



ELSEVIER

Contents lists available at ScienceDirect

Data in brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)



Data Article

# Complete genome sequence data of *Flavobacterium anhuiense* strain GSE09, a volatile-producing biocontrol bacterium isolated from cucumber (*Cucumis sativus*) root



Jin-Ju Jeong <sup>a</sup>, Siti Sajidah <sup>a</sup>, Ji Yeon Oh <sup>a</sup>, Mee Kyung Sang <sup>a,b</sup>,  
Kyoung-Su Kim <sup>c</sup>, Ki Deok Kim <sup>a,\*</sup>

<sup>a</sup> Laboratory of Plant Disease and Biocontrol, Department of Biosystems and Biotechnology, Korea University, Seoul, South Korea

<sup>b</sup> Division of Agricultural Microbiology, National Academy of Agricultural Science, Rural Development Administration, Wanju, South Korea

<sup>c</sup> Division of Bioresource Sciences, Bioherb Research Institute, Kangwon National University, Chuncheon, South Korea

## ARTICLE INFO

### Article history:

Received 15 April 2019

Received in revised form 9 July 2019

Accepted 9 July 2019

Available online 15 July 2019

### Keywords:

Biocontrol

Complete genome sequence

EggNOG analysis

*Flavobacterium anhuiense*

*Phytophthora capsici*

## ABSTRACT

*Flavobacterium anhuiense* (previously identified as *Flavobacterium johnsoniae*) strain GSE09 is a volatile-producing bacterium that exhibits significant biocontrol activity against an oomycete pathogen, *Phytophthora capsici*, on pepper plants. Here, we report the complete genome sequence data of strain GSE09, isolated from surface-sterilized cucumber root. The genome consists of a circular 5,109,718-bp chromosome with a G + C content of 34.30%. A total of 4,138 complete coding sequences including 15 rRNA, 66 tRNA, 3 ncRNA, and 51 pseudogene sequences were retrieved. Thus, the genome sequence data of *F. anhuiense* GSE09 may facilitate the elucidation of many biological traits related to the biocontrol against plant pathogens.

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

\* Corresponding author.

E-mail address: [kidkim@korea.ac.kr](mailto:kidkim@korea.ac.kr) (K.D. Kim).

Specification table

Subject area	Biology
More specific subject area	Microbiology and Genomics
Type of data	Complete genome sequence data of <i>Flavobacterium anhuiense</i> GSE09
How data was acquired	Genome sequencing using PacBio RS II at Theragen Etex Bio Institute, Suwon, South Korea
Data format	Raw and analyzed data
Experimental factors	DNA was extracted from <i>F. anhuiense</i> GSE09
Experimental features	Whole genome sequencing, assembly, and annotation.
Data source location	<i>F. anhuiense</i> GSE09 was isolated from the surface-sterilized root of a cucumber plant grown in a field in Gunsan, Korea.
Data accessibility	The genome sequence of <i>F. anhuiense</i> GSE09 has been deposited in DDBJ/ENA/GenBank under the accession number CP016907 ( <a href="https://www.ncbi.nlm.nih.gov/nucleotide/CP016907">https://www.ncbi.nlm.nih.gov/nucleotide/CP016907</a> )
Related research	M.K. Sang, K.D. Kim, The volatile-producing <i>Flavobacterium johnsoniae</i> strain GSE09 shows biocontrol activity against <i>Phytophthora capsici</i> in pepper. J. Appl. Microbiol. 113 (2012) [1]

