

Supplementary Materials

A type III polyketide synthase cluster in the phylum *Planctomycetota* is involved in alkylresorcinol biosynthesis

Lars Milke¹, Moses Kabuu², Renè Zschoche², Jochem Gätgens¹, Karin Krumbach¹, Kim-Loreen Carlstedt², Carmen E. Wurzbacher², Sven Balluff³, Christine Beemelmans^{3,4}, Christian Jogler², Jan Marienhagen^{1,5} and Nicolai Kallscheuer^{2,#}

¹ Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, 52425 Jülich, Germany

² Department of Microbial Interactions, Institute for Microbiology, Friedrich Schiller University, 07743 Jena, Germany

³ Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), 66123 Saarbrücken, Germany

⁴ Saarland University, Saarbrücken, Germany

⁵ Institute of Biotechnology, RWTH Aachen University, Worringer Weg 3, 52074 Aachen, Germany

#Corresponding author:

Nicolai Kallscheuer, phone +49 3641 949337, e-mail nicolai.kallscheuer@uni-jena.de

Running Title: Alkylresorcinol production by planctomycetal type III polyketide synthases

Keywords:

planctomycetes, *Corynebacterium glutamicum*, polyketide-derived compounds, hierridin, chemical mediators

Table S1. GenBank accession numbers of the analysed reference genomes

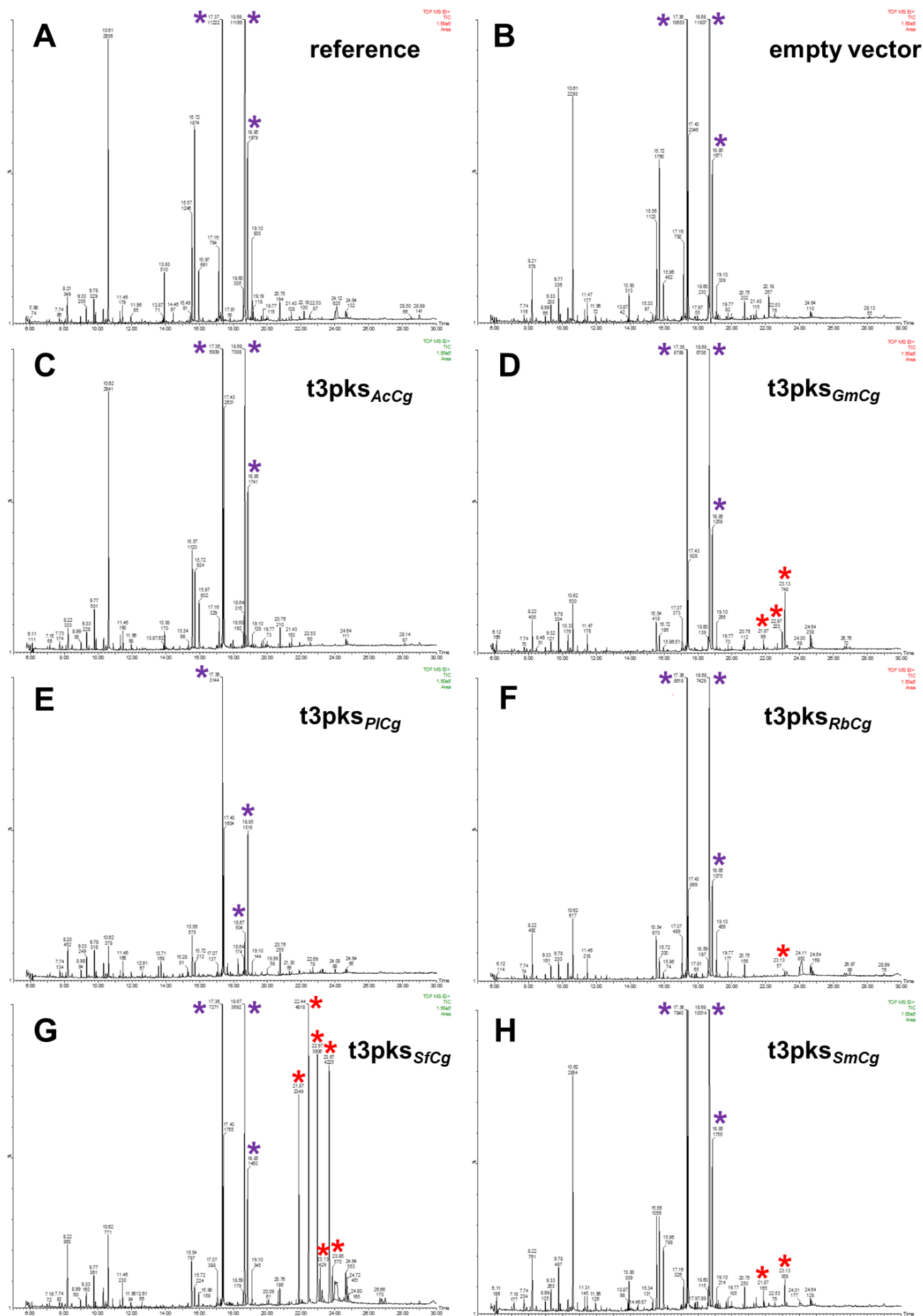
GenBank Acc. No.	Species name	Type strain
GCA_007741515.1	<i>Adhaereter mobilis</i>	HG15A2
GCA_007748035.1	<i>Aeoliella mucimassa</i>	Pan181
GCA_014207595.1	<i>Algisphaera agarilytica</i>	DSM 103725
GCA_007743815.1	<i>Alienimonas californiensis</i>	CA12
GCA_013036045.1	<i>Alienimonas chondri</i>	LzC2
GCA_002007645.1	<i>Anaerohalosphaera lusitana</i>	ST-NAGAB-D1
GCA_007747655.1	<i>Anatilimnocola aggregata</i>	ETA_A8
GCA_024256385.1	<i>Anatilimnocola floriformis</i>	PX40
GCA_008087625.1	<i>Aquisphaera giovannonii</i>	OJF2
GCA_011064615.1	<i>Aquisphaera insulae</i>	JC669
GCA_007752135.1	<i>Aureliella helgolandensis</i>	Q31a
GCA_000153105.1	<i>Blastopirellula marina</i>	DSM 3645
GCA_007859755.1	<i>Blastopirellula retiformator</i>	Enr8
GCA_020966695.1	<i>Blastopirellula sediminis</i>	JC733
GCA_007859955.1	<i>Botrimarina colliarenosi</i>	Pla108
GCA_007859815.1	<i>Botrimarina hoheduenensis</i>	Pla111
GCA_007753265.1	<i>Botrimarina mediterranea</i>	Spa11
GCA_013618625.1	<i>Bremerella alba</i>	FF15
GCA_003335485.1	<i>Bremerella cremea</i>	HEX-1 MG
GCA_007748115.1	<i>Bremerella volcania</i>	Pan97
GCA_008065115.1	<i>Bythopirellula goksoeyrii</i>	Pr1d
GCA_007860015.1	<i>Bythopirellula polymerisocia</i>	Pla144
GCA_007745435.1	<i>Calycomorphotria hydatis</i>	V22
GCA_001753675.2	<i>Candidatus Brocadia sapporoensis</i>	40
GCA_000949635.1	<i>Candidatus Brocadia sinica</i>	JPN1
GCA_000987375.1	<i>Candidatus Brocadia fulgida</i>	RU1
GCA_013360885.1	<i>Candidatus Jettenia caeni</i>	MAG_9
GCA_005524015.1	<i>Candidatus Jettenia ecosi</i>	J2
GCA_011066545.1	<i>Candidatus Kuenenia stuttgartiensis</i>	CSTR1
GCA_000786775.1	<i>Candidatus Scalindua brodae</i>	RU1
GCA_002443295.1	<i>Candidatus Scalindua japonica</i>	husup-a2
GCA_009002475.1	<i>Candidatus Uabimicrobium amorphum</i>	SRT547
GCA_007745175.1	<i>Caulifigura coniformis</i>	Pan44
GCA_007752935.1	<i>Crateriforma conspicua</i>	Mal65
GCA_012290005.1	<i>Crateriforma spongiae</i>	JC647
GCA_007748095.1	<i>Engelhardtia mirabilis</i>	Pla133
GCA_002197845.1	<i>Fimbrioglobus ruber</i>	SP5
GCA_013128195.2	<i>Frigoriglobus tundricola</i>	PL17
GCA_001983935.1	<i>Fuerstiella marisgermanici</i>	NH11
GCA_901538265.1	<i>Gemmata massiliana</i>	Soil9
GCA_901538385.1	<i>Gemmata obscuriglobus</i>	UQM 2246
GCA_017939745.1	<i>Gemmata palustris</i>	G18
GCA_007744675.1	<i>Gimesia alba</i>	Pan241w
GCA_007746795.1	<i>Gimesia algae</i>	Pan161
GCA_007748195.1	<i>Gimesia aquarii</i>	V144

GenBank Acc. No.	Species name	Type strain
GCA_008329715.1	<i>Gimesia chilikensis</i>	JC646
GCA_007754425.1	<i>Gimesia fumaroli</i>	Enr17
GCA_007747015.1	<i>Gimesia maris</i>	CA11
GCA_007748015.1	<i>Gimesia panareensis</i>	Pan110
GCA_015169395.1	<i>Humisphaera borealis</i>	M1803
GCA_000186345.1	<i>Isosphaera pallida</i>	ATCC 43644
GCA_007746075.1	<i>Lacipirellula limnantheis</i>	I41
GCA_009177095.1	<i>Lacipirellula parvula</i>	PX69
GCA_007751035.1	<i>Lignipirellula cremea</i>	Pla85_3_4
GCA_001999965.1	<i>Limihaloglobus sulfuriphilus</i>	SM-Chi-D1
GCA_008254045.1	<i>Limnoglobus roseus</i>	PX52
GCA_007747795.1	<i>Maioricimonas rarisocia</i>	Mal4
GCA_008087665.1	<i>Mariniblastus fucicola</i>	FC18
GCA_007748075.1	<i>Mucisphaera calidilacus</i>	Pan265
GCA_007860135.1	<i>Novipirellula artificiosorum</i>	Poly41
GCA_007860185.1	<i>Novipirellula aureliae</i>	Q31b
GCA_007860095.1	<i>Novipirellula galeiformis</i>	Pla52o
GCA_001956985.1	<i>Paludisphaera borealis</i>	PX4
GCA_011065895.1	<i>Paludisphaera rhizosphaeraeae</i>	JC665
GCA_011064595.1	<i>Paludisphaera soli</i>	JC670
GCA_000284115.1	<i>Phycisphaera mikurensis</i>	NBRC 102666
GCA_000025185.1	<i>Pirellula staleyi</i>	DSM 6068
GCA_007750855.1	<i>Pirellulimonas nuda</i>	Pla175
GCA_900113665.1	<i>Planctomicrobium piriforme</i>	DSM 26348
GCA_007752345.1	<i>Planctopirus ephydatiae</i>	spb1
GCA_001707835.1	<i>Planctopirus hydrillae</i>	JC280
GCA_000092105.1	<i>Planctopirus limnophila</i>	DSM 3776
GCA_007750395.1	<i>Polystyrenella longa</i>	Pla110
GCA_007747445.1	<i>Poriferisphaera corsica</i>	KS4
GCA_007859765.1	<i>Posidoniimonas corsicana</i>	KOR34
GCA_007859935.1	<i>Posidoniimonas polymericola</i>	Pla123a
GCA_007859945.1	<i>Pseudobythopirellula maris</i>	Mal64
GCA_023701485.1	<i>Rhodopirellula aestuarii</i>	ICT_H3.1
GCA_002727185.1	<i>Rhodopirellula bahusiensis</i>	SWK21
GCA_000196115.1	<i>Rhodopirellula baltica</i>	SH1
GCA_000346315.1	<i>Rhodopirellula europaea</i>	SH398
GCA_007860105.1	<i>Rhodopirellula heiligendammensis</i>	Poly21
GCA_001027925.1	<i>Rhodopirellula islandica</i>	K833
GCA_900182915.1	<i>Rhodopirellula lusitana</i>	DSM 25457
GCA_007859915.1	<i>Rhodopirellula pilleata</i>	Pla100
GCA_014192335.1	<i>Rhodopirellula rubra</i>	CECT 8075
GCA_000346505.1	<i>Rhodopirellula sallentina</i>	SM41
GCA_007859855.1	<i>Rhodopirellula solitaria</i>	CA85
GCA_007750655.1	<i>Rohdeia mirabilis</i>	Pla163
GCA_008312935.1	<i>Roseiconus lacunae</i>	JC635
GCA_008629675.1	<i>Roseiconus nitratreducens</i>	JC645
GCA_007741495.1	<i>Roseimaritima multifibrata</i>	FF011L
GCA_009618275.1	<i>Roseimaritima sediminicola</i>	JC651

