Supplementary material

Meta-transcriptomic analysis of the response of river biofilms to pharmaceutical products using anonymous DNA microarrays

Etienne Yergeau, John R. Lawrence, Marley J. Waiser, Darren R. Korber, and Charles W. Greer

Supplementary Table S1: Primers designed and used in this study.

Probe #	Gene	name	Sequence 5'-3'
13925	Chaperone protein DnaK	dnaK-1F	AGCCGAACGACAAGGTGCAAT
		dnaK-1R	ACTTGCCGTACATCATGCCGGT
8047	ppGpp synthehase I	ppGpp-1F	CGACTCCCAGGATGGGGGCA
		ppGpp-1R	CCATGCACACCGTCGCCGAA
14393	tonB-dependent receptor	tonB-2F	GCCGGGGCAGCAGAATACAGG
		tonB-2R	CAGTGCTGCCTGCGCTGGTT
5271	photosystem I reaction center subunit X	psrcX-2F	GCGCAGATCAATCCCAGGGCA
		psrcX-2R	GTGACACCGCCGACAGCCAA
14137	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	NAGCVT-2F	GCGAAAATGGCACCAGCGAGC
		NAGCVT-2R	CCTGCCCATGGCGGGCAA
2027	DNA-directed DNA polymerase B	DNApol-1F	CGCCCCAAGTCTACACAAACCGT
		DNApol-1R	ACGCTAGGGTACATGGATGCAACA
19275	acetyl-CoA carboxylase	ACoAC-2F	GCGGCGCAAGTGTTCATTCGG
		ACoAC-1R	ACGGCTTGTTGGGGTCGCTG

Supplementary Table S2: Most significantly affected probes by the sulfamethoxazole treatment for the Wascana Creek

Probe	CA1	Gene	Organism	Phylum	E-value	COG
		Positive response				
16111	5.28	hypothetical protein	*Monosiga brevicollis* MX1	*Choanozoa*	1.00E-13	
14393	4.78	**TonB-dependent receptor**	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	6.00E-22	
14137	4.76	**UDP-N-acetylglucosamine 1-carboxyvinyltransferase**	*Magnetospirillum gryphiswaldense* MSR-1	*Alphaproteobacteria*	6.00E-42	M
5565	4.54	hypothetical protein Ssed_2383	*Shewanella sediminis* HAW-EB3	*Gammaproteobacteria*	8.00E-05	
5587	4.00	RNA binding S1	*Cylindrospermopsis raciborskii* CS-505	*Cyanobacteria*	1.00E-18	K
7607	3.86	iron-sulfur cluster assembly protein	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	4.00E-17	S
5589	3.68	serine/threonine protein kinase	*Cyanothece* sp. PCC 7425	*Cyanobacteria*	1.00E-35	TKL
11677	3.41	serine hydroxymethyltransferase	*Synechococcus elongatus* PCC 7942	*Cyanobacteria*	1.00E-17	E
5321	3.37	death-on-curing family protein	*Denitrovibrio acetiphilus* DSM 12809	*Deferribacteres*	8.00E-50	R
15643	3.19	trimethylamine dehydrogenase	*Octadecabacter antarcticus* 238	*Alphaproteobacteria*	1.00E-38	C
847	3.16	b-glycosyltransferase	*Cytophaga hutchinsonii* ATCC 33406	*Bacteroidetes*	1.00E-25	M
13717	2.99	hypothetical membrane protein	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	3.00E-15	
14385	2.92	glycosyl transferase, group 1	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	8.00E-50	M
12455	2.92	Tetratricopeptide repeat protein	*Planctomyces maris* DSM 8797	*Planctomycetes*	3.00E-08	
1849	2.83	hybrid sensory kinase	*Lyngbya* sp. PCC 8106	*Cyanobacteria*	2.00E-15	T
3255	2.80	transporter	*Rhodopirellula baltica* SH 1	*Planctomycetes*	3.00E-07	
8271	2.78	protein serine/threonine phosphatases	*Trichodesmium erythraeum* IMS101	*Cyanobacteria*	1.00E-20	T
18967	2.75	cytosolic tRNA-Ala synthetase, putative	*Perkinsus marinus* ATCC 50983	*Dinoflagellata*	8.00E-06	J
16587	2.73	alpha/beta hydrolase fold protein	*Clostridium carboxidivorans* P7	*Firmicutes*	2.00E-07	I
11445	2.68	thiamine biosynthesis protein ThiS	*Solibacter usitatus* Ellin6076	*Acidobacteria*	2.00E-12	H
10047	2.62	hypothetical protein L8106_19046	*Lyngbya* sp. PCC 8106	*Cyanobacteria*	1.00E-11	
1155	2.59	diaminopimelate epimerase	*uncultured methanogenic archaeon* RC-I	*Euryarcheota*	4.00E-05	E
		Negative response				
5263	-3.95	cell shape determining protein MreB	*Acidobacterium capsulatum* ATCC 51196	*Acidobacteria*	3.00E-14	D
18133	-3.89	glycosyl transferase group 1	*Spirosoma linguale* DSM 74	*Bacteroidetes*	3.00E-09	M
2027	-3.75	**DNA-directed DNA polymerase B**	*Cyanothece* sp. PCC 7424	*Cyanobacteria*	2.00E-12	L
19275	-3.35	**acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)**	*Halogeometricum borinquense* DSM 11551	*Euryarcheota*	8.00E-40	I
1683	-3.31	hemolysin A	*Methylobacterium nodulans* ORS 2060	*Alphaproteobacteria*	2.00E-16	Q

