

Supplementary Materials

Genomic Analysis of the Halotolerant Hydrocarbon-Oxidizing Bacterium *Ectopseudomonas guguanensis* G3 From a Petroleum Reservoir

Alexey P. Ershov *, Tatyana P. Tourova, Diyana S. Sokolova, Ekaterina M. Semenova and Tamara N. Nazina

Winogradsky Institute of Microbiology, Research Center of Biotechnology, Russian Academy of Sciences, 119071 Moscow, Russia; e.alexey.mail@yandex.ru (A.P.E.); tptour@rambler.ru (T.P.T.); sokolovadiyana@gmail.com (D.S.S.); semenova_inmi@mail.ru (E.M.S.); nazina@inmi.ru (T.N.N.)

* Correspondence: e.alexey.mail@yandex.ru; Tel.: +7-499-135-03-41 (A.P.E.)

This file includes:

Figures S1 to S11

Table S1



Figure S1. Cell morphology of strain G3 grown in TEG medium at 30 °C for 5 days. The samples were examined under a scanning electron microscope (Quattro S, Thermo Fisher Scientific, Brno Černovice, Czech Republic) at an accelerating voltage of 15 kV. Bar, 5 μm.

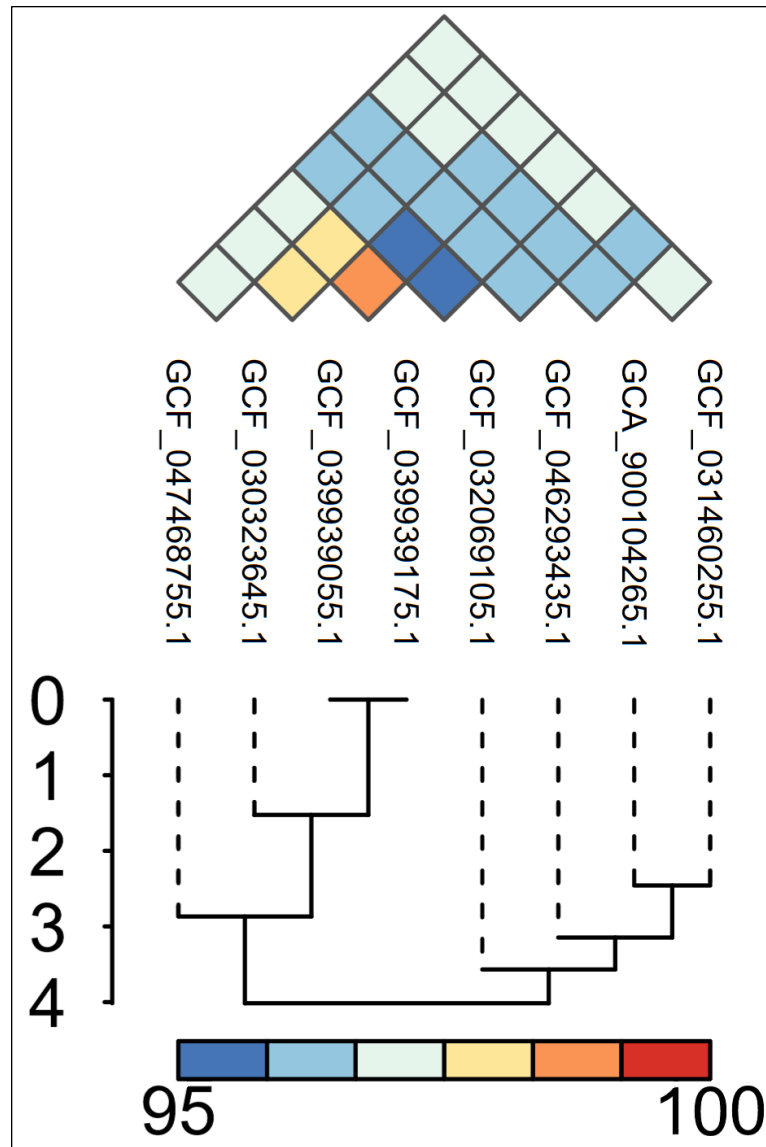


Figure S2. ANI matrix (%) of the *E. guguanensis* strain G3 and phylogenetically closely related whole-genome sequences from GenBank and their cladogram.

