Genome Sequence Resources of Two ‘*Candidatus Liberibacter asiaticus*’ Strains from Pakistan

Kehong Liu,¹ Sagheer Atta,² Xuejin Cui,¹ Chunhua Zeng,¹ Jianchi Chen,³ Changyong Zhou,^{1,†} and Xuefeng Wang^{1,†}

¹ National Citrus Engineering Research Center, Citrus Research Institute, Southwest University, Beibei, Chongqing 400712, China

² Department of Plant Protection, Faculty of Agricultural Sciences, Ghazi University Dera Ghazi Khan 32200, Pakistan

³ San Joaquin Valley Agricultural Sciences Center, Agricultural Research Services, United States Department of Agriculture, Parlier, CA 93648, U.S.A.

Abstract

‘*Candidatus Liberibacter asiaticus*’ (CLAs) is an unculturable, phloem-restricted α -Proteobacteria, associated with citrus Huanglongbing (HLB), which is one of the most destructive diseases in citrus production worldwide. Here, we present the genome sequences of CLAs strains PA19 and PA20 from HLB-affected kinnow trees in Multan, Punjab Province, Pakistan. The CLAs genomes of PA19 and PA20 comprise 1,224,156 bp and 1,226,225 bp, respectively, with an average GC content of 36.4%. Both harbored the Type 2 prophage. In this study, we report two CLAs genomes from Pakistan, which extends the sequence database of CLAs and will contribute to CLAs biology and HLB management.

‘*Candidatus Liberibacter asiaticus*’ (CLAs), associated with citrus Huanglongbing (HLB), is an uncultured, phloem-restricted, gram-negative, α -Proteobacteria and can infect almost all citrus varieties (Bové 2006; Jagoueix et al. 1994). The suspect HLB, initially called citrus dieback, was reported to be present on the Indian subcontinent including India and Pakistan during the early 1900s (Hussain and Nath 1927). The presence of HLB in citrus orchards of Pakistan was confirmed based on polymerase chain reaction (PCR) in 2007 (Chohan et al. 2007). Currently, HLB has severely limited citrus production in Punjab Province, where it is the most important citrus producing area in Pakistan. Recent research showed that CLAs strains from India were genetically distinct based on microsatellite analysis of CLAs strains from eight countries (Islam et al. 2012). Genomic information of CLAs strains from the Indian subcontinent are needed to better understand CLAs biology and its epidemiology in Asia. So far, 16 whole genome sequences of CLAs strains from the U.S.A., China, Japan, and Thailand have been released. In this study, CLAs genomes of PA19 and PA20 strains from Pakistan were reported.

In December 2018, leaves of Kinnow mandarin (*Citrus nobilis* Lour. $\times$ *C. deliciosa* Tenora) trees, showing typical blotchy mottling symptoms, were collected from Multan, Pakistan. Total DNA of strains PA19 and PA20 were extracted from the midrib of leaves using Biospin Omni Plant Genomic DNA Extraction Kit (BioFlux, China). CLAs presence was confirmed by real-time PCR (Li et al. 2006). Bacterial DNA was enriched and enlarged as described (Zheng et al. 2018), and then sequenced by HiSeq PE150 Illumina platform (Illumina Inc., San Diego, CA).

[†]Corresponding authors: C. Y. Zhou, zhoucy@cric.cn; and X. F. Wang; wangxuefeng@cric.cn

Kehong Liu and Sagheer Atta contributed to the paper equally.

The author(s) declare no conflict of interest.

Accepted for publication 21 February 2020.

Funding

This study was funded by the National Key R&D Program of China (2018YFD0201500), the National Natural Sciences Foundation of China (31871925, 31671992), Overseas Expertise Introduction Project for Discipline Innovation (B18044), Science and Technology Major Project of Guangxi (Gui Ke AA18118046) and Chongqing Research Program of Basic Research and Frontier Technology (cstc2017jcyjAX0150).

Keywords

‘*Candidatus Liberibacter asiaticus*’, citrus Huanglongbing, genome, Pakistan

Table 1. Mapping information of PA19 and PA20 strains

Reference sequence	Consensus size (bp)		Average nucleotide coverage (X)		Total mapped reads count	
	PA19	PA20	PA19	PA20	PA19	PA20
A4 (1,233,514 bp)	1,233,514	1,222,582	57.52	55.75	585,279	480,765
SC1 (40,048 bp)	30,429	18,732	309.18	25.18	93,650	7,919
SC2 (38,997 bp)	36,562	34,362	557.45	60.14	213,049	17,612
P-JXGC-3 (31,449 bp)	19,570	16,368	326.45	33.74	79,629	8,397

