

SUPPORTING INFORMATION

Identification of Putative Genes Involved in Bisphenol A degradation Using Differential Protein

*Abundance Analysis of *Sphingobium* sp. BiD32*

Nicolette A. Zhou^{†,‡}, Henrik Kjeldal[†], Heidi L. Gough[‡], Jeppe L. Nielsen^{*,†}

[†]Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H, DK-9220

Aalborg, Denmark

[‡]Department of Civil and Environmental Engineering, University of Washington, More Hall 201 Box

352700, Seattle, Washington 98195-2700, United States

*corresponding author, (+45) 99 40 85 06; fax (+45) 98 14 18 08; jln@bio.aau.dk

Submitted to Environmental Science & Technology

Table of contents

<i>Library preparation and genome sequencing</i>	S2
<i>Protein extraction</i>	S2
<i>In-solution digestion</i>	S3
<i>LC/MS analyses for metabolomics</i>	S3
Table S1. Summary of essential genes in <i>Sphingobium</i> sp. BiD32	S5
Table S2. DNA and protein sequences of <i>Sphingobium</i> sp. BiD32 genes and proteins of interest	S15
Figure S1. Phylogenetic relationship of <i>Sphingobium</i> sp. BiD32 proteins differentially expressed and functionally annotated to xenobiotics biodegradation and metabolism (blue), <i>Sphingobium</i> sp. BiD32 proteins functionally annotated to bisphenol degradation (green), <i>Sphingobium</i> sp. BiD32 proteins related to proteins responsible for BPA degradation by <i>Sphingomonas</i> sp. AO1 (purple), and <i>Sphingomonas</i> sp. AO1 proteins responsible for BPA degradation (orange).	S18

Library preparation and genome sequencing

DNA extraction was performed using the FastDNA Spin Kit for Soil (MP Biomedicals, Santa Ana, CA, USA). The DNA was cleaned up by the following process: mix 100 μ L of DNA 1:1 with phenol:chloroform:isoamyl alcohol (25:24:1, pH $\sim$ 7.5), vortex, centrifuge (10,000 $\times g$, 10 min), transfer the upper layer to a fresh tube, add 1 μ L of 20 mg/mL glycogen, 10 μ L of 7.5 M NH_4Ac , and 250 μ L 99% (v/v) ethanol, incubate for 5 min at room temperature, centrifuge (10,000 $\times g$, 15 min), discard supernatant, dry pellet 15 min, and dissolve in 100 μ L water. The library for paired-end sequencing was constructed using the Nextera DNA Sample Preparation Kit (Illumina, San Diego, CA, USA). The manufacturer's instructions were followed with the following modification: the clean-up of the tagmented DNA was performed using the MinElute Reaction Cleanup Kit (Qiagen Inc., Valencia, CA, USA). The library was sequenced using an Illumina HiSeq as previously described¹ generating paired-end reads of 150 bp in length.

Protein extraction

The cell pellet was resuspended in 10 mL PBS supplemented with protease inhibitors (Complete, Mini, EDTA-free Protease Inhibitor Cocktail Tablets; 1 Tablet in 50 mL solution; Boehringer Mannheim, Mannheim, Germany). The cells were disrupted by sonication (cycle 0.5, 10 min, 20W power output; Sonopuls HD2200, SH213G Booster Horn, TT13 sonotrode, Bandelin Electronic, Berlin, Germany), cell debris was removed by centrifugation (6000 $\times g$, 4 $^{\circ}\text{C}$, 4 min), and the supernatant was centrifuged (30,000 $\times g$, 4 $^{\circ}\text{C}$, 60 min). The supernatant (cytosolic fraction) was collected in a fresh tube ($\sim$ 10 mL). The pellet was resuspended in 10 mL PBS supplemented with protease inhibitor cocktail tablets, centrifuged (30,000 $\times g$, 4 $^{\circ}\text{C}$, 60 min), and the supernatant was discarded. The pellet (membrane-associated fraction) was resuspended in 5 mL PBS supplemented with the protease inhibitor cocktail tablets. The proteins were concentrated by acetone precipitation. The BCA Protein Assay (Thermo Fisher Scientific) was used to quantify the protein content prior to in-solution digestion.

In-solution digestion

Twenty (20) µg of protein was digested for each sample as follows: digestion buffer was added 1:1 with the protein samples for a final concentration of 1% (w/v) sodium deoxycholate and 50 mM triethylammonium bicarbonate. The mixture was heated to 99 °C for 10 minutes, then cooled to 37 °C. Protein extracts were reduced by adding tris(2-carboxyethyl)phosphine (TCEP) (1 µg TCEP/25 µg protein) and incubating samples at 37 °C for 30 min. Following reduction, proteins were alkylated by adding iodoacetic acid (1 µg iodoacetic acid/10 µg protein) and the proteins were incubated in the dark (37 °C, 20 min). Finally, proteins were mixed with trypsin in a ratio of 1:50 (w/w) and digested overnight at 37 °C. Samples were acidified and sodium deoxycholate was pelleted by centrifugation (10,000 x g, 4 °C, 10 min). The supernatant was transferred to a fresh tube, dried in a vacuum concentrator (Labconco, Buch & Holm, Herlev, Denmark), and resuspended in 0.1% (v/v) trifluoroacetic acid (TFA) and 0.005% (v/v) heptafluorobutyric Acid (HFBA).

LC/MS analyses for metabolomics

An Ultra High Definition Accurate-Mass Quadrupole Time-of-Flight coupled with an ultra-HPLC system (Agilent Technologies, Inc.; Santa Clara, CA, USA) was used to determine the presence of BPA and corresponding metabolites. Gradient elution was performed using an ACQUITY UPLC BEH C18 column (2.1 x 100 mm, 1.7 µm; Waters Co., Milford, MA, USA). The flow rate was set to 0.4 mL/min and the gradient was programmed as follows: 0.00-5.00 2 % B, 5.00-7.00 7 % B, 7.00-9.00 15% B, 9.00-11.50 20% B, 11.50-14.00 25% B, 14.00-15.50 30 % B, 15.50-17.00 35% B, 17.00-18.50 45% B, 18.50-20.00 60 % B, 20.00-27.00 100% B. Column equilibration to starting conditions was performed after gradient elution for 3 min. The sample injection volume was set to 5 µL. A jet stream ESI source running in negative ionization mode was used with the following settings: gas temperature: 320 °C; drying gas: 8 L/min; nebulizer gas: 35 psig; sheath gas temperature: 380 °C ; sheath gas flow: 11 L/min; VCap: 3000 V; nozzle voltage: 0 V; fragmentor: 100 V; skimmer: 40 V; OCT 1 RF Vpp: 750 V. Fullscan (100-600 m/z; 4 spectra/s) data dependent MS/MS (50-600 m/z; 4 spectra/s; collision energy: 22 eV) acquisition

was used for identification and relative quantification of the corresponding metabolites. Metabolites were identified after visual inspection and interpretation of the corresponding MS/MS spectra using MassHunter Workstation (Agilent Technologies, Inc.; Santa Clara, CA, USA) software Accelrys Draw Academic Version 4.1 (Accelrys, Inc.; San Diego, CA, USA). After that the corresponding metabolites were relatively quantified using full scan.

References

- (1) McIlroy, S. J.; Kristiansen, R.; Albertsen, M.; Karst, S. M.; Rossetti, S.; Nielsen, J. L.; Tandoi, V.; Seviour, R. J.; Nielsen, P. H. Metabolic model for the filamentous “*Candidatus Microthrix parvicella*” based on genomic and metagenomic analyses. *ISME J* **2013**, 7 (6), 1161–1172.

