

Supporting Information

Table of Contents

Figure S1 ARB neighbour joining trees of SAR11 16S rRNA gene sequences and internal transcribed spacer (ITS) regions. Separate tree files for this figure are provided:

Supplementary_Figure_S1_ITS_full_MSB.tree; Supplementary_Figure_S1_16S_MSB.tree;

Supplementary_Figure_S1_ITS_concise_MSB.tree

Figure S2 Normalised distribution of blast hits from genome-specific genes within metagenomes from polar, temperate and tropical biomes.

Figure S3 Shifts in sea surface temperature predicted from IPCC climate models.

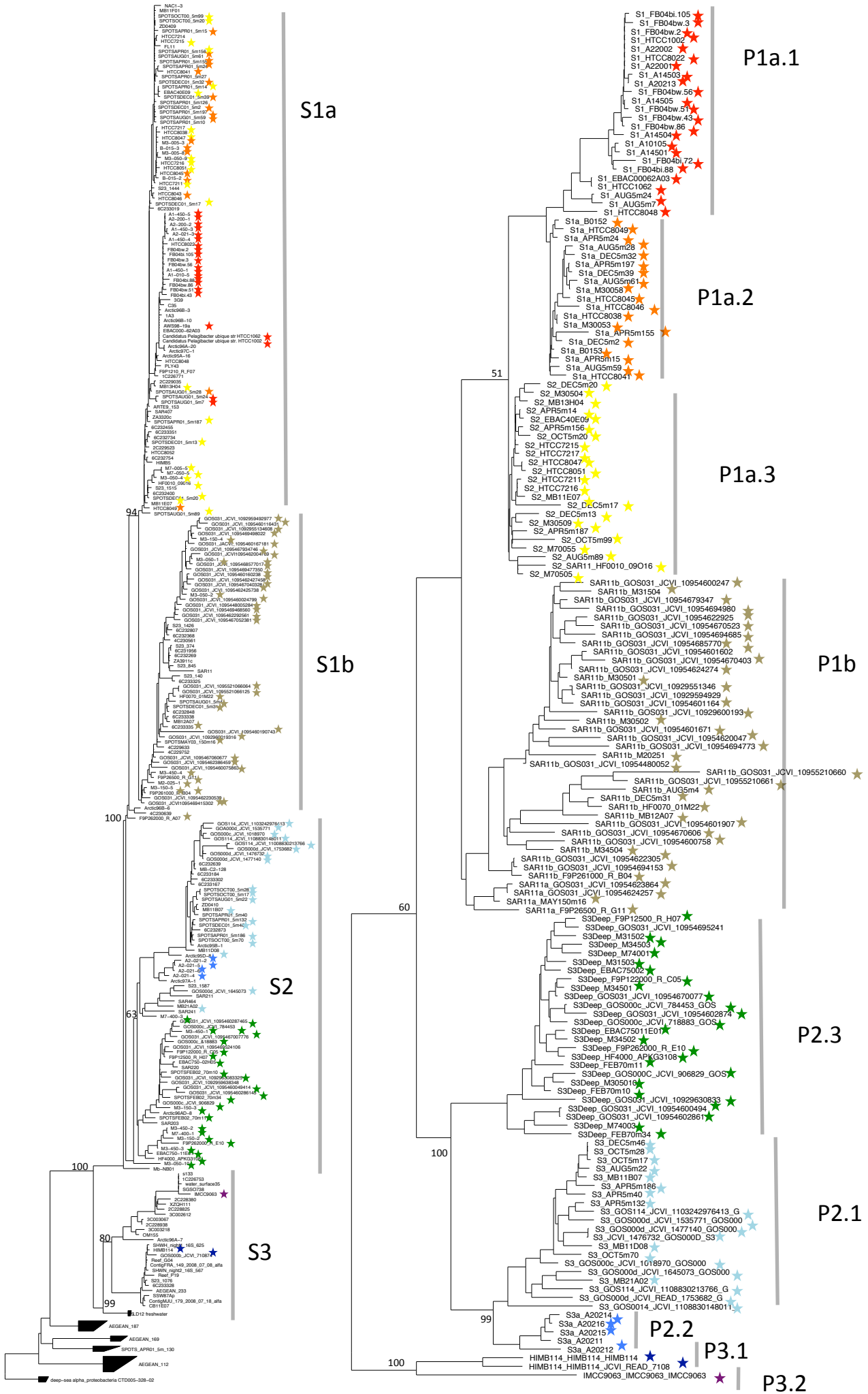
Table S1 Source of sequences making up the SAR11 ITS database.

Table S2 Metagenome sample sets used in this study.

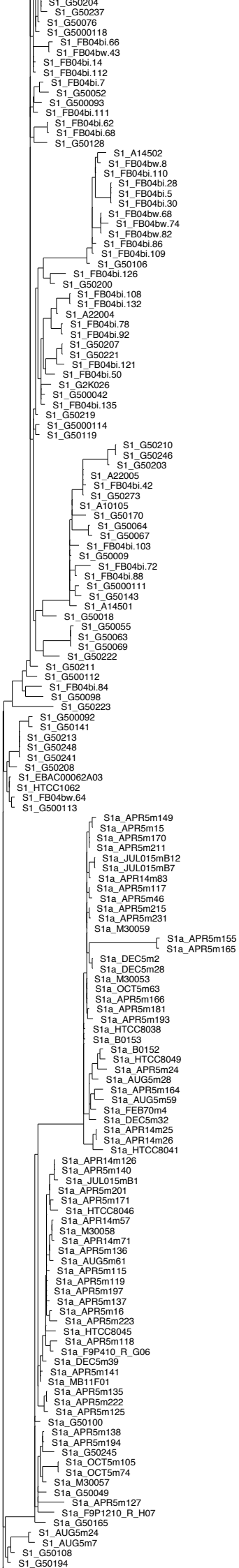
Table S3 Environmental data associated with metagenomes used in this study.

