Short communication

Complete genome sequence of a boxwood endophyte *Burkholderia* sp. SSG with broad biotechnological application potential

Ping Kong*, Chuanxue Hong

*Hampton Roads Agricultural Research and Extension Center, Virginia Tech, Virginia Beach, VA 23455, USA*

**Article Info**

Received 21 December 2019
Received in revised form 9 March 2020
Accepted 8 April 2020

**Abstract**

*Burkholderia* sp. strain SSG is a boxwood endophyte with potent antagonistic activities against a variety of plant pathogens. Here we present its complete genome sequence that is 8.6 Mb long with a GC content of 66.9%, 10,209 predicted protein-coding sequences, and 866 secondary metabolism gene clusters. Many of these genes and clusters involve antibiosis and other antagonistic activities against plant pathogens and insect pests as well as plant growth promoting traits but none for the *Burkholderia cepacia* epidemic strain marker. This genome sequence supports SSG as a potent biocontrol agent and source of other biotechnological applications.

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

The genus *Burkholderia* is diverse and widespread in the environment [1]. Of particular interest is *Burkholderia cepacia* complex (Bcc) due to the potential of its members to function as plant growth promoters, plant disease control agents, and bioremediators as well as their role as opportunistic human pathogens causing lung disease in immunocompromised individuals [2–6].

Unlike most plant-associated Bcc members that are most typically found in the rhizosphere, *Burkholderia* spp. strain SSG was isolated from boxwood leaves showing a resistant response to inoculation with *Calonectria pseudonaviculata* (Cps); the leaves initially produced water-soaked lesions at the inoculated sites but recovered a few days later [7]. Compared to other biocontrol agents evaluated to date, strain SSG provides superior protection of boxwood from the blight pathogen, *Cps* [7–11]. Although SSG grouped into the Bcc complex, the bacterium clusters separately in 16S and RecA phylogenetic comparisons with known *B. cepacia* species and exhibits distinct traits from clinical Bcc [7]. Here we report its complete genome sequence to provide data that could help resolve the species identity, clear the risk as a human pathogen, and elucidate the potential modes of action as a biocontrol agent and plant growth promoter.

SSG genome DNA was extracted from overnight cultures in nutrient broth (BD, Sparks, MD) at 28 °C using NucleoSpin® Microbial DNA-Macherey Nagel (TaKaRa Bio, Bethlehem, PA) and quantified using Quantus™ Fluorometer (Promega, Madison, WI). Sequencing was performed on a MiniION device (Oxford Nanopore Technologies, Oxford, United Kingdom). The sequencing library was prepared with the ligation sequencing kit (SQK-LSK109) according to the manufacturer’s instructions and run in a FLO-MIN106 (R9.4.1) flow cell. Sequence basecalling was performed using MinKnow (Oxford Nanopore, Oxford, United Kingdom) at Q-score of 11 and run option of Fast5 for 20 h. Fastq files with a total of 9.46 Gb bases from 1.19 million reads that passed the Q score were used for de novo genome assembly using Canu version 1.8 [12] with the default parameters for Nanopore data. After read correction and trimming, the final assembly from the retained single largest high-quality chunk of sequences resulted in a sequence with a total length of 8,571,737 bp and an average GC content of 66.9% arranged in six contigs. The genome coverage is 108.64-fold (N50 = 5,470,797) (Table 1). The assembly was annotated using Prokka 1.14.1 [13] and Rast 2.0 [14]. Prokka predicted 9039 protein coding sequences (CDS) and 76 tRNA, nine rRNA and one tmRNA. Rast predicted 10209 CDS, 67 tRNAs, 18 rRNAs and one tmRNA.