Table S1. Summary of essential genes in *Shingobium* sp. BiD32

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>dnaB</i>	Replicative DNA helicase	N1ML46	EBBID32_4900	Replicative DNA helicase
<i>dnaE</i>	DNA polymerase III, alpha subunit	N1MTG1	EBBID32_46350	DNA-directed DNA polymerase (EC 2.7.7.7)
<i>dnaG</i>	DNA primase	N1MQY3	EBBID32_37220	DNA primase (EC 2.7.7.-)
<i>dnaN</i>	DNA polymerase III, beta subunit	N1MMD2	EBBID32_27170	DNA polymerase III subunit beta (EC 2.7.7.7)
<i>dnaQ</i>	DNA polymerase III, epsilon subunit	N1MY79	EBBID32_45620	DNA polymerase III epsilon subunit (EC 2.7.7.7)
<i>dnaX</i>	DNA polymerase III, gamma and tau subunits	N1MN77	EBBID32_30090	DNA polymerase III subunits gamma and tau (EC 2.7.7.7)
<i>gyrA</i>	DNA gyrase, A subunit	N1MF61	EBBID32_1980	DNA gyrase subunit A (EC 5.99.1.3)
<i>gyrB</i>	DNA gyrase, B subunit	N1MX63	EBBID32_42330	DNA gyrase subunit B (EC 5.99.1.3)
<i>holA</i>	DNA polymerase III, delta subunit	N1MHJ1	EBBID32_10130	DNA polymerase III delta subunit
<i>holB</i>	DNA polymerase III, delta prime subunit	N1MG02	EBBID32_2260	DNA polymerase III delta prime subunit (EC 2.7.7.7)
<i>hupA</i>	DNA binding protein	-	-	-
<i>lig</i>	DNA ligase (NAD dependent)	N1MTF7	EBBID32_44150	DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD(+)])
<i>ssb</i>	SSB	N1MS91	EBBID32_42110	Single-stranded DNA-binding protein (SSB)
		N1MI11	EBBID32_11980	Single-stranded DNA-binding protein (SSB)
		N1MKF9	EBBID32_2740	Single-stranded DNA-binding protein
		N1MNZ8	EBBID32_16910	Single-stranded DNA-binding protein
		N1MQJ0	EBBID32_35470	Single-stranded DNA-binding protein
<i>nth</i>	Endonuclease III	N1MQ94	EBBID32_34580	Endonuclease III (EC 4.2.99.18) (DNA-(apurinic or apyrimidinic site) lyase)
<i>polA</i>	5'-3' exonuclease domain of DNA polymerase I	N1MIR1	EBBID32_9920	DNA polymerase I (EC 2.7.7.7)
<i>ung</i>	Uracil-DNA glycosylase	N1MPC7	EBBID32_18480	Uracil-DNA glycosylase (UDG) (EC 3.2.2.27)
<i>deaD</i>	ATP-dependent RNA helicase	N1MWC1	EBBID32_42370	ATP-dependent helicase, DEAD/DEAH box family, associated with Flp pilus assembly
<i>greA</i>	Transcription elongation factor	N1MRI7	EBBID32_37150	Transcription elongation factor GreA (Transcript cleavage factor GreA)
<i>nusA</i>	Transcription-translation coupling	N1MHH7	EBBID32_10030	Transcription termination/antitermination protein NusA
.	.	.	.	.

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>nusG</i>	Transcription antitermination protein	N1MYL4	EBBID32_46920	Transcription termination/antitermination protein NusG
<i>rpoA</i>	RNA polymerase, alpha subunit	N1MTL7	EBBID32_44800	DNA-directed RNA polymerase subunit alpha (RNAP subunit alpha) (EC 2.7.7.6) (RNA polymerase subunit alpha) (Transcriptase subunit alpha)
<i>rpoB</i>	RNA polymerase, beta subunit	N1MGG9	EBBID32_6430	DNA-directed RNA polymerase subunit beta (RNAP subunit beta) (EC 2.7.7.6) (RNA polymerase subunit beta) (Transcriptase subunit beta)
<i>rpoC</i>	RNA polymerase, beta prime subunit	N1MH76	EBBID32_6440	DNA-directed RNA polymerase subunit beta' (RNAP subunit beta') (EC 2.7.7.6) (RNA polymerase subunit beta') (Transcriptase subunit beta')
<i>rpoD</i>	RNA polymerase, major sigma subunit	N1MUQ0	EBBID32_37230	RNA polymerase sigma factor RpoD (Sigma-70)
<i>alaS</i>	Alanyl-tRNA synthase	N1MF76	EBBID32_2150	Alanine--tRNA ligase (EC 6.1.1.7) (Alanyl-tRNA synthetase)
<i>argS</i>	Arginyl-tRNA synthase	N1ML25	EBBID32_22890	Arginine--tRNA ligase (EC 6.1.1.19) (Arginyl-tRNA synthetase)
<i>asnS</i>	Asparaginyl-tRNA synthase	-	-	-
<i>aspS</i>	Aspartyl-tRNA synthase	N1MMZ3	EBBID32_24440	Aspartyl-tRNA synthetase @ Aspartyl-tRNA(Asn) synthetase (EC 6.1.1.12) (EC 6.1.1.23)
<i>cysS</i>	Cysteinyl-tRNA synthase	N1MI93	EBBID32_9890	Cysteine--tRNA ligase (EC 6.1.1.16) (Cysteinyl-tRNA synthetase)
<i>glnS</i>	Glutaminyl-tRNA synthase	-	-	-
<i>gltX</i>	Glutamyl-tRNA synthase	N1MMK7	EBBID32_11700	Glutamate--tRNA ligase (EC 6.1.1.17) (Glutamyl-tRNA synthetase)
		N1MQL4	EBBID32_35670	Glutamate--tRNA ligase (EC 6.1.1.17) (Glutamyl-tRNA synthetase)
<i>glyS</i>	Glycyl-tRNA synthase, b subunit	-	-	-
<i>hisS</i>	Histidyl-tRNA synthase	N1MRV5	EBBID32_40350	Histidine--tRNA ligase (EC 6.1.1.21) (Histidyl-tRNA synthetase)
<i>ileS</i>	Isoleucyl-tRNA synthase	N1MRC6	EBBID32_38780	Isoleucine--tRNA ligase (EC 6.1.1.5) (Isoleucyl-tRNA synthetase)
<i>leuS</i>	Leucyl-tRNA synthase	-	-	-
<i>lysS</i>	Lysyl-tRNA synthase	N1MMS3	EBBID32_28730	Lysine--tRNA ligase (EC 6.1.1.6) (Lysyl-tRNA synthetase)
<i>metS</i>	Methionyl-tRNA synthase	N1MF87	EBBID32_2250	Methionine--tRNA ligase (EC 6.1.1.10) (Methionyl-tRNA synthetase)
.	.	.	.	.