### Value of the data

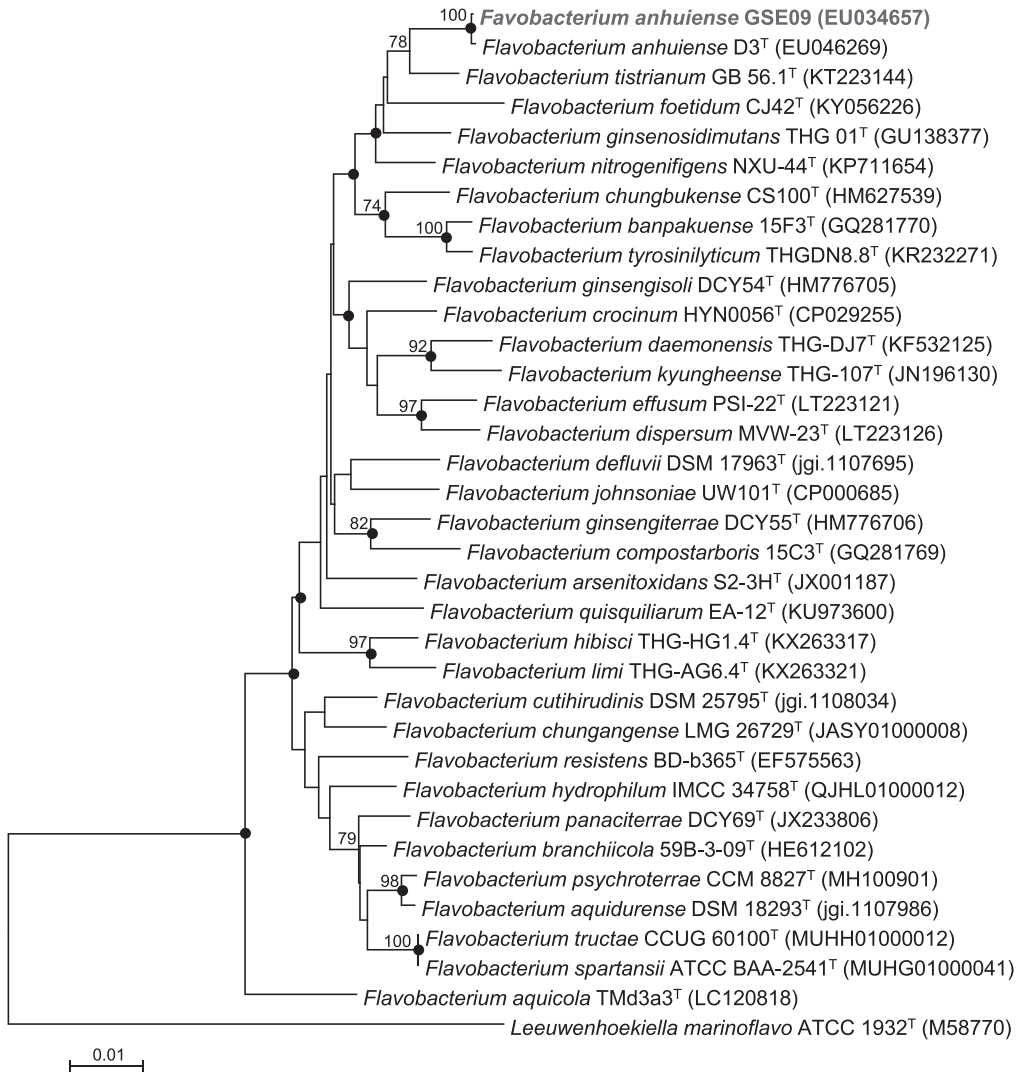
- The genome data of *F. anhuiense* GSE09 may be helpful in understanding biological traits such as plant resistance and plant growth promotion related to the biocontrol against plant pathogens.
- In *F. anhuiense* GSE09 genome, 61 defense mechanism genes and 5 cell motility genes were identified, which could possibly be significant traits of its biocontrol activity.
- The *F. anhuiense* GSE09 genome data will provide valuable information to reveal the phylogenetic diversity of the *Flavobacterium* species.

## 1. Data

Previous studies [1–3] showed that strain GSE09, isolated from the surface-sterilized root of a cucumber (*Cucumis sativus*) plant grown in a field in Gunsan, South Korea in 2002, significantly inhibited infections by a soilborne oomycete pathogen, *Phytophthora capsici*, the causal agent of Phytophthora blight of pepper. This strain produces the volatile organic compound (VOC), 2,4-di-*tert*-butylphenol, which inhibits mycelial growth, sporulation, and zoospore germination of the pathogen [1,3]. In a previous study [1], strain GSE09 was identified as *Flavobacterium johnsoniae*; however, it has now been re-identified as *Flavobacterium anhuiense* based on phylogenetic analysis of the 16S rRNA gene sequences (Fig. 1).

The genome of strain GSE09 consists of a circular 5,109,718-bp chromosome with a G + C content of 34.30%. In addition, 4,138 complete coding sequences (CDSs) were predicted, out of which 15 rRNA, 66 tRNA, 3 ncRNA, and 51 pseudogene sequences were retrieved. Moreover, 2,276 unigenes (56.46%) out of a total of 4,031 unigenes in GSE09 were assigned to 20 out of 24 functional groups based on the evolutionary genealogy of genes: non-supervised orthologous groups (eggNOG) analysis (Fig. 2). A total of 247 (6.13%), 233 (5.78%), and 218 (5.41%) unigenes were assigned to the ‘cell wall/membrane/envelope biogenesis’, ‘carbohydrate transport and metabolism’, and ‘amino acid transport and metabolisms’ groups, respectively; however, no unigenes were assigned to the ‘RNA processing and modification’, ‘nuclear structure’, ‘extracellular structures’, and ‘general function prediction’ groups. Furthermore, 61 unigenes were assigned to ‘defense mechanisms’ and 5 were assigned to ‘cell motility’, which may be related to the biocontrol traits of the strain (Fig. 2).

