

1 **Table S1.** Transcripts detected by *de novo* assembly RNA-Seq data

ID	Non-redundant	Description	Organism	Non-redundant length	Fraction of Alignment nr covered by alignment	% identity	Alignment e-value
8	319794846	hypothetical protein Varpa_4205	Variovorax paradoxus EPS	90	0.89	63	1.00E-23
48	323139488	SEC-C motif domain protein	Methylocystis sp. ATCC 49242	169	0.20	55	0.0004
48	323139488	SEC-C motif domain protein	Methylocystis sp. ATCC 49242	169	0.41	63	4.00E-18
90	307110660	hypothetical protein CHLNCDRAFT_50435	Chlorella variabilis	1535	0.05	24	4.5
90	307110660	hypothetical protein CHLNCDRAFT_50435	Chlorella variabilis	1535	0.05	24	4.6
120	309365749	hypothetical protein CBG_01089	Caenorhabditis briggsae AF16	812	0.12	32	0.7
121	310821854	hypothetical protein STAUR_4605	Stigmatella aurantiaca DW4/3-1	728	0.08	38	3.5
125	197103309	hypothetical protein PHZ_p0169	Phenylobacterium zucineum HLK1	223	0.55	60	8.00E-34
142	78213547	glycosyltransferase	Synechococcus sp. CC9605	1003	0.31	31	5.00E-14
142	325116388	tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA, related	Neospora caninum Liverpool	819	0.21	31	0.088
171	323136696	hypothetical protein Met49242DRAFT_1164	Methylocystis sp. ATCC 49242	92	0.64	53	6.00E-06
266	338532669	chloride channel	Myxococcus fulvus HW-1	628	0.72	63	3.00E-138
300	323136327	Patatin	Methylocystis sp. ATCC 49242	284	0.44	60	3.00E-38
303	227819692	transposase Y4ZB	Sinorhizobium fredii NGR234	493	0.35	42	8.00E-24
308	322440050	hypothetical protein AUBANDRAFT_66784	Aureococcus anophagefferens	1445	0.48	35	2.00E-22

308	323449050	hypothetical protein AURANDRAFT_66784	Aureococcus anophagefferens	1445	0.48	25	3.00E-33
314	341615616	two-component response regulator	Citromicrobium sp. JLT1363	257	0.15	59	0.32
339	149918868	site-specific recombinase, phage integrase family protein	Plesiocystis pacifica SIR-1	120	0.53	36	7.9
360	46153	endo-glucanase	Ruminococcus flavefaciens	680	0.09	38	4.6
370	167901659	hypothetical protein BpseN_05228	Burkholderia pseudomallei NCTC 13177	474	0.32	29	4.9
381	203288753	BdrQ-like protein	Borrelia duttonii Ly	239	0.67	30	4.00E-05
401	323457221	hypothetical protein AURANDRAFT_70536	Aureococcus anophagefferens	1999	0.06	33	2.1
431	221486147	conserved hypothetical protein	Toxoplasma gondii GT1	702	0.16	30	5.9
469	16519794	transposase number 3 of uncharacterized insertion sequence	Sinorhizobium fredii NGR234	511	0.31	62	7.00E-51
518	323135994	2-oxoglutarate dehydrogenase, E2 subunit, dihydrolipoamide succinyltransferase	Methylocystis sp. ATCC 49242	410	0.10	68	0.0002
523	332821691	PREDICTED: hypothetical protein LOC100609280	Pan troglodytes	329	0.43	28	0.92
542	299133059	Pseudomurein-binding repeat protein	Afipia sp. 1NLS2	686	0.28	96	2.00E-102
542	299133059	Pseudomurein-binding repeat protein	Afipia sp. 1NLS2	686	0.17	94	3.00E-59
569	118340377	putative granule bound starch synthase	Sorghum bicolor	608	0.08	45	5.9
570	316932709	RNA-binding S4 domain- containing protein	Rhodopseudomonas palustris DX-1	717	0.06	41	4.7
570	325116236	putative retinitis pigmentosa GTPase regulator	Neospora caninum Liverpool	1413	0.04	43	0.41
582	255714473	KLTH0E00704p	Lachancea thermotolerans	1389	0.03	42	2.6
712	25029285	hypothetical protein CE2729	Corynebacterium efficiens YS-314	251	0.26	43	0.0003
721	88812339	succinate dehydrogenase catalytic	Nitrococcus mobilis Nb-	259	0.75	69	4.00E-72