Probe	CA1	Gene	Organism	Phylum	E-value	COG
19299	-3.27	hypothetical protein VspiD_21975	*Verrucomicrobium spinosum* DSM 4136	*Verrucomicrobia*	2.00E-06	
12233	-3.27	protein-L-isoaspartate(D-aspartate) O-methyltransferase	*Rhodoferax ferrireducens* T118	*Betaproteobacteria*	8.00E-19	O
18099	-3.22	flagellar synthesis regulator FleN	*Pseudomonas aeruginosa* PAO1	*Gammaproteobacteria*	3.00E-32	D
19357	-3.15	hypothetical protein M446_1504	*Methylobacterium* sp. 4-46	*Alphaproteobacteria*	6.00E-16	
759	-3.08	Hypothetical protein CBG13434	*Caenorhabditis briggsae*	*Nematoda*	2.00E-08	
11731	-3.04	putative methyltransferase	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	3.00E-26	L
6171	-3.01	short chain enoyl-CoA hydratase / 3-hydroxyacyl-CoA dehydrogenase	*Polaromonas* sp. JS666	*Betaproteobacteria*	6.00E-58	I
16751	-2.94	glutamine amidotransferase	*alpha proteobacterium* BAL199	*Alphaproteobacteria*	1.00E-14	F
18653	-2.93	UbiA prenyltransferase	*Gallionella ferruginea* ES-2	*Betaproteobacteria*	2.00E-25	H
14509	-2.87	hypothetical protein PM8797T_03259	*Planctomyces maris* DSM 8797	*Planctomycetes*	5.00E-18	
8775	-2.82	Fatty acid/phospholipid synthesis protein PlsX	*Leptospira biflexa* 'Patoc 1 (Paris)'	*Spirochaetes*	8.00E-24	I
15591	-2.81	glutamine synthetase catalytic region	*Parvibaculum lavamentivorans* DS-1	*Alphaproteobacteria*	2.00E-26	E
8893	-2.76	Mg2+ transport protein	*Nodularia spumigena* CCY9414	*Cyanobacteria*	7.00E-17	P
365	-2.72	ribosomal protein L11	*Geobacter* sp. FRC-32	*Deltaproteobacteria*	3.00E-12	J
19763	-2.61	pyridine nucleotide-disulphide oxidoreductase	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	4.00E-30	C
13659	-2.56	MFS permease	*Agrobacterium vitis* S4	*Alphaproteobacteria*	4.00E-19	G
8427	-2.55	aldehyde Dehydrogenase	*Thermobaculum terrenum* ATCC BAA-798	NA	3.00E-16	C
19745	-2.55	putative membrane protein	*Lyngbya* sp. PCC 8106	*Cyanobacteria*	1.00E-10	
8585	-2.49	FHA domain containing protein	bacterium Ellin514	*Verrucomicrobia*	9.00E-09	T
13807	-2.47	DNA-directed RNA polymerase, beta subunit	*Microcoleus chthonoplastes* PCC 7420	*Cyanobacteria*	1.00E-52	K
3961	-2.44	TPR repeat protein	*Beggiatoa* sp. PS	*Gammaproteobacteria*	2.00E-10	S
6933	-2.44	acetate--CoA ligase	*Leptothrix cholodnii* SP-6	*Betaproteobacteria*	6.00E-26	I

Genes in boldface were used to design PCR primers or were depicted in Fig. 2.

COG: C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S3: Most significantly affected probes by the gemfibrozil treatment for the Wascana Creek

Probe	CA1	Gene	Organism	Phylum	E-value	COG
		Positive response				
5271	-3.70	**photosystem I reaction center subunit X**	*Synechococcus elongatus* PCC 6301	*Cyanobacteria*	2.00E-07	C
5107	-3.55	transaldolase	*Oligotropha carboxidovorans* OM5	*Alphaproteobacteria*	5.00E-27	G
19275	-3.35	**acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)**	*Halogeometricum borinquense* DSM 11551	*Euryarcheota*	8.00E-40	I
1683	-3.31	hemolysin A	*Methylobacterium nodulans* ORS 2060	*Alphaproteobacteria*	2.00E-16	Q
18099	-3.22	flagellar synthesis regulator FleN	*Pseudomonas aeruginosa* PAO1	*Gammaproteobacteria*	3.00E-32	D
19357	-3.15	hypothetical protein M446_1504	*Methylobacterium* sp. 4-46	*Alphaproteobacteria*	6.00E-16	
6171	-3.01	short chain enoyl-CoA hydratase / 3-hydroxyacyl-CoA dehydrogenase	*Polaromonas* sp. JS666	*Betaproteobacteria*	6.00E-58	I
14509	-2.87	hypothetical protein PM8797T_03259	*Planctomyces maris* DSM 8797	*Planctomycetes*	5.00E-18	
8775	-2.82	Fatty acid/phospholipid synthesis protein PlsX	*Leptospira biflexa* 'Patoc 1 (Paris)'	*Spirochaetes*	8.00E-24	I
15591	-2.81	glutamine synthetase catalytic region	*Parvibaculum lavamentivorans* DS-1	*Alphaproteobacteria*	2.00E-26	E
8477	-2.71	DNA gyrase, A subunit	*Methylotenera mobilis* JLW8	*Betaproteobacteria*	4.00E-07	L
		Negative response				
16111	5.28	hypothetical protein	*Monosiga brevicollis* MX1	*Choanozoa*	1.00E-13	
14393	4.78	**TonB-dependent receptor**	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	6.00E-22	
605	3.05	RNA polymerase, sigma-24 subunit, ECF subfamily	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	7.00E-17	K
13717	2.99	hypothetical membrane protein	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	3.00E-15	
14385	2.92	glycosyl transferase, group 1	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	8.00E-50	M
19439	2.89	TonB-dependent receptor	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	4.00E-16	
1849	2.83	hybrid sensory kinase	*Lyngbya sp.* PCC 8106	*Cyanobacteria*	2.00E-15	T
661	2.82	alpha amylase, catalytic domain subfamily, putative	*uncultured marine bacterium* Ant29B7	NA	2.00E-20	G

Genes in boldface were used to design PCR primers or were depicted in Fig. 2.