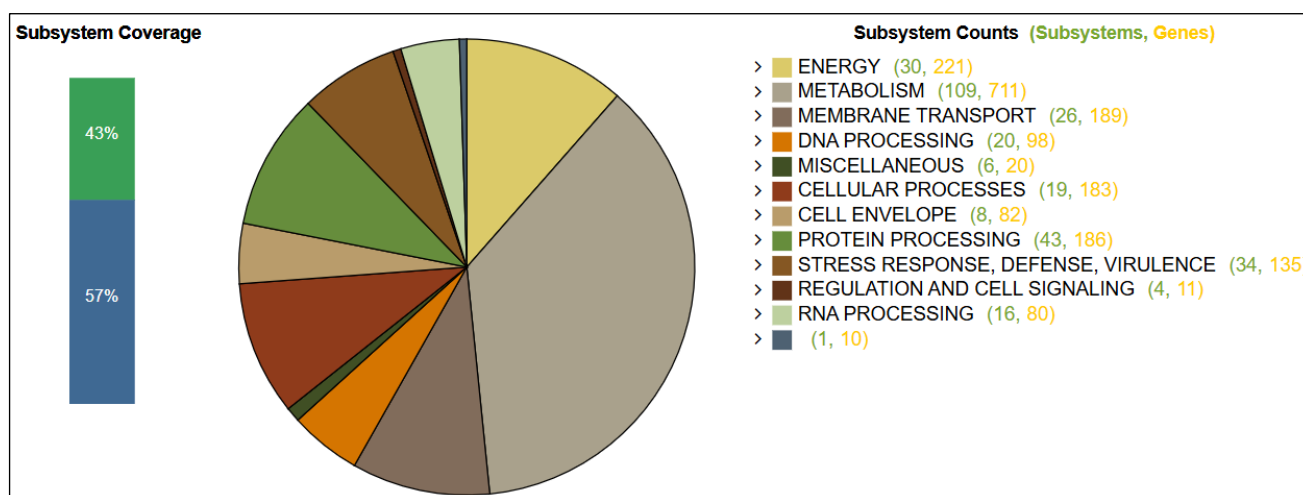


Figure S3. Distribution of genes in the genome of the *E. guguanensis* strain G3 among subsystem super classes and coverage of the genes in the subsystems.