		subunit	231				
737	222630362	hypothetical protein OsJ_17292	Oryza sativa Japonica Group	377	0.13	40	2
737	115462351	Os05g0171300	Oryza sativa Japonica Group	415	0.12	40	2
783	119174408	predicted protein	Coccidioides immitis RS	363	0.13	41	7.8
787	121710406	Ras GTPase activating protein, putative	Aspergillus clavatus NRRL 1	1671	0.04	35	5.9
788	323135945	molybdenum cofactor biosynthesis protein A	Methylocystis sp. ATCC 49242	346	0.09	87	1.00E-06
809	186510389	octicosapeptide/Phox/Bem1p domain-containing protein kinase	Arabidopsis thaliana	1117	0.09	30	2.7
810	83312534	hypothetical protein amb3435	Magnetospirillum magneticum AMB-1	86	0.38	61	0.029
810	83312534	hypothetical protein amb3435	Magnetospirillum magneticum AMB-1	86	0.42	58	0.006
823	154312862	hypothetical protein BC1G_05132	Botryotinia fuckeliana B05.10	315	0.13	44	7.9
829	304394601	conserved hypothetical protein	Ahrensia sp. R2A130	55	0.73	53	0.049
833	9631033	conotoxin-like protein	Lymantria dispar MNPV	92	0.68	38	7.9
844	298241184	conserved hypothetical protein	Ktedonobacter racemifer DSM 44963	204	0.16	55	2.1
895	198283398	XRE family transcriptional regulator	Acidithiobacillus ferrooxidans ATCC 53993	152	0.93	32	0.0004
909	315500368	transcriptional regulator, mucr family	Asticcacaulis excentricus CB 48	141	1.03	57	5.00E-36
929	23500301	IS66 family orf3	Brucella suis 1330	523	0.45	60	1.00E-70
929	288962634	transposase	Azospirillum sp. B510	420	0.29	63	6.00E-35
932	73980988	PREDICTED: similar to ALMS1	Canis familiaris	4146	0.01	42	7.9
933	307108552	expressed protein	Chlorella variabilis	838	0.18	27	0.41
954	299133061	DNA modification methyltransferase-related protein	Afipia sp. 1NLS2	917	0.16	94	8.00E-62
958	271965707	hypothetical protein Sros_4262	Streptosporangium roseum	238	0.20	40	6

			DSM 43021				
1074	323139155	replication protein C	Methylocystis sp. ATCC 49242	444	0.60	68	1.00E-94
1083	254502301	hypothetical protein SADFL11_2339	Labrenzia alexandrii DFL-11	993	0.12	83	3.00E-43
1097	85708226	putative DNA methylase	Erythrobacter sp. NAP1	909	0.66	76	0
1128	323139416	protein of unknown function DUF188	Methylocystis sp. ATCC 49242	152	0.34	85	9.00E-17
1128	323139416	protein of unknown function DUF188	Methylocystis sp. ATCC 49242	152	0.50	87	8.00E-18
1171	323139522	hypothetical protein Met49242DRAFT_3956	Methylocystis sp. ATCC 49242	444	0.60	91	2.00E-138
1200	299133058	Putative helicase A859L	Afipia sp. 1NLS2	408	0.48	81	1.00E-89
1250	224107413	PREDICTED: hypothetical protein, partial	Taeniopygia guttata	369	0.22	33	4.6
1259	225156592	conserved hypothetical protein succinate dehydrogenase, hydrophobic membrane anchor	Opitutaceae bacterium TAV2	275	0.17	35	3.5
1290	298293273	protein	Starkeya novella DSM 506	136	0.84	54	1.00E-24
1292	194761572	GF15723	Drosophila ananassae	514	0.14	37	4.5
1302	316934586	hypothetical protein Rpxd1_3258	Rhodopseudomonas palustris DX-1	253	0.27	72	1.00E-18
1310	154495963	hypothetical protein BACCAP_00246	Bacteroides capillosus ATCC 29799	166	0.23	49	4.5
1324	323137236	DNA gyrase, B subunit	Methylocystis sp. ATCC 49242	810	0.07	81	1.00E-15
1352	195442410	GK17748	Drosophila willistoni	421	0.10	45	5.9
1357	323137126	glycosyl transferase group 1	Methylocystis sp. ATCC 49242	408	0.14	70	2.00E-14
1369	323136028	major facilitator superfamily MFS_1	Methylocystis sp. ATCC 49242	528	0.15	65	1.00E-20
1374	90420316	phosphomethylpyrimidine kinase	Aurantimonas manganoxydans SI85-9A1	317	0.06	75	5.9

