Guangdong and Florida Populations of ‘Candidatus Liberibacter asiaticus’ Distinguished by a Genomic Locus with Short Tandem Repeats


First and sixth authors: U.S. Department of Agriculture–Agricultural Research Service, San Joaquin Valley Agricultural Sciences Center, 9611 S. Riverbend Ave., Parlier, CA 93648; second author: Laboratory of Citrus Huanglongbing Research, South China Agricultural University, Guangzhou, Guangdong, PR China; third and fourth authors: Florida Department of Agriculture and Consumer Services–Division of Plant Industry, Gainesville; and fifth author: United States Sugar Corporation/Southern Gardens Citrus, Clewiston, FL. Accepted for publication 22 February 2010.

ABSTRACT


Huanglongbing (HLB) (yellow shoot disease) is a highly destructive disease that threatens citrus production worldwide. The disease was first observed in Guangdong, P.R. China, over 100 years ago, and was found in Florida, United States, in 2005. ‘Candidatus Liberibacter asiaticus’ has been associated with HLB in many citrus-growing regions around the world, including Guangdong and Florida. The global epidemiology of HLB, as well as management of the disease, relies on knowledge of ‘Ca. L. asiaticus’ populations in different geographical regions around the world. In this study, we identified a genetic marker containing small tandem repeats in the genome of ‘Ca. L. asiaticus’ and comparatively analyzed the tandem repeat numbers (TRNs) in ‘Ca. L. asiaticus’ populations from Guangdong and Florida. Analyses of TRNs showed that the bacterial population in Guangdong was different from that in Florida. The Guangdong population consisted predominately of strains with a TRN of 5 (TRN5) at a frequency of 47.6%. The Florida population consisted predominately of strains with a TRN of 9 (TRN9) at a frequency of 84.4%. TRNs ranged from 3 to 16. The apparent absence of TRNs of 9, 10, 11, and 12 separated the bacterial strains into two groups: TRNs < 10 (TRN<10) and TRNs > 10 (TRN>10). In Florida, TRN>10 strains (103/109, or 94.5%) were widely distributed in all HLB-affected counties. TRN<10 strains (6/109, or 5.5%) were found in central Florida. This is the first report documenting the differentiation of ‘Ca. L. asiaticus’ populations between Asia and North America and the possible presence of two differentially distributed genotypes of ‘Ca. L. asiaticus’ in Florida.

Additional keywords: genomic variation.

Huanglongbing (HLB) (yellow shoot disease) is a highly destructive disease that threatens citrus production worldwide (4). The disease was first noted in Guangdong Province of the People’s Republic of China in the late 1800s (16) and is endemic there. HLB was recently reported in Florida (12) and is of great concern to the U.S. citrus industry. An α-proteobacterium, ‘Candidatus Liberibacter asiaticus’, was first identified from an HLB-affected tree from India (15). Since then, the bacterium has been associated with HLB in many parts of the world, including Guangdong (9,10,19) and Florida (12). The diversity of ‘Ca. L. asiaticus’ strains associated with HLB among different geographical regions around the world is of global epidemiological significance for disease management. Knowledge of the bacterial population diversity can facilitate the development of effective, economical, and sustainable disease-management strategies to mitigate losses due to HLB (4,7). Currently, very limited biological and genomic data are available regarding the diversity of ‘Ca. L. asiaticus’ populations among different geographical origins. Taxonomically, ‘Ca. L. asiaticus’ is defined by a signature sequence in the 16S rRNA gene (rrs) (15). Strain variation of ‘Ca. L. asiaticus’ in this genomic locus is very limited and the locus is not suitable for strain differentiation (1,7,8). Other genomic loci that have been used for bacterial identification include genes in the β operon of ribosomal RNA proteins (rpl) (13) and an outer-membrane protein gene (omp) (1). Similar to the rrs locus, sequences of rpl are highly conserved and reveal little strain diversity (7). Based on the polymerase chain reaction (PCR) restriction fragment length polymorphisms (RFLPs) of theomp locus, Bastianel et al. (1) differentiated strains of ‘Ca. L. asiaticus’ from 10 different countries, although each region was represented by only one to two strains. Deng et al. (7) recently performed a single-nucleotide polymorphism (SNP) analysis in theomp locus and found that the China-Guangdong strains of ‘Ca. L. asiaticus’ were grouped with bacterial strains from Thailand and Nepal but not from the Philippines and China-Behai. Tomimura et al. (20) analyzed the bacteriophage-type DNA polymerase sequences and revealed three clusters in the Southeast Asian ‘Ca. L. asiaticus’ population; one cluster correlated to Indonesian strains and the other clusters were not correlated with geographic distribution.