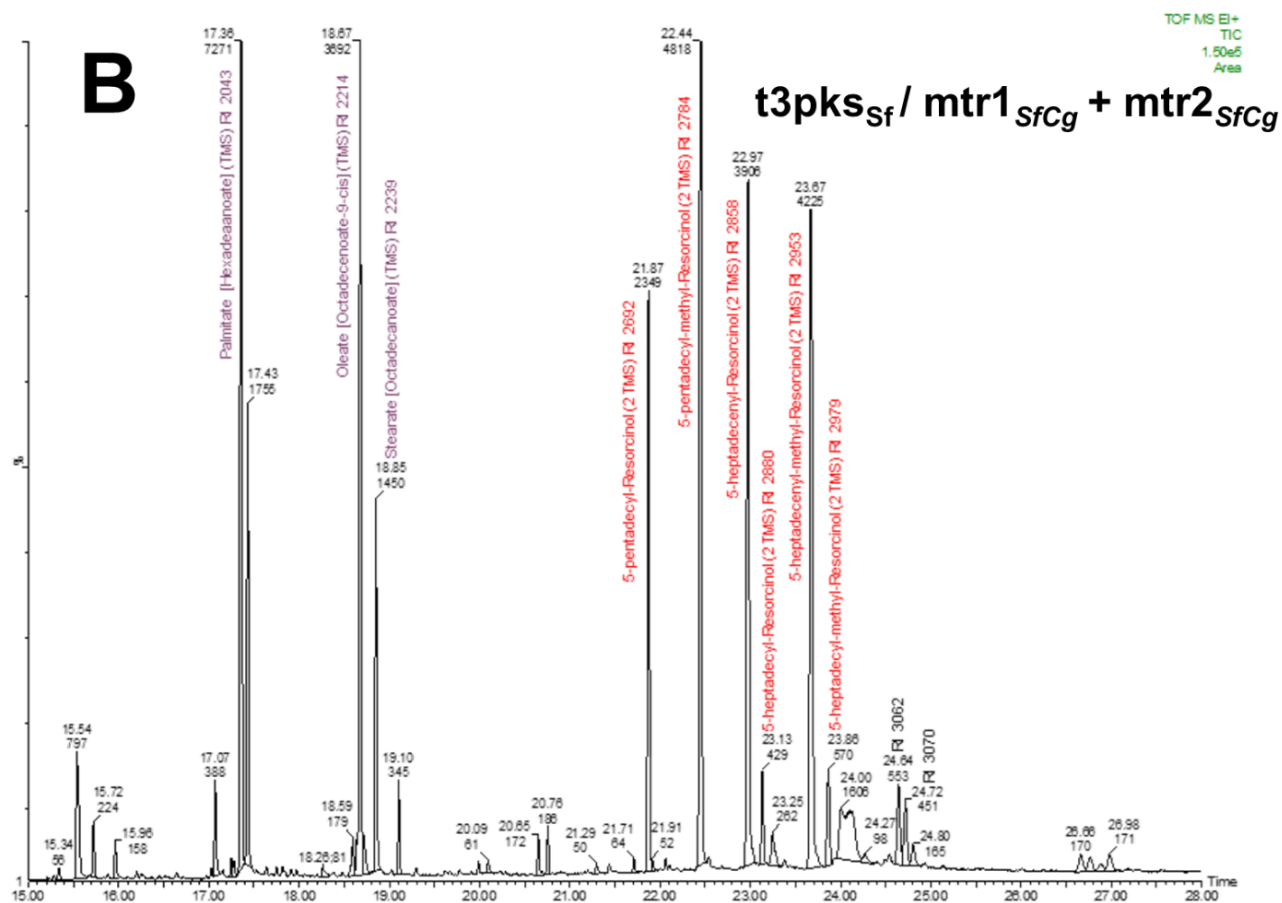
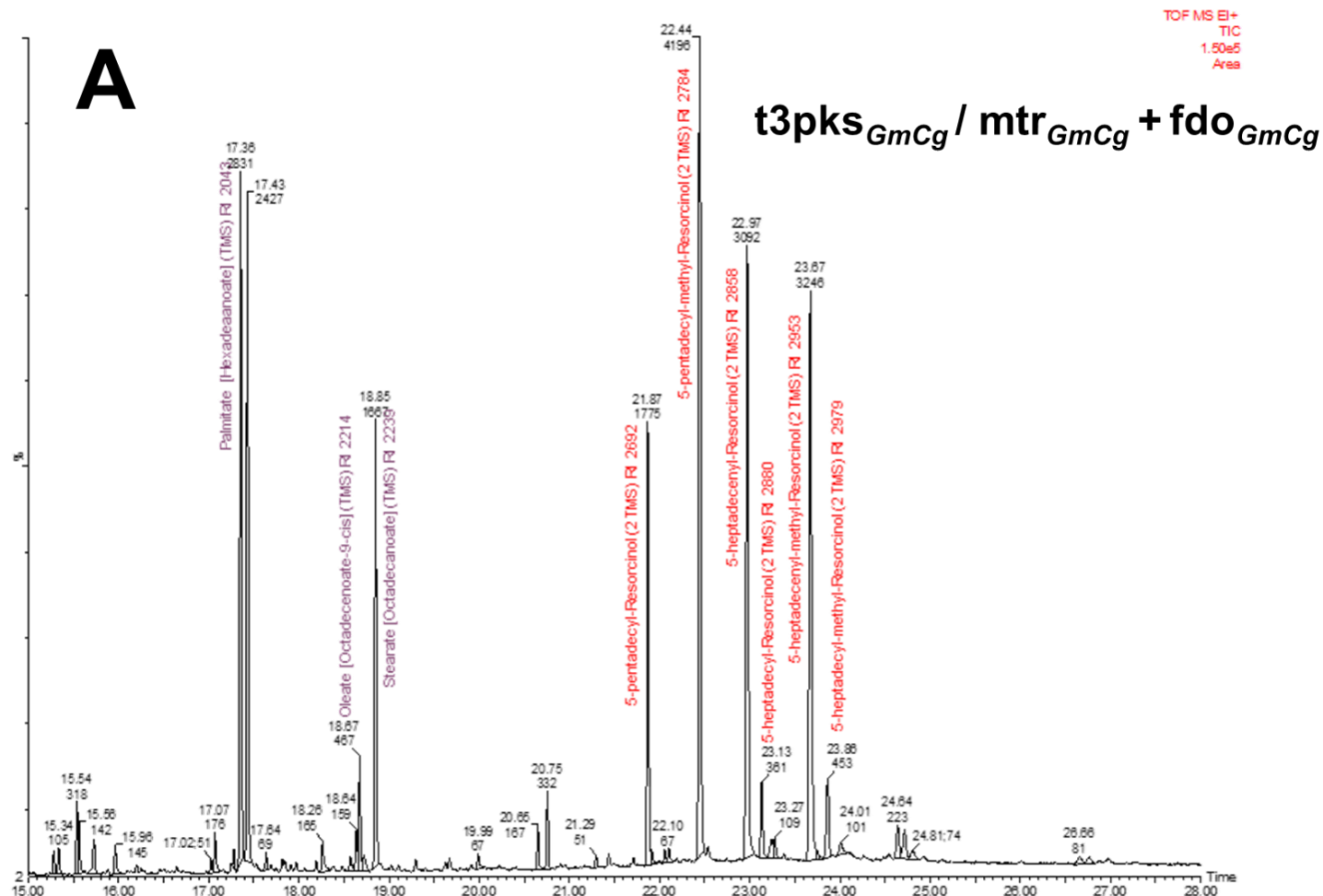
GenBank Acc. No.	Species name	Type strain
GCA_008065135.1	<i>Roseimaritima ulvae</i>	UC8
GCA_007753095.1	<i>Rosistilla carotiformis</i>	Poly24
GCA_007751715.1	<i>Rosistilla oblonga</i>	CA51
GCA_007741475.1	<i>Rosistilla ulvae</i>	EC9
GCA_000165715.3	<i>Rubinisphaera brasiliensis</i>	DSM 5305
GCA_007859715.1	<i>Rubinisphaera italica</i>	Pan54
GCA_022267515.1	<i>Rubinisphaera margarita</i>	ICM_H10
GCA_007859865.1	<i>Rubripirellula amarantea</i>	Pla22
GCA_007741535.1	<i>Rubripirellula lacrimiformis</i>	K22.7
GCA_008374075.1	<i>Rubripirellula obstinata</i>	LF1
GCA_007860175.1	<i>Rubripirellula reticaptiva</i>	Poly59
GCA_007860125.1	<i>Rubripirellula tenax</i>	Poly51
GCA_007751475.1	<i>Saltatorellus ferox</i>	Poly30
GCA_000255655.1	<i>Schlesneria paludicola</i>	DSM 18645
GCA_001997385.1	<i>Sedimentisphaera cyanobacteriorum</i>	L21-RPul-D3
GCA_002117005.1	<i>Sedimentisphaera salicampi</i>	ST-PulAB-D4
GCA_000242455.3	<i>Singulisphaera acidiphila</i>	DSM 18658
GCA_008035925.1	<i>Stieleria maiorica</i>	Mal15
GCA_007754155.1	<i>Stieleria neptunia</i>	Enr13
GCA_025060895.1	<i>Stieleria sedimenti</i>	ICT_E10.1
GCA_024129865.1	<i>Stieleria tagensis</i>	TO1_6
GCA_007860045.1	<i>Stieleria varia</i>	Pla52n
GCA_007744515.1	<i>Stratiformator vulcanicus</i>	Pan189
GCA_007747995.1	<i>Symmachiella dynata</i>	Mal52
GCA_007860075.1	<i>Symmachiella macrocystis</i>	CA54
GCA_009177065.1	<i>Tautonia marina</i>	JC650
GCA_007752535.1	<i>Tautonia plasticadhaerens</i>	EIP
GCA_012958305.1	<i>Tautonia rosea</i>	JC657
GCA_003977685.1	<i>Tautonia sociabilis</i>	GM2012
GCA_018398935.1	<i>Telmatocola sphagniphila</i>	SP2
GCA_007859735.1	<i>Thalassoglobus neptunius</i>	KOR42
GCA_007744255.1	<i>Thalassoglobus polymorphus</i>	Mal48
GCA_011634775.1	<i>Thalassoroseus pseudoceratinae</i>	JC658
GCA_013694095.1	<i>Thermogemmata fonticola</i>	2918
GCA_002277955.1	<i>Thermogutta terrifontis</i>	R1
GCA_901538355.1	<i>Tuwongella immobilis</i>	MBLW1
GCA_007747215.1	<i>Urbifossiella limnaea</i>	ETA_A1
GCA_000255705.1	<i>Zavarzinella formosa</i>	DSM 19928

Table S2. Protein sequence similarity matrix of type III PKSs from planctomycetes and *Cyanobium* sp. LEGE 06113. The similarity matrix was calculated from the same alignment, which was also used for the construction of the phylogenetic tree presented in Figure 5. Identity values are given in %.

