

Supplementary Information

Title: Functional metagenomic discovery of bacterial effectors in the human microbiome and isolation of commendamide, a GPCR G2A/132 agonist.

Authors: Louis J. Cohen^{1,2}, Hahk-Soo Kang¹, John Chu¹, Yun-Han Huang¹, Emma A. Gordon¹, Boojala Vijay B. Reddy¹, Melinda A. Ternei¹, Jeffrey W. Craig³, Sean F. Brady^{1*}

*To whom correspondence should be addressed: Email: sbrady@rockefeller.edu

Author Affiliation

¹ Laboratory of Genetically Encoded Small Molecules, Howard Hughes Medical Institute, The Rockefeller University 1230 York Avenue, New York, NY 10065

² Icahn School of Medicine at Mount Sinai, 1 Gustave Levy Place, New York, NY 10029

Current address:

³ Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.

Corresponding Author: Sean F. Brady

Contact

Laboratory of Genetically Encoded Small Molecules
The Rockefeller University
1230 York Avenue
New York, NY 10065

Phone: 212-327-8280

Fax: 212 327 8281

Email: sbrady@rockefeller.edu

Keywords

Functional Metagenomics, NF- κ B, N-acyl amino acids, 3-OH N-acyl palmitoyl glycine, commendamide, Cbeg, eDNA

1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6
7
8
9
0

Table of Contents

SI Materials and Methods

Supplementary Figures

Fig S1: ppGpp synthetase Cbegs

Fig S2: Additional analysis of Cbeg4, 5, and 6/7

Fig S3: Cbeg12 functional and phylogenetic characterization

Fig S4: Bioassay (HEK293:NF- κ B:GFP) guided purification of compounds 1- 4.

Fig S5: Z prime calculation and microplate layout

Fig S6: NMR Spectra for compounds 1 -4

Supplementary Tables

Table S1: Patient demographics

Table S2: A) Cosmid and Cbeg sequence analysis, B) Gene tables for each unique cosmid

C) Swissprot analysis

Table S3: NMR table for compounds 1 -4

1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6
7
8

SI Materials and Methods

Storage and Robotic Array of Metagenomic Clones in 384 Well Microplates

Metagenomic clones from the CD and HC libraries were directly plated on Luria broth (LB) agar supplemented with 15 µg/ml tetracycline. Clones were individually selected using a colony picker (CP7200 Hudson) and placed into wells of a 384 well microplate containing LB medium (tetracycline 15 µg/ml). Negative control clones (*E. coli* EC100 cells containing the pJWC1 cosmid vector) were plated into columns 2 and 11 of the microplate. Plates were grown overnight at 37 °C, glycerol was added to a final concentration of 25% and plates were stored at -80 °C. Metagenomic clones from the UC patient library clones were arrayed using a liquid dispensing robot (Multidrop™ ThermoScientific). An overnight culture was diluted to ~1-3 clones per 100 µl of media and 100 µl of media was dispensed into each well of the microplate well. Microplates were stored at -80 °C in the same fashion as the other libraries.

Z' for NF-κB Reporter Cells

HEK293:NF-κB:GFP cells were tested for robustness of NF-κB induced GFP expression using TNF α (Invitrogen) as a positive control and the TNF α activation inhibitor CAPE (caffeic acid phenethyl ester, Santa Cruz Biotech). Cells were incubated with TNF α or TNF α + CAPE for 24h and imaged using ImageXpress Micro XLS Widefield High Content Imaging System (Molecular Devices). Images were processed using MetaXpress (Molecular Devices) and the percent of cells showing GFP expression was calculated on a per well basis (Fig S5). At 10 ng/ml TNF α the near maximum GFP expression was observed. CAPE was added at different concentrations to wells containing 10 ng/ml TNF α. The maximum inhibition of TNF α mediated GFP expression was observed at a concentration of 10 µg/ml CAPE. At these concentrations the mean percent GFP expression of the TNF α positive control group was 74.1 with a standard deviation of 2.76. With the addition of 10 µg/ml CAPE the mean percent GFP expression was 24.3 with a standard deviation of 5.01. The Z prime score based on these values is 0.61 (Fig S5). (1)

16s PCR Analysis and Data Processing

1 100 ng of cosmid clone DNA from each of the three patient libraries was mixed with 0.5 μ M primers, 1x Buffer
2 G (Epicentre), 0.2 U Phusion polymerase, 2.5U rTaq polymerase (New England Biolabs), 3% DMSO. Primers
3 targeting the V4 rRNA region were used (F primer - 5'AYTGGYDTAAGNG, R primer -
4 5'TACNVGGGTATCTAATCC). (2) 454 sequencing adapters and unique barcodes were attached for
5 processing and sample identification. PCR conditions were 25 cycles of 10s at 98 °C, 30s at 55 °C, 30s at 72
6 °C followed by an extension step for 5min at 72 °C. PCR reactions were run on crystal violet stained gels and
7 amplicons of the correct size were gel purified (Qiagen MinElute). Purified PCR products were fluorometrically
8 quantified (PicoGreen Quant-iT, Invitrogen) and analyzed via capillary electrophoresis (DNA 7500; Agilent
9 Technologies). Each sample was diluted to $\sim 10^9$ molecules/ μ L. All samples were pooled for emulsion PCR
0 and pyrosequenced (454 GS-FLX Titanium). Raw 454 reads were processed using QIIME. (3) The general
1 workflow is summarized as follows: 1) sequences were denoised and quality filtered. 2) OTUs were assigned
2 at 97% identity and representative sequences identified in each OTU. 3) Each OTU representative sequence
3 was assigned taxonomy using Greengenes. 4) Per patient sample OTU tables were summarized by taxonomy
4 and diversity statistics calculated.

6 Subcloning of Cbeg12-1

7 The following primers were used to PCR amplify the Cbeg12-1 gene from cosmid Cbeg12 (35i19): F primer -
8 GATCAGTACTGCTTGGTAAACAGGAGGTTT R primer - GATCAGTACTTTATTTGGTCAGTATGGGAT. The
9 resulting PCR product was digested with *Sca*I and ligated into *Sca*I digested pJWC1 vector.

1 Isolation of N-acyl amino acids from Cbeg 12

2 *E. coli* transformed with the Cbeg12-1 (35i19) cosmid was and a negative control culture (*E. coli* EC100 with
3 empty pJWC1 cosmid vector) were grow in LB medium supplemented with tetracycline (15 μ g/ml) (4 days, 250
4 rpm, 30 °C). Both cultures were extracted with ethyl acetate 1:1 and the resulting extracts were dried in vacuo.
5 Extracts were fractionated by silica gel flash column chromatography (chloroform:methanol (0.1% TFA): 99:1
6 isocratic for 5 min, gradient from 99:1 to 70:30 over 10 min). Similar fractions were combined based on TLC
7 analysis and then combined fractions were assays for ability to induced GFP expression in HEK293:NF-
8 κ B:GFP cells (Fig S4). Only the combined fraction containing clones specific metabolites was found to induce

1 GFP. Compounds **1-4**, which were not seen in the corresponding fraction from the vector control culture, were
2 purified from this Cbeg12-1 fraction using reversed phase HPLC (Waters XBridge[™] C₁₈, 10 X 250 mm, 3
3 ml/min, isocratic 70% acetonitrile with 0.1% TFA) (Fig S4). Purified **1-4** were resuspended in DMSO and
4 assayed for the ability to induce GFP expression in HEK293:NF-κB:GFP cells.

5 6 Structure determination of compounds **1 - 4**

7 NMR spectroscopic data for naturally isolated compounds **1 – 4** appears in Figure S6, Table S3.

8
9 Structures of compounds **1 – 4** were determined by spectroscopic analysis including MS, and 1D and 2D NMR.

0 Compound **1** was isolated as the major compound. HRMS calculated for compound **1** C₁₈H₃₆NO₅⁺ 330.26,
1 found: 330.2859. The odd molecular weight suggested that compound **1** contains an odd number of nitrogens.

2 Analysis of the ¹H NMR spectrum of **1** in combination with DEPT-Q and HSQC spectra indicated the presence
3 of NH (δ_H 8.23, t), oxygenated methine (δ_H 3.78), two diastereotopic methylene (δ_H 3.78 and 3.85; δ_H 2.19 and
4 2.23) and largely overlapped aliphatic methylene (δ_H 1.24) and methyl (δ_H 3.78 and 3.85) protons. Analysis of
5 the COSY spectrum identified the presence of two separate spin systems. A COSY correlation between NH
6 and diastereotopic protons (H₂-2) together with an HMBC correlation from H₂-2 protons to a carbonyl carbon C-1
7 suggested that the first spin system arises from a glycine moiety. The presence of a fatty acid moiety was
8 apparent due to the appearance of a large methylene envelope at 1.24 ppm as well as a triplet methyl signal
9 (H₃-18, δ_H 1.85). Sequential COSY correlations were observed between H₂-4/H-5/H₂-6/H₂-7, and COSY
0 correlations between H-5 and OH indicated H-5 to be hydroxylated. In the COSY spectrum H₂-7 further
1 correlated with the largely overlapped aliphatic signals at 1.24 ppm, deducing the second spin system as a
2 hydroxylated fatty acid moiety. The glycine and hydroxylated fatty acid moiety can be combined through an
3 amide bond based on HMBC correlations from H₂-2 and NH to C-3, suggesting compound **1** to be an *N*-acyl
4 glycine. Finally, the length of a fatty acid moiety was determined as C₁₈ based on the predicted molecular
5 formula. Thus, the structure of **1** was determined to be 3-OH-*N*-palmitoyl glycine.

6
7 HRMS calculated for compound **2** C₁₈H₃₄NO₅⁺ 328.24, found: 328.2657. The ¹H and ¹³C NMR spectra of **2** are
8 almost identical to those of **1** except for the presence of signals for double bonds in the spectra of **2**. The

1 molecular weight difference of two suggested the same carbon chain length between compounds **1** and **2**.
2 COSY correlations between H-11/H-12 and H-10/H-13, and between H-10/H-13 and a largely overlapped
3 methylene signals indicated that a double bond is located on a fatty acid chain. The position of the double
4 bond was not be determined by a spectroscopic methods. The position of the double bond is predicted based
5 on the position that is most frequently seen in *E. coli* lipids and has been seen in all previous *N*-acyl amino
6 acids characterized produced by *E. coli*. (4)

7
8 HRMS calculated for compound **3** $C_{20}H_{38}NO_5^+$ 356.27, found: 356.2975. The 1H NMR spectrum of **3** is similar
9 to that of **2** except for the integration of the methylene envelope at δ_H 1.24. The ^{13}C NMR spectrum of **3**
0 displayed additional signals around δ_C 29.0. This together with a molecular weight difference of twenty eight
1 indicated a fatty acid moiety of **3** to be a 3-hydroxylated oleic acid (C18:1) instead of a 3-hydroxylated palmitic
2 acid that is seen in **2**.

3
4 HRMS calculated for compound **4** $C_{16}H_{32}NO_5^+$ 302.23, found: 302.2501. The 1H spectrum of **4** is similar to that
5 of **1**, however the integration of methylene envelope. The molecular weight difference of twenty-eight between
6 **1** and **4** suggested the fatty acid moiety of **4** as a 3-hydroxylated myristic acid instead of a 3-hydroxylated
7 palmitic acid found in **1**.

8 9 HEK293:NF- κ B:GFP bioassay of purified *N*-acyl-glycines

0 Each purified clone specific metabolite was resuspended at 0.45 mg/25 μ l methanol. 1 μ l of this dilution was
1 added to 80 μ l of sterile spent culture media from negative control clones and the mixture was added to 100 μ l
2 of HEK239:NF- κ B:GFP reporter cells in a 96-well microplate (8,000 cells per well). Each purified compound (**1**
3 - **4**) was assayed at 100 μ g/ml in the first well of microtiter plate and then serially diluted (two fold/well) across
4 the plate. GFP expression was monitored in the same fashion as described above.

5 6 *Bacteroides vulgatus* ATCC 8482 culture broth analysis

7 *B. vulgatus* was purchased from ATCC and grown for 14 days under anaerobic conditions in LY-BHI media.
8 The mature *B. vulgatus* culture was extracted 1:1 with ethyl acetate under neutral and acidic conditions (pH 2-

3). Crude extracts were separated by reversed phase flash chromatography [water:methanol gradient 80:20 to 0:100 (0.1% TFA)]. Fractions were analyzed by reversed phase LCMS (Waters XBridge™) using a binary solvent system (A/B: water/acetonitrile with 0.1% formic acid) from 5% solvent B to 95% solvent B over 40 minutes. Purified **1** was run under the same conditions as a control.

Synthesis of *N*-acyl amino acid derivatives

Glycine was purchased from Fisher Scientific Inc. (Waltham, MA). (Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate (PyBOP) and 6-Chloro-1-hydroxybenzotriazole (Cl-HOBt) were purchased from Peak Polypeptide Biosystems (Louisville, KY). All other reagents and solvents were purchased from Sigma-Aldrich (St. Louis, MO).

***N*-(3-hydroxypalmitoyl)glycine.** This compound was synthesized based on the same procedure as described for **1**. The desired product eluted at 76% solvent B and 0.7mg of purified material was obtained after lyophilization. NMR (CD₃OD-*d*₄): δ 0.90 (3H, t, *J* = 7.1 Hz), 1.1 ~ 1.3 (22H, m), 1.4 ~ 1.5 (4H, m), 2.33 (2H, dd, *J* = 14.2, 8.4 Hz), 2.38 (2H, dd, *J* = 14.2, 4.3 Hz), 3.80 (1H, d, 17.4 Hz), 3.96 (1H, m). HRMS calculated for C₁₈H₃₆NO₅⁺ 330.26, found: 330.2550

***N*-(3-hydroxydecanoyl)glycine.** This compound was synthesized based on the same procedure as described for **1**. The desired product eluted at 43% solvent B and 0.9mg of purified material was obtained after lyophilization. ¹H NMR (CD₃OD-*d*₄): δ 0.90 (3H, t, *J* = 7.2 Hz), 1.1 ~ 1.3 (8H, br), 1.4 ~ 1.5 (4H, m), 2.30 (2H, dd, *J* = 14.1, 8.5 Hz), 2.37 (2H, dd, *J* = 14.1, 4.0 Hz), 3.77 (1H, d, 17.4 Hz), 3.96 (1H, m). HRMS calcd for C₁₂H₂₄NO₅⁺ 246.16, found: 246.1530

***N*-(3-hydroxypalmitoyl)tyrosine.** dl-3-Hydroxypalmitic acid (10.0 mg, 37 μmol) was dissolved in 400 μL of anhydrous DMF and activated by PyBOP (19.1 mg, 37 μmol), Cl-HOBt (6.2 mg, 37 μmol), and DIPEA (37 μL, 2 M solution in NMP) for 10 min. The activated fatty acid was added to a solution of tyrosine (13.3 mg, 74 μmol) in 400 μL DMF premixed with DIPEA (74 μL, 2 M solution in NMP) and shaken at 200 rpm in a 37 °C incubator for 16 h. After removal of solvent by speed-vacuum, the resulting material was dissolved in methanol and

1 purified by HPLC using a binary solvent system (A/B: water/acetonitrile containing with 0.1% formic acid). The
2 desired product eluted at 76% solvent B and 0.4mg of purified material was obtained after lyophilization. ¹H
3 NMR (CD₃OD-*d*₄): δ 0.90 (3H, t, *J* = 7.1 Hz), 1.1 ~ 1.3 (22H, br), 1.4 ~ 1.5 (4H, m), 2.26 (2H, m), 2.85 (1H, m),
4 3.11 (1H, m), 3.85 (1H, m), 4.50 (1H, m), 6.66 (2H, d, *J* = 8.1 Hz), 7.05 (2H, d, *J* = 8.1 Hz). HRMS calcd for
5 C₂₅H₄₂NO₅⁺ 436.3057, found: 436.2973

7 **GPCR Screen**

8 Commendamide (natural and synthetic), natural products **2** and **4** and all synthesized acyl amide derivatives
9 were resuspended at 10 mM in DMSO and sent to DiscoverX for assay. Assays were performed per
0 DiscoverX. Initial screen against GPCR library was done only with commendamide at 10 μM concentration.
1 After identification of GPR 132/G2A activation by commendamide, other acyl amides were similarly screened
2 for G2A receptor activation.

4 **References:**

- 6 1. Shun TY, Lazo JS, Sharlow ER, & Johnston PA (2011) Identifying actives from HTS data sets:
7 practical approaches for the selection of an appropriate HTS data-processing method and quality
8 control review. *Journal of biomolecular screening* 16(1):1-14.
- 9 2. Cole JR, *et al.* (2009) The Ribosomal Database Project: improved alignments and new tools for
0 rRNA analysis. *Nucleic acids research* 37(Database issue):D141-145.
- 1 3. Caporaso JG, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data.
2 *Nat Methods* 7(5):335-336.
- 3 4. Brady SF, Chao CJ, & Clardy J (2002) New natural product families from an environmental DNA
4 (eDNA) gene cluster. *Journal of the American Chemical Society* 124(34):9968-9969.
- 5 5. Petersen TN, Brunak S, von Heijne G, & Nielsen H (2011) SignalP 4.0: discriminating signal
6 peptides from transmembrane regions. *Nat Methods* 8(10):785-786.
- 7

8 **Figure Legends:**

9 **Fig. S1. ppGpp synthetase.** Nine different effector genes are predicted to be ppGpp synthetases. (A) Effector
0 gene protein sequences were aligned to each other with MUSCLE (Macvector). ppGpp synthetase genes
1 have varying degrees of identity to each other. (B) A phylogenetic tree prepared of effector genes using the
2 UPMGA method and bootstrapped 999 times. ppGpp synthetase genes cluster according to species.
3
4
5
6

1 (CAPE 10 $\mu\text{g/ml}$) of HEK293:NF-kB:GFP cells (Table) a Z prime score is calculated with a value of 0.61. (C)
2 Each microplate was arranged as outlined. The yellow column contained 10 ng/ml of TNF α to ensure
3 HEK293:NF-kB:GFP cells continued to respond to stimuli by visual inspection. Blue columns are negative
4 controls and the red line shows excluded wells due to poor reproducibility of edge wells.

