MPMI Vol. XX, No. X, XXXX, pp. X–X, https://doi.org/10.1094/MPMI-05-21-0116-A

RESOURCE ANNOUNCEMENT

e-Xtra*

Whole-Genome Data from *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* Strains Associated with Tan Spot of Mungbean and Soybean Reveal Diverse Plasmid Profiles

Niloofar Vaghefi,[1,†] Dante L. Adorada,[1] Lauren Huth,[1] Lisa A. Kelly,[1,2] Barsha Poudel,[1] Anthony Young,[3] and Adam H. Sparks[1,4]

[1] Centre for Crop Health, University of Southern Queensland, Australia
[2] Department of Agriculture and Fisheries, Queensland, Australia
[3] School of Agriculture and Food Sciences, The University of Queensland, Australia
[4] Department of Primary Industries and Regional Development, Perth, WA 6983, Australia

Abstract

Despite the substantial economic impact of *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* on legume production worldwide, the genetic basis of its pathogenicity and potential host association is poorly understood. The production of high-quality reference genome assemblies of *C. flaccumfaciens* pv. *flaccumfaciens* strains associated with different hosts sheds light on the genetic basis of its pathogenic variability and host association. Moreover, the study of recent outbreaks of bacterial wilt and microevolution of the pathogen in Australia requires access to high-quality reference genomes that are sufficiently closely related to the population being studied within Australia. We provide the first genome assemblies of *C. flaccumfaciens* pv. *flaccumfaciens* strains associated with mungbean and soybean, which revealed high variability in their plasmid composition. The analysis of *C. flaccumfaciens* pv. *flaccumfaciens* genomes revealed an extensive suite of carbohydrate-active enzymes potentially associated with pathogenicity, including four carbohydrate esterases, 50 glycoside hydrolases, 23 glycosyl transferases, and a polysaccharide lyase. We also identified 11 serine peptidases, three of which were located within a linear plasmid, pCff119. These high-quality assemblies and annotations will provide a foundation for population genomics studies of *C. flaccumfaciens* pv. *flaccumfaciens* in Australia and for answering fundamental questions regarding pathogenicity factors and adaptation of *C. flaccumfaciens* pv. *flaccumfaciens* to various hosts worldwide and, at a broader scale, contribute to unraveling genomic features of gram-positive, xylem-inhabiting bacterial pathogens.

Curtobacterium flaccumfaciens pv. *flaccumfaciens* (NCBI: txid138532, *Actinobacteria*, *Microbacteriaceae*) is a gram-positive, xylem-inhabiting plant pathogen that causes bacterial wilt on a broad range of legume crops, including dry bean (*Phaseolus vulgaris*), cowpea (*Vigna unguiculata*), mungbean (*V. radiata*), and soybean (*Glycine max*). *C. flaccumfaciens* pv. *flaccumfaciens* enters the host through contaminated seed, wounds, and natural openings and, subsequently, infects the vascular tissue, impeding the translocation of water and nutrients (Hsieh et al. 2004; Osdaghi et al. 2020). In Australia, *C. flaccumfaciens* pv. *flaccumfaciens* is

†Corresponding author: N. Vaghefi; niloofar.vaghefi@usq.edu.au

*The *e*-Xtra logo stands for "electronic extra" and indicates that a supplementary file and supplementary table are published online.

The author(s) declare no conflict of interest.

Accepted for publication 27 June 2021.

Funding

This project was funded by the Grains Research and Development Corporation project DAQ00186 and the Broad Acre Cropping Initiative, a partnership between the Department of Agriculture and Fisheries of the Queensland Government and the University of Southern Queensland.

Keywords

gram-positive, *Microbacteriaceae*, xylem-inhabiting bacteria

known as the cause of tan spot of mungbean, incurring yield losses of up to 25% (Vaghefi et al. 2019; Wood and Easdown 1990). The symptoms of bacterial wilt on mungbean include irregular, papery, tan-brown interveinal necrosis with chlorotic margins (Kelly 2019). *C. flaccumfaciens* pv. *flaccumfaciens* is also occasionally isolated from similar symptoms on soybean crops in Australia, although its economic importance on soybean is not known. Variability in pathogenic ability and symptomology of *C. flaccumfaciens* pv. *flaccumfaciens* strains on different hosts has been reported (Chen et al. 2021; Osdaghi et al. 2020), although the genetic basis of such variation is unknown. High-quality genome assemblies of *C. flaccumfaciens* pv. *flaccumfaciens* strains from mungbean, soybean, and other legume hosts are needed for comparative analyses in order to elucidate the molecular determinants of pathogenicity.