To date, we are not aware of any comparative genetic evaluations of Chinese and U.S. populations of ‘Ca. L. asiaticus’. One important issue is the need for appropriate molecular markers. A shotgun whole-genome sequence of ‘Ca. L. asiaticus’ was recently reported (11). This provides accesses to more genomic loci for characterization of the bacterial population diversity. In the case of Xylella fastidiosa, a pathogen causing Pierce’s disease (PD) of grape and almond leaf scorch disease, analyses of whole-
genomic sequences led to the identification of a locus, PD0218
(psPB), a putative serine protease containing tandem repeats of
ACDCCA. The tandem repeat numbers (TRNs) in this genomic
locus were highly variable, allowing the separation of highly
similar G-genotype strains from grape and almond as defined by
Chen et al. (5). In this study, we applied a similar strategy to
characterize the variations in ‘Ca. L. asiaticus’ strains from
Guangdong, China and Florida, United States. We first identified
a locus containing short tandem repeats in the genome of ‘Ca. L.
asiaticus’. Analyses of TRNs in this locus showed that the bac-
terial population in Guangdong was significantly different from
that in Florida. The implications of TRN variation in the biology
and epidemiology and control of HLB are discussed.

MATERIALS AND METHODS

Bacterial sources. In this study, a strain of ‘Ca. L. asiaticus’
was defined and represented by bacterial DNA extracted from
symptomatic leaves. HLB-symptomatic citrus leaf samples were
collected by three laboratories: (i) Guangdong, P.R. China (South
China Agricultural University [SCAU]), (ii) Florida Department
of Agriculture and Consumer Services–Division of Plant Industry,
(FDACS-DPI), Gainesville; and (iii) U.S. Sugar Corporation/
Southern Gardens Citrus, (USCSSGC), Clewiston, FL. The
SCAU samples were collected from six groves in four prefecture
cities of Guangdong Province in 2007. Among the six locations,
three were in the western region, currently the major citrus
production area in the province, and three from the coastal central
and east regions, where HLB is historically endemic. citrus hosts
included Citrus grandis, C. sinensis, C. reticulata, C. sp., and
Poncirus trifoliata. In Florida, FDACS-DPI samples were col-
lected from 29 counties covering most of the citrus-production
regions between March 2008 and February 2009. The hosts
included C. aurantifolia, C. aurantium, C. limon, C. sinensis,
Citrus sp., Citrus × nobilis, Citrus × paradisi, and Citrus ×
tangelo. USCSSGC samples were collected from C. sinensis
in seven southern counties between July 2007 and January 2008.

DNA extraction. In Guangdong, midribs of citrus leaves were
excised and DNA was extracted using the cetyltrimethyl-
ammonium bromide method as previously described (7). FDACS-
DPI samples were extracted using 80 to 100 mg of midrib and
petiole cuttings, the BioSpec Mini-Beadbeater (Biospec, Bartles-
ville, OK), and the Plant DNeasy extraction system (Qiagen,
Valencia, CA). USCSSGC samples were extracted following the
procedure described previously except that the DNA pellet was
dissolved in 500 μl of distilled water (14). ‘Ca. L. asiaticus’ was
identified by PCR amplification with primer sets OI1/OI2c (15)
and ITSaf/ITSAr (8). Microbial-free DNA preparations were sent
to the research laboratory located at San Joaquin Valley Agri-
cultural Sciences Center, United States Department of Agricul-
ture–Agricultural Research Service, Parlier, CA, for analyses.

