

***Williamsia soli* sp. nov., isolated from thermal power plant in Yantai**

Ming-Jing Zhang · Xue-Han Li · Li-Yang Peng · Shuai-Ting Yun · Zhuo-Cheng
Liu · Yan-Xia Zhou*

Marine College, Shandong University, Weihai 264209, China

***Correspondence:**

Yan-Xia Zhou

Email: zhouyx@sdu.edu.cn

Table S1 General features of strain C17^T genome

Size (bp)	6072852
G+C content	64.7%
Gene number	5701
Average Gene length (bp)	972
tRNAs	46
rRNAs	6
sRNAs	2
Type III secretion system	353
Secondary metabolite	11
Terpene gene number	21
T1PKS gene number	28

Table S2 The distribution of BGCs of *Williamsia soli* C17^T based on “antiSMASH”

Gene type	Most similar biosynthetic gene cluster	Similarity(100 %)
NRPS	<i>Williamsia</i> sp. D3 Conting002	83
Ectoine	<i>Williamsia muralis</i> NBRC 105860	100
Bacteriocin	<i>Williamsia muralis</i> NBRC 105860	84
Terpene	<i>Williamsia</i> sp. D3 Conting001	89
T3pks-terpene	<i>Williamsia</i> sp. D3 Conting006	51

Table S3 Antibiotic gene clusters from CARD database

Antibiotic resistant	Antibiotic gene cluster
Fluoroquinolone	<i>Mfd</i> , <i>mfpA</i>
Tetracycline	<i>tetB</i> (P), <i>tetQ</i> , <i>tet44</i> , <i>tetI</i> , <i>tetW</i> , <i>tetS</i> , <i>tetM</i> , <i>tetO</i> , <i>otr</i> (A), <i>tet36</i> , <i>tet32</i>
Rifamycin	<i>RbpA</i> , <i>clbC</i> , <i>clbB</i> , <i>clbA</i> , <i>cipA</i> , <i>cfrA</i> , <i>cfiC</i> , <i>rphB</i> , <i>iri</i> , <i>rphA</i> , <i>arr-8</i> , <i>arr-7</i> , <i>arr-5</i> , <i>arr-4</i> , <i>arr-3</i> , <i>arr-2</i> , <i>arr-1</i>
Aminoglycoside	<i>Pseudomonas aeruginosa</i> <i>CpxR</i> , <i>baeR</i> , <i>adeR</i> , <i>vanRA</i> , <i>vanRL</i> , <i>arR</i> , <i>Escherichia coli</i> <i>CpxR</i> , <i>vanRM</i> , <i>vanRI</i> , <i>vanRB</i> , <i>vanRC</i> , <i>vanRD</i> , <i>vanRE</i> , <i>vanRF</i> , <i>vanRG</i> , <i>mtrA</i> , <i>kdpE</i> , <i>smeR</i> , <i>vanRN</i> , <i>vanRO</i> , <i>AAC</i> (2')-Ia, <i>AAC</i> (2')-Ic, <i>AAC</i> (2')-Ib, <i>AAC</i> (2')-Id
Glycopeptide	<i>Pseudomonas aeruginosa</i> <i>CpxR</i> , <i>baeR</i> , <i>adeR</i> , <i>vanRA</i> , <i>vanRL</i> , <i>arR</i> , <i>Escherichia coli</i> <i>CpxR</i> , <i>vanRM</i> , <i>vanRI</i> , <i>vanRB</i> , <i>vanRC</i> , <i>vanRD</i> , <i>vanRE</i> , <i>vanRF</i> , <i>vanRG</i> , <i>mtrA</i> , <i>kdpE</i> , <i>smeR</i> , <i>vanRN</i> , <i>vanRO</i> , <i>vanHA</i> , <i>vanHF</i> , <i>vanHM</i> , <i>vanTG</i> , <i>vanTE</i> , <i>vanTC</i> , <i>vanTrL</i> , <i>vanA</i> , <i>vanC</i> , <i>vanB</i> , <i>vanE</i> , <i>vanD</i> , <i>vanG</i> , <i>vanF</i> , <i>vanM</i> , <i>vanL</i> , <i>vanO</i> , <i>vanN</i>
Macrolide	<i>myrA</i> , <i>tlrB</i> , <i>chrB</i>
Peptide	<i>tsnr</i>
Phenicol	<i>cml</i> , <i>cmlv</i> , <i>cmrA</i> , <i>cmr</i> , <i>cmx</i>
Diaminopyrimidine	<i>suI2</i> , <i>suI3</i> , <i>dfiE</i>