COG: C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S4: Most significantly affected probes by the erythromycin treatment for the Wascana Creek

Probe	CA2	Gene	Organism	Phylum	E-value	COG
		Positive response				
13925	5.74	**chaperone protein DnaK**	*Spirosoma linguale* DSM 74	*Bacteroidetes*	1.00E-07	O
5897	4.68	hypothetical protein CAMRE0001_0799	*Campylobacter rectus* RM3267	*Epsilonproteobacteria*	4.00E-24	
13903	4.13	reverse transcriptase family protein	*Chlorobium chlorochromatii* CaD3	*Chlorobi*	1.00E-21	L
13229	3.92	TonB protein	*Phenylobacterium zucineum* HLK1	*Alphaproteobacteria*	1.00E-13	
17047	3.64	isoleucyl-tRNA synthetase	*Gordonia bronchialis* DSM 43247	*Actinobacteria*	3.00E-06	J
20051	3.45	Rickettsia 17 kDa surface antigen family	*Nitrosococcus oceani* AFC27	*Gammaproteobacteria*	2.00E-12	S
5517	3.24	uncharacterized protein	uncultured bacterium HF186_75m_14K15	*NA*	3.00E-23	
13885	3.16	hypothetical protein Phep_2786	*Pedobacter heparinus* DSM 2366	*Bacteroidetes*	8.00E-05	
1911	3.12	hypothetical protein RB2654_23343	*Rhodobacterales bacterium* HTCC2654	*Alphaproteobacteria*	6.00E-22	
6183	3.08	ATP synthase F0, A subunit	*Leptothrix cholodnii* SP-6	*Betaproteobacteria*	2.00E-17	C
8241	2.98	DNA protecting protein DprA	*Variovorax paradoxus* S110	*Betaproteobacteria*	5.00E-13	LU
14253	2.93	VWA domain-containing protein	*Hyphomonas neptunium* ATCC 15444	*Alphaproteobacteria*	2.00E-16	
12417	2.86	Sec-independent protein translocase TatC	*Solibacter usitatus* Ellin6076	*Acidobacteria*	1.00E-36	U
1077	2.86	hypothetical protein	*Curvibacter* sp.	*Betaproteobacteria*	7.00E-10	
16407	2.79	UDP-N-acetylmuramate--alanine ligase/undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	*Cellulomonas flavigena* DSM 20109	*Actinobacteria*	2.00E-10	M
7247	2.77	hypothetical protein	*Arthrobacter nicotinovorans*	*Actinobacteria*	3.00E-09	
16353	2.68	Type II restriction enzyme, methylase subunit; N6 adenine-specific DNA methyltransferase protein, N12 class	*Synechococcus* sp. PCC 7002	*Cyanobacteria*	9.00E-49	V
17657	2.67	resolvase, N-terminal	*Dechloromonas aromatica* RCB	*Betaproteobacteria*	3.00E-30	L
12659	2.67	DEAD/DEAH box helicase domain protein	*Desulfovibrio salexigens* DSM 2638	*Deltaproteobacteria*	3.00E-17	L
8047	2.59	**(p)ppGpp synthetase I, SpoT/RelA**	*Nostoc azollae* 0708	*Cyanobacteria*	3.00E-62	TK
13311	2.57	hypothetical protein all7633	*Nostoc* sp. PCC 7120	*Cyanobacteria*	4.00E-23	
13831	2.56	UmuD protein	*Desulfovibrio magneticus* RS-1	*Deltaproteobacteria*	6.00E-07	KT
17103	2.50	ribosomal protein L4/L1 family	*Oligotropha carboxidovorans* OM5	*Alphaproteobacteria*	4.00E-40	J
8129	2.50	hypothetical protein HNE_2611	*Hyphomonas neptunium* ATCC 15444	*Alphaproteobacteria*	4.00E-14	
8231	2.48	UDP-glucose 4-epimerase	*Dyadobacter fermentans* DSM 18053	*Bacteroidetes*	7.00E-09	M
5921	2.45	sensor protein	*Leptothrix cholodnii* SP-6	*Betaproteobacteria*	8.00E-08	T

