Complete genome sequence of *Bacillus velezensis* YYC, a bacterium isolated from the tomato rhizosphere

Yuechen Yan
Qiqihar University
Weihui Xu
Qiqihar University
Wenjing Chen
Qiqihar University
Yunlong Hu
Qiqihar University
zhigang Wang (wangzhigang@qqhru.edu.cn)
Qiqihar University https://orcid.org/0000-0002-0981-6597

Research Article

Keywords: Bacillus velezensis, PGPR, genome, biocontrol

Posted Date: September 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-903033/v1

Version of Record: A version of this preprint was published at Archives of Microbiology on December 21st, 2021. See the published version at https://doi.org/10.1007/s00203-021-02709-5.

Abstract

The *Bacillus velezensis* YYC strain was isolated from the tomato rhizosphere. In a previous experiment, it increased tomato growth and induced systemic resistance against *Ralstonia solanacearum*. However, information on its genomic content is lacking. The complete genome sequence of the bacterium was described in this study. The genome size was 3,973,236 bp and consisted of 4,034 genes in total, with a mean G + C content of 46.52%. 86 tRNAs and 27 ri-bosome RNAs were identified. 14 clusters of secondary metabolites were identified. The KEGG database analysis showed that 69 genes were related to quorum sensing, which were important for cross-kingdom communication. In addition, genes involved in promoting plant growth and triggering plant immunity were identified from the genome. Based on digital DNA–DNA hybridizations (dDDH), *B. velezensis* YYC was the most closely related with *B. velezensis* FZB42. However, compared with *B. velezensis* FZB42, the lantipeptide biosynthesis gene cluster was special and only existed in the genome of *B. velezensis* YYC. The complete genome data of *B. velezensis* YYC will provide a basis for explanation of its growth-promoting mechanism and biocontrol mechanism.

Introduction

Bacillus velezensis is an important member of plant growth-promoting rhizobacteria. This species was found to have multiple growth-promoting effects and to produce a variety of secondary metabolites with antibacterial activity (Chowdhury et al. 2015). *B. velezensis* can be widely isolated from diversified environments, such as plant rhizospheres, soil, rivers, human food, animal guts and seawater, and can easily be isolated and cultured (Ye et al. 2018). In our work, the *B. velezensis* YYC strain was isolated from the tomato rhizosphere in Heilongjiang Province, China. *B. velezensis* YYC increased tomato growth and induced systemic resistance against *Ralstonia solanacearum* (unpublished data). Strain YYC is a non-pathogenic bacterium. Genome sequencing of *B. velezensis* YYC will provide basic insight into the growth-promoting and biocontrol mechanism.

Data description

Genome sequencing, assembly and annotation

B. velezensis YYC strain was propagated in Luria-Bertani broth with shaking at 180 r/min overnight at 30°C. By alignments of the 16S ri-bosome RNA and housekeeping genes, it was identified as *Bacillus velezensis.* Bacterial genomic DNA extraction kit (Majorbio Bio-pharm Technology Co., Ltd., Shanghai, China) was used to extract genomic DNA. A TBS-380 fluorometer (Turner Bio Systems Inc., Sunnyvale, CA) was used to quantify the purified genomic DNA. PacBio RS II Single Molecule Real Time (SMRT) and Illumina sequencing platforms were used to sequence the genomic DNA. The sequencing yielded 170,436 reads, including 1,341,760,841 bp, with 337.7× sequence depth. A statistic of quality information was applied for quality trimming, by which the low-quality data could be removed to result in clean data. Using Unicycler (Version 0.4.7) (Wick et al. 2017), the reads were assembled into contigs. A complete genome

was generated by inspecting and completing the last circular step. Finally, using the Illumina reads, error correction of the PacBio assembly results was performed.