5

6 **Fig. S6: NMR Spectra for Compounds 1 – 4.**

7

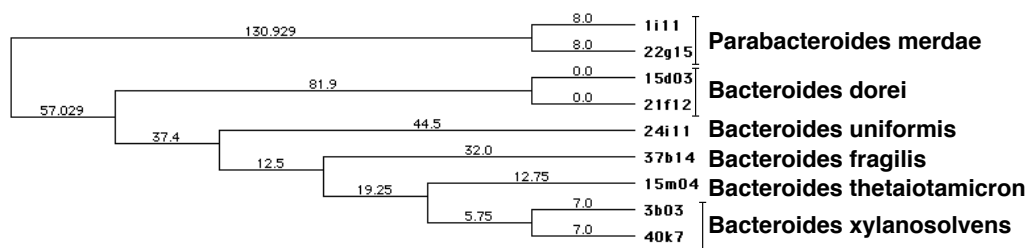
8

9

1 **Figure S1**

	3b03	15d03	15m04	24i11	1i11	22g15	21f12	37b14	40k7
3b03	100	78	96	87	59	54	78	91	98
15d03	88	100	77	76	59	54	100	77	77
15m04	98	88	100	87	59	54	77	91	97
24i11	94	88	94	100	60	55	76	88	88
1i11	75	76	76	76	100	90	59	60	59
22g15	69	70	69	69	99	100	55	54	54
21f12	88	100	88	88	76	70	100	77	77
37b14	95	89	94	94	76	70	89	100	91
40k7	99	99	94	93	76	69	88	96	100

B

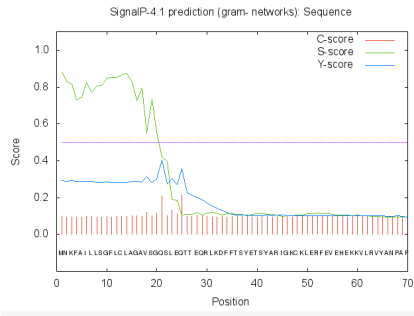


2
3
4

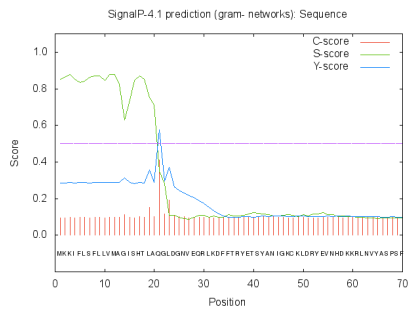
1 **Figure S2**

A

Cbeg4

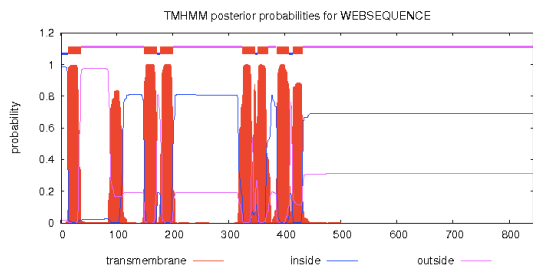


Cbeg5

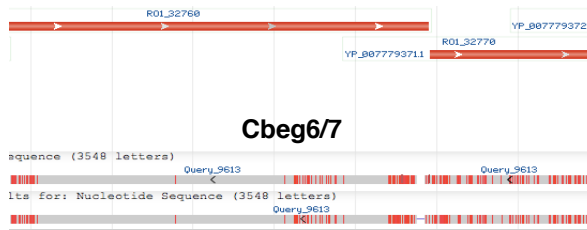


B

Cbeg6



Roseburia intestinalis



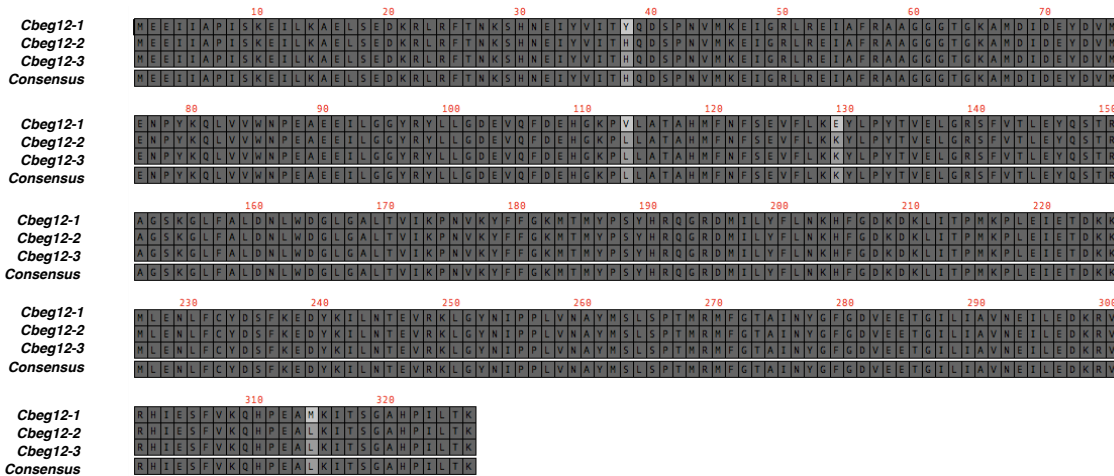
2

3

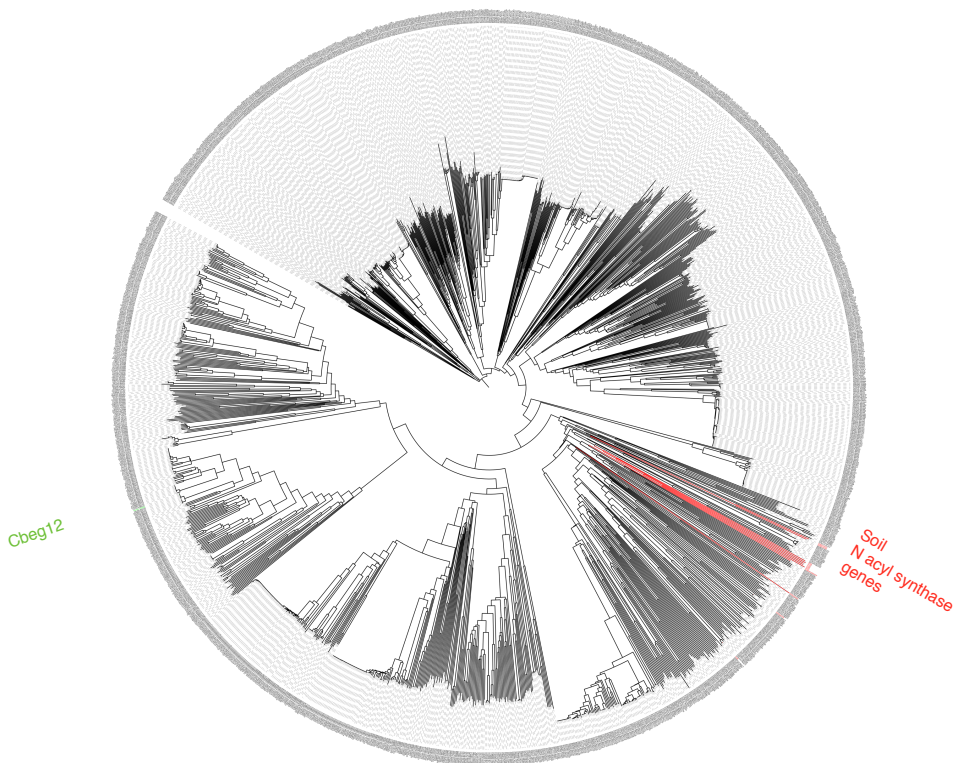
4

1 **Figure S3**

A



B



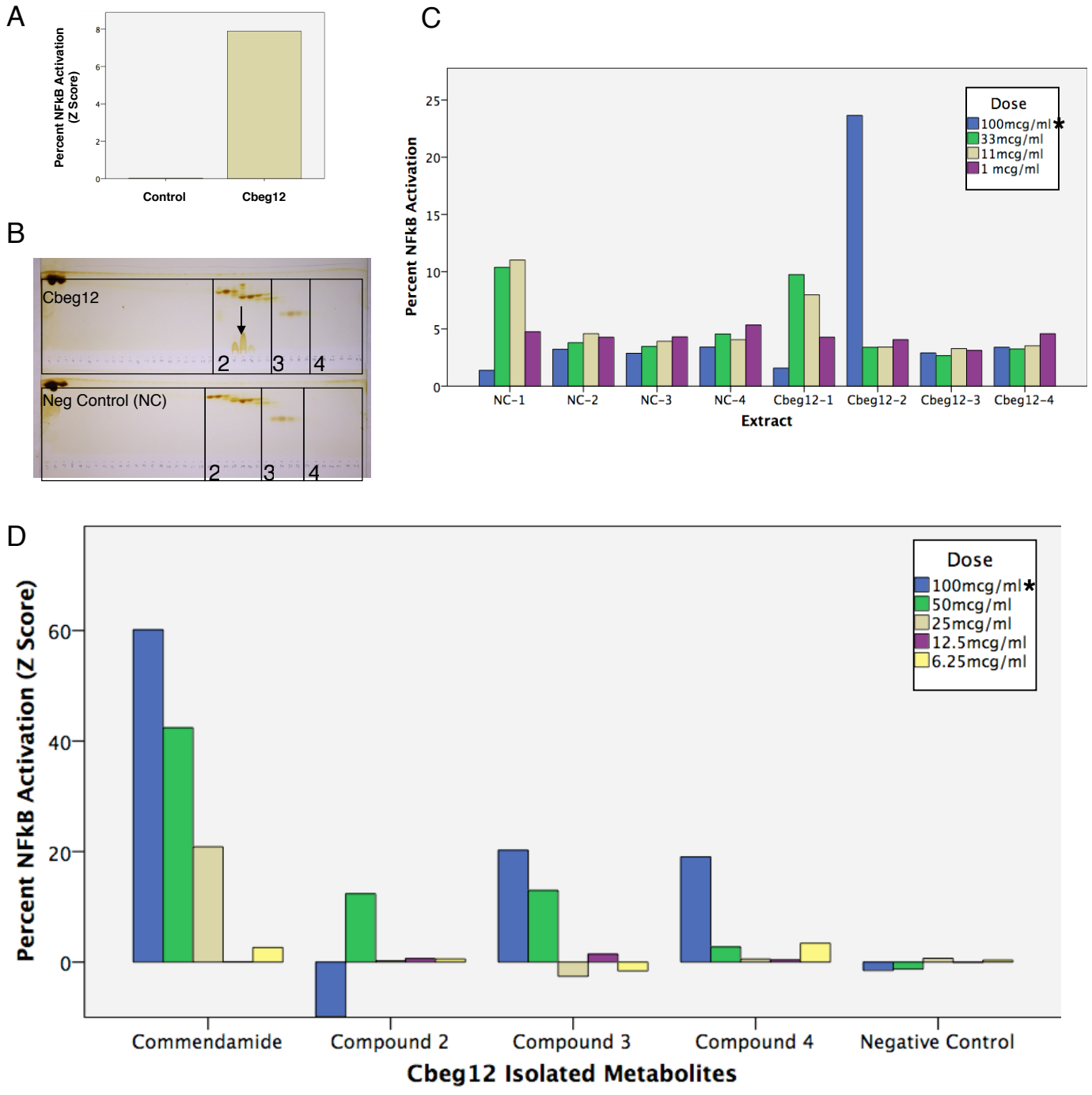
2

3

4

1 **Figure S4**

2

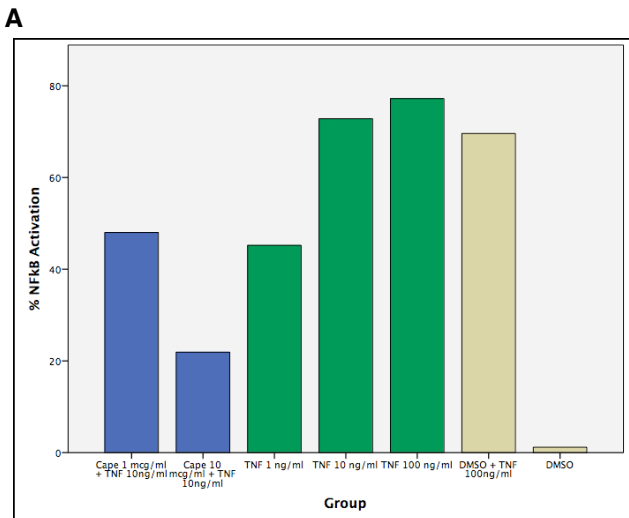


3

4

1 **Figure S5**

2

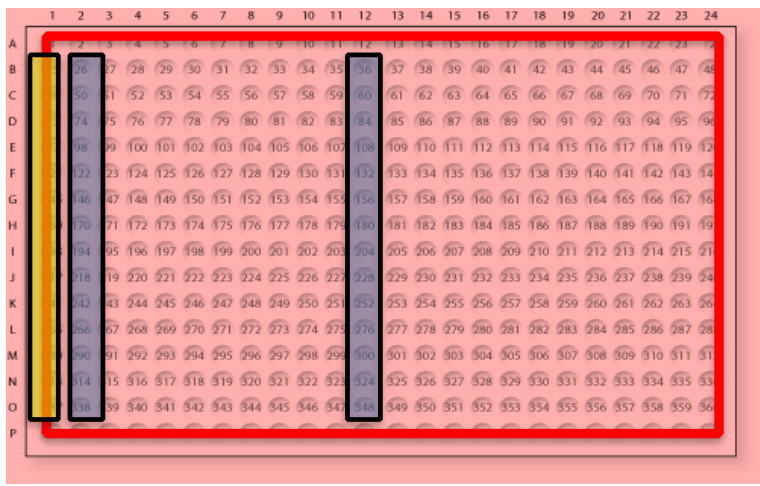


B

	Statistic	Std. Error	
Cape 10 mcg/ml + TNF 10ng/ml	Mean	21.8994	2.08894
	95% Confidence Interval Lower Bound for Mean	17.6390	
	Upper Bound	26.1599	
	5% Trimmed Mean	21.9716	
	Median	22.6405	
	Variance	139.637	
	Std. Deviation	11.81682	
	Minimum	0.00	
	Maximum	44.00	
	Range	44.00	
	Interquartile Range	11.05	
Skewness	-.506	.414	
Kurtosis	.021	.809	
TNF 10 ng/ml	Mean	72.8115	.98159
	95% Confidence Interval Lower Bound for Mean	70.8095	
	Upper Bound	74.8135	
	5% Trimmed Mean	73.2752	
	Median	73.3298	
	Variance	30.833	
	Std. Deviation	5.55271	
	Minimum	53.66	
	Maximum	81.67	
	Range	28.01	
	Interquartile Range	4.79	
Skewness	-1.587	.414	
Kurtosis	4.182	.809	

$$Z' = \frac{3(\sigma_p + \sigma_n)}{|\mu_p + \mu_n|} = 0.61$$

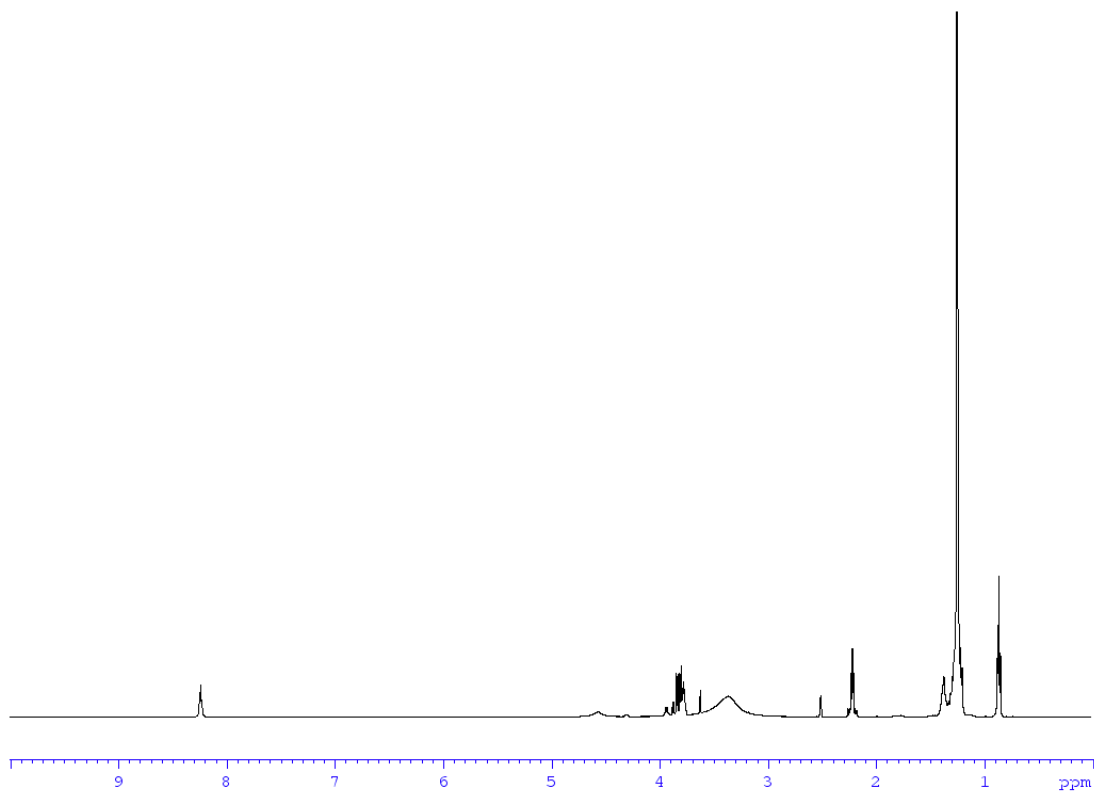
C



3
4
5

1 **Figure S6**

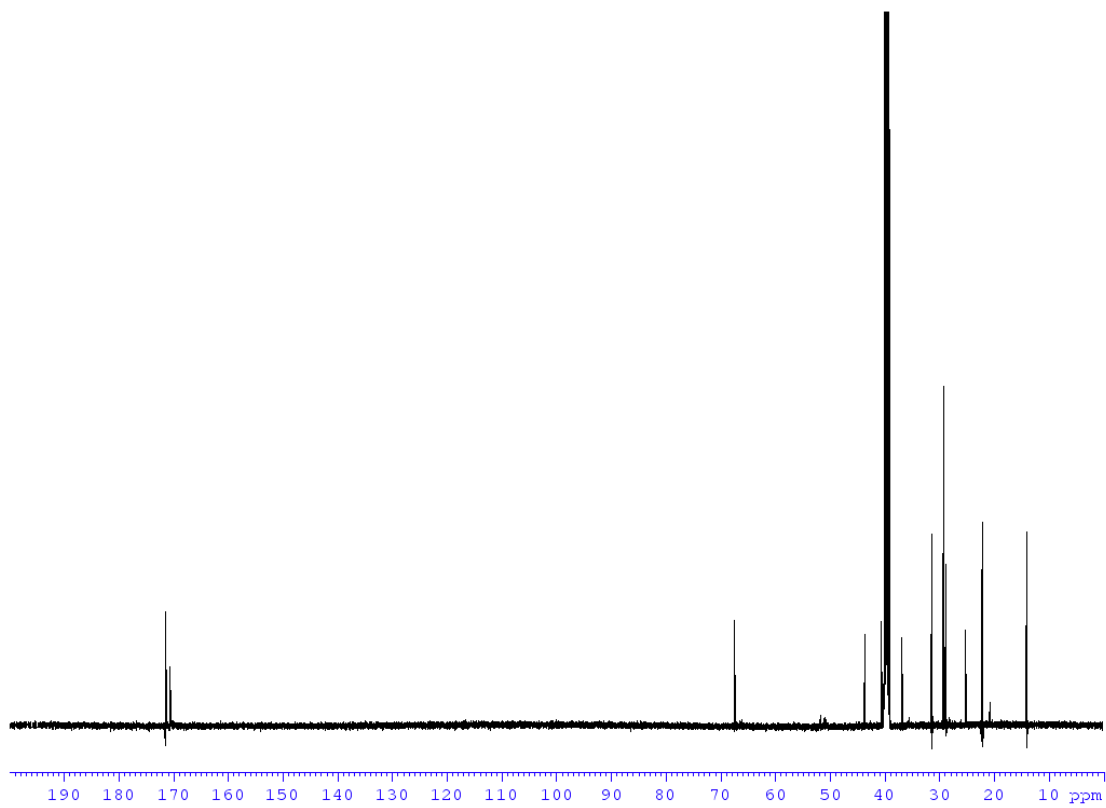
2



3

4

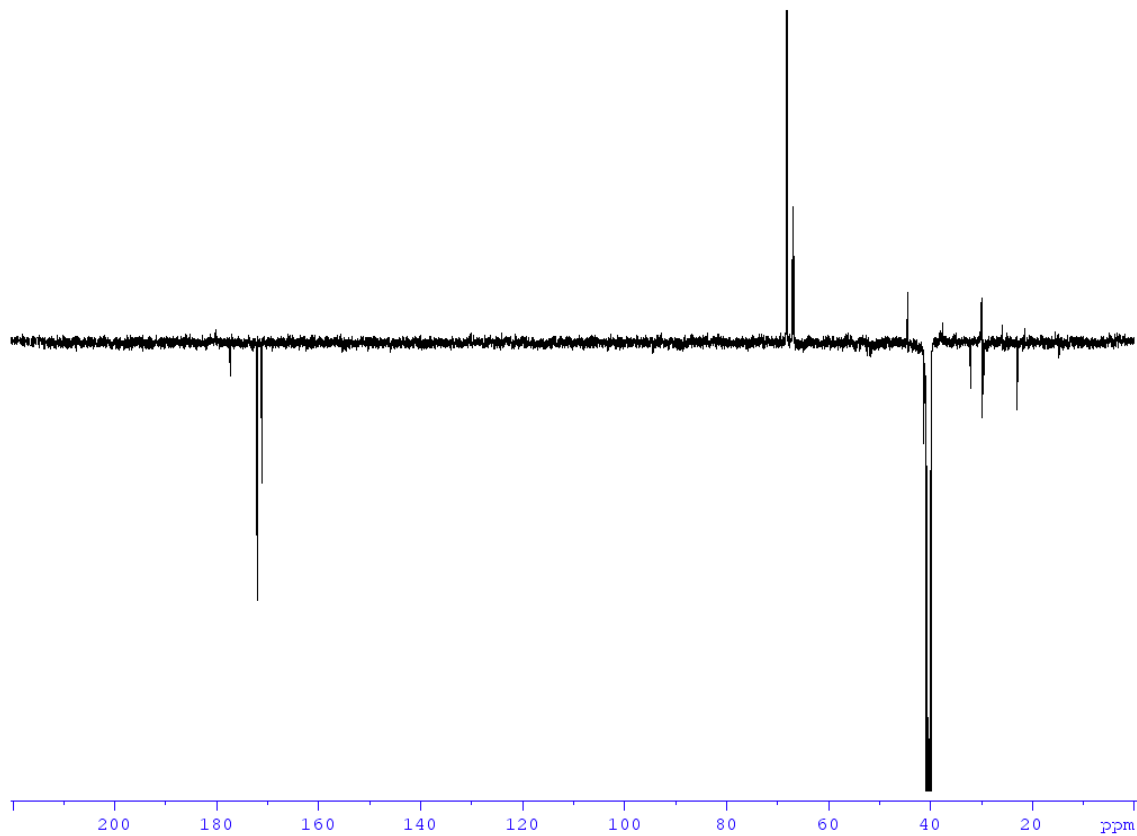
^1H NMR spectrum (DMSO- d_6 , 500 MHz) of compound **1**.