Species (type III PKS gene locus tag)	<i>Stieleria maiorica</i> (Mal15_23020)	<i>Rhodopirellula baltica</i> (RB8853)	<i>Gimesia maris</i> (PM8797T_03239)	<i>Alienimonas californiensis</i> (CA12_06310)	<i>Planctopirus limnophila</i> (Plim_0466)	" <i>Saltatorellus ferox</i> " (Poly30_13920)	<i>Cyanobium</i> sp. LEGE 06113 (IQ220_04565)
<i>Stieleria maiorica</i> (Mal15_23020)	100.0	42.7	57.7	46.7	45.2	33.0	50.6
<i>Rhodopirellula baltica</i> (RB8853)	42.7	100.0	43.5	40.0	40.3	28.2	43.8
<i>Gimesia maris</i> (PM8797T_03239)	57.7	43.5	100.0	47.9	46.5	30.3	54.8
<i>Alienimonas californiensis</i> (CA12_06310)	46.7	40.0	47.9	100.0	47.9	31.5	48.7
<i>Planctopirus limnophila</i> (Plim_0466)	45.2	40.3	46.5	47.9	100.0	27.9	47.4
" <i>Saltatorellus ferox</i> " (Poly30_13920)	33.0	28.2	30.3	31.5	27.9	100.0	33.2
<i>Cyanobium</i> sp. LEGE 06113 (IQ220_04565)	50.6	43.8	54.8	48.7	47.4	33.2	100.0



◁ **Figure S1. Total ion chromatograms of extracts from the *C. glutamicum* strains expressing genes encoding planctomycetal type III PKSs.** The used *C. glutamicum* strain without plasmid and without induction (reference) and the IPTG-induced strain carrying pMKEx2 (empty vector) served as negative controls. Peaks corresponding to fatty acids acting as starter units for the type III PKS-catalyzed reactions are highlighted with purple asterisks. Red asterisks denote peaks that correspond to the determined alkylresorcinols. Peak annotations: 17.37 min: palmitate [hexadecanoate] (1 TMS) RI 2044, 18.68 min: oleate [octadecenoate-9-*cis*] (1 TMS) RI 2215, 18.85 min: stearate [octadecanoate] (1 TMS) RI 2239, 21.87 min: 5-pentadecyl-resorcinol (2 TMS) RI 2692, 22.44 min: 5-pentadecyl-methyl-resorcinol (2 TMS) RI 2784, 22.97 min: 5-heptadecenyl-resorcinol (2 TMS) RI 2858, 23.13 min: 5-heptadecyl-resorcinol (2 TMS) RI 2880, 23.67 min: 5-heptadecenyl-methyl-resorcinol (2 TMS) RI 2953, 23.86 min: 5-heptadecyl-methyl-resorcinol (2 TMS) RI 2979. Abbreviations: *Ac*, *Alienimonas californiensis*; *Cg*, *Corynebacterium glutamicum*, *Gm*, *Gimesia maris*, *Pl*: *Planctopirus limnophila*, *Rb*: *Rhodopirellula baltica*; *Sf*, *Saltatorellus ferox*; *Sm*, *Stieleria maiorica*; RI, retention index; TMS, trimethylsilyl residues.



◁ **Figure S2. Total ion chromatograms of extracts from the *C. glutamicum* strains co-expressing genes encoding type III PKSs and hydroxylases/methyltransferases.** Peaks corresponding to fatty acids acting as starter units for the type III PKS-catalyzed reactions are highlighted with purple asterisks. Red asterisks denote peaks that correspond to the determined alkylresorcinols. Peak annotations: 17.37 min: palmitate [hexadecanoate] (1 TMS) RI 2044, 18.68 min: oleate [octadecenoate-9-*cis*] (1 TMS) RI 2215, 18.85 min stearate [octadecanoate] (1 TMS) RI 2239, 21.87 min: 5-pentadecyl-resorcinol (2 TMS) RI 2692, 22.44 min: 5-pentadecyl-methyl-resorcinol (2 TMS) RI 2784, 22.97 min: 5-heptadecenyl-resorcinol (2 TMS) RI 2858, 23.13 min: 5-heptadecyl-resorcinol (2 TMS) RI 2880, 23.67 min: 5-heptadecenyl-methyl-resorcinol (2 TMS) RI 2953, 23.86 min: 5-heptadecyl-methyl-resorcinol (2 TMS) RI 2979. Abbreviations: *Cg*, *Corynebacterium glutamicum*, *Gm*, *Gimesia maris*, *Sf*, *Saltatorellus ferox*; RI, retention index; TMS, trimethylsilyl residues.

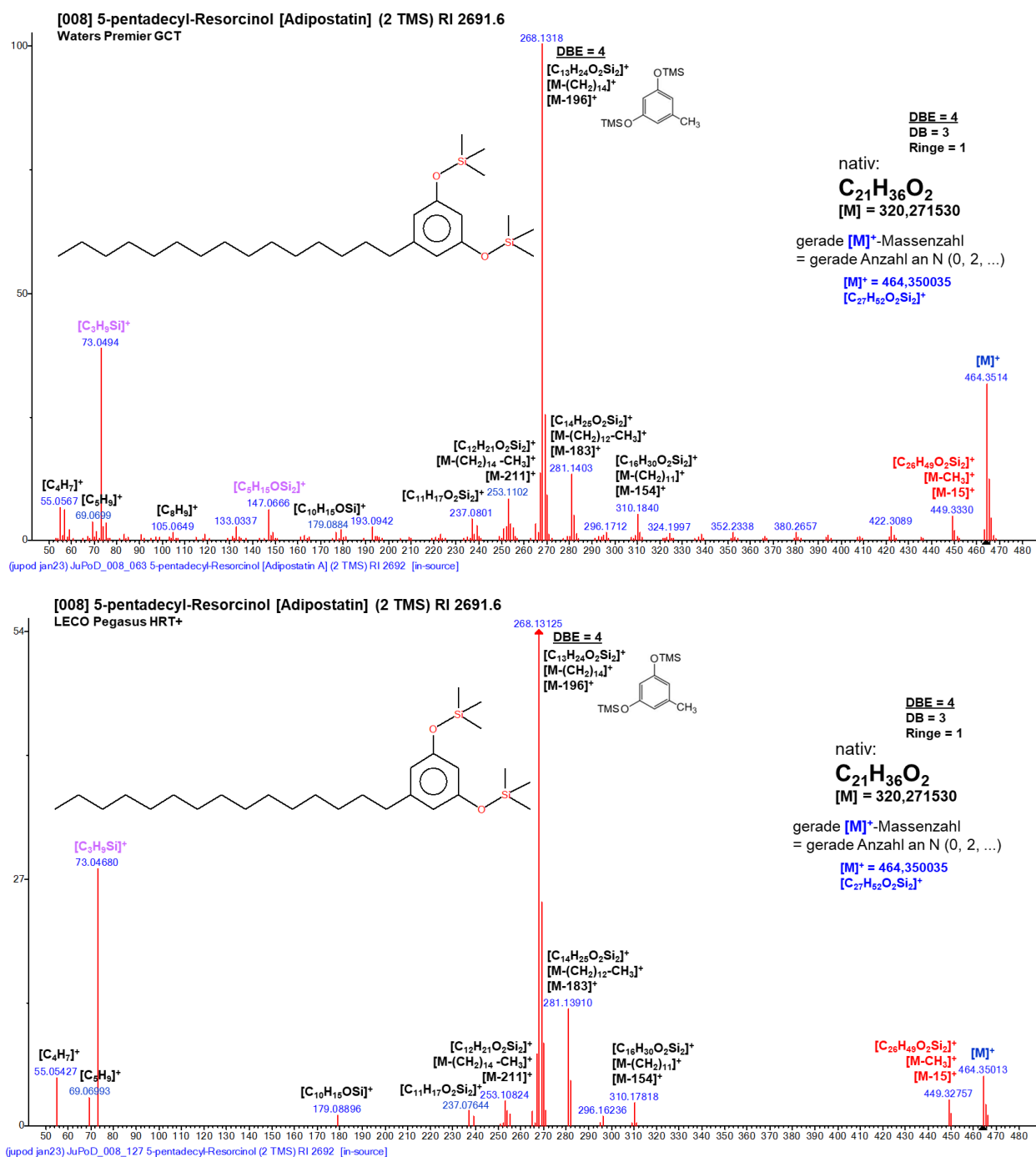


Figure S3. EI⁺ spectra of derivatized 5-pentadecylresorcinol (2 TMS). The spectra obtained with the Waters GCT Premier mass spectrometer (top panel) and with the LECO Pegasus HRT+ mass spectrometer (bottom panel) are shown. Abbreviations: DB, double bonds; DBE, double bond equivalents; RI: retention index, TMS: trimethylsilyl residue.

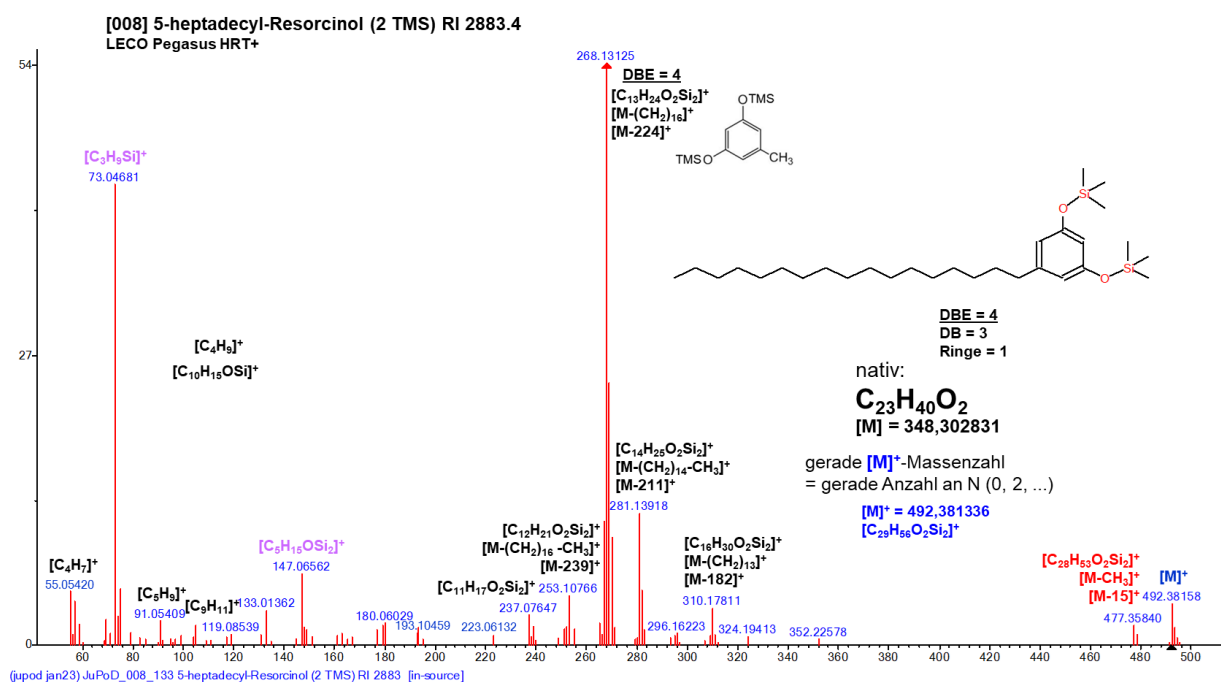
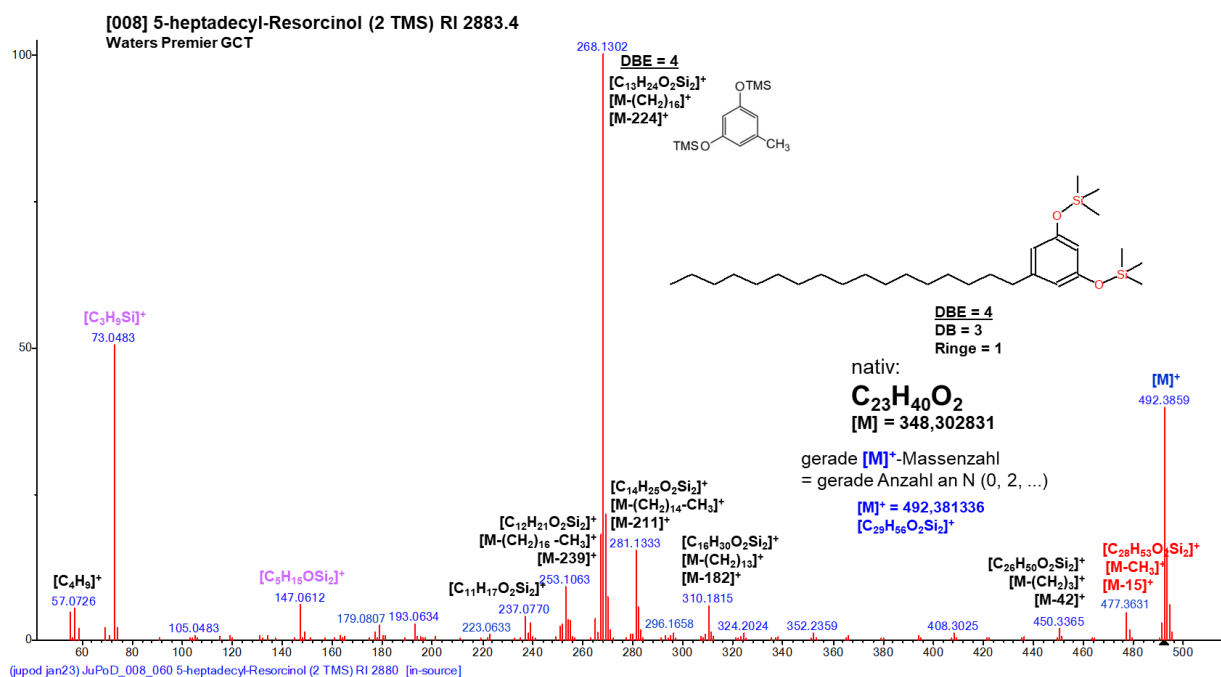


