

**Halotolerant *Gordonia* sp. DH2 Enhances Hydrocarbon Degradation and
Emulsification in High-Salinity Oil Reservoirs through cell-bound Biosurfactant
Production**

**Ji Gao¹, Yu Duan¹, Ting Xu², Ding Dong³, Dehao Lu⁴, Shuyuan Deng⁵, Jian Fu⁵,
Fan Zhang⁶, Hao Dong^{1,*}, Yuehui She^{5,*}**

¹ Hubei Engineering Research Centers for Clean Production and Pollution Control of Oil and Gas Fields, College of Chemistry and Environmental Engineering, Yangtze University, Jingzhou 434023, China

² College of Resources and Environment, Yangtze University, Wuhan, Hubei 430010, China

³ Karamay Xinaoda Petroleum Technological Co., Ltd, Xinjiang, 83400, China

⁴ Xinjiang Oilfield Company PetroChina, Karamay, Xinjiang, 834000, China

⁵ College of Petroleum Engineering, Yangtze University, Wuhan, Hubei 430010, China

⁶ The Key Laboratory of Marine Reservoir Evolution and Hydrocarbon Accumulation Mechanism, Ministry of Education, College of Energy Resources, China University of Geosciences (Beijing), Beijing 100083, China

*** Correspondence:** *Corresponding author.

E-mail: dong_hao2005@163.com; sheyuehui@163.com

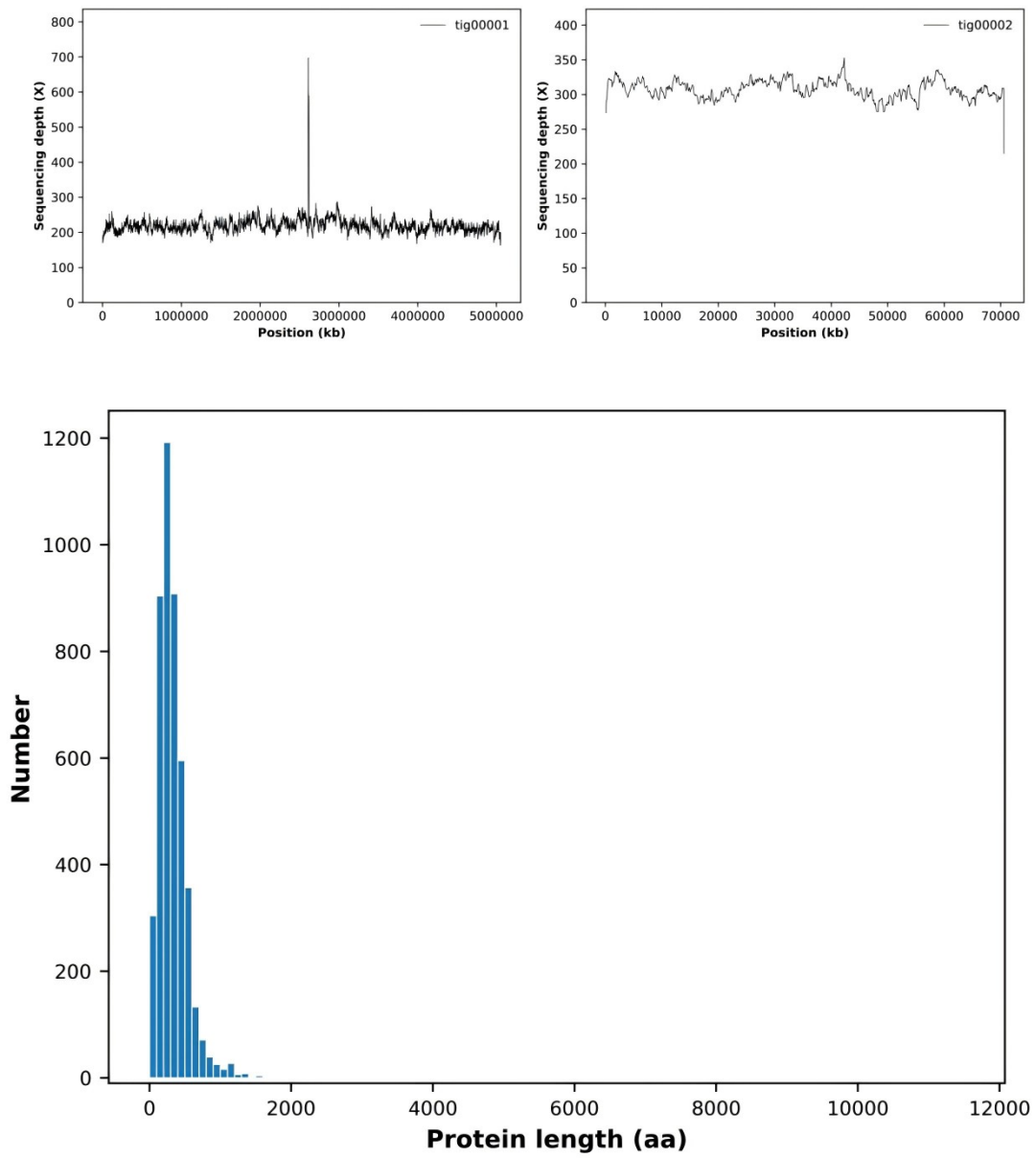


Fig. S1. (a) Distribution of genome sequencing depth (window=1000 bp); (b) Distribution of encoded protein lengths (window=100 aa).

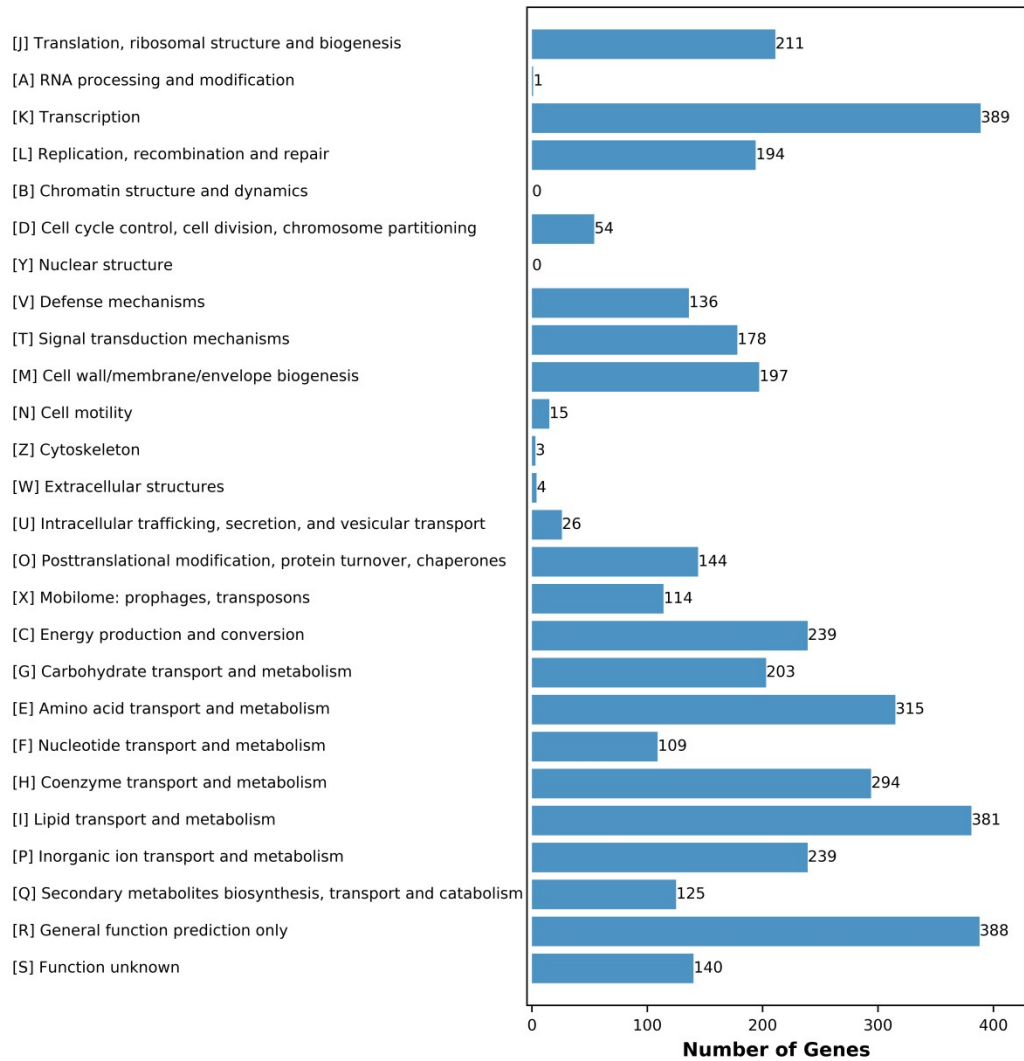


Fig. S2. COG functional classification statistics of encoded proteins in the genome.