5

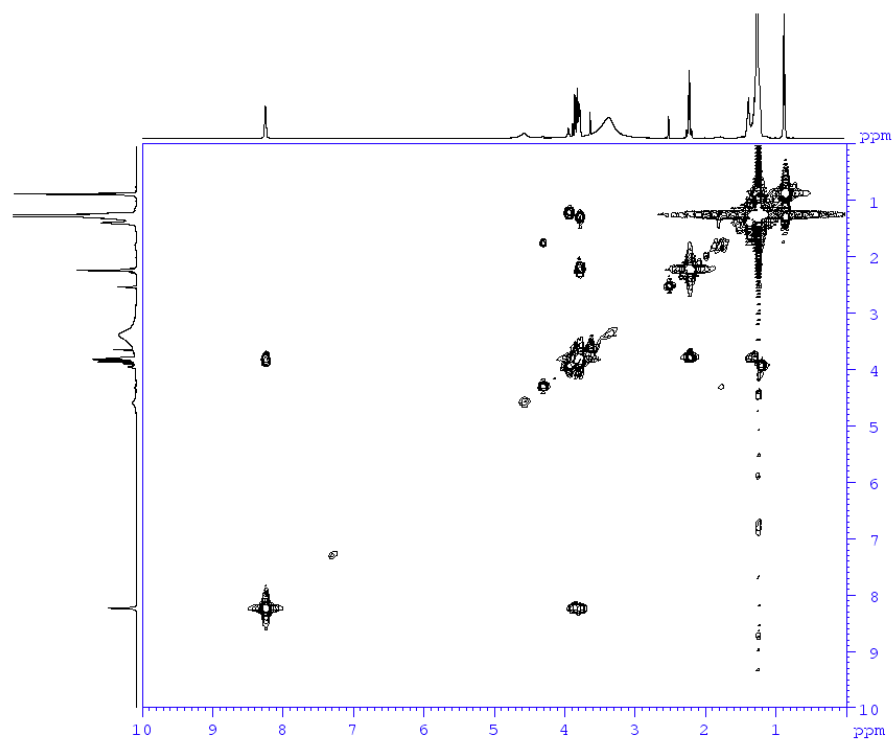
6

^{13}C NMR spectrum (DMSO- d_6 , 125 MHz) of compound **1**.



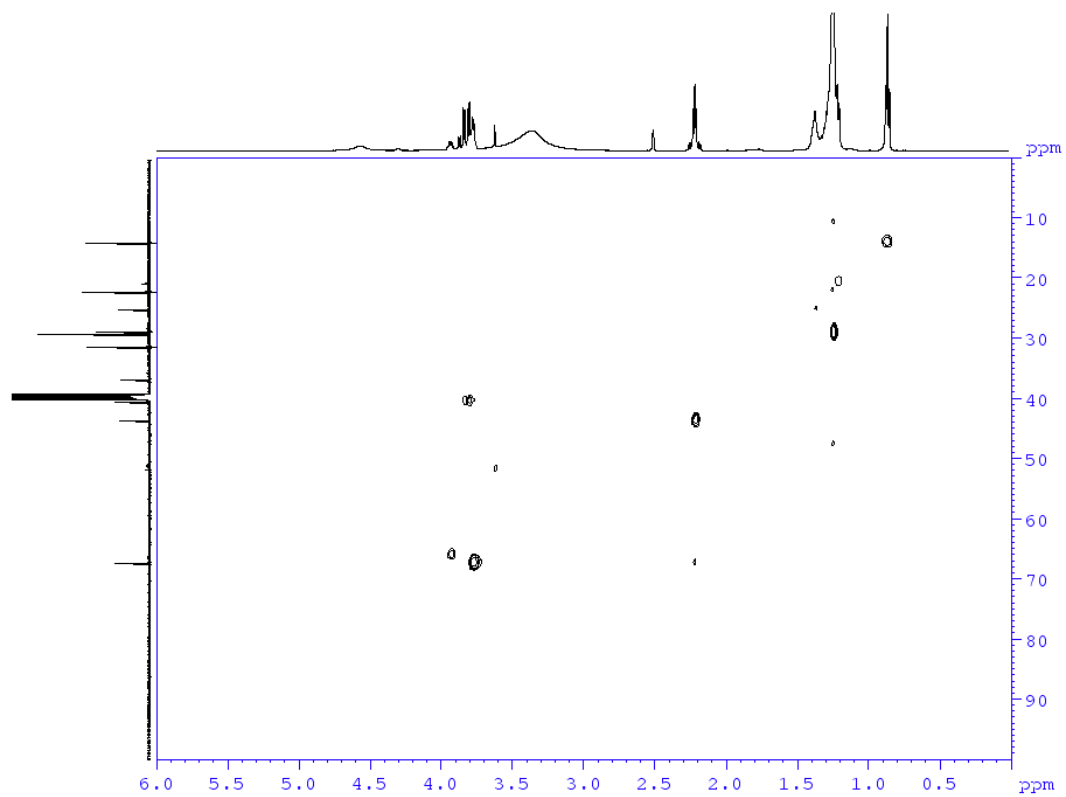
1
2

DEPT-Q spectrum (DMSO- d_6 , 125 MHz) of compound **1**.

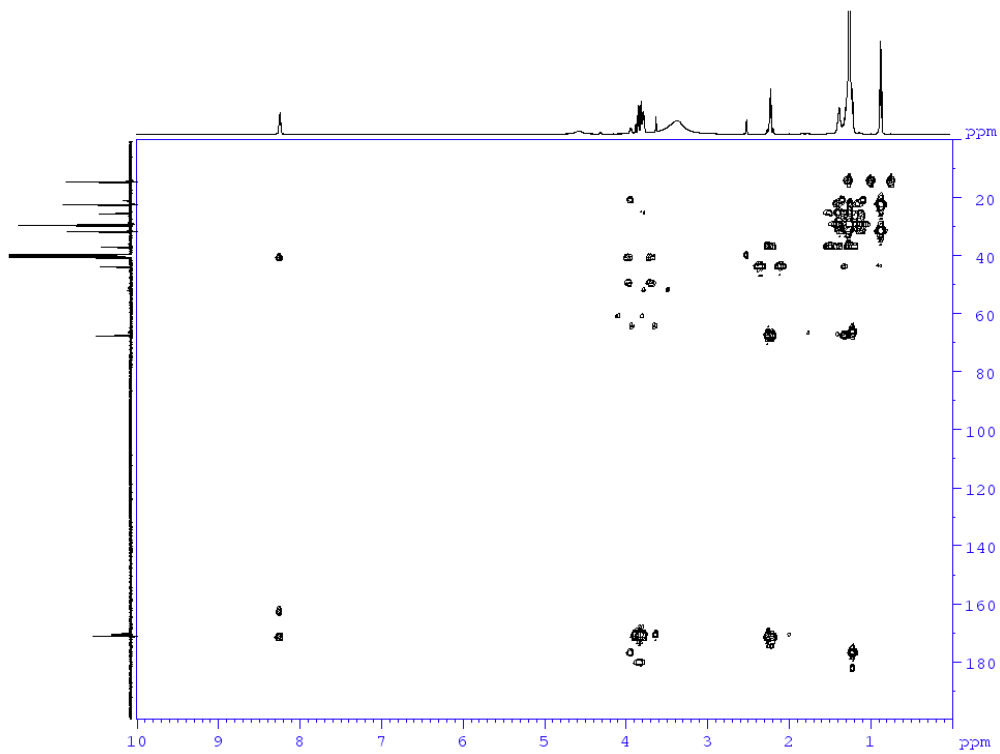


3
4

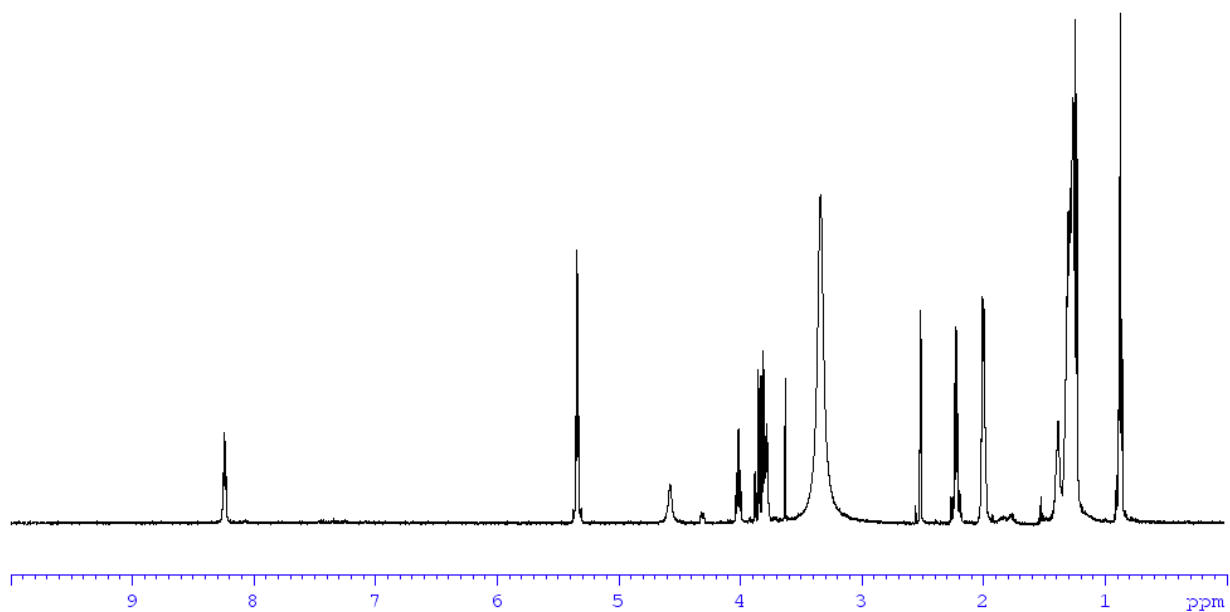
COSY spectrum (DMSO- d_6 , 500 MHz) of compound **1**.



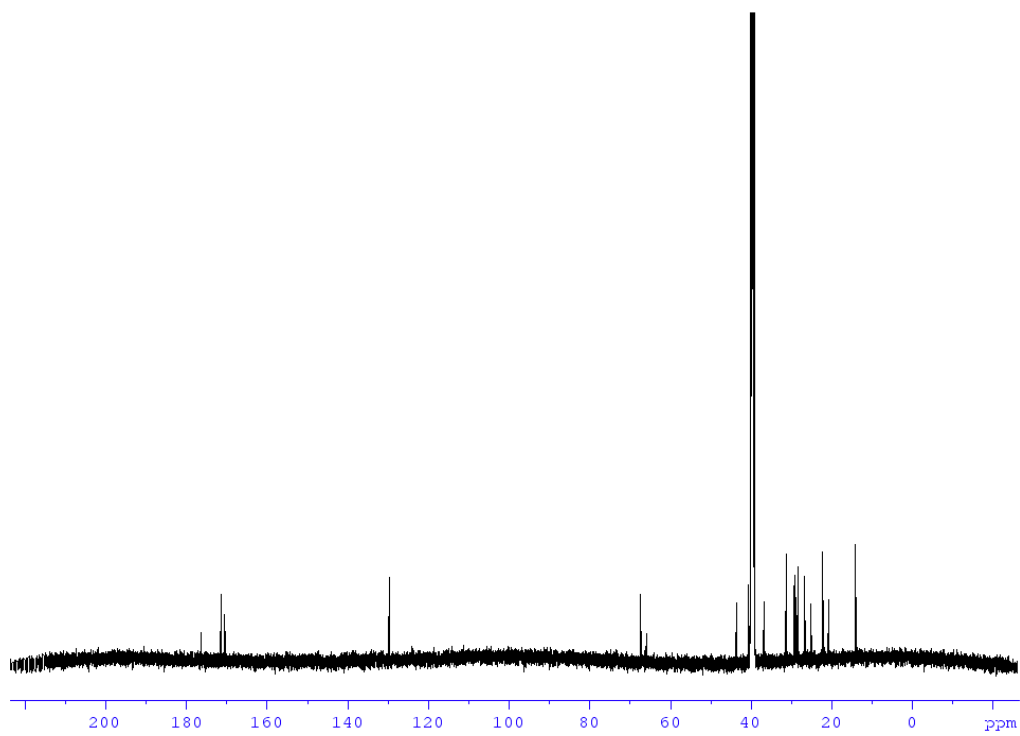
1
2 HMQC spectrum (DMSO-*d*₆, 500 MHz) of compound **1**.



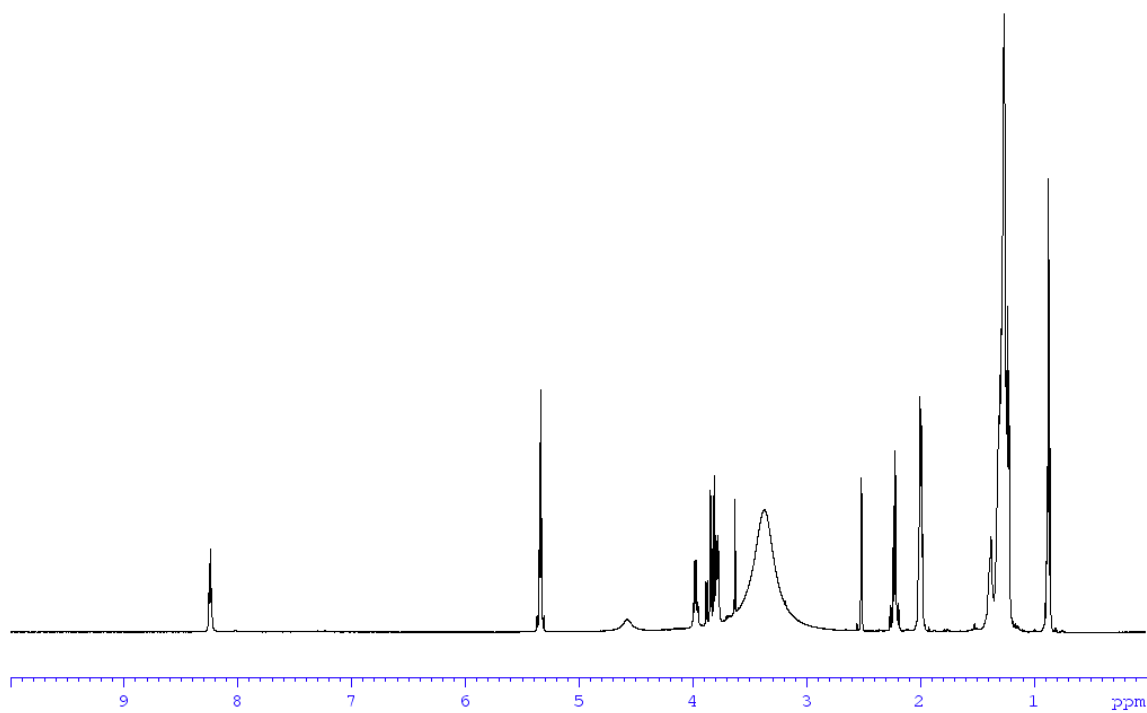
3
4 HMBC spectrum (DMSO-*d*₆, 500 MHz) of compound **1**.



1
2 ^1H NMR spectrum (DMSO- d_6 , 500 MHz) of compound **2**.

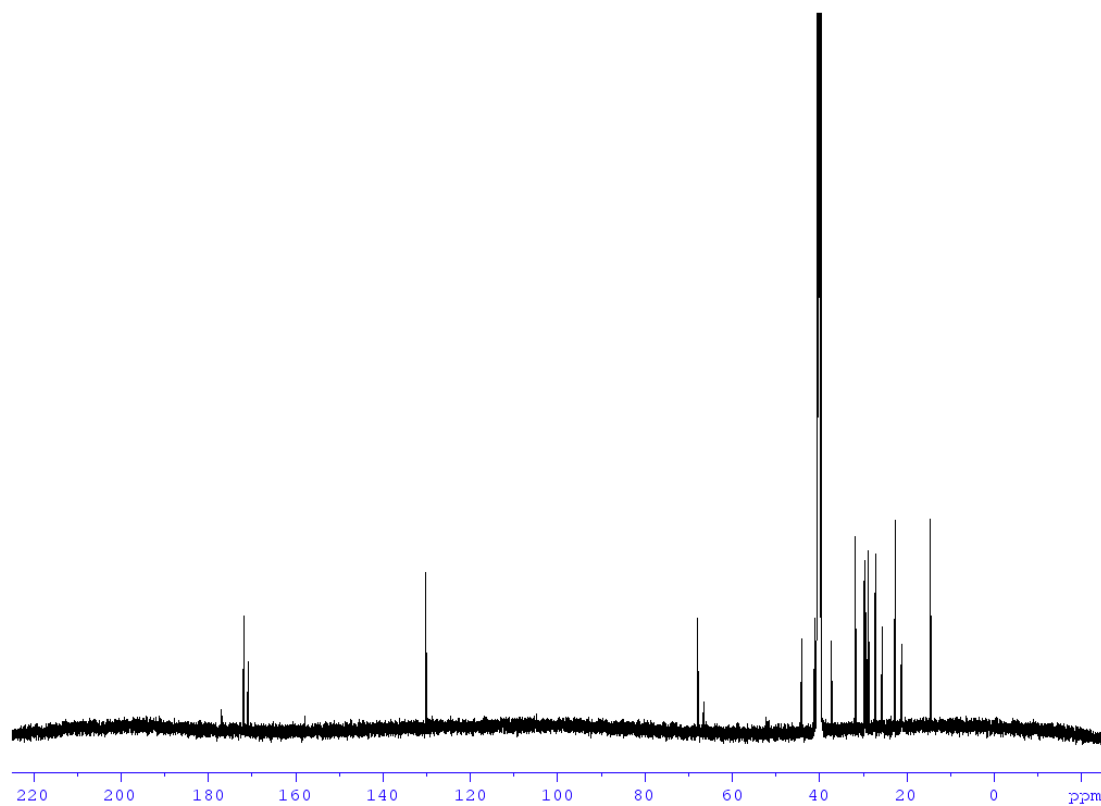


3
4 ^{13}C NMR spectrum (DMSO- d_6 , 125 MHz) of compound **2**.



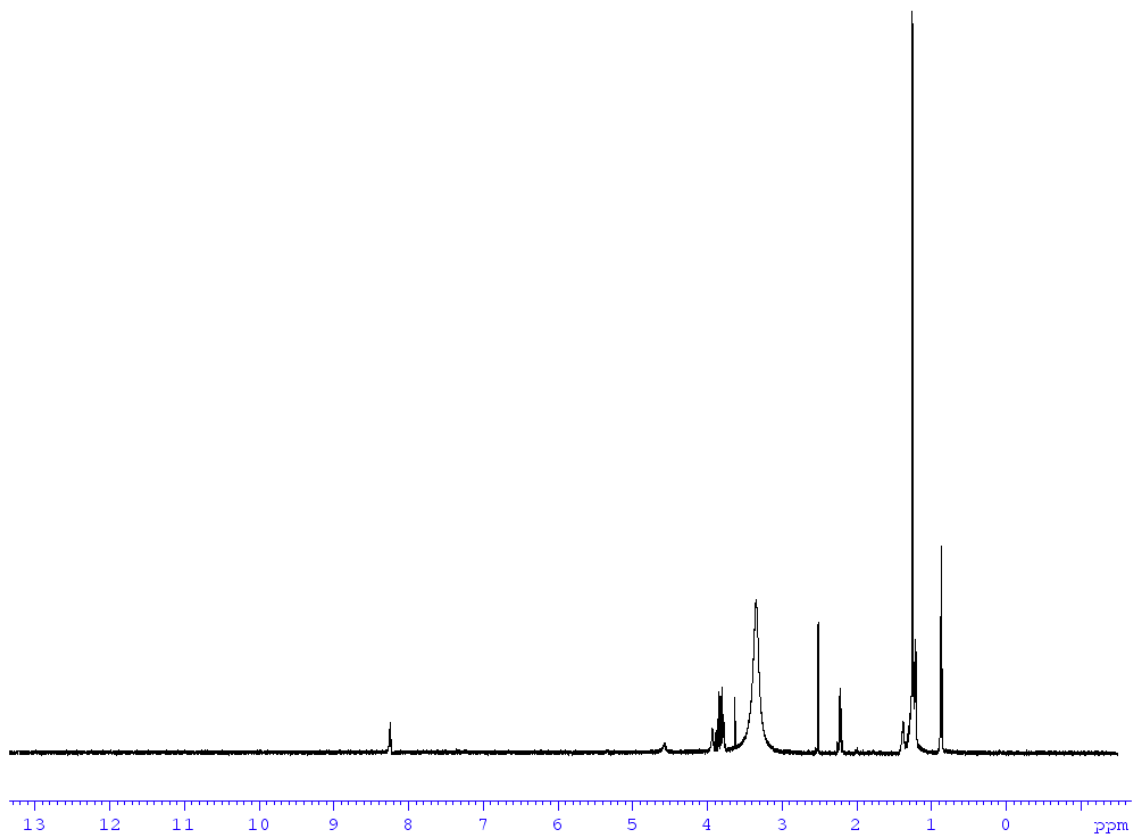
1
2

^1H NMR spectrum (DMSO- d_6 , 500 MHz) of compound **3**.