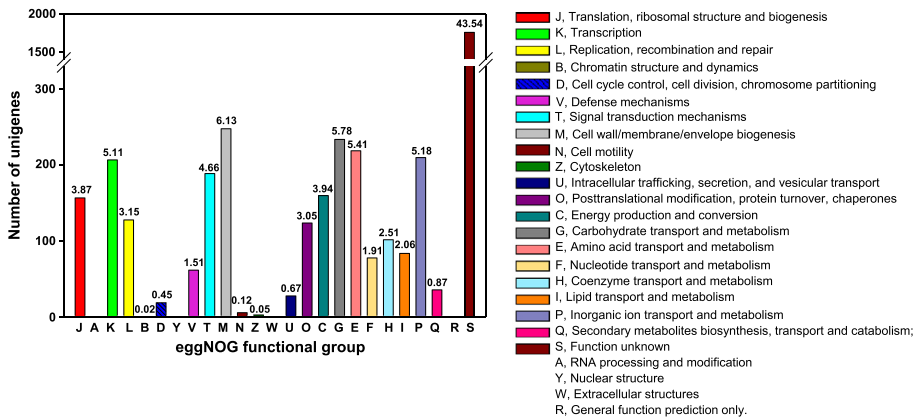
The genome of strain GSE09 also contained genes related to colonization (e.g., motility protein B, swarming motility protein SwrC, and gliding motility lipoprotein GldH precursor) [4–6] and antibacterial activity (e.g., thiazole synthase and polyketide cyclase/dehydrase and lipid transport) [7,8]. Furthermore, it contained various genes related to plant defense against plant pathogens and genes related to plant growth promotion, such as ammonia, siderophores, and indole-3-acetic acid (IAA) [9–11].



**Fig. 1.** Neighbor-joining phylogenetic tree, based on 16S rRNA gene sequences, showing the relationship between strain GSE09 (GenBank accession no. EU034657) and other species of the genus *Flavobacterium*. Filled circles on the branching points indicate that the corresponding nodes were also recovered in trees constructed using the maximum-likelihood and maximum-parsimony algorithms. Bootstrap values (≥70%) of 1000 analyses are shown at the branch points. *Leeuwenhoekiella marinoflavo* ATCC 1932<sup>T</sup> was used as an outgroup. Bar, 1 nt substitution per 100 nt of the 16S rRNA gene sequence.

## 2. Experimental design, materials and methods

Strain GSE09 was re-identified using phylogenetic analysis of the 16S rRNA gene sequence data (accession no. EU034657) in NCBI [1]. The analysis was performed with strain GSE09 and type strains of species belonging to the genus *Flavobacterium* using the molecular evolutionary genetics analysis (MEGA) version 6 software. Phylogenetic trees were constructed using neighbor-joining, maximum-likelihood, and maximum-parsimony algorithms [12–14]. The tree topology was confirmed via 1000 replications of bootstrap analysis [15].



**Fig. 2.** Number and percentage of *Flavobacterium anhuiese* strain GSE09 unigenes assigned to each functional group using the evolutionary genealogy of genes: non-supervised orthologous groups (eggNOG) analysis. The number on the bar represents the percentage of unigenes assigned to each functional group out of the total numbers of unigenes.

The genomic DNA of strain GSE09, which was prepared as described previously [1], was extracted using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocols. Genome sequencing of the strain was performed using the PacBio RS II instrument (Pacific Bioscience, Menlo Park, CA, USA) at the Theragen Exet Bio Institute (Suwon, South Korea). Single-molecule real-time (SMRT) sequencing with PacBio RS II was conducted with SMRTbell DNA template libraries that were generated using the standard 20-kb library protocols of the PacBio SMRTbell template prep kit (Pacific Bioscience). The sequence reads generated were used for *de novo* assembly using the PacBio Hierarchical Genome Assembly Process version 3 (HGAP3), which were then polished using Quiver [16]. Genome annotation of strain GSE09 was performed using the Prokaryotic Genome Annotation Pipeline (PGAP) service provided by NCBI. Furthermore, the eggNOG analysis of strain GSE09 genes was conducted using CLgenomics ver. 1.53 [17].