The first commercial mungbean crops in Australia were grown in the 1960s, and the pathogen was first reported on mungbean in Queensland in 1984 and, subsequently, in New South Wales in 1986 (Wood and Easdown 1990). The origin of these outbreaks in Australian mungbean crops is uncertain, although *C. flaccumfaciens* pv. *flaccumfaciens* may have been introduced on contaminated mungbean seed. Alternatively, *C. flaccumfaciens* pv. *flaccumfaciens* may have been transmitted from other cultivated or wild members of the Fabaceae family (Gonçalves et al. 2017), where a selective sweep resulting in the establishment of better-adapted genotypes may have occurred on mungbean crops in the 1980s. Population genomics has the potential to shed light on the origin of *C. flaccumfaciens* pv. *flaccumfaciens* in Australia as well as many unknown aspects of its biology and epidemiology. Such studies require access to an appropriate, high-quality reference genome. The complete genome sequences of a *C. flaccumfaciens* pv. *flaccumfaciens* strain isolated from dry bean in Turkey and another strain from pepper (*Capsicum annuum*) in Iran were released recently (Chen et al. 2021; O'Leary and Gilbertson 2020). However, the study of macroevolutionary events in the last 40 years in Australia requires a reference genome that is more closely related to the Australian populations (Straub et al. 2021). Here, we present complete genome assemblies of two *C. flaccumfaciens* pv. *flaccumfaciens* strains collected from symptomatic mungbean in Australia, constructed through hybrid assembly of Oxford Nanopore Technology (ONT) and Illumina MiSeq sequences. High-quality, near-complete draft assemblies of five additional *C. flaccumfaciens* pv. *flaccumfaciens* strains from mungbean and soybean were also produced using Illumina paired-end sequencing.

Six *C. flaccumfaciens* pv. *flaccumfaciens* strains from symptomatic mungbean leaves (BRIP 70601, BRIP 70606, BRIP 70607, BRIP 70610, BRIP 70614, and BRIP 70615) and one strain collected from symptomatic soybean (BRIP 70624) in Australia were isolated in 2018 (Table 1). The identity of all strains was confirmed using a *C. flaccumfaciens* pv. *flaccumfaciens*-specific PCR assay (Messenberg Guimaraës et al. 2001) and through Sanger sequencing of the 16S region of the ribosomal DNA using universal primer pair 27F/1492R (Lane 1991; Turner et al. 1999), which resulted in sequences identical to that of the *C. flaccumfaciens* pv. *flaccumfaciens* type strain (CFBP3418). Genomic DNA was extracted using a GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, Australia). Long-read sequencing was conducted for two strains only (BRIP 70606 and BRIP 70614). Oxford Nanopore MinION libraries were prepared from 400 ng of each DNA extract using a Rapid Sequencing kit (SQK-RAD004; Oxford Nanopore Technologies, Oxford, U.K.). Libraries were loaded onto MinION FLO-MIN106 R9.4.1 flow cells and sequenced for 6 and 12 h for BRIP 70606 and BRIP 70614, respectively. Guppy 3.1.5 (Wick et al. 2019) was used for base calling ONT MinION Fast5 read files to fastq. Read quality statistics were assessed using NanoPlot v.1.28.2 (De Coster et al. 2018). Nanopore sequencing of BRIP 70606 and BRIP 70614 resulted in approximately 1.7 and approximately 6.5 Gb of data, respectively, with a mean read length of approximately 6 kb. Porechop v.0.2.4 (Wick 2017) was used for adapter removal with an initial check of 100,000 reads (–check_reads 100,000, –discard_middle). The output from Porechop was further quality trimmed using nanofilt v.2.6.0 (De Coster et al. 2018) (–l,500, –q 10, –headcrop, 10 –tailcrop 10). Illumina short-read sequencing was conducted for all seven *C. flaccumfaciens* pv. *flaccumfaciens* strains. Illumina libraries were prepared from 200 ng of each DNA extract using a Nextera DNA Flex library prep kit and Nextera DNA CD Indexes (Illumina, Singapore). Pooled libraries were sequenced on an Illumina MiSeq platform using 600-cycle paired-end V3 reagents kits. Illumina sequencing of the seven strains yielded between 444,685 and 1,707,030 paired-end sequences (between 0.3 and 1.0 Gb of sequence data per strain), with a mean base

