Supplementary Material for:

Seasonal patterns in Arctic prasinophytes and inferred ecology of *Bathycoccus* unveiled in an Arctic winter metagenome

Nathalie Joli [1], Adam Monier[2], Ramiro Logares [3], Connie Lovejoy [1]

[1] Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Québec Océan, Université Laval, Québec, QC, Canada and Takuvik Joint International Laboratory, Centre National de la Recherche Scientifique (France, CNRS UMI 3376);

[2] Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK.;

[3] Department of Marine Biology and Oceanography, Institut de Ciències del Mar (CSIC), Barcelona, Spain.

<u>Table S1:</u> Sample characteristics and environmental parameters. Station name (STN); latitude (Lat); longitude (Long); sample water mass origin (WM), Polar Mixed Layer (PML), halocline (HC); sample collection depth (Z) in meters (m); percent (%) of regional ice (Ice) cover; day length (DL) in hours (h) when the sun was above the horizon; temperature (T); salinity (S); chlorophyll *a* (Chl) from calibrated relative fluorescence units; nitrate + nitrite (N); and phosphate (P); data not available (NA). November and December samples were collected in 2007 and January to July samples in 2008.

Date	**STN**	**Lat** °N	**Long** °E	**WM**	**Z** m	**Ice** %	**DL** h	**T** ◦C	**S**	**Chl** mg m^{-3}	**N** μM	**P** μM.
19 Nov	405	70.62	236.99	PML	10	95	1	-1.7	30.1	0.08	0.7	0.7
				HC	80			-1.4	32.3	0.05	5.4	1.2
07 Dec	5D	71.31	235.21	PML	10	95	0	-1.7	30.6	0.04	1.1	0.8
				HC	50			-0.9	31.4	0.03	5.2	1.1
03 Jan	14D	71.23	235.55	PML	10	95	0	-1.7	31.1	0.02	1.4	0.8
				HC	75			-1.6	32.0	0.02	5.0	0.9
18 Feb	22D	71.31	235.50	PML	12	96	8	-1.7	31.7	0.02	4.1	0.9
				HC	50			-1.7	32.0	0.02	7.8	1.2
10 Mar	29D	71.04	236.09	PML	12	97	11	-1.7	31.8	0.05	4.3	0.7
				HC	65			-1.5	32.5	0.01	11.0	0.9
17 Mar	D29	70.91	236.52	PML	10	97	12	-1.7	31.9	0.07	5.1	1.5
				HC	50			-1.6	32.2	0.00	10.6	2.0
09 Apr	D36	71.31	235.43	PML	10	90	16	-1.7	31.9	0.60	3.2	1.4
				HC	50			-1.7	32.0	0.36	5.0	1.5
30 Apr	D43	70.77	236.10	PML	11	80	19	-1.7	32.0	0.76	3.1	1.4
				HC	55			-1.7	32.0	0.78	11.1	2.1
19 May	405b	70.66	237.12	PML	12	30	24	-1.4	31.9	10.62	1.0	1.2
				HC	30			-1.3	32.3	3.12	4.4	1.6
14 Jun	FB01	69.99	234.45	PML	0	0	24	-1.3	31.8	0.58	1.5	N/A
				HC	22			-1.4	32.1	0.49	2.5	N/A
24 Jun	F7b	69.82	236.35	PML	5	0	24	2.1	31.3	0.70	1.3	N/A
				HC	33			-1.1	32.1	0.25	2.3	N/A
21 Jul	405-10a	70.70	237.07	PML	4	0	24	8.4	30.5	0.29	0.8	0.7
				HC	42			-1	32.2	4.64	14.6	1.5
Metagenome												
13 Dec	7D	71.62	233.93	PML	10	95	0	-1.6	30.56	NA	NA	NA

Table S2: Proportion of each size fraction that was pooled for amplicon sequencing. November and December samples collected in 2007 and January to July samples collected in 2008.