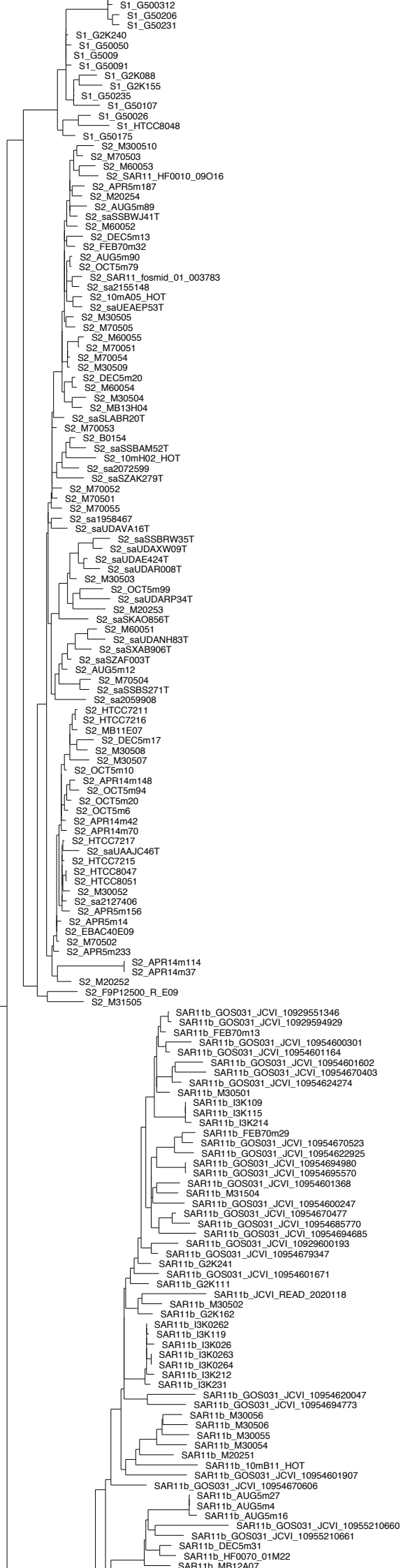
Table S4 Results of DISTLM fitting environmental data to relative abundances of SAR11 phylotypes in 128 surface marine metagenome samples.

Table S5 Specific genes, paralogous genes, and genes with signatures of positive selection in ACE_P1a.1, ACE_P3.2, and IMCC9063 genomes.