Probe	CA2	Gene	Organism	Phylum	E-value	COG
		Negative response				
5317	-5.09	conserved hypothetical protein	*Arthrospira maxima* CS-328	*Cyanobacteria*	5.00E-25	
5321	-4.13	death-on-curing family protein	*Denitrovibrio acetiphilus* DSM 12809	*Deferribacteres*	8.00E-50	R
1707	-4.03	Phosphoglycerate kinase	*Cyanothece* sp. PCC 7425	*Cyanobacteria*	3.00E-14	G
5345	-4.02	cysteine desulfurase, SufS subfamily	*Nostoc azollae* 0708	*Cyanobacteria*	1.00E-17	E
1293	-3.58	Methylated-DNA-(protein)-cysteine S-methyltransferase DNA binding	*Kosmotoga olearia* TBF 19.5.1	*Thermotogae*	7.00E-07	L
16249	-3.56	hypothetical protein all1772	*Nostoc* sp. PCC 7120	*Cyanobacteria*	8.00E-05	
4917	-3.19	protein RecA	*Prevotella veroralis* F0319	*Bacteroidetes*	6.00E-22	L
16447	-3.17	hypothetical protein RB2501_13699	*Robiginitalea biformata* HTCC2501	*Bacteroidetes*	8.00E-50	
14305	-3.15	hypothetical protein Lcho_2013	*Leptothrix cholodnii* SP-6	*Betaproteobacteria*	7.00E-07	
7495	-3.12	nucleoside recognition domain protein	*Sphingobacterium spiritivorum* ATCC 33300	*Bacteroidetes*	3.00E-52	R
2003	-3.12	Glycerophosphoryl diester phosphodiesterase	*Cylindrospermopsis raciborskii* CS-505	*Cyanobacteria*	6.00E-26	C
6765	-3.03	riboflavin synthase, alpha subunit	*Synechococcus* sp. PCC 7335	*Cyanobacteria*	9.00E-10	H
13905	-2.99	Aldehyde oxidase	*Erythrobacter* sp. NAP1	*Alphaproteobacteria*	3.00E-35	C
11725	-2.97	chromatic acclimation sensor	*Synechocystis* sp. PCC 6803	*Cyanobacteria*	2.00E-21	
5357	-2.96	hypothetical protein HCH_06244	*Hahella chejuensis* KCTC 2396	*Gammaproteobacteria*	8.00E-07	
9841	-2.93	mannose-6-phosphate isomerase	*Deinococcus geothermalis* DSM 11300	*Deinococcus-Thermus*	4.00E-13	G
16399	-2.92	acetyl-CoA synthetase	*Oceanicola batsensis* HTCC2597	*Alphaproteobacteria*	2.00E-06	C
8467	-2.80	protein of unknown function DUF323	*Haliangium ochraceum* DSM 14365	*Deltaproteobacteria*	3.00E-10	
739	-2.77	serine/threonine protein kinase	*Haliangium ochraceum* DSM 14365	*Deltaproteobacteria*	2.00E-07	TKL
19493	-2.75	Acyl-CoA dehydrogenase, C-terminal domain protein	*Alcanivorax* sp. DG881	*Gammaproteobacteria*	3.00E-30	I
15825	-2.73	hypothetical protein VspiD_11015	*Verrucomicrobium spinosum* DSM 4136	*Verrucomicrobia*	4.00E-26	
2129	-2.71	tRNA-i(6)A37 thiotransferase enzyme MiaB	*Pelobacter carbinolicus* DSM 2380	*Bacteroidetes*	4.00E-08	J
13769	-2.61	YD repeat protein	*Cyanothece* sp. PCC 7822	*Cyanobacteria*	4.00E-18	

Genes in boldface were used to design PCR primers or were depicted in Fig. 2.

COG: C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S5: Most significantly affected probes by the sulfamethazine treatment for the Wascana Creek

Probe	CA2	Gene	Organism	Phylum	E-value	COG
		Positive response				
5271	-7.46	**photosystem I reaction center subunit X**	*Synechococcus elongatus* PCC 6301	*Cyanobacteria*	2.00E-07	C
5317	-5.09	conserved hypothetical protein	*Arthrospira maxima* CS-328	*Cyanobacteria*	5.00E-25	
5263	-4.74	cell shape determining protein MreB	*Acidobacterium capsulatum* ATCC 51196	*Acidobacteria*	3.00E-14	D
5345	-4.02	cysteine desulfurase, SufS subfamily	*Nostoc azollae* 0708	*Cyanobacteria*	1.00E-17	E
1293	-3.58	Methylated-DNA-(protein)-cysteine S-methyltransferase DNA binding	*Kosmotoga olearia* TBF 19.5.1	*Thermotogae*	7.00E-07	L
14255	-3.32	hypothetical protein VspiD_21975	*Verrucomicrobium spinosum* DSM 4136	*Verrucomicrobia*	2.00E-24	
4917	-3.19	protein RecA	*Prevotella veroralis* F0319	*Bacteroidetes*	6.00E-22	L
14305	-3.15	hypothetical protein Lcho_2013	*Leptothrix cholodnii* SP-6	*Betaproteobacteria*	7.00E-07	
7495	-3.12	nucleoside recognition domain protein	*Sphingobacterium spiritivorum* ATCC 33300	*Bacteroidetes*	3.00E-52	R
2003	-3.12	Glycerophosphoryl diester phosphodiesterase	*Cylindrospermopsis raciborskii* CS-505	*Cyanobacteria*	6.00E-26	C
6765	-3.03	riboflavin synthase, alpha subunit	*Synechococcus* sp. PCC 7335	*Cyanobacteria*	9.00E-10	H
13407	-3.00	dihydrolipoyllysine-residue succinyltransferase, component of 2-oxoglutarate dehydrogenase complex	*Pedobacter* sp. BAL39	*Bacteroidetes*	3.00E-40	C
13905	-2.99	Aldehyde oxidase	*Erythrobacter* sp. NAP1	*Alphaproteobacteria*	3.00E-35	C
5357	-2.96	hypothetical protein HCH_06244	*Hahella chejuensis* KCTC 2396	*Gammaproteobacteria*	8.00E-07	
16399	-2.92	acetyl-CoA synthetase	*Oceanicola batsensis* HTCC2597	*Alphaproteobacteria*	2.00E-06	C
2027	-2.91	**DNA-directed DNA polymerase B**	*Cyanothece* sp. PCC 7424	*Cyanobacteria*	2.00E-12	L
1951	-2.85	heme-binding protein	bacterium Ellin514	*Verrucomicrobia*	7.00E-06	G
5257	-2.73	phosphoesterase PA-phosphatase related protein	*Halothiobacillus neapolitanus* c2	*Gammaproteobacteria*	2.00E-08	I
13707	-2.62	outer membrane protein assembly complex, YaeT protein	*Haliangium ochraceum* DSM 14365	*Deltaproteobacteria*	2.00E-44	M
13769	-2.61	YD repeat protein	*Cyanothece* sp. PCC 7822	*Cyanobacteria*	4.00E-18	
		Negative response				
1911	3.12	hypothetical protein RB2654_23343	*Rhodobacterales bacterium* HTCC2654	*Alphaproteobacteria*	6.00E-22	
12157	3.10	hypothetical protein FBALC1_15992	*Flavobacteriales bacterium* ALC-1	*Bacteroidetes*	5.00E-09	
8241	2.98	DNA protecting protein DprA	*Variovorax paradoxus* S110	*Betaproteobacteria*	5.00E-13	LU
14253	2.93	VWA domain-containing protein	*Hyphomonas neptunium* ATCC 15444	*Alphaproteobacteria*	2.00E-16	