Table 1. *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* genome assemblies produced in this study and comparison with publicly available complete genomes

Strain	Host	Status	Size (bp)	Coverage	Contigs	N_{50} (bp)[a]	GC (%)[b]	Genes[c]	BUSCOs[d]	Reference
BRIP 70606	*Vigna radiata*	Complete	3,930,926	>500	2	–	70.8	3,666	C:99.7% (355), F:0.3% (1)	This study
BRIP 70614	*V. radiata*	Complete	3,896,911	>500	2	–	70.9	3,632	C:99.7% (355), F:0.3% (1)	This study
BRIP 70601	*V. radiata*	Draft	3,804,919	106	37	151,068	71	3650	C:99.2% (353), F:0.8% (3)	This study
BRIP 70607	*V. radiata*	Draft	3,921,519	114	21	543,482	70.9	3,773	C:99.7% (355), F:0.3% (1)	This study
BRIP 70610	*V. radiata*	Draft	3,671,545	144	14	473,173	71	3,503	C:99.7% (355), F:0.3% (1)	This study
BRIP 70615	*V. radiata*	Draft	3,886,219	93	11	464,800	70.9	3,710	C:99.2% (353), F:0.8% (3)	This study
BRIP 70624	*Glycine max*	Draft	3,679,082	155	27	189,297	71.1	3,536	C:99.2% (353), F:0.8% (3)	This study
Cff1037	*Phaseolus vulgaris*	Complete	3,781,436	396	4	–	70.9	3,585	C:99.7% (355), F:0.3% (1)	O'Leary and Gilbertson (2020)
ICMP 22053	*Capsicum annuum*	Complete	3,931,704	Unknown	2	–	70.4	3,740	C:99.7% (355), F:0.3% (1)	Chen et al. (2021)

[a] For BRIP 70606 and BRIP 70614, complete assemblies were constructed consisting of a circular chromosome and a linear plasmid; hence, no N_{50} is given.
[b] This value refers to the GC content of the entire assembly (main chromosome and the plasmids).
[c] Total number of protein coding genes predicted in the current study using a local installation of the NCBI Prokaryotic Genome Annotation Pipeline (Tatusova et al. 2016).
[d] Benchmarking Universal Single-Copy Orthologs; C = complete BUSCOs, F = fragmented BUSCOs, and none of the assemblies had any missing BUSCOs.

quality of 31. The quality of the generated raw sequences was assessed using FastQC v.0.11.8 (Andrews 2010). The k-mer counting software Jellyfish v.2.2.3 (Marçais and Kingsford 2011) estimated the genome size to be approximately 3.9 Mb; hence, 77 to 250× coverage of the genomes was achieved via Illumina sequencing. Adapter and barcode removal and quality trimming of the short reads were conducted using Trim Galore v.0.6.3.