Eight hundred and sixty-six secondary metabolism gene clusters were detected through Rast analysis. 15 gene clusters related to antibiotic biosynthesis were detected with antiSMASH 5 [15], which included genes for nonribosomal peptide synthetase (NPRS), polyketide synthase (PKS), pyrrolnitrin and bacteriocin production (Table 2). These clusters accounted for 6% of the genome assembly. This genome capacity for antibiotic biosynthesis is more than twice that of other analyzed Bcc species [3]. This feature of SSG is consistent to its potent antagonism against oomycete, some bacterial and fungal pathogens (Kong et al, unpublished data). Interestingly, through manual annotation, we identified not only gene cluster for biosynthesis of terpene that has...
been used for pesticide (Table 2), but also genes for production of insecticidal photopexin and presqualene diphosphate synthase (hpnD) [16,17]. Many genes involving plant growth promoting traits were also identified (Table 3). These included genes for nitrogen fixation such as a nitrogenase gene (eg. *nifR*) [18] and a *hglE* cluster or heterocyst glycolipid synthase-like PKS involving nitrogen fixation in cyanobacteria heterocyst [19,20] as well as other genes for nitrogen fixation and regulation including *psn* and *glnB* [20,21]. There were also genes for phosphate solubilization (glucose dehydrogenase and pyrroloquinoline quinone (PQQ)) synthesis proteins for organic acid production [22,23], siderophore production for iron binding and transfer as well as genes for plant growth hormone production or modulation such as auxin biosynthase and ethylene metabolism associated 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase [24]. These results supported SSG as a possible potent biocontrol agent for plant diseases. They also indicated that SSG may also be a candidate biocontrol agent for insect pests and a biofertilizer.

SSG was identified as *B. cepacia* through genome-based identification on TrueBac™ ID [25]. The average nucleotide identity (ANI) between the genomes and the type strain of *B. cepacia* [26] was 98.4%. (ANI coverage of 94.8%). However, multilocus sequence typing (MLST) of the SSG genome sequence through https://pubmlst.org/bcc/ revealed that SSG contains only three of the seven loci that are used for differentiation of species in the Bcc [27,28]. Although SSG had the same allele number at *atpD* as two strains of Bcc (BCC0412, IST431) and the same allele number at *lep* as one strain (BCC0218) of Bcc in genomovar I, the overall SSG allelic profile did not match any Bcc that has been listed previously [27], indicating divergence of SSG from other species in this genomovar that uses *B. cepacia* as a representative.

Table 2

Predicted secondary metabolite clusters involving antibiotic biosynthesis.

| Cluster | Number | Contig | Average size (bp) | % in the genome | Examples | Potential applications |
|---------|--------|--------|------------------|----------------|---------|----------------------|
| Non-ribosomal peptide synthetase (NRPS) | 3 | 1 & 6 | 52601 | 1.84 | Pyochelin, ornibactin | Cytotoxic antibiotics |
| Polyketide synthase (PKS) | 2 | 1 | 46054 | 1.07 | Polyketide, myxochromide D, capsular polysaccharide | Antifungal, anticancer agents |
| tRNA-dependent cyclodipeptide synthases (CDPS) | 1 | 79 | 22042 | 0.26 | Cyclodipeptide | Antifungal, antiviral (influenza A), anti-multidrug resistant bacterial and anticancer agents |
| Terpene synthase | 5 | 1, 19, 79 | 21463 | 1.25 | Terpene | Pesticides |
| Aryl polyene | 1 | 19 | 41210 | 0.48 | Polyene | Anti-oxidants, antibiotics |
| Bacteriocin | 1 | 79 | 10758 | 0.13 | Protein TolQ, Colicin V synthase | Antibacterial drug |
| Phosphonate | 1 | 1 | 40578 | 0.47 | Phosphinothricin tripeptide | Antifungal and anti-oomycete agent |
| Other | 1 | 79 | 41082 | 0.48 | Pyrrolinirin | Antibacterial, antifungal and anti-oomycete agent |

Table 3

Predicted genes/products involving plant growth promotion traits (PGPT).