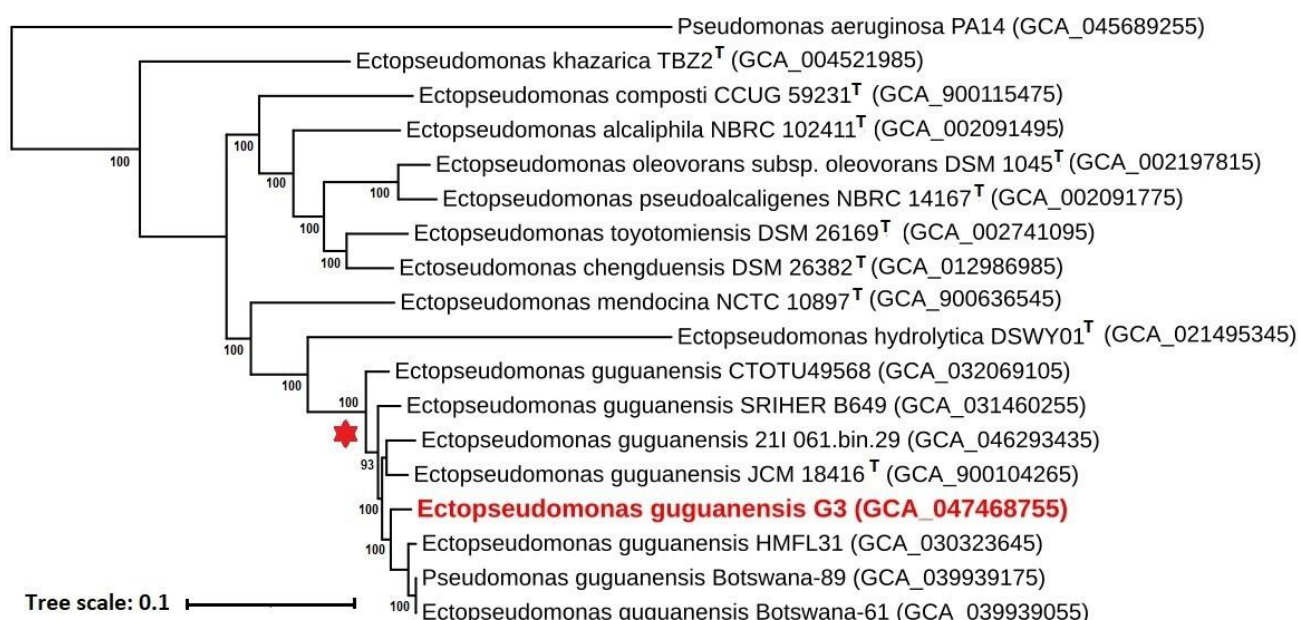


Figure S4. The maximum-likelihood phylogenetic tree derived from 500 single-copy proteins shows the position of strain G3 within the genus *Ectopseudomonas*. Bar: 0.1 amino acid substitutions per site. Bootstrap values >90 are listed as percentages at the branching points. The tree was rooted using *Pseudomonas aeruginosa* PA14 (GCA_045689255) as the outgroup. Accession numbers for the genomic assemblies are indicated in brackets. The name of the studied strain is marked by red boldface. The cluster of strains of *E. guguanensis* species is marked by a red asterisk.