Three approaches were adopted for de novo genome assembly of the two *C. flaccumfaciens* pv. *flaccumfaciens* strains with long-read sequence data (BRIP 70606 and BRIP 70614). First, the long-read genome assembler Canu v.1.8 (Koren et al. 2017) was used to correct and assemble the filtered ONT MinION reads with default options, followed by polishing the assemblies with Illumina MiSeq reads using Pilon v.1.23 (Walker et al. 2014) and bwa v.0.7.17 (Li and Durbin 2010). Second, Flye v.2.5 (Kolmogorov et al. 2019) assembler was used on the filtered ONT MinION data (–threads 40, –genome size 4 m, –plasmids), with the –plasmids option to improve assembly of potential plasmids, followed by polishing the assemblies with Illumina MiSeq reads using Pilon and bwa. Finally, Unicycler v.0.4.8 (Wick et al. 2017) was used, which is a hybrid assembly pipeline for bacterial genomes. Illumina short-read assemblies for the additional five strains (BRIP 70601, BRIP 70607, BRIP 70610, BRIP 70615, and BRIP 70624) were constructed using the short-read genome assembler SPAdes 3.14.0 (–careful), allowing the software to automatically choose k-mer values. Genome statistics were evaluated using QUAST v.5.0.2 (Gurevich et al. 2013). Completeness of the genome assemblies from all pipelines was assessed through identification of Benchmarking Universal Single-Copy Orthologs (BUSCO v.4.1.2) and *Actinobacteria* phylum dataset Odb10 (Simão et al. 2015). For BRIP 70606 and BRIP 70614, which had long-read sequence data, Unicycler produced the highest quality genomes, as indicated by 99.7% genome completeness and no missing BUSCOs (Supplementary Table S1). Therefore, subsequent genome annotation and secretome and carbohydrate-active enzyme (CAZyme) analyses were conducted for the Unicycler assemblies of these two strains. Other studies have also reported that accuracy and completeness of Unicycler hybrid assembly with Oxford Nanopore reads was superior to long-read assembly using Canu and Flye followed by polishing using short reads (De Maio et al. 2019).

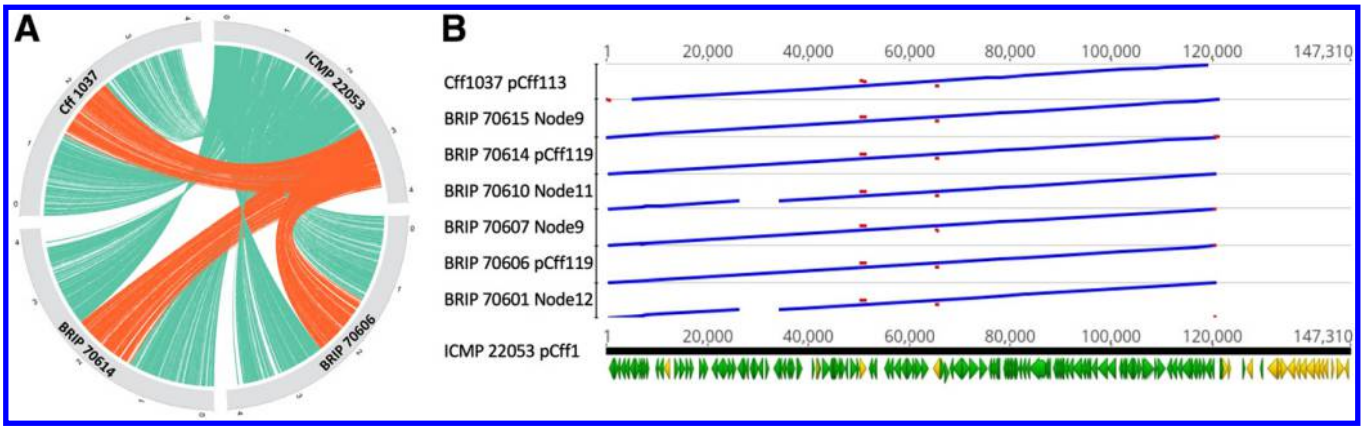

Fig. 1. A, Pairwise whole-chromosome alignment of *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* strains. Chromosomes of BRIP 70606 and BRIP 70614 from mungbean in Australia and Cff1037 from bean in Turkey (O'Leary and Gilbertson 2020) were aligned against that of ICMP 22053 isolated from pepper in Iran (Chen et al. 2021) using Mauve (Darling et al. 2004) and visualized using the R package *circlize* (Gu et al. 2014). Orange links represent an inverted region in the genome of ICMP 22053 compared with the other three genomes **B,** The gene-sparse region of the *C. flaccumfaciens* pv. *flaccumfaciens* plasmid (pCff1) (Chen et al. 2021) is lost in the *C. flaccumfaciens* pv. *flaccumfaciens* strains from Australia and Turkey. The last sequence on the bottom of alignment represents the pCff1 plasmid of strain ICMP 22053 isolated from pepper in Iran (Chen et al. 2021). Annotations show genes (green) and pseudogenes (yellow) predicted in pCff1 using the Prokaryotic Genome Annotation Pipeline (Tatusova et al. 2016) in the current study. The first sequence on the top of the alignment represents the plasmid pCff113 sequence of strain Cff1037 isolated from bean in Turkey (O'Leary and Gilbertson 2020), which is a truncated version of pCff1. The BRIP strains were sequenced in the current study. We constructed complete genome assemblies for BRIP 70606 and BRIP 70614, which identified plasmid pCff119; also, a truncated version of pCff1. The remaining strains (BRIP 70615, BRIP 70601, BRIP 70610, and BRIP 70615) were sequenced using the Illumina MiSeq platform only, resulting in near-complete but fragmented genome assemblies. Contigs (nodes) mapped against pCff1 with >99% similarity are shown. Strain BRIP 70624 lacked the plasmid and, thus, could not be included in the alignment. A dot plot for each strain is shown, with the color blue representing similarities and the color red inversions.