3
4

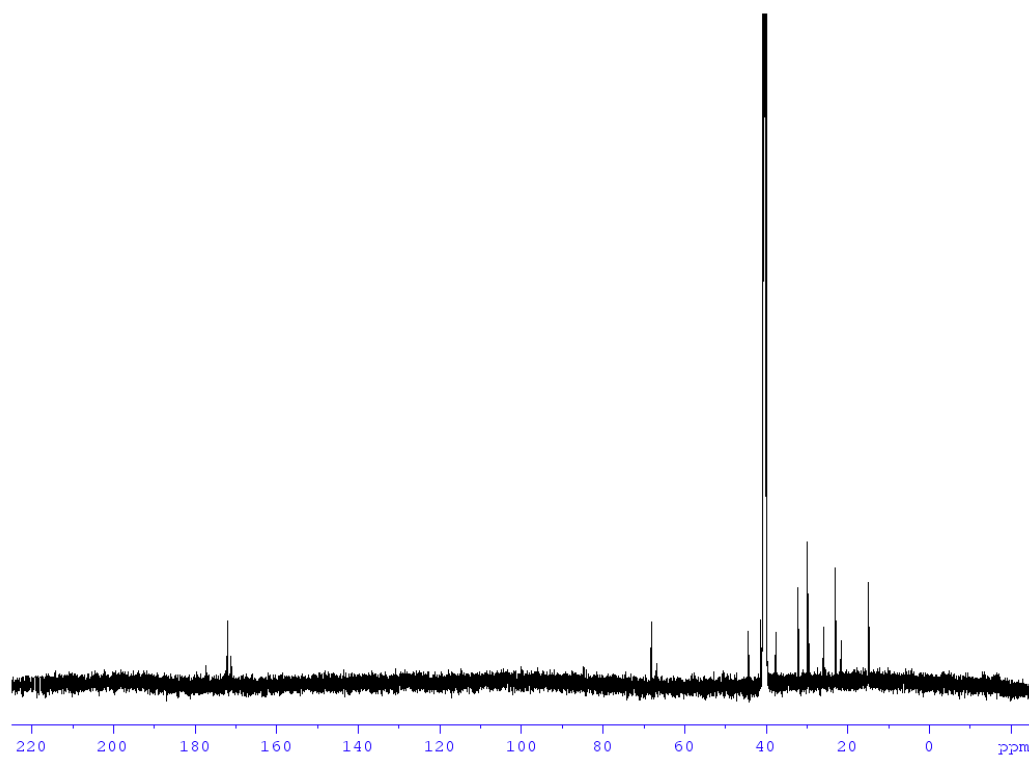
^{13}C NMR spectrum (DMSO- d_6 , 125 MHz) of compound **3**.



1

2

^1H NMR spectrum (DMSO- d_6 , 500 MHz) of compound **4**.



3

4

^{13}C NMR spectrum (DMSO- d_6 , 125 MHz) of compound **4**.

5

1 **Table S1**

2
3

4 **Table S1. Patient demographics**

5

	Crohns	Ulcerative Colitis	Healthy Control
Age (years)	38	34	47
Sex	M	M	F
Medical History*	Seizures	NA	NA
Disease Location	Ileocolonic, Perianal	Pancolitis	NA
Disease Activity^	Moderate	Moderate	NA
Medication	Mesalamine, Levetiracetam	Mesalamine, Cortenema®	None

6
7
8
9

^ Disease activity based on the Harvey Bradshaw Index (CD) and simple clinical colitis index (UC).
*All patients were free of abdominal surgery and antibiotic use > 6 months

0

1
2

1 **Table S2**

2

3 **Table S2A. Cosmid and Cbeg sequence analysis.** Each unique cosmid and the effector gene(s) on that
4 cosmid were independently aligned to the NCBI reference genome dataset using BlastN. For each unique
5 cosmid (blue column) and effector gene (red column) the most similar reference genome is recorded. Each
6 effector gene protein sequence was then aligned to the NCBI nr dataset (BlastP) to identify conserved domains
7 (green column). Conserved domains are presented in descending order (N to C terminus of predicted protein)
8 starting with conserved multidomains followed by individual domains.

9

Cbeg	Reference Genome						Effector Gene				Conserved Protein Domains						
	Unique Cosmid	Library	Metagenomic Insert Size (bp)	Closest Reference Genome	Reference genome %Coverage %Identity	Accession Number	Effector Gene Cluster Size (bp)	Closest Reference Genome	Reference Genome %Coverage %Identity	Accession Number	Conserved Protein Domains	Conserved Domains	pFAM	SMART	COG	TIGR	PRK
ppGpp Synthetase Gene	1111	Library 1	29,226	Parabacteroides merdae	100% coverage 99% identity	AAXE02000107.1	2295	Parabacteroides merdae	100% coverage 99% identity	WP 005634199.1	Multi-domain	Metal Dependent Phosphorylase 2.41 e-6	Metal Dependent Phosphorylase 3.35 e-35	Metal Dependent Phosphorylase 3.25 e-8	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 6.6 e-167
ppGpp Synthetase Gene	3b03	Library 1	31,096	Bacteroides xylanisolvens	71% coverage 95% identity	NC 021017.1	2241	Bacteroides ovatus	100% coverage 100% identity	WP 004301386	Multi-domain	Metal Dependent Phosphorylase 3.80 e-6	Metal Dependent Phosphorylase 2.72 e-54	Metal Dependent Phosphorylase 2.35 e-6	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 7.57 e-156
13	4h09	Library 1	24,150	Bacteroides uniformis	92% coverage 99% identity	AGX01000019.1	1815	Bacteroides uniformis	100% coverage 100% identity	WP 005833784.1	Multi-domain				Mdb8 Phosphoglycerol transferase 1.56 e-41		
12-2	10o08	Library 1	33,711	Bacteroides dorei	99% coverage 99% identity	CP 007619.1	981	Bacteroides dorei	100% coverage 100% identity	WP 00784762.1	Multi-domain						
ppGpp Synthetase Gene	15d03	Library 1	27,714	Bacteroides dorei	100% coverage 99% identity	CP 007619.1	2271	Bacteroides dorei	100% coverage 100% identity	WP 007836277.1	Multi-domain	Metal Dependent Phosphorylase 2.51 e-6	Metal Dependent Phosphorylase 2.95 e-55	Metal Dependent Phosphorylase 1.13 e-6	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 2.84 e-166
ppGpp Synthetase Gene	15m04	Library 1	26,634	Bacteroides thetaiotaomicron	99% coverage 99% identity	NC 004663.1	2244	Bacteroides thetaiotaomicron	100% coverage 100% identity	WP 008760913.1	Multi-domain	Metal Dependent Phosphorylase 7.72 e-7	Metal Dependent Phosphorylase 1.72 e-54	Metal Dependent Phosphorylase 3.7 e-7	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 8.68 e-156
ppGpp Synthetase Gene	24i11	Library 1	32,821	Bacteroides helcogenes	57% coverage 81% identity	NC 014933.1	2259	Bacteroides uniformis	100% coverage 99% identity	WP 005824059.1	Multi-domain	Metal Dependent Phosphorylase 3.26 e-6	Metal Dependent Phosphorylase 4.1 e-53	Metal Dependent Phosphorylase 1.9 e-6	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 0e+00	ppGpp Synthetase RelA/Spot 1.08 e-161
9	31110a	Library 1	6,063	Ruminococcus sp.	100% coverage 92% identity	NC 021014.1	2817	Dorea longicatena	99% coverage 98% identity	WP 028086444.1	Individual Domain	MIP-T3 Microtubule and TRAF3 binding protein 6.37 e-6			Uncharacterized protein		
											Individual Domain	MTA-like lytic transglycosylase 1.54 e-24	Zinc metalloproteinase 2.83 e-36		Uncharacterized protein	Membrane protein metalloendopeptidase 5.23 e-39	Cell wall hydrolase 6.28 e-29

Cbeq	Unique Cosmid	Library	Metagenomic Insert Size (bp)	Closest Reference Genome	Reference genome %Coverage %Identity	Accession Number	Effector Gene Cluster Size (bp)	Closest Reference Genome	Reference Genome %Coverage %Identity	Accession Number	Conserved Protein Domains	Conserved Domains	pFAM	SMART	COG	TIGR	PRK	
5	4f05	Library 1	27,194	Bacteroides plebius	24% coverage 97% identity	ABQC02000024.1	2874	Bacteroides plebius	100% coverage 99% identity	WP 007562580.1	Individual Domain	Carbohydrate Binding Module 1.15 e-33 Fibronectin Type III Domain 6.19 e-8				Fibronectin Type III Domain 3.4 e-5		
14	3e18	Library 1	31,518	Bacteroides plebius	66% coverage 96% identity	ABQC02000002.1	2490	Bacteroides plebius	100% coverage 98% identity	WP 022054304.1	Individual Domain		PM2 Membrane protein 1.08 e-6			Membrane protein 4.71 e-4		
15	2m03	Library 1	29,366	Bacteroides plebius	68% coverage 95% identity	ABQC02000004.1	1131	Bacteroides plebius	100% coverage 98% identity	WP 022054659.1	Multi-domain Individual Domain	MurG beta-N-acetylglucosamine transferase 1.01 e-146			MurG beta-N-acetylglucosamine transferase 5.85 e-106	MurG beta-N-acetylglucosaminyl transferase 1.21 e-101	MurG beta-N-acetylglucosaminyl transferase 4.29 e-148	
1	25h09	Library 1	30,129	Bacteroides dorei	82% coverage 99% identity	CP 007619.1	2358	Bacteroides dorei	100% coverage 99% identity	WP 007847083.1	Multi-domain Individual Domain	TonB Ligand gated channel 1.81 e-22			CirA Outer membrane receptor 1.38 e-27		TonB-hemimembrane receptor 3.15 e-9	
10	33g20	Library 2	27,092	Bacteroides fragilis	99% coverage 98% identity	NC 016776.1	1065	Bacteroides fragilis	100% coverage 100% identity	WP 032529456.1	Individual Domain				Beta lactamase superfamily 2.83 e-21	Metallo-beta lactamase 2.27 e-6	Zn dependent hydrolase 8.4 e-41	metal dependent hydrolase 7.25 e-16
2	3	18b09	Library 2	38,537	Bacteroides fragilis	99% coverage 99% identity	NC 003228.3	3372	Bacteroides fragilis	100% coverage 100% identity	WP 008770235.1	Multi-domain Individual Domain	Ligand gated channel 4.06 e-6			Cir A Outer membrane receptor 3.01 e-9	Omp Rag Sus C Family Membrane Transporter Oe+0	
							1620	Bacteroides fragilis	100% coverage 100% identity	EXY 11524.1	Individual Domain	SuD Superfamily 2.68 e-38			SuD Superfamily 1.23 e-67			
4	33g04	Library 2	35,265	Bacteroides dorei	99% coverage 96% identity	CP 008741.1	2871	Bacteroides vulgatus	100% coverage 100% identity	WP 032952522.1	Individual Domain	Carbohydrate Binding Module 3.73 e-36 Fibronectin Type III Domain 4.44 e-8			Fibronectin Type III Domain 2.92 e-5	Fibronectin Type III Domain 2.05 e-5		
12-1	35i19	Library 2	35,255	Bacteroides vulgatus	94% coverage 99% identity	NC 009614.1	981	Bacteroides vulgatus	100% coverage 100% identity	WP 005845337.1	Individual Domain				n-Acetyltransferase 2.11 e-15		hemolysin 2.43 e-3	
11	39k17	Library 2	26,355	Bacteroides fragilis	99% coverage 99% identity	NC 003228.3	744	Bacteroides fragilis	100% coverage 100% identity	WP_005797964.1	Multi-domain Individual Domain	acyl-ACP-TE acyl acp thioesterase 1.89 e-32			FatA Thioesterase Lipid Metabolism 3.06 e-19		4hbt thioesterase 4-chlorobenzoate degradation 8.37 e-14	Thioesterase 2.48 e-3
											Individual Domain	4hbt thioesterase 4-chlorobenzoate degradation 1.35 e-5					Thioesterase 6.87 e-3	
8	16f15	Library 2	32,631	Bacteroides fragilis	94% coverage 99% identity	NC 016776.1	900	Bacteroides fragilis	100% coverage 100% identity	WP 005786259.1	Multi-domain Individual Domain				acetyl glucosaminidase 2.59 e-43		Lysozyme 2.82 e-27	Autolysin 9.67 e-24
											Individual Domain	LysM 1.24 e-4			LysM 8.74 e-5		muramidase 1.04 e-25	
											Individual Domain	LysM 5.33 e-6			LysM 3.91 e-3		LysM 9.66 e-6	

1
2
3
4
5

Cbeg	Unique Cosmid	Library	Metagenomic Insert Size (bp)	Closest Reference Genome	Reference genome %Coverage %Identity	Accession Number	Effector Gene Cluster Size (bp)	Closest Reference Genome	Reference Genome %Coverage %Identity	Accession Number	Conserved Protein Domains	Conserved Domains	pFAM	SMART	COG	TIGR	PRK
12-3	8j18	Library 3	26,651	Bacteroides dorei	93% coverage 99% identity	CP_007619.1	981	Bacteroides dorei	100% coverage 100% identity	WP_007834762.1				n-acetyltransferase 3.52 e-15			
76	32o03	Library 3	24,342	Roseburia intestinalis	85% coverage 96% identity	NC_021012.1	1014	Roseburia intestinalis	100% coverage 96% identity	WP_015561211.1	Individual Domain	lytic transglycosylase 1.49 e-5 nlp superfamily	lysozyme like 2.4 e-72		Cell wall hydrolase 1.44 e-28	phage peptidase/hydrolase 3.38 e-7	secreted endopeptidase 9.96 e-34
							2532	Roseburia intestinalis	100% coverage 93% identity	WP_015521894.1	Individual Domain	Multidrug and toxic component extrusion protein 9.6 e-4	Unknown protein	ABC sugar transport 2.09 e-3			
ppGpp Synthetase Gene	22g15	Library 3	35,379	Parabacteroides distasonis	69% coverage 75% identity	NC_009615.1	2106	Parabacteroides merdae	97% coverage 99% identity	WP_005634199.1	Multi-domain			ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT 1.75 e-163	
ppGpp Synthetase Gene	21f12	Library 3	29,539	Bacteroides dorei	99% coverage 99% identity	CP008741.1	2271	Bacteroides dorei	100% coverage 100% identity	WP_007836277.1	Multi-domain			ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT 2.46 e-166	
ppGpp Synthetase Gene	37b14	Library 3	25,525	Bacteroides fragilis	95% coverage 99% identity	NC_006347.1	2244	Bacteroides fragilis	100% coverage 100% identity	WP_005784842.1	Multi-domain			ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT	
ppGpp Synthetase Gene	40k7	Library 3	31,623	Bacteroides xylanosolvens	83% coverage 98% identity		2241	Bacteroides xylanosolvens	100% coverage 99% identity		Multi-domain			ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT 0e+00	ppGpp Synthetase RelA/SpoT	

1
2

1
2
3
4
5
6
7

Table S2B. Gene tables for each unique cosmid. Assembled full length cosmids (with the exception of a single cosmid 31i10 which is only partially assembled) underwent gene calling with Metagenemark and individual genes aligned to the NCBI nr database using ngKlast. The effector gene identified through transposon mutagenesis is identified in red.

1i11



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	101	WP_005634251	MULTISPECIES: glycosyl hydrolase [Parabacteroides].	1.00E-38	100%
gene 2	490	WP_005634249	MULTISPECIES: oxidoreductase [Parabacteroides].	5.00E-297	100%
gene 3	107	WP_005634247	MULTISPECIES: membrane protein [Parabacteroides].	1.00E-41	100%
gene 4	516	WP_005634245	cytochrome C oxidase subunit II [Parabacteroides merdae].	1.00E-289	99.80%
gene 5	384	WP_005634243	cytochrome D ubiquinol oxidase, subunit II [Parabacteroidesmerdae].	4.00E-166	99.70%
gene 6	270	WP_022322920	ribonuclease Z [Parabacteroides merdae CAG:48].	3.00E-155	100%
gene 7	109	WP_005634237	hypothetical protein [Parabacteroides merdae].	2.00E-53	100%
gene 8	115	WP_005634235	hypothetical protein [Parabacteroides merdae].	3.00E-29	100%
gene 9	166	WP_005634233	MULTISPECIES: glucosaminidase [Parabacteroides].	9.00E-94	100%
gene 10	74	WP_005634231	MULTISPECIES: hypothetical protein [Parabacteroides].	2.00E-37	100%
gene 11	148	WP_022322918	putative DNA-binding protein [Parabacteroides merdae CAG:48].	4.00E-68	100%
gene 12	755	WP_005644530	hypothetical protein [Parabacteroides merdae].	0	99.60%
gene 13	178	WP_005644526	hypothetical protein [Parabacteroides merdae].	8.00E-55	97.40%
gene 14	1064	WP_005634215	hypothetical protein [Parabacteroides merdae].	0	98%
gene 15	102	WP_005634211	MULTISPECIES: hypothetical protein [Parabacteroides].	5.00E-51	100%
gene 16	254	WP_005634208	chromosome partitioning protein ParA [Parabacteroides merdae].	1.00E-135	100%
gene 17	295	WP_005634207	chromosome partitioning protein ParB [Parabacteroides merdae].	2.00E-151	100%
gene 18	252	WP_005634205	MULTISPECIES: hypothetical protein [Parabacteroides].	4.00E-135	100%
gene 19	516	WP_022321377	lysM domain protein [Parabacteroides merdae CAG:48].	9.00E-301	99.80%
gene 20	701	WP_005634199	GTP pyrophosphokinase [Parabacteroides merdae].	0	99.90%
gene 21	64	WP_005649494	GTP pyrophosphokinase [Parabacteroides merdae].	9.00E-27	100%
gene 22	87	WP_005634197	MULTISPECIES: hypothetical protein [Parabacteroides].	2.00E-35	100%
gene 23	327	WP_005634192	MULTISPECIES: peptidase M23 [Parabacteroides].	2.00E-178	100%
gene 24	875	WP_005634190	alanyl-tRNA synthetase [Parabacteroides merdae].	0	99.90%
gene 25	381	WP_022321829	3-dehydroquinase synthase [Parabacteroides merdae CAG:48].	9.00E-201	100%
gene 26	91	WP_005649490	MULTISPECIES: integration host factor [Parabacteroides].	7.00E-45	100%
gene 27	24	WP_008146034	hypothetical protein [Parabacteroides johnsonii].	5.00E-05	100%