1436	238023318	hypothetical protein bglu_2p0290	Burkholderia glumae BGR1	476	0.29	36	5.00E-13
1498	330966470	hypothetical protein PSYAC_17830	Pseudomonas syringae pv. actinidiae str. M302091	437	0.10	47	0.0009
1615	338820076	hypothetical protein Agau_P200245	Agrobacterium tumefaciens F2	783	0.11	47	6.00E-14
1691	332560244	Mg chelatase-related protein	Rhodobacter sphaeroides WS8N	512	0.11	38	4.5
1734	323137126	glycosyl transferase group 1	Methylocystis sp. ATCC 49242	408	0.24	77	9.00E-39
1787	299133059	Pseudomurein-binding repeat protein	Afipia sp. 1NLS2	686	0.13	99	9.00E-39
1787	299133059	Pseudomurein-binding repeat protein	Afipia sp. 1NLS2	686	0.09	98	9.00E-25
1793	323137126	glycosyl transferase group 1	Methylocystis sp. ATCC 49242	408	0.08	88	6.00E-08
1793	323137126	glycosyl transferase group 1	Methylocystis sp. ATCC 49242	408	0.08	88	6.00E-08
1836	150378510	tyrosine-protein phosphatase non-receptor type 22	Danio rerio	887	0.05	34	7.6
1856	17232905	hypothetical protein alr8532	Nostoc sp. PCC 7120	304	0.25	48	4.00E-11
1922	296534676	conserved hypothetical protein	Roseomonas cervicalis ATCC 49957	206	0.30	46	0.41
1947	163857359	LysR family transcriptional regulator	Bordetella petrii DSM 12804	295	0.15	68	7.00E-09
1953	332708113	penicillin-binding protein, family 1A	Lyngbya majuscula 3L	819	0.08	34	0.91
1961	323139158	integrase domain protein SAM domain protein	Methylocystis sp. ATCC 49242	347	0.10	74	2.00E-08
1970	154245910	succinate dehydrogenase cytochrome b556 subunit	Xanthobacter autotrophicus Py2	133	0.95	57	3.00E-32
1970	110635572	succinate dehydrogenase iron-sulfur subunit	Mesorhizobium sp. BNC1	259	0.25	88	2.00E-26
1976	197103308	hypothetical protein PHZ_p0168	Phenylobacterium zucineum HLK1	323	0.14	45	0.017
2036	291449786	LOW QUALITY PROTEIN:	Streptomyces albus J1074	880	0.11	36	3.5

FscRII

			Salmonella enterica subsp. enterica serovar Schwarzengrund str. CVM19633	321	0.17	49	6.00E-05
2044	194733871	antirestriction protein	Actinosynnema mirum DSM 43827	309	0.12	61	0.029
2048	256377452	AraC family transcriptional regulator	Actinosynnema mirum DSM 43827	309	0.12	61	0.029
2048	256377452	AraC family transcriptional regulator	Streptomyces sp. C	1616	0.03	43	0.71
2082	302539530	polyketide synthase type I	Afipia sp. 1NLS2	408	0.11	82	5.00E-13
2087	299133058	Putative helicase A859L	Chlamydomonas reinhardtii	439	0.08	51	7.8
2108	159489659	predicted protein	Roseomonas cervicalis ATCC 49957	206	0.32	38	4.6
2149	296534676	conserved hypothetical protein	Oscillatoria sp. PCC 6506	138	0.26	42	7.8
2151	300865706	conserved hypothetical protein	Gluconacetobacter diazotrophicus PA1 5	397	0.09	97	5.00E-05
2167	209544899	hypothetical protein Gdia_2780	peptidase U62 modulator of DNA Methylocystis sp. ATCC 49242	514	0.28	33	0.0002
2176	323138303	gyrase	Methylobacterium nodulans ORS 2060	173	0.23	75	6.00E-08
2236	220922655	hypothetical protein Mnod_2696	Rhodopseudomonas palustris DX-1	253	0.08	84	2.1
2392	316934586	hypothetical protein Rpxd1_3258	Methylocystis sp. ATCC 49242	150	0.11	88	1.2
2503	323139732	putative transposase	Dinoroseobacter shibae DFL 12	432	0.08	57	0.71
2606	159044846	hypothetical protein Dshi_2303	Desulfotomaculum acetoxidans DSM 771	377	0.10	74	3.00E-08
2614	258515430	hypothetical protein Dtox_2211	Oryza sativa Indica Group	532	0.08	44	2.7
2641	125524722	hypothetical protein OsI_00707	Populus trichocarpa	398	0.10	46	6
2690	222871727	predicted protein	Methylocystis sp. ATCC 49242	397	0.15	81	7.00E-20
2750	323136026	secretion protein HlyD family protein					