Probe	CA2	Gene	Organism	Phylum	E-value	COG
1077	2.86	hypothetical protein	*Curvibacter* sp.	*Betaproteobacteria*	7.00E-10	
2073	2.72	two-component system sensor histidine kinase/response regulator	*Bacteroides* sp. 9_1_42FAA	*Bacteroidetes*	6.00E-07	TK
17657	2.67	resolvase, N-terminal	*Dechloromonas aromatica* RCB	*Betaproteobacteria*	3.00E-30	L
8129	2.50	hypothetical protein HNE_2611	*Hyphomonas neptunium* ATCC 15444	*Alphaproteobacteria*	4.00E-14	

Genes in boldface were used to design PCR primers or were depicted in Fig. 2.

COG: C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S6: Most significantly affected probes by the erythromycin treatment for the South Sackatchewan River

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
Positive							
2411	2.91	-0.54	mandelate racemase/muconate lactonizing enzyme-like protein	*Opitutaceae* bacterium TAV2	*Verrucomicrobia*	4.00E-19	M
13925	2.86	-0.42	**chaperone protein DnaK**	*Spirosoma linguale* DSM 74	*Bacteroidetes*	1.00E-07	O
18719	0.23	-3.91	hypothetical protein RSc1496	*Ralstonia solanacearum* GMI1000	*Betaproteobacteria*	2.00E-14	
16357	0.10	3.15	thioredoxin	*Sphingobacterium spiritivorum* ATCC 33300	*Bacteroidetes*	7.00E-12	O
12663	-0.31	-2.80	Phage conserved hypothetical protein	*Comamonas testosteroni* CNB-2	*Betaproteobacteria*	2.00E-18	
2349	-0.81	-2.66	putative twin-arginine translocation pathway signal	*Flavobacteria bacterium* BBFL7	*Bacteroidetes*	7.00E-10	S
7007	-1.23	-3.10	carboxyphosphonoenolpyruvate phosphonomutase	uncultured crenarchaeote 31-F-01	*Crenarchaeota*	3.00E-20	G
11905	-1.45	-2.83	dihydrodipicolinate synthase	*Opitutaceae* bacterium TAV2	*Verrucomicrobia*	1.00E-26	EM
8909	-1.63	-2.94	thiamine-monophosphate kinase, putative	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	6.00E-30	H
19775	-2.88	4.84	protein of unknown function DUF433	*Cyanothece* sp. PCC 8802	*Cyanobacteria*	6.00E-23	
3063	-3.02	1.41	Transposase (class II)	*Synechococcus* sp. WH 5701	*Cyanobacteria*	1.00E-19	L
7775	-3.16	0.02	hypothetical protein MED134_14041	*Dokdonia donghaensis* MED134	*Bacteroidetes*	2.00E-64	
Negative							
16811	3.88	0.71	two-component hybrid sensor and regulator	*Nostoc* sp. PCC 7120	*Cyanobacteria*	2.00E-07	TK
9611	3.70	0.57	**3-isopropylmalate dehydratase, large subunit**	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	7.00E-25	E
17377	3.30	-2.14	hypothetical protein TRIADDRAFT_30805	*Trichoplax adhaerens*	*Placozoa*	5.00E-16	
9183	3.28	-1.83	Valyl-tRNA synthetase, class Ia	*Crocosphaera watsonii* WH 8501	*Cyanobacteria*	5.00E-43	J
18721	3.25	0.94	hypothetical protein Nmul_A0121	*Nitrosospira multiformis* ATCC 25196	*Nitrospirae*	8.00E-16	
4385	3.15	0.62	Protein kinase domain containing protein	*Brugia malayi*	*Nematoda*	2.00E-39	TKL
18507	2.99	-0.67	fructokinase	*Methanosarcina acetivorans* C2A	*Euryarchaeota*	5.00E-16	G
6609	1.93	3.44	**RecD/TraA family helicase**	*Methylobacterium* sp. 4-46	*Alphaproteobacteria*	2.00E-21	L
10833	1.15	4.75	ATPase, P-type (transporting), HAD superfamily, subfamily IC	*Mycobacterium* sp. JLS	*Actinobacteria*	2.00E-06	
10081	0.82	-3.27	epimerase/dehydratase	*Streptomyces* sp. AA4	*Actinobacteria*	9.00E-18	
18595	0.46	3.18	hypothetical protein DSM3645_02153	*Blastopirellula marina* DSM 3645	*Planctomycetes*	1.00E-15	