Hybrid genome assembly of BRIP 70606 and BRIP 70614 assembled the genomes into two molecules (Table 1): a circular replicon of approximately 3.7 Mb and a smaller linear replicon of approximately 119 kb. The two replicons have high but different G+C content: 71% for the circular replicon and 66% for the linear replicon. The large circular replicon is flanked by the *rnpA* gene on one side and the *dnaA*, *dnaN*, and *gyrAB* genes on the other; hence, this represents the main chromosome. The smaller linear replicon showed 99.9% similarity to pCff1, a known plasmid of strain ICMP 22053 isolated from pepper (pCff1) (Chen et al. 2021), and was identified as a plasmid, hereafter referred to as pCff119. The plasmid of ICMP 22053 (pCff1) contains an additional region of 27 kb at its 3′ end, resulting in its self complementarity and circularity, whereas pCff119 in the Australian strains has a linear structure, having lost a gene-sparse region of pCff1 (Fig. 1). A similar partial loss of plasmid has been reported in another *C. flaccumfaciens* pv. *flaccumfaciens* strain isolated from dry bean in Turkey (Cff1037) (O'Leary and Gilbertson 2020); Cff1037 seems to have also lost an additional 5-kb sequence from the 5′ end of pCff1 (O'Leary and Gilbertson 2020) (Fig. 1). The two additional plasmids (pCff2 and pCff3) reported in ICMP 22053 (Chen et al. 2021) were not detected in the Australian strains sequenced in the current study. Illumina sequencing and short-read assembly of the five additional *C. flaccumfaciens* pv. *flaccumfaciens* strains from mungbean and soybean resulted in near-complete assemblies (Table 1), ranging in size from 3.6 to 3.9 Mb. BUSCO analyses of the Illumina draft assemblies predicted 99.2 to 99.7% completeness, with no missing BUSCOs (Table 1). All assembled genomes had a GC content of 71%.

Gene annotations were conducted using a local installation of the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al. 2016) for the Australian strains as well as strains ICMP 22053 from pepper (Chen et al. 2021) and Cff1037 from dry bean (O'Leary and Gilbertson 2020) to allow for comparative analyses. Functional annotation of the predicted proteins was conducted using the combined results from DIAMOND v.2.0.4 (Buchfink et al. 2015) and InterProScan v.5.44-79.0 (Jones et al. 2014) in Blast2GO (Conesa et al. 2005). BRIP 70606 was predicted to encode 3,665 coding genes, 47 transfer RNAs (tRNAs), three complete ribosomal RNA (rRNA) operons (5S, 16S, and 23S), and 25 pseudogenes consisting of frameshifted and incomplete genes. In all, 78 putative CAZymes were predicted by all three tools in the dbCRAN meta server, eight of which contained a signal peptide. CAZymes are involved in degradation and metabolism of polysaccharides and, hence, have an essential role