2832	329889268	phosphotransferase domain-containing protein	Brevundimonas diminuta ATCC 11568	1279	0.03	54	0.01
2854	338778752	glutathione S-transferase	Achromobacter xylosoxidans AXX-A	201	0.22	73	5.00E-10
2901	323139705	hypothetical protein Met49242DRAFT_4133	Methylocystis sp. ATCC 49242	253	0.31	48	1.00E-11
2903	28897417	hypothetical protein VP0643	Vibrio parahaemolyticus RIMD 2210633	599	0.05	45	10
2975	323137126	glycosyl transferase group 1	Methylocystis sp. ATCC 49242	408	0.09	89	3.00E-11
2983	297620720	nitric oxide reductase, subunit B	Waddlia chondrophila WSU 86-1044	763	0.04	85	6.00E-10
3047	340030375	helix-turn-helix domain-containing protein	Paracoccus sp. TRP	81	0.64	60	6.00E-08
3093	302134922	hypothetical protein PsyrptN_26257	Pseudomonas syringae pv. tomato NCPPB 1108	988	0.04	53	0.92
3111	329889268	phosphotransferase domain-containing protein	Brevundimonas diminuta ATCC 11568	1279	0.03	51	0.022
3112	323139523	hypothetical protein Met49242DRAFT_3957	Methylocystis sp. ATCC 49242	482	0.09	88	7.00E-06
3123	297292023	PREDICTED: solute carrier family 22 member 3-like	Macaca mulatta	715	0.06	44	7.8
3128	334195582	putative osmotically inducible lipoprotein b1 transmembrane (osmB)	Ralstonia solanacearum Po82	296	0.17	45	3.5
3130	254500001	hypothetical protein SADFL11_39	Labrenzia alexandrii DFL-11	356	0.20	54	5.00E-12
3130	254500001	hypothetical protein SADFL11_39	Labrenzia alexandrii DFL-11	356	0.19	56	1.00E-11
3162	301780690	PREDICTED: lysine-specific demethylase 4D-like	Ailuropoda melanoleuca	487	0.08	49	7.8
3169	296123721	hypothetical protein Plim_3487	Planctomyces limnophilus DSM 3776	438	0.07	47	7.8
3190	149918280	RNA methyltransferase	Plesiocystis pacifica SIR-1	485	0.12	40	4.6
3190	149918280	RNA methyltransferase	Plesiocystis pacifica SIR-1	485	0.12	40	4.6

3191	91200065	conserved hypothetical protein hypothetical protein	Candidatus Kuenenia stuttgartiensis	221	0.24	46	0.11
3192	302813543	SELMODRAFT_427133 putative ADP-	Selaginella moellendorffii	270	0.15	48	0.7
3236	333027859	ribosylation/Crystallin J1	Streptomyces sp. Tu6071	893	0.06	44	7.6
3258	110347006	hypothetical protein Meso_4194	Mesorhizobium sp. BNC1	366	0.16	62	4.00E-12
3315	288963127	transposase	Azospirillum sp. B510	267	0.25	67	2.00E-16
3340	3059133	transposase IS1355	Methylobacterium extorquens DM4	179	0.22	87	4.00E-12
3398	83592164	CRISPR-associated helicase Cas3 family protein	Rhodospirillum rubrum ATCC 11170	752	0.05	69	5.00E-05
3602	323137054	hypothetical protein Met49242DRAFT_1521	Methylocystis sp. ATCC 49242	73	0.59	77	4.00E-09
3602	197103644	RNA polymerase sigma 54 subunit, RpoN	Phenylobacterium zucineum HLK1	499	0.11	44	2.7
3605	218667529	hypothetical protein AFE_2089	Acidithiobacillus ferrooxidans ATCC 23270	253	0.42	29	4.5
3656	121596846	RNA polymerase sigma factor	Burkholderia mallei SAVP1	231	0.26	51	3.00E-08
3677	299133058	Putative helicase A859L	Afipia sp. 1NLS2	408	0.27	85	2.00E-28
3742	90424677	plasmid maintenance system killer protein	Rhodopseudomonas palustris BisB18	93	0.31	41	1.4
3786	229828371	hypothetical protein GCWU000342_00430	Shuttleworthia satelles DSM 14600	298	0.15	40	7.9
3786	229828371	hypothetical protein GCWU000342_00430	Shuttleworthia satelles DSM 14600	298	0.15	40	7.7
3891	220922092	NnrS family protein	Methylobacterium nodulans ORS 2060	392	0.09	68	2.00E-05
3901	221487851	conserved hypothetical protein Predicted metal-dependent amidohydrolase with the TIM-	Toxoplasma gondii GT1	3229	0.02	39	4.5
3981	85711049	barrel fold	Idiomarina baltica OS145	558	0.23	26	3.5
3981	89067955	type I secretion target repeat protein	Oceanicola granulosus	818	0.17	28	4.2