	19 Nov	07 Dec	03 Jan	18 Feb	10 Mar	17 Mar	09 Apr	30 Apr	19 May	14 Jun	24 Jun	21 Jul
Small	41%	41%	50%	44%	50%	72%	62%	15%	10%	51%	65%	85%
Large	59%	59%	50%	56%	50%	28%	38%	85%	90%	49%	35%	15%

Table S3: Arctic metagenome coverage by contigs and reads of reference genomes: *Bathycoccus* Ban7 (RCC1105, isolated from Mediterranean, Genbank FO082278), *Micromonas* RCC299 (isolated from Equatorial Pacific, GCF_000090985.2), *Micromonas* CCMP1545 (isolated from English Channel, GCA_000151265.1). Also shown is the Arctic metagenome coverage by contigs of transcriptomes available from the MMETSP (see Keeling et al. 2014): *Bathycoccus* RCC716 (isolated from Indian Ocean; CAM_SMPL_003310) and *Micromonas* CCMP2099 (isolated from Arctic Ocean; CAM_SMPL_003148 and CAM_SMPL_002821). Note that genome size of *Micromonas* RCC299 is 20Mb and *Bathycoccus* Ban7 is 15Mb.

	Target	**% Coverage by contigs**	**% Coverage by reads**
Bathycoccus Ban7	Genome	88.8	95.4
Bathycoccus RCC716	Transcriptome	53.0	-
Micromonas RCC299	Genome	21.5	21.6
Micromonas CCMP1545	Genome	2.5	-
Micromonas CCMP2099	Transcriptome	3.8	-

Table S4: Original source and sizes in megabases (Mb) of the *Bathycoccus* reference Ban7 genome and metagenome assembled genomes (MAGs) referred to in the main text.

	Isolated	**Genome or MAG size**	**Accession number**
B. prasinos Ban7	Mediterranean	15 Mb	FO082258, FO082259, FO082278-FO082260
B. prasinos Pacific 1	Chilean coast, Pacific	9 Mb	CAFX01000001- CAFX01015049
B. prasinos Pacific 2	Chilean coast, Pacific	13 Mb	CAFY01000001-CAFY01022174
B. prasinos Atlantic	Tropical Atlantic	7.1 Mb	AFUW01000001:AFUW01000185
*B. prasinos*Arctic	Amundsen Gulf, Arctic	12.5 Mb	SAMN05514161[a]

[a] NCBI Biosample number which contains all the metagenomic contigs. A note saying « Bathycoccus-like » had been added to each contig taxonomically binned as Bathycoccus using Metawatt.

Table S5A: Relative abundance of the total chlorophytes reads as a percentage of microbial eukaryote reads (excluding Fungi and Metazoa) in the separate samples. The proportions of chlorophyte reads among Eukaryote detected over the study are given in Table S5B. Sample water mass origin (WM), Polar Mixed Layer (PML), halocline (HC) as defined in the main text.

Template	WM	11 Nov	07 Dec	03 Jan	18 Feb	10 Mar	17 Mar	9 Apr	30 Apr	19 May	14 Jun	24 Jun	21 Jul
DNA	PML	4.50%	1.60%	0.50%	1.00%	0.70%	0.80%	4.80%	6.70%	2.60%	16.90%	13.20%	2.70%
	HC	1.20%	0.70%	0.30%	0.30%	0.20%	0.50%	3.60%	5.10%	3.50%	6.20%	2.80%	4.10%
RNA	PML	3.40%	1.50%	0.80%	2.20%	5.20%	7.20%	8.50%	8.50%	1.30%	41.90%	14.00%	1.70%
	HC	2.20%	1.30%	2.20%	2.00%	1.00%	0.80%	11.10%	13.40%	0.30%	7.40%	0.40%	3.90%

<u>Table S5B</u>: Proportion of individual chlorophytes taxa (all prasinophytes) out of total eukaryotic reads from DNA or RNA templates. Water mass (WM), Polar Mixed Layer (PML), halocline (HC).