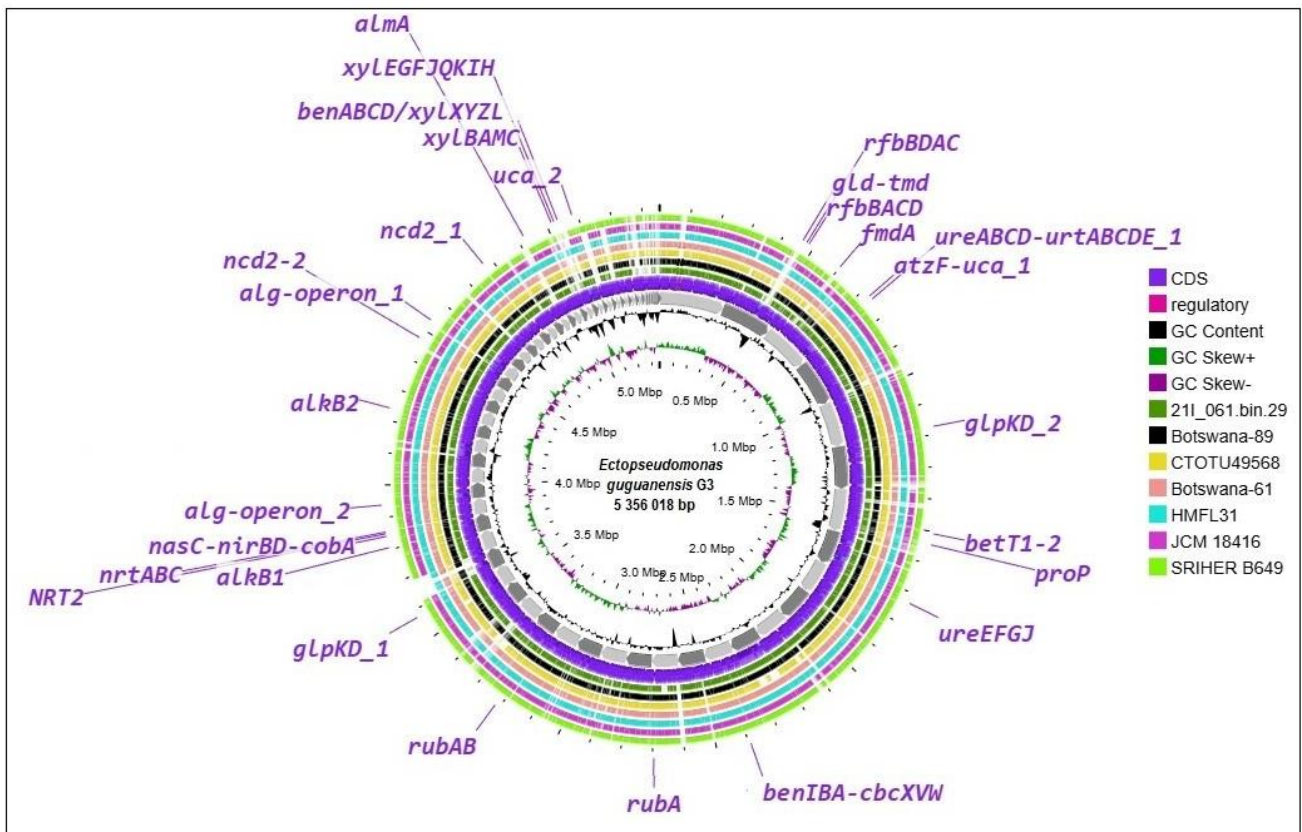


Figure S5. The circular genome map of strain G3 with BLAST comparison with genomes of the of seven other *E. guguanensis* strains. Abbreviations for genes: *rfbBACD*, *rfbBDAC*, clusters of dTDP-L-rhamnose biosynthesis; *glt-tmd*, the tandem of GDP-6-deoxy-D-talose biosynthesis; *fmdA*, formamidase; *ureABCD-urtABCDE-1*, urease operon I; *uca-atzF*, the tandem urea carboxylase and allophanate hydrolase; *urtABCDE*, urea transmembrane transporters; *glpKD_1*, *glpKD_2*, glycerol kinase and glycerol-3-phosphate dehydrogenase; *betT1-2*, choline transporter; *proP*, L-proline/glycine betaine transporter; *ureEFGJ*, operon II of accessory urease genes; *benIBA-cbcXVW*, choline dehydrogenase and choline/glycine betaine ABC transporter; *rubA*, *rubAB*, rubredoxin and rubredoxin reductase; *alkB1*, *alkB2*, *almA*, alkane 1-monooxygenase; *NRT2*, *nrtABC* nitrate/nitrite transporters; *nasC-nirBD-cobA*, the assimilatory nitrate reductase, the assimilatory nitrite reductase, and uroporphyrin-III C-methyltransferase; *alg-operon_1* and *alg-operon_2*, two putative alginate biosynthesis operons; *ncd2_1*, *ncd2_2*, nitronate monooxygenase; *benABCD/xyLXYZL*, benzoate/toluene 1,2-dioxygenase and dihydroxycyclohexadiene carboxylate dehydrogenase; *xyLBAMC*, benzyl alcohol dehydrogenase, toluene methyl-monooxygenase, and benzaldehyde dehydrogenase; *xyLEGFJQKIH*, the operon of the catechol *meta*-cleavage pathway; *uce_2*, urea carboxylase.