			HTCC2516				
4000	323136327	Patatin	Methylocystis sp. ATCC 49242	284	0.44	61	3.00E-37
4020	328541621	hypothetical protein SL003B_p0053	Polymorphum gilvum SL003B-26A1	435	0.43	58	2.00E-56
4035	108763646	putative lipoprotein	Myxococcus xanthus DK 1622	447	0.13	35	7.8
4187	288957870	transposase	Azospirillum sp. B510	460	0.09	43	5.9
4195	335035794	hypothetical protein AGRO_3125	Agrobacterium sp. ATCC 31749	323	0.09	62	0.55
4205	90424421	transposase IS3/IS911	Rhodopseudomonas palustris BisB18	66	0.38	68	1.6
4305	329928833	hypothetical protein HMPREF9412_3689	Paenibacillus sp. HGF5	694	0.05	94	1.00E-10
4311	167621675	IS66 Orf2 family protein	Caulobacter sp. K31	117	0.35	59	3.00E-05
4338	209886756	hypothetical protein OCAR_7650	Oligotropha carboxidovorans OM5	362	0.10	64	2.00E-05
4381	284989640	hypothetical protein Gobs_1064	Geodermatophilus obscurus DSM 43160	285	0.11	59	2.7
4492	13475179	hypothetical protein mlr6199	Mesorhizobium loti MAFF303099	240	0.18	72	1.00E-09
4538	238061000	5'-nucleotidase domain-containing protein	Micromonospora sp. ATCC 39149	616	0.13	35	7.4
4646	302522431	LOW QUALITY PROTEIN: ABC transporter ATP-binding protein	Streptomyces sp. SPB78	564	0.15	30	4.5
4866	299133061	DNA modification methyltransferase-related protein	Afipia sp. 1NLS2	917	0.14	90	3.00E-58
4903	299741957	endoprotease	Coprinopsis cinerea okayama7#130	613	0.07	44	4.5
4908	323139523	hypothetical protein Met49242DRAFT_3957	Methylocystis sp. ATCC 49242	482	0.08	97	2.00E-14
4992	323135620	hypothetical protein Met49242DRAFT_0090	Methylocystis sp. ATCC 49242	153	0.23	91	2.00E-11

5003	299133061	DNA modification methyltransferase-related protein	Afipia sp. 1NLS2	917	0.04	95	6.00E-14
5069	316934586	hypothetical protein Rpx1_3258	Rhodopseudomonas palustris DX-1	253	0.39	88	1.00E-44
5103	312883142	hypothetical protein VIBC2010_07634	Vibrio caribbenthicus ATCC BAA-2122	138	0.85	38	1.00E-14
5192	16125845	sulfate ABC transporter ATP- binding protein	Caulobacter crescentus CB15	359	0.12	55	0.007
5192	86739180	ABC transporter-like protein	Frankia sp. Ccl3	385	0.13	51	0.017
5210	323139955	transposase IS204/IS1001/IS1096/IS1165 family protein	Methylocystis sp. ATCC 49242	549	0.05	54	4.7
5248	170743671	putative transposase	Methylobacterium sp. 4-46	373	0.08	71	0.001
5420	323139641	protein of unknown function DUF1403	Methylocystis sp. ATCC 49242	345	0.10	61	0.9

1
2
3