Table S4 ANI, dDDH and AAI values between strain C17^T and neighbour type strains

Reference genome (Accession no.)	ANI (%)	dDDH (%)	AAI (%)
<i>Williamsia soli</i> C17 ^T (JAFRDR000000000)	100	100	100
<i>Williamsia faeni</i> N1350 ^T (QJF000000000)	77.5	20.7	77.9
<i>Williamsia limnetica</i> L1505 ^T (QJSP000000000)	86.1	30.7	89.9
<i>Williamsia muralis</i> CIP 106979^T (RBKV000000000)	77.2	20.3	78.8
<i>Williamsia marianensis</i> MT8 ^T (QEOM000000000)	77.2	20.3	79.1
<i>Williamsia herbipolensis</i> ARP1 ^T (JXYP000000000)	73.6	19.1	65.8
<i>Gordonia soli</i> CC-AB07 ^T (BANX000000000)	72.9	20.3	64.1
<i>Gordonia oryzae</i> RS15-1S ^T (RKMH000000000)	72.5	19.8	63.1

*Strains in bold letters indicate type species of each genus.

Fig. S1 Maximum-likelihood phylogenetic tree based on 16S rRNA gene sequences of strain C17^T and high similarity of genus. Bootstrap support values (1000 replications) above 50% are shown at nodes. The filled circles indicate that the corresponding node is also restored in the maximum likelihood tree. *Absiella argi* N6H1-5^T was used as an outgroup. Bar, 0.05 substitutions per nucleotide position.