	Template	**WM**	**19 Nov**	**07 Dec**	**03 Jan**	**18 Feb**	**10 Mar**	**17 Mar**	**09 Apr**	**30 Apr**	**19 May**	**14 Jun**	**24 Jun**	**21 Jul**
Micromonas CCMP2099	**DNA**	**PML**	0.30%	0.20%	0.00%	0.20%	0.20%	0.20%	2.60%	3.50%	0.10%	14.10%	12.80%	0.70%
		HC	0.10%	0.10%	0.00%	0.10%	0.10%	0.40%	2.40%	3.70%	0.80%	5.30%	2.50%	0.80%
	RNA	**PML**	1.20%	0.30%	0.40%	0.80%	2.60%	3.50%	6.00%	3.80%	0.30%	39.40%	13.70%	0.30%
		HC	0.40%	0.60%	1.60%	0.90%	0.70%	0.60%	9.80%	11.10%	0.30%	6.50%	0.30%	1.60%
Bathycoccus Ban7	**DNA**	**PML**	3.20%	0.90%	0.30%	0.20%	0.30%	0.30%	0.50%	0.40%	0.00%	0.10%	0.10%	1.10%
		HC	0.90%	0.30%	0.10%	0.00%	0.00%	0.00%	0.80%	0.50%	0.20%	0.20%	0.10%	0.00%
	RNA	**PML**	1.00%	0.50%	0.20%	0.20%	0.60%	0.40%	0.10%	0.10%	0.00%	0.10%	0.00%	0.20%
		HC	0.70%	0.10%	0.10%	0.40%	0.10%	0.10%	0.20%	0.40%	0.00%	0.10%	0.00%	0.00%
Mantoniella	**DNA**	**PML**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	0.20%	0.00%	0.00%	0.00%	0.00%
		HC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.20%	0.20%	0.00%	0.00%	0.00%	0.00%
	RNA	**PML**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		HC	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%	0.10%	0.10%	0.00%	0.00%	0.00%	0.00%
Micromonas CCMP1195	**DNA**	**PML**	0.60%	0.20%	0.10%	0.10%	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.10%
		HC	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	RNA	**PML**	0.70%	0.50%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.00%	0.00%	0.00%	0.10%
		HC	0.50%	0.00%	0.10%	0.00%	0.00%	0.00%	0.50%	0.20%	0.00%	0.00%	0.00%	0.00%
Pterosperma	**DNA**	**PML**	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.40%
		HC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.20%
	RNA	**PML**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
		HC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.20%
Pyramimonas	**DNA**	**PML**	0.20%	0.30%	0.10%	0.30%	0.30%	0.20%	1.40%	2.50%	2.40%	2.50%	0.10%	0.20%
		HC	0.10%	0.20%	0.20%	0.20%	0.00%	0.10%	0.20%	0.70%	2.40%	0.70%	0.20%	3.00%
	RNA	**PML**	0.40%	0.10%	0.10%	0.90%	1.50%	3.20%	2.30%	4.50%	0.90%	2.40%	0.20%	0.40%
		HC	0.30%	0.40%	0.20%	0.50%	0.10%	0.00%	0.50%	1.50%	0.00%	0.60%	0.10%	2.00%

Table S6: Assembly statistics of the Arctic metagenome from metaQUAST (http://bioinf.spbau.ru/quast)

# contigs (>= 0 bp)	1687512
# contigs (>= 1000 bp)	130257
Total length (>= 0 bp)	881683496
Total length (>= 1000 bp)	333707177
Largest contig	200608
Total length	881683496
GC (%)	42.81
N50	615
N75	315

Table S7: Meiosis related genes targeted by BLASTn against the *Bathycoccus*-like contigs from the metagenomic data. References genes based on Derelle et al. 2006 and Worden et al. 2009. Function are genes or gene families; Query column lists Genbank gi numbers. Meiotic recombination protein (DMC1), DNA mismatch repair and recombination (MLH1), DNA repair/recombination protein (Mre11), DNA repair/recombination protein (RAD), DNA recombinase (RECA), DNA topoisomerase VI, subunit A (SPO11), chloroplast envelope protein translocase family (TOC) and DNA topoisomerase VI (TOP).

Function	**Query**	**Bathycoccus-like contigs**	**% Identity**	**Evalue**	**Coverage**
DMC1	gi|255086168|	contig_41_26911450_length_1761_multi_4_in_0_out_0	69.7%	1.02e-50	42.13%
DMC1	gi|303284846|	contig_55_6029503_length_522_multi_4_in_0_out_0	73.4%	1.47e-23	17.11%
DMC1	gi|303284846|	contig_41_26911450_length_1761_multi_4_in_0_out_0	72.3%	1.79e-60	39.7%
MLH1	gi|303283871|	contig_57_2795483_length_2883_multi_4_in_0_out_0	67.4%	5.73e-26	20.21%
Mre11	gi|255083255|	contig_69_494867_length_2019_multi_4_in_0_out_0	66.1%	2.94e-43	27.11%
MRE11	gi|303288633|	contig_69_494867_length_2019_multi_4_in_0_out_0	65.9%	2.21e-38	27.33%
MRE11	gi|693499007|	contig_69_494867_length_2019_multi_4_in_0_out_0	68.9%	1.01e-35	19.97%
RAD50	gi|255071308|	contig_51_8008664_length_2279_multi_4_in_0_out_0	75.4%	2.61e-90	12.36%
RAD50	gi|255071308|	contig_41_2430054_length_3324_multi_4_in_0_out_0	69.5%	7.99e-59	14.83%
RAD50	gi|693501461|	contig_41_2430054_length_3324_multi_4_in_0_out_0	73.6%	1.21e-94	13.8%
RAD50	gi|693501461|	contig_51_8008664_length_2279_multi_4_in_0_out_0	69.5%	7.14e-66	13.95%
RAD51	gi|255075466|	contig_41_26911450_length_1761_multi_4_in_0_out_0	67.2%	1.45e-10	19.71%
RAD51	gi|255075466|	contig_47_6166815_length_1105_multi_4_in_0_out_0	68.2%	8.58e-39	41.93%
RAD51	gi|255081794|	contig_41_26911450_length_1761_multi_4_in_0_out_0	67.2%	1.75e-10	16.43%
RAD51	gi|255081794|	contig_47_6166815_length_1105_multi_4_in_0_out_0	68.2%	1.03e-38	34.94%
RAD51	gi|303279525|	contig_41_26911450_length_1761_multi_4_in_0_out_0	66.1%	2.83e-09	14.62%
RAD51	gi|303279525|	contig_47_6166815_length_1105_multi_4_in_0_out_0	71.4%	7.10e-55	26.07%
RAD51A	gi|303279167|	contig_41_26911450_length_1761_multi_4_in_0_out_0	66.1%	2.47e-09	16.69%
RAD51A	gi|303279167|	contig_47_6166815_length_1105_multi_4_in_0_out_0	71.4%	6.20e-55	29.76%
RAD51C	gi|303272032|	contig_49_9069687_length_1612_multi_4_in_0_out_0	64.5%	5.58e-16	38.03%
RAD51C	gi|303273073|	contig_41_2430054_length_3324_multi_4_in_0_out_0	72.8%	1.06e-83	11.13%
RAD51D	gi|255079113|	contig_47_7431291_length_1342_multi_4_in_0_out_0	68.6%	2.22e-07	17.97%
RAD54	gi|255082184|	contig_55_5843141_length_2538_multi_5_in_0_out_0	73.5%	3.46e-56	10.91%
RAD54	gi|255082184|	contig_37_4621688_length_1168_multi_4_in_0_out_0	65.2%	5.87e-28	16.47%
RAD54	gi|303279421|	contig_37_4621688_length_1168_multi_4_in_0_out_0	68.1%	9.97e-57	19.85%
RAD54	gi|303279421|	contig_55_5843141_length_2538_multi_5_in_0_out_0	68.5%	1.48e-73	23.17%
RECA	gi|255080473|	contig_29_66625928_length_996_multi_4_in_0_out_0	72.7%	2.68e-130	46.26%
RECA	gi|303271542|	contig_29_66625928_length_996_multi_4_in_0_out_0	75.0%	9.12e-173	66.59%
SPO11	gi|693499744|	contig_73_1282211_length_1398_multi_4_in_0_out_0	76.3%	0	73.05%
TOC	gi|303289255|	contig_59_5833480_length_3632_multi_4_in_0_out_0	70.7%	9.09e-51	13.06%
TOC	gi|303289255|	contig_71_1892218_length_2815_multi_5_in_0_out_0	63.3%	1.11e-11	17.59%
TOC	gi|693498025|	contig_59_4385194_length_1832_multi_4_in_0_out_0	72.4%	2.65e-82	18.83%
TOP6A	gi|255070110|	contig_73_1282211_length_1398_multi_4_in_0_out_0	71.0%	1.45e-112	71.68%
TOP6B	gi|255073850|	contig_33_18016919_length_864_multi_4_in_0_out_0	68.7%	1.51e-64	31.34%
TOP6B	gi|303276025|	contig_33_18016919_length_864_multi_4_in_0_out_0	71.1%	2.73e-93	29.25%

Table S8: List of *Bathycoccus*-specified KEGG pathway variant genes (Gene affected). The type of variant refers to specific variation that would have a high effect on the final putative protein. The search was carried out was against *Bathycoccus prasinos* pathways implemented in KEGG.

Bathycoccus-specific KEGG pathway	Gene affected	Type of variant
bpg01110 Biosynthesis of secondary metabolites	K00611 OTC; ornithine carbamoyltransferase [EC:2.1.3.3]	frameshift
	K00640 cysE; serine O-acetyltransferase [EC:2.3.1.30]	start loss, stop gain
	K01899 LSC1; succinyl-CoA synthetase alpha subunit [EC:6.2.1.4 6.2.1.5]	stop gain
	K01918 panC; pantoate--beta-alanine ligase [EC:6.3.2.1]	stop gain
	K11778 DHDDS; ditrans,polycis-polyprenyl diphosphate synthase [EC:2.5.1.87]	start loss, stop gain
	K13789 GGPS; geranylgeranyl diphosphate synthase, type II [EC:2.5.1.1 2.5.1.10 2.5.1.29]	stop gain
	K15634 gpmB; probable phosphoglycerate mutase [EC:5.4.2.12]	stop gain
bpg00480 Glutathione metabolism	K00799 GST; glutathione S-transferase [EC:2.5.1.18]	stop gain
bpg03010 Ribosome	K02875 RP-L14e; large subunit ribosomal protein L14e	stop gain
	K02925 RP-L3e; large subunit ribosomal protein L3e	frameshift
	K02927 RP-L40e; large subunit ribosomal protein L40e	start loss, stop gain
	K02929 RP-L44e; large subunit ribosomal protein L44e	stop gain
	K02964 RP-S18e; small subunit ribosomal protein S18e	stop gain
	K02997 RP-S9e; small subunit ribosomal protein S9e	stop gain
bpg01200 Carbon metabolism	K00640 cysE; serine O-acetyltransferase [EC:2.3.1.30]	start loss, stop gain
	K01899 LSC1; succinyl-CoA synthetase alpha subunit [EC:6.2.1.4 6.2.1.5]	stop gain
	K15634 gpmB; probable phosphoglycerate mutase [EC:5.4.2.12]	stop gain
bpg01230 Biosynthesis of amino acids	K00611 OTC; ornithine carbamoyltransferase [EC:2.1.3.3]	frameshift
	K00640 cysE; serine O-acetyltransferase [EC:2.3.1.30]	start loss, stop gain
	K15634 gpmB; probable phosphoglycerate mutase [EC:5.4.2.12]	stop gain
bpg04141 Protein processing in endoplasmic reticulum	K10597 UBE4B; ubiquitin conjugation factor E4 B [EC:2.3.2.27]	stop gain
	K13993 HSP20; HSP20 family protein	stop gain
	K14004 SEC13; protein transport protein SEC13	start loss, stop gain
bpg03015 mRNA surveillance pathway	K14404 CPSF4; cleavage and polyadenylation specificity factor subunit 4	frameshift, splice region
	K14408 CSTF3; cleavage stimulation factor subunit 3	splice region, stop loss
	K15542 PFS2; polyadenylation factor subunit 2	stop gain
bpg03018 RNA degradation	K12580 CNOT3; CCR4-NOT transcription complex subunit 3	start loss, stop gain
	K12598 MTR4; ATP-dependent RNA helicase DOB1 [EC:3.6.4.13]	stop gain
	K12623 LSM4; U6 snRNA-associated Sm-like protein LSm4	stop gain
bpg04120 Ubiquitin mediated proteolysis	K10586 BIRC6; baculoviral IAP repeat-containing protein 6 (apollon) [EC:2.3.2.23]	splice region, stop loss
	K10597 UBE4B; ubiquitin conjugation factor E4 B [EC:2.3.2.27]	stop gain
	K10598 PPIL2; peptidyl-prolyl cis-trans isomerase-like 2 [EC:5.2.1.8]	stop gain
bpg03013 RNA transport	K03243 EIF5B; translation initiation factor 5B	stop gain
	K13137 STRAP; serine-threonine kinase receptor-associated protein	stop gain
	K14004 SEC13; protein transport protein SEC13	start loss, stop gain
bpg00190 Oxidative phosphorylation	K02126 ATPeF0A; F-type H+-transporting ATPase subunit a	stop gain
	K02151 ATPeV1F; V-type H+-transporting ATPase subunit F	stop gain
	K02267 COX6B; cytochrome c oxidase subunit 6b	stop gain
bpg03040 Spliceosome	K12623 LSM4; U6 snRNA-associated Sm-like protein LSm4	stop gain
	K12822 RBM25; RNA-binding protein 25	frameshift
	K12850 PRPF38B; pre-mRNA-splicing factor 38B	stop gain
bpg00564 Glycerophospholipid metabolism	K14676 NTE; lysophospholipid hydrolase [EC:3.1.1.5]	stop gain
bpg03030 DNA replication	K01972 E6.5.1.2; DNA ligase (NAD+) [EC:6.5.1.2]	stop gain
	K02212 MCM4; DNA replication licensing factor MCM4 [EC:3.6.4.12]	stop gain
bpg03420 Nucleotide excision repair	K01972 E6.5.1.2; DNA ligase (NAD+) [EC:6.5.1.2]	stop gain
	K10842 MNAT1; CDK-activating kinase assembly factor MAT1	stop gain
bpg00900 Terpenoid	K11778 DHDDS; ditrans,polycis-polyprenyl diphosphate synthase [EC:2.5.1.87]	start loss, stop gain
	K13789 GGPS; geranylgeranyl diphosphate synthase, type II [EC:2.5.1.1 2.5.1.10 2.5.1.29]	stop gain

backbone biosynthesis		
bpg00270 Cysteine and methionine metabolism	K00640 cysE; serine O-acetyltransferase [EC:2.3.1.30]	start loss, stop gain
bpg04144 Endocytosis	K07937 ARF1; ADP-ribosylation factor 1	stop gain
	K11824 AP2A; AP-2 complex subunit alpha	stop gain
bpg00920 Sulfur metabolism	K00640 cysE; serine O-acetyltransferase [EC:2.3.1.30]	start loss, stop gain
bpg03440 Homologous recombination	K08991 MUS81; crossover junction endonuclease MUS81 [EC:3.1.22.-]	stop gain
	K10875 RAD54L; DNA repair and recombination protein RAD54 and RAD54-like protein [EC:3.6.4.-]	stop gain
bpg00260 Glycine, serine and threonine metabolism	K15634 gpmB; probable phosphoglycerate mutase [EC:5.4.2.12]	stop gain
bpg00010 Glycolysis / Gluconeogenesis	K15634 gpmB; probable phosphoglycerate mutase [EC:5.4.2.12]	stop gain
bpg03008 Ribosome biogenesis in eukaryotes	K14572 MDN1; midasin	stop gain
bpg00590 Arachidonic acid metabolism	K15730 PTGES3; cytosolic prostaglandin-E synthase [EC:5.3.99.3]	stop gain
bpg04145 Phagosome	K02151 ATPeV1F; V-type H+-transporting ATPase subunit F	stop gain
bpg03430 Mismatch repair	K01972 E6.5.1.2; DNA ligase (NAD+) [EC:6.5.1.2]	stop gain
bpg00970 Aminoacyl-tRNA biosynthesis	K03341 SEPSECS; O-phospho-L-seryl-tRNASec:L-selenocysteinyl-tRNA synthase [EC:2.9.1.2]	stop gain
bpg00640 Propanoate metabolism	K01899 LSC1; succinyl-CoA synthetase alpha subunit [EC:6.2.1.4 6.2.1.5]	stop gain
bpg00450 Selenocompound metabolism	K03341 SEPSECS; O-phospho-L-seryl-tRNASec:L-selenocysteinyl-tRNA synthase [EC:2.9.1.2]	stop gain
bpg04140 Regulation of autophagy	K08333 PIK3R4; phosphoinositide-3-kinase, regulatory subunit 4 [EC:2.7.11.1]	stop gain
bpg00410 beta-Alanine metabolism	K01918 panC; pantoate--beta-alanine ligase [EC:6.3.2.1]	stop gain
bpg00770 Pantothenate and CoA biosynthesis	K01918 panC; pantoate--beta-alanine ligase [EC:6.3.2.1]	splice region, stop gain
bpg04070 Phosphatidylinositol signaling system	K13024 PPIP5K; inositol-hexakisphosphate/diphosphoinositol-pentakisphosphate 1-kinase [EC:2.7.4.24]	stop gain
bpg03022 Basal transcription factors	K10842 MNAT1; CDK-activating kinase assembly factor MAT1	stop gain
bpg00195 Photosynthesis	K03541 psbR; photosystem II 10kDa protein	frameshift
bpg00561 Glycerolipid metabolism	K00679 E2.3.1.158; phospholipid:diacylglycerol acyltransferase [EC:2.3.1.158]	start loss, stop gain
bpg03410 Base excision repair	K01972 E6.5.1.2; DNA ligase (NAD+) [EC:6.5.1.2]	stop gain
bpg00220 Arginine biosynthesis	K00611 OTC; ornithine carbamoyltransferase [EC:2.1.3.3]	frameshift
bpg00020 Citrate cycle (TCA cycle)	K01899 LSC1; succinyl-CoA synthetase alpha subunit [EC:6.2.1.4 6.2.1.5]	stop gain