A total of 91,828,062 and 86,678,874 paired reads (150 bp per read) for PA19 and PA20 were generated, respectively. These reads were mapped against A4 (GenBank accession no. CP010804), SC1 (HQ377372), SC2 (HQ377373), and P-JXGC-3 (KY661963) using CLC Genomics Workbench 11.0.1 (Qiagen) with default settings. PCR reactions using prophage specific primers were performed to target three types of prophage (Zheng et al. 2018). These results indicated that PA19 and PA20 lacked Type 1 and Type 3 prophages but harbored the Type 2 prophage (designated as P-PA19-2 and P-PA20-2, respectively) (Table 1). Citrus genome sequences were excluded from HiSeq reads using *Citrus sinensis* (GenBank accession AJPS000000000), *Citrus clementina* (AMZM000000000), and *Citrus unshiu* (BDQV000000000) as references as possible by Bowtie2 2.2.5 with default settings (Langmead and Salzberg 2012). Then, de novo assembly was performed with 26,494,090 reads for PA19 and 31,542,958 reads for PA20 using Velvet 1.2.10 (kmer = 85 for PA19, kmer = 93 for PA20, min_contig_lgth = 1,000) (Zerbino and Birney 2008). Consensus sequence was obtained by mapping HiSeq reads against reference sequence, and contigs were attained by de novo assembly. The genome sequence was assembled by combining both consensus sequence and contigs through BLAST+ 2.9.0 (Camacho et al. 2009). Furthermore, read mapping method was used to optimize assembly of prophage sequence, and the predicted sequences were identified by PCR with specific primers and confirmed using Sanger's sequencing method. The genome of PA19 contains 5 contigs ranging from 31,502 to 657,537 with a genome size of 1,224,156 bp and an average GC content of 36.4%. The genome of PA20 consists of 7 contigs ranging from 9,907 to 595,991 with genome size of 1,226,225 bp and an average GC content of 36.4%. The online version of the RAST annotation server (<https://rast.nmpdr.org/>) (Aziz et al. 2008) was used to annotate the genomes. After annotation, 1,203 open reading frames (ORFs), 51 RNA genes, 157 subsystems for PA19, and 1,200 ORFs, 53 RNA genes, 157 subsystems for PA20 were found. Two contigs respectively connecting both 5' and 3' end of P-PA19-2 and P-PA20-2 sequences were found and amplified, so P-PA19-2 and P-PA20-2 were confirmed as Type 2-like circular prophages, containing 31,502 bp and 38,427 bp, respectively, with an average GC content of 38.9%.

The present study reports the CLas genome sequences of PA19 and PA20 strains for the first time from Pakistan. The information on genome of PA19 and PA20 will enrich the CLas genome database, contribute to CLas population diversity, and aid future research designing efficient management strategy to control HLB.

Accession Number

These Whole Genome Shotgun projects have been deposited at DDBJ/ENA/GenBank under the accession WOXD000000000 for PA19 and WOUN000000000 for PA20. The versions described in this paper are version WOXD010000000 for PA19 and WOUN010000000 for PA20.

Literature Cited

- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., and Zagnitko, O. 2008. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 9:75.
- Bové, J. M. 2006. Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *J. Plant Pathol.* 88:7-37.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.
- Chohan, S. N., Qamar, R., Sadiq, I., Azam, M., Holford, P., and Beattie, A. 2007. Molecular evidence for the presence of huanglongbing in Pakistan. *Australas. Plant Dis. Notes* 2:37-38.
- Hussain, M. A., and Nath, D. 1927. The citrus psylla (*Diaphorina citri* Kuwayama) [Psyllidae:Homoptera]. *Memories of the Department of Agriculture in India. Entomological Series* 10:1-27.

- Islam, M. S., Glynn, J. M., Bai, Y., Duan, Y. P., Coletta-Filho, H. D., Kuruba, G., Civerolo, E. L., and Lin, H. 2012. Multilocus microsatellite analysis of '*Candidatus Liberibacter asiaticus*' associated with citrus Huanglongbing worldwide. *BMC Microbiol.* 12:39.
- Jagoueix, S., Bove, J. M., and Garnier, M. 1994. The phloem-limited bacterium of greening disease of citrus is a member of the α subdivision of the Proteobacteria. *Int. J. Syst. Bacteriol.* 44:379-386.
- Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357-359.
- Li, W. B., 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.
- Zerbino, D. R., and Birney, E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821-829.
- Zheng, Z., Bao, M., Wu, F., Van, H. C., Chen, J., and Deng, X. 2018. A type 3 prophage of '*Candidatus Liberibacter asiaticus*' carrying a restriction-modification system. *Phytopathology* 108:454-461.