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
19493	0.44	4.15	**Acyl-CoA dehydrogenase, C-terminal domain protein**	*Alcanivorax* sp. DG881	*Gammaproteobacteria*	9.00E-31	I
5767	0.09	3.35	hypothetical protein RB2150_01719	*Rhodobacterales* bacterium HTCC2150	*Alphaproteobacteria*	8.00E-06	
13213	-0.10	3.04	glutamine synthetase catalytic region	*Parvibaculum lavamentivorans* DS-1	*Alphaproteobacteria*	2.00E-22	E
16817	-0.71	3.90	GCN5-related N-acetyltransferase	*Shewanella woodyi* ATCC 51908	*Gammaproteobacteria*	4.00E-05	
17773	-1.22	3.59	GCN5-related N-acetyltransferase	*Rhizobium leguminosarum* WSM1325	*Alphaproteobacteria*	2.00E-07	
4505	-1.30	3.06	hypothetical protein MED134_13151	*Dokdonia donghaensis* MED134	*Bacteroidetes*	3.00E-30	
8121	-2.78	2.01	N-acetylmuramoyl-L-alanine amidase	*Sphingobacterium spiritivorum* ATCC 33861	*Bacteroidetes*	3.00E-06	M
19615	-3.02	-0.77	hypothetical protein AM1_2725	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	1.00E-11	
3545	-3.18	0.95	ATP-dependent Clp protease ATP-binding subunit ClpC	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	5.00E-35	O
18683	-4.14	-0.31	hypothetical protein MAE_42380	*Microcystis aeruginosa* NIES-843	*Cyanobacteria*	4.00E-20	

Genes in boldface are similar to genes from the Wascana Creek (Table S2-S5).

COG : C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S7: Most significantly affected probes by the sulfamethazine treatment for the South Sackatchewan River

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
Positive							
19813	4.40	1.63	hypothetical protein	*Trypanosoma cruzi* strain CL Brener	*Euglenozoa*	2.00E-05	
16811	3.88	0.71	two-component hybrid sensor and regulator	*Nostoc* sp. PCC 7120	*Cyanobacteria*	2.00E-07	TK
9611	3.70	0.57	**3-isopropylmalate dehydratase, large subunit**	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	7.00E-25	E
17377	3.30	-2.14	hypothetical protein TRIADDRAFT_30805	*Trichoplax adhaerens*	*Placozoa*	5.00E-16	
18721	3.25	0.94	hypothetical protein Nmul_A0121	*Nitrosospira multiformis* ATCC 25196	*Nitrospirae*	8.00E-16	
4385	3.15	0.62	Protein kinase domain containing protein	*Brugia malayi*	*Nematoda*	2.00E-39	TKL
17233	2.99	-1.06	ArsR family transcriptional regulator	*Gemella haemolysans* ATCC 10379	*Firmicutes*	2.00E-07	K
18507	2.99	-0.67	fructokinase	*Methanosarcina acetivorans* C2A	*Euryarchaeota*	5.00E-16	G
4085	2.95	-1.36	hypothetical protein Cflav_PD2800	bacterium Ellin514	NA	1.00E-33	
9471	2.94	-1.15	M61 family peptidase	*Flavobacterium psychrophilum* JIP02/86	*Bacteroidetes*	1.00E-18	R
6597	2.87	0.28	nucleolar RNA-binding protein	*Leishmania major* strain Friedlin	*Euglenozoa*	9.00E-07	J
19153	2.87	-1.76	NUDIX hydrolase	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	2.00E-06	F
13925	2.86	-0.42	**chaperone protein DnaK**	*Spirosoma linguale* DSM 74	*Bacteroidetes*	1.00E-07	O
6609	1.93	3.44	**RecD/TraA family helicase**	*Methylobacterium* sp. 4-46	*Alphaproteobacteria*	2.00E-21	L
10833	1.15	4.75	ATPase, P-type (transporting), HAD superfamily, subfamily IC	*Mycobacterium* sp. JLS	*Actinobacteria*	2.00E-06	P
9437	0.67	3.46	predicted protein	*Thalassiosira pseudonana* CCMP1335	*Bacillariophyta*	2.00E-08	
5767	0.09	3.35	hypothetical protein RB2150_01719	*Rhodobacterales bacterium* HTCC2150	*Alphaproteobacteria*	8.00E-06	
16817	-0.71	3.90	GCN5-related N-acetyltransferase	*Shewanella woodyi* ATCC 51908	*Gammaproteobacteria*	4.00E-05	R
18937	-0.77	-2.80	hypothetical protein Msil_1356	*Methylocella silvestris* BL2	*Alphaproteobacteria*	2.00E-10	
4505	-1.30	3.06	hypothetical protein MED134_13151	*Dokdonia donghaensis* MED134	*Bacteroidetes*	3.00E-30	
13251	-1.60	2.99	putative selenocysteine synthase	uncultured bacterium	NA	7.00E-15	E
Negative							
7011	-0.05	-3.91	hypothetical protein ALOHA_HF4000APKG2J17ctg1g16	uncultured marine microorganism HF4000_APKG2J17	NA	3.00E-09	
3545	-3.18	0.95	ATP-dependent Clp protease ATP-binding subunit ClpC	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	5.00E-35	O
35	-1.86	-2.71	**protein of unknown function DUF490**	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	4.00E-14	
12663	-0.31	-2.80	Phage conserved hypothetical protein	*Comamonas testosteroni* CNB-2	*Betaproteobacteria*	2.00E-18	
7007	-1.23	-3.10	carboxyphosphonoenolpyruvate phosphonomutase	uncultured crenarchaeote 31-F-01	*Crenarchaeota*	3.00E-20	G