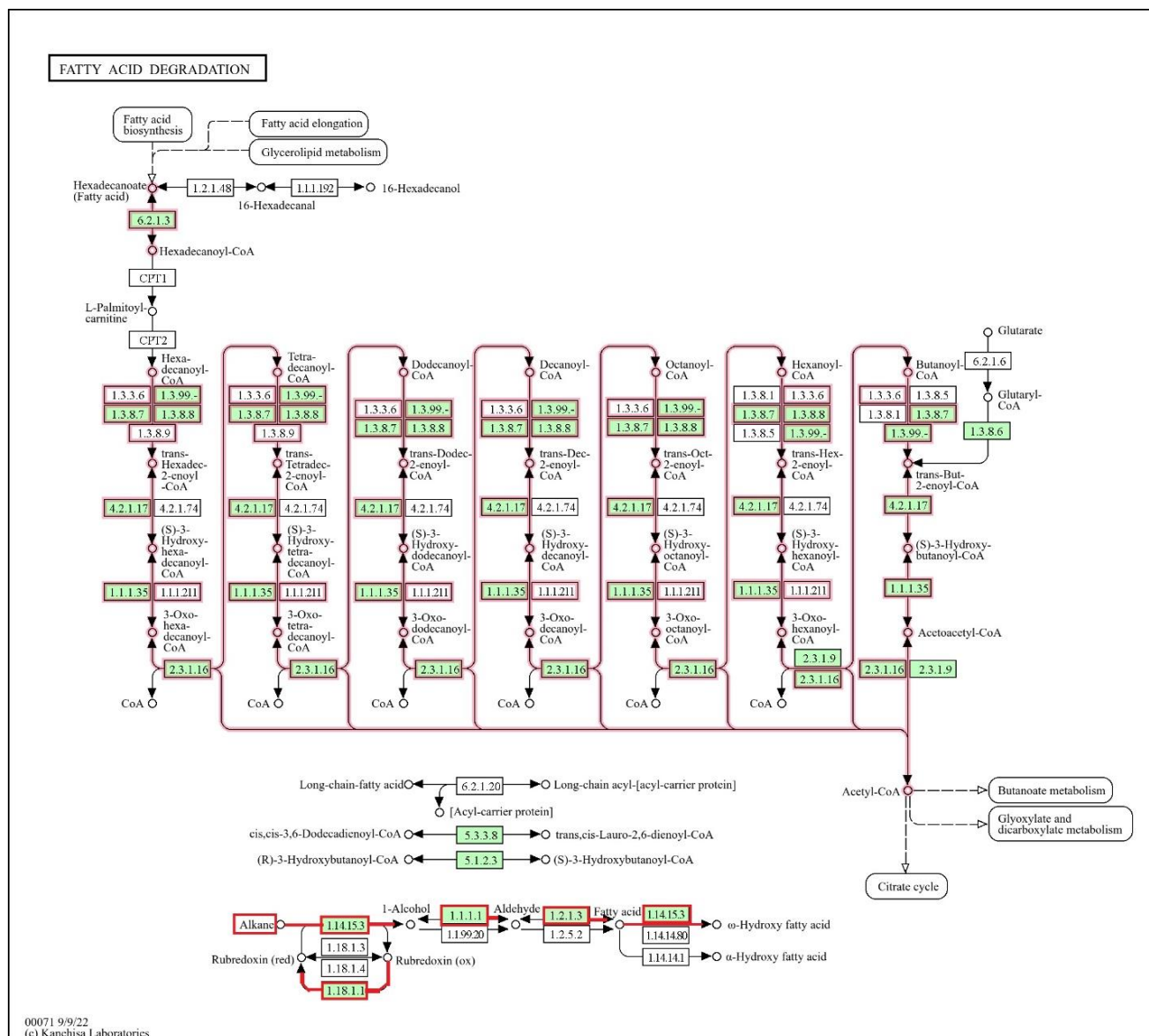


Figure S6. KEGG map of fatty acid degradation pathways based on the genome analysis of strain G3. The enzymes encoded by the genes annotated in the genome are highlighted in green. The presumptive modules for fatty acid beta-oxidation are highlighted in pink and for *n*-alkane oxidation in red.