Figure S4. EI⁺ spectra of derivatized 5-heptadecylresorcinol (2 TMS). The spectra obtained with the Waters GCT Premier mass spectrometer (top panel) and with the LECO Pegasus HRT+ mass spectrometer (bottom panel) are shown. Abbreviations: DB, double bonds; DBE, double bond equivalents; RI: retention index, TMS: trimethylsilyl residue.

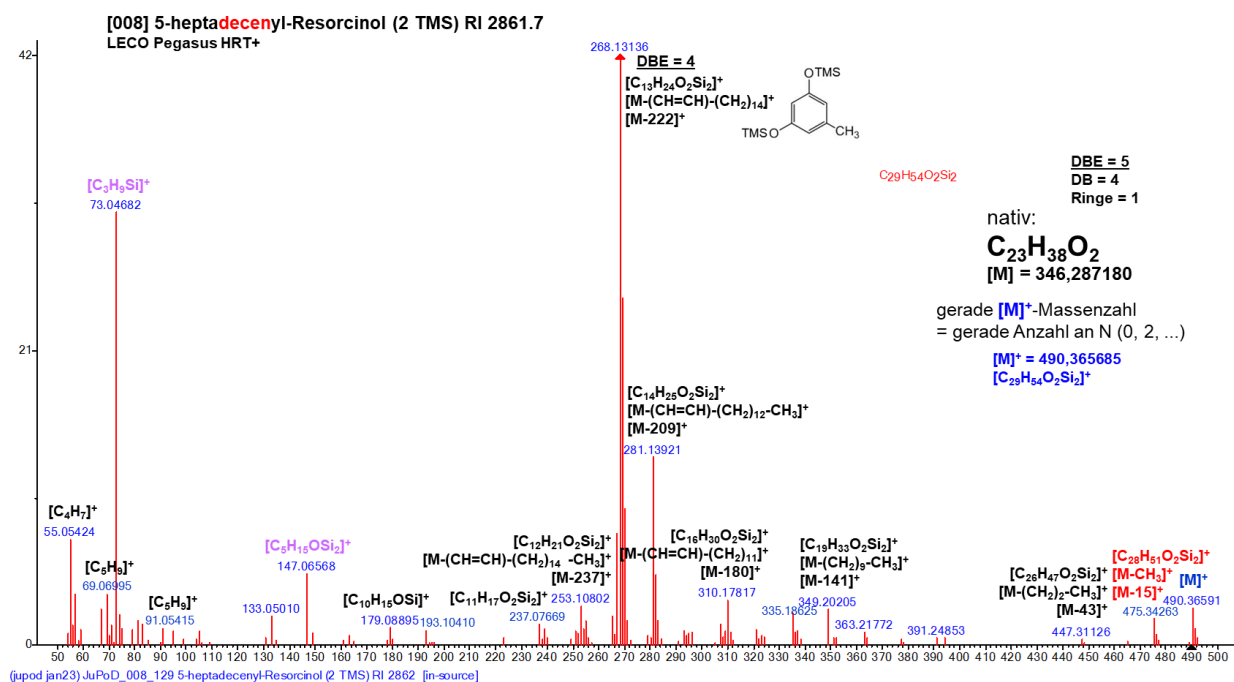
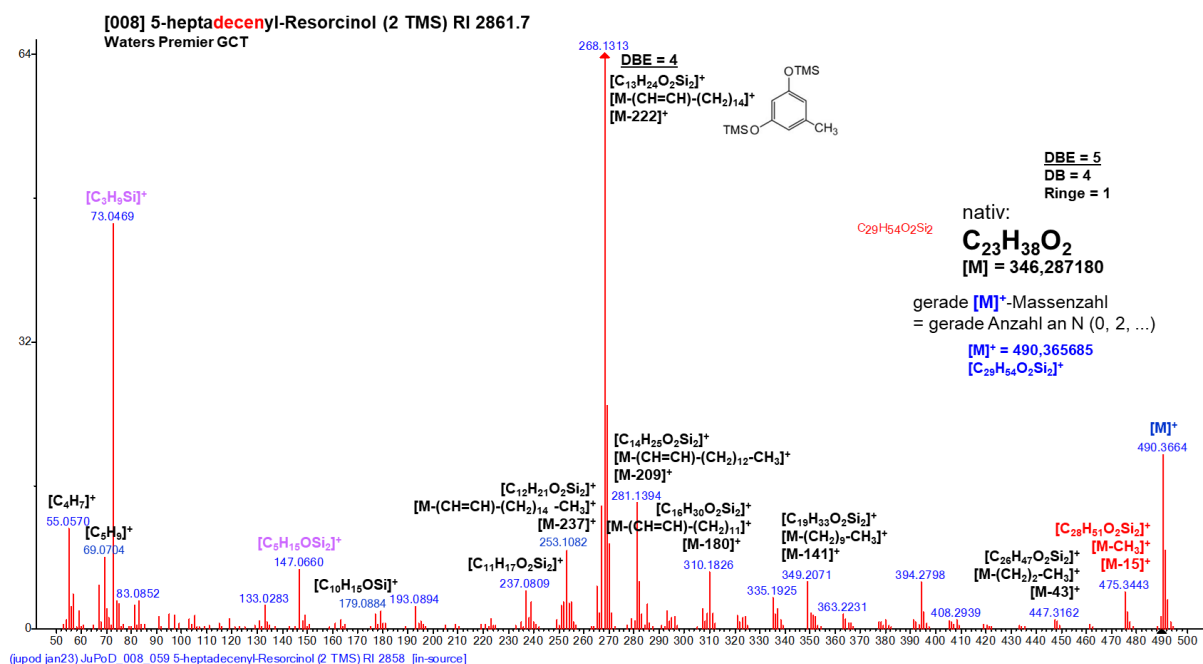


Figure S5. EI⁺ spectra of derivatized 5-hepta~~dec~~enylresorcinol (2 TMS). The spectra obtained with the Waters GCT Premier mass spectrometer (top panel) and with the LECO Pegasus HRT+ mass spectrometer (bottom panel) are shown. Abbreviations: DB, double bonds; DBE, double bond equivalents; RI: retention index, TMS: trimethylsilyl residue.

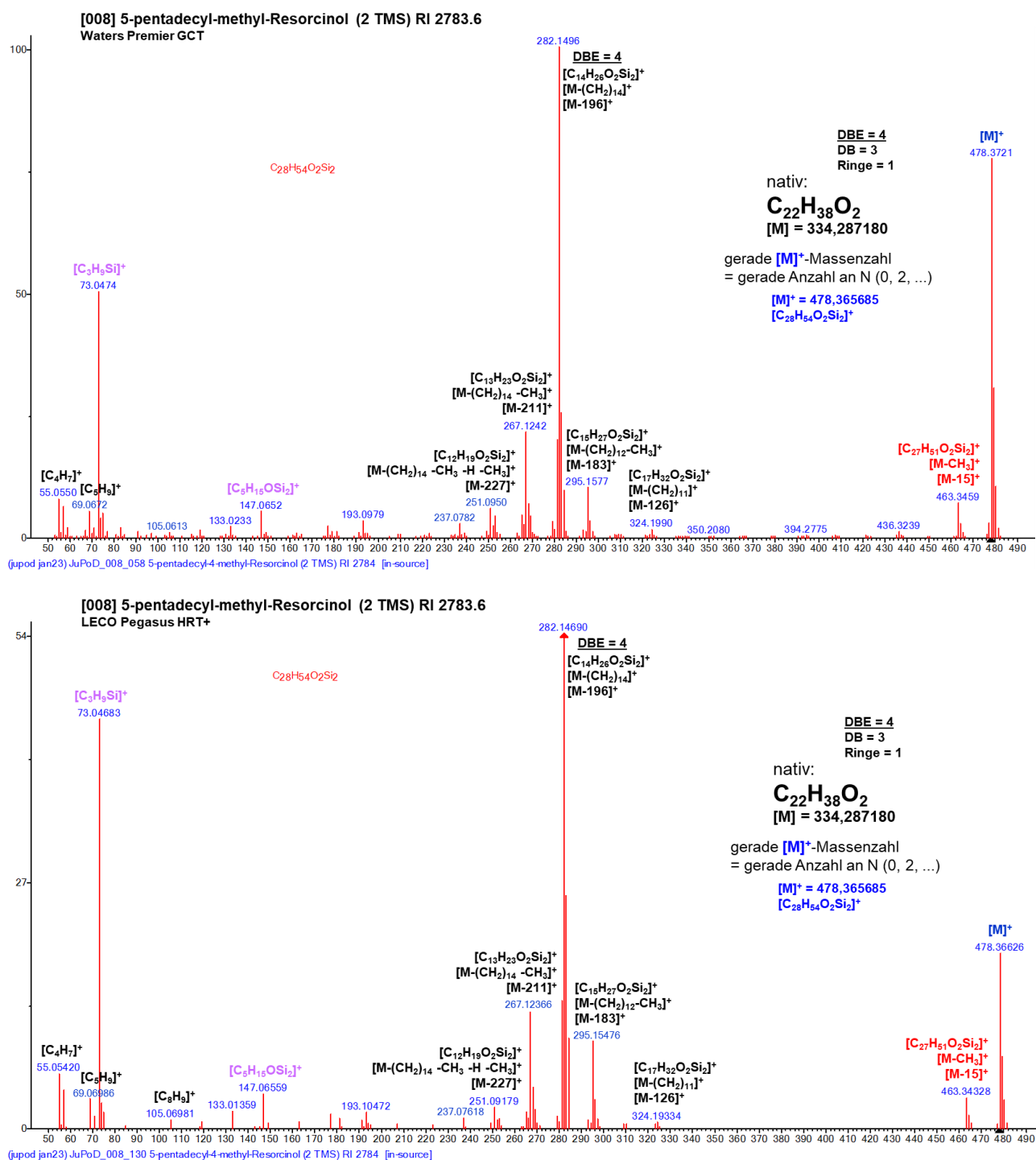


Figure S6. EI⁺ spectra of derivatized 5-pentadecyl-methylresorcinol (2 TMS). The spectra obtained with the Waters GCT Premier mass spectrometer (top panel) and with the LECO Pegasus HRT+ mass spectrometer (bottom panel) are shown. Abbreviations: DB, double bonds; DBE, double bond equivalents; RI: retention index, TMS: trimethylsilyl residue.