for plant-pathogenic bacteria, giving them the ability to break down plant cell walls and facilitating colonization (Thapa et al. 2019). Of the 78 CAZymes in BRIP 70606, 50 belonged to the glycoside hydrolase (GH) and 23 belonged to the glycosyl transferase (GT) families. The same discovery has been made for another gram-positive, xylem-inhabiting bacterium, *Clavibacter michiganensis* subsp. *michiganensis*, which has a large suite of diverse polysaccharide-degrading proteins (Thapa et al. 2017). Another striking similarity to the *C. michiganensis* genome is the presence of one copy of PL3 in the genome of BRIP 70606, which belongs to the polysaccharide lyase family, reported to degrade pectin (Thapa et al. 2017). The predicted PL3-encoding protein was located on the linear plasmid pCff119, suggesting that the plasmid may have a role in pathogenicity. PGAP annotation of BRIP 70614 led to similar results: a total of 3,632 predicted coding genes, 47 tRNAs, three complete rRNA operons, and 25 pseudogenes consisting of frameshifted and incomplete genes. In all, 79 putative CAZymes were predicted in dbCRAN, nine of which contained a signal peptide. Of these, 51 and 23 CAZymes belonged to the GH and GT families, respectively. Like BRIP 70606, BRIP 70614 also harbors one copy of a PL3-encoding gene in its plasmid (Supplementary File S1).

Identification of a pathogen's secretome is important because many small, secreted proteins (effectors) are known to have an essential role in pathogenicity of phytopathogens (Galán 2009). Unlike gram-negative bacteria that are known to use the type III secretion system (T3SS) for transferring effectors into the host cells, gram-positive bacteria do not have a T3SS (Thapa et al. 2019). Secreted proteins were predicted by identifying proteins involved in the general secretory (Sec) pathway and twin arginine translocation (Tat) system, as predicted by SignalP v.5.0 (Almagro Armenteros et al. 2019), and filtered for lipoproteins and transmembrane proteins, as predicted by LipoP v.1.0 (Juncker et al. 2003) and TMHMM v.2.0 (Sonnhammer et al. 1998). In total, 86 and 84 secreted proteins were detected in BRIP 70606 and BRIP 70614, respectively, the majority of which belonged to the Sec pathway (60 Sec and 26 Tat proteins in BRIP 70606; and 56 Sec and 28 Tat proteins in BRIP 70614). Functional annotation of these secreted proteins predicted multiple carbohydrate-degrading enzymes, hydrolases, and peptidases (Supplementary File S1).

PGAP annotation identified 146 coding sequences (CDS) within plasmid pCff1 of strain ICMP 22053, which represented 126 genes and 20 pseudogenes. Plasmid pCff119 in BRIP 70606 and BRIP 70614 consisted of 126 CDS (122 genes and four pseudogenes). Therefore, the partial loss of plasmid in BRIP 70606 and BRIP 70614 resulted in the loss of 20 CDS, 16 of which belonged to pseudogenes and only four of which were predicted to encode for proteins (Fig. 1). One of these proteins was predicted to contain a helix-turn-helix motif, suggesting a DNA-binding function, whereas no functions were predicted for the remaining three proteins. Function prediction of the 122 proteins predicted to be encoded by pCff119 identified multiple enzymes, including two short-chain dehydrogenases, three serine peptidases, 10 hydrolases, and two lytic transglycolases, suggesting its potential role in pathogenicity.

Mapping the Illumina-based near-complete assemblies of the remaining five *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* strains to pCff1 revealed that BRIP 70607, BRIP 70614, and BRIP 70615 from mungbean harbored the same plasmid (pCff119) as BRIP 70606 and BRIP 70614 (Fig. 1). The isolates BRIP 70601 and BRIP 70610 carry a further truncated version of the plasmid, having lost an additional 8-kb sequence from the middle of pCff1, resulting in the loss of three predicted nucleases, an Lsr2 DNA-binding domain containing protein, and an additional four predicted CDS with unknown functions. Interestingly, no contigs from BRIP 70624 were mapped to pCff1. In order to ensure that this was not due to misassembly of short reads, raw Illumina paired-end reads from BRIP 70624 were mapped back to pCff1, which confirmed the absence of this plasmid in BRIP 70624. Mapping the raw Illumina reads from all the other strains also confirmed the plasmid boundaries detected after the assemblies. Many bacterial plasmids are known to harbor antimicrobial resistance and virulence determinants, which are accessory to bacterial survival and are only beneficial in specific environments (Williams and Thomas 1992). It is plausible that partial and complete loss of plasmids in the Australian *C. flaccumfaciens* pv. *flaccumfaciens* strains may be due to the lack of selective advantage of these loci under Australian environmental conditions, or in association with mungbean and soybean. Loss of such environment-specific accessory genes reduces the cost of DNA replication and gene expression and, hence, is favored by the process of selection. The fact that mostly gene-sparse regions of pCff1 are lost in the Australian strains is further in favor of this hypothesis. In other pathosystems, partial or complete loss of plasmids