| Gene/Product | Number of genes (>1) | Example | Contig | PGPT Trait | Potential application |
|--------------|----------------------|---------|--------|-------------|----------------------|
| Coenzymes pyrroloquinoline quinone (PQQ) | 5 | *pqdB,C,D,E* | 1, 79 | Plant defense, production of glucose dehydrogenases (GDHs) | Plant stress resistant elicitor, glucoacidic production, antioxidant, antineuroinflammatory drug production |
| Hydrogen cyanide synthase | 6 | *HcnB, C* | 1 | Regulating availability of phosphate | Biofertilizer |
| Proteins in butanediol metabolic process | 2 | *BudC* | 19, 2, 19 | Plant defense | Plant resistant elicitor |
| Nitrogen metabolism and transport | 4 | *gdh, glnB, psn, ureaA-1, allA, alc, pucl* | 1, 19, 79 | Regulating nitrogen utilization | Biofertilizer |
| Urea degradation | 20 | *acr, pucl, acdS* | 1 | Reducing plant ethylene levels | Plant growth regulator |
| 1-aminocyclopropane-1-carboxylate deaminase (ACC) | 1 | *trPA, B* | 1 | Auxin production | Plant growth regulator |
| Tryptophan synthase | 2 | *accB, C* | 1 | Seed development | Plant seed production |
| Glucoc acid production | 5 | *GDHs, gdh, IV* | 1, 19, 79 | Phosphate solubilization | Biofertilizer |
| Siderophore biosynthesis, transport and liberation of iron | 102 | *yurV, TonB* | All 6 | Iron uptake, phosphate solubilization by production of chelating substance | Plant growth regulator |
indicating absence of BCESM, which is consistent with PCR results presented in the previous study [7]. Together with the presence of genes involved in nitrogen fixation and production of bacteriocin, traits that are uncommon in Bcc clinical strains [29,30], SSG is as a unique member of the Bcc which is distinct from clinical strains and appears to have great promise for agriculture and biotechnology applications.

Nucleotide sequence accession numbers
This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession WTPQ8000000. The version described in this paper is version WTPQ800100000. BioSample SAMN13541113; SRA accession: PRJNA594935. The SSG strain is stored at the Virginia Tech Collection of Phytophthora and Beneficial Microbes (VTC) of the World Data Center for Microorganism (WDCM1197).

Declaration of Competing Interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement
Ping Kong: Conceptualization, Methodology, Data curation, Writing - original draft. Chuanxue Hong: Resources, Writing - review & editing.

Acknowledgements
This research was supported by the U.S. Department of Agriculture (USDA) Agricultural Marketing Service (AMS) [Agreement #: 418607]. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA AMS.

References
[1] S. Compton, J. Nowak, T. Coenye, C. Clement, E.A. Barka, Diversity and occurrence of Burkholderia spp. in the natural environment, FEMS Microbiol. Rev. 32 (2008) 607–626.
[2] T. Coenye, P. Vandamme, Diversity and significance of Burkholderia species occupying diverse ecological niches, Environ. Microbiol. 5 (2003) 719–729.
[3] E. Depooter, M.J. Bull, C. Peeters, T. Coege, P. Vandamme, E. Mahenthiralingam, Burkholderia: an update on taxonomy and biotechnological potential as antibiotic producers, Appl. Microbiol. Biotechnol. 100 (2016) 5215–5229.
[4] E. Mahenthiralingam, T.A. Urban, J.B. Goldberg, The multifarious, multireplicon Burkholderia cepacia complex, Nature Reviews Microbiology 3 (2005) 144–156.
[5] J.L. Parke, D. Gurian-Sherman, Diversity of the Burkholderia cepacia complex and implications for risk assessment of biological control strains, Annu. Rev. Phytopathol. 39 (2001) 225–258.
[6] Z.R. Suárez-moreno, J. Caballero-mellado, B.G. Coutinho, L. Mendonça-previaito, E.K. James, V. Venturi, Common Features of Environmental and Potentially Beneficial Plant-Associated Burkholderia, Microb. Ecol. 63 (2012) 249–266.
[7] P. Kong, C. Hong, A potent Burkholderia endophyte against boxwood blight caused by Calonectria pseudonaviculata, Microorganisms 8 (2) (2020) 310,