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
11905	-1.45	-2.83	dihydrodipicolinate synthase	*Opitutaceae* bacterium TAV2	*Verrucomicrobia*	1.00E-26	EM
19775	-2.88	4.84	protein of unknown function DUF433	*Cyanothece* sp. PCC 8802	*Cyanobacteria*	6.00E-23	
3063	-3.02	1.41	Transposase (class II)	*Synechococcus* sp. WH 5701	*Cyanobacteria*	1.00E-19	L
7775	-3.16	0.02	hypothetical protein MED134_14041	*Dokdonia donghaensis* MED134	*Bacteroidetes*	2.00E-64	
8909	-1.63	-2.94	thiamine-monophosphate kinase, putative	*Acaryochloris marina* MBIC11017	*Cyanobacteria*	6.00E-30	H
9095	0.02	-2.76	delta-aminolevulinic acid dehydratase	*Roseiflexus* sp. RS-1	*Chloroflexi*	7.00E-47	H
3211	-2.91	-0.53	Stage II sporulation protein E	*Microcoleus chthonoplastes* PCC 7420	*Cyanobacteria*	4.00E-58	TK
13463	-1.47	-3.22	**DEAD/DEAH box helicase domain protein**	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	1.00E-49	LKJ
18697	-1.03	-3.88	SpoVR family protein	*Haliangium ochraceum* DSM 14365	*Deltaproteobacteria*	1.00E-66	S

Genes in boldface are similar to genes from the Wascana Creek (Table S2-S5).

COG : C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S8: Most significantly affected probes by the gemfibrozil treatment for the South Sackatchewan River

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
Positive							
16411	-0.32	3.14	isopentenyl pyrophosphate isomerase	*Nostoc punctiforme* PCC 73102	*Cyanobacteria*	3.00E-31	C
4505	-1.30	3.06	hypothetical protein MED134_13151	*Dokdonia donghaensis* MED134	*Bacteroidetes*	3.00E-30	
8121	-2.78	2.01	N-acetylmuramoyl-L-alanine amidase	*Sphingobacterium spiritivorum* ATCC 33861	*Bacteroidetes*	3.00E-06	M
3545	-3.18	0.95	ATP-dependent Clp protease ATP-binding subunit ClpC	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	5.00E-35	O
9183	3.28	-1.83	Valyl-tRNA synthetase, class Ia	*Crocosphaera watsonii* WH 8501	*Cyanobacteria*	5.00E-43	J
35	-1.86	-2.71	protein of unknown function DUF490	*Chitinophaga pinensis* DSM 2588	*Bacteroidetes*	4.00E-14	
11905	-1.45	-2.83	dihydrodipicolinate synthase	*Opitutaceae* bacterium TAV2	*Verrucomicrobia*	1.00E-26	EM
3827	-0.34	-3.08	iojap-related protein	*Prevotella melaninogenica* ATCC 25845	*Bacteroidetes*	4.00E-06	S
10081	0.82	-3.27	epimerase/dehydratase	*Streptomyces* sp. AA4	*Actinobacteria*	9.00E-18	
18719	0.23	-3.91	hypothetical protein RSc1496	*Ralstonia solanacearum* GMI1000	*Betaproteobacteria*	2.00E-14	
Negative							
19493	0.44	4.15	**Acyl-CoA dehydrogenase, C-terminal domain protein**	*Alcanivorax* sp. DG881	*Gammaproteobacteria*	9.00E-31	I
6609	1.93	3.44	**RecD/TraA family helicase**	*Methylobacterium* sp. 4-46	*Alphaproteobacteria*	2.00E-21	L
16357	0.10	3.15	thioredoxin	*Sphingobacterium spiritivorum* ATCC 33300	*Bacteroidetes*	7.00E-12	O
13213	-0.10	3.04	glutamine synthetase catalytic region	*Parvibaculum lavamentivorans* DS-1	*Alphaproteobacteria*	2.00E-22	E
19813	4.40	1.63	hypothetical protein	*Trypanosoma cruzi* strain CL Brener	*Euglenozoa*	2.00E-05	
6597	2.87	0.28	nucleolar RNA-binding protein	*Leishmania major* strain Friedlin	*Euglenozoa*	9.00E-07	J
13755	-2.87	-0.18	TPR repeat-containing protein	*Chloroherpeton thalassium* ATCC 35110	*Chlorobi*	3.00E-22	S
13925	2.86	-0.42	**chaperone protein DnaK**	*Spirosoma linguale* DSM 74	*Bacteroidetes*	1.00E-07	O
19935	2.89	-0.97	hypothetical protein Rru_A2454	*Rhodospirillum rubrum* ATCC 11170	*Alphaproteobacteria*	5.00E-20	
17233	2.99	-1.06	ArsR family transcriptional regulator	*Gemella haemolysans* ATCC 10379	*Firmicutes*	2.00E-07	K
8341	3.14	-1.61	polysaccharide polymerase	*Sorangium cellulosum* 'So ce 56'	*Deltaproteobacteria*	3.00E-08	M
12663	-0.31	-2.80	Phage conserved hypothetical protein	*Comamonas testosteroni* CNB-2	*Betaproteobacteria*	2.00E-18	
7007	-1.23	-3.10	carboxyphosphonoenolpyruvate phosphonomutase	uncultured crenarchaeote 31-F-01	*Crenarchaeota*	3.00E-20	G

Genes in boldface are similar to genes from the Wascana Creek (Table S2-S5).

COG : C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.