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>pheS</i>	Phenylalanyl-tRNA synthase, a subunit	N1MMW9	EBBID32_11610	Phenylalanine--tRNA ligase alpha subunit (EC 6.1.1.20) (Phenylalanyl-tRNA synthetase alpha subunit)
<i>pheT</i>	Phenylalanyl-tRNA synthase, b subunit	N1MMK2	EBBID32_11600	Phenylalanine--tRNA ligase beta subunit
<i>proS</i>	Prolyl-tRNA synthase	N1MVW1	EBBID32_40560	Proline--tRNA ligase (EC 6.1.1.15) (Prolyl-tRNA synthetase)
<i>serS</i>	Seryl-tRNA synthase	N1MWF7	EBBID32_42780	Serine--tRNA ligase (EC 6.1.1.11) (Seryl-tRNA synthetase) (Seryl-tRNA(Ser/Sec) synthetase)
<i>thrS</i>	Threonyl-tRNA synthase	N1MQ37	EBBID32_20910	Threonine--tRNA ligase (EC 6.1.1.3) (Threonyl-tRNA synthetase)
<i>trpS</i>	Tryptophanyl-tRNA synthase	N1MLA8	EBBID32_20750	Tryptophanyl-tRNA synthetase (EC 6.1.1.2)
<i>tyrS</i>	Tyrosyl-tRNA synthase	N1MNH0	EBBID32_31200	Tyrosine--tRNA ligase (EC 6.1.1.1) (Tyrosyl-tRNA synthetase)
<i>valS</i>	Valyl-tRNA synthase	-	-	-
<i>iscS</i>	Cysteine desulfurase-NifS homolog	-	-	-
<i>mnmA</i>	tRNA (5-methylaminomethyl-2-thiouridylate) methyl-transferase	N1MJD7	EBBID32_12170	tRNA-specific 2-thiouridylase MnmA (EC 2.8.1.-)
<i>mnmE</i>	GTP binding protein involved in biosynthesis of 5-methylaminomethyl-2-thiouridine	-	-	-
<i>mnmG</i>	Glucose-inhibited division protein A, involved in biosynthesis of 5-methylaminomethyl-2-thiouridine	N1MWE3	EBBID32_42620	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG (Glucose-inhibited division protein A)
<i>pth</i>	Peptidyl-tRNA hydrolase	-	-	-
<i>rnpA</i>	Protein component of Rna P	N1MS04	EBBID32_28250	Ribonuclease P protein component (RNase P protein) (RNaseP protein) (EC 3.1.26.5) (Protein C5)
<i>rplA</i>	50S ribosomal protein L1	N1MTK2	EBBID32_46900	50S ribosomal protein L1
<i>rplB</i>	50S ribosomal protein L2	N1MU12	EBBID32_32250	50S ribosomal protein L2
<i>rplC</i>	50S ribosomal protein L3	N1MNQ3	EBBID32_32220	50S ribosomal protein L3
<i>rplD</i>	50S ribosomal protein L4	N1MPK6	EBBID32_32230	50S ribosomal protein L4
<i>rplE</i>	50S ribosomal protein L5	N1MT72	EBBID32_32340	50S ribosomal protein L5
<i>rplF</i>	50S ribosomal protein L6	N1MNR8	EBBID32_32370	50S ribosomal protein L6
<i>rplI</i>	50S ribosomal protein L9	N1MH13	EBBID32_5530	50S ribosomal protein L9
<i>rplJ</i>	50S ribosomal protein L10	N1MM45	EBBID32_21480	50S ribosomal protein L10
<i>rplK</i>	50S ribosomal protein L11	N1MXT1	EBBID32_46910	50S ribosomal protein L11
<i>rplL</i>	50S ribosomal protein L12	N1MQV1	EBBID32_21470	50S ribosomal protein L7/L12

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>rplM</i>	50S ribosomal protein L13	N1MKZ8	EBBID32_6050	50S ribosomal protein L13
<i>rplN</i>	50S ribosomal protein L14	N1MNR3	EBBID32_32320	50S ribosomal protein L14
<i>rplO</i>	50S ribosomal protein L15	N1MQ71	EBBID32_32410	50S ribosomal protein L15
<i>rplP</i>	50S ribosomal protein L16	N1MT67	EBBID32_32290	50S ribosomal protein L16
<i>rplQ</i>	50S ribosomal protein L17	N1MSE1	EBBID32_44810	50S ribosomal protein L17
<i>rplR</i>	50S ribosomal protein L18	N1MPM0	EBBID32_32380	50S ribosomal protein L18
<i>rplS</i>	50S ribosomal protein L19	-	-	-
<i>rplT</i>	50S ribosomal protein L20	N1MHX8	EBBID32_11630	50S ribosomal protein L20
<i>rplU</i>	50S ribosomal protein L21	N1MXS9	EBBID32_46860	50S ribosomal protein L21
<i>rplV</i>	50S ribosomal protein L22	N1MNQ6	EBBID32_32270	50S ribosomal protein L22
<i>rplW</i>	50S ribosomal protein L23	N1MT64	EBBID32_32240	50S ribosomal protein L23
<i>rplX</i>	50S ribosomal protein L24	N1MPL6	EBBID32_32330	50S ribosomal protein L24
<i>rpmA</i>	50S ribosomal protein L27	N1MTJ8	EBBID32_46850	50S ribosomal protein L27
<i>rpmB</i>	50S ribosomal protein L28	-	-	-
<i>rpmC</i>	50S ribosomal protein L29	N1MU16	EBBID32_32300	50S ribosomal protein L29
<i>rpmE</i>	50S ribosomal protein L31	N1MK28	EBBID32_19160	50S ribosomal protein L31
<i>rpmF</i>	50S ribosomal protein L32	N1MR51	EBBID32_24830	50S ribosomal protein L32
<i>rpmG</i>	50S ribosomal protein L33	N1MNB6	EBBID32_28140	50S ribosomal protein L33
<i>rpmH</i>	50S ribosomal protein L34	N1MND0	EBBID32_28240	50S ribosomal protein L34
<i>rpmI</i>	50S ribosomal protein L35	N1MIP7	EBBID32_11640	50S ribosomal protein L35
<i>rpmJ</i>	50S ribosomal protein L36	N1MVY0	EBBID32_40770	50S ribosomal protein L36
<i>rpsB</i>	30S ribosomal protein S2	N1MLE7	EBBID32_19050	30S ribosomal protein S2
<i>rpsC</i>	30S ribosomal protein S3	N1MPL1	EBBID32_32280	30S ribosomal protein S3
<i>rpsD</i>	30S ribosomal protein S4	N1MS81	EBBID32_39520	30S ribosomal protein S4
<i>rpsE</i>	30S ribosomal protein S5	N1MT75	EBBID32_32390	30S ribosomal protein S5
<i>rpsF</i>	30S ribosomal protein S6	N1MLB9	EBBID32_5550	30S ribosomal protein S6
<i>rpsG</i>	30S ribosomal protein S7	N1MPK3	EBBID32_32180	30S ribosomal protein S7
<i>rpsH</i>	30S ribosomal protein S8	N1MQ66	EBBID32_32360	30S ribosomal protein S8
<i>rpsI</i>	30S ribosomal protein S9	N1MH47	EBBID32_6040	30S ribosomal protein S9
<i>rpsJ</i>	30S ribosomal protein S10	N1MQ50	EBBID32_32210	30S ribosomal protein S10
<i>rpsK</i>	30S ribosomal protein S11	N1MXW9	EBBID32_44790	30S ribosomal protein S11
<i>rpsL</i>	30S ribosomal protein S12	N1MNP8	EBBID32_32170	30S ribosomal protein S12
<i>rpsM</i>	30S ribosomal protein S13	N1MX42	EBBID32_44780	30S ribosomal protein S13
<i>rpsN</i>	30S ribosomal protein S14	N1MU22	EBBID32_32350	30S ribosomal protein S14
<i>rpsO</i>	30S ribosomal protein S15	N1MWU7	EBBID32_44080	30S ribosomal protein S15
<i>rpsP</i>	30S ribosomal protein S16	-	-	-
<i>rpsQ</i>	30S ribosomal protein S17	N1MQ61	EBBID32_32310	30S ribosomal protein S17
<i>rpsR</i>	30S ribosomal protein S18	N1MKW0	EBBID32_5540	30S ribosomal protein S18

