Genetic Variation of ‘*Candidatus Liberibacter solanacearum*’ Haplotype C and Identification of a Novel Haplotype from *Trioza urticae* and Stinging Nettle

M. Haapalainen,[†] J. Wang, S. Latvala, M. T. Lehtonen, M. Pirhonen, and A. I. Nissinen

First, second, and fifth authors: University of Helsinki, Department of Agricultural Sciences, P.O. Box 27, FI-00014 University of Helsinki, Finland; third and sixth authors: Natural Resources Institute Finland (Luke), Natural Resources, Tietotie, FI-31600 Jokioinen, Finland; and fourth author: Finnish Food Safety Authority Evira, FI-00790 Helsinki, Finland.
Accepted for publication 28 March 2018.

ABSTRACT

‘*Candidatus Liberibacter solanacearum*’ (CLso) haplotype C is associated with disease in carrots and transmitted by the carrot psyllid *Trioza apicalis*. To identify possible other sources and vectors of this pathogen in Finland, samples were taken of wild plants within and near the carrot fields, the psyllids feeding on these plants, parsnips growing next to carrots, and carrot seeds. For analyzing the genotype of the CLso-positive samples, a multilocus sequence typing (MLST) scheme was developed. CLso haplotype C was detected in 11% of the *T. anthrisci* samples, in 35% of the *Anthriscus sylvestris* plants with discoloration, and in parsnips showing leaf discoloration. MLST revealed that the CLso in *T. anthrisci* and most

A. sylvestris plants represent different strains than the bacteria found in *T. apicalis* and the cultivated plants. CLso haplotype D was detected in 2 of the 34 carrot seed lots tested, but was not detected in the plants grown from these seeds. Phylogenetic analysis by unweighted-pair group method with arithmetic means clustering suggested that haplotype D is more closely related to haplotype A than to C. A novel, sixth haplotype of CLso, most closely related to A and D, was found in the psyllid *T. urticae* and stinging nettle (*Urtica dioica*, Urticaceae), and named haplotype U.

Additional keywords: *Pastinaca sativa*.

Gram-negative bacteria belonging to the genus *Candidatus Liberibacter* are associated with several plant diseases (reviewed by Haapalainen 2014). These bacteria live as phloem-limited obligate parasites in plants and are persistently transmitted by psyllids, which feed on the plant phloem sap. In Finland, ‘*Candidatus Liberibacter solanacearum*’ (CLso) has been detected in the carrot psyllids (*Trioza apicalis* Förster) and the psyllid-infested symptomatic carrots (Haapalainen et al. 2017; Munyaneza et al. 2010a, b; Nissinen et al. 2014). After the first discovery in Finland, CLso was subsequently found in symptomatic carrots in several other countries in Europe (Alfaro-Fernández et al. 2012; Holeva et al. 2017; Loiseau et al. 2014; Munyaneza et al. 2012a, b, 2015). In Spain, CLso was also detected in symptomatic celery, parsnip, and parsley (Alfaro-Fernández et al. 2017a; Teresani et al. 2014). In France, CLso was detected in the apiaceous crops of celery, chervil, fennel, parsley, and parsnip, in addition to carrot, and the infected carrot, celery, parsley, and chervil plants had severe symptoms (Hajri et al. 2017).

These recent findings of new apiaceous hosts of CLso in France and Spain suggest that there might be host plants other than carrot in northern Europe as well. Unlike in North America, where CLso has been found in several wild plant species of the family Solanaceae (Torres et al. 2015; Wen et al. 2009), no wild host plants are known for CLso in Europe. However, this may be due to a lack of research, since CLso was detected in the seed specimens of the wild carrot *Daucus carota* and the related species *D. aureus* collected from the wild in Lebanon, Western Asia, in 1990 (Monger and Jeffries 2017).

While *T. apicalis* is acting as the vector of CLso in northern Europe, in the Mediterranean region the main vector is *Bactericera trigonica* Hodkinson, feeding on both carrot and celery (Antolínez et al. 2017). Although some of the *Bactericera* species, including *B. cockerelli* (not found in Europe) and *B. nigricornis*, are polyphagous, most of the psyllid species have a narrow host specificity and they seldom feed on more than one host plant genus (Hodkinson 1974). For example, the plant genera *Apium*, *Chaerophyllum*, *Foeniculum*, *Petroselinum*, and *Pastinaca*, which belong to family Apiaceae, are not mentioned at all in the host plant list by Hodkinson (2009), where the life history parameters for 342 psyllid species were analyzed. Thus, the host range of many European psyllid species is still poorly known and there might be more vectors and host plants of CLso to be discovered.

Several distinct haplotypes of CLso have been defined based on single nucleotide polymorphisms in the 50S ribosomal protein gene region *rplJ/rplL*, the 16S-23S rRNA intergenic spacer region and the 16S rRNA (Nelson et al. 2011, 2013; Teresani et al. 2014). The different haplotypes occur in different geographical regions: haplotypes A and B in the North and Central America, haplotype A in New Zealand, haplotype C in northern Europe, and haplotypes D and E in southern Europe and southern Mediterranean region. The different haplotypes have different psyllid vectors, whose host plant range determine the plants that can be infected with CLso. Haplotypes A and B, transmitted by potato/tomato psyllid *B. cockerelli*,

[†]Corresponding author: M. Haapalainen; E-mail: minna.haapalainen@helsinki.fi

First and second authors contributed equally to this research.

Funding: This work was partially supported by funding from the European Union's Horizon 2020 Pest Organisms Threatening Europe research and innovation program under grant agreement number 635646 and the Ministry of Agriculture and Forestry of Finland project numbers 1842/03.01.02/2013 and 1506/03.01.02/2016, and personal grants were provided for M. Haapalainen by Marjatta & Eino Kolli Foundation and for J. Wang by China Scholarship Council.

Sequence deposition: the ribosomal 50S protein gene region *rplJ/rplL*, 23S-16S intergenic spacer region and 16S rRNA gene, GenBank accession numbers from MG701014 to MG701054, and the seven MLST genes, GenBank accession numbers from MG704922 to MG705180.

*The e-Xtra logo stands for “electronic extra” and indicates that one supplementary figure, two supplementary tables, and two supplementary files are published online.

are associated with diseases of solanaceous crops, and the most severe problems are caused to potato cultivation due to zebra chip disease (Munyanenza 2012). In addition to carrots, haplotype C has also been detected in asymptomatic potatoes (Haapalainen et al. 2018) and in the seeds of several cultivars of celeriac (Monger and Jeffries 2017), suggesting that despite the lack of reports on CLso-associated disease in celeriac this species can also be infected. Haplotypes D and E are associated with diseases of several apiaceous crops (Alfaro-Fernández et al. 2017a; EPPO 2017; Hajri et al. 2017; Nelson et al. 2013; Tahzima et al. 2014; Teresani et al. 2014), and haplotype E was also found to infect potato at low level (Antolínez et al. 2017).

The haplotyping scheme presented by Nelson et al. (2011) allows fast identification of the type of CLso detected in the plant and insect samples. However, these DNA fragments are not ideal for phylogenetic studies. The 16S rRNA gene sequence is highly conserved and thus contains insufficient amount of sequence variation to conclusively resolve the phylogeny of closely related bacteria (Williams et al. 2007). The other two ribosomal gene regions contain noncoding intergenic sequences, which are prone to base substitutions and where small-scale repeats and indels tend to accumulate (Higgins and Lemey 2009). Moreover, there are three copies of the ribosomal RNA operon in the CLso genome and these copies are not all completely identical to each other (Wang et al. 2017). The *rplJ/rplL* DNA fragment contains partial coding regions of two different genes and noncoding DNA between them, which complicates the use of this fragment in phylogenetic analyses. Gaps (or indels) of one to three nucleotides were observed in both the noncoding region and the protein-coding region of *rplL*, even when only three closely related CLso haplotypes were compared (Nelson et al. 2011). Using this mosaic fragment increases the risk of ambiguous sequence alignment which may result in bias in the subsequent phylogenetic analysis. Thus, instead of the rRNA gene regions, protein-coding regions of single copy genes should be used for genotyping and phylogenetic analyses. The whole genomic sequences of three different haplotypes of CLso, A, B, and C, have been published (Lin et al. 2011; Thompson et al. 2015; Wang et al. 2017), and can be used for searching new genetic markers. Thereafter, the chosen marker genes can be used for phylogenetic analyses of new samples, instead of whole genome sequencing which is laborious with these unculturable parasitic bacteria.

To reveal possible new sources of CLso infection in Finland, samples of different psyllids and plants were collected and tested for CLso by PCR. Samples included imported carrot seeds used for carrot production by commercial carrot growers, and plant and psyllid samples collected from the fields in several regions of Finland: carrots, parsnips, wild plants growing as weeds within the carrot fields and on the edges of the fields, and the psyllid species feeding on these plants. The sampling was focused in those areas where CLso was previously found in carrots (Haapalainen et al. 2017). For analyzing the genotype of the CLso positive samples, a multilocus sequence typing (MLST) scheme was developed following the guidelines set by previous studies on the genetic variation in bacterial populations (Feil and Spratt 2001; Maiden 2006; Pérez-Losada et al. 2013; Scally et al. 2005). PCR primers were designed for seven single-copy protein-coding genes that represent different essential metabolic functions (Tatusov et al. 2000), are conserved in all the sequenced *Liberibacter* genomes, and have a suitable gene sequence variation (dN/dS) and an adequate sequence length.