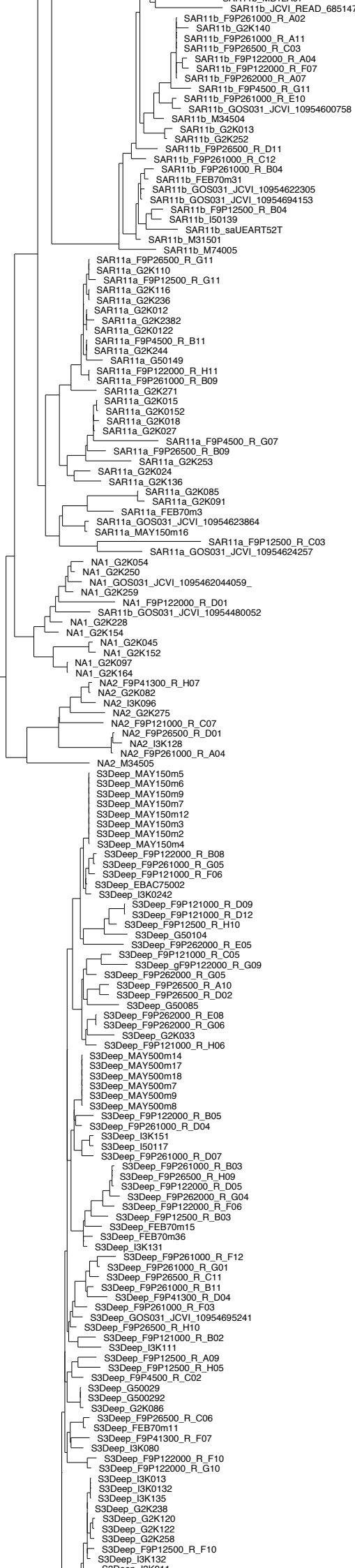


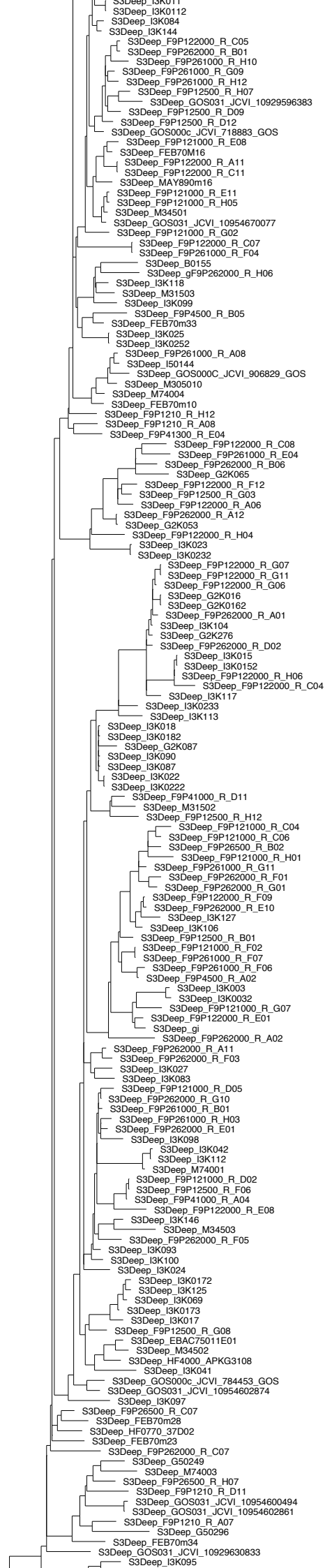
S1_FB04bi.105
S1_FB04bi.131
S1_231_Clade1
S1_FB04bi.101
S1_FB04bi.32
S1_FB04bi.55
S1_FB04bi.90
S1_G50120
S1_A10104
S1_G50133
S1_G50092
S1_A14503
S1_A20213
S1_G50163
S1_A22001
S1_A22002
S1_A22003
S1_G50001
S1_G50110
S1_G50117
S1_G50152
S1_HTCC1002
S1_HTCC8010
S1_HTCC8022
S1_FB04bi.125
S1_FB04bi.34
S1_FB04bi.89
S1_FB04bi.94
S1_FB04bi.91
S1_FB04bw.2
S1_FB04bw.44
S1_FB04bi.35
S1_FB04bw.3
S1_FB04bw.59
S1_FB04bw.41
S1_FB04bi.46
S1_FB04bi.115
S1_FB04bi.96
S1_FB04bi.77
S1_FB04bw.67
S1_G50027
S1_A14505
S1_G50059
S1_FB04bi.102
S1_FB04bi.79
S1_FB04bw.27
S1_FB04bw.51
S1_FB04bi.21
S1_G500182
S1_G5000115
S1_G50093
S1_G50209
S1_G500212
S1_G50101
S1_G5000119
S1_G50031
S1_G500012
S1_G500017
S1_G50081
S1_G50186
S1_FB04bi.80
S1_FB04bi.134
S1_G500213
S1_FB04bi.38
S1_FB04bi.71
S1_G50011
S1_G50096
S1_G500072
S1_G5000110
S1_G50062
S1_FB04bi.67
S1_G5000121
S1_G50014
S1_G500013
S1_G500214
S1_FB04bi.52
S1_FB04bi.60
S1_F9P1210_R_F07
S1_G5000112
S1_FB04bi.127
S1_FB04bi.141
S1_FB04bw.46
S1_G50007
S1_FB04bw.56
S1_FB04bw.79
S1_G50216
S1_A10101
S1_FB04bw.89
S1_G50077
S1_G50247
S1_G500014
S1_G50065
S1_FB04bw.40
S1_A14504
S1_FB04bi.122
S1_FB04bw.47
S1_G5000117
S1_G500016
S1_FB04bi.37
S1_G50083
S1_G50148
S1_APR5m175
S1_FB04bi.74
S1_FB04bi.43
S1_JuI5m5
S1_G2K279
S1_G5000120
S1_G50043
S1_G50142
S1_G50295
S1_FB04bi.100
S1_FB04bi.99
S1_FB04bi.98
S1_FB04bi.18
S1_FB04bi.19
S1_APR14m121
S1_G50056
S1_G50226
S1_G500015
S1_G500073
S1_G500142
S1_G500019
S1_FB04bi.130
S1_G50018
S1_G50079
S1_G50172
S1_A10102
S1_FB04bi.10
S1_G50158
S1_G50121
S1_FB04bi.118
S1_FB04bi.85
S1_FB04bi.40
S1_FB04bi.69
S1_G50036
S1_FB04bi.106
S1_FB04bi.128
S1_G50097
S1_FB04bi.137
S1_FB04bw.86
S1_FB04bi.26
S1_FB04bi.11
S1_FB04bi.54
S1_FB04bw.81
S1_G50070
S1_G50173
S1_A10103
S1_G500074
S1_G50197
S1_G50217
S1_G5000116
S1_G50090





SAR11b_JCVI_READ_685147





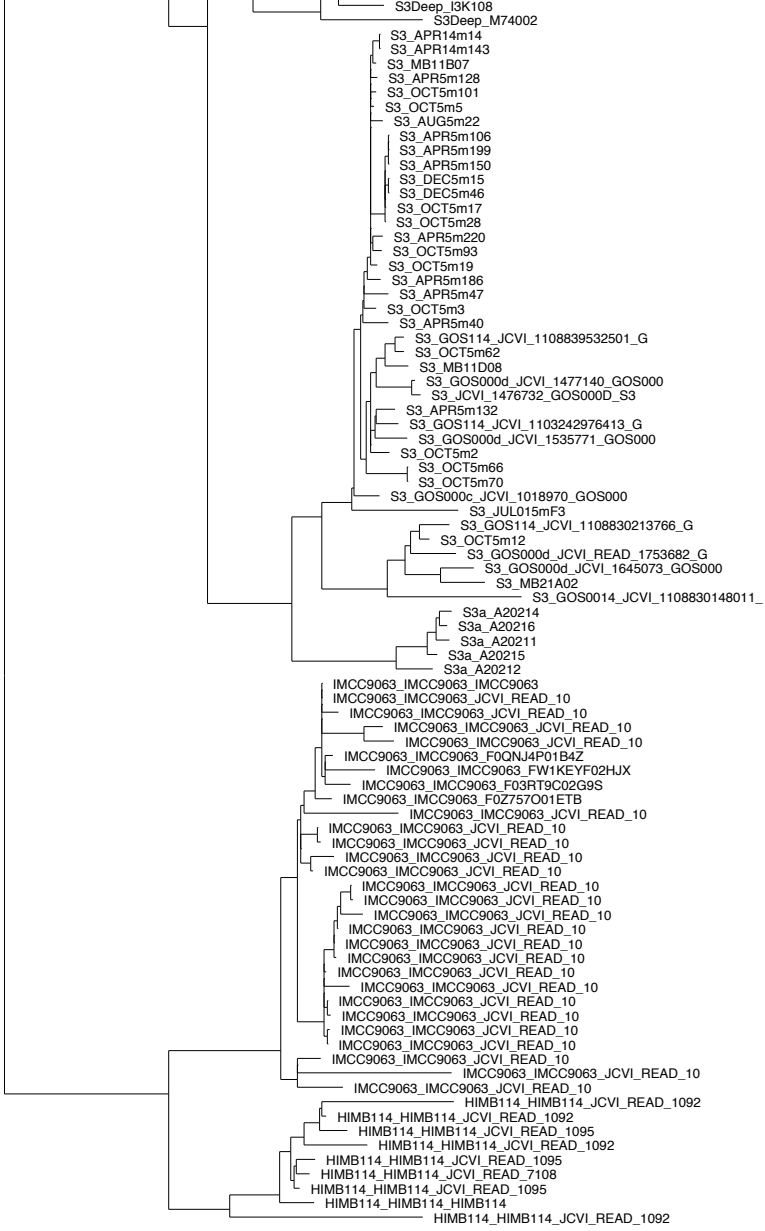


Figure S1 ARB neighbour joining trees of SAR11 16S rRNA gene sequences and internal transcribed spacer (ITS) regions. Top panel: Comparison of ARB neighbour joining trees generated from 16S rRNA gene sequences (left) and ITS regions (right) from clones which contain both regions. Bootstrap values >50% are shown. Bottom panel: ARB neighbour joining tree generated from all 865 ITS sequences used to compile the database.