8
9

2m03



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	84	WP_016660809	hypothetical protein [Bacteroides stercoris].	2.00E-28	100%
gene 2	481	WP_005642186	MULTISPECIES: transposase [Bacteroidales].	6.00E-276	100%
gene 3	660	WP_009123684	energy transducer TonB [Bacteroides fluxus].	0	99.80%
gene 4	372	WP_009123685	MULTISPECIES: hypothetical protein [Bacteroides].	7.00E-198	100%
gene 5	407	WP_009123686	MULTISPECIES: integrase [Bacteroides].	8.00E-236	100%
gene 6	311	WP_022054647	putative uncharacterized protein [Bacteroides plebeius CAG:211].	2.00E-151	96.80%
gene 7	161	WP_007559889	MULTISPECIES: ferritin [Bacteroides].	8.00E-87	99.40%
gene 8	226	WP_022054648	putative uncharacterized protein [Bacteroides plebeius CAG:211].	1.00E-130	98.70%
gene 9	189	WP_007559893	cobalamin adenosyltransferase [Bacteroides plebeius].	4.00E-94	99.50%
gene 10	73	WP_002558131	MULTISPECIES: hypothetical protein [Bacteroidales].	2.00E-36	100%
gene 11	615	WP_022307353	putative uncharacterized protein [Alistipes sp. CAG:268].	0	90.90%
gene 12	158	WP_022054650	putative uncharacterized protein [Bacteroides plebeius CAG:211].	1.00E-77	100%
gene 13	892	WP_022054651	putative uncharacterized protein [Bacteroides plebeius CAG:211].	0	99%
gene 14	321	WP_007559905	MULTISPECIES: DNA-cytosine methyltransferase [Bacteroides].	2.00E-172	100%
gene 15	319	WP_007559905	MULTISPECIES: DNA-cytosine methyltransferase [Bacteroides].	4.00E-181	99.70%
gene 16	84	WP_007559907	MULTISPECIES: 30S ribosomal protein S20 [Bacteroides].	5.00E-30	100%
gene 17	171	WP_022054652	putative uncharacterized protein [Bacteroides plebeius CAG:211].	6.00E-79	96.50%
gene 18	210	WP_007559912	DNA repair protein RecO [Bacteroides plebeius].	2.00E-117	100%
gene 19	151	WP_007559914	aspartyl-tRNA amidotransferase subunit B [Bacteroides plebeius].	1.00E-75	100%
gene 20	436	WP_022354669	cell division protein FtsZ [Bacteroides sp. CAG:875].	6.00E-189	99.10%
gene 21	525	WP_022054656	putative uncharacterized protein [Bacteroides plebeius CAG:211].	2.00E-231	100%
gene 22	209	WP_022054657	putative uncharacterized protein [Bacteroides plebeius CAG:211].	3.00E-106	99%
gene 23	458	WP_022054658	uDP-N-acetylmuramate--L-alanine ligase [Bacteroides plebeiusCAG:211].	2.00E-264	99.60%
gene 24	376	WP_007559923	UDP-diphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase [Bacteroides plebeius].	2.00E-197	98.10%
gene 25	416	WP_022054660	putative uncharacterized protein [Bacteroides plebeius CAG:211].	5.00E-219	97.60%
gene 26	329	WP_022054661	uDP-N-acetylmuramoylalanine--D-glutamate ligase [Bacteroidesplebeius CAG:211].	2.00E-191	100%

1
2
3

3b03



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	331	WP_015532350	Domain of Unknown Function (DUF349) [Bacteroides xylanisolvens].	1.00E-159	100%
gene 2	119	WP_004301405	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-50	100%
gene 3	718	WP_004301403	cell division protein FtsH [Bacteroides ovatus].	0	100%
gene 4	282	WP_004301401	phosphatidate cytidyltransferase [Bacteroides ovatus].	3.00E-149	99.60%
gene 5	260	WP_032845960	hypothetical protein [Bacteroides ovatus].	2.00E-140	100%
gene 6	378	WP_004301397	lipid-A-disaccharide synthase [Bacteroides ovatus].	9.00E-217	100%
gene 7	277	WP_004301396	stationary phase survival protein SurE [Bacteroides ovatus].	4.00E-164	99.60%
gene 8	254	WP_004301394	MULTISPECIES: chromosome partitioning protein ParA [Bacteroides].	3.00E-135	100%
gene 9	296	WP_004320990	chromosome partitioning protein ParB [Bacteroides ovatus].	4.00E-160	100%
gene 10	295	WP_004301390	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-165	100%
gene 11	431	WP_004301388	MULTISPECIES: lytic transglycosylase [Bacteroides].	7.00E-251	100%
gene 12	746	WP_004301386	GTP pyrophosphokinase [Bacteroides ovatus].	0	100%
gene 13	114	WP_004301384	MULTISPECIES: transcriptional regulator [Bacteroides].	6.00E-58	100%
gene 14	286	WP_004301381	peptidase M23 [Bacteroides ovatus].	2.00E-163	100%
gene 15	295	WP_004320991	hypothetical protein [Bacteroides ovatus].	4.00E-160	100%
gene 16	283	WP_004320992	hypothetical protein [Bacteroides ovatus].	2.00E-160	100%
gene 17	91	WP_004320993	hypothetical protein [Bacteroides ovatus].	5.00E-37	100%
gene 18	74	WP_004320994	hypothetical protein [Bacteroides ovatus].	4.00E-38	100%
gene 19	57	WP_015531800	hypothetical protein [Bacteroides xylanisolvens].	5.00E-24	96.50%
gene 20	59	WP_004323037	hypothetical protein [Bacteroides ovatus].	3.00E-25	94.90%
gene 21	133	WP_015531798	hypothetical protein [Bacteroides xylanisolvens].	2.00E-73	100%
gene 22	150	WP_004296549	MULTISPECIES: hypothetical protein [Bacteroides].	4.00E-78	96%
gene 23	52	WP_004320996	hypothetical protein [Bacteroides ovatus].	2.00E-21	100%
gene 24	71	WP_004320997	hypothetical protein [Bacteroides ovatus].	5.00E-32	100%
gene 25	62	WP_004320998	hypothetical protein [Bacteroides ovatus].	2.00E-29	100%
gene 26	262	WP_004321000	hypothetical protein [Bacteroides ovatus].	6.00E-141	100%
gene 27	356	WP_004321002	penicillin-binding protein 2B [Bacteroides ovatus].	8.00E-202	100%
gene 28	143	WP_004321003	hypothetical protein [Bacteroides ovatus].	6.00E-10	100%
gene 29	113	WP_004321004	hypothetical protein [Bacteroides ovatus].	4.00E-45	100%
gene 30	153	WP_004321005	hypothetical protein [Bacteroides ovatus].	6.00E-64	100%
gene 31	148	WP_022021199	uncharacterized protein [Bacteroides sp. CAG:661].	2.00E-53	64.20%
gene 32	195	WP_004324654	tyrosine recombinase [Bacteroides ovatus].	1.00E-105	99%
gene 33	90	WP_004321010	hypothetical protein [Bacteroides ovatus].	1.00E-44	100%
gene 34	338	WP_005681910	hypothetical protein [Bacteroides caccae].	5.00E-184	99.40%
gene 35	201	WP_004323732	hypothetical protein [Bacteroides ovatus].	3.00E-116	100%
gene 36	287	WP_004323733	hypothetical protein [Bacteroides ovatus].	6.00E-169	100%
gene 37	177	WP_004323734	hypothetical protein [Bacteroides ovatus].	8.00E-103	99.40%

1
2
3

3e18

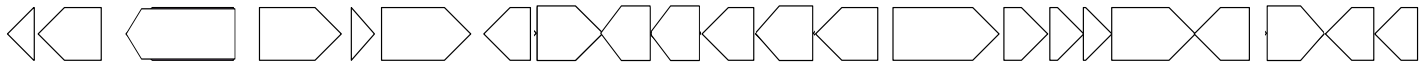


Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	159	WP_003228230	levansucrase [Bacillus subtilis].	2.00E-85	100%
gene 2	432	WP_007900085	MULTISPECIES: hypothetical protein [Bacteroidales].	2.00E-236	98.10%
gene 3	274	WP_022054294	tonB-dependent receptor plug [Bacteroides plebeius CAG:211].	8.00E-133	97%
gene 4	680	WP_022054295	methionine--tRNA ligase [Bacteroides plebeius CAG:211].	0	99.70%
gene 5	481	WP_007559421	lipopolysaccharide biosynthesis protein [Bacteroides plebeius].	4.00E-272	99.20%
gene 6	312	WP_007559422	glycosyltransferase family 2 [Bacteroides plebeius].	3.00E-180	98.40%
gene 7	295	WP_007559424	LICD [Bacteroides plebeius].	1.00E-170	96.90%
gene 8	353	WP_007559425	hemolysin activation protein [Bacteroides plebeius].	1.00E-213	99.20%
gene 9	358	WP_007559426	glycosyl transferase [Bacteroides plebeius].	5.00E-196	97.20%
gene 10	420	WP_007559427	glycosyl transferase [Bacteroides plebeius].	3.00E-235	98.10%
gene 11	254	WP_007559428	glycosyl transferase family 2 [Bacteroides plebeius].	2.00E-146	99.20%
gene 12	874	WP_007559429	membrane protein [Bacteroides plebeius].	0	99%
gene 13	223	WP_022054305	putative uncharacterized protein [Bacteroides plebeius CAG:211].	1.00E-126	100%
gene 14	161	WP_007559432	MULTISPECIES: S-ribosylhomocysteinase [Bacteroides].	2.00E-89	100%
gene 15	808	WP_022054306	putative uncharacterized protein [Bacteroides plebeius CAG:211].	0	98.90%
gene 16	222	WP_022054307	uncharacterized protein [Bacteroides plebeius CAG:211].	5.00E-122	98.60%
gene 17	620	WP_022054308	uncharacterized protein [Bacteroides plebeius CAG:211].	0	99.80%
gene 18	1030	WP_022054309	putative uncharacterized protein [Bacteroides plebeius CAG:211].	0	99.80%
gene 19	1361	WP_022054310	putative uncharacterized protein [Bacteroides plebeius CAG:211].	0	96.60%
gene 20	261	WP_032726907	levansucrase [Bacillus subtilis].	1.00E-132	99.20%

1

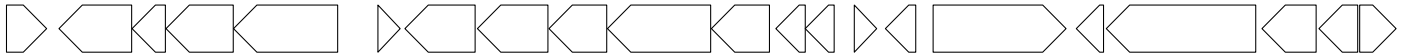
2

4h09



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	153	WP_022401844	uncharacterized protein [Bacteroides uniformis CAG:3].	7.00E-65	98.70%
gene 2	359	WP_009037388	oxidoreductase [Bacteroides sp. D20].	4.00E-212	98.90%
gene 3	612	WP_005833760	MULTISPECIES: antibiotic ABC transporter ATP-binding protein[Bacteroides].	0	99.70%
gene 4	466	WP_005833763	hypothetical protein [Bacteroides uniformis].	5.00E-226	100%
gene 5	131	WP_005833765	hypothetical protein [Bacteroides uniformis].	1.00E-66	100%
gene 6	508	WP_005833767	hypothetical protein [Bacteroides uniformis].	1.00E-294	100%
gene 7	265	WP_005833769	hypothetical protein [Bacteroides uniformis].	2.00E-158	100%
gene 8	367	WP_005833771	glycosyltransferase family 1 [Bacteroides uniformis].	1.00E-204	100%
gene 9	272	WP_005833773	glycosyl transferase [Bacteroides uniformis].	2.00E-152	99.60%
gene 10	278	WP_005833775	glycosyl transferase [Bacteroides uniformis].	2.00E-161	100%
gene 11	294	WP_005833777	hypothetical protein [Bacteroides uniformis].	1.00E-174	100%
gene 12	321	WP_005833780	hypothetical protein [Bacteroides uniformis].	4.00E-187	100%
gene 13	354	WP_005833782	CDP-glycerol glycerophosphotransferase [Bacteroides uniformis].	5.00E-211	100%
gene 14	604	WP_005833784	hypothetical protein [Bacteroides uniformis].	0	100%
gene 15	249	WP_005825949	MULTISPECIES: hypothetical protein [Bacteroides].	5.00E-144	100%
gene 16	193	WP_005833786	phosphoheptose isomerase [Bacteroides uniformis].	1.00E-103	100%
gene 17	162	WP_005833788	D,D-heptose 1,7-bisphosphate phosphatase [Bacteroides uniformis].	9.00E-91	100%
gene 18	476	WP_005833790	RfaE, domain I [Bacteroides uniformis].	2.00E-244	100%
gene 19	314	WP_005833792	hypothetical protein [Bacteroides uniformis].	5.00E-183	100%
gene 20	305	WP_005825959	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-179	100%
gene 21	275	WP_005833794	MULTISPECIES: LicD [Bacteroides].	1.00E-162	100%
gene 22	243	WP_005825963	MULTISPECIES: hypothetical protein [Bacteroides].	8.00E-125	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	257	WP_032726907	levansucrase [Bacillus subtilis].	9.00E-133	99.20%
gene 2	466	WP_022052955	putative uncharacterized protein [Bacteroides plebeius CAG:211].	1.00E-262	99.60%
gene 3	213	WP_007563484	ATP-dependent DNA helicase [Bacteroides plebeius].	9.00E-115	98.10%
gene 4	431	WP_022052957	uPF0597 protein BACPLE_03244 [Bacteroides plebeius CAG:211].	4.00E-239	99.30%
gene 5	668	WP_022052958	putative uncharacterized protein [Bacteroides plebeius CAG:211].	0	99.90%
gene 6	142	WP_007758564	mobilization protein [Bacteroides finegoldii].	2.00E-52	96.30%
gene 7	450	WP_005832685	multidrug transporter MatE [Bacteroides uniformis].	4.00E-87	50%
gene 8	452	WP_026983296	RND transporter [Flavobacterium sp. URHB0058].	1.00E-63	31.60%
gene 9	370	WP_012505954	multidrug transporter AcrB [Prosthecochloris aestuarii].	4.00E-90	53.10%
gene 10	662	WP_012466369	multidrug transporter AcrB [Chlorobium limicola].	8.00E-152	56%
gene 11	372	WP_026812990	secretion protein HlyD [Arenibacter certesii].	6.00E-51	32%
gene 12	188	WP_008999156	DNA-binding protein [Bacteroides sp. D2].	3.00E-24	37.50%
gene 13	185	WP_008147829	hypothetical protein [Parabacteroides johnsonii].	3.00E-74	75%
gene 14	143	WP_008147828	MULTISPECIES: nitroreductase [Bacteroidales].	1.00E-71	94.40%
gene 15	194	WP_022054717	putative uncharacterized protein [Bacteroides plebeius CAG:211].	9.00E-84	98%
gene 16	853	WP_007562587	beta-N-acetylglucosaminidase [Bacteroides plebeius].	0	98.70%
gene 17	174	WP_007562582	diguanylate cyclase [Bacteroides plebeius].	5.00E-95	99.40%
gene 18	959	WP_007562580	xanthan lyase [Bacteroides plebeius].	0	98.70%
gene 19	347	WP_007562578	ketol-acid reductoisomerase [Bacteroides plebeius].	3.00E-198	100%
gene 20	248	WP_007562576	acyl-ACP thioesterase [Bacteroides plebeius].	3.00E-139	98.40%
gene 21	233	WP_003228230	levansucrase [Bacillus subtilis].	5.00E-118	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	233	WP_003228230	levansucrase [Bacillus subtilis].	1.00E-118	98.60%
gene 2	444	WP_007834746	UDP-N-acetylmuramoylalanine--D-glutamate ligase [Bacteroidesdorei].	1.00E-259	100%
gene 3	422	WP_007834748	MULTISPECIES: phospho-N-acetylmuramoyl-pentapeptide-transferase[Bacteroides].	5.00E-187	100%
gene 4	484	WP_007834750	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase[Bacteroides dorei].	4.00E-267	100%
gene 5	726	WP_007834752	MULTISPECIES: penicillin-binding protein [Bacteroides].	0	100%
gene 6	124	WP_007847173	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-48	100%
gene 7	304	WP_008674321	16S rRNA methyltransferase [Bacteroides sp. 9_1_42FAA].	3.00E-168	100%
gene 8	154	WP_007834758	MULTISPECIES: cell division protein MraZ [Bacteroides].	6.00E-84	100%
gene 9	279	WP_007834760	glycerol acyltransferase [Bacteroides dorei CAG:222].	1.00E-161	100%
gene 10	326	WP_007834762	MULTISPECIES: hemolysin [Bacteroides].	2.00E-187	100%
gene 11	443	WP_007834767	MULTISPECIES: dehydrogenase [Bacteroides].	7.00E-254	100%
gene 12	144	WP_005845350	MULTISPECIES: deoxyuridine 5'-triphosphate nucleotidohydrolase[Bacteroides].	2.00E-76	100%
gene 13	585	WP_007842003	hypothetical protein [Bacteroides dorei].	0	100%
gene 14	277	WP_007842008	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-153	100%
gene 15	233	WP_008655581	peptidase M23 [Bacteroides sp. 3_1_33FAA].	4.00E-104	100%
gene 16	216	WP_007834770	MULTISPECIES: peptidase M23 [Bacteroides].	3.00E-90	100%
gene 17	452	WP_007834771	hypothetical protein [Bacteroides dorei].	1.00E-247	100%
gene 18	122	WP_007834772	MULTISPECIES: RNA-binding protein [Bacteroides].	3.00E-39	100%
gene 19	316	WP_007842013	MULTISPECIES: glutaminase [Bacteroides].	3.00E-164	100%
gene 20	614	WP_007842021	MULTISPECIES: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase[Bacteroides].	0	99.80%
gene 21	169	WP_007834776	MULTISPECIES: phosphoribosylaminoimidazole carboxylase[Bacteroides].	1.00E-72	100%
gene 22	126	WP_007834777	MULTISPECIES: glycine cleavage system protein H [Bacteroides].	2.00E-66	100%
gene 23	218	WP_007855043	membrane protein [Bacteroides dorei].	4.00E-116	100%
gene 24	76	WP_007847161	RNA polymerase sigma54 factor [Bacteroides dorei].	1.00E-35	100%
gene 25	524	WP_007834779	MULTISPECIES: RNA polymerase sigma54 factor [Bacteroides].	1.00E-142	99.20%

1

2



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	320	WP_008655581	peptidase M23 [Bacteroides sp. 3_1_33FAA].	8.00E-139	100%
gene 2	277	WP_007847169	hypothetical protein [Bacteroides dorei].	5.00E-154	100%
gene 3	585	WP_007847170	MULTISPECIES: hypothetical protein [Bacteroides].	0	100%
gene 4	144	WP_005845350	MULTISPECIES: deoxyuridine 5'-triphosphate nucleotidohydrolase[Bacteroides].	2.00E-76	100%
gene 5	443	WP_007834767	MULTISPECIES: dehydrogenase [Bacteroides].	7.00E-254	100%
gene 6	326	WP_007834762	MULTISPECIES: hemolysin [Bacteroides].	2.00E-187	100%
gene 7	279	WP_007834760	glycerol acyltransferase [Bacteroides dorei CAG:222].	1.00E-161	100%
gene 8	154	WP_007834758	MULTISPECIES: cell division protein MraZ [Bacteroides].	6.00E-84	100%
gene 9	304	WP_008674321	16S rRNA methyltransferase [Bacteroides sp. 9_1_42FAA].	3.00E-168	100%
gene 10	124	WP_007847173	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-48	100%
gene 11	726	WP_007834752	MULTISPECIES: penicillin-binding protein [Bacteroides].	0	100%
gene 12	484	WP_007834750	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase[Bacteroides dorei].	4.00E-267	100%
gene 13	422	WP_007834748	MULTISPECIES: phospho-N-acetylmuramoyl-pentapeptide-transferase[Bacteroides].	5.00E-187	100%
gene 14	444	WP_007847176	MULTISPECIES: UDP-N-acetylmuramoylalanine--D-glutamate ligase[Bacteroides].	1.00E-259	100%
gene 15	465	WP_007834744	MULTISPECIES: rod shape-determining protein RodA [Bacteroides].	6.00E-222	100%
gene 16	376	WP_007834743	MULTISPECIES: UDP-diphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase [Bacteroides].	1.00E-182	100%
gene 17	458	WP_007841992	MULTISPECIES: UDP-N-acetylmuramate--alanine ligase [Bacteroides].	9.00E-257	100%
gene 18	245	WP_007834741	MULTISPECIES: cell division protein FtsQ [Bacteroides].	1.00E-129	100%
gene 19	503	WP_008674315	cell division protein FtsA [Bacteroides sp. 9_1_42FAA].	8.00E-255	100%
gene 20	434	WP_016275184	MULTISPECIES: cell division protein FtsZ [Bacteroidales].	2.00E-193	100%
gene 21	149	WP_007847182	aspartyl-tRNA amidotransferase subunit B [Bacteroides dorei].	7.00E-56	99.30%
gene 22	51	n/a	n/a	n/a	n/a
gene 23	241	WP_007834736	MULTISPECIES: DNA repair protein RecO [Bacteroides].	7.00E-137	100%
gene 24	84	WP_005838940	MULTISPECIES: 30S ribosomal protein S20 [Bacteroides].	2.00E-30	100%
gene 25	654	WP_007834731	MULTISPECIES: DNA-cytosine methyltransferase [Bacteroides].	0	99.80%
gene 26	47	WP_007855026	hypothetical protein [Bacteroides dorei CAG:222].	2.00E-19	97.90%
gene 27	892	WP_007834727	4-alpha-glucanotransferase [Bacteroides dorei].	0	100%
gene 28	451	WP_007847194	major facilitator transporter [Bacteroides dorei].	3.00E-248	99.80%
gene 29	96	WP_008655562	transcriptional regulator [Bacteroides sp. 3_1_33FAA].	3.00E-46	100%