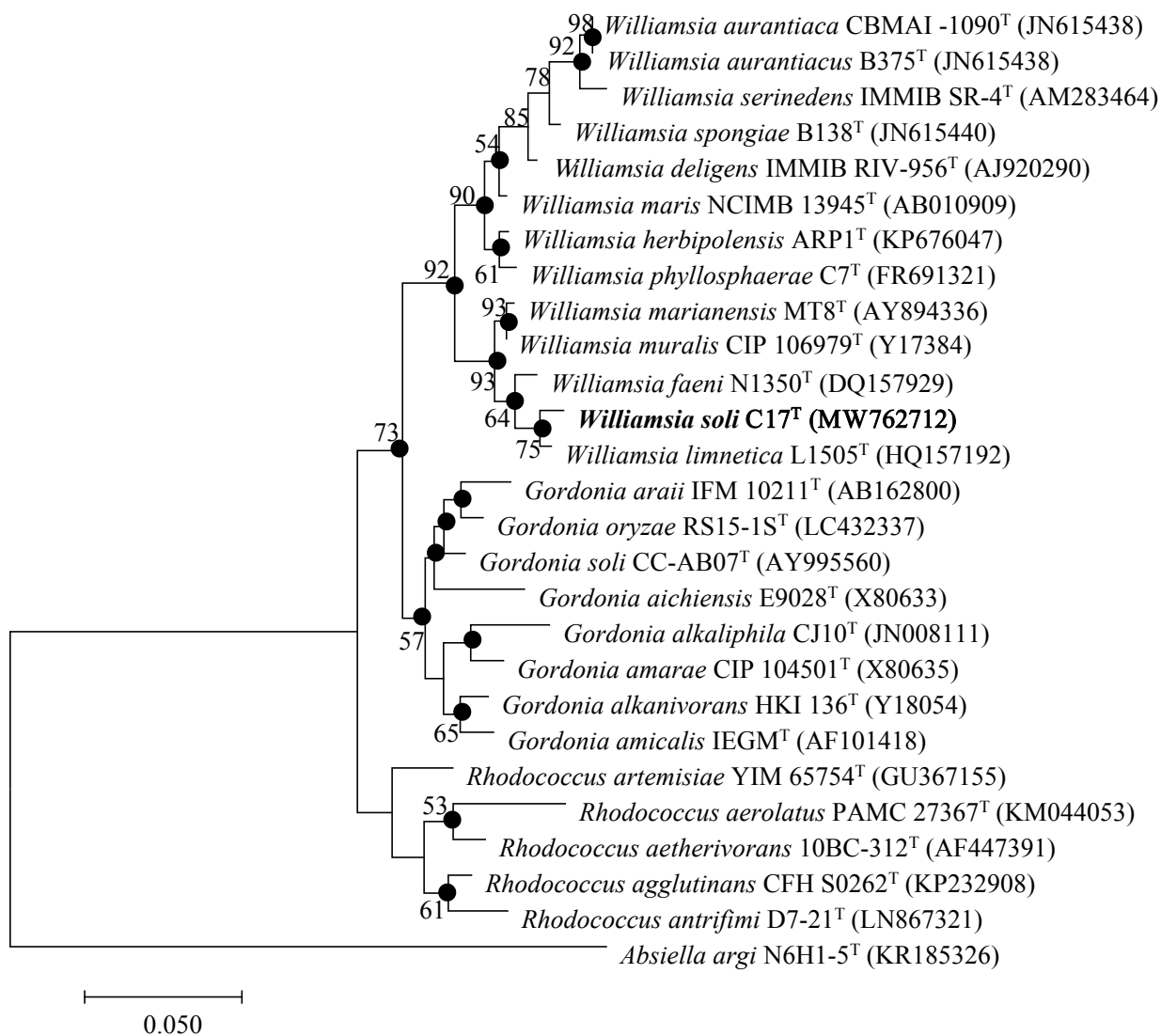


Fig. S2 After incubating the bacterium on LB at 25 °C for 3 days, cell morphology of strain C17^T was observed by an electron microscope. Bar, 1 μ m.

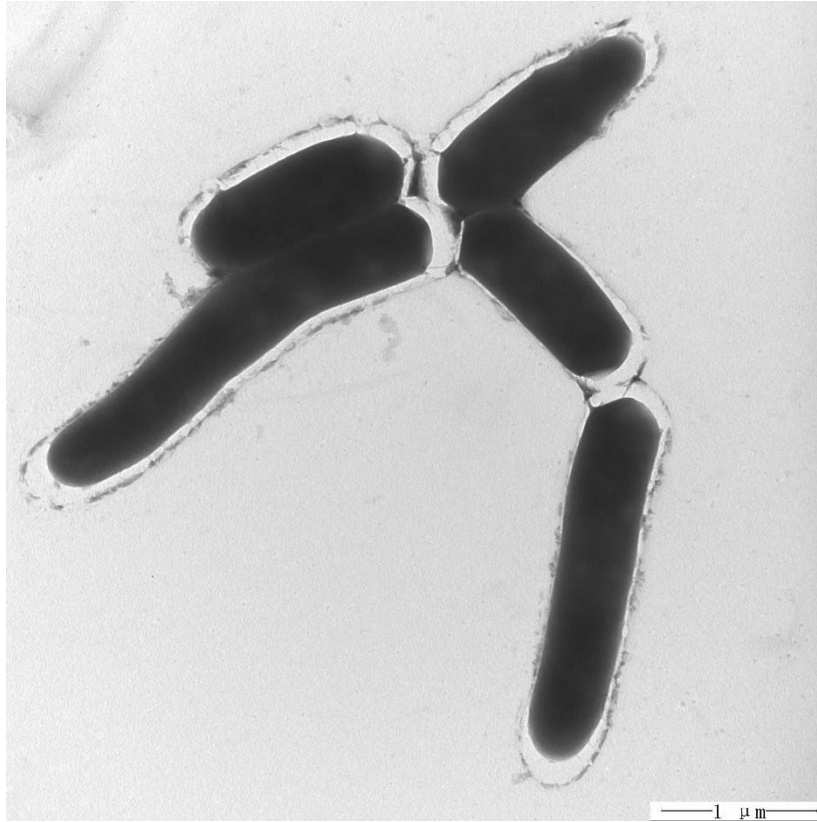


Fig. S3 Two-dimensional Thin Layer Chromatography (TLC) pattern of the total polar lipids from strain C17^T.

The plate is sprayed with 5% molybdato-phosphoric acid to show all lipids.

L, unidentified lipid; PE, phosphatidylethanolamine; DPG, diphosphatidylglycerol; PI, phosphatidylinositol; GL, unidentified glycolipid.