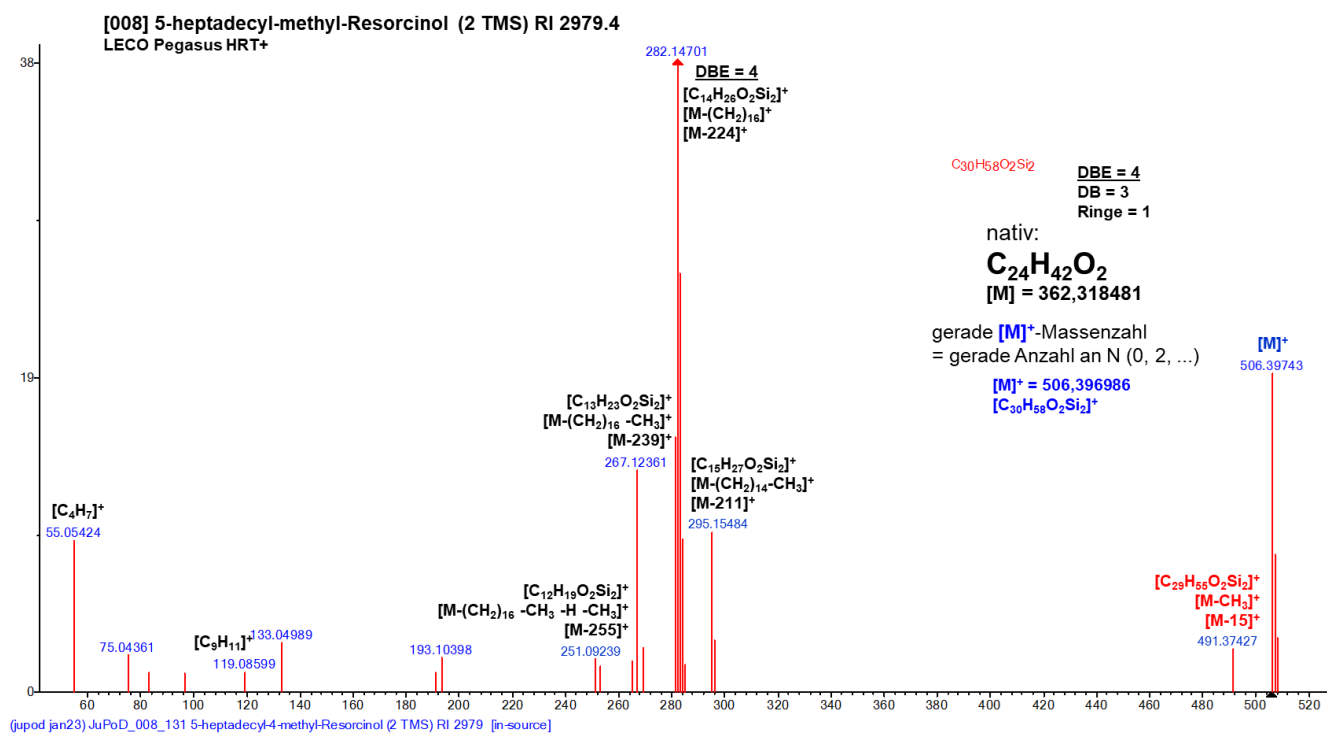
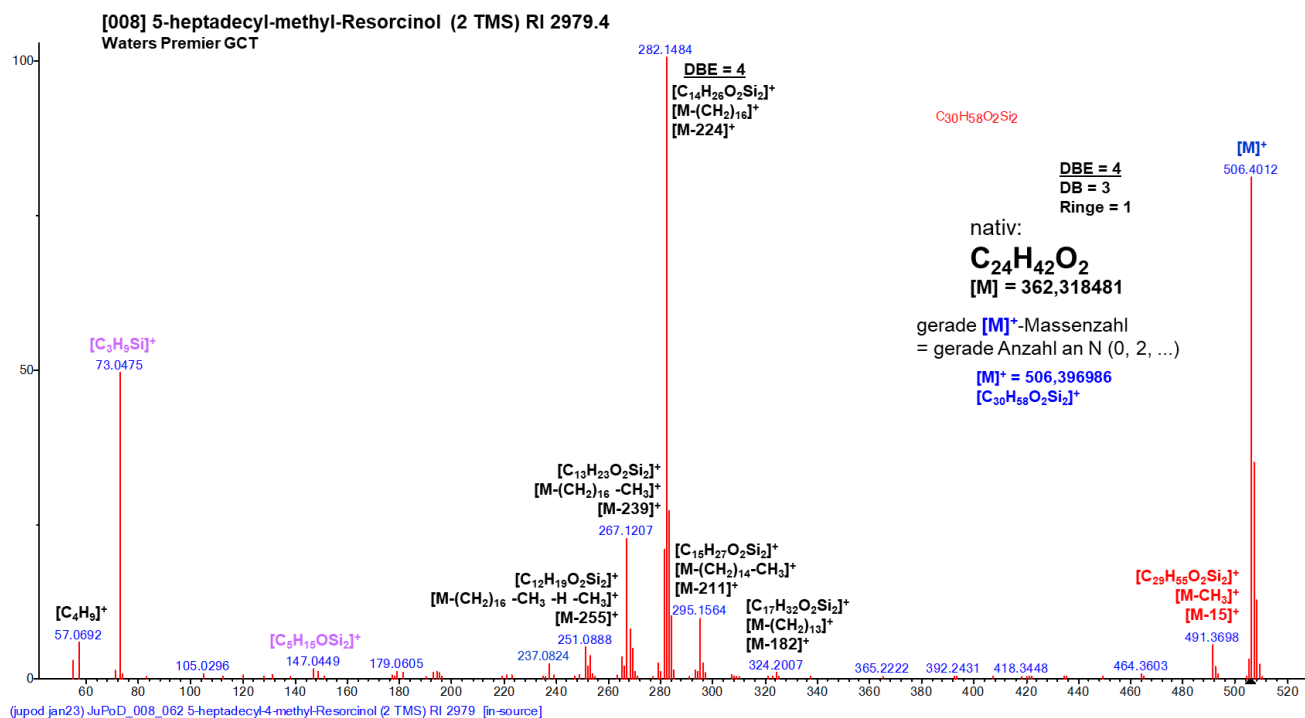


Figure S7. EI⁺ spectra of derivatized 5-heptadecyl-methylresorcinol (2 TMS). The spectra obtained with the Waters GCT Premier mass spectrometer (top panel) and with the LECO Pegasus HRT+ mass spectrometer (bottom panel) are shown. Abbreviations: DB, double bonds; DBE, double bond equivalents; RI: retention index, TMS: trimethylsilyl residue.

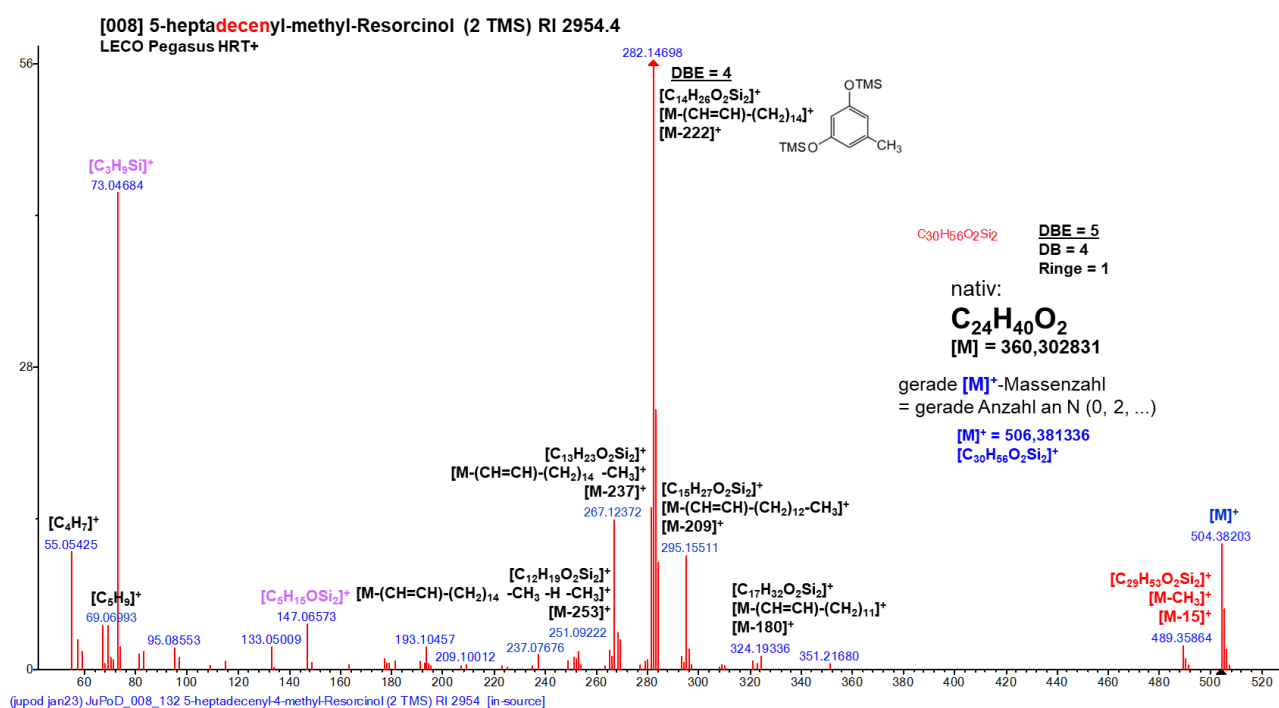
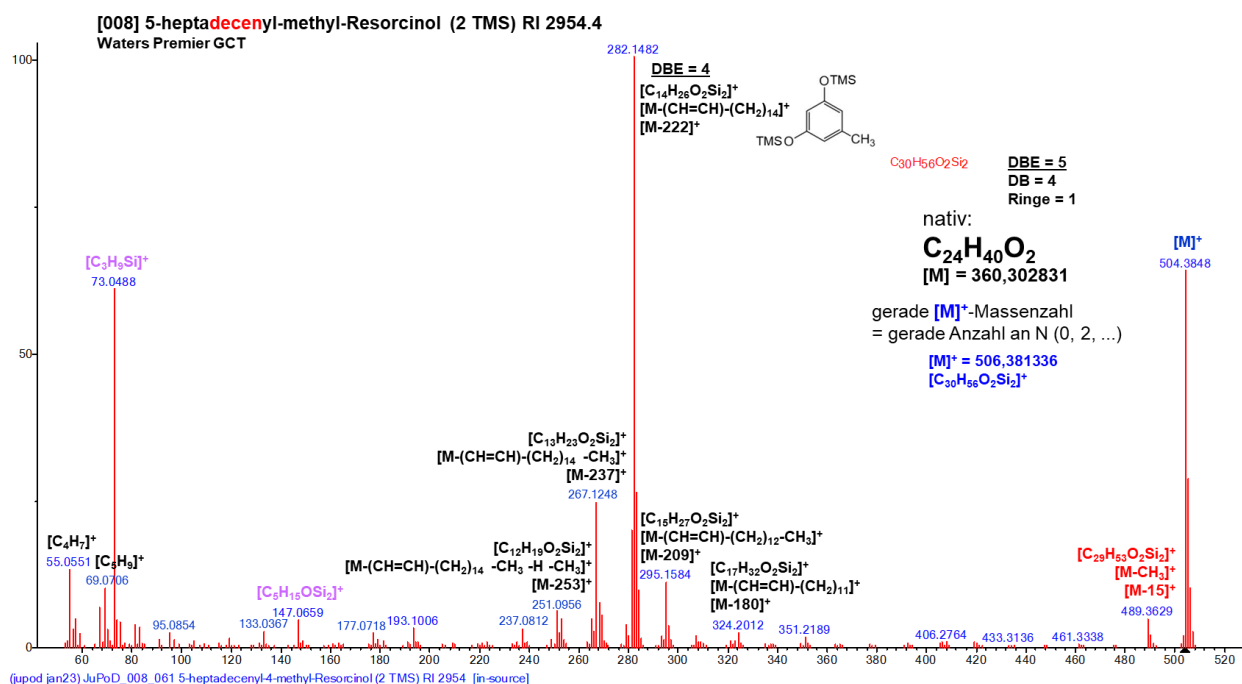


Figure S8. EI⁺ spectra of derivatized 5-heptadecenyl-methylresorcinol (2 TMS). The spectra obtained with the Waters GCT Premier mass spectrometer (top panel) and with the LECO Pegasus HRT+ mass spectrometer (bottom panel) are shown. Abbreviations: DB, double bonds; DBE, double bond equivalents; RI: retention index, TMS: trimethylsilyl residue.