Primers and PCR procedure for bacterial population
analyses. The sequence of ‘Ca. L. asiaticus’ strain psy62
(NC_012985.2) was obtained from the GenBank database at the
National Center for Biotechnology Information (http://www.ncbi.
nlm.nih.gov). The locus of CLIBASIA_01645 was selected due to
the presence of tandem repeats. PCR primers flanking the repeat
region of CLIBASIA_01645 were designed with Primer 3
software (18). Primers LapGP-1f (5'-GACATTTCACG
GTATCGAC-3') and LapGP-1r (5'-GGCGACATACTCTACT
CTT-3') were obtained and their specificity to ‘Ca. L. asiaticus’
was checked by in silico analysis through BLASTn search against
the GenBank database that included many known citrus patho-
genotypes. The primer set was used for DNA amplification following
the procedures described previously (6). Briefly, reactions were
carried out with a TaKaRa Hot Start Taq system (Clontech
Laboratories, Inc., Madison, WI) in 25-μl mixtures containing
2.5 μl of 10× DNA polymerase buffer, 2.5 μl of 2.5 mM
deoxyribonucleotide triphosphates (dNTPs), 1 μl of template DNA,
0.2 μl of Taq DNA polymerase at 5 U/μl, and 18.3 μl of H2O. An
MJ Research DNA Engine Tetrad 2 was programmed with an
initial denaturing step of 96°C for 1 min and 35 cycles of the
following conditions: 96°C for 30 s, 55°C for 30 s, and 72°C for
30 s. A final extension of 72°C for 4 min was done after the last
cycle. PCR products were electrophoresed in a 1.5% agarose gel
and visualized by ethidium bromide staining under UV light.

DNA sequencing and analysis. PCR amplicons were
sequenced directly in both orientations using BigDye Terminator
v3.1 Cycle Sequencing Kit in a 3130xl Genetic Analyzer
(Applied Biosystems, Inc.) with the same primers used for PCR.
Tandem repeats were counted manually. Frequencies of different
TRN genotypes were calculated. Bacterial population difference
was analyzed by a χ2 test and Wilcoxon signed-rank test.

RESULTS

In total, 174 ‘Ca. L. asiaticus’ strains were analyzed, 65 from
Guangdong (5 C. grandis, 19 C. sinensis, 33 C. reticulata, 2
Citrus sp., and 6 P. trifoliata) and 109 from Florida. Among
the Florida strains, 54 (1 C. aurantifolia, 1 C. aurantium, 1 C. limon,
30 C. sinensis, 13 Citrus sp., 1 Citrus × nobilis, 5 Citrus ×
paradisi, and 2 Citrus × tangelo) were collected from 29 counties,
including south, central, and north Florida by FDACS-DPI and 55
(C. sinensis) were from 7 counties in southern Florida by
USCSSGC. Primer BLASTn search identified only to sequence of
CLIBASIA_01645 of ‘Ca. L. asiaticus’. All PCR-positive
samples with primer set LapGP-1f/LapGP-1r were also positive
with primer sets ITSaf/ITSAr and OI1/OI2c (data not shown).
Weak amplification with ‘Ca. L. asiaticus’-specific primers was
also correlated with weak amplification using primer set LapGP-
1f/LapGP-1r (Fig. 1), evidence that primer set LapGP-1f/LapGP-
1r specifically amplified DNA from ‘Ca. L. asiaticus’. In most
cases, primer set LapGP-1f/LapGP-1r generated similar ampli-
cions of approximately 700 bp. In addition, larger amplicons
were also seen (Fig. 1).

Variations of the LapGP-1f/LapGP-1r amplicons were further
demonstrated by nucleotide sequence analyses. The
CLIBASIA_01645 locus of ‘Ca. L. asiaticus’ strain psy62 con-
tained five tandem repeats of AGACACA. Examination of the 174
strains showed that TRNs at this locus varied from 3 to 16 (Fig. 2;
Table 1). Among the Florida strains, the TRN of 5 (TRN₅) predominated with a frequency of 84.4%, followed by the TRN of 4 (TRN₄) with a much smaller frequency of 4.5%. In contrast, TRN of 7 (TRN₇) was the most frequent (47.6%) among the Guangdong strains, followed by TRN of 6 (TRN₆), with a frequency of 29.2%. Due to the presence of some small expected values in the $\chi^2$ test, data in Table 1 were regrouped into three categories: TRN $\leq$ 5, TRN = 6, and TRN $\geq$ 7. Test results showed that the within-population distribution frequencies of TRNs in the Guangdong and Florida populations were significantly different ($P < 0.001$). Results of the Wilcoxon signed-rank test with all nine TRN categories (Table 1) also showed a significant difference ($P < 0.001$) between the two populations. Distribution of TRNs among ‘Ca. L. asiaticus’ in different citrus hosts are shown in Figure 2. Sweet orange (C. sinensis) strains with TRN₅ and TRN₇ were used as an example. Among the 85 Florida sweet orange strains, 73 (85.6%) were TRN₅ and none (0%) was TRN₇. Among the 19 Guangdong sweet orange strains, 5 (26.3%) were TRN₅ and 7 (36.8%) were TRN₇. Apparently, sweet orange did not particularly favor any of the TRN variants identified in this study.