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>rpsS</i>	30S ribosomal protein S19	N1MQ55	EBBID32_32260	30S ribosomal protein S19
<i>rpsT</i>	30S ribosomal protein S20	N1MLB3	EBBID32_20800	30S ribosomal protein S20
<i>cspR</i>	Ribosomal methyltransferase	-	-	-
<i>engA</i>	GTP binding protein	N1ML93	EBBID32_20540	GTP-binding protein EngA
<i>era</i>	GTP binding protein	N1MPH8	EBBID32_29830	GTPase Era
<i>ksgA</i>	Dimethyladenosine transferase	N1MNJ0	EBBID32_15010	Ribosomal RNA small subunit methyltransferase A (EC 2.1.1.182) (16S rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase) (16S rRNA dimethyladenosine transferase) (16S rRNA dimethylase) (S-adenosylmethionine-6-N', N'-adenosyl(rRNA) dimethyltransferase)
<i>obg</i>	GTP binding protein	N1MH61	EBBID32_9070	GTPase Obg (GTP-binding protein Obg)
<i>rbfA</i>	Ribosome binding factor A	N1MIB5	EBBID32_10090	Ribosome-binding factor A
<i>ychF</i>	GTP binding protein	-	-	-
<i>efp</i>	Elongation factor P	N1MSN2	EBBID32_30260	Elongation factor P (EF-P)
<i>fusA</i>	Elongation factor G	N1MT59	EBBID32_32190	Elongation factor G (EF-G)
<i>frr</i>	Ribosome-recycling factor	N1MPI6	EBBID32_19080	Ribosome-recycling factor (RRF) (Ribosome-releasing factor)
<i>hemK</i>	N5-glutamine methyltransferase, modulation of release factor activity	-	-	-
<i>infA</i>	Initiation factor IF-1	N1MLI5	EBBID32_21700	Translation initiation factor 1
<i>infB</i>	Initiation factor IF-2	N1MMH1	EBBID32_10060	Translation initiation factor IF-2
<i>infC</i>	Initiation factor IF-3	N1MQQ8	EBBID32_20920	Translation initiation factor IF-3
<i>lepA</i>	GTP binding elongation factor	N1MPM8	EBBID32_32480	Elongation factor 4 (EF-4) (EC 3.6.5.n1) (Ribosomal back-translocase LepA)
<i>prfA</i>	Peptide chain release factor 1 (RF1)	N1ML74	EBBID32_7030	Peptide chain release factor 1 (RF-1)
<i>smpB</i>	tmRNA binding protein	N1MT03	EBBID32_28860	SsrA-binding protein
<i>tsf</i>	Elongation factor Ts	N1MK19	EBBID32_19060	Elongation factor Ts (EF-Ts)
<i>tufA</i>	Elongation factor Tu	N1MU06	EBBID32_32200	Elongation factor Tu (EF-Tu)
<i>pnp</i>	Polyribonucleotide nucleotidyltransferase	N1MTF1	EBBID32_44100	Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) (Polynucleotide phosphorylase)
<i>rnc</i>	Ribonuclease III	N1MSI7	EBBID32_29860	Ribonuclease 3 (EC 3.1.26.3) (Ribonuclease 3) (Ribonuclease III)
<i>map</i>	Methionine aminopeptidase	N1MNN0	EBBID32_14170	Methionine aminopeptidase (MAP) (MetAP) (EC 3.4.11.18) (Peptidase M)
.	.	.	.	.

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>pepA</i>	Aminopeptidase A/I	N1MJP9	EBBID32_15050	Probable cytosol aminopeptidase (Leucine aminopeptidase) (Leucyl aminopeptidase)
<i>dnaJ</i>	Hsp70 cochaperone	N1MUJ9	EBBID32_36680	Chaperone protein DnaJ
<i>dnaK</i>	Chaperone Hsp70	N1MKR0	EBBID32_21690	Chaperone protein DnaK (HSP70) (Heat shock 70 kDa protein) (Heat shock protein 70)
<i>groEL</i>	Class I heat shock protein	N1MXP7	EBBID32_44190	60 kDa chaperonin (GroEL protein) (Protein Cpn60)
<i>groES</i>	Class I heat shock protein	N1MWW0	EBBID32_44180	10 kDa chaperonin (GroES protein) (Protein Cpn10)
<i>grpE</i>	Hsp70 cochaperone	N1MSJ2	EBBID32_27410	Protein GrpE (HSP-70 cofactor)
<i>ffh</i>	Protein component of signal recognition particle	-	-	-
<i>ftsY</i>	Signal recognition particle receptor	-	-	-
<i>secA</i>	Preprotein translocase subunit (ATPase)	N1MJR3	EBBID32_2320	Protein translocase subunit SecA
		N1MLZ7	EBBID32_25510	Protein translocase subunit SecA
<i>secE</i>	Membrane-embedded preprotein translocase subunit	N1MU79	EBBID32_46930	Protein translocase subunit SecE
<i>secY</i>	Membrane-embedded preprotein translocase subunit	N1MNS2	EBBID32_32420	Protein translocase subunit SecY
<i>gcp</i>	Probable O-sialoglycoprotein endopeptidase	-	-	-
<i>hflB</i>	ATP-dependent protease	-	-	-
<i>lon</i>	ATP-dependent protease La	N1MSL6	EBBID32_30110	Lon protease (EC 3.4.21.53) (ATP-dependent protease La)
<i>ftsZ</i>	Cytoskeletal cell division protein	N1MN60	EBBID32_29940	Cell division protein FtsZ
<i>pitA</i>	Low-affinity inorganic phosphate transporter	N1MKL0	EBBID32_18270	Probable low-affinity inorganic phosphate transporter
<i>ptsG</i>	PTS glucose-specific enzyme II	N1MPW1	EBBID32_36160	PTS system, N-acetylglucosamine-specific IIA component / PTS system, N-acetylglucosamine-specific IIB component / PTS system, N-acetylglucosamine-specific IIC component (EC 2.7.1.69)
<i>ptsH</i>	Histidine-containing phosphocarrier protein of PTS	-	-	-
<i>ptsI</i>	PTS enzyme I	-	-	-
<i>eno</i>	Enolase	N1MPM3	EBBID32_17470	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase)

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>fbaA</i>	Fructose-1,6-bisphosphate aldolase	N1MFY7	EBBID32_2110	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)
<i>gapA</i>	Glyceraldehyde-3-phosphate dehydrogenase	N1MGI2	EBBID32_2090	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-)
<i>gpmA</i>	Phosphoglycerate mutase	N1MKX3	EBBID32_4340	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (BPG-dependent PGAM) (PGAM) (Phosphoglyceromutase) (dPGM) (EC 5.4.2.11)
<i>ldh</i>	L-lactate dehydrogenase	N1MX13	EBBID32_41730	Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit YkgF
		N1MST7	EBBID32_41740	Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE
<i>pfkA</i>	6-Phosphofructokinase	N1MJW0	EBBID32_13780	Phosphofructokinase
<i>pgi</i>	Glucose-6-phosphate isomerase	N1MN52	EBBID32_29890	Glucose-6-phosphate isomerase (GPI) (EC 5.3.1.9) (Phosphoglucose isomerase) (Phosphohexose isomerase)
<i>pgk</i>	Phosphoglycerate kinase	N1MF71	EBBID32_2100	Phosphoglycerate kinase (EC 2.7.2.3)
<i>pykA</i>	Pyruvate kinase	N1MKP8	EBBID32_21540	Pyruvate kinase (EC 2.7.1.40)
<i>tpiA</i>	Triose-phosphate isomerase	N1MPN0	EBBID32_19530	Triosephosphate isomerase (TIM) (EC 5.3.1.1) (Triose-phosphate isomerase)
<i>atpA</i>	ATP synthase alpha chain	N1MH59	EBBID32_4450	ATP synthase subunit alpha (EC 3.6.3.14) (ATP synthase F1 sector subunit alpha) (F-ATPase subunit alpha)
<i>atpB</i>	ATP synthase A chain	N1MU63	EBBID32_35320	ATP synthase subunit a (ATP synthase F0 sector subunit a) (F-ATPase subunit 6)
<i>atpC</i>	ATP synthase epsilon chain	N1MHA6	EBBID32_9420	ATP synthase epsilon chain (ATP synthase F1 sector epsilon subunit) (F-ATPase epsilon subunit)
<i>atpD</i>	ATP synthase beta chain	N1MGP1	EBBID32_4470	ATP synthase subunit beta (EC 3.6.3.14) (ATP synthase F1 sector subunit beta) (F-ATPase subunit beta)
<i>atpE</i>	ATP synthase C chain	N1MQH3	EBBID32_35310	ATP synthase C chain (EC 3.6.3.14)
<i>atpF</i>	ATP synthase B chain	N1MPN2	EBBID32_35300	ATP synthase subunit b (ATP synthase F(0) sector subunit b) (ATPase subunit I) (F-type ATPase subunit b)
		N1MR14	EBBID32_35290	ATP synthase subunit b (ATP synthase F(0) sector subunit b) (ATPase subunit I) (F-type ATPase subunit b)
.	.	.	.	.