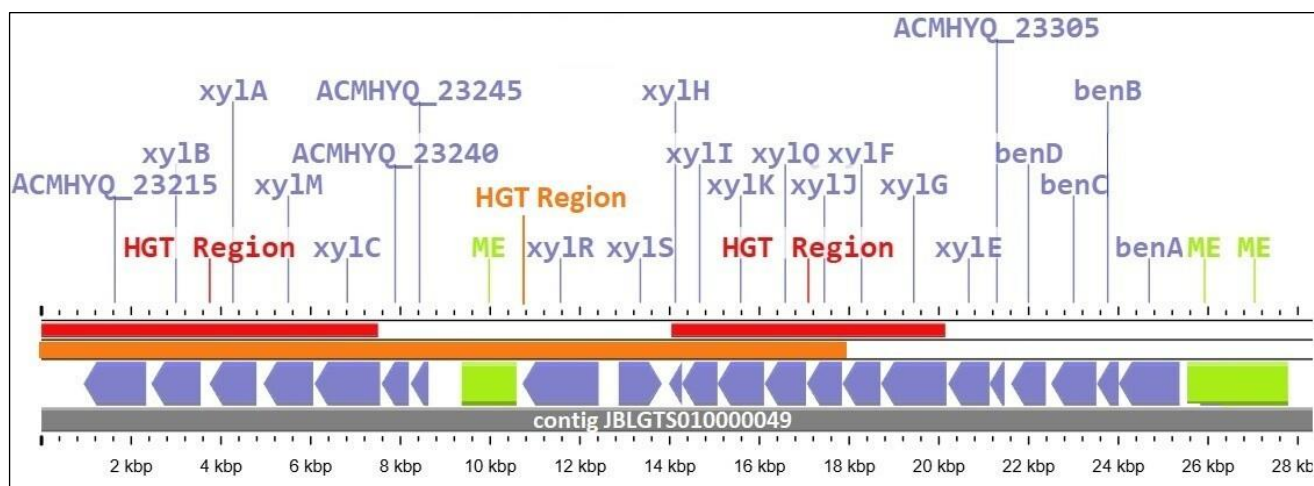


Figure S8. Localization of benzoate, catechol, and xylene degradation genes in the G3 genome. Genes of mobile elements (ME) are highlighted in green. HGT (horizontal gene transfer) regions are shown according to the results of the Alien Hunter module of the Proksee server (red) and the IslandViewer server (orange). Abbreviations: *benA*, benzoate 1,2-dioxygenase alpha subunit; *benB*, benzoate 1,2-dioxygenase beta subunit; *benC*, benzoate 1,2-dioxygenase, ferredoxin reductase component; *benD*, 1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase; *xyiE*, catechol 2,3-dioxygenase; *xyiG*, 2-hydroxymuconate-6-semialdehyde dehydrogenase; *xyiF*, 2-hydroxymuconic semialdehyde hydrolase; *xyiJ*, 2-hydroxyhexa-2,4-dienoate hydratase; *xyiQ*, acetaldehyde dehydrogenase; *xyiK*, 4-hydroxy-2-oxovalerate aldolase; *xyiI*, 4-oxalocrotonate decarboxylase; *xyiH*, 2-hydroxymuconate tautomerase; *xyiS*, transcriptional activator protein; *xyiR*, transcriptional regulator protein; *xyiC*, benzaldehyde dehydrogenase (NAD); *xyiM*, toluene methyl-monooxygenase; *xyiA*, toluene methyl-monooxygenase electron transfer component; *xyiB*, aryl-alcohol dehydrogenase.

