Korean Journal of Microbiology (2022) Vol. 58, No. 3, pp. 179-182
DOI https://doi.org/10.7845/kjm.2022.2030

pISSN 0440-2413
eISSN 2383-9902

Complete genome sequence of a highly effective diazotroph *Ensifer sesbaniae* SDT23 isolated from soybean root nodules in Vietnam

Thi Hieu Thu Nguyen, Thuy Hang Dinh, and Cao Son Trinh*

Institute of Microbiology and Biotechnology, Vietnam National University, Hanoi 122863, Vietnam

(Received May 12, 2022; Revised June 16, 2022; Accepted June 21, 2022)

The diazotroph *Ensifer sesbaniae* SDT23 was isolated from root nodules of healthy soybean plants in the Hanoi countryside. The strain possesses several agricultural beneficial characteristics, i.e. fixing nitrogen, producing Indole-3-acetic acid (IAA) and Extracellular polysaccharides (EPS), and thus has good potential in the application for organic agriculture and soil amendment. Here we show the draft genome sequence of strain SDT23 which is composed of 6,870,960 bp. with a GC content of 62.11% and approximately 6427 protein-coding sequences. The genome annotation indicated nitrogen fixation- and nodulation-related genes, along with genes involved in IAA, siderophore, and EPS biosynthesis, confirming the physiological properties of the strain.

Keywords: *Ensifer sesbaniae*, Diazotroph, draft genome, organic agriculture

The diazotrophic microorganisms perform either symbiotic (endophytic) or free-living lifestyles, converting atmospheric N_2 into usable form ammonium, supporting up to 60% of plant needs globally (Menéndez *et al*., 2017). The diazotrophs have been exploited in microbial formulas for use in agriculture, such as *Gluconacetobacter diazotrophicus*, *Azorhizobium caulinodans*, and *Bradyrhizobium* sp. (Azotic Technologies, Pivot Bio, etc.). In special cases, several diazotrophs also possess useful properties promoting plant growth such as enhancing nutrient uptake (via solubilizing K, P, and Fe species in soil) and producing phytohormones or compounds against phytopathogens (Choudhary *et al*., 2011; García-Faile *et al*., 2015).

The genus *Ensifer* (syn. *Sinorhizobium*) consists of both symbiotic and nonsymbiotic species, has been attracting attention in studying the evolution of symbiotic potential (Fagorzi *et al*., 2020). *Ensifer* species make up more than 1% of bacterial endophytes isolated, showing their potential in promoting organic agriculture. Heading to this line, the strain *E. sesbaniae* SDT23 (deposited at VTCC under accession number VTCC60024) owing several traits beneficial for organic agriculture was subjected to analysis of its whole genome.

The genomic DNA of *Ensifer sesbaniae* SDT2.3 was extracted using the E.Z.N.A Bacterial DNA kit (OMEGA D3350-01). The whole-genome sequencing of strain SDT23 was conducted at Macrogen Inc. using the Illumina Hiseq 2500-150PE system. Raw reads were processed as described previously (Chaitankar *et al*., 2016). Quality control was done by FastQC, and short and low-quality reads were trimmed by Trimmomatic (Bolger *et al*., 2014). After being quality-filtered, the *de novo* assembly was operated by SPAdes (Bankevich *et al*., 2012). Whole-genome annotation was accomplished by Prokka, and additional analyses of hypothesizing gene function were performed by Rapid Annotation using Subsystem Technology (RAST) with the SEED database (Aziz *et al*., 2008). The genome was compared against another *E. sesbaniae* genome, strain CCBAU 65729 (GCA_013283665.1), and several other *Ensifer* species, including *E. glycinis* (GCA_001651865.1), *E. aridi* (GCA_000510685.1), and *E. psoraleae* (GCA_013283645.1).

The coverage of the draft genome sequencing for strain SDT23 was 101.6×. The whole genome consists of 89 contigs with a total size of 6,870,960 bp, the G + C content of 62.11%, and approximately 6275 protein-coding sequences (Table 1). The 16S rRNA sequence of strain SDT23 is 1,345 bp long and has 100% similarity with *Ensifer sesbaniae* (OM570606). The

***For correspondence.** E-mail: caoson@vnu.edu.vn;
Tel.: +84-2437547407; Fax: +84-2437547407

Table 1. Genomic features of *E. sesbaniae* SDT23 in comparison with other *Ensifer* species

Genomic features	*E. sesbaniae* SDT23	*E. sesbaniae* CCBAU 65729	*E. glycinis* CCBAU 23380	*E. aridi* TW10	*E. psoraleae* CCBAU 65732
Genome size (bp)	6,870,960	6,897,201	6,039,294	6,802,256	7,427,611
Contigs	89	12	68	57	67
G + C content (%)	62.11	62.1	62.4	61.7	61.3
Total gene number	6,560	6,573	5,536	6,232	7,028
CDS	6,275	6,192	5,327	5,946	6,531
rRNA genes	3	3	3	3	3
tRNA genes	49	53	48	50	53
Other RNA	4	4	1	4	4
Pseudo-genes	229	321	157	229	438

whole genome sequences of *Ensifer sesbaniae* SDT23 showed high similarity to that of other *Ensifer* species, i.e. *E. aridi*, *E. psoraleae*, *E. glycinis*, and *E. sesbaniae* CCBAU 65729, with the average nucleotide identity (ANI) calculated as 80.27%, 80.21%, 81.98%, and 98.86% respectively (ANI calculator, Ezbiocloud). In particular, the high G + C content of around 62% seems to be the genus-specific characteristic of the *Ensifer* species (Table 1).