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>atpG</i>	ATP synthase gamma chain	N1MFW9	EBBID32_4460	ATP synthase gamma chain (ATP synthase F1 sector gamma subunit) (F-ATPase gamma subunit)
<i>atpH</i>	ATP synthase delta chain	N1MKY8	EBBID32_4440	ATP synthase subunit delta (ATP synthase F(1) sector subunit delta) (F-type ATPase subunit delta)
<i>yidC</i>	Essential for proper integration of ATPase into the membrane	N1MP13	EBBID32_28270	Membrane protein insertase YidC (Foldase YidC) (Membrane integrase YidC) (Membrane protein YidC)
<i>rpe</i>	Ribulose-phosphate 3-epimerase	N1MJA7	EBBID32_11920	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
<i>rpiA</i>	Ribose 5-phosphate isomerase	N1MNY2	EBBID32_30000	Ribose 5-phosphate isomerase B (EC 5.3.1.6)
<i>tkt</i>	Transketolase	N1MIJ1	EBBID32_13890	Transketolase (EC 2.2.1.1)
		N1MKB0	EBBID32_2080	Transketolase (EC 2.2.1.1)
<i>cdsA</i>	Phosphatidate cytidylyltransferase	N1MLF5	EBBID32_19100	Phosphatidate cytidylyltransferase (EC 2.7.7.41)
<i>fadD</i>	Acyl-CoA synthase	-	-	-
<i>gpsA</i>	<i>sn</i> -Glycerol-3-phosphate dehydrogenase	N1MMY4	EBBID32_11760	Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94) (NAD(P)H-dependent glycerol-3-phosphate dehydrogenase)
<i>plsB</i>	<i>sn</i> -Glycerol-3-phosphate acyltransferase	-	-	-
<i>plsC</i>	1-Acyl- <i>sn</i> -glycerol-3-phosphate acyltransferase	N1MQV6	EBBID32_34760	1-acyl- <i>sn</i> -glycerol-3-phosphate acyltransferase (EC 2.3.1.51)
		N1MLM1	EBBID32_8480	1-acyl- <i>sn</i> -glycerol-3-phosphate acyltransferase (EC 2.3.1.51)
<i>psd</i>	Phosphatidylserine decarboxylase	N1MJA8	EBBID32_550	Phosphatidylserine decarboxylase proenzyme (EC 4.1.1.65)
		N1MK16	EBBID32_19010	Phosphatidylserine decarboxylase proenzyme (EC 4.1.1.65)
<i>pssA</i>	Phosphatidylserine synthase	-	-	-
<i>adk</i>	Adenylate kinase	N1MPM4	EBBID32_32430	Adenylate kinase (AK) (EC 2.7.4.3) (ATP-AMP transphosphorylase) (ATP:AMP phosphotransferase) (Adenylate monophosphate kinase)
<i>dcd</i>	dCTP deaminase	N1MKN3	EBBID32_4940	Deoxycytidine triphosphate deaminase (dCTP deaminase) (EC 3.5.4.13)
<i>gmK</i>	Guanylate kinase	N1MQ35	EBBID32_32060	Guanylate kinase (EC 2.7.4.8) (GMP kinase)
<i>hpt</i>	Hypoxanthine phosphoribosyltransferase	-	-	-
<i>ndk</i>	Nucleoside diphosphate kinase	N1MP03	EBBID32_15070	Nucleoside diphosphate kinase (NDK) (NDP kinase) (EC 2.7.4.6) (Nucleoside-2-P kinase)
.	.	.	.	.

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>nrdE</i>	Ribonucleoside diphosphate reductase (major subunit)	N1MEU4	EBBID32_480	Ribonucleoside-diphosphate reductase (EC 1.17.4.1)
<i>nrdF</i>	Ribonucleoside diphosphate reductase (minor subunit)	N1MK05	EBBID32_460	Ribonucleoside-diphosphate reductase subunit beta (EC 1.17.4.1)
<i>ppa</i>	Inorganic pyrophosphatase	N1MSH1	EBBID32_40330	Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-hydrolase)
<i>prsA</i>	Phosphoribosylpyrophosphate synthase	N1MHY2	EBBID32_11680	Ribose-phosphate pyrophosphokinase (RPPK) (EC 2.7.6.1) (5-phospho-D-ribosyl alpha-1-diphosphate) (Phosphoribosyl diphosphate synthase) (Phosphoribosyl pyrophosphate synthase)
<i>pyrG</i>	CTP synthase	N1MM39	EBBID32_10000	CTP synthase (EC 6.3.4.2) (CTP synthetase) (UTP--ammonia ligase)
<i>thyA</i>	Thymidylate synthase	N1MQ23	EBBID32_33950	Thymidylate synthase (EC 2.1.1.45)
<i>tmk</i>	Thymidylate kinase	N1MJR1	EBBID32_2270	Thymidylate kinase (EC 2.7.4.9) (dTMP kinase)
<i>trxA</i>	Thioredoxin	N1MF95	EBBID32_2350	Thioredoxin
		N1MTF3	EBBID32_46250	Thioredoxin
<i>trxB</i>	Thioredoxin reductase	N1MIV3	EBBID32_14840	Thioredoxin reductase (EC 1.8.1.9)
<i>upp</i>	Uracil phosphoribosyltransferase	-	-	-
<i>coaA</i>	Pantothenate kinase	-	-	-
<i>coaD</i>	4'-Phosphopantetheine adenylyltransferase	N1MPL5	EBBID32_19380	Phosphopantetheine adenylyltransferase (EC 2.7.7.3) (Dephospho-CoA pyrophosphorylase) (Pantetheine-phosphate adenylyltransferase)
<i>coaE</i>	Dephospho-CoA kinase	N1MTX4	EBBID32_45630	Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase)
<i>dfp</i>	Phosphopantothenate cysteine ligase and 4'-Phospho-pantothenyl-L-cysteine decarboxylase	N1MQ34	EBBID32_20860	Phosphopantothenoylcysteine decarboxylase / Phosphopantothenoylcysteine synthetase (EC 4.1.1.36) (EC 6.3.2.5)
<i>folA</i>	Dihydrofolate reductase	N1MRZ6	EBBID32_38760	Dihydrofolate reductase (EC 1.5.1.3)
<i>glyA</i>	Glycine hydroxymethyltransferase	N1MN65	EBBID32_29990	Serine hydroxymethyltransferase (SHMT) (Serine methylase) (EC 2.1.2.1)
<i>metK</i>	Methionine adenosyltransferase	N1MRL3	EBBID32_42250	S-adenosylmethionine synthase (EC 2.5.1.6)
<i>nadR</i>	Adenylyltransferase	-	-	-
<i>nadV</i>	Nicotinamide phosphoribosyltransferase	-	-	-
<i>pdxY</i>	Pyridoxal kinase	-	-	-
.	.	.	.	.