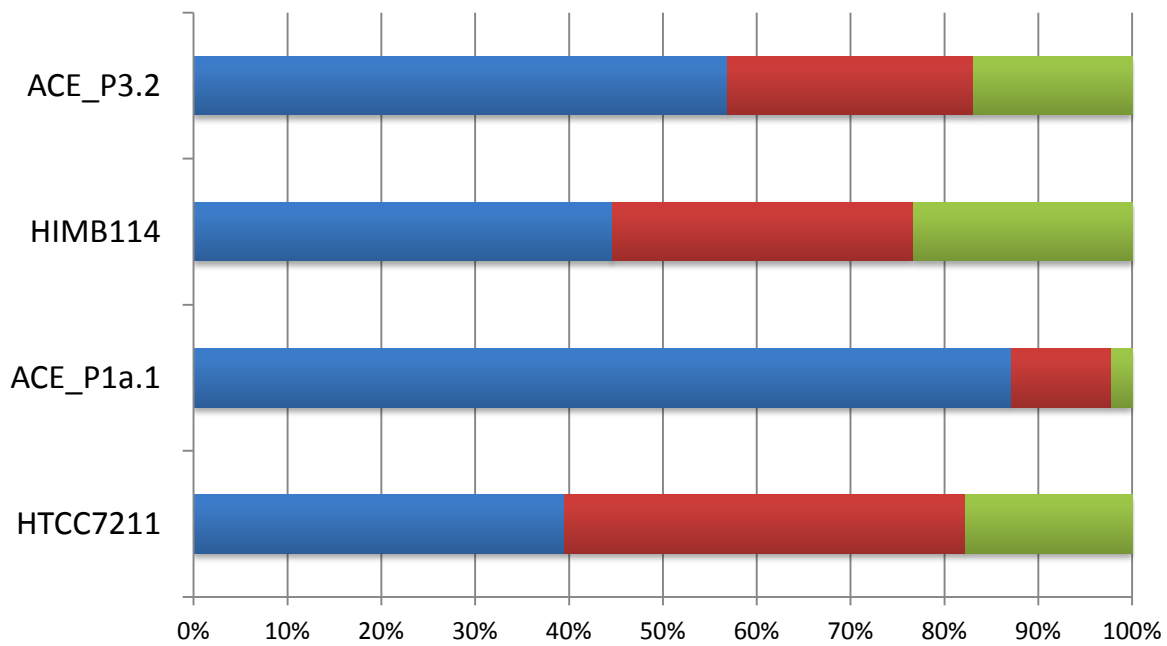


Figure S2 Normalised distribution of blast hits from genome-specific genes within metagenomes from polar, temperate and tropical biomes. The normalized distribution was calculated by dividing the number of genome-specific protein hits to each biome by the total number of hits to all biomes. Polar <10°C (blue), temperate 10-20°C (green), tropical >20°C (red).