MATERIALS AND METHODS

Psyllid and plant samples. To identify potential vectors of CLso, in addition to the known vector *T. apicalis*, samples of *T. anthrisci* and *T. urticae* were collected in Tavastia Proper and Satakunta regions in Finland during the summer of 2015. These two *Trioza* species commonly occur in the yellow sticky traps in the

carrot fields, used to measure *T. apicalis* flight intensity. The psyllids were captured from their primary host plants, *T. anthrisci* from *Anthriscus sylvestris* and *T. urticae* from *Urtica dioica*, at the carrot field edges or ditch banks in the areas where CLso had been previously detected in carrots (Haapalainen et al. 2017). The psyllid samples were collected in the end of May and early June, during the flight peaks of these two species, by shaking the infested plants in 20-liter plastic bags. The sample bags were stored at -20°C , until the psyllids were inspected and transferred individually into sterile Eppendorf tubes for subsequent DNA extraction.

To study the possibility that the wild host plants of *T. anthrisci* and *T. urticae* could function as reservoirs of CLso, plant samples were collected in the end of the growing season in 2015 and 2016. Thirty-seven samples of cow parsley (syn. wild chervil, *A. sylvestris*) and 19 samples of stinging nettle (*U. dioica*) were collected in Tavastia Proper and Satakunta regions, and cow parsley also in South Savonia, at the sites where the psyllids had been observed earlier in the season, near CLso-infected carrot fields. The sampling was not random but targeted: only those plants that showed yellow or purple discoloration in the leaf margins, resembling the symptoms displayed by CLso-infected carrots (Munyanenza et al. 2010a), were collected as samples. In addition to cow parsleys and nettles, six samples of thistle (*Cirsium arvense*) showing leaf discoloration symptoms and growing beside CLso-infected carrot fields and 13 samples of hairy nightshade (*Solanum physalifolium*) growing as weeds within the infected carrot fields were collected in Tavastia Proper and in Satakunta region. In addition, one asymptomatic CLso-positive black nightshade (*S. nigrum*) plant found in a carrot field in Satakunta region in western Finland in 2014 was included in the MLST study. These solanaceous weeds, *S. nigrum* and *S. physalifolium*, were tested for CLso even though they were asymptomatic, because they could form a potential risk to potato health, if infected. *S. nigrum* and *S. physalifolium* are not common species in Finland but only found at sporadic locations. Three of the previously found CLso-positive volunteer potato samples (Haapalainen et al. 2018), and for reference, two CLso-positive carrots from the same region, Satakunta, were also included in the MLST study. To find out if CLso infection had spread to other apiaceous crop plants, samples were taken of parsnips showing yellow or purple discoloration in the leaf margins and growing beside carrots in the same field plot in Tavastia Proper region in 2016.

Commercial carrot growers in Finland use imported carrot seeds, since there is no domestic carrot seed production. To find out if CLso occurring in carrots in Finland would originate from the imported carrot seeds, growers were contacted and requested for samples from the seed lots they had used for sowing their fields in 2016. Thirty-four commercial carrot seed lots were received and tested for CLso by PCR. Later in the season, follow-up carrot samples were collected from the fields that had been sown with the seed tested CLso positive (cultivar Maestro) and CLso negative (cultivar Natalja) on the same farm in Northern Ostrobothnia, the region previously found free of CLso (Haapalainen et al. 2017). Six symptomless plants and six plants displaying leaf discoloration symptom were sampled from both carrot varieties for testing for CLso. None of these plants had leaf curling symptom, which is the typical symptom associated with carrot psyllid feeding (Nissinen et al. 2014).

DNA extraction. The carrot seeds were washed with 0.5% Triton X-100 (Sigma) for 30 min on a platform rocker and subsequently rinsed three times with sterile distilled water to remove the seed dressing. For each sample, 1.0 g of carrot seeds (approximately 500 seeds) was ground to fine powder in Bioreba extraction bags using Homex 6 homogenizer (Bioreba) and suspended in 10 ml of PBST buffer, pH 7.2 (137 mM NaCl, 8 mM Na_2HPO_4 , 1 mM KH_2HPO_4 , 3 mM KCl, and 0.05% Tween 20). DNA was extracted from 200 μl of the suspension, using KingFisher Flex robot with QuickPick Plant DNA purification kit, and eluted in 100 μl of elution buffer (BioNobile). For fresh or frozen plant samples, DNA was extracted

separately from the shoots and the roots, of 0.1-g tissue samples, using FastPrep with lysing matrix A (MP Biomedicals) and DNeasy Plant Mini kit (Qiagen), and eluted in 100 µl of distilled water. Psyllid DNA was extracted from each individual separately using DNeasy Blood and Tissue kit (Qiagen), and eluted in 30 µl of distilled water.

MLST target loci and primer design. Seven genes were chosen for MLST according to the following principles: (i) the genes are present throughout the sequenced ‘*Liberibacter*’ species as single-copy protein-coding genes; (ii) the encoded proteins play a role in different essential metabolic functions; (iii) the sequence length is adequate to obtain an amplicon of 400 to 700 nucleotides; and (iv) there is evidence of stabilizing selection within CLso haplotypes A, B, and C (ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site $dN/dS < 1$), and the nucleotide diversity is similar ($0.005 < \pi < 0.020$). To determine a set of single-copy protein-coding genes that are present in all the ‘*Liberibacter*’ species, complete genome assemblies (Supplementary Table S1) of *Liberibacter crescens* (BT-0, BT-1), ‘*Ca. Liberibacter asiaticus*’ (A4, Gxpsy, Ishi-1, psy62), ‘*Ca. Liberibacter americanus*’ (Sao Paulo), ‘*Ca. Liberibacter africanus*’ (PTSAPSY), CLso (ZC1) and two good quality draft genome assemblies of CLso (NZ1 and FIN114) were compared. Protein coding genes were clustered into ortholog groups using OrthoMCL v2.0.9 (Li et al. 2003) as previously described (Wang et al. 2017). Single-copy ortholog groups with an average size less than 150 codons (450 nucleotides) were filtered out, and the rest were assigned to clusters of orthologous groups (COGs) according to functional categories (Tatusov et al. 2000). Ortholog groups assigned to COGs categories C (energy production and conversion), D (cell cycle control, cell division, chromosome partitioning), E (amino acid transport and metabolism), F (nucleotide transport and metabolism), G (carbohydrate transport and metabolism), L (replication, recombination, and repair), and O (posttranslational modification, protein turnover, chaperones) were extracted. The pairwise dN/dS values of CLso haplotypes A (NZ1), B (ZC1), and C (FIN114) were determined using the method of Yang and Nielsen (2000) in yn00 model of coding sequence evolution from the PAML 4.9b package (Yang 2007) and nucleotide diversity (π) values were calculated using R package PopGenome (Pfeifer et al. 2014). In total, seven genes were selected (Table 1), and primers binding to conserved sequence regions and having equal melting temperatures (60°C with Phusion PCR mix, Thermo Scientific) were designed with Primer3 (Koressaar and Remm 2007).

PCR methods. CLso detection was performed by conventional PCR with 16S internal primers OA2 (GCGCTTATTTTAAATAGG AGCGGCA) (Liefing et al. 2009) and Lsc2 (GCCTCACGACT TCGCAACCCAT) (Haapalainen et al. 2017). Haplotyping was based on this 16S sequence and the *rplJ/rplL* gene sequences, amplified with primers rp01F/CL514F) (CTCTAAGATTTCGGT TGGTT) and rp01R/CL514R) (TATATCTATCGTTGCACCAG) (Liefing et al. 2009), and for part of the samples also on the 16S-23S intergenic spacer sequence, amplified with primers LpFrag4-161 1F (GGTTGATGGGGTCATTTGAG) and LpFrag4-480R (CAC GGTACTGGTTCACTATCGGTC) (Hansen et al. 2008), as previously described (Haapalainen et al. 2017). Carrot seed testing was performed with real-time PCR primers CaLspF/CaLspR and CaLsoP probe (Teresani et al. 2014) with following modifications: Maxima Probe qPCR Master Mix (Thermo Scientific) containing 300 nM of each primer, 100 nM probe, and 5 µl of the purified DNA in a total volume of 25 µl was run on a C1000 thermal cycler equipped with CFX96 real-time PCR detection system (Bio-Rad). For determining the relative CLso titer in the plant samples, real-time PCR was performed with primers LsoF/HLBr and COXf/COXr and probes HLBp and COXp (Li et al. 2006, 2009) and Maxima Probe qPCR Master Mix (Thermo Scientific). The program, 95°C for 10 min and 40 cycles of 95°C for 15 s and 60°C for 1 min, was run on LightCycler480 (Roche). The relative titer was calculated by the Pfaffl method (Pfaffl 2001). For CLso genotyping with the MLST primers (Table 1), Phusion polymerase (Thermo Scientific) was used with the supplied high-fidelity reaction buffer, 200 µM dNTP and 500 nM primers in 50 µl reaction volume. The PCR program was 98°C for 30 s, 35 cycles of 98°C for 10 s, 63°C for 30 s, 72°C for 30 s, and 72°C for 10 min.