has been reported to result in altered fitness and virulence of bacterial strains (Vivian et al. 2001). In *Clavibacter michiganensis*, plasmid-free strains were reported to cause localized lesions but not systemic disease in tomato (Chalupowicz et al. 2017). Similarly, plasmid-cured strains of *Pseudomonas syringae* pv. *phaseslicola* lost virulence toward bean (Jackson et al. 1999). In the current study, the only *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* strain that completely lacked any plasmids was isolated from soybean. Partial or complete loss of plasmids in *C. flaccumfaciens* pv. *flaccumfaciens* strains may also be associated with host specificity. Considering the importance of plasmids as carriers of plant-pathogenic traits, saprophytic fitness and aggressiveness studies on *C. flaccumfaciens* pv. *flaccumfaciens* strains are required to elucidate the potential fitness costs or benefits of plasmids in *C. flaccumfaciens* pv. *flaccumfaciens* populations.

Data Availability

All genome sequencing data have been deposited at the NCBI GenBank database under BioProject accession PRJNA687999.

Author-Recommended Internet Resources

BioProject: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA687999
Trim Galore v.0.6.3: https://github.com/FelixKrueger/TrimGalore

Acknowledgments

We thank E. Adorada for excellent technical support and the growers and agronomists in Queensland and New South Wales for granting us access to their paddocks for sampling.

Literature Cited

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 37:420-423.

Andrews, S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Buchfink, B., Xie, C., and Huson, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12:59-60.

Chalupowicz, L., Barash, I., Reuven, M., Dror, O., Sharabani, G., Gartemann, K. H., Eichenlaub, R., Sessa, G., and Manulis-Sasson, S. 2017. Differential contribution of *Clavibacter michiganensis* ssp. *michiganensis* virulence factors to systemic and local infection in tomato. Mol. Plant Pathol. 18:336-346.

Chen, G., Khojasteh, M., Taheri-Dehkordi, A., Taghavi, S. M., Rahimi, T., and Osdaghi, E. 2021. Complete genome sequencing provides novel insight into the virulence repertories and phylogenetic position of dry beans pathogen *Curtobacterium flaccumfaciens* pv. *flaccumfaciens.* Phytopathology 111:268-280.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674-3676.

Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14:1394-1403.

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and Van Broeckhoven, C. 2018. NanoPack: Visualizing and processing long-read sequencing data. Bioinformatics 34:2666-2669.

De Maio, N., Shaw, L. P., Hubbard, A., George, S., Sanderson, N. D., Swann, J., Wick, R., AbuOun, M., Stubberfield, E., Hoosdally, S. J., Crook, D. W., Peto, T. E. A., Sheppard, A. E., Bailey, M. J., Read, D. S., Anjum, M. F., Walker, A. S., and Stoesser, N.; On Behalf of the Rehab Consortium. 2019. Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. Microb. Genomics 5:e000294.

Galán, J. E. 2009. Common themes in the design and function of bacterial effectors. Cell Host Microbe 5:571-579.

Gonçalves, R. M., Schipanski, C. A., Koguishi, L., Soman, J. M., Sakate, R. K., Silva Júnior, T. A. F., and Maringoni, A. C. 2017. Alternative hosts of *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*, causal agent of bean bacterial wilt. Eur. J. Plant Pathol. 148:357-365.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. 2014. *circlize* Implements and enhances circular visualization in R. Bioinformatics 30:2811-2812.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075.

Hsieh, T. F., Huang, H. C., and Conner, R. L. 2004. Bacterial wilt of bean: Current status and prospects. Pages 181-206 in: Recent Research Development in Plant Science, Vol. 2. S. G. Pandalai, ed. Research Signpost, Trivandrum, India.