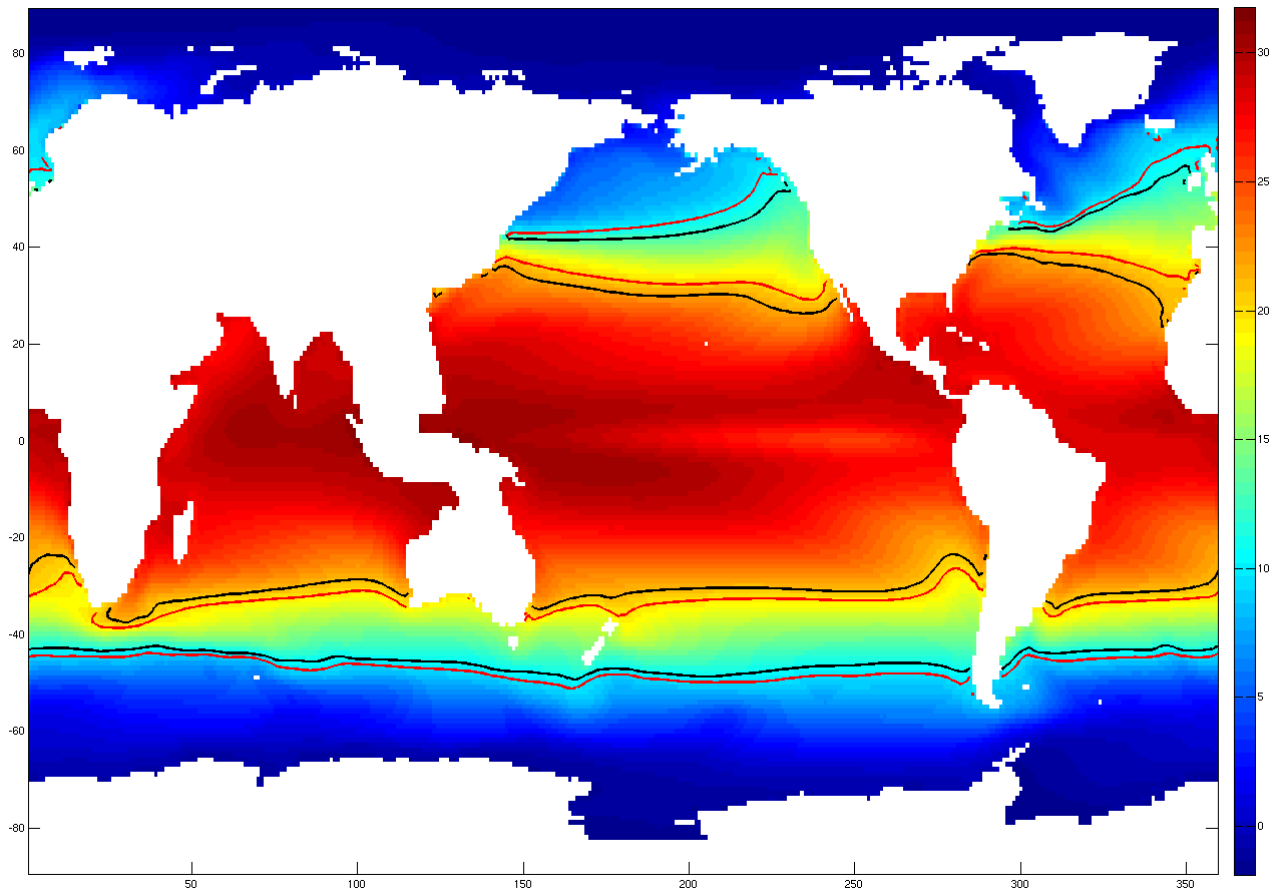


Figure S3 Shifts in sea surface temperature predicted from IPCC climate models. Changes in the location of the 10°C and 20°C sea surface temperature isotherms from 1990 (black) and 2090 (red) calculated as a multi-model mean from 15 IPCC models.

Table S1. Source of sequences making up the SAR11 ITS database.

Source	Location	Depth (m)	Latitude	Longitude	No. of sequences
P1a.1					
ACE lake	Antarctic continent	11.5	68 28.33S	78 11.29E	1
Antarctic seawater	Antarctic near ice shelf between Scotia and Weddell seas	10	60 31.111S	48 18.788W	5
Antarctic seawater	Antarctic near ice shelf between Scotia and Weddell seas	450	60 31.111S	48 18.788W	5
Antarctic seawater	Antarctic ocean off the Antarctic Peninsula	21	63 24.5S	56 40.717W	1
Antarctic seawater	Antarctic ocean off the Antarctic Peninsula	200	63 24.5S	56 40.717W	5
Temperate Coastal seawater	San Pedro Ocean Time Series, California	14	33 33N	118 24W	1
Temperate Coastal seawater	San Pedro Ocean Time Series, California	5	33 33N	118 24W	4
Temperate Coastal seawater	Monterey Bay, California	5			1
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	1
Arctic sea ice	Franklin Bay	0.025	70.0N	126.3W	74
Arctic sea water	Franklin Bay	under ice surface	70.0N	126.3W	23
Arctic sea water	North Atlantic, Greenland Sea	2000	72.9N	4.7E	5
Arctic sea water	North Atlantic, Greenland Sea	50	72.9N	4.7E	120
Temperate Coastal seawater	Oregon				5
P1a.2					
Temperate Coastal seawater	San Pedro Ocean Time Series, California	14	33 33N	118 24W	6
Temperate Coastal seawater	San Pedro Ocean Time Series, California	5	33 33N	118 24W	46
Temperate Coastal seawater	Gulf of Biscay pff Cudellero (Spain)	15	43 42N	6 09W	2
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97N	13.67W	1
Sub-arctic seawater	North subarctic Pacific seawater	500	48.65N	126.67W	1
Temperate Coastal seawater	San Pedro Ocean Time Series, California	70	33 33N	118 24W	1
Arctic sea water	North Atlantic, Greenland Sea	50	72.9N	4.7E	4

Temperate Coastal seawater	Oregon coast				5
Mediterranean seawater	Mediterranean off Malaga	5	36 14.610N	4 15.260W	4
Temperate Coastal seawater	Monterey Bay, California	5	35.44 N	124.89 W	1
P1a.3					
Sub-tropical seawater	Hawaiian Ocean Time Series	10	25 45N	158W	3
Temperate Coastal seawater	San Pedro Ocean Time Series, California	14	33 33N	118 24W	5
Temperate Coastal seawater	San Pedro Ocean Time Series, California	5	33 33N	118 24W	16
Temperate Coastal seawater	Gulf of Biscay pff Cudellero (Spain)	15	43 42N	6 09W	1
Temperate Coastal seawater	Monterey Bay, California	5	36.78N	122.48W	1
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	1
Temperate Coastal seawater	San Pedro Ocean Time Series, California	70	33 33N	118 24W	1
Sub-tropical seawater	Sargasso Sea				4
Temperate Coastal seawater	Oregon coast				2
Mediterranean seawater	Mediterranean off Almunecar (Spain)	25	36 42.904N	3 44.444W	3
Mediterranean seawater	Mediterranean off Malaga (Spain)	5	36 14.610N	4 15.260W	1
Mediterranean seawater	Mediterranean off Malaga (Spain)	50	36 14.610N	4 15.260W	7
Mediterranean seawater	Mediterranean off Malaga (Spain)	150	36 14.610N	4 15.260W	1
Mediterranean seawater	Transect from Barcelona to Majorca (Spain) Slope	5	41 08.760N	2 28.010E	5
Mediterranean seawater	Transect from Barcelona to Majorca (Spain) Abyssal plain	5	40 40.270N	2 51.980E	5
Mediterranean seawater	Transect from Barcelona to Majorca (Spain) Abyssal plain	50	40 40.270N	2 51.980E	5
Temperate Coastal seawater	Monterey Bay, California	5	35.44N	124.89W	1
Temperate Coastal seawater	Monterey Bay, California	5			1
Sub-tropical seawater	Sargasso Sea, Bermuda Atlantic Time Series	5	32 10N	64 30W	22
Temperate Coastal seawater	Western Channel Observatory, Plymouth	5	50.25N	4.212W	1