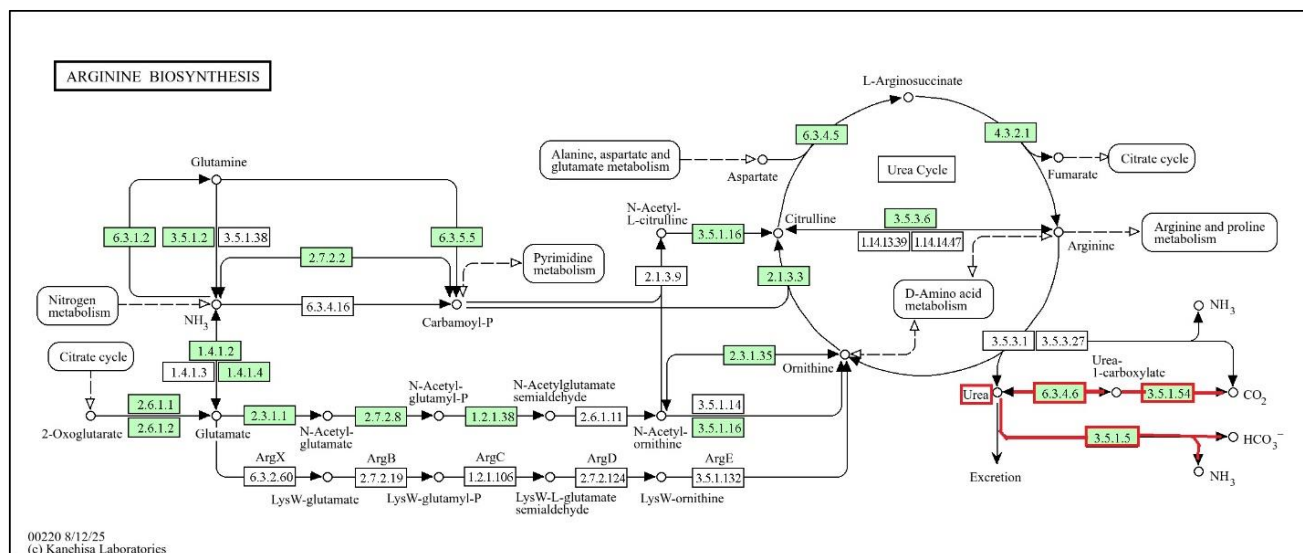


Figure S11. KEGG map of arginine biosynthesis pathways based on the genome analysis of strain G3. The enzymes encoded by the genes annotated in the genome are highlighted in green. The presumptive modules for urea degradation are highlighted in red.

Table S1. Comparison of functional genes of alkane oxidation and osmoprotectant metabolism of strain G3 and phylogenetically closely related strains/MAGs of the *E. guguanensis* species.

Strain/MAG*	Genome assembly	Status	Isolation source	Geographic location	Annotated functional genes
21I_061.bin.29*	GCF_046293435.1	WGS	Hot spring water	China: Tibet	<i>alkB2</i> , <i>benABCD</i> , <i>betAB</i>
Botswana-61	GCF_039939055.1	WGS	Portable suction machine	Botswana: Gaborone	<i>alkB1</i> , <i>alkB2</i> , <i>betAB</i>
Botswana-89	GCF_039939175.1	WGS	Sink drain	Botswana: Gaborone	<i>alkB1</i> , <i>alkB2</i> , <i>betAB</i>
CC-G9A ^T	GCA_900104265.1	WGS	Hot spring water	Thailand	<i>alkB2</i> , <i>betAB</i>
CTOTU49568*	GCF_032069105.1	WGS	Urban	USA	<i>alkB2</i> , <i>benABCD</i> , <i>betAB</i>
G3	GCF_047468755.1	WGS	Oil reservoir	Kazakhstan	<i>alkB1</i>, <i>alkB2</i>, <i>benABCD</i>, <i>betAB</i>, <i>xylCMAB</i>, <i>xylEGFJQKIH</i>
HMFL31	GCF_030323645.1	Complete	Sludge from paper mill effluent treatment	Chile: Biobio Region, Arauco	<i>alkB1</i> , <i>alkB2</i> , <i>betAB</i> , <i>xylEGFJQKIH</i>
SRIHER B649	GCF_031460255.1	WGS	Oil-contaminated marine water	India: Chennai Harbour	<i>alkB1</i> , <i>alkB2</i> , <i>benABCD</i> , <i>betAB</i>