The number of protein coding sequences (CDSs) in the *B. velezensis* YYC genome was predicted by Glimmer (version 3.02) (http://ccb.jhu.edu/software/glimmer/index.shtml) (Delcher et al. 2007) and GeneMarkS software (version 4.3) (Besemer et al. 2005). The transfer RNA (tRNA) gene was analyzed by tRNAscan-SE v2.0 software (Version 2.0) (http://trna.ucsc.edu/software) (Chan et al. 2019). Barrnap software (Version 0.8) (https://github.com/tseemann/barrnap) was utilized to predict ri-bosome RNA genes. By aligning reads with the Nonredundant (NR), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2016), Gene Ontology (GO) (Ashburner et al. 2000), Cluster of Orthologous Groups of proteins (COG) (Galperin et al. 2015) and protein families (Pfam) (Finn et al. 2014) databases, all genes were annotated. The bioactive secondary metabolites were predicted by antiSMASH software (Version 4.0.2) (Weber et al. 2015).

General genome features of B. velezensis YYC

Whole-genome sequencing showed that the *B. velezensis* YYC strain contained a genome size of approximately 3,973,236 bp, with an average G + C content of 46.52%. The Glimmer program predicted that the number of protein coding sequences (CDSs) was 4,034, and the average gene length was 877.29 bp. Furthermore, a total of 86 tRNA and 27 ri-bosome RNA genes were identified and analyzed in the genome. By aligning the genome to sequences from diverse databases, including the NR, Swiss-Prot, Pfam, COG, GO and KEGG databases, the numbers of identified genes were 4,034, 3,533, 3,337, 3,013, 2,668, and 2,163, respectively.

The KEGG database analysis showed a great number of two-component systems (113 genes) and ABC transporters (117 genes). Meanwhile, 69 genes were related to quorum sensing, which were important for cross-kingdom communication (Schikora et al. 2016).

Secondary metabolites related to biocontrol

AntiSMASH version 4.0.2 analysis identified 14 clusters of secondary metabolites (Fig. 1). Seven clusters of secondary metabolites were related to the synthesis of bacillaene, macrolactin, bacilysin, fengycin, difficidin, surfactin and bacillibactin. Some of these substances have antagonistic effects on bacteria, fungi, and viruses (Moldenhauer et al. 2007; Chowdhury et al.

2015; Wu et al. 2015; Zhang et al. 2018). Fengycin was found to induce resistance to plant diseases (Farzand et al. 2019).

Comparative genomics of diverse Bacillus strains

In addition, based on digital DNA–DNA hybridizations (dDDH), *B. velezensis* YYC was the most closely related with *B. velezensis* FZB42 (Table 1). And it shared 96.80% identity with the strain that was used as biofertilizer and biocontrol agent (*B. velezensis* FZB42). Compared with FZB42, lantipeptide biosynthesis

gene cluster was special and only existed in the genome of *B. velezensis* YYC. It is engaged in the synthesis of locillomycin, which was related to hemolytic activity, swarming motility, biofilm formation, and colony morphology (Luo et al. 2019). In addition, four clusters encoding new metabolites with no reported description previously. Comparative analysis of secondary metabolite clusters of these strains were summarized for comparisons (Supplementary Table S1).

Table 1
Pairwise comparisons of *B. velezensis* YYC with type strain genomes

Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G + C content difference (in %)
Bacillus velezensis FZB42	96.8	[95.1–98.0]	90.8	[88.5–92.6]	97.5	[96.3–98.4]	0.05
Bacillus velezensis NRRL B-41580	91.8	[88.8–94.0]	85.4	[82.7–87.7]	93.3	[91.0–95.0]	0.21
Bacillus methylotrophicus KACC 13105	95	[92.7–96.6]	84.6	[81.9–87.0]	95.5	[93.7–96.8]	0.09
Bacillus siamensis KCTC 13613	89.1	[85.8–91.8]	56.9	[54.2–59.7]	85.4	[82.3–88.1]	0.19
Bacillus vanillea XY18	89	[85.6–91.7]	56.9	[54.2–59.7]	85.3	[82.2–88.0]	0.2
Bacillus amyloliquefaciens DSM 7	82.7	[78.8–86.0]	56	[53.2–58.7]	79.8	[76.3–82.8]	0.44
Bacillus nakamurai NRRL B-41091	73.4	[69.5–77.1]	31	[28.6–33.5]	61.2	[57.9–64.4]	1.26
Bacillus tequilensis NCTC13306	30.9	[27.5–34.5]	21.3	[19.0–23.7]	27.5	[24.6–30.6]	2.54
Bacillus spizizenii TU-B-10	33.6	[30.2–37.2]	21	[18.8–23.4]	29.2	[26.3–32.3]	2.7
Bacillus subtilis NCIB 3610	32.5	[29.1–36.0]	20.9	[18.7–23.3]	28.4	[25.5–31.5]	3.13

The dDDH values were provided along with their confidence intervals (C.I.).