***S. maiorica* Δ Mal15_23020::cat**



***S. maiorica* Δ Mal15_51000::cat**

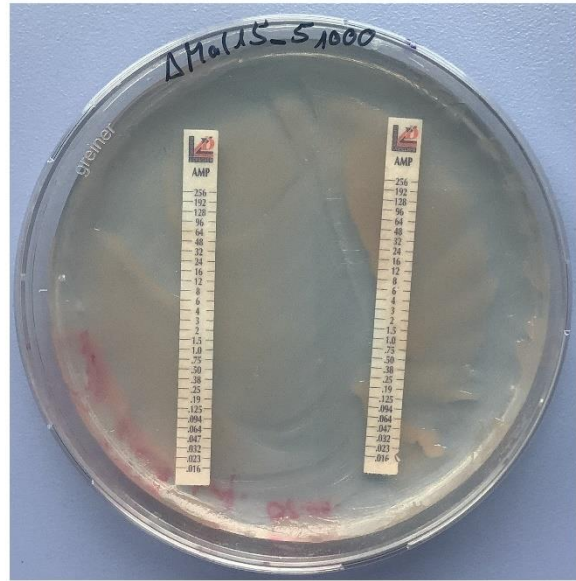


Figure S9. Plate assay for the determination of minimal inhibitory concentrations for ampicillin with constructed *S. maiorica* gene deletion mutants. The two constructed strains were streaked on M3H NAG ASW plates with 34 mg/L chloramphenicol. The plates were incubated with MIC stripes containing ampicillin (0.016-256 mg/L) for two days at 28 °C. Abbreviations: cat, chloramphenicol acetyltransferase (chloramphenicol resistance gene).

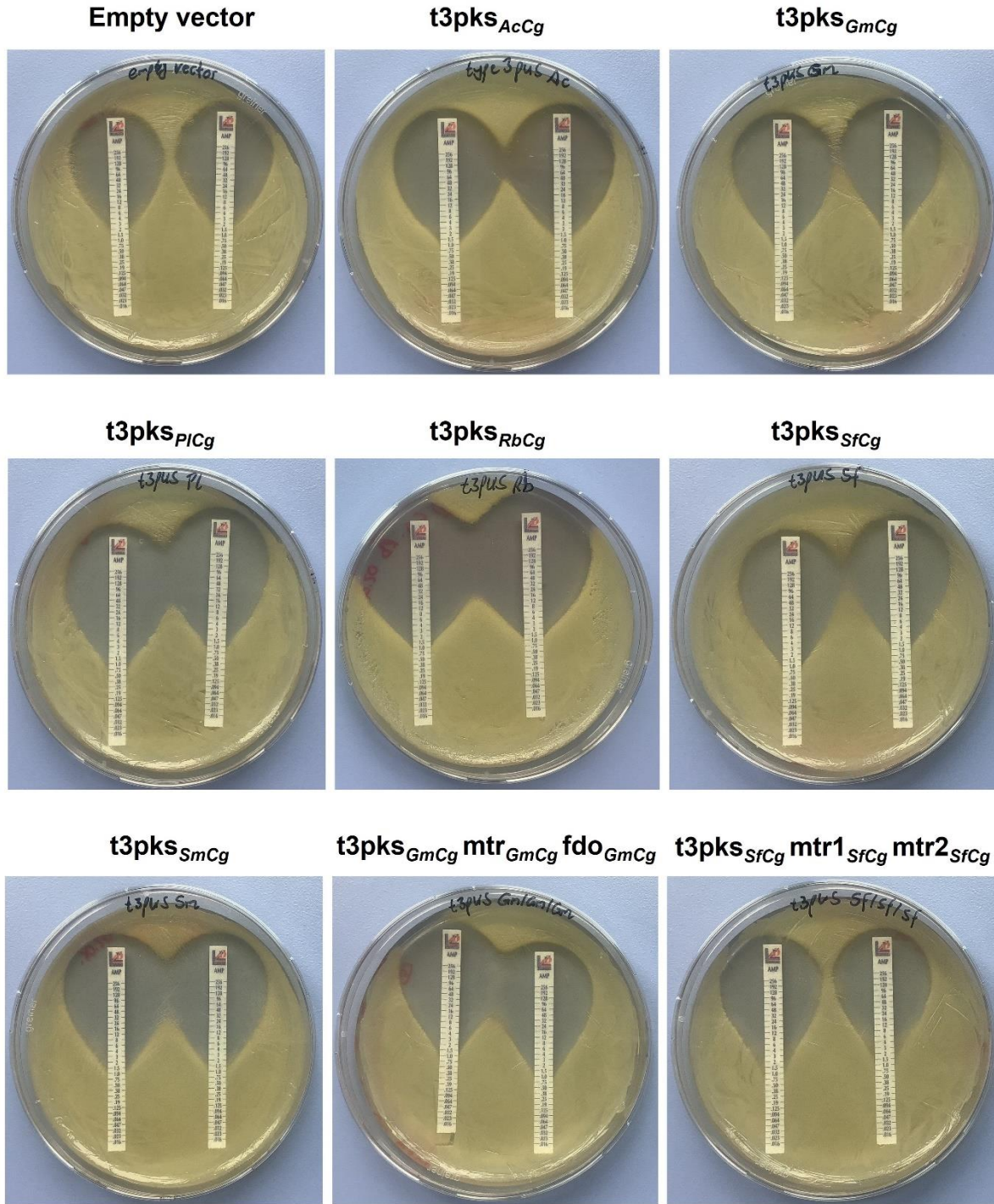


Figure S10. Plate assay for the determination of minimal inhibitory concentrations for ampicillin with constructed *C. glutamicum* strains expressing planctomyete-derived genes. Cell material of the induced *C. glutamicum* cultures (72 h after inoculation) was streaked on CGXII agar with 4% (w/v) glucose, the required antibiotics for plasmid maintenance and 1 mM IPTG. The plates were incubated with MIC stripes containing ampicillin (0.016-256 mg/L) overnight at 28 °C. Abbreviations: Ac, *Alienimonas californiensis*; Cg, *Corynebacterium glutamicum*, Gm, *Gimesia maris*, Pl: *Planctopirus limnophila*, Rb: *Rhodopirellula baltica*; Sf, *Saltatorellus ferox*; Sm, *Stieleria maiorica*.

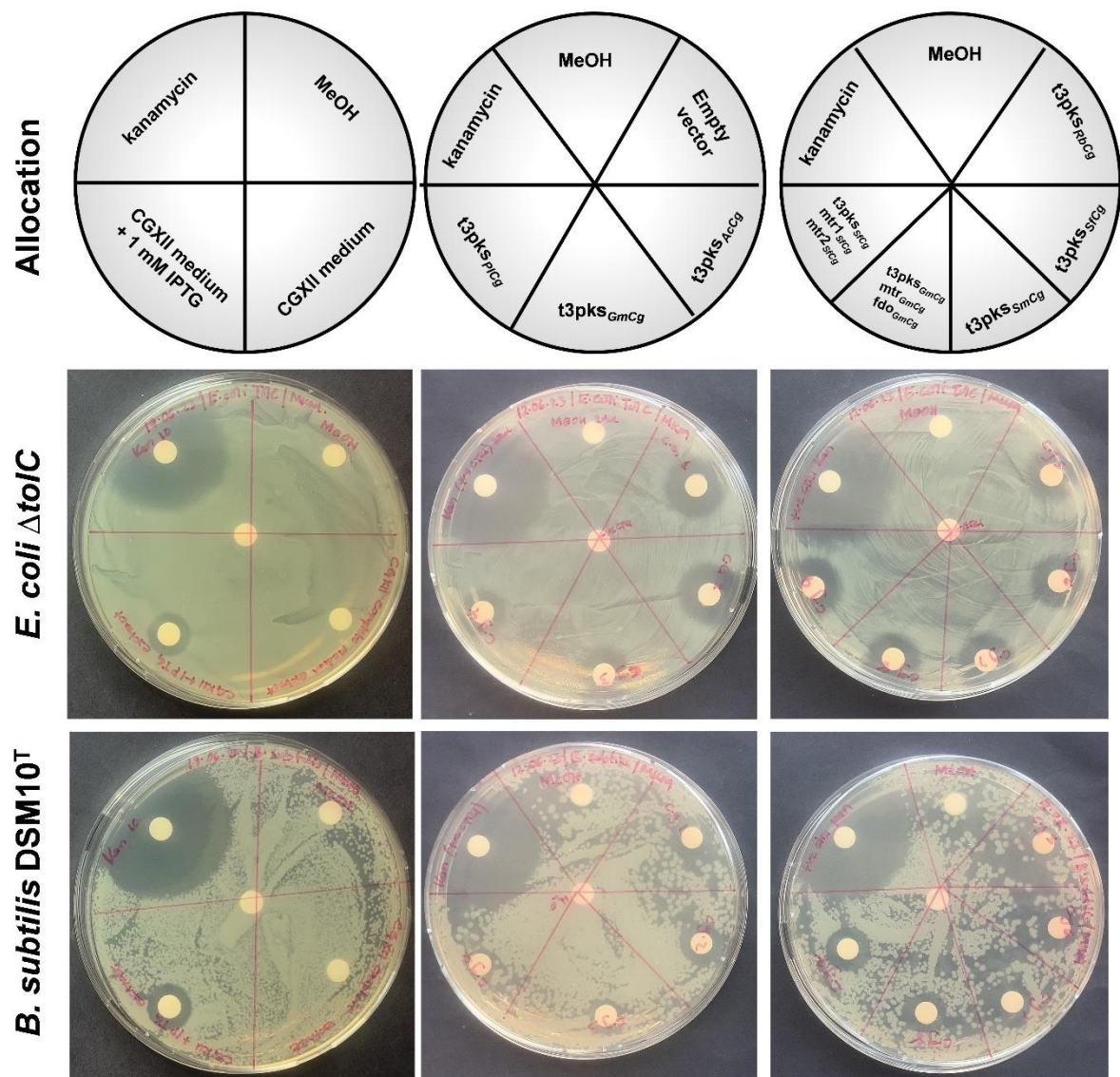


Figure S11. Bioactivity assay with culture extracts of *C. glutamicum* strains expressing planctomycete-derived genes. Fifty mL of the each of the induced *C. glutamicum* cultures (72 h after inoculation) was extracted. The extracts were tested for bioactivity against *E. coli* $\Delta tolC$ and *B. subtilis*. Kanamycin was used as a positive control. Abbreviations: Ac, *Alienimonas californiensis*; Cg, *Corynebacterium glutamicum*; Gm, *Gimesia maris*; Pl: *Planctopirus limnophila*, Rb: *Rhodopirellula baltica*; Sf, *Saltatorellus ferox*; Sm, *Stieleria maiorica*.