Jackson, R. W., Athanassopoulos, E., Tsiamis, G., Mansfield, J. W., Sesma, A., Arnold, D. L., Gibbon, M. J., Murillo, J., Taylor, J. D., and Vivian, A. 1999. Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola.* Proc. Natl. Acad. Sci. U.S.A. 96:10875-10880.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R., and Hunter, S. 2014. InterProScan 5: Genome-scale protein function classification. Bioinformatics 30:1236-1240.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., and Krogh, A. 2003. Prediction of lipoprotein signal peptides in gram-negative bacteria. Protein Sci. 12:1652-1662.

Kelly, L. A. 2019. Integrated disease management strategies in mungbean. Grains Research and Development Corporation Update Papers. https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2019/08/integrated-disease-management-strategies-in-mungbean

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. 2019. Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 37:540-546.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017. Canu: Scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res. 27:722-736.

Lane, D. J. 1991. 16S/23S rRNA sequencing. Pages 115-175 in: Nucleic acid techniques in bacterial systematics. E. Stackebrandt and M. Goodfellow, eds. John Wiley and Sons, New York, NY, U.S.A.

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

Marçais, G., and Kingsford, C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27:764-770.

Messenberg Guimaraĕs, P., Palmano, S., Smith, J. J., Grossi de Sá, M. F., and Saddler, G. S. 2001. Development of a PCR test for the detection of *Curtobacterium flaccumfaciens* pv. *flaccumfaciens.* Antonie Leeuwenhoek 80: 1-10.

O'Leary, M. L., and Gilbertson, R. L. 2020. Complete genome sequence resource of a strain of *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*, the causal agent of bacterial wilt of common bean, from Turkey. Phytopathology 110: 2010-2013.

Osdaghi, E., Young, A. J., and Harveson, R. M. 2020. Bacterial wilt of dry beans caused by *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*: A new threat from an old enemy. Mol. Plant Pathol. 21:605-621.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210-3212.

Sonnhammer, E. L., von Heijne, G., and Krogh, A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6:175-182.

Straub, C., Colombi, E., and McCann, H. C. 2021. Population genomics of bacterial plant pathogens. Phytopathology 111:23-31.

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., and Ostell, J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44:6614-6624.

Thapa, S. P., Davis, E. W., 2nd, Lyu, Q., Weisberg, A. J., Stevens, D. M., Clarke, C. R., Coaker, G., and Chang, J. H. 2019. The evolution, ecology, and mechanisms of infection by gram-positive, plant-associated bacteria. Annu. Rev. Phytopathol. 57:341-365.

Thapa, S. P., Pattathil, S., Hahn, M. G., Jacques, M. A., Gilbertson, R. L., and Coaker, G. 2017. Genomic analysis of *Clavibacter michiganensis* reveals insight into virulence strategies and genetic diversity of a gram-positive bacterial pathogen. Mol. Plant-Microbe Interact. 30:786-802.

Turner, S., Pryer, K. M., Miao, V. P. W., and Palmer, J. D. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J. Eukaryot. Microbiol. 46:327-338.

Vaghefi, N., Adorada, D., Adorada, E., Kelly, L. A., Young, A., and Sparks, A. H. 2019. What can we learn from population genomics studies of *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*, the cause of tan spot on mungbean? In: Proc. Australas. Plant Pathol. Soc. Conf. 25-28 November 2019. Melbourne, Australia. http://era.daf.qld.gov.au/id/eprint/7161/

Vivian, A., Murillo, J., and Jackson, R. W. 2001. The roles of plasmids in phytopathogenic bacteria: Mobile arsenals? Microbiol. Read. 147:763-780.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., and Earl, A. M. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963.

Wick, R. R. 2017. Porechop: Adapter trimmer for Oxford Nanopore reads. https://github.com/rrwick/Porechop

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput. Biol. 13:e1005595.

Wick, R. R., Judd, L. M., and Holt, K. E. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol. 20:129.

Williams, D. R., and Thomas, C. M. 1992. Active partitioning of bacterial plasmids. J. Gen. Microbiol. 138:1-16.

Wood, B. A., and Easdown, W. J. 1990. A new bacterial disease of mung bean and cowpea for Australia. Australas. Plant Pathol. 19:16-21.