RAST analyses predicated 1692/6427 (25%) genes with functions categorized into groups in SEED viewer format. We found 11 nitrogen fixation-related genes in the nitrogen metabolism category, 5 siderophore antharachelin biosynthesis genes in the ion acquisition and metabolism category, 5 auxin biosynthesis genes in the secondary metabolism category, 4 EPS biosynthesis genes and 10 nodulation protein-encoding genes (Table 2).

The findings on the genome of strain SDT23 were intimately correlated to its physiological properties beneficial to agriculture, i.e. fixing nitrogen by forming plant root nodules, and producing IAA and EPS. The bacterium is a plant endophyte, it is expected that its colonization in plant compartments would lead to the plant growth enhancement effects that strain SDT23 displayed upon *in planta* conditions.

Nucleotide sequence accession number

The draft genome sequence of *Ensifer sesbanise* SDT23 was registered on NCBI GenBank under the accession number GCF_023108975.1 (WGS: JALMMB010000001-JALMMB010000089).

Acknowledgments

The study was supported by Grant No NVQG-2021/DT.07 from the Vietnam Ministry of Science and Technology.

Conflict of Interest

The authors have no conflict of interest to report.

References

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, *et al.* 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Sergey I, Nikolenko S, Pham S, *et al.* 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477.

Bolger AM, Lohse M, and Usadel B. 2014. Trimomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Chaitankar V, Karakülah G, Ranapriya R, Giuste FO, Brooks MJ, and Swaroop A. 2016. Next generation sequencing technology and genomewide data analysis: perspective for retinal research. *Prog. Retin. Eye Res.* **55**, 1–31.

Choudhary DK, Sharma KP, and Gaur RK. 2011. Biotechnological perspectives of microbes in agro-ecosystems. *Biotechnol. Lett.* **33**,1905–1910.

Fagorzi C, Ilie A, Decorosi F, Cangioli L, Viti C, Mengoni A, and diCenzo G. 2020. Symbiotic and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping.

Table 2. Gene loci related to the agriculture-beneficial traits of strain *E. sesbaniae* SDT23

No.	Gene	Start–Stop	Size (bp)	Strand	Predicted function
Nitrogen fixing					
1	*nifW*	1639–1187	453	-	Nitrogenase stabilizing/protective protein
2	*nifT*	2752–2522	231	-	NifT protein
3	*nifZ*	3072–2749	324	-	NifZ protein
4	*nifB*	4863–3382	1482	-	Nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly protein
5	*nifA*	6734–5088	1647	-	Nitrogenase (molybdenum-iron)-specific transcriptional regulator NifA
6	*nifX*	14567–13593	975	-	Nitrogenase FeMo-cofactor carrier protein
7	*nifN*	15973–14564	1410	-	Nitrogenase FeMo-cofactor scaffold and assembly protein
8	*nifE*	17473–15983	1491	-	Nitrogenase FeMo-cofactor scaffold and assembly protein
9	*nifH*	235–1125	891	+	Nitrogenase (molybdenum-iron) reductase and maturation protein
10	*nifX2*	283–771	489	+	NifX-associated protein
11	*nifQ*	13115–13834	720	+	Nitrogenase FeMo-cofactor synthesis molybdenum delivery protein
Nodulation					
1	*nod*	3495–5288	1793	+	Nodulation protein
2	*nodD*	5833–6492 13653–13868	659 215	+ +	Nodulation protein D (transcriptional regulator, LysR family
3	*nodD2*	4596–5534 13527–12562	939 966	+ -	Nodulation protein D2
4	*nodC*	8885–9532	647	+	Nodulation protein C
5	*nodB*	9529–10119	590	+	Nodulation protein B
6	*nodA*	10527–11150	623	+	Nodulation protein A
7	*nodN*	188.427–188.978	551	+	Nodulation protein N
8	*nolO*	3448–3561 5399–5527	113 128	+ +	Nodulation protein nolO
9	*nolB*	17626–18348	722	+	Nodulation protein nolB
10	*nolU*	16106–15480	627	-	Nodulation protein NolU
Growth phytohormone biosynthesis					
1	*APRT*	282.790–283.803	1014	+	Anthranilate phosphoribosyltransferase
2	*AAD*	379.204–380.622	1419	+	Aromatic-L-amina-acid decarboxylase
3	*TSa*	49.200–48.361	840	-	Tryptophan synthase alpha chain
4	*TSb*	50.425–49.205	1221	-	Tryptophan synthase beta chain
5	*PRAI*	51.096–50.422	675	-	Phosphoribosylanthranilate isomerase
Siderophore production					
1	*FhuF*	297.463–298.218	756	+	Ferric reductase
2	*HumS*	623.951–625.009	1059	+	Hemin transport protein
3	*HBP*	625.031–625.981	951	+	Periplasmic hemin-binding protein
4	*PP_7*	625.998–627.086	1089	+	Hemin ABC transporter, permease protein
5	*ATPb*	627.095–627.886	792	+	ABC-type hemin transport system, ATPase component
EPS production					
1	*GP*	176.288–178.825	2538	+	Glycogen phosphorylase
2	*GBr*	178.822–181.032	2211	+	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type
3	*GAT*	181.153–182.415	1263	+	Glucose-1-phosphate adenylyl-transferase
4	*GS*	182.427–183.869	1443	+	Glycogen synthase, APD-glucose trans-glucosylase

Genome Biol. Evol. **12**, 2521–2534.

García-Fraile P, Menéndez E, and Rivas R. 2015. Role of bacterial biofertilizers in agriculture and forestry. *AIMS Bioeng.* **2**,183–205.

Menéndez E, Martínez-Hidalgo P, and Silva LR. 2017. Recent advances in the active biomolecules involved in Rhizobia-Legume symbiosis, pp. 45–74. *In* Khan MS, Zaidi A, and Mussarrat J. (eds), Microbes for Legume Improvement, Springer, Cham, Switzerland.