P1b					
Sub-tropical seawater	Hawaiian Ocean Time Series	10	25 45N	158W	2
Mediterranean Seawater	Ionian Sea	3000	36.50N	15.83E	11
Mediterranean Seawater	Ionian Sea	50	36.50N	15.83E	1
Mediterranean seawater	Mediterranean off Almunecar (Spain)	25	36 42.904N	3 44.444W	1
Mediterranean Seawater	Mediterranean off Malaga (Spain)	5	36 14.610N	4 15.260W	3
Mediterranean Seawater	Mediterranean off Malaga (Spain)	50	36 14.610N	4 15.260W	3
Mediterranean Seawater	Mediterranean off Malaga (Spain)	150	36 14.610N	4 15.260W	2
Mediterranean Seawater	Mediterranean off Malaga (Spain)	450	36 14.610N	4 15.260W	1
Temperate Coastal seawater	Monterey Bay, California		35.44 N	124.89 W	1
Arctic sea water	North Atlantic, Greenland Sea	2000	72.9N	4.7E	25
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	6
Sub-arctic seawater	North subarctic Pacific seawater	500	50.00 N	145.00 W	12
Sub-arctic seawater	North subarctic Pacific seawater	500	48.65 N	126.67 W	3
Temperate Coastal seawater	San Pedro Ocean Time Series, California	5	33 33N	118 24W	5
Temperate Coastal seawater	San Pedro Ocean Time Series, California	70	33 33N	118 24W	3
Sub-tropical seawater	Sargasso Sea, Bermuda Atlantic Time Series		32.174835	-64.01017	2
Sub-tropical seawater	Sargasso Sea, Bermuda Atlantic Time Series	5	32 10N	64 30W	1
Mediterranean seawater	Transect from Barcelona to Majorca (Spain) Abyssal plain	400	40 40.270N	2 51.980E	1
Tropical seawater - coastal upwelling	Upwelling, Fernandina Island	12	-0.3011111	91.651665	30
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	3
Sub-arctic seawater	North subarctic Pacific seawater	500	50.00 N	145.00 W	3
Sub-arctic seawater	North subarctic Pacific seawater	500	48.65 N	126.67 W	2
Temperate Coastal seawater	San Pedro Ocean Time Series, California	70	33 33N	118 24W	1
Arctic sea water	North Atlantic,	2000	72.9N	4.7E	18

	Greenland Sea				
Arctic sea water	North Atlantic, Greenland Sea	50	72.9N	4.7E	1
Tropical seawater - coastal upwelling	Upwelling, Fernandina Island	12	-0.3011111	91.651665	2
Temperate Coastal seawater	San Pedro Ocean Time Series, California	150	33 33N	118 24W	2
P2.1					
Tropical seawater	00 Miles west of the Seychelles in the Indian Ocean	5	-4.990278	64.97667	3
Temperate Coastal seawater	Monterey Bay, California		35.44N	124.89W	3
Temperate Coastal seawater	San Pedro Ocean Time Series, California	14	33 33N	118 24W	2
Temperate Coastal seawater	San Pedro Ocean Time Series, California	5	33 33N	118 24W	25
Sub-tropical seawater	Sargasso Sea, Bermuda Atlantic Time Series		32.174835	-64.01017	1
Sub-tropical seawater	Sargasso Sea, Bermuda Atlantic Time Series		31.175	-64.32433	5
Temperate Coastal seawater	South of Charleston, SC		32.506943	-79.263885	1
P2.2					
Antarctic seawater	Antarctic ocean off the Antarctic Peninsula	21	68 24.5S	61 40.717W	5
P2.3					
Temperate Coastal seawater	Gulf of Biscay pff Cudellero (Spain)	15	43 42N	6 09W	1
Sub-tropical seawater	Hawaiian Ocean Time Series	770	25 45N	158W	1
Sub-tropical seawater	Hawaiian Ocean Time Series	4000	25 45N	158W	1
Mediterranean	Ionian Sea	3000	36.50N	15.83E	53

Seawater					
Mediterranean Seawater	Ionian Sea	50	36.50N	15.83E	2
Mediterranean Seawater	Mediterranean off Malaga (Spain)	50	36 14.610N	4 15.260W	1
Mediterranean Seawater	Mediterranean off Malaga (Spain)	150	36 14.610N	4 15.260W	2
Mediterranean Seawater	Mediterranean off Malaga (Spain)	450	36 14.610N	4 15.260W	3
Temperate Coastal seawater	Monterey Bay, California	750			2
Arctic sea water	North Atlantic, Greenland Sea	2000	72.9N	4.7E	12
Arctic sea water	North Atlantic, Greenland Sea	50	72.9N	4.7E	6
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	56
Sub-arctic seawater	North subarctic Pacific seawater	500	50.00 N	145.00 W	48
Sub-arctic seawater	North subarctic Pacific seawater	500	48.65 N	126.67 W	11
Temperate Coastal seawater	San Pedro Ocean Time Series, California	70	33 33N	118 24W	9
Temperate Coastal seawater	San Pedro Ocean Time Series, California	150	33 33N	118 24W	8
Temperate Coastal seawater	San Pedro Ocean Time Series, California	500	33 33N	118 24W	6
Temperate Coastal seawater	San Pedro Ocean Time Series, California	890	33 33N	118 24W	1
Sub-tropical seawater	Sargasso Sea, Bermuda Atlantic Time Series		32.174835	-64.01017	3
Mediterranean	Transect from Barcelona	400	40 40.270N	2 51.980E	4

seawater	to Majorca (Spain) Abyssal plain				
Tropical seawater - coastal upwelling	Upwelling, Fernandina Island	12	-0.3011111	91.651665	7
P3.1					
Tropical seawater	201 miles from F. Polynesia, Tropical South Pacific	50	-10.131389	-135.44945	1
Tropical coastal seawater	Devil's Crown, Floreana Island		-1.2161111	-90.422775	2
Sub-tropical coastal seawater	Kaneohe Bay, Hawaii		0.27222222	-91.63333	1
Tropical coastal seawater	Northeast of Colon, Caribbean Sea		10.716389	-80.25445	1
Sub-tropical seawater	Sargasso Sea Station 11		31.175	-64.32433	1
Tropical coastal seawater	Warm seep, Roca Redonda		27222222	-91.63333	1
Tropical coastal seawater	Wolf Island	1.7	1.3891667	-91.81695	2
P3.2					
ACE lake	Antarctic continent	5	68 28.33S	78 11.29E	1
Temperate coastal seawater	Cape May, NJ	1	38.94	-74.685	1
Temperate Estuarine	Chesapeake Bay, MD	13.2	38.946945	-76.41722	11
Temperate Estuarine	Delaware Bay, NJ	1	39.417778	75.504166	3
Temperate coastal seawater	Gulf of Maine	1	42.503056	-67.24	1
Temperate coastal seawater	Off Nags Head, NC	2.1	36.003887	-75.39472	5