DNA sequencing. The products of conventional PCR were purified using QIAquick PCR Purification kit (Qiagen), or when necessary, using QIAquick Gel Purification kit (Qiagen). The purified DNA fragments were sequenced in both directions at the Natural Resources Institute Finland, Jokioinen, using an automated sequencer (Applied Biosystems3500xL Genetic Analyzer, Foster City, CA). The forward and reverse reads were assembled into consensus sequences with preliminary quality trimming using Pregap4 and Gap4 from the Staden package (Bonfield et al. 1995).

Multilocus sequence typing. To determine the MLST of the CLso samples studied, all variations in the seven gene sequences were analyzed. If the nucleotide sequences of the same gene, an MLST locus, from different samples were identical, they were marked with the same allele number (identifier). If there was even

TABLE 1. Multilocus sequence typing target genes and primer sequences

Gene name and locus tag ^a	Encoded protein and functional category ^b	dN/dS^c	π^d	Primers	Primer sequences 5'-3'	PCR product length (bp)
<i>adk</i> (AYJ09_RS05135)	Adenylate kinase (F)	0.2119	0.0198	adk-F adk-R adk-1F adk-1R	CGTGCAGAAGTTAGTAAGGGTA ACTGTATCCGCATCTAACATTC CCTCAATTATCAACGGGTGATA ACAGCATTTCCCTCTATAGAA	443 518
<i>atpA</i> (AYJ09_RS00170)	F0F1-type ATP synthase, alpha subunit (C)	0.1089	0.0109	atpA-F atpA-R	GGTAAAGGTCCCATTGAATGTA AAGCAGAAACATCATTAACCTG	661
<i>fbpA</i> (AYJ09_RS02615)	Fructose-1,6-bisphosphate aldolase (G)	0.2417	0.0118	fbpA-F fbpA-R	ATGCGATCAATGTAGAATCAAC AGTATTCGTATGGTCTTGTTCC	631
<i>ftsZ</i> (AYJ09_RS01140)	Cell division GTPase (D)	0.2450	0.0126	ftsZ-F ftsZ-R	ATTCTTGGTGCTACATTTGATG AATGAGACTGACGACGTAAAA	640
<i>glyA</i> (AYJ09_RS02965)	Glycine/serine hydroxymethyl-transferase (E)	0.2882	0.0118	glyA-F glyA-R	ACGACAACACACAAGTCTTTAA AGCAAAATCATAGGTAGGGAAA	591
<i>groEL</i> (AYJ09_RS03890)	Chaperonin, HSP60 family (O)	0.0341	0.0088	groEL-1F groEL-1R	AGTTGGAAGTCGTGAAGAAATC TTCATAATACGAGCCTGAAC	634
<i>gyrB</i> (AYJ09_RS02005)	Type IIA topoisomerase, B subunit (L)	0.0549	0.0077	gyrB-F gyrB-R	GTTGGTGATGATTGTAGAGAGG CCATAGAGAGGTGGTTTATGA	677

^a Locus tags of the genes of CLso haplotype C genome FIN114 (GenBank accession number NZ_LWEB00000000.1).

^b Functional category according to Tatusov et al. (2000).

^c dN/dS : ratio of nonsynonymous and synonymous substitution rates in pairwise comparison of the protein-coding sequences of CLso haplotypes A (NZ1), B (ZC1), and C (FIN114).

^d π : nucleotide diversity of the protein-coding sequences in CLso haplotypes A (NZ1), B (ZC1), and C (FIN114).

one nucleotide difference, the sequences were marked with different identifiers. Then, for each sample, the combination of the seven allelic identifier numbers formed the allelic profile, from here on called the sequence type (ST). Further, all these different STs were clustered into ST complexes (groups) by goeBURST distance algorithm applied to the multilocus sequence typing data, performed in PhyloViz v2.0 (Francisco et al. 2009).

Phylogenetic analyzes. To analyze phylogenetic relatedness, the assembled sequences of *rplJ/rplL*, 16S-23S rRNA intergenic spacer (IGS), partial 16S rRNA and each individual MLST locus were aligned with the previously sequenced CLso haplotypes and other ‘*Liberibacter*’ species using MAFFT v7.310 (Katoh and Standley 2013) after primer trimming, to observe candidate single nucleotide polymorphisms (SNPs), and to detect putative insertions or deletions (indels). To avoid ambiguous, codon-breaking nucleotide alignments caused by indels, the nucleotide alignments of these protein-coding sequences were corrected via translated amino acid alignments. The MLST nucleotide sequences were first translated into amino acid sequences, using a codon usage table for bacteria, and these sequences were aligned in MAFFT. These alignments were then used as scaffolds for adjusting the corresponding nucleotide alignments. Recombination analyzes of the MLST loci were performed using RDP v4.94beta (Martin et al. 2015) with

MaxChi method. For the *rplJ/rplL* sequence alignment, the sequences were first partitioned into *rplJ*-coding, noncoding and *rplL*-coding regions. Sequences of the noncoding region were directly aligned, but for the two protein-coding regions the alignment was made according to the amino acid sequence and then converted into a nucleotide alignment, similarly as for the MLST sequences. Finally, the alignments of the three parts were concatenated to form the *rplJ/rplL* alignment. Maximum parsimony (MP) trees were constructed based on the alignments of *rplJ/rplL* and 16S-23S rRNA IGS sequences in TNT v1.5 package (Goloboff et al. 2008). Gaps (indels) were not removed from the alignments, in order to avoid losing phylogenetic information. All the nucleotide characters were treated with equal weights and the gaps in the alignments were regarded as a fifth character state. Heuristic parsimony search of 1,000 replications was performed using random addition of sequences and followed by tree bisection reconnection (TBR) branch swapping, retaining the shortest tree of each replication. Strict consensus tree was then calculated from several equally parsimonious trees. To assess the reliability of the MP trees, bootstrap support (1,000 replicates) and Bremer support (decay index) were calculated for each branch. The codon alignments of individual MLST loci were concatenated and only the unique sequence types (STs) of CLso were retained and used for calculating pair-wise genetic distances. Polymorphic sites were extracted from the concatenated alignment of these STs using R package ‘adegenet’ v2.1.0 (Jombart and Ahmed 2011), and the Nei’s genetic distance (Nei 1978) was calculated for all the pairs of STs using R package ‘Poppr’ v2.5.0 (Kamvar et al. 2014). Pair-wise genetic distance matrix was calculated according to six different nucleotide substitution models: JC69 (Jukes and Cantor 1969), K80 (Kimura 1980), K81 (Kimura 1981), F81 (Felsenstein 1981), F84 (Felsenstein 1984), and TN93 (Tamura and Nei 1993) with R package ‘ape’ v4.1 (Paradis et al. 2004). The corresponding gene fragments derived from the genomic sequences of ‘*Ca. Liberibacter asiaticus*’ strains psy62 and Ishi1 were included as outgroups. To manage gaps in the sequence alignments, a pair-wise deletion option was used when computing the pair-wise distances. From the pair-wise genetic distance matrixes calculated by the six models described above, phylogenetic trees were reconstructed using unweighted-pair group method with arithmetic means (UPGMA) clustering method with R package ‘phangorn’ (Schliep 2011).

RESULTS

CLso-positive samples and the haplotypes identified. A total of 11 of the 105 *T. anthrisci* samples from Tavastia Proper and two out of the nine samples from Satakunta region were CLso positive. Accordingly, CLso was also detected in cow parsley, the principal host plant of *T. anthrisci*, in 13 of the 37 symptomatic plants tested. Of the cow parsley samples collected in 2015, two of the nine samples from Tavastia Proper were CLso positive, while the four samples from Satakunta were negative. In 2016, 7 of the 14 samples from Tavastia Proper and 4 of the 10 samples from South Savonia were CLso positive. In six positive cow parsleys, CLso was detected in both the leaf petioles and the roots, indicating that these plants were systemically infected. In the positive cow parsley petiole samples, the CLso cycle threshold (Ct) values varied from 22.9 to 33.1, and accordingly, the relative titers, expressed as the DNA/DNA ratio of the CLso 16S rRNA gene and the plant mitochondrial COX gene, varied from 150 to 175,690 (Supplementary Table S2). The CLso found in *T. anthrisci* and cow parsley samples was identified as haplotype C, based on the sequences of the 50S ribosomal protein *rplJ/rplL* gene region and the 23S-16S intergenic spacer region. Two samples from *T. anthrisci* (15-12 and 15-175) had one SNP in the 23S-16S intergenic spacer region, at the same nucleotide (59: G/A) in both of the samples (Table 2). Intrahaplotype variation of the ribosomal sequences has been previously reported for CLso haplotype D: two samples from Spain had a SNP in the *rplJ/rplL* sequence at site 209 in comparison with

TABLE 2. Polymorphic sites in the 50S ribosomal protein *rplJ/rplL* gene region, the 23S-16S intergenic spacer region, and the 16S rRNA gene of ‘*Candidatus Liberibacter solanacearum*’^a

Gene region	Nucleotide position ^b	Haplotype of ‘ <i>Candidatus Liberibacter solanacearum</i> ’				
		U	A	D	E	C
50S ribosomal proteins <i>rplJ/rplL</i> gene region (637–640 nt)	43	G	G	G	G	C
	82	A	A	G	A	A
	100	C	C	C	C	T
	149	T	C	T	T	C
	151	G	G	G	G	T
	160	A	A	G	A	A
	182	G	G	G	A	G
	209	C	C	A/C	C	C
	239	A	–	A	A/–	A
	310	T	T	C	C	T
	370	C	T	C	C	C
	380–382	–	–	–	–	GCT
	402	T	C	C	C	C
	419	G	G	G	G	T
	425	T	G	G	G	G
	451	G	T	G	G	G
	457	A	A	A	A	G
	505	A	G	G	G	A
	532	C	C	T/C	C	C
	574	T	C	C	C	C
	586	A	G	A	A	A
23S-16S intergenic spacer region (945–947 nt)	59	G	A	A	A	A/G
	112	A	A	G	A	A
	118	C	C	T	C	C
	228–231	-ATT	-ATT	-AT-	-ATT	GATT
	244	A	A	A	G	A
	291	T	T	T	T	C
	452	G	G	A	nd ^c	G
	645	C	T	T	nd	T
	859	A	G	G	nd	G
	918	C	C	C/T	T	C
16S internal fragment (1,168 nt)	998	G	A	G	A	G
	1,032	G	G	A	G	G
	1,078	G	G	G/A	G	G

^a The novel haplotype U was compared with closely related haplotypes A, C, D, and E.