1

2

15d03



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	195	WP_007836312	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-110	98%
gene 2	106	WP_008674893	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-46	100%
gene 3	106	WP_007836307	MULTISPECIES: XRE family transcriptional regulator [Bacteroides].	2.00E-52	100%
gene 4	78	WP_007836305	MULTISPECIES: hypothetical protein [Bacteroides].	6.00E-38	98.70%
gene 5	286	WP_007836303	MULTISPECIES: transcriptional regulator [Bacteroides].	2.00E-163	100%
gene 6	353	WP_007836297	MULTISPECIES: 3-dehydroquinase synthase [Bacteroides].	7.00E-205	100%
gene 7	134	WP_008656260	hypothetical protein [Bacteroides sp. 3_1_33FAA].	3.00E-61	100%
gene 8	449	WP_007836292	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-238	100%
gene 9	530	WP_007836292	MULTISPECIES: hypothetical protein [Bacteroides].	6.00E-300	100%
gene 10	478	WP_007836290	MULTISPECIES: cardiolipin synthetase [Bacteroides].	3.00E-270	100%
gene 11	186	WP_007836288	methyltransferase [Bacteroides dorei].	4.00E-103	100%
gene 12	272	WP_007836286	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-148	100%
gene 13	228	WP_007836284	MULTISPECIES: hypothetical protein [Bacteroides].	9.00E-120	100%
gene 14	472	WP_007836283	ATP-dependent endonuclease [Bacteroides dorei].	1.00E-263	99.80%
gene 15	32	n/a	n/a	n/a	n/a
gene 16	448	WP_007836280	MULTISPECIES: alanyl-tRNA synthetase [Bacteroides].	5.00E-252	100%
gene 17	443	WP_007836280	MULTISPECIES: alanyl-tRNA synthetase [Bacteroides].	1.00E-237	99.50%
gene 18	322	WP_007840678	MULTISPECIES: peptidase M23 [Bacteroides].	4.00E-184	100%
gene 19	115	WP_007836278	MULTISPECIES: transcriptional regulator [Bacteroides].	3.00E-58	100%
gene 20	756	WP_007836277	MULTISPECIES: GTP pyrophosphokinase [Bacteroides].	0	100%
gene 21	420	WP_007836276	MULTISPECIES: lytic transglycosylase [Bacteroides].	5.00E-245	100%
gene 22	235	WP_032934275	MULTISPECIES: hypothetical protein [Bacteroides].	5.00E-122	100%
gene 23	298	WP_005845804	MULTISPECIES: chromosome partitioning protein ParB [Bacteroides].	8.00E-150	100%
gene 24	278	WP_032936379	chromosome partitioning protein ParA [Bacteroides dorei].	2.00E-136	100%
gene 25	257	WP_007836273	MULTISPECIES: stationary phase survival protein SurE [Bacteroides].	3.00E-150	100%
gene 26	379	WP_007836272	MULTISPECIES: lipid-A-disaccharide synthase [Bacteroides].	3.00E-218	100%
gene 27	244	WP_007845865	MULTISPECIES: hypothetical protein [Bacteroides].	7.00E-138	100%
gene 28	72	WP_007836270	phosphatidate cytidyltransferase [Bacteroides dorei].	2.00E-32	98.60%

1
2
3

15m04



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	339	WP_008764224	MULTISPECIES: hypothetical protein [Bacteroides].	3.00E-197	100%
gene 2	384	WP_008764225	MULTISPECIES: phospholipase [Bacteroides].	3.00E-207	100%
gene 3	476	WP_008760902	MULTISPECIES: endo-beta-N-acetylglucosaminidase [Bacteroides].	6.00E-262	100%
gene 4	444	WP_008764227	peptidoglycan-binding protein [Bacteroides sp. 1_1_6].	3.00E-253	100%
gene 5	144	WP_008760904	MULTISPECIES: hypothetical protein [Bacteroides].	4.00E-80	100%
gene 6	741	WP_008760905	alpha-1 2-mannosidase [Bacteroides sp. 1_1_14].	0	99.90%
gene 7	746	WP_016269483	hypothetical protein [Bacteroides thetaiotaomicron].	0	100%
gene 8	336	WP_032814424	anti-sigma factor [Bacteroides thetaiotaomicron].	4.00E-176	99.70%
gene 9	184	WP_008760908	MULTISPECIES: RNA polymerase sigma70 factor [Bacteroides].	3.00E-92	100%
gene 10	743	WP_008764233	alpha-1 2-mannosidase [Bacteroides sp. 1_1_6].	0	99.90%
gene 11	872	WP_011109057	alanyl-tRNA synthetase [Bacteroides thetaiotaomicron].	0	99.90%
gene 12	322	WP_008760911	MULTISPECIES: peptidase M23 [Bacteroides].	7.00E-185	100%
gene 13	113	WP_008760912	MULTISPECIES: transcriptional regulator [Bacteroides].	5.00E-57	100%
gene 14	747	WP_008760913	MULTISPECIES: GTP pyrophosphokinase [Bacteroides].	0	100%
gene 15	431	WP_008764235	MULTISPECIES: lytic transglycosylase [Bacteroides].	2.00E-250	100%
gene 16	295	WP_008760915	MULTISPECIES: hypothetical protein [Bacteroides].	3.00E-148	100%
gene 17	296	WP_008760916	MULTISPECIES: chromosome partitioning protein ParB [Bacteroides].	1.00E-159	100%
gene 18	315	WP_008760917	chromosome partitioning protein ParA [Bacteroides thetaiotaomicron].	7.00E-171	100%
gene 19	259	WP_008760918	MULTISPECIES: stationary phase survival protein SurE [Bacteroides].	1.00E-152	100%
gene 20	198	WP_008760919	lipid-A-disaccharide synthase [Bacteroides sp. 1_1_14].	3.00E-112	100%

1
2
3



Gene	Length	Accession	Hit Definition	E Value	Identity
gene 1	28	WP_005800760	cytidine deaminase [Bacteroides fragilis].	8.00E-07	100%
gene 2	299	WP_005786259	MULTISPECIES: N-acetylmuramoyl-L-alanine amidase [Bacteroides].	7.00E-171	100%
gene 3	415	WP_032528499	ABC transporter permease [Bacteroides fragilis].	2.00E-216	100%
gene 4	421	WP_005786264	MULTISPECIES: ABC transporter permease [Bacteroides].	5.00E-215	100%
gene 5	422	WP_005800756	ABC transporter permease [Bacteroides fragilis].	4.00E-228	99.80%
gene 6	435	WP_005816631	MULTISPECIES: ABC transporter permease [Bacteroides].	2.00E-238	100%
gene 7	170	WP_032567596	ABC transporter permease [Bacteroides fragilis].	3.00E-93	100%
gene 8	266	WP_005800752	ABC transporter permease [Bacteroides fragilis].	2.00E-125	100%
gene 9	435	WP_005816632	MULTISPECIES: ABC transporter permease [Bacteroides].	1.00E-230	100%
gene 10	423	WP_032528500	ABC transporter permease [Bacteroides fragilis].	2.00E-247	100%
gene 11	299	WP_005786279	MULTISPECIES: transcriptional regulator [Bacteroides].	1.00E-169	100%
gene 12	221	WP_032528501	macrolide ABC transporter ATP-binding protein [Bacteroides fragilis].	6.00E-120	100%
gene 13	424	WP_005816636	MULTISPECIES: ABC transporter permease [Bacteroides].	9.00E-198	100%
gene 14	416	WP_032528502	ABC transporter permease [Bacteroides fragilis].	6.00E-231	100%
gene 15	221	WP_005795276	MULTISPECIES: macrolide ABC transporter ATP-binding protein [Bacteroides].	4.00E-119	100%
gene 16	214	WP_005795275	MULTISPECIES: hypothetical protein [Bacteroides].	3.00E-117	100%
gene 17	432	WP_014298553	ABC transporter permease [Bacteroides fragilis].	1.00E-226	100%
gene 18	414	WP_005816642	MULTISPECIES: ABC transporter permease [Bacteroides].	7.00E-211	100%
gene 19	821	WP_005795272	MULTISPECIES: sensor histidine kinase [Bacteroides].	0	100%
gene 20	490	WP_005795271	MULTISPECIES: membrane protein [Bacteroides].	1.00E-251	100%
gene 21	449	WP_032529652	chemotaxis protein CheY [Bacteroides fragilis].	7.00E-249	100%
gene 22	315	WP_005786305	MULTISPECIES: sensor histidine kinase [Bacteroides].	2.00E-156	100%

1
2
3

18b09



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	310	WP_005798453	MULTISPECIES: anti-sigma factor [Bacteroides].	4.00E-175	100%
gene 2	192	WP_005791112	MULTISPECIES: RNA polymerase sigma70 factor [Bacteroides].	9.00E-105	100%
gene 3	194	WP_005791113	MULTISPECIES: GTP cyclohydrolase [Bacteroides].	1.00E-105	100%
gene 4	149	WP_032580447	cell division protein [Bacteroides fragilis].	9.00E-72	100%
gene 5	267	WP_005791115	MULTISPECIES: triosephosphate isomerase [Bacteroides].	3.00E-128	100%
gene 6	443	WP_032529314	DoxX family protein [Bacteroides fragilis].	1.00E-246	99.80%
gene 7	180	WP_005791117	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-95	100%
gene 8	293	WP_008770039	MULTISPECIES: peptidase M23 [Bacteroides].	3.00E-166	100%
gene 9	154	WP_008770040	MULTISPECIES: nucleoside diphosphate kinase [Bacteroides].	2.00E-82	100%
gene 10	698	WP_005817637	MULTISPECIES: ATP-dependent DNA helicase RecG [Bacteroides].	0	100%
gene 11	204	WP_005802112	MULTISPECIES: 2-C-methyl-D-erythritol 4-phosphatecytidyltransferase [Bacteroides].	1.00E-111	100%
gene 12	183	WP_005791124	MULTISPECIES: thiazole biosynthesis protein Thij [Bacteroides].	1.00E-93	100%
gene 13	280	WP_008657688	MULTISPECIES: cell envelope biogenesis protein TonB [Bacteroides].	2.00E-60	100%
gene 14	140	WP_010993550	biopolymer transporter ExbD [Bacteroides fragilis].	3.00E-72	100%
gene 15	240	WP_005791129	MULTISPECIES: biopolymer transporter ExbB [Bacteroides].	5.00E-121	100%
gene 16	237	WP_005791132	MULTISPECIES: pyridoxamine 5'-phosphate oxidase [Bacteroides].	6.00E-123	100%
gene 17	290	WP_005791134	MULTISPECIES: inorganic polyphosphate/ATP-NAD kinase [Bacteroides].	6.00E-163	100%
gene 18	46	WP_005799295	hypothetical protein [Bacteroides fragilis].	2.00E-09	94.10%
gene 19	160	WP_032497521	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-68	100%
gene 20	165	WP_032599763	RNA polymerase sigma70 factor [Bacteroides fragilis].	4.00E-74	100%
gene 21	322	WP_005791673	MULTISPECIES: anti-sigma factor [Bacteroides].	6.00E-173	100%
gene 22	1146	WP_008770235	MULTISPECIES: TonB-linked outer membrane protein [Bacteroides].	0	100%
gene 23	539	WP_005791678	MULTISPECIES: membrane protein [Bacteroides].	2.00E-303	100%
gene 24	124	WP_032529351	hypothetical protein [Bacteroides fragilis].	4.00E-66	100%
gene 25	662	WP_032529350	endonuclease [Bacteroides fragilis].	0	100%
gene 26	730	WP_022348127	MULTISPECIES: uncharacterized protein [Bacteroides].	0	100%
gene 27	124	WP_005791690	MULTISPECIES: transcriptional regulator [Bacteroides].	3.00E-63	100%
gene 28	1095	WP_005821947	peptidase S41 [Bacteroides fragilis].	0	100%
gene 29	277	WP_005791692	MULTISPECIES: malate dehydrogenase [Bacteroides].	3.00E-139	100%

1
2
3

21f12



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	270	WP_007836303	MULTISPECIES: transcriptional regulator [Bacteroides].	2.00E-152	99.60%
gene 2	353	WP_007836297	MULTISPECIES: 3-dehydroquinase synthase [Bacteroides].	7.00E-205	100%
gene 3	134	WP_008656260	hypothetical protein [Bacteroides sp. 3_1_33FAA].	3.00E-61	100%
gene 4	965	WP_007836292	MULTISPECIES: hypothetical protein [Bacteroides].	0	100%
gene 5	478	WP_007836290	MULTISPECIES: cardiolipin synthetase [Bacteroides].	3.00E-270	100%
gene 6	186	WP_007836288	methyltransferase [Bacteroides dorei].	4.00E-103	100%
gene 7	272	WP_007836286	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-148	100%
gene 8	228	WP_007836284	MULTISPECIES: hypothetical protein [Bacteroides].	9.00E-120	100%
gene 9	184	WP_007840675	ATP-dependent endonuclease [Bacteroides dorei].	4.00E-74	100%
gene 10	281	WP_008674890	ATP-dependent endonuclease [Bacteroides sp. 9_1_42FAA].	1.00E-156	100%
gene 11	32	n/a	n/a	n/a	n/a
gene 12	872	WP_007836280	MULTISPECIES: alanyl-tRNA synthetase [Bacteroides].	0	99.90%
gene 13	322	WP_007840678	MULTISPECIES: peptidase M23 [Bacteroides].	4.00E-184	100%
gene 14	115	WP_007836278	MULTISPECIES: transcriptional regulator [Bacteroides].	3.00E-58	100%
gene 15	756	WP_007836277	MULTISPECIES: GTP pyrophosphokinase [Bacteroides].	0	100%
gene 16	420	WP_022185205	uncharacterized protein [Bacteroides dorei CAG:222].	9.00E-245	100%
gene 17	235	WP_032936421	hypothetical protein [Bacteroides dorei].	5.00E-122	100%
gene 18	298	WP_005845804	MULTISPECIES: chromosome partitioning protein ParB [Bacteroides].	8.00E-150	100%
gene 19	278	WP_007840682	MULTISPECIES: chromosome partitioning protein ParA [Bacteroides].	3.00E-136	100%
gene 20	257	WP_007836273	MULTISPECIES: stationary phase survival protein SurE [Bacteroides].	3.00E-150	100%
gene 21	97	WP_007836272	MULTISPECIES: lipid-A-disaccharide synthase [Bacteroides].	3.00E-48	100%
gene 22	245	WP_007836272	MULTISPECIES: lipid-A-disaccharide synthase [Bacteroides].	1.00E-135	100%
gene 23	244	WP_007845865	MULTISPECIES: hypothetical protein [Bacteroides].	7.00E-138	100%
gene 24	279	WP_007836270	phosphatidate cytidylyltransferase [Bacteroides dorei].	4.00E-143	99.60%
gene 25	668	WP_007836269	MULTISPECIES: cell division protein FtsH [Bacteroides].	0	100%
gene 26	119	WP_007836264	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-62	100%
gene 27	198	WP_007836261	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-106	100%