Sub-arctic coastal seawater	Outside Halifax, Nova Scotia	2	44.137222	-63.644444	1
Sub-arctic coastal seawater	Outside Halifax, Nova Scotia	2	44.137222	-63.644444	2
Antarctic seawater	Southern Ocean	2	-55.25	74.25613	1
Antarctic coastal seawater	Southern Ocean	2	-64.831	80.724	1
Antarctic seawater	Southern Ocean	2	-60.00007	141.23353	1
Arctic coastal seawater	Svalbard, Norway	5	79 00.07N	11 18.33E	1
NA2					
Arctic sea water	North Atlantic, Greenland Sea	2000	72.9N	4.7E	9
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	1
Tropical seawater - coastal upwelling	Upwelling, Fernandina Island	12	-0.3011111	91.651665	1
Tropical seawater - coastal upwelling	Upwelling, Fernandina Island	12	-0.3011111	91.651665	2
NA2					
Mediterranean Seawater	Ionian Sea	3000	36.50N	15.83E	2
Mediterranean seawater	Mediterranean off Malaga (Spain)	450	36 14.610N	4 15.260W	1
Arctic sea water	North Atlantic, Greenland Sea	2000	72.9N	4.7E	2
Sub-arctic seawater	North subarctic Pacific seawater	500	48.97 N	130.67 W	1
Sub-arctic seawater	North subarctic Pacific seawater	500	50.00 N	145.00 W	2

Table S2. Metagenome sample sets used in this study. * sample sets generated during this study.

Sample name	MG-RAST ID	Sample date	Environment	Location	Latitude	Longitude	Sequence platform
GS399*	N/A	18/11/09	Sea-Ice	Ross Sea, McMurdo Sound	77 42.43S	166 0.7E	454
GS400*	N/A	21/11/09	Sea-Ice	Ross Sea, McMurdo Sound	77 42.43S	166 0.7E	454
GS401*	N/A	22/11/09	Sea-Ice	Ross Sea, McMurdo Sound	77 42.43S	166 0.7E	454
GS402*	N/A	23/11/09	Ice edge	Ross Sea, McMurdo Sound	77 42.43S	166 0.7E	454
GS369*	N/A	24/01/09	Open Ocean	Ross Sea, McMurdo Sound	77 40.23S	165 60.46E	454
GS370*	N/A	28/01/09	Sea-Ice	Ross Sea, McMurdo Sound	77 42.43S	166 4.52E	454
GS371*	N/A	28/01/09	Under Ice	Ross Sea, McMurdo Sound	77 42.43S	166 0.7E	454
GS372*	N/A	30/01/09	Under Ice	Ross Sea, McMurdo Sound	77 40.27S	166 0.7E	454
GS353*	N/A	30/12/07	Open Ocean	Southern Ocean	67 02.56S	144 40.06E	454
GS389*	N/A	22/10/08	Open Ocean	Southern Ocean	64 48.10S	112 22.40E	454
GS390*	N/A	30/10/08	Open Ocean	Southern Ocean	64 48.93S	080 43.22E	454
GS388*	N/A	20/10/08	Open Ocean	Southern Ocean	63 48.54S	115 09.70E	454
GS392*	N/A	13/12/08	Open Ocean	Southern Ocean	64 10.95S	076 27.22E	454
GS387*	N/A	19/10/08	Open Ocean	Southern Ocean	60 29.58S	120 02.44E	454
GS391*	N/A	12/12/08	Open Ocean	Southern Ocean	68 23.39S	076 39.91E	454
GS349*	N/A	27/12/07	Open Ocean	Southern Ocean	66 33.49S	142 19.00E	454

GS355*	N/A	3/01/08	Open Ocean	Southern Ocean	66 44.85S	144 19.53E	454
GS352*	N/A	29/12/07	Open Ocean	Southern Ocean	66 45.45S	143 17.22E	454
GS347*	N/A	23/12/07	Open Ocean	Southern Ocean	66 0.64S	142 39.48E	454
GS360*	N/A	13/01/08	Open Ocean	Southern Ocean	66 33.95S	140 51.93E	454
GS351*	N/A	28/12/07	Open Ocean	Southern Ocean	66 33.26S	143 19.62E	454
GS348*	N/A	24/12/07	Open Ocean	Southern Ocean	66 20.18S	142 59.14E	454
GS235*	N/A	1/01/07	Coastal	Southern Ocean	66 16.1S	110 32E	454
GS357*	N/A	5/01/08	Open Ocean	Southern Ocean	66 10.16S	142 55.56E	454
GS236*	N/A	7/01/07	Coastal	Southern Ocean	63 52.72S	112 4.2E	454
GS358*	N/A	9/01/08	Open Ocean	Southern Ocean	64 18.00S	150 00.17E	454
GS359*	N/A	12/01/08	Open Ocean	Southern Ocean	66 10.71S	143 29.25E	454
GS362*	N/A	19/01/08	Open Ocean	Southern Ocean	65 32.10S	140 42.70E	454
GS231	4443684.3	21/12/06	Saline lake	Ace Lake	68.473°S	78.189°E	Sanger/454
GS232	4443682.3	21/12/06	Saline lake	Ace Lake	68.473°S	78.189°E	Sanger/454
GS012	4441584.3	18/12/03	Estuary	Chesapeake Bay	38.9469°N	76.4172W	Sanger
GS393*	N/A	15/12/08	Open Ocean	Southern Ocean	55 15.44S	074 15.18E	454
GS386*	N/A	17/10/08	Open Ocean	Southern Ocean	54 56.45S	129 37.11E	454
GS394*	N/A	17/12/08	Open Ocean	Southern Ocean	53 00.74S	073 22.25E	454