^b Nucleotide position in haplotype C sequence (*rplJ/rplL* sequence starting from CAGTCAATTGAGAT, 23S-16S intergenic sequence starting from GGTG ATGGGGTCAT, and 16S sequence starting from GCGCTATTTTTAAT).

^c nd = no data available.

the other haplotype D samples (Hajri et al. 2017). Another SNP at site 532 in the *rplJ/rplL* sequence of haplotype D samples and a single nucleotide deletion at site 239 in one of the haplotype E samples were identified by sequence comparisons (Table 2).

Five parsnips were tested for CLso infection because these plants showed mild leaf discoloration symptoms. The root samples of all five plants were CLso positive, and the petiole samples were positive for three of the plants. Of these petiole samples the titer of CLso was determined by real-time PCR, and Ct values 23.7, 24.5, and 27.2 were obtained, corresponding to DNA/DNA ratios 37,630, 43,590, and 7,650, respectively, indicating high infection levels. The CLso detected in the symptomatic parsnips was identified as haplotype C, based on the sequence of the 50S ribosomal protein *rplJ/rplL* gene region and the 23S-16S intergenic spacer region. This is the first time haplotype C has been detected in parsnip; previously haplotypes D and E have been found infecting parsnip in Spain and France (Alfaro-Fernández et al. 2017a; Hajri et al. 2017). Haplotype C was also confirmed in the single positive black nightshade, growing as a weed in a carrot field in Satakunta region. The samples of the other weeds, thistle (*C. arvense*) and hairy nightshade (*S. physalifolium*), were all CLso negative.

CLso was detected in 3 of the 34 carrot seed lots tested. Two of these seed samples (cultivars Ceres and Maestro) had high contamination levels and one (cultivar Musico) had a low contamination level, based on the real-time PCR Ct values for CLso: 23.6, 19.8, and 32.4, respectively. The samples of cultivars Ceres and Musico found CLso positive were the only samples of these cultivars, whereas the positive sample of cultivar Maestro was one of the three samples of this cultivar. Of the two seed samples with a high contamination level, haplotype was determined based on the 16S rRNA sequence and the 50S ribosomal protein *rplJ/rplL* gene region, and both of the samples were identified as haplotype D (Nelson et al. 2013). However, the plants grown from the CLso-positive Maestro seeds tested negative for CLso at the end of the growing season.

A total of 5 of the 59 *T. urticae* samples and 2 of the 13 symptomatic nettle plants collected in Tavastia Proper region and tested for CLso were positive. In contrast, none of the 46 *T. urticae* samples or the six symptomatic nettles collected in Satakunta region were CLso positive. Sequence comparisons of the *rplJ/rplL* gene region and the 23S-16S intergenic spacer region indicated that the CLso found in the psyllid *T. urticae* and its host plant stinging nettle in Tavastia Proper represent a new, previously uncharacterized haplotype (Table 2). SNPs at nucleotide positions 402 (T), 425 (T), 574 (T) within the *rplJ/rplL* region and at positions 645 (C) and 859 (A) within the 23S-16S intergenic spacer region were found exclusively in the CLso from the psyllid *T. urticae* and its host plant stinging nettle. This putative new haplotype of CLso was designated as “U” after the host plant genus *Urtica*, since this haplotype was not detected in any of the other plant or psyllid samples tested.

Phylogenetic analyzes based on the ribosomal loci. Phylogenetic analyzes of the two ribosomal loci showed that the *rplJ/rplL* region is a better genetic marker for resolving the CLso haplotypes than the 16S-23S IGS region. The 16S-23S IGS region did not produce a proper clustering of the samples representing haplotype A (Supplementary Fig. S1). Moreover, the sequence data available for haplotype D 16S-23S IGS region was scarce, apparently because for most of haplotype D samples, including the two haplotype D positive seed samples found in this study, the primer pair LpFrag4-1611F/LpFrag4-480R failed to give a PCR product. For haplotype E, only a part of the sequence corresponding to this PCR fragment is available in GenBank for comparison. The sequence alignments of the *rplJ/rplL* region and 16S-23S IGS region were based on the full-length PCR amplicons after primer trimming. By using the maximum parsimony (MP) method, the gaps were not deleted from the sequence alignments, but were regarded as extra information (fifth character states). In the *rplJ/rplL* MP strict consensus tree the CLso samples were clustered with reliable branch supports into five groups, corresponding to

haplotypes A, B, U, D+E, and C (Fig. 1). The haplotypes D and E were not separable by the *rplJ/rplL* sequence. This consensus tree mostly agrees with the previously reported phylogenetic analysis based on the alignment of partial *rplJ/rplL* sequences of CLso samples (Hajri et al. 2017). The analysis result also supports our hypothesis that the putative haplotype U of CLso found in the psyllid *T. urticae* and its host plant *U. dioica* represents a distinct new haplotype.

MLST. To study genetic variation between the CLso samples from different sources within Finland, and especially, to find out whether the CLso haplotype C harbored by the psyllid *T. anthrisci* represents the same or a different strain in comparison with the bacteria harbored by *T. apicalis*, a new MLST scheme was developed. Complete DNA sequences were obtained for all the seven MLST loci of 37 CLso-positive samples, including 24 plant samples of carrot, parsnip, potato, cow parsley, nettle and black nightshade, and 13 psyllid samples of three *Trioza* species (*T. apicalis*, *T. anthrisci*, and *T. urticae*). In addition, the corresponding gene fragments from six genomic sequences of CLso were included in the analysis (Table 3). Only one of the three CLso-positive carrot seed samples, 16-004 (Table 3), was included in the MLST analysis, because the other positive sample with a high contamination level (16-011) was found to contain at least two different sequence types (STs) of haplotype D, one of which was the same ST as in the sample 16-004 (Supplementary File S1). Because the commercial seed lots are usually mixtures of seeds from several producers, the CLso DNA amplified by PCR may represent several different sequence types, and resolving the sequence of a heterogeneous PCR product can be difficult or even impossible without cloning. The third positive sample had a low CLso titer, and the quality of the DNA sequences was not adequate for analysis. The recently published genomic sequence D1 showed two SNPs in the primer atpA-F binding region and one SNP in the fbpA-R binding region. Thus, at least for some strains of haplotype D, the PCR performance could be improved by using alternative primers: atpA-1F (GCGCTTGGTAATCCTATT GAT) giving a 682-bp product with the primer atpA-R, and fbpA-1R (TGCTATACCTTGGACAGACAA) giving a 668-bp product with the primer fbpA-F. For haplotype U samples, alternative primers adk-1F and adk-1R (Table 1) were used for amplifying the *adk* gene fragment, because of two SNPs in the primer adk-F binding region: CGTG(C/A)AGA(A/G)GTTAGTAAGGGTA. Of the five CLso-positive *T. urticae* samples, three were included in the MLST analysis, and the other two, which had a lower CLso titer, were excluded due to inadequate sequence quality. However, the sequence data obtained confirmed that these two samples from *T. urticae* also represented CLso haplotype U. One CLso-positive *T. anthrisci* sample (15-165) had a divergent *ftsZ* sequence with a deletion (Supplementary File S2) that could not be fitted in the analysis model, since the only other sequence found to have the same deletion is from ‘*Ca. L. americanus*’, while all the other six MLST fragments were identical to the CLso ST3 sequences (Table 3). Nevertheless, the CLso carrying this divergent *ftsZ* allele, possibly resulting from a recombination event, can be considered to represent another distinct strain. The *ftsZ* gene fragment showed more variation than the other gene fragments probably because the primers ftsZ-F and ftsZ-R amplified a sequence encoding the alphaproteobacterial C-terminal extension domain and a region preceding this domain. The deletion seen in ‘*Ca. Liberibacter americanus*’ and in the CLso sample 15-165 is in a loop region between the FtsZ C-terminal domain (domain 2) and the alphaproteobacterial C-terminal extension domain (domain 3). Because only two of the seven MLST loci, *ftsZ* and *gyrB*, showed variation between the 33 haplotype C samples from different plants and psyllids, 10 additional samples of *T. apicalis* and *T. anthrisci* were analyzed using these two gene fragments only (Table 4).