Genes involved in promoting plant growth and triggering plant immunity

In addition to producing secondary metabolites with antifungal or antibacterial activity. *B. velezensis* YYC contains a various of genes implicated in biofilm formation and root colonization (Table 2). *B. velezensis* YYC contained the genes encoding acetolactate synthase (*ilvH*, *ilvB*), acetolactate decarboxylase (*alsD*),

and butanediol dehydrogenase (*butB*), which have plant growth–promoting effects including stimulating root formation and increasing systemic disease resistance (He et al. 2013; Jayakumar et al. 2020). *B. velezensis* YYC also has the genes required for synthesis of 2,3-butanediol (*alsD*), the compound reported to trigger systemic resistance (He et al., 2013).

Table 2
Representative genes of *B. velezensis* YYC probably involved in plant bacterium interactions

Gene	Position	Protein	Description
spo0A	2432908 – 2432108	Sporulation transcription factor Spo0A	Biofilm formation
sinR	2466273– 2466614	Transcriptional regulator	Biofilm formation
abrB	45929 – 45645	Transition state regulator Abh	Biofilm formation
resE	2260204– 2261985	Sensor histidine kinase	Biofilm formation
lytS	2757501– 2759282	Sensor histidine kinase	Biofilm formation
ycbA	248181– 249473	Sensor histidine kinase	Biofilm formation
sacB	3926734– 3928215	Levan sucrase	Root adhesion
efp	2452165– 2452722	Elongation factor P	Essential for swarming motility
comP	3033195– 3035495	Histidine kinase	Regulator of surfactin production
fliD	3414679– 3416199	Flagellar capping protein	Elicitation of plant basal defence
flgK	3419847– 3421364	Flagellar hook-associated protein FlgK	Elicitation of plant basal defence
xynB	1846654– 1848069	1,4-beta-D-xylan xylohydrolase	Carbohydrate metabolism
lacR	1216343– 1217104	Lactose phosphotransferase system repressor	Lactose metabolism
lacG	1214690– 1216090	6-phospho-beta-galactosidase	Hydrolyzation of phospholactose
lacE	1212603– 1214300	Phosphotransferase system	Cellobiose degradation
lacF	1214312– 1214626	Phosphotransferase system	Cellobiose degradation
alsD	3488161– 3488928	Acetolactate decarboxylase	Synthesis of 2,3-butanediol
pta	3639358– 3640329	Phosphotransacetylase	Strongly upregulated by root exudate

Gene	Position	Protein	Description
ilvH	2697631–2698149	Acetolactate synthase	Promote plant growth
ilvB	2698146–2699963	Acetolactate synthase	Promote plant growth
alsD	3488161–3488928	Acetolactate decarboxylase	Promote plant growth
butB	629733–630773	Butanediol dehydrogenase	Promote plant growth

Nucleotide sequence accession number

The complete genome sequence of *B. velezensis* YYC was deposited in GenBank under accession number CP075055. (BioProject: PRJNA728388, BioSample: SAMN19079027).

Declarations

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 31870493) and the Basic Research Fees of Universities in Heilongjiang Province, China (No. 135409103).

Compliance with ethical standards

Conflict of interest The author(s) declare no conflict of interest.