Table S1. (continued)

Genes from Gil et al. 2004			Genes from this study	
Gene	Protein function	UniProt accession number	Gene	UniProt protein name
<i>ribF</i>	Riboflavin kinase	N1MQK4	EBBID32_38770	Riboflavin biosynthesis protein (EC 2.7.1.26) (EC 2.7.7.2)
<i>yloS</i>	Thiamine pyrophosphokinase	-	-	-
<i>mesJ</i>	Conserved hypothetical protein	-	-	-
<i>mraW</i>	Methyltransferase	-	-	-
<i>ybeY</i>	Conserved hypothetical protein	N1MTZ0	EBBID32_34790	Endoribonuclease YbeY (EC 3.1.-.-)
<i>ycfF</i>	HIT family	-	-	-
<i>ycfH</i>	Putative deoxyribonuclease, <i>tatD</i> family	N1MGJ8	EBBID32_2240	Putative deoxyribonuclease YcfH
<i>yoaE</i>	Conserved hypothetical protein	-	-	-
<i>yqgF</i>	Conserved hypothetical protein	-	-	-
<i>yraL</i>	Conserved hypothetical protein	-	-	-

Table S2. DNA and protein sequences of *Sphingobium* sp. BiD32 genes and proteins of interest

Accession number	DNA sequence	Protein sequence
N1MWA7	TTGTTTCTGGCGGGCGATGCAGCGCATATCGTGCCTCCCACCGGCGCCAAGGGCTTAAATCTGGCA GCGTCGGATGTGATCATGCTTAGTCAAGCCCTGCGCGAATTTTATCGCGACCATTCTTCCGCAGGC ATCGACAGCTACTCTGAACGCGCACTGTCGAGGGTCTGGAAGGCGGAGCGCTTTTCGTGGTGGTTT ACGTCTATCACGCACCGCTTTCCCAACATGGATGGGTTCGATCGTCGCATCCAGATGGCGGAGCTG GATTACATCCGCGGCTCGATAACAGCGCAGCGATCGTTGGCCGAAAATTATGTCGGCCTACCTTTG GAGACGATATAA	MFLAGDAAHIVPPTGAKGLNL AASDVIMLSQALREFYRDHSSA GIDSYSERALSrvWKAERFSW WFTSITHRFPNMDGFDRIQMA ELDYIRGSITAQRSLAENYVGLP LETI
N1MT47	ATGAGCCTGATTATTGATTGCCATGGCCATTATACGGTGTGCCCCAAGGGGCATGATGCCTGGCGC GAGGAGCAGAAGGCGGCGTTCAAGGCCGGGACGCCGTGCCCGCCCTATCCCGATATTAGCGACGC AGAGATTCTGTGAGACGATCGAGGCGAACCAGTTGCGGCTGATCAAGGAGCGTGGCGCGGACCTGA CGATCTTTTCGCCTCGCGCATCGGCGATGGCACC GCATGTCGGCGATGAAGCGATGGCCAGGGAAT GGGCGATCCGCTGCAACGACCTGATCGCGCGGGTGGTGGCGATGTTCCCCGAAACCTTTGTCGGCG TATGCATGTTGCCGAGTCGCCCCAAGGCGGACATGCGCAATAGCATCGCCGAGTTGGAGCGCTGTG TCACTGAACTGGGTTTCATCGGCTGCAACCTCAATCCCGATCCGGGCGGCGGGCATTTCACCCACC CGCCGCTGACCGACCGCTATTGGTATCCCTTCTACGAGAAAATGGTCGAGCTGGACGTCCCGGCGA TGATCCATGTTTCGGGCAGCTGCAACCCGGCGATGCACGCGACCGGCGGCTATTATATCGCGGCGG ACACGATCGCTTTCATGCAGTTGCTGGAAGGCGACCTGTTCAAGGATTTCCCGACGCTGCGTTTCA TCATCCCGCATGGCGGCGGCGCGGTGCCCTATCATTGGGGGCGTTATCGCGGCTTGGCCGACATGC TCAAAAAGCCCGATCTGTGACGCATCTGATGAACAACATCTATTTTCGACACCTGCGTCTATCATC AGCCGGGCGTCGATCTGCTGGCCGATGTGATCGACAACAAGAATATCCTGTTTCGGGTTCGGAAATG GTCGGCGCGGTGCGCGGGATCGACCCACCACCGGGCAATATTTTCGATGACACCAAGCGCTATGTC GATGCGCTGGACATTAGCGCGGCCGAGCGCGCCGCCATTTTCGAGGGCAATGCCCCGCCGCTCTTT CCCCGCTGGACGCGCAATTAAAGGAGAGGGGACTATGA	MSLIIDCHGHYTVLPKGHDAW REEQKAAFKAGTPCPPYDISD AEIRETIEANQLRLIKERGADLTI FSRASAMAPHVGDAMAREW AIRCNDLIARVVAMFPETFVGV CMLPQSPKADMRNSIAELERC TELGFICNLNPDPGGGHFTHP LTDYRWYPFYEKMVELDVPAM IHVSGSCNPAMHATGGYYIAAD TIAFMQLLEGDLFKDFPTLRFIIP HGGGAVPYHWGRYRGLADML KKPDLSTHLMNNIYFDTCVYHQ PGVDLLADVIDNKNILFGSEMV GAVRGIDPTTGQYFDDTKRYV DALDISAAERAAIFEGNARRVFP RLDAQLKERGL
N1MSB0	ATGATGACAGAAGATCCCAAGCATCAGGATATTCACGAATATCTGGCTGAGTTCGAGGATATTCCC GGCACCCGCGTCTATACGGCGGCACGCGCGCGCAAGGGCTATCATATGAACCAAGTTTTCGATGAG CCTGATGAAGGACGAAAATCGCACCCGGTTCAAGGCCGACGAGCGCGCCTATATGGACGAATGGC CGTTGACCGAGGACCAGAAGCTGGCGTTGCTGGCGCGCGACTATAATCGGTGCCTGGACCTGGGC GGCAATGTCTATTTCCCTGTCCAAGCTGTTTTGACCGACGGGATTCCGTTCCCGAGGCGGTCAGC ACCATGACCGGCATGAGCTTTCCCGAATATCGGGAAATGATGATGAATGGTGGCCGTTACCCGAA GGCCAGCGGTTCGATCAGGGAGAAAAGATGA	MMTEDPKHQDIHEYLAEFEDIP GTRVYTAARARKGYHMNQFA MSLMKDENRTRFKADERAYM DEWPLTEDQKLALLARDYNRC LDLGGNVYFLSKLFSTDGIPFAE AVSTMTGMSFPEYREMMNG GRSPEGQRSIREKR

.

Table S2. (continued)