The absence of TRN 9, 10, 11, and 12 genotypes naturally separated the ‘Ca. L. asiaticus’ strains from both locations into two groups (Table 1): a large TRN group (TRN > 10 [TRN>10]) and a small TRN group (TRN < 10 [TRN<10]). In the Guangdong population, two (3.0%) TRN>10 strains were detected, whereas 63 strains were TRN<10 (mean = 6.4 ± 0.9). In all, 6 (11.1%) TRN>10 strains were present in the FDACS-DPI samples, all of which

Fig. 2. Graphical illustration of distribution of tandem repeat numbers (TRNs) in ‘Candidatus Liberibacter asiaticus’ strains from different citrus hosts in Guangdong, China, and Florida, United States. USSCSGC: U.S. Sugar Corporation/Southern Gardens Citrus, Clewiston, FL; FDACS-DPI: Florida Department of Agriculture and Consumer Services–Division of Plant Industry, Gainesville.

<table>
<thead>
<tr>
<th>Location</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>13</th>
<th>15</th>
<th>16</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guangdong</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boluo-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>9</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Boluo-2</td>
<td></td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>8</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>Guangzhou</td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Luoding</td>
<td></td>
<td></td>
<td>2</td>
<td>8</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Sihui-1</td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Sihui-2</td>
<td></td>
<td></td>
<td>2</td>
<td>8</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1</td>
<td>9</td>
<td>19</td>
<td>31</td>
<td>3</td>
<td>2</td>
<td></td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td>1.5</td>
<td>13.8</td>
<td>29.2</td>
<td>47.6</td>
<td>4.6</td>
<td></td>
<td></td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td></td>
<td>3</td>
<td>50</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>USSCSGCb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>54</td>
</tr>
<tr>
<td>FDACSb</td>
<td>2</td>
<td>2</td>
<td>42</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>109</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>5</td>
<td>92</td>
<td>42</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>1.8</td>
<td>4.5</td>
<td>84.4</td>
<td>3.6</td>
<td></td>
<td>1.8</td>
<td>2.7</td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a Numbers of total strains and frequencies are highlighted in bold.
*b USSCSGC = U.S. Sugar Corporation/Southern Gardens Citrus. Samples were collected from Broward, Charlotte, Collier, Desoto, Highlands, Martin, and Okeechobee Counties in southern Florida.
*c FDACS = Florida Department of Agriculture and Consumer Services. Samples were collected from 29 counties from north, central, and south Florida.
originated from central Florida (Fig. 3), and 48 strains were \( \text{TRN}_{<10} \) (mean = 4.9 ± 0.5). All strains from the USSCSGC samples were \( \text{TRN}_{<10} \) (mean = 5.0 ± 0.3). The larger standard deviation of Guangdong strains indicated that the population of ‘Ca. L. asiaticus’ in Guangdong was more diverse than that in Florida. Standard deviations of \( \text{TRN}_{<10} \) from both Guangdong and Florida were not calculated due to the small sample size.

CLIBASIA_01645 was annotated as a bacteriophage repressor protein C1 (11). The gene has a length of 672 bp (223 amino acids). The tandem repeat region spanned from nucleotide position 401 to 436. A peptidase S24 LexA-like region was predicted to be encoded by nucleotide position 429 to 654. Because the repeat unit (AGACACA) in CLIBASIA_01645 had 7 bp, only the addition or deletion of three units (i.e. 21 nucleotides, encoding RHKTQDT) could keep the gene in frame for translation. Sequences with TRNs of 4, 6, and 7 led to peptide truncations at nucleotide positions 468, 462, and 486, respectively. Only three strains from Guangdong were in-frame variants with TRN of 8 (Table 1). No in-frame variant was identified in the Florida strains.

**DISCUSSION**

Population diversity of HLB-associated ‘Ca. L. asiaticus’ has been a research challenge for many years since the bacterium was first described. Detection of genomic variations among the HLB-associated ‘Ca. Liberibacter’ strains from different geographical origins based on \( \text{rrn} \) and \( \text{rpl} \) sequences has met with no or limited success (7). Variation in the \( \text{omp} \) locus is limited to a few SNPs (1,7) and may not be efficient and sufficient for population analyses. In this study, we referenced our previous success in differentiating closely related bacterial strains (5) and identified a locus with tandem repeats in the genome of ‘Ca. L. asiaticus’. The data presented herein indicate that the within-population...
distribution frequencies of TRNs of the Guangdong and Florida populations are different.