Supplementary Table S9: Most significantly affected probes by the sulfamethoxazole treatment for the South Sackatchewan River

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
Positive							
13213	-0.10	3.04	glutamine synthetase catalytic region	*Parvibaculum lavamentivorans* DS-1	*Alphaproteobacteria*	2.00E-22	E
9095	0.02	-2.76	delta-aminolevulinic acid dehydratase	*Roseiflexus* sp. RS-1	*Chloroflexi*	7.00E-47	H
12663	-0.31	-2.80	Phage conserved hypothetical protein	*Comamonas testosteroni* CNB-2	*Betaproteobacteria*	2.00E-18	
3827	-0.34	-3.08	iojap-related protein	*Prevotella melaninogenica* ATCC 25845	*Bacteroidetes*	4.00E-06	S
8121	-2.78	2.01	N-acetylmuramoyl-L-alanine amidase	*Sphingobacterium spiritivorum* ATCC 33861	*Bacteroidetes*	3.00E-06	M
9183	3.28	-1.83	Valyl-tRNA synthetase, class Ia	*Crocosphaera watsonii* WH 8501	*Cyanobacteria*	5.00E-43	J
10081	0.82	-3.27	epimerase/dehydratase	*Streptomyces* sp. AA4	*Actinobacteria*	9.00E-18	
7011	-0.05	-3.91	hypothetical protein ALOHA_HF4000APKG2J17ctg1g16	uncultured marine microorganism HF4000_APKG2J17	NA	3.00E-09	
8341	3.14	-1.61	polysaccharide polymerase	*Sorangium cellulosum* 'So ce 56'	*Deltaproteobacteria*	3.00E-08	M
10047	4.57	-0.43	**conserved hypothetical protein**	*Arthrospira maxima* CS-328	*Cyanobacteria*	6.00E-13	
15039	4.02	-0.59	membrane lipoprotein lipid attachment site	*Polynucleobacter necessarius* subsp. necessarius STIR1	*Betaproteobacteria*	3.00E-10	R
18683	-4.14	-0.31	hypothetical protein MAE_42380	*Microcystis aeruginosa* NIES-843	*Cyanobacteria*	4.00E-20	
19935	2.89	-0.97	hypothetical protein Rru_A2454	*Rhodospirillum rubrum* ATCC 11170	*Alphaproteobacteria*	5.00E-20	
4085	2.95	-1.36	hypothetical protein Cflav_PD2800	bacterium Ellin514	NA	1.00E-33	
4385	3.15	0.62	Protein kinase domain containing protein	*Brugia malayi*	*Nematoda*	2.00E-39	TKL
17377	3.30	-2.14	hypothetical protein TRIADDRAFT_30805	*Trichoplax adhaerens*	*Placozoa*	5.00E-16	
18507	2.99	-0.67	fructokinase	*Methanosarcina acetivorans* C2A	*Euryarchaeota*	5.00E-16	G
7007	-1.23	-3.10	carboxyphosphonoenolpyruvate phosphonomutase	uncultured crenarchaeote 31-F-01	*Crenarchaeota*	3.00E-20	G
negative							
9437	0.67	3.46	predicted protein	*Thalassiosira pseudonana* CCMP1335	*Bacillariophyta*	2.00E-08	
16817	-0.71	3.90	GCN5-related N-acetyltransferase	*Shewanella woodyi* ATCC 51908	*Gammaproteobacteria*	4.00E-05	R
4505	-1.30	3.06	hypothetical protein MED134_13151	*Dokdonia donghaensis* MED134	*Bacteroidetes*	3.00E-30	
3545	-3.18	0.95	ATP-dependent Clp protease ATP-binding subunit ClpC	*Gemmatimonas aurantiaca* T-27	*Gemmatimonadetes*	5.00E-35	O
11905	-1.45	-2.83	dihydrodipicolinate synthase	*Opitutaceae* bacterium TAV2	*Verrucomicrobia*	1.00E-26	EM
18719	0.23	-3.91	hypothetical protein RSc1496	*Ralstonia solanacearum* GMI1000	*Betaproteobacteria*	2.00E-14	
19775	-2.88	4.84	protein of unknown function DUF433	*Cyanothece* sp. PCC 8802	*Cyanobacteria*	6.00E-23	
16411	-0.32	3.14	isopentenyl pyrophosphate isomerase	*Nostoc punctiforme* PCC 73102	*Cyanobacteria*	3.00E-31	C

Probe	CA1	CA2	Gene	Organism	Phylum	E-value	COG
5767	0.09	3.35	hypothetical protein RB2150_01719	*Rhodobacterales* bacterium HTCC2150	*Alphaproteobacteria*	8.00E-06	
6609	1.93	3.44	**RecD/TraA family helicase**	*Methylobacterium* sp. 4-46	*Alphaproteobacteria*	2.00E-21	L
10833	1.15	4.75	ATPase, P-type (transporting), HAD superfamily, subfamily IC	*Mycobacterium* sp. JLS	*Actinobacteria*	2.00E-06	P
16357	0.10	3.15	thioredoxin	*Sphingobacterium spiritivorum* ATCC 33300	*Bacteroidetes*	7.00E-12	O
13559	-2.73	1.87	N-acetyl-gamma-glutamyl-phosphate reductase	*Curvibacter* putative symbiont of Hydra magnipapillata	*Betaproteobacteria*	3.00E-17	E
13755	-2.87	-0.18	TPR repeat-containing protein	*Chloroherpeton thalassium* ATCC 35110	*Chlorobi*	3.00E-22	S
18773	3.01	0.52	GUN4-like	*Crocosphaera watsonii* WH 8501	*Cyanobacteria*	3.00E-09	TKL
17773	-1.22	3.59	GCN5-related N-acetyltransferase	*Rhizobium leguminosarum* bv. trifolii WSM1325	*Alphaproteobacteria*	2.00E-07	R

Genes in boldface are similar to genes from the Wascana Creek (Table S2-S5).

COG : C, Energy production and conversion; D, Cell division and chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell envelope biogenesis, outer membrane; O, Posttranslational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms.