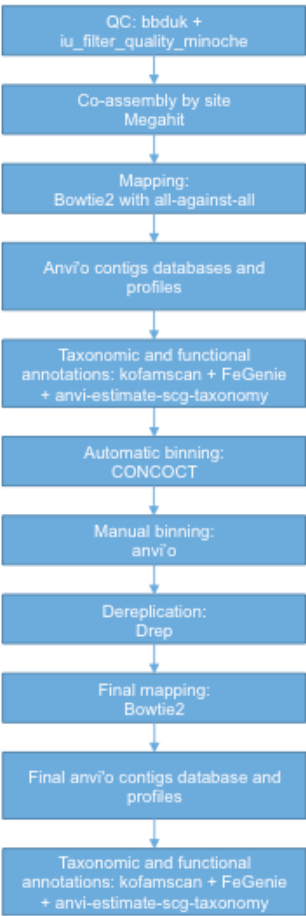


Supplementary Information



Additional file 1. Scheme of the bioinformatics pipeline. (PNG 45kb)

Site	Sample		Total pairs raw reads	Total pairs trimmed reads	Total pairs filtered reads (iu. minoché)	Percent of all passed pairs	Number of contigs > 1kb	RE3	RE7	RE5	RE6	RE12	RE13
Rainbow	RE12		129,416,702	129,416,134	119,375,209	92	53,608	14.63%	15.06%	44.31%	35.52%	83.35%	91.99%
	RE13		119,361,816	119,361,557	111,192,767	93							
TAG	RE5	RE5_SEQ1	39,419,336	148,102,297	105,351,977	71	64,689	19.99%	18.30%	68.14%	52.79%	30.28%	35.21%
		RE5_SEQ2	108,684,665										
	RE6	RE6_SEQ1	34,052,972	126,914,271	75,047,098	59							
		RE6_SEQ2	92,866,098										
Snake Pit	RE7		38,308,167	38,298,694	29,570,786	77	34,823	37.35%	39.48%	35.23%	32.86%	20.34%	21.32%
	RE3		36,460,685	36,458,247	22,233,459	61							

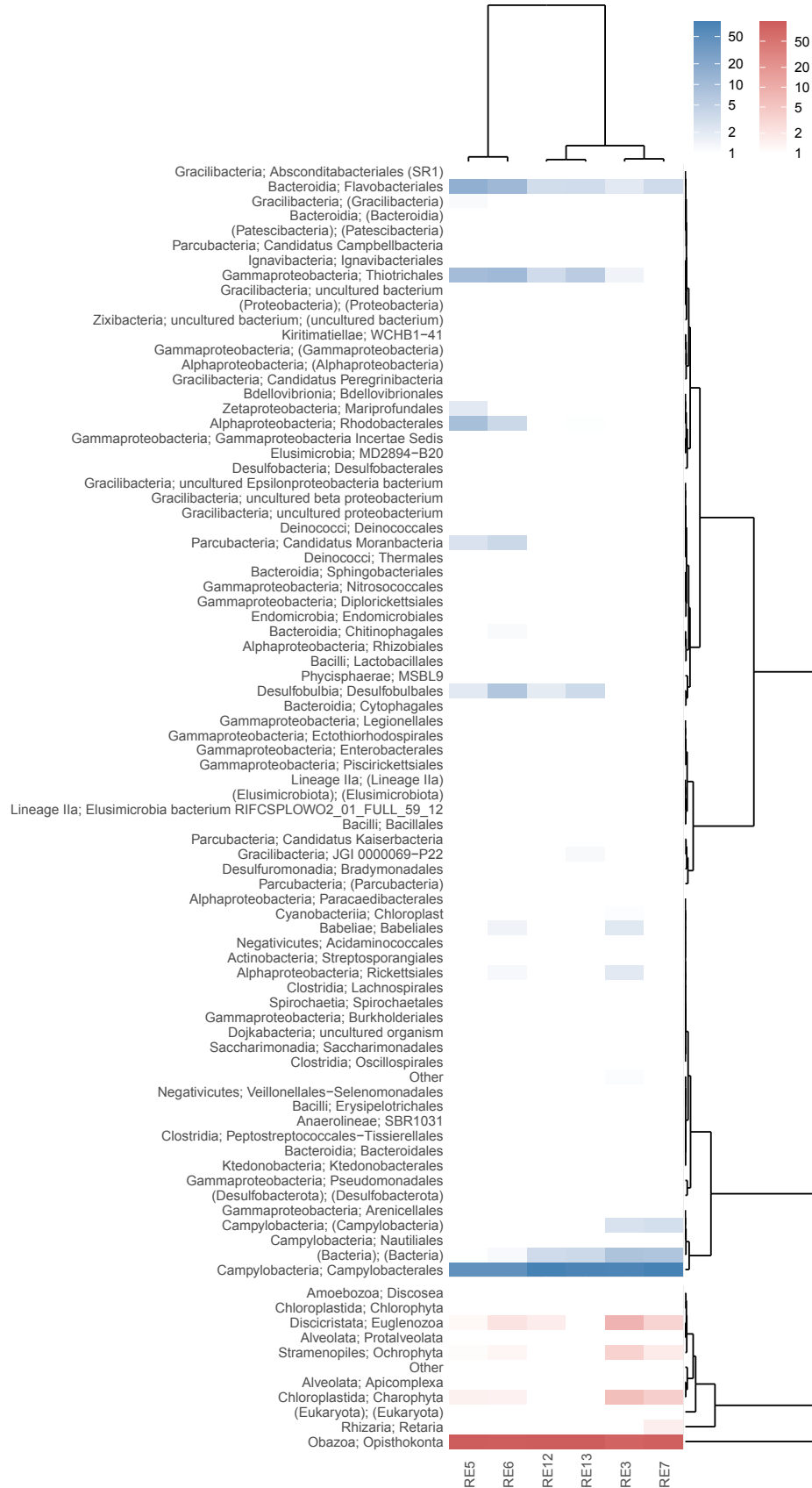
Additional file 2. Quality trimming and filtering statistics for each individual, number of assembled contigs longer than 1 kbp for each site, and number of reads recruited by these contigs for each metagenome. (XLSX 11 kb)

Name	% completion	% redundancy	number of contigs	total length	classification
RB_MAG_00001	98.59	0	195	2287451	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__Flavobacteriaceae;g__LPB0138;s__
RB_MAG_00002	97.18	0	73	1692263	d__Bacteria;p__Campylobacterota;c__Campylobacteria;o__Campylobacteriales;f__Sulfurovaceae;g__Nitratifactor;s__
RB_MAG_00003	97.18	2.82	181	2850013	d__Bacteria;p__Bacteroidota;c__Ignavibacteria;o__Ignavibacteriales;f__Melioribacteraceae;g__1-14-2-50-31-20;s__
RB_MAG_00004	94.37	1.41	230	3591863	d__Bacteria;p__Desulfobacterota;c__Desulfobulbia;o__Desulfobulbales;f__Desulfocapsaceae;g__Desulforhopalus;s__
RB_MAG_00005	97.18	4.23	163	3532123	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__g__s__
RB_MAG_00006	92.96	0	328	3829142	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Chitinophagales;f__Saprospiraceae;g__s__
RB_MAG_00007	94.37	2.82	335	2876824	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__UBA10066;g__s__
RB_MAG_00008	92.96	2.82	162	1853475	d__Bacteria;p__Proteobacteria;c__Zetaproteobacteria;o__Mariprofundales;f__Mariprofundaceae;g__Ghiorsea;s__
RB_MAG_00009	92.96	2.82	214	2497075	d__Bacteria;p__Verrucomicrobiota;c__Kiritimatiellae;o__UBA8416;f__g__s__
RB_MAG_00010	88.73	1.41	269	1902201	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__Flavobacteriaceae;g__s__
RB_MAG_00011	94.37	7.04	148	1980870	d__Bacteria;p__Campylobacterota;c__Campylobacteria;o__Campylobacteriales;f__Sulfurimonadaceae;g__s__
RB_MAG_00012	91.55	5.63	75	2807279	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__UBA3824;g__s__
RB_MAG_00013	84.51	2.82	47	618108	d__Bacteria;p__Patescibacteria;c__Paceibacteria;o__UBA9983_A;f__UBA1006;g__UBA1006;s__
RB_MAG_00014	87.32	5.63	524	2669979	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Marinosulfonomonas;s__
RB_MAG_00015	80.28	0	188	1686342	d__Bacteria;p__Campylobacterota;c__Campylobacteria;o__Campylobacteriales;f__Sulfurovaceae;g__s__
RB_MAG_00016	83.1	2.82	383	2620804	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Marinosulfonomonas;s__
RB_MAG_00017	78.87	0	154	978606	d__Bacteria;p__Patescibacteria;c__Gracilibacteria;o__BD1-5;f__UBA6164;g__UBA6489;s__
RB_MAG_00018	83.1	4.23	403	2787078	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__s__
RB_MAG_00019	76.06	0	54	884939	d__Bacteria;p__Patescibacteria;c__Gracilibacteria;o__GCA-2401425;f__g__s__
RB_MAG_00020	74.65	1.41	128	903611	d__Bacteria;p__Patescibacteria;c__Gracilibacteria;o__UBA4473;f__UBA4473;g__s__
RB_MAG_00021	74.65	2.82	134	2453758	d__Bacteria;p__Desulfobacterota;c__Desulfobulbia;o__Desulfobulbales;f__Desulfocapsaceae;g__Desulfocapsa;s__
RB_MAG_00022	71.83	1.41	22	624325	d__Bacteria;p__Patescibacteria;c__Paceibacteria;o__UBA9983_A;f__GCA-2747955;g__s__
RB_MAG_00024	66.2	0	65	425625	d__Bacteria;p__Patescibacteria;c__Paceibacteria;o__UBA9983_A;f__UBA1006;g__UBA1006;s__
RB_MAG_00025	70.42	5.63	309	1526674	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Thiotrichales;f__Thiotrichaceae;g__s__
RB_MAG_00026	63.38	0	412	1485315	d__Bacteria;p__Mycococcota;c__Bradimonadia;o__Bradymonadales;f__g__s__
RB_MAG_00027	64.79	2.82	114	640890	d__Bacteria;p__Patescibacteria;c__Gracilibacteria;o__BD1-5;f__UBA6164;g__s__
RB_MAG_00028	54.93	2.82	275	2568175	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__UBA3824;g__s__
RB_MAG_00029	46.48	0	280	2059101	d__Bacteria;p__Planctomycetota;c__SZUA-567;o__SZUA-567;f__g__s__
TAG_MAG_00001	100	4.23	203	3078853	d__Bacteria;p__Proteobacteria;c__Ignavibacteria;o__Ignavibacteriales;f__Melioribacteraceae;g__1-14-2-50-31-20;s__
TAG_MAG_00002	92.96	0	200	1454508	d__Bacteria;p__Campylobacterota;c__Campylobacteria;o__Campylobacteriales;f__Sulfurovaceae;g__Nitratifactor;s__
TAG_MAG_00003	91.55	0	203	2239806	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__s__
TAG_MAG_00004	94.37	2.82	113	2753715	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__UBA3824;g__s__
TAG_MAG_00005	91.55	1.41	332	2636589	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Marinosulfonomonas;s__
TAG_MAG_00006	95.77	5.63	220	3156654	d__Bacteria;p__Desulfobacterota;c__Desulfobulbia;o__Desulfobulbales;f__Desulfocapsaceae;g__Desulfocapsa;s__
TAG_MAG_00007	88.73	0	71	1201735	d__Bacteria;p__Patescibacteria;c__Paceibacteria_A;o__Moranbacterales;f__GCA-2747515;g__s__
TAG_MAG_00008	87.32	0	222	3066709	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__s__
TAG_MAG_00009	90.14	2.82	191	2717585	d__Bacteria;p__Deinococcota;c__Deinococci;o__Deinococcales;f__Trueperaceae;g__s__
TAG_MAG_00010	84.51	0	339	3037540	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Chitinophagales;f__Saprospiraceae;g__s__
TAG_MAG_00011	85.92	1.41	309	2210069	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__Flavobacteriaceae;g__s__
TAG_MAG_00012	87.32	2.82	256	1807310	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__Flavobacteriaceae;g__LPB0138;s__
TAG_MAG_00013	90.14	7.04	311	2630970	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__UBA3824;g__s__
TAG_MAG_00014	94.37	2.82	204	1625457	d__Bacteria;p__Proteobacteria;c__Zetaproteobacteria;o__Mariprofundales;f__Mariprofundaceae;g__Ghiorsea;s__
TAG_MAG_00015	81.69	1.41	320	2139890	d__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Marinosulfonomonas;s__
TAG_MAG_00016	81.69	1.41	223	2695452	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__UBA10066;g__s__
TAG_MAG_00017	77.46	1.41	322	2228076	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Chitinophagales;f__Saprospiraceae;g__s__
TAG_MAG_00018	74.65	0	124	841694	d__Bacteria;p__Patescibacteria;c__Gracilibacteria;o__GCA-2401425;f__g__s__
TAG_MAG_00019	71.83	0	396	1614305	d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Thiotrichales;f__Thiotrichaceae;g__s__
TAG_MAG_00020	66.2	1.41	53	775413	d__Bacteria;p__Patescibacteria;c__Gracilibacteria;o__GCA-2401425;f__g__s__
TAG_MAG_00021	61.97	0	139	1781918	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Flavobacteriales;f__Flavobacteriaceae;g__s__

27 **Additional file 3.** MAG collection, estimates of completion and redundancy calculated
 28 based on the occurrence of Single-copy Core Genes (SCG), number of contigs, total bin
 29 length, and taxonomic affiliation from TAG (above) and Rainbow (below) determined by
 30 the Genome Taxonomy Database (GTDB). (XLSX 49 kb)

bins	RE12	RE13	RE5	RE6
RB_MAG_00001	156.4473	158.9698	64.03198	18.61443502
RB_MAG_00002	77.57325	75.72071	8.825798	7.55483338
RB_MAG_00003	22.24228	37.88509	2.713664	1.883120182
RB_MAG_00004	31.31677	60.61314	3.365832	3.169692998
RB_MAG_00005	13.4421	28.16437	0.343008	0.219799537
RB_MAG_00006	17.06394	51.59705	13.85091	10.42769163
RB_MAG_00007	7.183664	7.200721	0.18203	0.053155494
RB_MAG_00008	26.65745	29.77692	17.63489	2.047864147
RB_MAG_00009	25.57316	58.14429	16.25311	12.6193186
RB_MAG_00010	22.20553	34.91068	24.17086	7.485347763
RB_MAG_00011	104.8823	114.8834	0.449283	0.287015806
RB_MAG_00012	52.08004	73.96002	9.976201	5.441842083
RB_MAG_00013	5.599381	58.47288	1.229871	0.547645395
RB_MAG_00014	19.37331	63.77732	453.0661	69.05732742
RB_MAG_00015	304.217	295.915	25.47229	14.16088848
RB_MAG_00016	7.051459	12.37643	44.02728	7.693223911
RB_MAG_00017	0.82676	18.34323	0.04853	0.012800862
RB_MAG_00018	32.07764	43.50831	27.03028	3.411206288
RB_MAG_00019	33.90195	23.46798	1.329736	0.713297753
RB_MAG_00020	3.839349	6.045185	2.944681	0.242795849
RB_MAG_00021	6.097366	12.36709	0.139577	0.30905941
RB_MAG_00022	68.93723	50.36893	19.19539	22.39274897
RB_MAG_00024	6.776127	10.67942	0.768865	0.239765051
RB_MAG_00025	551.5875	858.8045	321.3752	261.5608421
RB_MAG_00026	0.904733	5.652469	0.018581	0.006348148
RB_MAG_00027	81.1845	40.48893	0.810655	0.13889435
RB_MAG_00028	0.900961	10.66044	0.05966	0.022457192
RB_MAG_00029	3.360645	7.38431	0.014344	0.012720114
TAG_MAG_00001	2.124775	3.643299	23.13496	9.919818192
TAG_MAG_00002	12.68722	11.69342	19.09132	17.08611503
TAG_MAG_00003	0.075278	0.352676	76.3824	22.92086636
TAG_MAG_00004	4.022307	5.232831	177.2273	90.58007964
TAG_MAG_00005	2.973642	5.262176	81.50766	14.29566345
TAG_MAG_00006	167.5042	268.17	113.8137	242.3805739
TAG_MAG_00007	0.490094	0.507284	361.5977	205.7078857
TAG_MAG_00008	14.22142	19.77789	130.633	14.83413131
TAG_MAG_00009	0.147581	2.671891	26.91085	14.74947573
TAG_MAG_00010	0.48259	0.749427	19.20688	15.60296852
TAG_MAG_00011	8.723665	7.379724	33.70549	13.32555952
TAG_MAG_00012	27.06787	26.574	149.3246	41.8877586
TAG_MAG_00013	1.701309	0.980936	14.2533	23.9962147
TAG_MAG_00014	2.522233	2.560178	155.3632	19.13003297
TAG_MAG_00015	10.23453	33.24781	1273.575	180.4320619
TAG_MAG_00016	0.274866	0.282543	22.57392	2.346675066
TAG_MAG_00017	1.831955	3.357543	42.7589	33.20564963
TAG_MAG_00018	10.20827	10.59338	273.3632	94.93974889
TAG_MAG_00019	276.4497	426.6543	598.6925	489.6683161
TAG_MAG_00020	0.413676	0.308379	33.341	21.71860028
TAG_MAG_00021	0.406056	0.401242	60.98623	17.30322495

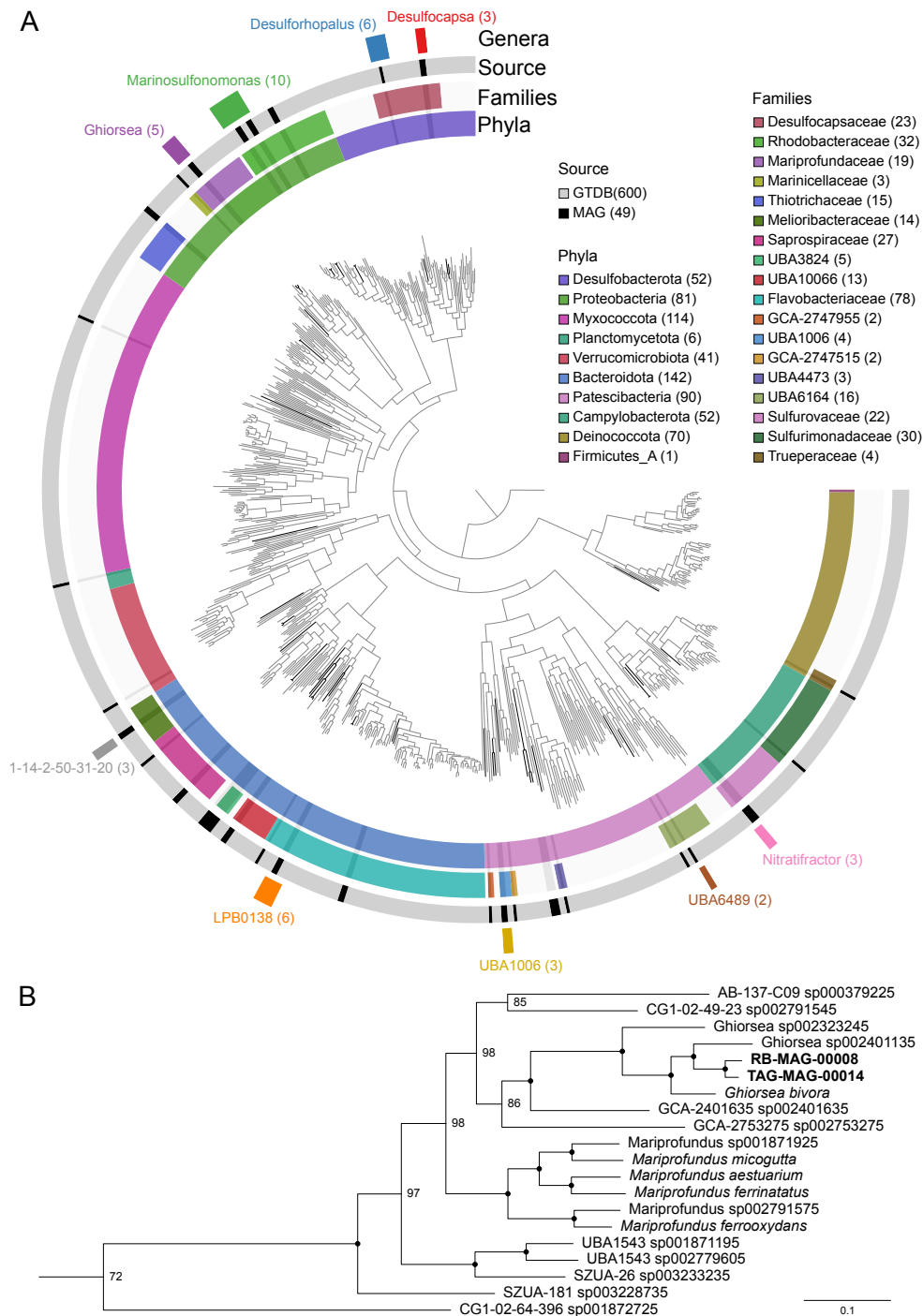
- 33 **Additional file 4.** Average coverage of MAGs at Rainbow (RB) and TAG sites after
34 MAG dereplication. (XLSX 59 kb)
35
36
37 **Additional file 5.** Percent of total reads mapping to each MAG. (XLSX 31 kb)



Additional File 6. PhyloFlash analysis results. Heatmap of taxonomic assignments (rows) for small-subunit rRNA reads in the six individual metagenomes (columns). The plot was generated using with the comparison script provided with phyloFlash. Color intensities represent the percentage of reads mapping to a given taxon, separated by prokaryotes (blue) and eukaryotes (red). Samples are clustered by their similarity in terms of taxonomic content and taxa are clustered by their co-occurrence across samples. Clustering are based on the euclidean distance and on the Ward's minimum variance method. (PDF 39 kb).

query acc.ver	subject acc.ver	% identity	alignment length	mismatches	gap opens	q. start	q. end	s. start	s. end	evalue	bit score
RE12.PFspades_2_2.719641	RE13.PFspades_3_2.987402	99.935	1538	1	0	1	1538	1	1538	0	2835
RE12.PFspades_2_2.719641	RE6.PFspades_7_5.799881	97.745	1552	21	6	1	1538	1	1552	0	2660
RE12.PFspades_2_2.719641	RE5.PFspades_8_47.722283	97.745	1552	21	6	1	1538	1	1552	0	2660
RE12.PFspades_2_2.719641	TAG_MAG_00014_16SrRNAs	97.729	1541	21	6	9	1535	1	1541	0	2639
RE13.PFspades_3_2.987402	RE12.PFspades_2_2.719641	99.935	1538	1	0	1	1538	1	1538	0	2835
RE13.PFspades_3_2.987402	RE6.PFspades_7_5.799881	97.809	1552	20	6	1	1538	1	1552	0	2665
RE13.PFspades_3_2.987402	RE5.PFspades_8_47.722283	97.809	1552	20	6	1	1538	1	1552	0	2665
RE13.PFspades_3_2.987402	TAG_MAG_00014_16SrRNAs	97.794	1541	20	6	9	1535	1	1541	0	2645
RE5.PFspades_8_47.722283	RE6.PFspades_7_5.799881	100	1552	0	0	1	1552	1	1552	0	2867
RE5.PFspades_8_47.722283	TAG_MAG_00014_16SrRNAs	100	1541	0	0	9	1549	1	1541	0	2846
RE5.PFspades_8_47.722283	RE13.PFspades_3_2.987402	97.809	1552	20	6	1	1552	1	1538	0	2665
RE5.PFspades_8_47.722283	RE12.PFspades_2_2.719641	97.745	1552	21	6	1	1552	1	1538	0	2660
RE6.PFspades_7_5.799881	RE5.PFspades_8_47.722283	100	1552	0	0	1	1552	1	1552	0	2867
RE6.PFspades_7_5.799881	TAG_MAG_00014_16SrRNAs	100	1541	0	0	9	1549	1	1541	0	2846
RE6.PFspades_7_5.799881	RE13.PFspades_3_2.987402	97.809	1552	20	6	1	1552	1	1538	0	2665
RE6.PFspades_7_5.799881	RE12.PFspades_2_2.719641	97.745	1552	21	6	1	1552	1	1538	0	2660
TAG_MAG_00014_16SrRNAs	RE6.PFspades_7_5.799881	100	1541	0	0	1	1541	9	1549	0	2846
TAG_MAG_00014_17SrRNAs	RE5.PFspades_8_47.722283	100	1541	0	0	1	1541	9	1549	0	2846
TAG_MAG_00014_18SrRNAs	RE13.PFspades_3_2.987402	97.794	1541	20	6	1	1541	9	1535	0	2645
TAG_MAG_00014_19SrRNAs	RE12.PFspades_2_2.719641	97.729	1541	21	6	1	1541	9	1535	0	2639

Additional File 7. BLAST of 16S rRNA genes for *Zetaproteobacteria*. BLASTN search were performed between the 16S rRNA sequences affiliated to the *Zetaproteobacteria* assembled using SPAdes in phyloFlash and the 16S rRNA sequence retrieved from TAG_MAG_00014 (DOC 128 kb).



Additional File 8: A. Maximum-likelihood tree based on concatenated marker proteins according to the GTDB-Tk genome phylogeny visualized using anvi'o. Tree includes 600 genomes from GTDB and 49 MAGs covering mostly unknown genera, highlighting the importance of lineages lacking representatives. A single Firmicutes was used to root the

tree. The bars in the innermost circular layer show the phylum affiliation of each genome. The second layer represents the family affiliation. The third layer marks genomes as either MAGs from our study (49, black) or genomes from GTDB (grey). The outermost layer shows the genus affiliation (10) or the lack thereof (19) of our MAGs. Only the families and genera observed in the MAGs are shown. **B.** Zoom inset of the *Zetaproteobacteria* phylogenetic relationships visualized using FigTree. Nodes represented by a dot indicate a bootstrap value of 100; lower values are specified. (PDF 87 kb).

Additional File 9. Key gene predictions from Zetaproteobacterial MAGs

76

genome (NCBI accession number GCF_000744415.1) using RAST and FeGenie

(indicated by a star). NA: ‘Not Available’. (XLSX 45 kb)

Additional File 10. Table containing the names of all genes found per MAG using

KEGG annotation. (XLSX 4.3Mb)

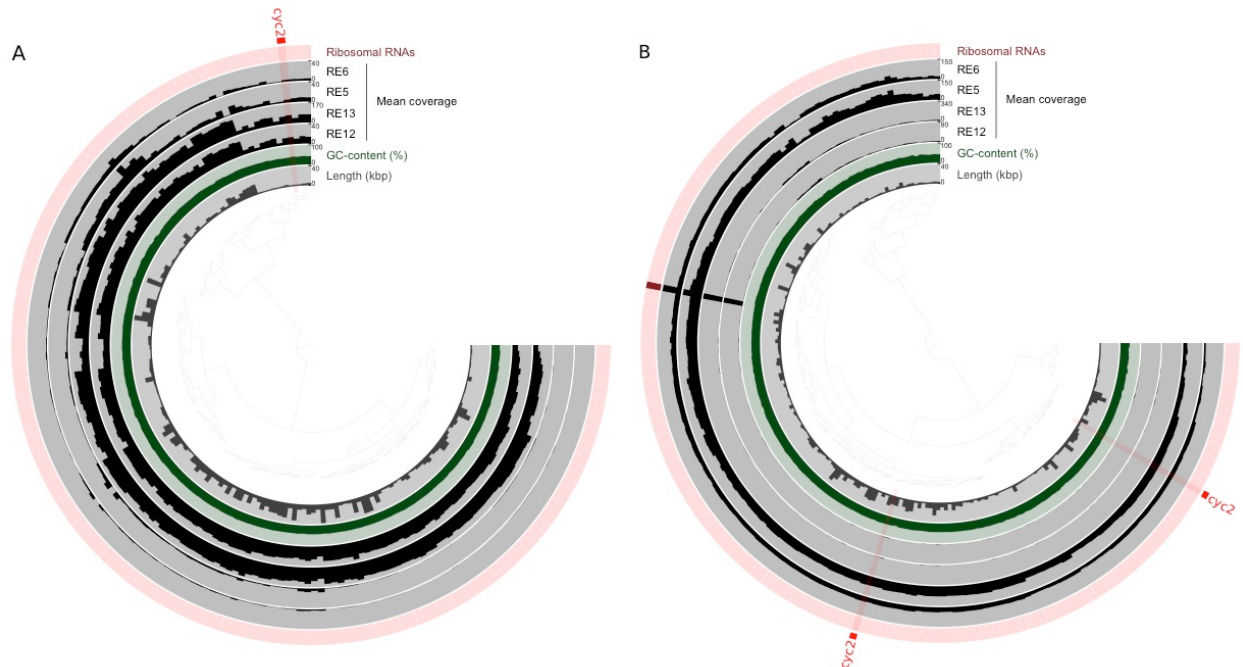
	iron_aquisition- iron_transport	iron_aquisition- heme_transport	iron_aquisition- heme_oxygenase	iron_aquisition- siderophore_ synthesis	iron_aquisition- siderophore_ transport	iron_gene_ regulation	iron_oxidation	possible_iron_ oxidation_and_ possible_iron_ reduction	probable_iron_ reduction	iron_reduction	iron_storage	magnetosome_ formation
Ghiorsea_bivora.fa	0	0	0	0	9	1	2	0	0	0	0	0
Mariprofundus.fa	2	0	0	0	14	1	1	0	0	0	0	0
RB_MAG_00001-contigs.fa	2	0	0	0	11	5	0	0	0	0	0	0
RB_MAG_00002-contigs.fa	5	0	0	0	7	0	0	0	0	0	0	0
RB_MAG_00003-contigs.fa	2	0	0	0	13	0	0	0	0	2	1	0
RB_MAG_00004-contigs.fa	3	1	0	0	14	1	0	0	0	0	3	0
RB_MAG_00005-contigs.fa	2	1	0	0	9	4	0	0	0	0	2	0
RB_MAG_00006-contigs.fa	5	1	0	0	17	4	0	0	0	0	0	0
RB_MAG_00007-contigs.fa	2	2	0	0	9	3	0	0	0	0	1	0
RB_MAG_00008-contigs.fa	0	0	0	0	14	2	1	0	0	0	0	0
RB_MAG_00009-contigs.fa	2	1	0	0	12	3	0	0	0	0	1	0
RB_MAG_00010-contigs.fa	4	1	0	0	8	1	0	0	0	0	1	0
RB_MAG_00011-contigs.fa	6	1	0	0	13	0	0	0	0	0	0	0
RB_MAG_00012-contigs.fa	0	0	0	0	3	5	0	0	0	0	0	0
RB_MAG_00014-contigs.fa	2	0	0	0	5	1	0	0	0	0	0	0
RB_MAG_00015-contigs.fa	2	0	0	0	3	1	0	0	0	0	0	0
RB_MAG_00016-contigs.fa	4	0	0	0	2	2	0	0	0	0	0	0
RB_MAG_00018-contigs.fa	2	0	0	0	2	1	0	0	0	0	1	0
RB_MAG_00019-contigs.fa	0	0	0	0	0	1	0	0	0	0	0	0
RB_MAG_00020-contigs.fa	0	0	0	0	0	3	0	0	0	0	0	0
RB_MAG_00021-contigs.fa	3	1	0	0	15	0	0	0	0	0	1	0
RB_MAG_00025-contigs.fa	0	0	0	0	7	0	0	0	0	0	0	0
RB_MAG_00026-contigs.fa	0	0	0	0	3	1	0	0	0	0	0	0
RB_MAG_00028-contigs.fa	0	0	0	0	5	1	0	0	0	0	0	0
RB_MAG_00029-contigs.fa	2	0	0	0	9	1	0	0	0	0	1	0
TAG_MAG_00001-contigs.fa	2	0	0	0	13	4	0	0	0	3	1	0
TAG_MAG_00002-contigs.fa	4	0	0	0	4	0	0	0	0	0	0	0
TAG_MAG_00003-contigs.fa	2	0	0	0	21	0	0	0	0	0	0	0
TAG_MAG_00004-contigs.fa	0	0	0	0	5	4	0	0	0	0	1	0
TAG_MAG_00005-contigs.fa	4	0	0	0	2	2	0	0	0	0	0	0
TAG_MAG_00006-contigs.fa	6	1	0	0	7	1	0	0	0	0	2	0
TAG_MAG_00007-contigs.fa	0	0	0	0	2	1	0	0	0	0	0	0
TAG_MAG_00008-contigs.fa	2	0	0	0	4	1	0	0	0	0	1	0
TAG_MAG_00009-contigs.fa	10	1	0	0	6	2	0	0	0	0	0	0
TAG_MAG_00010-contigs.fa	4	1	0	0	6	9	0	0	0	0	0	0
TAG_MAG_00011-contigs.fa	2	0	0	0	9	3	0	0	0	0	0	0
TAG_MAG_00012-contigs.fa	2	0	0	0	7	4	0	0	0	0	0	0
TAG_MAG_00013-contigs.fa	0	0	0	0	0	3	0	0	0	0	0	0
TAG_MAG_00014-contigs.fa	0	0	0	0	9	0	2	0	0	0	0	0
TAG_MAG_00015-contigs.fa	2	0	0	0	2	2	0	0	0	0	0	0
TAG_MAG_00016-contigs.fa	0	1	0	0	3	0	0	0	0	0	1	0
TAG_MAG_00017-contigs.fa	7	1	0	0	11	2	0	0	0	0	0	0
TAG_MAG_00019-contigs.fa	2	0	0	0	8	1	0	0	0	0	0	0
TAG_MAG_00020-contigs.fa	2	0	0	0	0	1	0	0	0	0	0	0
TAG_MAG_00021-contigs.fa	2	0	0	0	5	3	0	0	0	0	0	0

Additional file 11. Iron genes and gene clusters identified by FeGenie for the 49 MAGs

and reference genomes *Ghiorsea bivora* and *Mariprofundus ferrooxydans* PV-1 (NCBI

accession number GCF_000744415.1 and GCF_000153765.1). *Cyc2* genes were retrieved

in both MAGs, confirming they have the potential to oxidize Fe (II) (XLSX 41 kb)



Additional File 12. Differential coverage of contigs within *Zetaproteobacteria* bins.

Static image from the anvi'o refine display for **A.** RB_MAG_00008 and **B.**

TAG_MAG_00014. From inner to outer layers: clustering based on sequence

composition and differential coverage with Euclidian distance and Ward clustering

method, length layer (shows the actual length of a split), auxiliary layer with information

about contigs stored in the contig database (GC-content), four view layers with

information about MAGs across samples stored in the profile database (mean coverage),

and Ribosomal RNA presence. Splits containing the *cyc2* genes are highlighted in red.

(PDF 129kb).

Additional File 13. Analysis supplement: analysis details of metabolic potential among *Rimicaris* epibionts as depicted in Figure 5 (DOCX)

Based on the 49 obtained MAGs, we identified the following capabilities in the shrimp microbiome.

Autotrophic carbon fixation

Eight MAGs belonging to *Gamma*-, *Zeta*-, and *Alphaproteobacteria* showed the capacity to fix carbon through the CBB cycle at both sites. In contrast, only two *Campylobacteria* *Sulfurovaceae* MAGs from Rainbow and TAG harbored the complete set of genes for the rTCA cycle. *Sulfurovaceae* TAG MAG 00002 did not contain rTCA cycle genes, nor did *Sulfurimonadaceae* RB MAG 00011. This is probably due to incomplete MAGs, as both lineages are known to share this cycle. The *Desulfocapsaceae* MAGs were the only ones harboring the Wood-Ljungdahl pathway genes for carbon assimilation at both TAG and Rainbow.

Glycolysis, TCA cycle and carbon degradation

Almost all the MAGs had the capacity for glycolysis and 43 showed the potential for the TCA cycle. We noted the potential for carbon degradation in 35 MAGs from both sites using glucoamylase (1 MAG), D-galacturonate epimerase (20 MAGs), D-galacturonate isomerase (4 MAGs), chitinase (5 MAGs), beta-N-acetylhexosaminidase (22 MAGs), beta-glucosidase (1 MAG) and pullulanase enzymes (6 MAGs). It is notable that chitinases could be used to graze on degrading chitin shrimp molts under heterotrophic

metabolism, but bacteria can also secrete these enzymes for cell adhesion during pathogenesis and symbiosis [1].

Sulfur metabolism

Dissimilatory sulfate reduction/sulfur oxidation. The *aprAB* genes for adenosine phosphosulfate reductase and the *dsr* genes for dissimilatory sulfite reductase were found in four and nine MAGs, respectively. All these genes may also be involved in sulfur oxidation.

SOX system. Twelve MAGs from both TAG and Rainbow were observed to be capable of thiosulfate oxidation (using *sox* genes). The capacity for alternative thiosulfate oxidation (*tsdA* genes) was found in *Sulfurocaceae* MAGs from both sites. Thiosulfate polysulfide reductase (reverse), possibly converting thiosulfate into sulfite and hydrogen sulfide was identified in ten MAGs. We also identified the *sor* genes encoding sulfite dehydrogenase in ten MAGs and *soe* genes encoding sulfite:quinone oxidoreductase, both involved in sulfur oxidation, in eight MAGs. The capacity for sulfide oxidation using *sqr* genes was found in 25 MAGs. The capacity for potentially encoding DMSO reductase (using *dmsABC* genes), was identified in nine MAGs, although no phylogenetic analyses were performed.

Iron metabolism

Both *Zetaproteobacteria* MAGs showed the potential to oxidize Fe (II) using *Cyc2* genes [2,3].

150

151 **Hydrogen metabolism**

152 Hydrogen-metabolizing organisms use [NiFe]-hydrogenase to catalyze hydrogen
153 oxidation. The *hoxHFUY* genes for NifFe NAD-reducing hydrogenase that couples
154 reduction of NAD⁺ to the oxidation of hydrogen were identified in six MAGs from both
155 sites and a total of 18 MAGs encoded the *hyaABC* genes for NiFe hydrogenase Hyd-1
156 (Hydrogenase 1).

157

158 **Nitrogen metabolism**

159 **Denitrification.** The potential for dissimilatory nitrate reduction using *nar* or *nap* genes
160 was identified 15 MAGs from both sites. In addition, the capacity for the denitrification
161 step of nitrite reduction requiring *nir* genes was observed in 14 MAGs, capacity for nitric
162 oxide reduction requiring *nor* genes in 19 MAGs and capacity for nitrous oxide reduction
163 requiring *nos* genes in 18 MAGs.

164

165 **DNRA.** Genes *nir* or *nrf* for dissimilatory nitrate reduction to ammonium (DNRA) were
166 retrieved in six MAGs.

167

168 **Nitrification.** The key enzyme for nitrite oxidation, nitrite oxidoreductase (*nxr*), was
169 identified in six MAGs. In addition, nitrogen fixation genes (*nif* genes) were retrieved in
170 the two *Desulfocapsaceae* MAGs only, in agreement with the recent study of Jiang et al.
171 (2020).

172

Respiration and oxygen sensing

Genes for cytochrome c oxidase (*cox*), ubiquinol cytochrome c reductase (*pet;fbc*), cytochrome c oxidase cbb3 type (*cco*) and cytochrome bd complex (*cyd*), which show different affinities for oxygen, were found in 18, 12, 29, and 22 MAGs from both sites, respectively.

Symbiont-host colonization

The potential capacity for biofilm exopolysaccharide PGA synthesis using *pgaABCD*, encoding an adhesin implicated in biofilm maintenance, was identified in one *Desulfocapsaceae* MAG only and the potential for sulfolipid biosynthesis was identified in one MAG from the *Trueperaceae* family. In addition, we observed an enrichment of genes coding for type I secretion systems within 11 MAGs from the *Rhodobacteraceae*, *Flavobacteriaceae*, and *Marinicellaceae* families. Type II encoding pili-related proteins were also observed in 18 MAGs. Overall, type I and type II mediate the secretion of a large variety of protein substrates (unfolded and folded) often associated with virulence [5,6], which could be used by symbionts to colonize hosts. Virulence factors such as hemolysins and chitinases exported by the type II secretion system have also been shown to be critical for initial establishment of symbiont *Aeromonas veronii* in the leech gut [7] or in the active invasion of the rice fungus *Rhizopus microsporus* by its endosymbionts *Burkholderia rhizoxinica* [8]. A type III secretion system was possibly identified in one *Mariprofundaceae* MAG, yet not all subunit genes were retrieved. These secretion systems are usually found in pathogenic bacteria, promoting the transfer of bacterial effector proteins to eukaryotic cells and promoting bacterial invasion and colonization

[5,6]. Type IV secretion systems possibly involved in the dissemination of mobile genetic elements in addition to effector molecules were solely found in *Rhodobacteraceae* (4 MAGs) and *Flavobacteriaceae* (1 MAG) at both TAG and Rainbow. Type VI secretion systems were observed in the same *Rhodobacteraceae* MAGs as type IV. These secretion systems are reported to transfer toxic effector proteins into eukaryotic and prokaryotic target cells, to have an important role in pathogenesis and to defend bacteria against competing organisms [1]. The type IV secretion system also encodes pili that enable directional crawling (twitching motility), biofilm formation, and adhesion at the initial stages of colonization [9]. Twin arginine targeting (Tat) systems, another type of bacterial secretion system, documented to transport fully folded protein substrates and to be vital to bacteria [10], were identified in 36 MAGs.

Metal transporters and detoxification

Cobalt transporter genes *cbi* and *cor* were identified in five and four MAGs, respectively, belonging to *Desulfocapsaceae* and *Sulfurimonadaceae*. Copper transporters and ferrous iron transporters were found in 15 and 27 MAGs, respectively. Strikingly, a total of 41 MAGs showed the potential for dissimilatory arsenic reduction, with eight of them displaying most of the *ars* operon *arsRABC* (more than 75% pathway completeness). This latter allows the reduction (*arsC*), export (*arsB*, helped by *arsA*) and regulation of arsenic level (*arsR*). A potential role of detoxification through polymer chelation has also been reported [11].

Vitamin biosynthesis and transporters

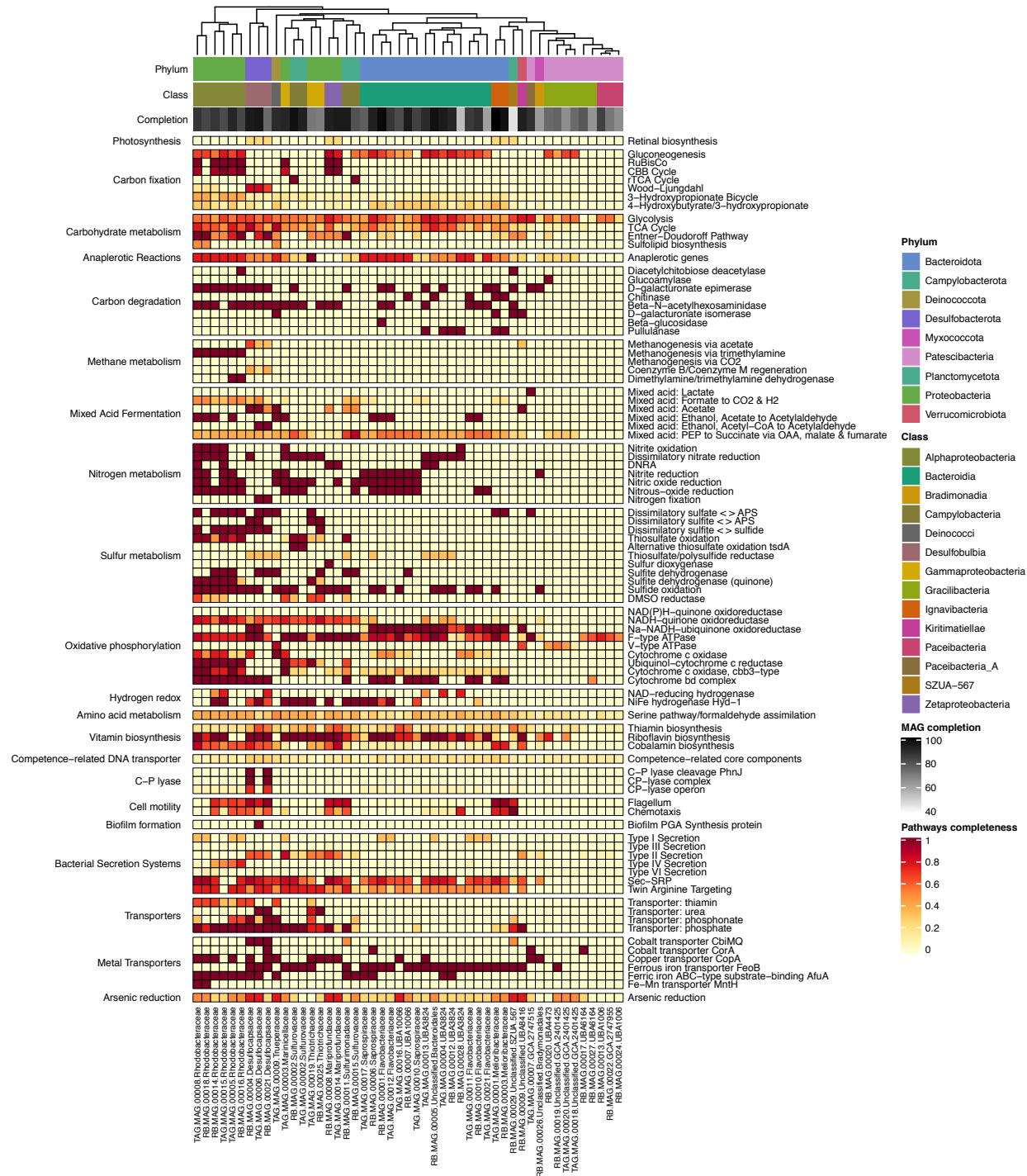
A total of 48 MAGs showed the capacity to encode some of the genes required for thiamin biosynthesis and 37 could synthesize riboflavin and cobalamin, suggesting the importance of vitamin synthesis for the epibiont community. In addition, the presence of thiamin transporters (*tbpA*, *thiP*) in nine MAGs confirmed the capacity to export newly synthesized vitamins. The presence of genes coding for phosphate transporters (*pstABCS*) in 24 MAGs across more than six families suggests the capacity and importance of phosphate uptake in the *Rimicaris* holobiont.

Flagellum and chemotaxis

Flagellum biosynthesis genes (represented by a subset of the *Fli*, *Flh* and *Flg* genes) were identified in 13 MAGs belonging to *Rhodobacteraceae*, *Desulfocapsaceae*, *Melioribacteraceae*, *Sulfurimonadaceae* and *Mariprofundaceae*. The capacity for chemotaxis was identified in 19 MAGs belonging to the same families.

1. Yang Y, Sun J, Sun Y, Kwan YH, Wong WC, Zhang Y, et al. Genomic, transcriptomic, and proteomic insights into the symbiosis of deep-sea tubeworm holobionts. *ISME J.* 2020;14:135–50.
2. Barco RA, Emerson D, Sylvan JB, Orcutt BN, Jacobson Meyers ME, Ramírez GA, et al. New insight into microbial iron oxidation as revealed by the proteomic profile of an obligate iron-oxidizing chemolithoautotroph. *Appl Environ Microbiol.* 2015;81:5927–37.
3. Mori JF, Scott JJ, Hager KW, Moyer CL, Küsel K, Emerson D. Physiological and ecological implications of an iron- or hydrogen-oxidizing member of the Zetaproteobacteria, *Ghiorsea bivora*, gen. nov., sp. *ISME J.* 2017;11:2624–36.
4. Jiang L, Liu X, Dong C, Huang Z, Cambon-Bonavita M-A, Alain K, et al. Candidatus *Desulfobulbus rimicarensis*, an uncultivated deltaproteobacterial epibionts from the deep-sea hydrothermal vent shrimp *Rimicaris exoculata*. *Appl Environ Microbiol.* 2020;86(8):e02549-19.

5. Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, et al. Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nat Rev Microbiol* 2015;13:343–59.
6. Li Y, Liles MR, Halanych KM. Endosymbiont genomes yield clues of tubeworm success. *ISME J.* 2018;12:2785–95.
7. Maltz M, Graf J. The type II secretion system is essential for erythrocyte lysis and gut colonization by the leech digestive tract symbiont *aeromonas veronii*. *Appl Environ Microbiol.* 2011;77:597–603.
8. Moebius N, Üzümlü Z, Dijksterhuis J, Lackner G, Hertweck C. Active invasion of bacteria into living fungal cells. *Elife.* 2014;3:e03007.
9. Grohmann E, Christie PJ, Waksman G, Backert S. Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol Microbiol.* 2018;107(4):455-471.
10. Pickering BS, Oresnik IJ. The twin arginine transport system appears to be essential for viability in *Sinorhizobium meliloti*. *J Bacteriol.* 2010;192:5173–80.
11. Gupta P, Diwan B. Bacterial Exopolysaccharide mediated heavy metal removal: A Review on biosynthesis, mechanism and remediation strategies. *Biotechnol Reports;* 2017;13:58–71.



Additional File 14. KEGG Decoder heat map representing metabolic pathway

completeness of the MAGs based on the presence or absence of genes as determined by

KEGG Decoder. The dendrogram at the top represents the similarity between the MAGs

272 based on their metabolic pathways, using Euclidean distance and complete linkage
273 clustering. Taxonomic affiliations at the class and phylum levels are indicated and
274 represented by the different colors. NAD-reducing hydrogenase: *hox*HFUY. NiFe
275 hydrogenase Hyd-1: *hya*ABC. (PDF 31kb).

276

277

278

279

280

281

282

283

284

285