The coexistence of different TRN genotypes suggests that ‘Ca. L. asiaticus’ harbors a mechanism for tandem repeat generation. The molecular mechanism of TRN variation in bacteria is, in general, believed to be related to DNA strand slippage (2,21,22) and is associated with bacterial environmental adaptation (3). The lack of TRN 9, 10, 11, and 12 genotypes (Table 1; Fig. 2) suggests that there could be two mechanisms of TRN variations in ‘Ca. L. asiaticus’. One involves the change of one unit at a given time, resulting in the continuous TRN variations within the TRN_{>10} and TRN_{<10} genotypes. The other mechanism involves duplication or triplication of several units to create a TRN_{>10} strain from a TRN_{<10} strain. Although the true mechanisms need to be confirmed, it remains to be determined when (how frequently) and where (i.e., other plant hosts, insect vectors, and so on) TRN variations initially occur. Furthermore, the presence of predominant ‘Ca. L. asiaticus’ TRN types in both Guangdong and Florida suggests selection pressure (action of unknown bio- or abio-factors) for TRN_{<} and TRN_{>}, genotypes, respectively, on the bacterial populations.

That the ‘Ca. L. asiaticus’ population in Guangdong is more diverse than that in Florida is consistent with the current knowledge that HLB has been present in Guangdong much longer than in Florida. This also supports the conclusion that strains of ‘Ca. L. asiaticus’. TRN strains are currently identified. These strains could possibly form a second group of more localized strains but many more strains would need to be analyzed before this group would be a definitive group. We speculated that the current ‘Ca. L. asiaticus’ populations in Florida may be due to at least two different introductions. The TRN_{<10} strains, predominated by TRN_{<} strains, could have been introduced from an unknown location and spread in citrus trees throughout Florida. The TRN_{>10} strains could have been introduced into central Florida later, as suggested by the limited spread. Alternatively, both genotypic groups could have been introduced into central Florida, at either the same time or different times, but spread at different rates. In any case, further evidence is needed and the presumptive times of bacterial introduction remain unknown.

Although TRN is a sensitive indicator for bacterial strain genetic variation, it should be noted that the stability of the selected tandem repeat locus is imperative for epidemiological analyses (17,21). Without the in vitro cultivation of ‘Ca. L. asiaticus’, evaluation of tandem repeat stability is difficult. It is theoretically possible that a mixture of ‘Ca. L. asiaticus’ TRN strains could have been introduced into Florida and environmental conditions in Florida favored the selection of TRN_{<} strains. If so, identification of factors (e.g., host plants or environmental factors) affecting TRN variation would help in understanding the epidemiology of HLB hosts. However, sweet orange does not particularly favor any of the TRN variants identified in this study.

It has been speculated that a TRN change will lead to a modification of the expressed protein and, possibly, alteration of its stability, substrate specificity, and enzyme activity (17,21). Although the true coding status of CLIBASIA_01645 remains to be experimentally verified, it is interesting to note that the region of tandem repeats (position 401 to 436) overlapped with the predictive peptidase domain (position 429 to 654). Other than TRN_{<}, all of the TRN variants lead to truncated peptides. The frame-shift variation at the CLIBASIA_01645 locus of the non-TRN_{<} and TRN_{>} strains could be interpreted as a negative regulation on this phage repressor gene. Alternatively, the locus of CLIBASIA_01645 may not be a coding region and its biological function remains cryptic.

In summary, by analyzing TRNs at the CLIBASIA_01645 locus, we found genomic evidence that Guangdong (China) and Florida (United States) populations of ‘Ca. L. asiaticus’ associated with citrus HLB are different. Preliminary evidence showed that there could be two different genotypic groups of ‘Ca. L. asiaticus’ in Florida, one widely spread through Florida and the other found, by far, in central Florida. The analysis technique described here is potentially useful for identification, characterization, and understanding of the diversity of ‘Ca. L. asiaticus’ populations in different geographical regions and various biological habitats.

ACKNOWLEDGMENTS

Part of this research was supported by a California Citrus Research Board grant and Modern Agricultural Technology System of China. We thank C. Ledbetter and M. Sitverson for their assistance in statistical analysis and G. Phillips, S. Lee, and T. Takeda for their technical support.

LITERATURE CITED