References

1. Ashburner M, Ball CA, Blake JA et al (2000) Gene ontology: Tool for the unification of biology. Nat Genet 25:25–29. https://doi.org/10.1038/75556
2. Jayakumar A, Nair IC, Krishnankutty RE (2020) Environmental adaptations of an extremely plant beneficial *Bacillus subtilis* Dcl1 identified through the genomic and metabolomic analysis. Microb Ecol 81:687–702. https://doi.org/10.1007/s00248-020-01605-7
3. Besemer J, Borodovsky M (2005) GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33:W451–W454. https://doi.org/10.1093/nar/gki487
4. Chowdhury SP, Hartmann A, Gao X, Borriss R (2015) Biocontrol mechanism by root-associated *Bacillus amyloliquefaciens* FZB42-a review. Front Microbiol 6:780. https://doi.org/10.3389/fmicb.2015.00780
5. Chan PP, Lowe TM (2019) tRNAscan-SE: Searching for tRNA genes in genomic sequences. Methods Mol Biol 1962:1–14. https://doi.org/10.1007/978-1-4939-9173-01

6. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679. https://doi.org/10.1093/bioinformatics/btm009
7. Finn RD, Bateman A, Clements J (2014) Pfam: The protein families database. Nucleic Acids Res 42:D222–D230. https://doi.org/10.1002/047001153X.g306303
8. Farzand A, Moosa A, Zubair M, Khan AR, Massawe VC, Tahir HAS, Sheikh TMM, Ayaz M, Gao X (2019) Suppression of sclerotinia sclerotiorum by the induction of systemic resistance and regulation of antioxidant pathways in tomato using fengycin produced by *Bacillus amyloliquefaciens* FZB42. Biomolecules 9:613. https://doi.org/10.3390/biom9100613
9. Galperin MY, Makarova KS, Wolf YI, Koonin EV (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43:D261–D269. https://doi.org/10.1093/nar/gku1223
10. He PF, Hao K, Blom J et al (2013) Genome sequence of the plant growth promoting strain *Bacillus amyloliquefaciens* subsp. plantarum B9601-Y2 and expression of mersacidin and other secondary metabolites. J Biotechnol 164:281–291. https://doi.org/10.1016/j.jbiotec.2012.12.014
11. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457–D462. https://doi.org/10.1093/nar/gkv1070
12. Luo C, Liu X, Zhou H, Wang X, Chen Z (2015) Nonribosomal peptide synthase gene clusters for lipopeptide biosynthesis in *Bacillus subtilis* 916 and their phenotypic functions. Appl Environ Microb 81:422–431. https://doi.org/10.1128/AEM.02921-14
13. Moldenhauer J, Chen X, Borriss R, Piel J (2007) Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family. Angew Chem 46:8195–8197. https://doi.org/10.1002/anie.200703386
14. Schikora A, Schenk ST, Hartmann A (2016) Beneficial effects of bacteria-plant communication based on quorum sensing molecules of the N-acyl homoserine lactone group. Plant Mol Biol 90:605–612. https://doi.org/10.1007/s11103-016-0457-8
15. Wick RR, Judd LM, Gorrie CL, Holt KE (2017) Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 3:000132. https://doi.org/10.1101/160614
16. Weber T, Blin K, Duddela S (2015) antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437
17. Wu L, Wu H, Chen L, Yu X, Borriss R, Gao X (2015) Difficidin and bacilysin from *Bacillus amyloliquefaciens* FZB42 have antibacterial activity against Xanthomonas oryzae rice pathogens. Sci Rep-UK 5:463–471. http://dx.doi.org/10.1038/srep12975
18. Ye M, Tang X, Yang R, Zhang H, Li F, Tao F, Li F, Wang Z (2018) Characteristic and application of a novel species of *Bacillus*: *Bacillus velezensis*. Acs Chem Biol 13:500–505. https://pubs.acs.org/doi/pdf/10.1021/acschembio.7b00874

19. Zhang L, Sun C (2018) Fengycins, cyclic lipopeptides from marine *Bacillus subtilis* strains, kill the plant-pathogenic fungus magnaporthe grisea by inducing reactive oxygen species production and chromatin condensation. Appl Environ Microb 84:e00445. https://doi.org/10.1128/AEM.00445-18