Phylogenetic analyzes based on the MLST sequences. In total, based on the seven gene fragments, 10 different STs of CLso, forming five ST complexes, were identified by MLST and

goeBURST analyzes. These ST complexes correspond to haplotypes A, B, C, D, and U determined by the *rplJ/rplL* sequence comparison (Table 3). The ST complex ‘0’ consists of the first four STs, which equally represent the previously reported haplotype C. No evidence of potential recombination event in any MLST loci was detected in the RDP4 program. The CLso found in carrots and carrot psyllids represented haplotype C ST1 and ST2, and the CLso found in parsnips, volunteer potatoes, cow parsleys, and the single black nightshade represented ST1. The MLST sequences derived from six CLso-positive carrot psyllid samples and three infected carrots also represented ST1 (Tables 3 and 4). In the carrot sample 16-857 (Table 3), the *ftsZ* gene had one SNP at position 278 (G instead of T)—the position determined after removal of the primer sequences—and thus this sample represented haplotype C ST2. This same SNP was found in the *ftsZ* sequence of three *T. apicalis* samples tested (one of which contained a mixture of ST1 and ST2, Table 4), which indicates that CLso ST2 is also transmitted by the carrot psyllids. Altogether, CLso ST1 was found in 19 samples and ST2 in four samples. The majority of the *ftsZ* gene sequences derived from cow parsley and all the sequences from the psyllid *T. anthrisci* showed two SNPs in comparison with ST1. While ST1 *ftsZ* gene fragment has T at the positions 278 and 352, ST3 and ST4 have G and C at these positions. One of the ten cow parsleys, sample 16-20 (Table 3), had the ST1 type *ftsZ* sequence. One SNP was also found in the *gyrB* locus: six samples of *T. anthrisci* and one cow parsley had A at the position 454 where all the other haplotype C samples had C. These seven samples represented ST3, while five samples from *T. anthrisci* and eight cow parsley samples represented ST4 (Tables 3 and 4).

Applying the UPGMA clustering method to the different genetic distance matrixes that were calculated using the Nei’s D or nucleotide

substitution models (JC69, K80, K81, F81, F84, and TN93) generated the same tree topology with all the matrixes. The clustering results obtained by the Nei’s D and TN93 substitution models that gave the best resolution between the four STs of haplotype C are presented (Fig. 2A and B). In agreement with the goeBURST results, the UPGMA clustering based on the concatenated MLST loci also shows five groups, corresponding to the different sequence complexes ‘0’, ‘1’, ‘2’, ‘3’, and ‘4’ or haplotypes C, U, D, A, and B, respectively. This result, together with the phylogenetic analysis based on the ribosomal *rplJ/rplL* loci, confirms that the putative haplotype U represents a new haplotype of CLso. The haplotype C (ST complex ‘0’) is divided into two subgroups, the first one including ST1 and ST2, and the second one including ST3 and ST4. The short branch lengths on these subgroups suggest recent lineage divergences within haplotype C. For haplotype D, the MLST allele profile shows that the sample 16-004 from carrot seeds and the genomic sequence D1 from Israel represent different sequence types (ST6 and ST7, Table 3) separated by five SNPs, four in *atpA* and one in *ftsZ*, while the *rplJ/rplL* sequence of 16-004 is identical to D1. The UPGMA clustering results indicate genetic closeness of haplotype C with the group consisting of haplotypes U (ST5), D (ST6 and ST7), and A (ST8 and ST9). Within this group, haplotype U is more closely related to A than to D. The haplotype B (ST10) was found more diverged from the other CLso haplotypes, which agrees with the result of the previous whole genome sequence comparison (Wang et al. 2017).

DNA sequences. The sequenced DNA fragments of the 50S ribosomal protein gene region *rplJ/rplL*, the 23S-16S intergenic spacer region and 16S rRNA gene of the new samples of CLso haplotypes C, D, and U from Finland were deposited at GenBank as accession numbers MG701014 to MG701054. The DNA sequences

Fig. 1. Maximum parsimony strict consensus phylogenetic tree based on the 50S ribosomal protein *rplJ/rplL* gene region of ‘*Candidatus Liberibacter solanacearum*’. The tree was midpoint rooted, and bootstrap value and decay index were calculated for each branch. Tree length, L = 191; consistency index, CI = 0.914; and retention index, RI = 0.957. Numbers on the branches are bootstrap percentage values followed by Bremer support. The corresponding sequence from ‘*Candidatus Liberibacter asiaticus*’ strain psy62 was used as an outgroup. A scale bar representing 5 base pair changes is shown below, and the haplotype symbols on the right side.

of the different alleles of the seven MLST genes of CLso haplotypes C, D, and U were deposited in GenBank as accession numbers MG704922 to MG705180.

DISCUSSION

In this study, a new MLST tool was developed to analyze the sequence variation of CLso found in different plant and psyllid species in Finland. The 50S ribosomal *rplJ/rplL* gene region, thus far used for CLso haplotyping and found very useful in rough and rapid identification, is not well suited for phylogenetic analyzes. This DNA fragment is a mosaic of two different coding regions and the intergenic region, which makes its use in the phylogenetic analyzes technically challenging. When the sequence was partitioned before aligning, and the protein-coding regions were aligned according to the amino acid sequences, several SNPs were located

close to the primer binding sites. Differently from the previous reports (Nelson et al. 2011, 2013; Teresani et al. 2014), the insertion of three nucleotides in the protein-coding region of *rplL* of haplotype C was concluded to be GCT, which results in the insertion of one alanine residue in the protein. In this study, the inference from maximum parsimony phylogeny based on the *rplJ/rplL* fragment clearly clustered CLso sequences into five distinct groups (Fig. 1). However, this analysis did not produce a highly resolved bifurcating phylogram with robust resampling scores on every internal branch, and thus it could not define the genetic relatedness of the different groups. Moreover, this analysis based on a single DNA fragment did not provide high enough resolution to separate possible different strains of CLso within haplotype C. To improve the resolution, seven fragments of different protein coding genes were amplified, sequenced, and used for analyzing sequence variation between the CLso samples from different sources.

TABLE 3. Results of the multilocus sequence typing of ‘*Candidatus Liberibacter solanacearum*’-positive samples

Sample source and identity	Allele identifier ^a							Sequence type (ST)	ST complex	Sample haplotype ^b
	<i>adk</i>	<i>atpA</i>	<i>fbpA</i>	<i>ftsZ</i>	<i>glyA</i>	<i>groEL</i>	<i>gyrB</i>			
<i>Triozia apicalis</i> 11-H40 (SS) ^c	1	1	1	1	1	1	1	1	0	C
<i>T. apicalis</i> 13-12 (TP)	1	1	1	2	1	1	1	2	0	C
<i>T. apicalis</i> 13-25 (TP)	1	1	1	2	1	1	1	2	0	C
<i>T. apicalis</i> 15-P24a (TP)	1	1	1	1	1	1	1	1	0	C
Carrot 13-180, petioles (SS)	1	1	1	1	1	1	1	1	0	C
Carrot 14-866, petioles (SS)	1	1	1	1	1	1	1	1	0	C
Carrot 16-857, root (SK)	1	1	1	2	1	1	1	2	0	C
Carrot 16-860, root (SK)	1	1	1	1	1	1	1	1	0	C
Parsnip 16-1, petioles (TP)	1	1	1	1	1	1	1	1	0	C
Parsnip 16-2, petioles (TP)	1	1	1	1	1	1	1	1	0	C
Parsnip 16-5, petioles (TP)	1	1	1	1	1	1	1	1	0	C
<i>T. anthrisci</i> 15-12 (TP)	1	1	1	3	1	1	2	3 ^d	0	C
<i>T. anthrisci</i> 15-214 (TP)	1	1	1	3	1	1	2	3	0	C
<i>T. anthrisci</i> 13-544A (SK)	1	1	1	3	1	1	1	4	0	C
<i>T. anthrisci</i> 15-167 (TP)	1	1	1	3	1	1	1	4	0	C
<i>T. anthrisci</i> 15-196 (TP)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 16-KL2, shoot (SS)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 16-KL3, shoot (SS)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 16-KL8, shoot (SS)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 16-KL9, shoot (SS)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 15-21, shoot (TP)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 15-31, shoot (TP)	1	1	1	3	1	1	2	3	0	C
<i>A. sylvestris</i> 16-9, shoot (TP)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 16-20, shoot (TP)	1	1	1	1	1	1	1	1	0	C
<i>A. sylvestris</i> 16-24, shoot (TP)	1	1	1	3	1	1	1	4	0	C
<i>A. sylvestris</i> 16-25, shoot (TP)	1	1	1	3	1	1	1	4	0	C
Volunteer potato 12-230, stolons (SK)	1	1	1	1	1	1	1	1	0	C
Volunteer potato 12-96, stolons (SK)	1	1	1	1	1	1	1	1	0	C
Volunteer potato 13-355, stolons (SK)	1	1	1	1	1	1	1	1	0	C
Black nightshade 14-1, roots (SK)	1	1	1	1	1	1	1	1	0	C
<i>Triozia urticae</i> 15-108 (TP)	2	2	2	4	2	2	3	5	1	U
<i>Triozia urticae</i> 15-122 (TP)	2	2	2	4	2	2	3	5	1	U
<i>Triozia urticae</i> 15-125 (TP)	2	2	2	4	2	2	3	5	1	U
<i>Urtica dioica</i> 16-7 (N3), shoot (TP)	2	2	2	4	2	2	3	5	1	U
<i>Urtica dioica</i> 16-26 (N1), shoot (TP)	2	2	2	4	2	2	3	5	1	U
Carrot seeds 16-004	3	3	3	5	3	3	4	6	2	D
<i>T. apicalis</i> 12-111 (TP), genomic ^e	1	1	1	1	1	1	1	1	0	C
<i>T. apicalis</i> 12-114 (TP), genomic	1	1	1	1	1	1	1	1	0	C
<i>B. trigonica</i> D1 (Israel 2017), genomic	3	4	3	6	3	3	4	7	2	D
<i>B. cockerelli</i> NZ1 (New Zealand 2010), genomic	4	5	4	7	4	4	5	8	3	A
<i>B. cockerelli</i> HenneA (Texas USA 2012), genomic	4	5	4	7	4	4	5	8	3	A
<i>Solanum lycopersicum</i> Tomato R1 (California USA 2014), genomic	4	5	4	7	4	4	6	9	3	A
<i>B. cockerelli</i> ZC1 (Texas USA 2009), genomic	5	6	5	8	5	5	7	10	4	B
<i>T. anthrisci</i> 15-165 (TP)	1	1	1	9	1	1	2	nd ^f	nd	C