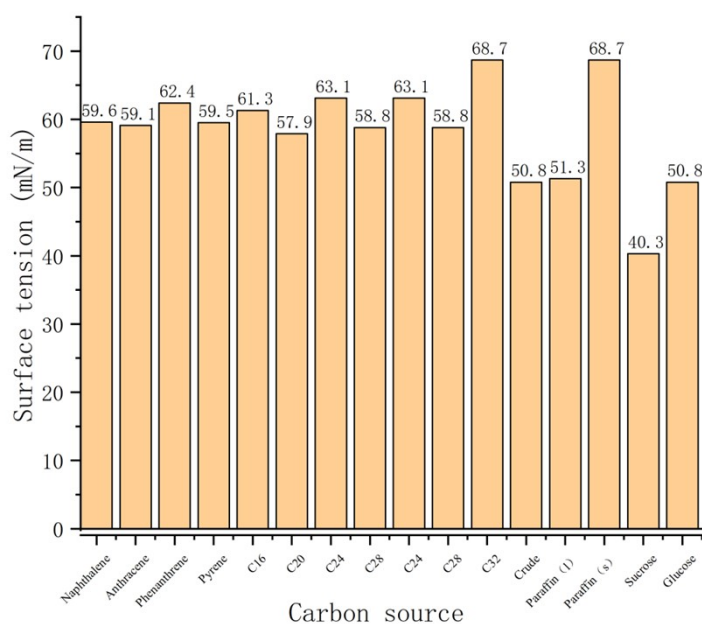


Fig. S3 Surface Tensions of Different Carbon Sources

The genome consists of a circular chromosome of 5,054,758 bp with a GC content of 67.38%, containing 4,604 predicted coding sequences (CDSs) (Fig S1b), 51 tRNAs, and 12 rRNAs (4 copies of 16S rRNA, 4 copies of 23S rRNA, and 4 copies of 5S rRNA) (Table S1). Additionally, the genome includes a 70,604 bp plasmid. In Fig S2, 389 genes are involved in transcription, 239 genes are related to energy production and conversion, and 315 genes are allocated for amino acid transport and metabolism. These findings suggest the high metabolic capacity of *Gordonia* sp. DH2, its potential to thrive in extreme environments, and its capability for biosurfactants production.

Tab. S1. Statistics of genomic structure prediction.

Type	Number	Length (bp)	% genome
tRNA	51	3,879	0.08
rRNA (16S)	4	6,090	0.12
rRNA (23S)	4	12,540	0.24
rRNA (5S)	4	412	0.01
CDS	4,604	4,671,234	91.14
CRISPR	0	0	0
genomic_island	15	412,240	8.04
prophage_region	2	52,032	1.02

Tab. S2 Halotolerance genes within the genome of strain DH2.

Gene	Function
<i>nhaA</i>	Na ⁺ /H ⁺ antiporter NhaA
<i>nhaP</i>	NhaP-type Na ⁺ /H ⁺ or K ⁺ /H ⁺ antiporter
<i>kefB</i>	Kef-type K ⁺ transport system%2C membrane component KefB
<i>plsC</i>	Major Facilitator Superfamily
<i>sulP</i>	Sulfate permease or related transporter%2C MFS superfamily
<i>gluP</i>	Major Facilitator Superfamily
<i>fucP</i>	Major Facilitator Superfamily
<i>gph</i>	MFS/sugar transport protein
<i>sulI</i>	Sulfate permease or related transporter%2C MFS superfamily
<i>araJ</i>	Major Facilitator Superfamily
<i>trkH</i>	Trk-type K ⁺ transport system%2C membrane component
<i>trkA</i>	Ion transport protein
<i>nhaP</i>	NhaP-type Na ⁺ /H ⁺ and K ⁺ /H ⁺ antiporter with C-terminal TrkAC and CorC domains
<i>kdpD</i>	K ⁺ -sensing histidine kinase KdpD
<i>gltD</i>	NADPH-dependent glutamate synthase beta chain or related oxidoreductase
<i>gltB1</i>	Glutamate synthase central domain
<i>betB</i>	Choline-glycine betaine transporter
<i>betA</i>	Choline dehydrogenase or related flavoprotein
<i>glnA</i>	Glutamine synthetase
<i>etcA</i>	Synthesis of ectoine
<i>etcB</i>	Synthesis of ectoine
<i>etcC</i>	Synthesis of ectoine
<i>proB</i>	Synthesis of proline
<i>metlD</i>	Synthesis of mannitol
<i>CAT1</i>	Decomposition of H ₂ O ₂
<i>HSP70</i>	Prevention of protein denaturation under salt stress
<i>AQP</i>	Regulation of water permeability
<i>MscS</i>	Osmotic shock protection

Tab. S3 Biosurfactant-producing genes and gene clusters within the genome of strain DH2.