Accession number	DNA sequence	Protein sequence
N1MLS9	ATGCCGAAACTGATCGTGGTCAACCGTGCGGGCGAAGAACAGGCTGTGCGATGGCGACAATGGCCTGTCGGTGATGGAAGTCATCCGGGACAATGGTTTTGACGAACTGCTGGCCTTGTGCGGCGGCTGCTGCTCCTGCGCCACCTGCCATGTCTATATCGACCCGGCTTTCGCCGATTGCTTCCCGCGATGACCGAGGACGAAAACGACCTGCTCGACAGCTCCGACCATCGCAACGACACCAGCCGCTGTCTTGCCAGGTTGTGCTGAGCGACGCGCTGGACGGCCTGCGGGTCAACCATCGCACCCGAAGATTGA	MPKLIVNVRAGEEQAVDGDNGLSVMEVIRDNGFDELLALCGGCSCATCHVYIDPAFADSLPAMTEDENDLLDSSDHRNDTSRLSCQVVLSDALDGLRVTIAPED
N1MMX0	ATGCGTATTGCCCTGGCTGGTGCAGGCGCCTTTGGCGAAAAGCATCTCGACGGCCTGAAGAATATCGATGGCGTGAGCGTGACGTCGCTGGTTGGCCGACGGCTGGATGCGACGACGAGCGGATTGCCGACAAATATGGCATTGGTCATGCGACCACCGAGTTGGCGGACACGCTGGCGCGCGATGATGTTGATGCGGTGATCCTGTGCACCCCGACGCAGATGCACGCAGCGCAGGCGATTGCTGTCATGGAGGCGGGCAAGCATGTTTCAGGTGGAGATTCCGCTGGCCGATAGCTGGGCGGACGCGCAGGCGGTGCTGGCCAAGCAACAGGAAACCGGCCTGACCTGCATGGTCGGGCATACGCGGCGGTTCAACCCAGCCATCAATATGTCATAATCTGATCACGGCAGGTGATCTGAACATCCAGCAGATGGATGTGCAGACCTATTTCTTCCGCGCAAGAATATCAATGCCAAGGGCGAAGCGCGGTCTGGACCGACCATCTGCTGTGGCACCATGCGCGCATACGGTCGATCTGTTTCGCCTATCAGGCGGGCAGGATCGTGTGCGCGCACGCGATGACGAGGGCCGATCCACCCGACCTGGGCATTGCGATGGACATGTCTGATCCAGTTGAAGAGCGAGAGCGGGGCGATCTGCACCCTGTCGCTGTCAATTAACAATGACGGGCGCTGGGCACCTTCTTCCGCTATATTGGCGACACCGGCACCTATATCGCGCGCTACGACGATCTGGTGAACGGCAAGGAAGAGCCGATCGACGTGTCCAAGGTGGCAGTGTGATGAACGGCATCGAATTGCAGGACCGCGAATTCATCAGCGCGATCCGCGAAGGCCGCGAACCCTCGAGCGTCGCAAAGGTGTTGATTGCTATCGTGTGCTGGGCGAGCTGGAGCAGCAGCTGGCGTAA	MRIALAGAGAFGEKHLDDLKNI DGVSVTSLVGRRLDATQAIADK YGIGHATTELADTLARDDVDA VILCTPTQMHAQAIAACMEAG KHVQVEIPLADSWADAQAVLA KQETGLTCMVGHTRRFNPSH QYVHNLITAGDLNIQQMDVQT YFFRRKNINAKGEARSWTDHLL WHHAAHTVDLFAYQAGRIVSA HAMQGPPIHPDLGIAMDMSIQLK SESAGICTLSLSFNNDGPLGTFF RYIGDTGTIYIARYDDLVDNGKEE PIDVSKVAVSMNGIELQDREFIS AIREGREPNSSVAKVFDCYRVL GELEQQLA
N1MNM9	ATGATGGCGAGGATTACTGCTGGCGTGGCGTCGAGCCATGTGCCGCTGCTGGGCGTGGCGTCGGACTTCAACAAGGATCGTGACGAATATTTGCGGGCCGATCTTTGCAGGGTATGACTGGACGCGCGAATGGAGAAGGCGGAAAAGCCCGATGTCATCATTTCTGGTCTATAATGACCATGCGTCCGCCTTTGACATGAAGATCATCCCGACCTTCGCGATCGGTTGCGGCGAGCGCTACAAATCGGCGGACGAAGGCTGGGGCCGCGCGCGGTGCCGGACGTGGAGGGGCATCCCGACCTGGCCTGGCATATTGCGCAGAGCCTGGTGCTAGACGAGTTCGACATGACCATCATCAACGAGATGGACGTCGACCATGGGCTGACCGTGCCGCTCAGCATGATGTTTCGGCAAGCCCGACGCTGGCCGTGCAAGGTCTGTGCCGCTGGCCGTCAATGTCGTACCTATCCCCCGCCGTGCGGCAATCGCTGCTGGGCGCTGGGCGAAGCGATCAAGAAGGCGGTGCAAGGCTTTCCCCGAAGATCTGAACGTCCAGATCTGGGGCACCGGCGGCATGAGCCACCAAGTTGAGGGACCGCGCGCGGGGCTGATCAACAAGGAATGGGACAATATGTTTCTGGATGGGCTGGTCTGGTGAGAGCGACCATCTGCGCCGATTCCGCATATCGAATATCTGCGCGAGACGGGCAGCGAAGGCATTGAAATGGTGATGTGGCTGGTGATGCGCGGGGCGCTGGGGAAAAAGACCCGCGCGCTGCACCGCCATTATCATGTGCCCTGCTCCAACACCGCGATCGGCCATATGGTGCTGCGCCCGGATAATGGCGAAGGGCTGGATATGACCGGGAGCGATACGGTGACGGCGATCGCAGCGGAGTAA	MMARITAGVASSHVPLLGVAS DFNKDRDEYFGPIFAGYDWTRE WEKAEPDVIILVYNDHASAFD MKIIPTFAIGCGERYKSADEGW GPRAVPDVEGHPDLAWHIAQSL VLDEFDMTIINEMDVDHGLTVP LSMMFGKPDAPWPKVVPPLAVN VVTYPPPSGNRCWALGEAIKKA VESFPEDLNVQIWGTGGMSHQL QGPRAGLINKEDNMFLDGLV GESDHLRRIPHIEYLRETGSEGIE MVMWLVMRGALGKKTRALHR HYHVPSCNTAIGHMVLRPDNG EGLDMTGSDTVAIAAA

.

Table S2. (continued)