1
2
3



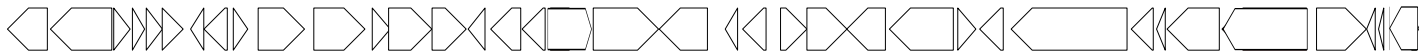
Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	296	WP_022322466	oxidoreductase 2-nitropropane dioxygenase family protein[Parabacteroides merdae CAG:48]	3.00E-152	100%
gene 2	137	WP_005634170	MULTISPECIES: membrane protein [Parabacteroides].	5.00E-57	100%
gene 3	185	WP_005634172	MULTISPECIES: RNA polymerase sigma70 factor [Parabacteroides].	1.00E-97	99.50%
gene 4	119	WP_005634173	MULTISPECIES: hypothetical protein [Parabacteroides].	1.00E-49	99.20%
gene 5	597	WP_005649481	30S ribosomal protein S1 [Parabacteroides merdae].	0	100%
gene 6	87	WP_005634177	MULTISPECIES: hypothetical protein [Parabacteroides].	1.00E-36	100%
gene 7	129	WP_005634180	chemotaxis protein CheY [Parabacteroides merdae].	2.00E-65	100%
gene 8	91	WP_005649490	MULTISPECIES: integration host factor [Parabacteroides].	2.00E-44	98.90%
gene 9	354	WP_022321829	3-dehydroquinase synthase [Parabacteroides merdae CAG:48].	8.00E-201	100%
gene 10	875	WP_005634190	alanyl-tRNA synthetase [Parabacteroides merdae].	0	99.90%
gene 11	327	WP_005644517	peptidase M23 [Parabacteroides merdae].	1.00E-178	100%
gene 12	118	WP_005634196	MULTISPECIES: transcriptional regulator [Parabacteroides].	3.00E-59	100%
gene 13	131	WP_005634197	MULTISPECIES: hypothetical protein [Parabacteroides].	1.00E-62	100%
gene 14	64	WP_005649494	GTP pyrophosphokinase [Parabacteroides merdae].	9.00E-27	100%
gene 15	701	WP_005634199	GTP pyrophosphokinase [Parabacteroides merdae].	0	99.70%
gene 16	516	WP_022321377	lysM domain protein [Parabacteroides merdae CAG:48].	3.00E-301	100%
gene 17	252	WP_005634205	MULTISPECIES: hypothetical protein [Parabacteroides].	4.00E-135	100%
gene 18	295	WP_005634207	chromosome partitioning protein ParB [Parabacteroides merdae].	2.00E-151	100%
gene 19	254	WP_005634208	chromosome partitioning protein ParA [Parabacteroides merdae].	1.00E-135	100%
gene 20	102	WP_005634211	MULTISPECIES: hypothetical protein [Parabacteroides].	5.00E-51	100%
gene 21	48	WP_005634215	hypothetical protein [Parabacteroides merdae].	3.00E-14	100%
gene 22	548	WP_005634215	hypothetical protein [Parabacteroides merdae].	6.00E-285	94.20%
gene 23	468	WP_005634215	hypothetical protein [Parabacteroides merdae].	9.00E-247	96.90%
gene 24	134	WP_005644526	hypothetical protein [Parabacteroides merdae].	5.00E-25	95.20%
gene 25	755	WP_005634221	hypothetical protein [Parabacteroides merdae].	0	99.50%
gene 26	61	WP_005634223	hypothetical protein [Parabacteroides merdae].	5.00E-03	100%
gene 27	148	WP_022322918	putative DNA-binding protein [Parabacteroides merdae CAG:48].	8.00E-68	99.30%
gene 28	74	WP_005634231	MULTISPECIES: hypothetical protein [Parabacteroides].	2.00E-37	100%
gene 29	175	WP_005634233	MULTISPECIES: glucosaminidase [Parabacteroides].	9.00E-94	100%
gene 30	115	WP_005634235	hypothetical protein [Parabacteroides merdae].	3.00E-29	100%
gene 31	109	WP_005634237	hypothetical protein [Parabacteroides merdae].	2.00E-53	100%
gene 32	270	WP_022322920	ribonuclease Z [Parabacteroides merdae CAG:48].	3.00E-155	100%
gene 33	384	WP_005644540	cytochrome D ubiquinol oxidase, subunit II [Parabacteroidesmerdae].	2.00E-168	99.70%
gene 34	516	WP_005634245	cytochrome C oxidase subunit II [Parabacteroides merdae].	2.00E-290	100%
gene 35	107	WP_005634247	MULTISPECIES: membrane protein [Parabacteroides].	1.00E-41	100%
gene 36	490	WP_005634249	MULTISPECIES: oxidoreductase [Parabacteroides].	5.00E-297	100%
gene 37	284	WP_005634251	MULTISPECIES: glycosyl hydrolase [Parabacteroides].	2.00E-144	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	262	WP_032944525	thiamine biosynthesis protein ThiF [Bacteroides vulgatus].	1.00E-149	100%
gene 2	790	WP_032944607	histidine kinase [Bacteroides vulgatus].	0	99.90%
gene 3	435	WP_032944527	chemotaxis protein CheY [Bacteroides vulgatus].	2.00E-247	100%
gene 4	138	WP_032949291	hypothetical protein [Bacteroides vulgatus].	1.00E-73	100%
gene 5	128	WP_032944609	transcriptional regulator [Bacteroides vulgatus].	7.00E-54	100%
gene 6	92	WP_032944529	excisionase [Bacteroides vulgatus].	4.00E-44	96.70%
gene 7	103	WP_022161963	MULTISPECIES: uncharacterized protein [Bacteroides].	2.00E-51	98.10%
gene 8	447	WP_022161962	uncharacterized protein [Bacteroides sp. CAG:633].	2.00E-240	97.10%
gene 9	411	WP_022161961	tyrosine type site-specific recombinase [Bacteroides sp. CAG:633].	8.00E-235	99.80%
gene 10	352	WP_005810485	hypothetical protein [Bacteroides fragilis].	2.00E-201	100%
gene 11	257	WP_005824094	hypothetical protein [Bacteroides uniformis].	3.00E-130	100%
gene 12	176	WP_009037187	DNA-binding protein [Bacteroides sp. D20].	1.00E-94	100%
gene 13	353	WP_005835147	MULTISPECIES: 3-dehydroquinase synthase [Bacteroides].	2.00E-194	100%
gene 14	137	WP_009037184	membrane protein [Bacteroides sp. D20].	2.00E-60	100%
gene 15	974	WP_009037183	hypothetical protein [Bacteroides sp. D20].	0	99.70%
gene 16	479	WP_005824073	MULTISPECIES: cardiolipin synthetase [Bacteroides].	1.00E-277	100%
gene 17	177	WP_005824070	MULTISPECIES: methyltransferase [Bacteroides].	3.00E-98	100%
gene 18	259	WP_005835157	hypothetical protein [Bacteroides uniformis].	2.00E-147	100%
gene 19	228	WP_005824067	MULTISPECIES: hypothetical protein [Bacteroides].	4.00E-120	99.60%
gene 20	472	WP_005824065	MULTISPECIES: ATP-dependent endonuclease [Bacteroides].	7.00E-268	100%
gene 21	269	WP_005835166	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-153	100%
gene 22	872	WP_005824062	MULTISPECIES: alanyl-tRNA synthetase [Bacteroides].	0	100%
gene 23	322	WP_005824061	MULTISPECIES: peptidase M23 [Bacteroides].	5.00E-185	100%
gene 24	129	WP_005824060	MULTISPECIES: transcriptional regulator [Bacteroides].	2.00E-66	100%
gene 25	752	WP_005824059	MULTISPECIES: GTP pyrophosphokinase [Bacteroides].	0	99.90%
gene 26	431	WP_005824058	MULTISPECIES: lytic transglycosylase [Bacteroides].	1.00E-249	99.80%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	284	WP_007846654	integrase [Bacteroides dorei].	2.00E-148	98.60%
gene 2	436	WP_007846652	integrase [Bacteroides dorei].	3.00E-232	99.30%
gene 3	116	WP_005649903	tyrosine recombinase [Parabacteroides merdae].	2.00E-30	72.40%
gene 4	86	WP_032543834	hypothetical protein [Bacteroides fragilis].	2.00E-38	98.80%
gene 5	114	WP_032543834	hypothetical protein [Bacteroides fragilis].	6.00E-57	98.20%
gene 6	148	WP_032543835	hypothetical protein [Bacteroides fragilis].	5.00E-71	98.60%
gene 7	93	WP_013612210	MULTISPECIES: DNA-binding protein [Bacteroidales].	1.00E-45	98.90%
gene 8	165	WP_021645363	hypothetical protein [Bacteroides pyogenes].	9.00E-70	95.40%
gene 9	102	WP_021645362	hypothetical protein [Bacteroides pyogenes].	3.00E-51	99%
gene 10	329	WP_013612208	virulence-associated E family protein [Odoribacter splanchnicus].	6.00E-192	99.10%
gene 11	363	WP_008157037	mobilization protein [Parabacteroides johnsonii].	6.00E-199	97.20%
gene 12	127	WP_016661720	hypothetical protein [Bacteroides stercoris].	1.00E-62	99.20%
gene 13	307	WP_005645891	mobilization protein [Parabacteroides merdae].	2.00E-144	95.60%
gene 14	242	WP_032594609	hypothetical protein [Bacteroides fragilis].	2.00E-119	92.10%
gene 15	119	WP_009126239	hypothetical protein [Bacteroides fluxus].	4.00E-51	98.30%
gene 16	213	WP_005790176	hypothetical protein [Bacteroides fragilis].	6.00E-112	96.70%
gene 17	163	WP_022217865	uncharacterized protein [Bacteroides clarus CAG:160].	2.00E-77	97.10%
gene 18	281	WP_005790180	LuxR family transcriptional regulator [Bacteroides fragilis].	1.00E-133	95.30%
gene 19	464	WP_032496997	MULTISPECIES: ArsR family transcriptional regulator [Bacteroides].	2.00E-258	95.60%
gene 20	342	WP_009229000	hypothetical protein [Prevotella sp. oral taxon 299].	8.00E-111	58.40%
gene 21	87	WP_007834964	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-43	98.90%
gene 22	166	WP_007834963	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-94	100%
gene 23	187	WP_008656811	alkaline phosphatase [Bacteroides sp. 3_1_33FAA].	5.00E-99	100%
gene 24	284	WP_007851359	alkaline phosphatase [Bacteroides dorei].	4.00E-151	100%
gene 25	274	WP_007851360	competence protein ComEA [Bacteroides dorei].	5.00E-146	100%
gene 26	454	WP_007834957	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-226	100%
gene 27	128	WP_007834955	MULTISPECIES: hypothetical protein [Bacteroides].	3.00E-71	100%
gene 28	167	WP_007834952	MULTISPECIES: RNA polymerase sigma factor [Bacteroides].	4.00E-91	100%
gene 29	824	WP_007847083	MULTISPECIES: TonB-dependent receptor [Bacteroides].	0	99.80%
gene 30	156	WP_005845492	MULTISPECIES: adenylate cyclase [Bacteroides].	1.00E-85	100%
gene 31	61	WP_008781241	MULTISPECIES: hypothetical protein [Bacteroides].	9.00E-27	98.40%
gene 32	371	WP_007834946	MULTISPECIES: peptide chain release factor 2 [Bacteroides].	5.00E-204	100%
gene 33	601	WP_007842133	long-chain fatty acid-CoA ligase [Bacteroides dorei].	0	100%
gene 34	354	WP_007834942	MULTISPECIES: acetylornithine deacetylase [Bacteroides].	2.00E-200	99.70%
gene 35	60	WP_005849515	hypothetical protein [Bacteroides vulgatus].	1.00E-08	100%
gene 36	36	WP_007842131	hypothetical protein [Bacteroides dorei].	6.00E-12	100%
gene 37	182	WP_007842122	MULTISPECIES: capsular polysaccharide biosynthesis protein CapD[Bacteroides].	2.00E-98	100%

1
2
3

31i10



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	394	WP_028086446	hypothetical protein [Dorea longicatena].	7.00E-196	100%
gene 2	117	WP_012744294	hypothetical protein [Eubacterium rectale].	3.00E-42	88.80%
gene 3	933	WP_028086444	peptidase M23 [Dorea longicatena].	0	98.60%
gene 4	534	WP_024720789	MULTISPECIES: conjugal transfer protein TraE [unclassifiedClostridiales (miscellaneous)].	4.00E-297	99.60%



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	211	WP_003127055.1	MULTISPECIES: transcriptional regulator (Firmicutes)	5.00E-152	100%
gene 2	283	WP_015542745.1	Putative phage replication protein RstA (Ruminococcus obeum)	0	100%
gene 3	67	WP_002610247.1	MULTISPECIES: excisionase (Bacteria)	9.00E-40	100%
gene 4	397	WP_003436632.1	integrase (Clostridium difficile)	0.00E+00	100%
gene 5	222	WP_015529634.1	propionate CoA-transferase (Ruminococcus torques)	4.00E-117	99%
gene 6	643	WP_015529633.1	NADH:flavin oxidoreductases, Old Yellow Enzyme family (ruminococcus torques)	0	100%
gene 7	434	WP_015529632.1	hypothetical protein (ruminococcus torques)	0.00E+00	100%
gene 8	359	WP_015529631.1	hypothetical protein (ruminococcus torques)	0.00E+00	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	282	WP_032726907	levansucrase [Bacillus subtilis].	6.00E-134	99.20%
gene 2	149	WP_022111720	putative uncharacterized protein [Roseburia intestinalis CAG:13].	1.00E-51	100%
gene 3	343	WP_015521896	hypothetical protein [Roseburia intestinalis].	1.00E-169	93%
gene 4	177	WP_005348419	MULTISPECIES: quinolinate synthetase [Eubacterium].	0.88	16.80%
gene 5	312	WP_015561211	Cell wall-associated hydrolases (invasion-associated proteins)[Roseburia intestinalis].	7.00E-180	97.40%
gene 6	849	WP_015521894	hypothetical protein [Roseburia intestinalis].	0	92.50%
gene 7	87	WP_015521893	hypothetical protein [Roseburia intestinalis].	7.00E-43	98.90%
gene 8	733	WP_015521893	hypothetical protein [Roseburia intestinalis].	0	98.50%
gene 9	143	WP_015521892	hypothetical protein [Roseburia intestinalis].	7.00E-78	100%
gene 10	85	WP_015521891	hypothetical protein [Roseburia intestinalis].	5.00E-14	100%
gene 11	38	n/a	n/a	n/a	n/a
gene 12	526	WP_015521890	Putative phage replication protein RstA [Roseburia intestinalis].	2.00E-281	99%
gene 13	127	WP_015521889	hypothetical protein [Roseburia intestinalis].	2.00E-52	97.60%
gene 14	447	WP_015521888	DNA segregation ATPase FtsK/SpoIIIE and related proteins [Roseburia intestinalis].	4.00E-248	100%
gene 15	153	WP_015521887	Bacterial protein of unknown function (DUF961) [Roseburia intestinalis].	1.00E-81	99.30%
gene 16	115	WP_015521885	Bacterial protein of unknown function (DUF961) [Roseburia intestinalis].	2.00E-58	100%
gene 17	177	WP_015561495	Bacterial protein of unknown function (DUF961) [Roseburia intestinalis].	9.00E-97	100%
gene 18	245	WP_015561206	hypothetical protein [Roseburia intestinalis].	3.00E-126	93.60%
gene 19	73	WP_015561492	hypothetical protein [Roseburia intestinalis].	5.00E-25	80.80%
gene 20	74	WP_015561491	hypothetical protein [Roseburia intestinalis].	4.00E-34	95.90%
gene 21	63	WP_015521883	hypothetical protein [Roseburia intestinalis].	3.00E-30	98.40%
gene 22	73	WP_015561492	hypothetical protein [Roseburia intestinalis].	2.00E-34	100%
gene 23	74	WP_015561491	hypothetical protein [Roseburia intestinalis].	5.00E-36	100%
gene 24	95	WP_015561490	hypothetical protein [Roseburia intestinalis].	3.00E-38	100%
gene 25	60	WP_015561489	hypothetical protein [Roseburia intestinalis].	8.00E-23	100%
gene 26	52	WP_015561488	hypothetical protein [Roseburia intestinalis].	2.00E-06	100%
gene 27	60	WP_015561487	hypothetical protein [Roseburia intestinalis].	4.00E-27	100%
gene 28	88	WP_015561486	hypothetical protein [Roseburia intestinalis].	7.00E-21	100%
gene 29	68	WP_015521878	Helix-turn-helix [Roseburia intestinalis].	2.00E-29	98.50%
gene 30	71	WP_015521877	hypothetical protein [Roseburia intestinalis].	2.00E-33	100%
gene 31	96	WP_015521876	hypothetical protein [Roseburia intestinalis].	1.00E-49	100%
gene 32	1134	WP_006856044	hypothetical protein [Roseburia intestinalis].	9.00E-260	81.50%
gene 33	213	WP_003228230	levansucrase [Bacillus subtilis].	2.00E-92	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	449	WP_011965543	peptidase S41 [Bacteroides vulgatus].	3.00E-261	99.80%
gene 2	546	WP_005846096	MULTISPECIES: aspartate aminotransferase [Bacteroides].	0	100%
gene 3	142	WP_005848641	transporter [Bacteroides vulgatus].	5.00E-46	100%
gene 4	424	WP_005846098	transporter [Bacteroides vulgatus].	9.00E-208	99.80%
gene 5	273	WP_032946848	hypothetical protein [Bacteroides vulgatus].	4.00E-156	99.60%
gene 6	725	WP_005846102	hypothetical protein [Bacteroides vulgatus].	0	100%
gene 7	394	WP_005846104	membrane protein [Bacteroides vulgatus].	9.00E-215	99.70%
gene 8	1145	WP_022508443	MULTISPECIES: transcription-repair coupling factor [Bacteroides].	0	100%
gene 9	228	WP_005848631	MULTISPECIES: cell envelope biogenesis protein TonB [Bacteroides].	1.00E-82	100%
gene 10	253	WP_032952525	amidohydrolase [Bacteroides vulgatus].	2.00E-147	100%
gene 11	538	WP_032952524	indolepyruvate ferredoxin oxidoreductase [Bacteroides vulgatus].	1.00E-305	100%
gene 12	193	WP_005846114	MULTISPECIES: indolepyruvate oxidoreductase [Bacteroides].	6.00E-93	100%
gene 13	435	WP_005848625	MULTISPECIES: aminotransferase [Bacteroides].	4.00E-232	100%
gene 14	433	WP_005846118	MULTISPECIES: phenylacetate-CoA ligase [Bacteroidales].	2.00E-249	100%
gene 15	138	WP_005846121	MULTISPECIES: amino acid-binding protein [Bacteroides].	1.00E-70	100%
gene 16	524	WP_016270074	V-type H(+)-translocating pyrophosphatase [Bacteroides vulgatus].	5.00E-255	100%
gene 17	202	WP_008671042	potassium transporter [Bacteroides sp. 4_3_47FAA].	3.00E-71	100%
gene 18	642	WP_022508447	uncharacterized protein [Bacteroides vulgatus CAG:6].	0	99.40%
gene 19	105	WP_005846128	MULTISPECIES: hypothetical protein [Bacteroides].	3.00E-52	100%
gene 20	975	WP_032952522	xanthan lyase [Bacteroides vulgatus].	0	100%
gene 21	323	WP_005848617	alpha-1,3-galactosidase B [Bacteroides vulgatus].	7.00E-177	96.50%
gene 22	282	WP_032952521	alpha-1,3-galactosidase B [Bacteroides vulgatus].	1.00E-164	100%
gene 23	185	WP_005848615	MULTISPECIES: acyltransferase [Bacteroides].	3.00E-107	100%
gene 24	82	WP_005846136	MULTISPECIES: hypothetical protein [Bacteroides].	2.00E-37	100%
gene 25	547	WP_005846138	MULTISPECIES: diphosphate--fructose-6-phosphate1-phosphotransferase [Bacteroides].	0	100%
gene 26	288	WP_005848612	MULTISPECIES: glycosyl hydrolase [Bacteroides].	4.00E-148	100%
gene 27	122	WP_005846143	MULTISPECIES: membrane protein [Bacteroides].	2.00E-64	100%
gene 28	281	WP_005848610	MULTISPECIES: 50S ribosomal protein L11 methyltransferase[Bacteroides].	9.00E-143	100%
gene 29	86	WP_005846148	MULTISPECIES: glutaredoxin [Bacteroides].	6.00E-44	100%

1
2
3

33g20



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	46	WP_032600916	hypothetical protein, partial [Bacteroides fragilis].	1.00E-17	100%
gene 2	438	WP_005817824	aminotransferase [Bacteroides fragilis].	7.00E-238	100%
gene 3	377	WP_032529462	aldo/keto reductase [Bacteroides fragilis].	5.00E-220	100%
gene 4	330	WP_005787282	MULTISPECIES: NADPH-flavin oxidoreductase [Bacteroides].	9.00E-188	100%
gene 5	283	WP_009292263	MULTISPECIES: 2,5-diketo-D-gluconic acid reductase [Bacteroides].	7.00E-162	100%
gene 6	84	WP_007481879	AraC family transcriptional regulator [Bacteroides salyersiae].	3.00E-12	94.60%
gene 7	151	WP_032590016	aldo/keto reductase [Bacteroides fragilis].	2.00E-81	100%
gene 8	97	WP_032556675	hypothetical protein [Bacteroides fragilis].	5.00E-52	100%
gene 9	288	WP_005794287	hypothetical protein [Bacteroides fragilis].	1.00E-171	100%
gene 10	477	WP_032529461	cation transporter [Bacteroides fragilis].	1.00E-271	100%
gene 11	387	WP_032529460	cystathionine beta-lyase [Bacteroides fragilis].	4.00E-207	100%
gene 12	314	WP_032529459	aldo/keto reductase [Bacteroides fragilis].	9.00E-180	100%
gene 13	337	WP_032529503	amidohydrolase [Bacteroides fragilis].	7.00E-194	100%
gene 14	53	WP_014298850	short-chain dehydrogenase [Bacteroides fragilis].	7.00E-18	100%
gene 15	231	WP_008660000	MULTISPECIES: short-chain dehydrogenase [Bacteroides].	4.00E-125	100%
gene 16	270	WP_010992873	AraC family transcriptional regulator [Bacteroides fragilis].	7.00E-155	100%
gene 17	277	WP_005787306	gamma-carboxymuconolactone decarboxylase [Bacteroides fragilis].	3.00E-150	99.60%
gene 18	190	WP_032529457	flavodoxin [Bacteroides fragilis].	4.00E-105	100%
gene 19	125	WP_010992875	acetyltransferase [Bacteroides fragilis].	3.00E-66	100%
gene 20	354	WP_032529456	beta-lactamase [Bacteroides fragilis].	2.00E-202	100%
gene 21	136	WP_014298855	hypothetical protein [Bacteroides fragilis].	4.00E-73	99.30%
gene 22	250	WP_032529455	oxidoreductase [Bacteroides fragilis].	4.00E-138	100%
gene 23	387	WP_032574206	aldo/keto reductase [Bacteroides fragilis].	2.00E-202	100%
gene 24	217	WP_032529453	5-amino-6-(5-phosphoribosylamino)uracil reductase [Bacteroides fragilis].	8.00E-122	100%
gene 25	356	WP_032529452	alpha/beta hydrolase [Bacteroides fragilis].	6.00E-204	100%
gene 26	411	WP_032529451	membrane protein [Bacteroides fragilis].	1.00E-208	100%
gene 27	338	WP_032529450	aldo/keto reductase [Bacteroides fragilis].	3.00E-195	100%
gene 28	161	WP_032529449	membrane protein [Bacteroides fragilis].	8.00E-90	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	279	WP_005849599	MULTISPECIES: glycerol acyltransferase [Bacteroides].	7.00E-162	100%
gene 2	326	WP_005845337	MULTISPECIES: hemolysin [Bacteroides].	2.00E-187	100%
gene 3	443	WP_008781261	MULTISPECIES: dehydrogenase [Bacteroides].	2.00E-254	100%
gene 4	144	WP_005845350	MULTISPECIES: deoxyuridine 5'-triphosphate nucleotidohydrolase[Bacteroides].	2.00E-76	100%
gene 5	585	WP_005845352	MULTISPECIES: hypothetical protein [Bacteroides].	0	100%
gene 6	277	WP_016271912	hypothetical protein [Bacteroides vulgatus].	6.00E-154	100%
gene 7	444	WP_016271913	hypothetical protein [Bacteroides vulgatus].	3.00E-192	100%
gene 8	69	WP_008668465	hypothetical protein [Bacteroides vulgatus].	4.00E-30	100%
gene 9	123	WP_005845358	MULTISPECIES: RNA-binding protein [Bacteroides].	3.00E-39	100%
gene 10	304	WP_005849590	MULTISPECIES: glutaminase [Bacteroides].	4.00E-164	100%
gene 11	614	WP_032952213	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase [Bacteroidesvulgatus].	0	100%
gene 12	169	WP_005845363	MULTISPECIES: phosphoribosylaminoimidazole carboxylase[Bacteroides].	9.00E-73	100%
gene 13	126	WP_005845365	MULTISPECIES: glycine cleavage system protein H [Bacteroides].	3.00E-66	100%
gene 14	218	WP_005845367	membrane protein [Bacteroides vulgatus].	3.00E-116	100%
gene 15	489	WP_005845369	MULTISPECIES: RNA polymerase sigma54 factor [Bacteroides].	6.00E-263	100%
gene 16	88	WP_032952214	xaa-Pro aminopeptidase [Bacteroides vulgatus].	3.00E-38	100%
gene 17	302	WP_032952215	ferredoxin [Bacteroides vulgatus].	2.00E-177	100%
gene 18	808	WP_032952216	formate acetyltransferase [Bacteroides vulgatus].	0	100%
gene 19	241	WP_005845380	MULTISPECIES: ATPase AAA [Bacteroides].	2.00E-137	100%
gene 20	74	WP_005845380	MULTISPECIES: ATPase AAA [Bacteroides].	9.00E-32	100%
gene 21	342	WP_004323659	MULTISPECIES: mobilization protein [Bacteroidales].	5.00E-187	99.40%
gene 22	95	WP_004310673	MULTISPECIES: transcriptional regulator [Bacteroidales].	3.00E-46	100%
gene 23	302	WP_007479974	MULTISPECIES: hypothetical protein [Bacteroidales].	1.00E-164	100%
gene 24	428	WP_008668453	MULTISPECIES: mannose-1-phosphate guanylyltransferase[Bacteroides].	8.00E-191	100%
gene 25	135	WP_005845399	MULTISPECIES: HIT family hydrolase [Bacteroides].	3.00E-71	100%
gene 26	154	WP_005845398	MULTISPECIES: transcription elongation factor GreA [Bacteroides].	1.00E-76	100%
gene 27	415	WP_005845396	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-209	100%
gene 28	728	WP_011965175	MULTISPECIES: polynucleotide phosphorylase [Bacteroides].	0	100%
gene 29	54	WP_005845392	MULTISPECIES: rubredoxin [Bacteroides].	3.00E-27	100%
gene 30	259	WP_032952218	NAD-dependent protein deacylase [Bacteroides vulgatus].	3.00E-129	100%
gene 31	194	WP_005845388	MULTISPECIES: peptidylprolyl isomerase [Bacteroides].	9.00E-105	100%
gene 32	288	WP_005845386	MULTISPECIES: peptidylprolyl isomerase [Bacteroides].	9.00E-141	100%