Sequences of synthesized genes used in this study. The open reading frames of the codon-optimized genes coding for the type III PKS (t3pks), oxidoreductase (fdo) and methyltransferase (mtr) are provided below. Abbreviations: *Ac*, *Alienimonas californiensis*; *Cg*, *Corynebacterium glutamicum*, *Gm*, *Gimesia maris*, *Pl*: *Planctopirus limnophila*, *Rb*: *Rhodopirellula baltica*; *Sf*, *Saltatorellus ferox*; *Sm*, *Stieleria maiorica*

>t3pks_{AcCg}

ATGTCGGGTATTGGAACCTGCCTTGCCAGCTGGATGGATCGAACAGGGAGATACTGGTGTAT
CTGGCCGGACGGACTGGACCGCCAGCTGGTCCCGACGGAGAAGCGGGCCGTCGCCGTATC
GCAGCCCTCTACCGCAAGGTGGGAGTGAAGAAACGTCACCTGGTTCTTGTGGACCCAGAT
TCGTCTGGTACGGAACCCGATCGCGTCCCATTCTATCCCCAGGACCGGGTACTCATGCA
GGAGCGCCACGACTGGTGATCGGATTATCGCGTACGAAAAACACGCCGGCCCTTTGGCA
GTACGCGCCGAGCTGCCGCACTGGCAGATGCTGGAGTGCATCCAGGCCGTGTAACCTCAG
AGCGTTACTGTCTCTTGCACGGGTTTTGCTGCCCCCGGAGTCGATTGTGCATTGATCGAA
GACCTCGGTCTTCCGCGCTCGGTGGGTGCGACCCACGTGGGATTTATGGGATGCCATGCT
GCTCTGAACGGACTTCGTGTGGCCGGTGCTTTGTGTGCTGCGGACCCATCTGCGGCCGTG
CTGCTTACCGCCGTGGAATTGTGCAGCCTGCACCATCAGTACACCCCCGACCCTGTAGCT
GATTCTGGTCAAGTTATTGCCAACGCCCTTTTTGCGGACGGAGCAGCCGCGATGTTGTGC
ACTGGCGCAGACTTCCGGCCTGATCCCCCTACCGCGGAGTGTGCGACGAACCAAGACT
GCCACGGCTCCCTACCGCCTGCTCGGAAGCGGATCTGTTGTTTTGGCAGATACGGCCGAC
GCAATGACCTGGCGTGTGGGCGACGCGGGCTTCTTGATGACTCTTGAGCTTGCTGTTCCC
GTAGCAATCCGCGCTCATCTGGCAGGTTGGATGGATGAATGGCTCTCTACTTTTCGGTTTG
ACCCGTGCTGAGGTGCGCACTTGGGCAGTGCACCCGGGCGGACCCGCGATTTTGGACGCC
GCAGCGGAAGCACTCGGACTCGGCGAACGTGAACCTCGACCCGTCGCGCGCCCTTTTGGCT
GAAGTGGGCAACCTCTCCTCCCCGACTATCCTCTTTCTTTTGGACCGCCTCCGCTCTGCT
CGGGGTCTGGCCGGCGCAAAGCCTCGGACGGCCCTGTAGTCGCAGTGGGATTTGGACCT
GGACTTACGGTTGAAGCTGCCCTGTTTGGTTAA

>t3pks_{GmCg}

ATGAGCTTCGAAATTTTGGGTATCGGCACCACTAACCCGAACATTCAATCCAACAATCG
GAGGCCGCCGTCCACGCCCAATCACTCTCATGTACGCAGAAAGCACTGAACAACAGCGT
CGTCTTTTGCCTGTACTGTATCGCCGTGCGGGAGTAAAGACGCGTCATTCCGTAGTCCTC
GAAAATTCTCTGGCGAAGAAGCCGTTTCGGCAAACGTTTTACCAACCAGCGGATAGCCCC
GTGGACCTGGGTCCAGCAACGTCTACGCGTATGCAGGAATATGAGAAACATGCTGCCGAG
CTGGCAATCACTGCTGTACGCGAAGCACTGGATTTCGGCTAACGTGAACCCGGCGGAAGTG
ACTCAACTTGTGACCGTTAGCTGTAGCGGCTTTAGCGCGCCCGGATTCGATCTTCAGGTC
TTGAAACAGCCGGGTTTTTACCCGACGTGTCGCGGACCCACATCGGCTTCATGGGCTGT
CACGGTGCCTGAACGGTCTGCGTGTGCAAAGTCCTTTACGGATAATGACCCCGAGGCC
TGTGTTGTAGTCTGCGCGGTGGAGTTGTGTTCACTGCACCAGCAATACGGATGGTGCCCC
GACAAGATTGTTGAAATGCACTCTTTGCTGACGGTGCCGCCGCGGTAGTTGGTAAACAG
AGCCAGTCATCATCTTCGGACCACTGGAAGTTGGTTGCCTCAGGATCAACGGTGGTGCCT
GACTCAGAGGAAATGATGAGCTGGCGTATCGGTGATCATGGATTTCGAAATGACGCTCTCC
CCCTTGATCCCCGATTTGATTAAATCGCGCTTGCGTCTTGCTTGAGAACTGGCTCGCG
GGTCAGGGTACGTCTATTGAGGAGATCCGTTCTTGGGCGATCCATCCCGGTGGCCCCCGC
ATTCTTACGGCTGTCTCTGAAGCTGTGGGATTCGATGAGGAGAAATTGGTCCCTTCTCGG
GCAATTCTTGCTGAGTTTGGTAACATGTCATCTCCAACCGTGCTGTTTCATCTGCAACGC
CTTCAGGCGACTCGCACCGCTCGCCCTTGTTGTTATGTTGGGCTTCGCCCCGGCCTTACT
ATTGAGGCTGCTCTTGTCAAATAA

>t3pks_{P1Cg}

ATGGCAATGCTGATCTCAGGTATCGGTACTGCCGCACCTGCATCCCGTATCCCCCAAATG
GATGCTTACAGCGCTATGTCTGAATACTTGTGTGATTTCGGATGAACATCGCCGTGTTATG
GAGATGATTTACCAAGGTTCCGGCGTCAAGCACCGTGGCTCAGTTCTCGTAGGCGCCGAG
CCCGAGGACTTTATTTCGTGCTGATGACCGCGACGAGTTCTTTTTTTACGAGCGCGAGAAC
GAAGAGGATTACGGTCCATCGACCCAGGAGCGGATGCGGCAATATGAACAGCACGCACTG
CCATTGGCTATCAAGTCATCCGTCGCGGCACTTCGGGATGCCGGACAGGACGCACGGGAG
ATTACTCATATCGTACGGTTTTTCGTGCTCAGGCTTCAATAGCCCCGGCGTCGATATGGGA
CTGATTGAAGAACTGGGCCTTAACCGGAATACTTCACGCACTCATGTTGGCTTCATGGGT
TGCCACGGCTCTTTCAACGGATTGCGTGTTGCGCACGGATACACTGCGTCTGACCCCGAT
GCTGTCTGTGCTCATGTGCTCTGTCTGAACCTCTGCTCTCTTACCACATCATTACGGATGGACC
ACCGACAAAGTAATTGCTAATGCACTGTTTGTGATGGCAGCGGCGCGGTAATTTGCCCG
TCGAGCAAAATCGAAGACCGTAAGCCTGCTTACCAGTTGATTTCGGAGCGGTTCCCTTTTG
GTTCCAAACACCAAATTTGCAATGTCTTGGCGCATTTGGAAATCACGGATTTGAGATGACG
TTGTCTCAAAAAGTGCCCGCTCTGATTGAGGAAAATTTGGTCTCTTGGCTCACTCCATTT
CTTGCTCGCGAAGGTTTGACTATTAAAGACATCGCGGGTTGGGCGATCCACCCTGGCGGC
CCTCGGATTCTCGACTCTTGTCTTGCCGCTCTCTCTCTCTCCAACAGCCACGTGGCACCT
GCTCGTGAAGTCCTCGAAAAGCACGGTAACATGTCTGTCTAGCACCATTTTGTGTTGTACTG
GATCGCATGCGTGAGGAACGTATTAATGGTCCAGTCGTGGCCCTTGATTTCGGCCCAGGC
CTGGCTATCGAAGCAATGTTGCTGAAAGCCGAGAACATCTAA

>t3pks_{RbCg}

ATGACCGCTCAGATTCTGTCTGATTGCCACTGCACAACCAACGCATAAAGCCGATTTGGCA
TTCTCTACTCATTGCGCGCAGAGCATGTCTTGTGAAGACGAAAGCCAAGCGGTAAAGTTG
GCTAAGCTGTACCGTCGCACCGGTGTTGACACTCGCGGATCGGTACTCGTTGAGAAAGGT
GACGCTGGCGAACTTACTCAATCATTCTACCCGCCGATGCAAGACGGATGCGATCGTGGC
CCCACTATGGCGACGCGTAACGAGCGTTTCGCGGAAGAAGCGCCAGCGCTCGCGTGTCTGG
GCAGGCCAGGGTGCTTTGGATGGCTCCGGAATCAGCTCTGACGACGTGACTCACGTTGTA
ACCGTGACCTGCACCGGCTTCACTGCGCCCGGTATCGATGTCCAACCTTATCGATAAACTG
GGTCTCCCGATTACGACTCAACGTATCCAAGTCGGCTTCATGGGATGCCACGGACTTGTA
AACGCGCTCCGTACCGCGCGCGGACTTGTTAGCCGCTGACAGCGATGCCGTCGTCCTTATT
GTCTGTATCGAGCTTTGTTCCCTCCATTATCAATATGGTTATGACGCCAGCGTATCGTA
TCTGGCTCCCTGTTTCGCGGATGGATCAGCAGGATTGATTGTAGCCGCTGACGATTGTCCC
GCCTCGGGCGTTGATACGCCCCCTCGGTGAGATTGTGTCCGTAGGTTCCCTGTTTGATTCCGG
GATTCTCATGACGCAATGACGTGGCGGATCGGCGACAATGGTTTCATCATGACTTTGGAA
GCGTCGGTGCCCGGTTTTATTGAAACTAATCTTCGTGAGTTTCTCGTCCCTTGGCTCGCC
AAGTCCGGTCTCGATTTGGATTCTATTGGCGGCTGGGCAGTCCATCCTGGCGGAGTGCGT
ATTTTGCAGTCTGTTGAGACCGCCCTGGAGTTGCCGTCGGGCGCTTTGGACGTAAGCCGG
GAAGTCCTCCGCGAACACGGTAACATGTCCTCTGCCACTTTGGGATTCGTTCTGCAAAAA
TTTCAACAGCGGAACGTCCCTGGCCCCCTGGTTGATGCTGGGTTTCGGTCCTGGCTTGAA
ATCGAGGTTGCAGTTGTCCGCTAA