## Acknowledgments

Jin-Ju Jeong was supported by the Global Ph.D. program through the National Research Foundation of Korea funded by the Ministry of Education (2015-034526), South Korea. Siti Sajidah was supported by the Korean Government Scholarship Program (KGSP) during her M.S. study at Korea University, Seoul, South Korea.

## Conflict of interest

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.

## References

- [1] M.K. Sang, K.D. Kim, The volatile-producing *Flavobacterium johnsoniae* strain GSE09 shows biocontrol activity against *Phytophthora capsici* in pepper, *J. Appl. Microbiol.* 113 (2012) 383–398.
- [2] M.K. Sang, S.C. Chun, K.D. Kim, Biological control of *Phytophthora* blight of pepper by antagonistic rhizobacteria selected from a sequential screening procedure, *Biol. Control* 46 (2008) 424–433.
- [3] M.K. Sang, J.D. Kim, B.S. Kim, K.D. Kim, Root treatment with rhizobacteria antagonistic to *Phytophthora* blight affects anthracnose occurrence, ripening, and yield of pepper fruit in the plastic house and field, *Phytopathology* 101 (2011) 666–678.

- [4] Y. Luo, Y. Cheng, J. Yi, Z. Zhang, Q. Luo, D. Zhang, Y. Li, Complete genome sequence of industrial biocontrol strain *Paenibacillus polymyxa* HY96-2 and further analysis of its biocontrol mechanism, *Front. Microbiol.* 9 (2018) 1520.
- [5] D.B. Kearns, F. Chu, R. Rudner, R. Losick, Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility, *Mol. Microbiol.* 52 (2004) 357–369.
- [6] M.J. McBride, T.F. Braun, J.L. Brust, *Flavobacterium johnsoniae* GldH is a lipoprotein that is required for gliding motility and chitin utilization, *J. Bacteriol.* 185 (2003) 6648–6657.
- [7] M. Kaspady, V.K. Narayanaswamy, M. Raju, G.K. Rao, Synthesis, antibacterial activity of 2, 4-disubstituted oxazoles and thiazoles as bioisosteres, *Lett. Drug Des. Discov.* 6 (2009) 21–28.
- [8] G.F. Kremmydas, A.P. Tampakaki, D.G. Georgakopoulos, Characterization of the biocontrol activity of *Pseudomonas fluorescens* strain X reveals novel genes regulated by glucose, *PLoS One* 8 (2013) e61808.
- [9] A.M. Briones, S. Okabe, Y. Umemiya, N.B. Ramsing, W. Reichardt, H. Okuyama, Ammonia-oxidizing bacteria on root biofilms and their possible contribution to N use efficiency of different rice cultivars, *Plant Soil* 250 (2003) 335–348.
- [10] X. Yu, C. Ai, L. Xin, G. Zhou, The siderophore-producing bacterium, *Bacillus subtilis* CAS15, has a biocontrol effect on Fusarium wilt and promotes the growth of pepper, *Eur. J. Soil Biol.* 47 (2011) 138–145.
- [11] E.E. Idris, D.J. Iglesias, M. Talon, R. Borriss, Tryptophan-dependent production of indole-3-acetic acid (IAA) affects level of plant growth promotion by *Bacillus amyloliquefaciens* FZB42, *Mol. Plant Microbe Interact.* 20 (2007) 619–626.
- [12] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4 (1987) 406–425.
- [13] J. Felsenstein, Evolutionary trees from DNA sequences: a maximum likelihood approach, *J. Mol. Evol.* 17 (1981) 368–376.
- [14] W.M. Fitch, Toward defining the course of evolution: minimum change for a specific tree topology, *Syst. Zool.* 20 (1971) 406–416.
- [15] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, *Evolution* 39 (1985) 783–791.
- [16] C.S. Chin, D.H. Alexander, P. Marks, A.A. Klammer, J. Drake, C. Heiner, A. Clum, A. Copeland, J. Huddleston, E.E. Eichler, S.W. Turner, J. Korlach, Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data, *Nat. Methods* 10 (2013) 563–569.
- [17] L.J. Jensen, P. Julien, M. Kuhn, C. von Mering, J. Muller, T. Doerks, P. Bork, eggNOG: automated construction and annotation of orthologous groups of genes, *Nucleic Acids Res.* 36 (2008) D250–D254.