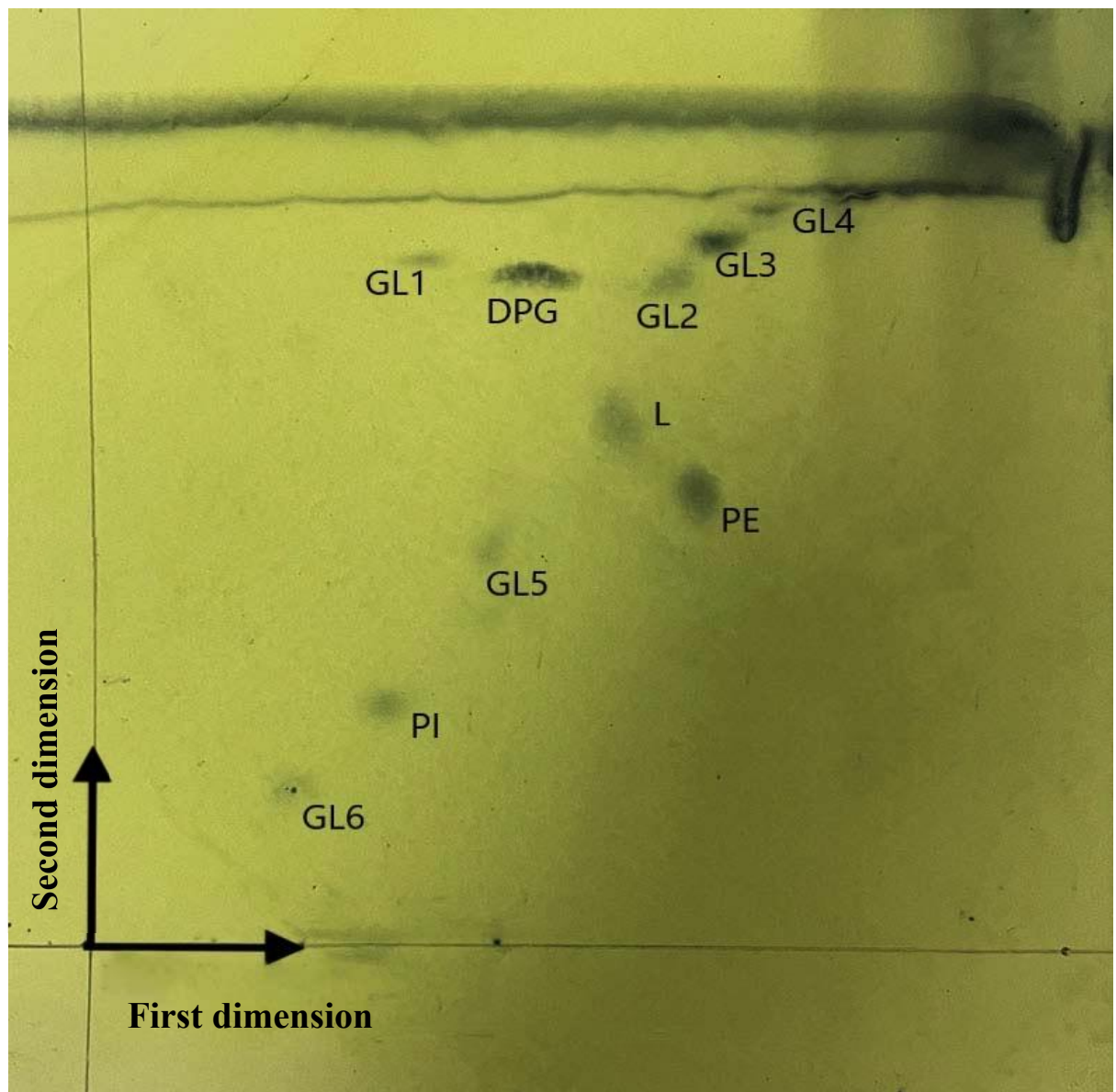


Fig. S4 Predicted gene length profiles

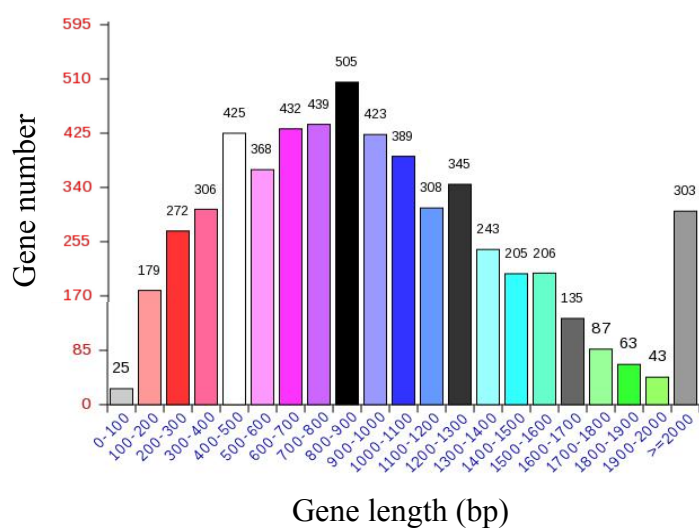


Fig. S5 Statistical chart of COG classification