>t3pks_{SfCg}

ATGCGTATCGCTTCCGTCGGTCGTGCCCTGCCAGAACATGTAGCCTCACAGCAAGAGGTA
ACGGCGGCTTTGATGCAGATTTGGTCTTCGCGGAAATCCGTCACTGACCGTCTGCCGAAG
TTGCTGGAAAACACGCGGGTAGAGCAGCGCCACCTGGTGATGCCGATCGAAAAGTACGGA
GAGATCGGCACGTTCCGGTGAATGTAACGATTTCTGGATCGAGTGCGCCAAGGATCTGGGT
GAAAAAGCGATCTTGAGGGCCTGGAGAAGGTTGGTTTGGGCCCTCAAGATATTGATGCA
ATCTTCACTGTCACGGTAACCGGTCTTGCGTCACCCCTCCCTTGATGCCCGCTTGGTAAAC
CGTATGGGACTGCGTGACGATGTGAAGCGTACTCCCATCTTTGGCCTCGGTTGCGTGCGG
GGAGTTGCTGGACTTAGCCGTGCTGCGGACTATGTCAAAGCCTATCCAGATCAGGTAGCT
GTACTGCTCTCTGTCGAGTTGTGCAGCTTGACCTTTCAACCAGGAGACCATTTCGGTTGCG
AACTTGATCAGCTCGGGCCTGTTTGGAGACGGAGCTGCCGCGGTGGTTCTCGTGGGAGAG
GAGCGTGCGCGGCGTATGGGTCTTGAGGGCGGTCTCCGTGTGCTCGACACGCGGAGCATT
TTTTACAAGGACACGGAGAACATTATGGGATGGCGGATCTCAGAAGGAGGATTCCAAATT
GTACTTTTACCCGCTGTTCCGGAAGTTGCAAAAGAACGGCTTTCTCCTGGCGTGACGCG
ATGCTTGAGAAACATGGAATGCAACGGGACGATGTGCAAGCTGGATTTGTACCCAGGT
GGACCGAAGGTTCTGTGCGCGATGCAGGAGGGTCTTGTTTGGACGATTCAAAAGTACAA
CACTCGTGGGACGCACTTGCTCAGCTTGGAATTTGTCCAGCGCATCTGCACTGATGGTA
TTGCGTGCTCATCTGGATGCGGGCTCACCCGGCCAATTCAAGCGGGGCCTGGCTCTTGCT
ATGGGACCCGCTTCTGTTTCAAGATTGCTTCTGTTTGAAGTGGTGCTAA

>t3pks_{SmCg}

ATGATCATCGACGGTTTGGGCACTGCAACTCCTACTCACGGAATTTTCGAGTCTAGCGCC
GCACGCATTGCCCAGACGTTGTGTTGTGATAACCCCAAGCAGAGCCAACTCCTCTCCATT
CTCTATCAACGCTCAGGTGTGAATTACCGTCATTCAGTAGCTCTGGAGTCTTCGGAGCCT
CGGTCTGTATCGGACCGTTCCCATCGCGTTGCCGTACAGGCGGCAGGTGGCCGTGTACCG
GCGGCTGCTGATGTTGCGGTGGAAACCCAACCGGTTTCGGCAAACCTTTCTTTCCACCAGCA
AGCTCACCTGAGGACTGTGGTCCCACCACTGCTCAACGGATGGAGCTTTATGAGCGCCAC
GCCGGACGTCTCGGCGCGCAGTCTGTGCAAAAGCACTCGACCAGTCCTCCTGTAACGCG
GGCGACATCGACCATTTGATTACGATCTCCTGCACCGGATTCTATGCCCCAGGACTTGAC
GTTTCCATTATTAATCGCTGCGGCTTGAGCCCCTCCGTGCGTGGACCCATATTGGATTCT
ATGGGCTGCCACGGAGCGCTCAACGGCCTCCGGGTGGCACAGTCTCTGTACAGTCTGAAT
CCGCGTCTGCGGGTTTTGCTTTGCGCAGTCTGAGCTGTGCACTCTTCACCAACAGTACGGA
TGGCATCCAGAACGTGTGTCGAACGCGCTTTTTGCGGATGGTGCGGCAGCTCTTGTA
GCACGCACCTCGTCAAGCCCCCGTCAGGACCTAAGGTGATCGCGTCAGGTTCAACCATC
TTGCCCCGATAGCGAGGATCTCATGCGGTGGCGCATCGGAGATCATGGTTTTGAAATGACG
CTTAGCCACGCGTCCCGACTGTAATCCGGCAGAGCCTGCAACCTTGGTTGCGGGATTGG
CTGGCCAAACATCATTTGGCTATCGAGGACATCCGTGGCTGGGCTATCCATCCAGGAGGC
CCCAAATTTCTGCAAGCCTGCGGCGACGCATTGCAATTTTCAGCGGGATGAGCTGAGCATT
TCCCGTTCTATTCTTGCGGAATATGGCAATATGTCATCGCCAACTGTCCTTTTTATCCTG
CAACGCCTGCTCCAAGAACCTGACCGTCTCCCTTGTGTTCTCTTGGGTTTCGGACCTGGT
CTCTCTATTGAAGCCGCATTGATTGGATAA

>mtr_{GmCg}

ATGAATTTCAAAATCCGTCAGCGTGAGCCAGAATTGATGGATCAACCTGGCCTTTCCGCA
GGCGAGCATGGACGTGCTCTGGCCGGACTGGGCCGCATTAACCTGGTGGAGCCGCTCCGAT
GCCATTGTCTGGCCTGCTGTGTTGGAATCGGCACGCCGCGGTGCGGAAAGCCTTTGCAA
ATCCTCGATATTGCGTCTGGCGGCGGAGATGTAGCACTTTCAATCGCTGCTCGCGCTGAA
CGTGCTGGCATTCCCGTGAAATCGATGGTTGTGATATCAGCCCTTTCGCTATTAGCTAC
GCTACCCAGCAGGACGAGCTCGCGGCCTCGAACAGGTACGGTTCTTTGAATGTGATGTA
CTGACTCAATCCCTCGACAAACAGTACGATGTGGTAATGTGCTCACTCTTCTTGATCAT
CTTGATGAGGAACCAGCCAGCAGCTGCTCAAGATTATGGCCGAATCGTCTCGCGAACTG
ATCCTCGTAAATGATCTTTGCCGTTCCCGCGCCGGTTATTGCCTGGCATGGACGGCGTGT
CGTCTCTTGACGCGGTCTCCTGTAGTGATGCGGATGGACCGCTCTCTGTCGCTGGTGCT
TTCACTGTCAACGAGATTACCGAACTTGACGCGCAGGACGGCCTTTCGGGTTTTTCAGGTG
ACCCGTCACTGGCCACAGCGTTGGTTGTTGAAATGGAGCCGTAAATAA

>fdo_{GmCg}

ATGACTTTTGACGTCTTCACTGAATCGTGAGACTGCATGTGGTCAGTCTTGGGATGTTATC
GTCATCGGAGCGGGCCCGGCTGGAGCCGTAGCCGCCCGGCAGCTTGCGCTGCAAAAAGCTC
CGCACCTTGTGATTGAGCGTAAACTTTCCCGCGTTATAAGGTTTGCGGTTGCTGTCTT
AACCAACGGGCGATCAATGCTTTGCAACAGATCGGCTTGGGCCAGCTGCTGGAATCGGCT
GGCGCCGTGCCTCTGACGCAATTCCGTATGCAGTACCAGGGACGCAAAGCAGCAGTTTCG
CTTCCCGGAGGCCTCGCCATTTCTCGTGAGGTATTCGATACCCTTCTTGTCCAAAGCGCA
ATCGACGCTGGTGCGGCGTATCTCCAGGAAACCTCCGCTACCATTGTGGACTCGGCAGAG
AAGCGTTCGTGCGGTACCGTGGAACCTGAATTTCCAAGGTCATACGGCCGGTAGCGTCAGC
GCCAAAATCGTGCTCGCGGCGGACGGCCTCGGTTCATCCTTCGTTGCAGAATTGCGCGGGA
TTTGACATGCAAGTAGCTTCAGACTCTCGCATTGGTGTGGAGCGGCCTTCGTAGTTAAA
AACCAGTCGAGCTATCCCGCAGGTTTGATCCACATGGCCATTTCATAAGTCAGGATACGTA
GGCCTGGTTCGCGTTGAAGGAGACCTGCTGAACCTGGCGGCAGCTATTGACCGTGACTTC
ATGAAGCAATGTGCTGGTCCCGCAGACGCGGTACTTCAAATTCTTCGCGCTTCAGAGTCA
CCCCTCCTCCCGGAAATGACTGAATCCAACCTTCAAGGGCACCCCTTCCCGTCACGCGGAAG
ACCTTGCGGCCAGTCGGTAACCGCCTTTTCATCCTCGGTGACGCAGCGGGATACCTTGAG
CCTTTACACCGGAGAGGGAATGGCTAATGCCATTACTGAGGGAATGGCAGTAGCACAGTTG
GTCCACAAGGACTGCGGCATTGGGATTGAGCACTCGAACAGAAATGGTTGAAAGCCCAT
TACGCGCTGACGGAGGGTCTGCTTTTTCTGGTGCCGGTGGCTTGACAGCCTTCTCCGGTGC
CCAGCAGCTGTGGGTGGGGTCTCCGCTTGTTCATCTTTTTCCCGGTATGGCACGCCCT
ATCGTAGCACGGCTTAACCATTA

>mtr1_{sfcg}

ATGGCTCAAGCACAGAATGACTTGGAGATTTACGATCGTGAGGCGGATGCCTGGTGGGAT
GGTTCTGATCGTTTTTTCCGGTCCCTTCGTAGCGTTAAAGAGTTCCATTTGGAGCTGCTG
TTGGAACGCTGGGCGAGCGAGCCATCGGGCGCACGTGTAGTTGATCTCGGATGCGGCGGA
GGTTTGTTCTCGATGGCTCTCGCGGATCGGGGTGCCGATGTAGTTGGTGTGGACCTTAGC
GGTCCCTCCCTCGGTTCCGCGGCACGTGAGGCGGGACGGCGCGGTTCATGCAGCTCAATTT
GTGCGGGGTGACTTGACTTCGTGCGCCAGTACGTGATGGTTGGGCTGACCTTGTTCTCTTG
TCAGATGTCCTTGAACATGTGGAAGATCCGAAGGCCGCTGTGATCGAGGCTGGCCGTCTG
CTGCGGCCCCGGTGGCACGCTCTTGTTGAATACCTTCGATCGTCGGCCGGGTGCGGGACTC
CTCGTTGTTACTGTGCGCCGAGGTCTCGGTCTCGTCCCAAAGGGAACGCACGACCTGCA
ATGTTCTGTCCTGCTGAGGTTGATGACGCAGCTCGTGAAGCCGGTATGACTCGCACT
CACTTGTTTGGGAGGCGCCATCATTGGTGAAGACGGTCCGGACCTGGACCATCCACATG
CGCCGCGCACCGCGCGGATTCGCGTATACCGCGTTGTATGAGAAGGCACAGGCATAA

>mtr2_{sfcg}

ATGACGGAGATTGCCCTTCGCCAACGCGGCCACGTGATACTGCGCTTGATCCTTCATTG
TCCACCTTGGCACGCCGTGTGATCGCCGGCACGTACGGCTTGGTCAATCACGGCTCCTTC
GTCGTAGCTGTCGCCGCTATGGCATGGGGCCTTTACCACGGTATGTCGCAGGAGGGAGTG
CCGCACCTGCAAGTCGGACTCGAAGGAGGATGGCGCTGGCTTTGGAACGCCCTTCTTCTT
TTGCAATTTCTCTCATTCATAGCTGGCTTTTGACGGCTCGGGGTGCGCGTTTCATGGCC
CGGCTCGCACCTTTTCGGCACCGGACGCACGCTGGGTGCGACCATCTATACTACGTTGGCG
TCCTGGCAGATCTTGTTTGCGTTCTCCGCTTGGGTTCGGATCGGTAAGGCCTCTTGGTCC
CCTGAGGGTTGGCTCCTCCATGCTTGGCAAGTTTTGTTTCGAGGTTGTTGGTTGTTTCTC
GTTAAAGCGCTTAAGGACGGAGGAATGACCCTCCAGACTGGATCTCTGGGTGGATGGCC
TTGTGGAAAGGCGAGCGGCCACGTTTTCCGGGCTTGCGTACGCAGGGCACCTTTGCGGCG
TGCCGCCAACCTATTTACCTCGCCTTTGCATCTCTTTTGTGGACTGGCCCTGTTTGGACC
GCTGATCATTTGGCGGTGGCCGTCTCTGGACCGTATACTGCCTGGTAGGACCAATTCAC
AAAGAGCGGCGGTTCCAATCCATGTACGGTGATGCCTTTATTACTTATAAGTCGCGTATC
CCCTATTTCTGTGCCCCGTCTGAAAGGACGGTCTCGCACGCCCCAGCGGCCGTAA