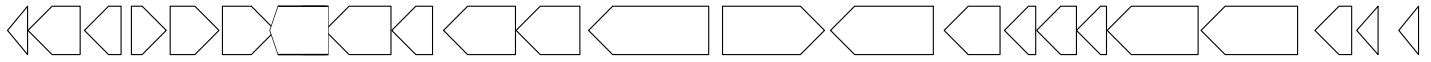
1
2
3

37b14



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	210	WP_005784848	chromosome partitioning protein ParB [Bacteroides fragilis].	2.00E-110	100%
gene 2	288	WP_005784846	MULTISPECIES: hypothetical protein [Bacteroides].	5.00E-149	100%
gene 3	439	WP_010992147	lytic transglycosylase [Bacteroides fragilis].	2.00E-244	100%
gene 4	747	WP_005784842	MULTISPECIES: GTP pyrophosphokinase [Bacteroides].	0	100%
gene 5	452	WP_008667781	MULTISPECIES: transposase [Bacteroides].	1.00E-259	99.60%
gene 6	121	WP_005784840	MULTISPECIES: transcriptional regulator [Bacteroides].	8.00E-62	100%
gene 7	322	WP_005796222	MULTISPECIES: peptidase M23 [Bacteroides].	1.00E-184	100%
gene 8	872	WP_005796226	MULTISPECIES: alanyl-tRNA synthetase [Bacteroides].	0	99.90%
gene 9	284	WP_005784835	MULTISPECIES: AraC family transcriptional regulator [Bacteroides].	1.00E-163	100%
gene 10	372	WP_005784833	MULTISPECIES: MFS transporter [Bacteroides].	4.00E-192	100%
gene 11	738	WP_008658550	MULTISPECIES: alpha-1 2-mannosidase [Bacteroides].	0	100%
gene 12	184	WP_005784829	MULTISPECIES: RNA polymerase sigma70 factor [Bacteroides].	2.00E-94	100%
gene 13	342	WP_010992141	anti-sigma factor [Bacteroides fragilis].	3.00E-188	100%
gene 14	755	WP_032595539	alpha-mannosidase [Bacteroides fragilis].	0	99.70%
gene 15	741	WP_014298248	alpha-1 2-mannosidase [Bacteroides fragilis].	0	99.90%
gene 16	766	WP_005784823	membrane protein [Bacteroides fragilis].	0	99.70%
gene 17	191	WP_010992139	membrane protein [Bacteroides fragilis].	2.00E-106	100%

1
2
3



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	122	WP_008657495	hypothetical protein [Bacteroides sp. 3_2_5].	1.00E-67	100%
gene 2	132	WP_005802341	hypothetical protein [Bacteroides fragilis].	2.00E-68	100%
gene 3	228	WP_008657494	MULTISPECIES: hypothetical protein [Bacteroides].	7.00E-131	100%
gene 4	600	WP_005790703	MULTISPECIES: dihydroxy-acid dehydratase [Bacteroides].	0	100%
gene 5	565	WP_005790701	MULTISPECIES: acetolactate synthase [Bacteroides].	0	100%
gene 6	186	WP_032496973	MULTISPECIES: acetoxyacid synthase small subunit[Bacteroides].	8.00E-91	100%
gene 7	247	WP_005797964	MULTISPECIES: acyl-ACP thioesterase [Bacteroides].	9.00E-130	100%
gene 8	194	WP_008657489	MULTISPECIES: hypothetical protein [Bacteroides].	5.00E-108	100%
gene 9	347	WP_008769939	ketol-acid reductoisomerase [Bacteroides fragilis].	2.00E-189	100%
gene 10	639	WP_008657488	MULTISPECIES: hypothetical protein [Bacteroides].	0	100%
gene 11	634	WP_008657487	MULTISPECIES: helicase [Bacteroides].	0	100%
gene 12	747	WP_032529063	aconitate hydratase [Bacteroides fragilis].	0	100%
gene 13	396	WP_005790683	MULTISPECIES: isocitrate dehydrogenase [Bacteroides].	2.00E-218	100%
gene 14	447	WP_005790680	MULTISPECIES: citrate synthase [Bacteroides].	2.00E-230	100%
gene 15	253	WP_005797955	3-oxo-5-alpha-steroid 4-dehydrogenase [Bacteroides fragilis].	9.00E-143	100%
gene 16	410	WP_032529062	NADH:flavin oxidoreductase [Bacteroides fragilis].	1.00E-231	100%
gene 17	290	WP_005790670	oxidoreductase [Bacteroides fragilis].	4.00E-157	100%
gene 18	329	WP_011203422	6-phosphofructokinase [Bacteroides fragilis].	6.00E-173	100%
gene 19	303	WP_005797949	MULTISPECIES: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase[Bacteroides].	7.00E-163	100%
gene 20	216	WP_005790663	cytidylate kinase [Bacteroides fragilis CAG:47].	1.00E-105	100%
gene 21	227	WP_025813996	energy transducer TonB [Bacteroides fragilis].	2.00E-98	100%
gene 22	324	WP_005790657	MULTISPECIES: isoprenyl synthetase [Bacteroides].	4.00E-182	100%
gene 23	125	WP_005790656	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-64	100%

1
2
3

40k7



Gene	Length	Accession Number	Hit Definition	E Value	Identity
gene 1	54	WP_008020625	MULTISPECIES: lytic transglycosylase [Bacteroides].	4.00E-22	98.10%
gene 2	746	WP_009039031	GTP pyrophosphokinase [Bacteroides sp. D22].	0	99.70%
gene 3	114	WP_004301384	MULTISPECIES: transcriptional regulator [Bacteroides].	6.00E-58	100%
gene 4	286	WP_004315137	MULTISPECIES: peptidase M23 [Bacteroides].	2.00E-163	100%
gene 5	872	WP_008642071	alanyl-tRNA synthetase [Bacteroides sp. 1_1_30].	0	100%
gene 6	732	WP_015532344	alpha-1,2-mannosidase, putative [Bacteroides xylanisolvens].	0	99.70%
gene 7	185	WP_008776778	RNA polymerase sigma70 factor [Bacteroides sp. 2_2_4].	1.00E-101	98.90%
gene 8	335	WP_004301377	MULTISPECIES: anti-sigma factor [Bacteroides].	7.00E-189	100%
gene 9	768	WP_004315134	MULTISPECIES: alpha-1 2-mannosidase [Bacteroides].	0	100%
gene 10	755	WP_004315133	MULTISPECIES: alpha-1 2-mannosidase [Bacteroides].	0	99.20%
gene 11	63	WP_008020615	hypothetical protein [Bacteroides xylanisolvens].	8.00E-28	100%
gene 12	218	WP_032851020	Fic family protein [Bacteroides sp. D22].	3.00E-88	96.40%
gene 13	434	WP_004315131	MULTISPECIES: phage tail protein [Bacteroides].	4.00E-252	99.30%
gene 14	495	WP_008020609	endo-beta-N-acetylglucosaminidase [Bacteroides xylanisolvens].	1.00E-277	99.60%
gene 15	386	WP_008020607	phospholipase [Bacteroides xylanisolvens].	3.00E-223	100%
gene 16	356	WP_008020604	hypothetical protein [Bacteroides xylanisolvens].	6.00E-204	98.60%
gene 17	504	WP_008771072	MULTISPECIES: hypothetical protein [Bacteroides].	1.00E-299	100%
gene 18	1130	WP_004312951	MULTISPECIES: membrane protein [Bacteroides].	0	99.90%
gene 19	367	WP_004314589	MULTISPECIES: hypothetical protein [Bacteroides].	7.00E-188	95.90%
gene 20	390	WP_004314590	MULTISPECIES: hypothetical protein [Bacteroides].	6.00E-224	99.70%
gene 21	442	WP_004314591	MULTISPECIES: putative lipoprotein [Bacteroides].	4.00E-253	99.30%

1
2
3

1 **Table S2C. Swissprot analysis.** Each effector gene was aligned to the Swissprot database using ngKlast.
 2 The top Swissprot hit is recorded along with the e value, percent identity (over the full effector gene), percent
 3 identity (over the portion aligned), and taxonomy.
 4

Cbeg	Effector Gene	Protein Length	Hit Accession	Hit Definition	E Value	Identity	Align. Length	Align Id	Organism	Taxonomy
ppGpp	1i11	701	RELA_MYXXA	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	6.00E-117	45%	538	62.50%	Myxococcus xanthus	Proteobacteria
15	2m03	376	MURG_BACV8	RecName: Full=UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase {ECO:0000255 HAMAP-Rule:MF_00033};EC=2.4.1.227 {ECO:0000255 HAMAP-Rule:MF_00033};AltName: Full=Undecaprenyl-PP-MurNAc-pentapeptide-UDPGlcNAc GlcNAc transferase {ECO:0000255 HAMAP-Rule:MF_00033};	3.00E-165	83.70%	374	87.20%	Bacteroides vulgatus ATCC 8482	Bacteroidetes
ppGpp	3b03	746	RELA_MYXXA	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	2.00E-108	42.30%	742	52.70%	Myxococcus xanthus	Proteobacteria
14	3e18	874	SECY_STRSC	RecName: Full=Protein translocase subunit SecY {ECO:0000255 HAMAP-Rule:MF_01465};	0.31	28.70%	87	46%	Streptomyces scabiei	Actinobacteria
13	4h09	604	YVGJ_BACSU	RecName: Full=Lipoteichoic acid synthase-like YvgJ;Contains:RecName: Full=Uncharacterized protein YvgJ;Contains:RecName: Full=Processed uncharacterized protein YvgJ;	1.00E-26	24.20%	322	41.30%	Bacillus subtilis subsp. subtilis str. 168	Firmicutes
5	4i05	959	XANLY_BACGL	RecName: Full=Xanthan lyase {ECO:0000312 EMBL:BAB21059.1};EC=4.2.2.12 {ECO:0000269 PubMed:9758797};Flags: Precursor;	1.00E-05	36.40%	88	52.30%	Bacillus sp. GL1	Firmicutes
12-2	8j18	326	ASCC3_RAT	RecName: Full=Activating signal cointegrator 1 complex subunit 3;EC=3.6.4.12;	1	22.20%	135	41.50%	Rattus norvegicus	Eukaryota
12-3	10o08	326	ASCC3_RAT	RecName: Full=Activating signal cointegrator 1 complex subunit 3;EC=3.6.4.12;	1	22.20%	135	41.50%	Rattus norvegicus	Eukaryota
ppGpp	15d03	756	RELA_MYXXA	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	8.00E-113	38.90%	737	52.90%	Myxococcus xanthus	Proteobacteria
ppGpp	15m04	747	RELA_STAAN	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	1.00E-104	45.40%	723	53.50%	Staphylococcus aureus subsp. aureus N315	
8	16f15	299	ALYS_ENTFA	RecName: Full=Autolysin;EC=3.2.1.-;AltName: Full=Beta-glycosidase;AltName: Full=Peptidoglycan hydrolase;Flags: Precursor;	5.00E-17	27.80%	295	47.50%	Enterococcus faecalis V583	Firmicutes
2	18b09_1	1146	SUSC_BACTN	RecName: Full=TonB-dependent receptor SusC;AltName: Full=Starch-utilization system protein C;Flags: Precursor;	1.00E-92	28.50%	1063	45%	Bacteroides thetaiotaomicron VPI-5482	Bacteroidetes
3	18b09_2	539	SUSD_BACO1	RecName: Full=SusD-like protein BACOVA_02651;Flags: Precursor;	8.00E-10	26.40%	197	45.70%	Bacteroides ovatus ATCC 8483	Bacteroidetes
ppGpp	21f12	756	RELA_MYXXA	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	8.00E-113	38.90%	737	52.90%	Myxococcus xanthus	Proteobacteria

5
6
7

Cbeg	Effector Gene	Protein Length	Hit Accession	Hit Definition	E Value	Identity	Align. Length	Align Id	Organism	Taxonomy
ppGpp	24i11	752	RELA_MYXXA	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	9.00E-109	40.30%	740	51.90%	Myxococcus xanthus	Proteobacteria
1	25h09	824	BTUB_VIBVU	RecName: Full=Vitamin B12 transporter BtuB {ECO:0000255} HAMAP-Rule:MF_01531;AltName: Full=Cobalamin receptor {ECO:0000255} HAMAP-Rule:MF_01531;AltName: Full=Outer membrane cobalamin translocator {ECO:0000255} HAMAP-Rule:MF_01531;Flags: Precursor;	7.00E-06	21.20%	510	39.40%	Vibrio vulnificus CMCP6	Proteobacteria
9	31i10	933	P60_LISSE	RecName: Full=Probable endopeptidase p60;EC=3.4.-.-;AltName: Full=Invasion-associated protein p60;Flags: Precursor;	2.00E-22	46.80%	124	62.10%	Listeria seeligeri	Firmicutes
7	32o03_1	312	YDDH_BACSU	RecName: Full=Probable endopeptidase YddH;EC=3.4.-.-;Flags: Precursor;	2.00E-50	35.40%	314	54.80%	Bacillus subtilis subsp. subtilis str. 168	Firmicutes
6	32o03_2	849	YDDG_BACSU	RecName: Full=Uncharacterized membrane protein YddG;Flags: Precursor;	1.00E-13	22.70%	300	48%	Bacillus subtilis subsp. subtilis str. 168	Firmicutes
4	33g04	975	XANLY_BACGL	RecName: Full=Xanthan lyase {ECO:0000312} EMBL: BAB21059.1;EC=4.2.2.12 {ECO:0000269} PubMed:9758797;Flags: Precursor;	2.00E-06	31.90%	138	46.40%	Bacillus sp. GL1	Firmicutes
10	33g20	354	Y906_MYCTO	RecName: Full=Uncharacterized protein MT0929;Flags: Precursor;	1.00E-49	37.50%	301	55.10%	Mycobacterium tuberculosis CDC1551	Actinobacteria
12	35i19	326	ASCC3_RAT	RecName: Full=Activating signal cointegrator 1 complex subunit 3;EC=3.6.4.12;	1.31	22.20%	135	41.50%	Rattus norvegicus	Eukaryota
ppGpp	37b14	747	RELA_BACSU	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	3.00E-107	38.70%	762	49%	Bacillus subtilis subsp. subtilis str. 168	Firmicutes
11	39k17	247	FATA_CORSA	RecName: Full=Oleoyl-acyl carrier protein thioesterase, chloroplastic;EC=3.1.2.14;AltName: Full=18:0-acyl-carrier protein thioesterase;Short=18:0-ACP thioesterase;AltName: Full=Acyl-[acyl-carrier-protein] hydrolase;Flags: Precursor; Fragment;	1.00E-12	24.20%	219	45.70%	Coriandrum sativum	Eukaryota
ppGpp	40k7	746	RELA_MYXXA	RecName: Full=GTP pyrophosphokinase;EC=2.7.6.5;AltName: Full=(p)ppGpp synthase;AltName: Full=ATP:GTP 3'-pyrophosphotransferase;AltName: Full=ppGpp synthase I;	6.00E-107	42.30%	742	52.60%	Myxococcus xanthus	Proteobacteria

1

2

Table S3: NMR spectroscopic data for compounds **1** – **4** in DMSO-*d*₆.

No.	compound 1			compound 2			compound 3			compound 4		
	δ_C^b	δ_H^a	<i>J</i> in Hz	δ_C^b	δ_H^a	<i>J</i> in Hz	δ_C^b	δ_H^a	<i>J</i> in Hz	δ_C^b	δ_H^a	<i>J</i> in Hz
1	170.4			170.4			170.4			170.4		
2	40.5	3.78	dd (17.1, 6.2)	40.5	3.78	dd (17.1, 6.2)	40.5	3.78	dd (17.1, 6.2)	40.5	3.78	dd (17.1, 6.2)
		3.85	dd (17.1, 6.2)		3.85	dd (17.1, 6.2)		3.85	dd (17.1, 6.2)		3.85	dd (17.1, 6.2)
3	171.4			171.4			171.4			171.4		
4	43.6	2.19	dd (14.0, 6.0)	43.5	2.19	dd (14.0, 6.0)	43.5	2.19	dd (14.0, 6.0)	43.6	2.19	dd (14.0, 6.0)
		2.23	dd (14.0, 6.0)		2.23	dd (14.0, 6.0)		2.23	dd (14.0, 6.0)		2.23	dd (14.0, 6.0)
5	67.3	3.78	m	67.3	3.78	m	67.3	3.78	m	67.3	3.78	m
6	36.7	1.28	m	36.7	1.28	m	36.7	1.28	m	36.7	1.28	m
7	25.1 ^c	1.24	m	24.9 ^c	1.24	m	25.1 ^c	1.24	m	25.1	1.24	m
8	29.1	1.24	m	28.7 ^c	1.24	m	29.1 ^c	1.24	m	29.1	1.24	m
9	29.1	1.24	m	29.2 ^c	1.29	m	28.9 ^c	1.24	m	29.1	1.24	m
10	29.1	1.24	m	26.6	1.99	m	28.6 ^c	1.24	m	29.1	1.24	m
11	29.1	1.24	m	129.6	5.33	m	19.1	1.29	m	29.1	1.24	m
12	29.1	1.24	m	129.7	5.33	m	26.6	1.99	m	29.1	1.24	m
13	29.1	1.24	m	26.6	1.99	m	129.6	5.33	m	29.0	1.24	m
14	29.1	1.24	m	29.1 ^c	1.29	m	129.7	5.33	m	31.3	1.24	m
15	29.0 ^c	1.24	m	28.3 ^c	1.24	m	26.6	1.99	m	22.1	1.26	m
16	31.3	1.24	m	31.1	1.24	m	29.1	1.29	m	13.8	1.85	t (7.1)
17	22.1	1.26	m	22.1	1.26	m	28.3 ^c	1.24	m			

18	13.8	1.85	t (7.1)	13.9	1.85	t (7.1)	31.1	1.24	m			
19							22.1	1.26	m			
20							13.9	1.85	t (7.1)			
NH		8.23	t (6.2)		8.23	t (6.2)		8.23	t (6.2)		8.23	t (6.2)
OH		4.57	brs		4.57	brs		4.57	brs		4.57	brs

^arecorded at 600 MHz, ^brecorded at 150 MHz, ^{a,b}signals were referenced to the DMSO-*d*₆ solvent signals (δ_{H}

2.50 and δ_{H} 39.51), ^cambiguously assigned by comparing with predicted chemical shifts from ChemBioDraw

13.0.

1
2
3
4
5
6
7
8
9
0
1
2
3
4