of “*Candidatus Liberibacter asiaticus*” from Guangdong, China. *Genome Announc.* **2014**, *2*, doi:10.1128/genomeA.00273-14.
49. Zhou, L.; Powell, C.A.; Hoffman, M.T.; Li, W.; Fan, G.; Liu, B.; Lin, H.; Duan, Y. Diversity and Plasticity of the Intracellular Plant Pathogen and Insect Symbiont “*Candidatus Liberibacter asiaticus*” as Revealed by Hypervariable Prophage Genes with Intragenic Tandem Repeats. *Appl. Environ. Microbiol.* **2011**, *77*, 6663–6673, doi:10.1128/AEM.05111-11.
50. Katoh, H.; Miyata, S.; Inoue, H.; Iwanami, T. Unique Features of a Japanese ‘*Candidatus Liberibacter asiaticus*’ Strain Revealed by Whole Genome Sequencing.’ *PLoS One* **2014**, *9*, e106109, doi:10.1371/journal.pone.0106109.
51. Wu, F.; Zheng, Z.; Deng, X.; Cen, Y.; Liang, G.; Chen, J. Draft Genome Sequence of “*Candidatus Liberibacter asiaticus*” from Diaphorina citri in Guangdong, China. *Genome Announc.* **2015**, *3*, doi:10.1128/genomeA.01316-15.
52. Wu, F.; Kumagai, L.; Liang, G.; Deng, X.; Zheng, Z.; Keremane, M.; Chen, J. Draft Genome Sequence of “*Candidatus Liberibacter asiaticus*” from a Citrus Tree in San Gabriel, California. *Genome Announc.* **2015**, *3*, doi:10.1128/genomeA.01508-15.
53. Zheng, Z.; Deng, X.; Chen, J. Draft Genome Sequence of “*Candidatus Liberibacter asiaticus*” from California. *Genome Announc.* **2014**, *2*, doi:10.1128/genomeA.00999-14.
54. Lin, H.; Pietersen, G.; Han, C.; Read, D.A.; Lou, B.; Gupta, G.; Civerolo, E.L. Complete Genome Sequence of “*Candidatus Liberibacter africanus*,” a Bacterium Associated with Citrus Huanglongbing. *Genome Announc.* **2015**, *3*, doi:10.1128/genomeA.00733-15.
55. Wulff, N.A.; Zhang, S.; Setubal, J.C.; Almeida, N.F.; Martins, E.C.; Harakava, R.; Kumar, D.; Rangel, L.T.; Foissac, X.; Bové, J.M.; et al. The Complete Genome Sequence of ‘*Candidatus Liberibacter americanus*’, Associated with Citrus Huanglongbing.’ *Mol. Plant-Microbe Interact.* **2014**, *27*, 163–176, doi:10.1094/MPMI-09-13-0292-R.
56. Lin, H.; Coletta-Filho, H.D.; Han, C.S.; Lou, B.; Civerolo, E.L.; Machado, M.A.; Gupta, G. Draft Genome Sequence of “*Candidatus Liberibacter americanus*” Bacterium Associated with Citrus Huanglongbing in Brazil. *Genome Announc.* **2013**, *1*, doi:10.1128/genomeA.00275-13.
57. Lin, H.; Lou, B.; Glynn, J.M.; Doddapaneni, H.; Civerolo, E.L.; Chen, C.; Duan, Y.; Zhou, L.; Vahling, C.M. The Complete Genome Sequence of ‘*Candidatus Liberibacter solanacearum*’, the Bacterium Associated with Potato Zebra Chip Disease. *PLoS One* **2011**, *6*, e19135, doi:10.1371/JOURNAL.PONE.0019135.
58. Thompson, S.M.; Johnson, C.P.; Lu, A.Y.; Frampton, R.A.; Sullivan, K.L.; Fiers, M.W.E.J.; Crowhurst, R.N.; Pitman, A.R.; Scott, I.A.W.; Wen, A.; et al. Genomes of ‘*Candidatus Liberibacter solanacearum*’ Haplotype A from New Zealand and the United States Suggest Significant Genome Plasticity in the Species.’ *Phytopathology* **2015**, *105*, 863–871, doi:10.1094/PHYTO-12-14-0363-FI.
59. Wang, J.; Haapalainen, M.; Schott, T.; Thompson, S.M.; Smith, G.R.; Nissinen, A.I.; Pirhonen, M. Genomic sequence of “*Candidatus Liberibacter solanacearum*” haplotype C and its comparison with haplotype A and B genomes. *PLoS One* **2017**, *12*, e0171531, doi:10.1371/journal.pone.0171531.
60. Wu, F.; Deng, X.; Liang, G.; Wallis, C.; Trumble, J.T.; Prager, S.; Chen, J. De Novo Genome Sequence of “*Candidatus Liberibacter solanacearum*” from a Single Potato Psyllid in California. *Genome Announc.* **2015**, *3*, doi:10.1128/genomeA.01500-15.

-
61. Zheng, Z.; Clark, N.; Keremane, M.; Lee, R.; Wallis, C.; Deng, X.; Chen, J. Whole-Genome Sequence of “*Candidatus Liberibacter solanacearum*” Strain R1 from California. *Genome Announc.* **2014**, *2*, doi:10.1128/genomeA.01353-14.
 62. Leonard, M.T.; Fagen, J.R.; Davis-Richardson, A.G.; Davis, M.J.; Triplett, E.W. Complete genome sequence of *Liberibacter crescens* BT-1. *Stand. Genomic Sci.* **2012**, *7*, 271–283, doi:10.4056/sigs.3326772.