<i>Gene</i>	<i>Function</i>
<i>EntF</i>	<i>Non-ribosomal peptide synthetase</i>
<i>EntF2</i>	<i>Non-ribosomal peptide synthetase</i>
<i>PapA3</i>	<i>PKS-related acyltransferase</i>
<i>SrfAC</i>	<i>Surfactin synthetase complex component</i>
<i>FenE</i>	<i>Fengycin synthetase .</i>
<i>RfaB</i>	<i>Glycosyltransferase involved in lipopolysaccharide core synthesis</i>
<i>WcaA</i>	<i>Polyprenol phosphate mannose synthase, catalyzing mannosyl transfer in glycolipids</i>
<i>wax</i>	<i>Trehalose dimycolate (TDM) synthase, producing trehalose lipid surfactants</i>
<i>YjhB</i>	<i>Mannosyltransferase, involved in glycosyl modification of glycolipids</i>
<i>GlfT1</i>	<i>Galactosyltransferase involved in galactolipid synthesis</i>
<i>GlfT2</i>	<i>Galactosyltransferase involved in galactolipid synthesis</i>
<i>FabG</i>	<i>Providing precursors of fatty acid chains for surfactants .</i>
<i>AcpP</i>	<i>Carrying fatty acid chains for glycolipid synthesis</i>
<i>FadD</i>	<i>Activation of fatty acid chains</i>
<i>ComP</i>	<i>Regulation of biosurfactant synthesis gene cluster expression</i>
<i>DrrA</i>	<i>Involvement in biosurfactant secretion</i>
<i>DrrB</i>	<i>Involvement in biosurfactant secretion</i>

Tab. S4 Hydrocarbon-degrading genes within the genome of strain DH2.

<i>Gene</i>	<i>Function</i>
<i>alkB</i>	<i>Catalytic terminal hydroxylation of alkanes</i>
<i>LadA</i>	<i>Degradation of long-chain alkanes</i>
<i>AlmA</i>	<i>Degradation of long-chain alkanes</i>
<i>EthA</i>	<i>Catalytic initial oxidation of short-chain alkanes</i>
<i>EthB</i>	<i>Catalytic initial oxidation of short-chain alkanes</i>
<i>BenA</i>	<i>Catalytic dihydroxylation of benzene ring</i>
<i>BenB</i>	<i>Catalytic dihydroxylation of benzene ring</i>
<i>XylA</i>	<i>Participation in toluene/xylenes metabolic pathway</i>
<i>XylB</i>	<i>Participation in toluene/xylenes metabolic pathway</i>
<i>DmpK</i>	<i>Degradation of phenol and its derivatives</i>
<i>DmpB</i>	<i>Degradation of phenol and its derivatives</i>
<i>LeuA</i>	<i>Degradation of phenol and its derivatives</i>
<i>mhpF</i>	<i>Degradation of phenol and its derivatives</i>
<i>dmpG</i>	<i>Degradation of phenol and its derivatives</i>
<i>mhpE</i>	<i>Degradation of phenol and its derivatives</i>
<i>PhnA</i>	<i>Catalytic initial oxidation of naphthalene, phenanthrene</i>
<i>PhnB</i>	<i>Catalytic initial oxidation of naphthalene, phenanthrene</i>
<i>NidA</i>	<i>Pyrene-degrading dioxygenase</i>
<i>NidB</i>	<i>Pyrene-degrading dioxygenase</i>
<i>NirB</i>	<i>Pyrene-degrading dioxygenase</i>
<i>hcaD</i>	<i>Pyrene-degrading dioxygenase</i>
<i>PaaJ</i>	<i>Degradation of phenylacetic acid</i>
<i>PaaD</i>	<i>Degradation of phenylacetic acid</i>
<i>paaF</i>	<i>Degradation of phenylacetic acid</i>
<i>PaaK</i>	<i>Degradation of phenylacetic acid</i>

<i>beta</i>	<i>Degradation of phenylacetic acid</i>
<i>Fdx</i>	<i>Degradation of phenylacetic acid</i>
<i>FadD</i>	<i>Participation in β-oxidation of alkane derivatives</i>
<i>FadL</i>	<i>Participation in β-oxidation of alkane derivatives</i>
<i>Cypx</i>	<i>Oxidation of both aromatic and aliphatic hydrocarbons</i>
<i>AdhE</i>	<i>Oxidation of alkane-derived primary alcohols to carboxylic acids</i>

Tab. S5 Metabolites of dibenzothiophene detected by GC-MS.

Product peak	Retention time (min)	MW and m/z	Metabolic intermediates
P1	8.99	136 (136,91,65)	Phenylacetic acid
P2	25.61	184 (184, 158,139,92)	DBT
P3	28.08	198 (198,165,151,98)	4-MDBT