Accession number	DNA sequence	Protein sequence
N1MI62	ATGAGCGAGGAACAGGCGATGACCAAATTGGCGGAGCGCGACCTGCTCGCAACCGGATCGATGAC AGATCCGCTGATTTCGGCGCGGCCGCGTGACTTTTACGCCGCGATGCGCCGGGACGACCCGGTGCA TTGGGACGAGAAGCTGGGCATGTATCTGGTGTGCGGGTACGAGGATATCGCGACGATCCAGCAGG ACCCGATCACCTTTTCGGTGAACAAGGGCTATCACACGCAACAGGCCAAGGGGTTCAGCAGGAA TTTCAGGAGATATTGGAGCGTGAGGGCGGGCGGTTACTTTCCCGACGCGATCATGTCCGACCCGCC TATCACACACGTATCCGCAAGCTGATGGAGAAAAGCCTTTACCGCGCATCGCGTAAAGGCGCTGGA ACCGCGGATCGAAAAGGTGGTGGTCGACCTGGTCGACAGCGTCGCAGCGCGGGGCGAGGCGGACG CGATCCATGATTTTCGCGGTGCCGCTGACGGTGC GGATCATCTGCGAGCAGCTGGGCCTGGACTGGG ACATGAAGGACCGCATCCAGCGCTGGTCGATCGCGGTGACGGCGCAGATCGGGCGGATGCAGGAT CGCGAACAGATGCTGGGTACGCGCGGGAGATTTGCGACCTGCAACATTATCTGATCGCCAAGAT GCAGGAGCGGCAGGCCGATCCGCGCGAGGACATGATTTCCGACCTGGTCCATGCCCGCGATGCCG ATGGCGAGGGGTTGAGCTTTGCCGAGGCGGTGTCGCTGGTCCGCGCGCTGCTGATCGCCGGGAAC GAGACGACCACGACGGCGCTGGGCAATCTGTTCTTCATTCTCGCCACCCGGCCGGAGATCGCCACC CTGTTGCAGGAGTCGGCGGATGACGACCGGCTGATGAACCGTTTCGTCGAGGAATTGTTGCGGATC GAGCCGCGGTGCGCGGCCTGTCGCGCATGACCACGCGCAGGTCGAGCTGGGCGGACGACGCT GCCTGCGGGCGCGCATCTGCTGCTGATGTCGTCGGCCAATGATCAGGAGGATGTGTTCCCCCA TCCGCGCGATTTTGACCTCAACCGTCCCAATATCGGGCGGCACCTGTCCTTTGGCGGCGGGGTGCA TCGTTGCGTCGGGCTGGCGCTGGCGCGGATGGAGATCAAGGTGGCGGCGCGGAGATCGTCAAGC GCATCGGCAACATCGAGCTGGCCATTGCGGAGGCGGACATCCGCTATCTGCCACGGTAGCGACTC AATCGATCGAGCGACTGCCGATCCGTTTCACGCGGCGGGGCTGA	MSEEQAMTKLAERDLLATGSM TDPLISARPRDFYAAMRRDDPV HWDEKLGMYLVSRYEDIATIQ QDPITFSVNKG YHTQQAKGFQQ EFQEILERE GGYFPDAIMSDPP YHTRIRKLMEKAFTAHRVKALE PRIEKVVVDLVDSVAARGEAD AIHDFAVPLTVRIICEQLGLDWD MKDRIQRWSIAVTAQIGRMQD REQMLGHAREICDLQHYLIAK MQERQADPREDMISDLVHARD ADGEGLSFAEAVSLVRALLIAG NETTTTALGNLFFILATRPEIATL LQESADDDRLMNRFVEELLRIE PPVRGLSRMTTREVELGGRTLP AGAHLLMYASANDQEDVFPH PRDFDLNRPNIGRHLSFGGGVH RCVGLALARMEIKVAAREIVKR IGNIELAIAEADIRYLPTVATQSI ERLPIRFTRRG
N1MS32	ATGATCCATCTTCTGTCCGCAACCGTCGAACGCTGGCCGGTTGCCGGCGCGTTCATCATCAGTCGC GGGGCCAAGACGGACGTGGACGTGGTGGTCTGCACCGTGGGCGATGGCACCCATGTCGGCCGGGG CGAGGGGACGGCGATCTATTATGAGGGCGAGACGGCGGAAGGCTGCGCGGCGGGCGATCCGCGCCT ATGCCGGGCGGCTCGACCGGCAGGCGCTGCTGACGCAGATGCCGCGCGGCGCGGCGCGCAATGCG CTCGACTGTGCGCTCTGGGATCTGGAGGCGAAGCGCGCAGACGTGCCGGTCTGGCAACTCGCCGG GCTGAGTGCGCCCGCGCCGTTGCCCACTGCCTTCACCATCAGCCTGGGCGATCCCGCGACCATGGA GGCGGCCGCCCCGTACGGCGGGCGGGCGCGGTTTTGCCCTGCTCAAATGCAAACCTGACCGGGGAGG GTGACCGGGCGCGCATCGCCGCCGTGCGCGCGGGTGGCCCCGACGCGCGACTGATCGTCGATGCC AATGAAAGCTGGCATGATCTGGATATTGCGGCAGAGGCGCAGGCGCTGGCGGACCTTGGTGTGCA AATGGTTCGAGCAACCGGTTATCCATGGCCGGGAGGCGGCGCTGGCCGGGGTGGCGCGCGCGCTGC CGCTTTGTGCCGACGAAAGCTGCCACACCAGCGCGGACCTTGATCGGCTTGGTGA CTTTGATGCGG TGAACATCAAGCTGGACAAGGCAGGCGGCCTGACCGAGGCGCTGGTACTGCGTCAGGCAGCGCGG GAGCGGGGGTTCCGGGTCATGGTCGGCTGTATGCTCGGCACCTCGCTGGGCATCGCCCCGGCCGCG CTGGTGGCGCAGGGTGCCGACTGGATCGACCTTGACGGCGCGCTCCTGCTGGCGCGGGATCGCGA CGGCGCTCTGGCGCTGGACGATGGGCTGCTCCAGCCCGGCAACCTGTGGGGCATGGGCGAATGA	MIHLLSATVERWPVAGAFIISRG AKTDVDVVVCTVGDGTHVGR GEGTAIYYEGETAEGCAAIRA YAGPLDRQALLTQMPRGAARN ALDCALWDLEAKRADVPVWQ LAGLSAPAPLPTAFTISLGPAT MEAAARTAAGRGFALLKCKLT GEGDRARIAAVRAGAPDARLIV DANESWHDLDIAAEAQALADL GVEMVEQPIHGREAAAGVR APLPLCADESCHTSADLDRLGD FDAVNIKLDKAGGLTEALVLRQ AARERGFVRVMVGCMLGTSLGI APAALVAQGADWIDLGALLL ARDRDGALALDDGLLQPGNLW GMGE

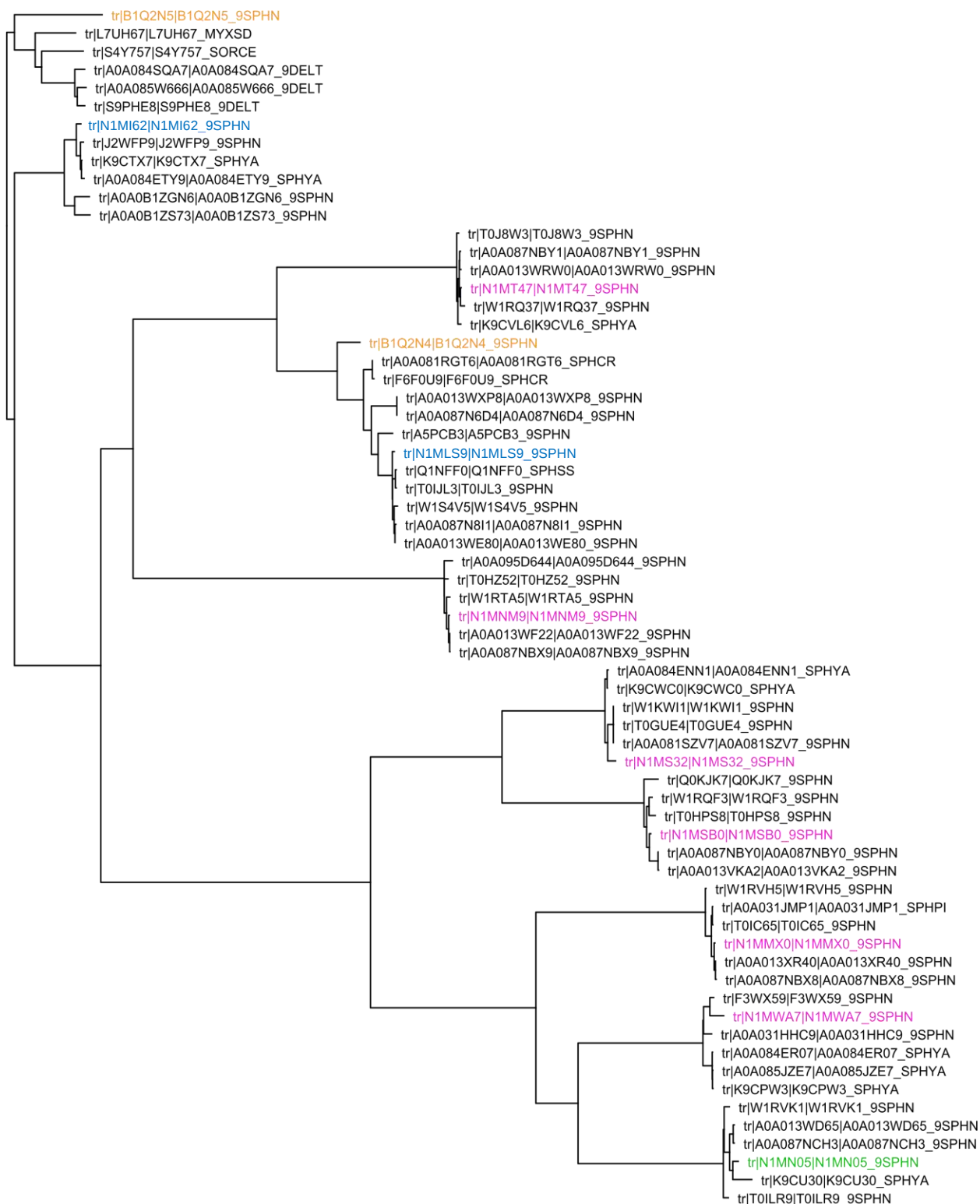


Figure S1. Phylogenetic relationship of *Sphingobium* sp. BiD32 proteins differentially expressed and functionally annotated to xenobiotics biodegradation and metabolism (pink), *Sphingobium* sp. BiD32 proteins functionally annotated to bisphenol degradation (green), *Sphingobium* sp. BiD32 proteins related to proteins responsible for BPA degradation by *Sphingomonas* sp. AO1 (blue), and *Sphingomonas* sp. AO1 proteins responsible for BPA degradation (orange).