^a Compared with the sequence type 1 of haplotype C.

^b Haplotyping based on comparison of the 50S ribosomal proteins *rplJ/rplL* gene region (Nelson et al. 2011, 2013).

^c Sampling regions in Finland: SS, South Savonia; SK, Satakunta; TP, Tavastia Proper.

^d This sample had a single nucleotide polymorphism in the 23S-16S intergenic spacer region.

^e The GenBank accession numbers for the genomic sequences are as follows: NZ_LVWB000000000.1 and NZ_LWEB000000000.1 for FIN111 and FIN114, NZ_JMTK000000000.1 for NZ1, NZ_JQIG000000000.1 for HenneA, NZ_JNVH000000000.1 for R1, NC_014774.1 for ZC1, and NZ_PKRU000000000.1 for D1 (project).

^f nd, not defined.

By the MLST approach four different sequence types were identified within the CLso haplotype C. Of these sequence types, or strains, two (ST1 and ST2) were harbored by the carrot psyllid *T. apicalis* and the other two (ST3 and ST4) by the psyllid *T. anthrisci*, which also harbored a CLso strain distinguished by a SNP in the 16S-23S IGS region. The larger variation of CLso sequences derived from the psyllid *T. anthrisci* suggests a longer history of these populations in Finland, in comparison with the CLso harbored by *T. apicalis*. All the carrot psyllid samples collected in both western and eastern Finland represented either ST1 or ST2 or a mixture of them, and both sequence types were found in all the three regions included in this study. The CLso detected in carrots, potatoes and parsnips represented the same two sequence types found in the carrot psyllids. This suggests that *T. apicalis* is the main vector of CLso on the cultivated plants in those areas where carrots are intensively cultivated and the carrot psyllid populations are high.

At first, finding the CLso haplotype C in cow parsley (*A. sylvestris*), a perennial wild plant closely related to carrot, aroused concern that there might be a persistent reservoir of CLso on the edges of the fields. Cow parsley is very common both in the natural and agricultural habitats in all the regions of Finland, and it forms large stands on meadows, field ditch banks and road verges. Flowering of

TABLE 4. Results of the sequence typing of ‘*Candidatus Liberibacter solanacearum*’ haplotype C found in the psyllids *Trioza apicalis* and *T. anthrisci*, based on *ftsZ* and *gyrB* gene sequences

Sample source and identity	Allele identifier ^a		Haplotype C sequence type
	<i>ftsZ</i>	<i>gyrB</i>	
<i>T. apicalis</i> 13-34 (TP) ^b	1	1	1
<i>T. apicalis</i> 15-P28a (TP)	1	1	1
<i>T. apicalis</i> 11-H80 (SS)	1	1	1
<i>T. apicalis</i> 11-H88 (SS)	1 and 2	1	1 and 2
<i>T. anthrisci</i> 15-20 (TP)	3	2	3
<i>T. anthrisci</i> 15-211 (TP)	3	2	3
<i>T. anthrisci</i> 15-217 (TP)	3	2	3
<i>T. anthrisci</i> 15-175 (TP)	3	2	3 ^c
<i>T. anthrisci</i> 15-24 (TP)	3	1	4
<i>T. anthrisci</i> 15-35 (TP)	3	1	4

^a Compared with the sequence type 1 of haplotype C.
^b Sampling regions: TP, Tavastia Proper; SS, South Savonia.
^c This sample had a single nucleotide polymorphism in the 23S-16S intergenic spacer region.

Fig. 2. Phylogenetic trees based on the concatenated multilocus sequence typing loci of the different sequence types of ‘*Candidatus Liberibacter solanacearum*’ (CLso). **A**, Unweighted-pair group method with arithmetic means (UPGMA) phylogenetic tree of the 10 sequence types (STs), using pairwise Nei’s D (Nei 1978). **B**, UPGMA phylogenetic tree of the 10 STs and ‘*Candidatus Liberibacter asiaticus*’ (CLas) strains psy62 and Ishi1 as outgroups, calculated using pairwise genetic distance with nucleotide substitution model TN93 (Tamura and Nei 1993).

cow parsley starts in late May or early June, when cultivated carrot seedlings have just emerged. Some of the CLso-positive symptomatic cow parsleys contained a high titer of the bacteria, indicating that not all of these wild hosts were resistant to CLso. Because the symptomless cow parsleys at the sampling sites were not tested, it is not known whether those plants were infected with CLso or not. It is possible that some of them were infected with a low titer of bacteria and therefore did not display symptoms. In carrots, the development of visible symptoms correlates with a high titer of CLso (Nissinen et al. 2014). Since less than 10% of the cow parsley plants at the sampling sites had foliar discoloration, it is also possible that cow parsleys are not as susceptible to CLso as the cultivated carrots. In comparison, the related species ‘*Ca. L. asiaticus*’ is a severe pathogen of cultivated citrus trees, while many wild Rutaceae family trees are either resistant or tolerant (Ramadugu et al. 2016). Also, despite the high occurrence of CLso on several cultivated plants in Spain, CLso was not detected in any of the weed samples tested (Alfaro-Fernández et al. 2017b). In Finland, it seems that the *T. apicalis*-associated CLso strains in carrots and the *T. anthrisci*-associated CLso strains in cow parsley form separate populations, which is probably due to the host specificity of these two psyllid species. Previous observations (Kainulainen et al. 2002) suggested that the host preference of the two related *Trioza* species may be based on different volatile cues from their host plants. It is not known whether *T. apicalis* is able to transmit CLso into cow parsley or *T. anthrisci* into carrot. However, in the survey 2011 to 2013, some *T. anthrisci* individuals were identified in the sweep net samples from the carrot fields (A. I. Nissinen, *personal observation*). Although no *T. apicalis* has been detected among the psyllid samples collected from cow parsleys (A. I. Nissinen, *personal observation*), the CLso ST1 strain found in one of the symptomatic cow parsleys suggests that carrot psyllids may also occasionally feed on cow parsleys. It seems likely that the CLso found in *T. anthrisci* and *A. sylvestris* do not currently form a major source of CLso infections in carrots, parsnips, or potato.

CLso haplotype D was confirmed in 2 of the 34 imported carrot seed lots tested in 2016. However, haplotype D has not been detected in carrots grown in Finland (Haapalainen et al. 2017; this study), and thus it seems that in Finland the carrot seeds are not a significant source of CLso infections. It may be that the bacteria detected in the seeds were either nonviable or unable to move from the seed coat into the developing embryo: the suspensor connecting the embryo to the funiculus does not contain phloem, as illustrated by Stadler et al. (2005), and CLso is restricted to the phloem sieve cells (Nissinen et al. 2014). Our results agree with a recent study on vertical transmission of CLso by carrot seeds, concluding that seed is not a major pathway of transmission (Loiseau et al. 2017). Also for the solanaceous host plants of CLso, it was concluded that apparently CLso is not transmitted through the true seed from infected plants (Munyanza 2012). Because haplotype D has not been detected in any other plant species or psyllid samples in Finland, it is not known if any of the psyllid species in this country could transmit this haplotype and if this haplotype is able to survive in the cool climate conditions. If haplotype D was able to survive and spread, then it would be expected to be found in the carrots in Finland, considering the long history of carrot cultivation. Although the field-scale carrot cultivation only started in Finland in the late nineteenth century, there are indications that carrots have already been grown in the kitchen gardens since the seventeenth century (Grotenfelt 1922). Moreover, seedlings of wild carrot (*D. carota* subsp. *carota*) have been occasionally found near the harbor areas in Finland (Lampinen and Lahti 2017), but they do not survive over the cold winters. This, together with the results that seed transmission of CLso is a very unlikely event on carrots (Loiseau et al. 2017), suggests that the seeds of wild carrot, assumed to have been brought into this country within the hay seed lots, are not likely to form a significant infection source.