Figures

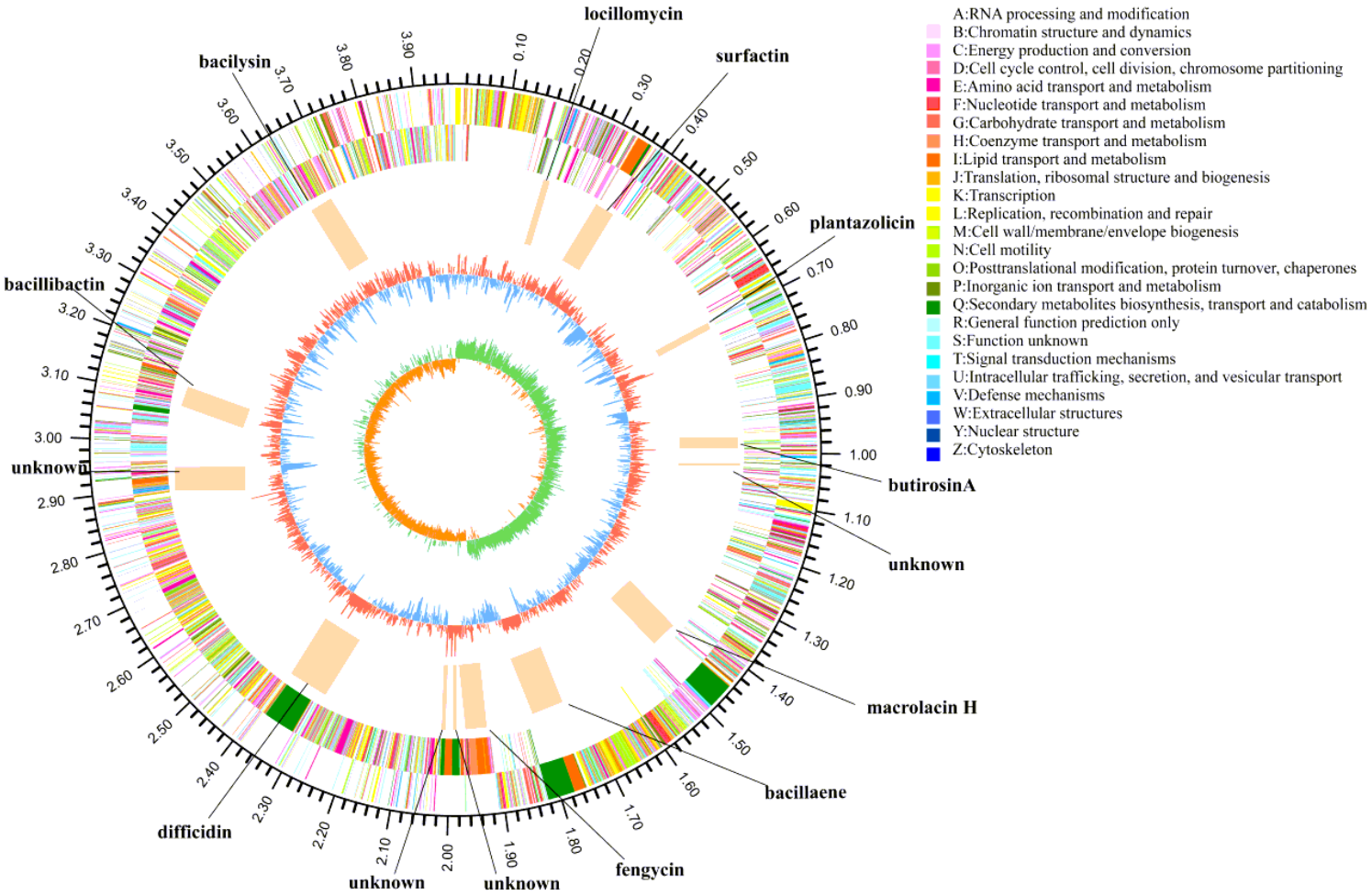

Figure 1

Circular representation of Bacillus velezensis YYC genome. First circle represents the genome size of B. velezensis YYC; Second and third circles are positive and negative chain genes, respectively, all genes are color coded according to their COG annotation functions; Fourth circle indicates nomenclature and locations of predictive secondary metabolite clusters excluding microcins; Fifth circle is GC content (red indicates greater than average value and blue indicates less than average).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- renamedac28a.docx