A novel haplotype of CLso was found in the psyllid *T. urticae* and its host plant *U. dioica*, and called “U” after *Urtica*. This is the first

report on CLso in a plant that belongs to neither Solanaceae nor Apiaceae but to the family Urticaceae. In addition to nettle, this plant family includes plants like *Parietaria*, *Boehmeria*, and *Laportea*. In Finland, only the genus *Urtica* is native, with two species *U. dioica* and *U. urens*. The phylogenetic analysis based on seven gene sequences revealed that haplotype U is distinct from all the previously characterized haplotypes. The UPGMA clustering of the concatenated MLST sequence alignment indicated that haplotype U is more closely related to haplotypes A and D than to C. This result supports the conclusion that haplotype phylogeny, based on the CLso core genome sequences, does not correlate with the relatedness of the host plants of the bacteria (Wang et al. 2017). Probably, the host range is determined by exchangeable genes that allow the bacteria to interact successfully with certain species of psyllids and their host plants. Further, the result suggests that different haplotypes of CLso can remain separate even within the same region, if they are transmitted by different psyllid species that feed on different host plants. Since haplotype U was not found in any other plants than nettles, it is not known if this haplotype poses a risk of disease to the cultivated plants in Europe.

The UPGMA clustering also resolved the genetic relatedness between haplotypes A, C, and D, suggesting that haplotype D is more closely related to A than to C. This result is different from the previous studies based on the 16S ribosomal rRNA sequence (Nelson et al. 2013; Teresani et al. 2014), but agrees with a recent analysis of partial *rplJ/rplL* sequences (Hajri et al. 2017). The disagreement between different sequence analyses highlights the effect of both the chosen genetic loci and the calculation models on the topology of the resulting phylogenetic tree, when the sequence comparisons are done within a species, between closely related haplotypes and strains. MLST is a powerful tool for comparing the genetic structure of plant pathogen populations in different geographic regions, and for reconstruction of evolution scenario (Saville et al. 2017). However, geographically extensive sampling and proper sampling size from the same population are necessary for that purpose. With a proper sample size of haplotype C population, two subgroups within haplotype C were clearly identified (Fig. 2A and B). However, phylogenetic relations of haplotype U to A and to D still lack strong resampling support on the internal branch (Fig. 2B). This is probably due to the limited number of samples and especially, different samples representing different sequence types of haplotypes A and D, and thus, to improve the result, more samples and sequence types should be included in the analysis in the future. In this study, only the CLso samples collected in Finland were studied and compared with the genomic sequences deposited in GenBank. Within haplotype C, the genes *ftsZ* and *gyrB* proved valuable in separating bacterial populations that represent different strains, whereas haplotype D showed variation in *atpA* and *ftsZ*. For the other haplotypes of CLso, different genes may have to be selected for MLST to separate the strains. Recently, a multilocus sequence analysis of seven highly conserved genes, *dnaG*, *gyrB*, *mutS*, *nusG*, *rplA*, *rpoB*, and *tufB*, was performed to resolve the phylogenetic tree of the '*Ca. Liberibacter*' species (Morris et al. 2017). Whether this MLSA scheme would be suitable for resolving the phylogeny within a species, needs further study. If suitable genes and primers are found to obtain strain-specific resolution, then MLST can be a powerful tool for epidemiology. The implications of phylogenetic analysis to epidemiology strongly depend on the resolution of the method used. For example, if the CLso-positive *T. anthrisci* and cow parsley samples were classified at the haplotype level only, then the risk of transmission from *T. anthrisci* to carrots and the other cultivated plants would be estimated to be high. In contrast, with the MLST approach that identified unique CLso strains in *T. anthrisci* and cow parsley, the epidemiological risk can be estimated to be relatively low.

ACKNOWLEDGMENTS

We thank A. Eskola, O. Järvinen, S. Räsänen, and S. Tuominen (Luke) for technical assistance and A. Virta (Luke) for DNA sequencing.

LITERATURE CITED

- Alfaro-Fernández, A., Cebrián, M. C., Villaescusa, F. J., Hermoso de Mendoza, A., Ferrándiz, J. C., Sanjuán, S., and Font, M. I. 2012. First report of '*Candidatus Liberibacter solanacearum*' in carrot in mainland Spain. *Plant Dis.* 96: 582.
- Alfaro-Fernández, A., Hernández-Llopis, D., and Font, M. I. 2017a. Haplotypes of '*Candidatus Liberibacter solanacearum*' identified in Umbelliferous crops in Spain. *Eur. J. Plant Pathol.* 149:127-131.
- Alfaro-Fernández, A., Verdeguez, M., Rodríguez-León, F., Ibáñez, I., Hernández, D., Teresani, G. R., Bertolini, E., Cambra, M., and Font, M. I. 2017b. Search for reservoirs of '*Candidatus Liberibacter solanacearum*' and mollicutes in weeds associated with carrot and celery crops. *Eur. J. Plant Pathol.* 147:15-20.
- Antolinez, C. A., Feres, A., and Moreno, A. 2017. Risk assessment of '*Candidatus Liberibacter solanacearum*' transmission by the psyllids *Bactericera trigonica* and *B. tremblayi* from Apiaceae crops to potato. *Sci. Rep.* 7:45534.
- Bonfield, J. K., Smith, K. F., and Staden, R. 1995. A new DNA sequence assembly program. *Nucleic Acids Res.* 23:4992-4999.
- EPPO. 2017. First report of '*Candidatus Liberibacter solanacearum*' on carrots in Israel. EPPO Reporting Service no. 01-2017, article no. 2017/020. <https://gd.eppo.int/reporting/article-5988>
- Feil, E. J., and Spratt, B. G. 2001. Recombination and the population structures of bacterial pathogens. *Annu. Rev. Microbiol.* 55:561-590.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* 17:368-376.
- Felsenstein, J. 1984. Distance methods for inferring phylogenies: A justification. *Evolution* 38:16-24.
- Francisco, A. P., Bugalho, M., Ramirez, M., and Carriço, J. A. 2009. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics* 10:152.
- Goloboff, P. A., Farris, J. S., and Nixon, K. C. 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24:774-786.
- Grotenfelt, G. 1922. Suomalainen Peltokasvivijelys, osa II, Maatalouden Tietokirja. Otava, Helsinki, Finland.
- Haapalainen, M. 2014. Biology and epidemics of '*Candidatus Liberibacter* species, psyllid-transmitted plant-pathogenic bacteria. *Ann. Appl. Biol.* 165:172-198.
- Haapalainen, M., Kivimäki, P., Latvala, S., Rastas, M., Hannukkala, A., Jauhainen, L., Lemmetty, A., Pirhonen, M., Virtanen, A., and Nissinen, A. I. 2017. Frequency and occurrence of the carrot pathogen '*Candidatus Liberibacter solanacearum*' haplotype C in Finland. *Plant Pathol.* 66: 559-570.
- Haapalainen, M., Latvala, S., Rastas, M., Wang, J., Hannukkala, A., Pirhonen, M., and Nissinen, A. 2018. Carrot pathogen '*Candidatus Liberibacter solanacearum*' haplotype C detected in symptomless potato plants in Finland. *Potato Res.* 1-20.
- Hajri, A., Loiseau, M., Cousseau-Suhard, P., Renaudin, I., and Gentit, P. 2017. Genetic characterization of '*Candidatus Liberibacter solanacearum*' haplotypes associated with apiaceous crops in France. *Plant Dis.* 101:1383-1390.
- Hansen, A. K., Trumble, J. T., Stouthamer, R., and Paine, T. D. 2008. A new Huanglongbing species, '*Candidatus Liberibacter psyllauros*' found to infect tomato and potato, is vectored by the psyllid *Bactericera cockerelli* (Sulc.). *Appl. Environ. Microbiol.* 74:5862-5865.
- Higgins, D., and Lemey, P. 2009. Multiple sequence alignment. Pages 68-108 in: *The Phylogenetic Handbook*, 2nd Ed. P. Lemey, M. Salemi, and A.-M. Vandamme, eds. Cambridge University Press, New York.
- Hodkinson, I. D. 1974. The biology of the Psylloidea (Homoptera): A review. *Bull. Entomol. Res.* 64:325-338.
- Hodkinson, I. D. 2009. Life cycle variation and adaptation in jumping plant lice (Insecta: Hemiptera: Psylloidea): A global synthesis. *J. Nat. Hist.* 43: 65-179.
- Holeva, M. C., Glynos, P. E., and Karafla, C. D. 2017. First report of '*Candidatus Liberibacter solanacearum*' on carrot in Greece. *Plant Dis.* 101:1819.
- Jombart, T., and Ahmed, I. 2011. Adegnet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070-3071.
- Jukes, T., and Cantor, C. 1969. Evolution of protein molecules. *Mamm. Protein Metab.* 3:21-132.
- Kainulainen, P., Nissinen, A., Piirainen, A., Tiilikka, K., and Holopainen, J. K. 2002. Essential oil composition in leaves of carrot varieties and preference of specialist and generalist sucking insect herbivores. *Agric. For. Entomol.* 4:211-216.
- Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281.
- Katoh, K., and Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30:772-780.

- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111-120.
- Kimura, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci.* 78:454-458.
- Koressaar, T., and Remm, M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23:1289-1291.
- Lampinen, R., and Lahti, T. 2017. Kasviatlas 2016. Helsingin Yliopisto, Luonnontieteellinen keskusmuseo, Helsinki. <http://koivu.luomus.fi/kasviatlas/>
- Li, L., Stoeckert, C. J., and Roos, D. S. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13:2178-2189.
- Li, W., Abad, J. A., French-Monar, R. D., Rascoe, J., Wen, A., Gudmestad, N. C., Secor, G. A., Lee, I. M., Duan, Y., and Levy, L. 2009. Multiplex real-time PCR for detection, identification and quantification of '*Candidatus Liberibacter solanacearum*' in potato plants with zebra chip. *J. Microbiol. Methods* 78:59-65.
- Li, W., Hartung, J. S., and Levy, L. 2006. Quantitative real-time PCR for detection and identification of '*Candidatus Liberibacter* species associated with citrus huanglongbing. *J. Microbiol. Methods* 66:104-115.
- Liefting, L. W., Weir, B. S., Pennycook, S. R., and Clover, G. R. G. 2009. '*Candidatus Liberibacter solanacearum*', associated with plants in the family Solanaceae. *Int. J. Syst. Evol. Microbiol.* 59:2274-2276.
- Lin, H., Lou, B. H., Glynn, J. M., Doddapaneni, H., Civerolo, E. L., and Chen, C. W. 2011. The complete genome sequence of '*Candidatus Liberibacter solanacearum*', the bacterium associated with potato zebra chip disease. *PLoS One* 6:e19135.
- Loiseau, M., Garnier, S., Boirin, V., Merieau, M., Leguay, A., Renaudin, I., Renvoisé, J.-P., and Gentit, P. 2014. First report of '*Candidatus Liberibacter solanacearum*' in carrots in France. *Plant Dis.* 98:839.
- Loiseau, M., Renaudin, I., Cousseau-Suhard, P., Lucas, P.-M., Forveille, A., and Gentit, P. 2017. Lack of evidence of vertical transmission of '*Candidatus Liberibacter solanacearum*' by carrot seeds suggests that seed is not a major transmission pathway. *Plant Dis.* 101:2104-2109.
- Maiden, M. C. J. 2006. Multilocus sequence typing of bacteria. *Annu. Rev. Microbiol.* 60:561-588.
- Martin, D. P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1:1-5.
- Monger, W. A., and Jeffries, C. J. 2017. A survey of '*Candidatus Liberibacter solanacearum*' in historical seed from collections of carrot and related Apiaceae species. *Eur. J. Plant Pathol.* 150:803-815.
- Morris, J., Shiller, J., Mann, R., Smith, G., Yen, A., and Rodoni, B. 2017. Novel '*Candidatus Liberibacter*' species identified in the Australian eggplant psyllid, *Acizzia solanicola*. *Microb. Biotechnol.* 10:833-844.
- Munyaneza, J. E. 2012. Zebra chip disease of potato: Biology, epidemiology, and management. *Am. J. Potato Res.* 89:329-350.
- Munyaneza, J. E., Fisher, T. W., Sengoda, V. G., Garczynski, S. F., Nissinen, A., and Lemmetty, A. 2010a. First report of '*Candidatus Liberibacter solanacearum*' associated with psyllid-affected carrots in Europe. *Plant Dis.* 94:639.
- Munyaneza, J. E., Fisher, T. W., Sengoda, V. G., Garczynski, S. F., Nissinen, A., and Lemmetty, A. 2010b. Association of '*Candidatus Liberibacter solanacearum*' with the psyllid, *Triozia apicalis* (Hemiptera: Trioziidae) in Europe. *J. Econ. Entomol.* 103:1060-1070.
- Munyaneza, J. E., Sengoda, V. G., Stegmark, R., Arvidsson, A. K., Anderbrant, O., Yuvaraj, J. K., Rämert, B., and Nissinen, A. 2012a. First report of '*Candidatus Liberibacter solanacearum*' associated with psyllid-affected carrots in Sweden. *Plant Dis.* 96:453.
- Munyaneza, J. E., Sengoda, V. G., Sundheim, L., and Meadow, R. 2012b. First report of '*Candidatus Liberibacter solanacearum*' associated with psyllid-affected carrots in Norway. *Plant Dis.* 96:454.
- Munyaneza, J. E., Swisher, K. D., Hommes, M., Willhauck, A., Buck, H., and Meadow, R. 2015. First report of '*Candidatus Liberibacter solanacearum*' associated with psyllid-infested carrots in Germany. *Plant Dis.* 99:1269.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583-590.
- Nelson, W. R., Fisher, T. J., and Munyaneza, J. E. 2011. Haplotypes of '*Candidatus Liberibacter solanacearum*' suggest long-standing separation. *Eur. J. Plant Pathol.* 130:5-12.
- Nelson, W. R., Sengoda, V. G., Alfaro-Fernandez, A. O., Font, M. I., Crosslin, J. M., and Munyaneza, J. E. 2013. A new haplotype of '*Candidatus Liberibacter solanacearum*' identified in the Mediterranean region. *Eur. J. Plant Pathol.* 135:633-639.
- Nissinen, A. I., Haapalainen, M., Jauhiainen, L., Lindman, M., and Pirhonen, M. 2014. Different symptoms in carrots caused by male and female carrot psyllid feeding and infection by '*Candidatus Liberibacter solanacearum*'. *Plant Pathol.* 63:812-820.
- Paradis, E., Claude, J., and Strimmer, K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289-290.
- Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., and Crandall, K. A. 2013. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infect. Genet. Evol.* 16:38-53.
- Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:45e.
- Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., and Lercher, M. J. 2014. PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* 31:1929-1936.
- Ramadugu, C., Keremane, M. L., Hallbert, S. E., Duan, Y. P., Roose, M. L., Stover, E., and Lee, R. F. 2016. Long-term field evaluation reveals huanglongbing resistance in *Citrus* relatives. *Plant Dis.* 100:1858-1869.
- Saville, A., Charles, M., Chavan, S., Muñoz, M., Gómez-Alpizar, L., and Ristaino, J. B. 2017. Population structure of *Pseudocercospora fijiensis* in Costa Rica reveals shared haplotype diversity with Southeast Asian populations. *Phytopathology* 107:1541-1548.
- Scalli, M., Schuenzel, E. L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Appl. Environ. Microbiol.* 71:8491-8499.
- Schliep, K. P. 2011. Phangorn: Phylogenetic analysis in R. *Bioinformatics* 27:592-593.
- Stadler, R., Lauterbach, C., and Sauer, N. 2005. Cell-to-cell movement of green fluorescent protein reveals post-phloem transport in the outer integument and identifies symplastic domains in *Arabidopsis* seeds and embryos. *Plant Physiol.* 139:701-712.
- Tahzima, R., Maes, M., Achbani, E. H., Swisher, K. D., Munyaneza, J. E., and De Jonghe, K. 2014. First report of '*Candidatus Liberibacter solanacearum*' on carrot in Africa. *Plant Dis.* 98:1426.
- Tamura, K., and Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10:512-526.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. 2000. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28:33-36.
- Teresani, G. R., Bertolini, E., Alfaro-Fernandez, A., Martinez, C., Tanaka, F. A. O., Kitajima, E. W., Rosello, M., Sanjuan, S., Ferrandiz, J. C., Lopez, M. M., Cambra, M., and Font, M. I. 2014. Association of '*Candidatus Liberibacter solanacearum*' with a vegetative disorder of celery in Spain and development of a real-time PCR method for its detection. *Phytopathology* 104:804-811.
- 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., Gudmestad, N. C., and Smith, G. R. 2015. Genomes of '*Candidatus Liberibacter solanacearum*' haplotype A from New Zealand and the United States suggest significant genome plasticity in the species. *Phytopathology* 105:863-871.
- Torres, G. L., Cooper, W. R., Horton, D. R., Swisher, K. D., Garczynski, S. F., Munyaneza, J. E., and Barcenas, N. M. 2015. Horizontal transmission of '*Candidatus Liberibacter solanacearum*' by *Bactericera cockerelli* (Hemiptera:Trioziidae) on *Convolvulus* and *Ipomoea* (Solanales: Convolvulaceae). *PLoS One* 10:e0142734.
- Wang, J., Haapalainen, M., Schott, T., Thompson, S., Smith, G., Nissinen, A., and Pirhonen, M. 2017. Genomic sequence of '*Candidatus Liberibacter solanacearum*' haplotype C and its comparison with haplotype A and B genomes. *PLoS One* 12:e0171531.
- Wen, A., Mallik, I., Alvarado, V. Y., Pasche, J. S., Wang, X., Li, W., Levy, L., Lin, H., Scholthof, H. B., Mirkov, T. E., Rush, C. M., and Gudmestad, N. C. 2009. Detection, distribution, and genetic variability of '*Candidatus Liberibacter*' species associated with zebra complex disease of potato in North America. *Plant Dis.* 93:1102-1115.
- Williams, K. P., Sobral, B. W., and Dickerman, A. W. 2007. A robust species tree for the Alphaproteobacteria. *J. Bacteriol.* 189:4578-4586.
- Yang, Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586-1591.
- Yang, Z., and Nielsen, R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* 17:32-43.