GS346*	N/A	20/12/07	Open Ocean	Southern Ocean	59 17.86S	142 26.88E	454
GS363*	N/A	22/01/08	Open Ocean	Southern Ocean	59 59.50S	141 13.51E	454
GS364*	N/A	23/01/08	Open Ocean	Southern Ocean	56 41.36S	141 52.06E	454
GS366*	N/A	24/01/08	Open Ocean	Southern Ocean	52 01.20S	144 03.47E	454
PLY3704	4443704.3	19/05/06	Coastal	Norwegian Fjord	60 16.10N	5 13.20E	454
GS013	4441585.3	19/12/03	Coastal	Nags Head, NC	36 0'14N	75 23'41W	Sanger
GS008	4441583.3	16/11/03	Coastal	Newport Harbour, RI	41 29'9N	71 21'4W	Sanger
Apr_night	4445067.3	22/04/08	Coastal	English Channel	50.2518N	4.2089W	454
PLY3705	4443705.3	19/05/06	Coastal	Norwegian Fjord	60 16.10N	5 13.20E	454
PLY3702	4443702.3	13/05/06	Coastal	Norwegian Fjord	60 16.10N	5 13.20E	454
PLY3703	4443703.3	13/05/06	Coastal	Norwegian Fjord	60 16.10N	5 13.20E	454
Apr_day	4444077.3	22/04/08	Coastal	English Channel	50.2518N	4.2089W	454
Jan_night	4445070.3	28/01/08	Coastal	English Channel	50.2518N	4.2089W	454
GS367	N/A	25/01/08	Open Ocean	Tasman Sea	48 13.96S	145 47.65E	Sanger
GS009	4441143.3	17/11/03	Coastal	Block Island, NY	41 5'28N	71 36'8W	Sanger
GS011	4441658.3	18/11/03	Estuary	Delaware Bay, NJ	39 25'4N	75 30'15W	Sanger
MontBay5	4443716.3	15/05/01	Coastal	Monterey Bay, CA, USA	37 0N	122 0W	
GS006	4441582.3	23/08/03	Estuary	Bay of Fundy, Nova Scotia, CAN	45 6'42N	64 56'48W	Sanger

GS003	4441580.3	21/08/03	Coastal	Browns Bank, Maine, CAN	42 51'10N	66 13'2W	Sanger
GS010	4441144.3	18/11/03	Coastal	Cape May, NJ, USA	38 56'24N	74 41'6W	Sanger
MontBay1	4443713.3	15/10/00	Coastal	Monterey Bay, CA, USA	36.75 N	122 0W	454
MontBay2	4443712.3	15/10/00	Coastal	Monterey Bay, CA, USA	36.75 N	122 0W	454
MontBay3	4443714.3	25/04/01	Coastal	Monterey Bay, CA, USA	36.75 N	122 0W	454
MontBay4	4443715.3	25/04/01	Coastal	Monterey Bay, CA, USA	36.75 N	122 0W	454
MontBay6	4443717.3	15/05/01	Coastal	Monterey Bay, CA, USA	36.75 N	122 0W	454
GS368	N/A	26/01/08	Open Ocean	Southern Ocean	44 42.54S	145 45.15E	Sanger
GS005	4441581.3	22/08/03	Embayment	Bedford Basin, Nova Scotia, CAN	44 41'25N	63 38'14W	Sanger
Aug4am	4445066.3	26/08/08	Coastal	English Channel	50.2518N	4.2089W	454
Aug10pm	4445065.3	26/08/08	Coastal	English Channel	50.2518N	4.2089W	454
Aug4pm	4445068.3	26/08/08	Coastal	English Channel	50.2518N	4.2089W	454
BBAY01	4443688.3	18/10/06	Coastal	Botany Bay, Australia	33 59'27S	151 13'55E	Sanger
GS004	4441152.3	22/08/03	Coastal	Halifax, Nova Scotia, CAN	44 8'14N	63 38'40W	Sanger
GS007	4441153.3	25/08/03	Coastal	Northern Gulf of Maine, Canada	43 37'56N	66 50'50W	Sanger
GS002	4441579.3	21/08/03	Coastal	Gulf of Maine, Canada	42 30'11N	67 14'24W	Sanger
BBAY02	4443689.3	18/10/06	Coastal	Botany Bay, Australia	33 59'27S	151 13'55E	Sanger
GS014	4441659.3	20/12/03	Coastal	South of Charleston, SC	32 30'25N	79 15'50W	Sanger

GS031	4441597.3	10/02/04	Coastal Upwelling	Fernandina Island, Galapagos Islands	0 18'4S	91 39'6W	Sanger
GS000c	4441574.3	25/02/03	Open Ocean	Sargasso Sea, Station 3	32 10'29.4N	64 00'36.6W	Sanger
GS000d	4441575.3	26/02/03	Open Ocean	Sargasso Sea, Station 13	31 10'30N	64 19'27.6W	Sanger
NASB1741	4443725.3	N/A	Open Ocean	Sargasso Sea	32 10'29.4N	64 00'36.6W	454
NASB1742	4443729.3	N/A	Open Ocean	Sargasso Sea	32 10'29.4N	64 00'36.6W	454
NASB1791	4443731.3	N/A	Open Ocean	Sargasso Sea	32 10'29.4N	64 00'36.6W	454
NASB1792	4443732.3	N/A	Open Ocean	Sargasso Sea	32 10'29.4N	64 00'36.6W	454
GS122a	4441615.3	30/09/05	Open Ocean	Indian Ocean	30 53'54S	40 25'13E	Sanger
GS123	4441616.3	1/10/05	Open Ocean	Indian Ocean	32 23'57S	36 35'31E	Sanger
GS000b	4441572.3	26/02/03	Open Ocean	Sargasso Sea, Station 11	31 10'30N	64 19'27.6W	Sanger
GS035	4441601.3	1/03/04	Coastal	Wolf Island, Galapagos Islands	1 23'21N	91 49'1W	Sanger
GS120	4441135.3	27/09/05	Open Ocean	Madagascar, Indian Ocean	26 2'6S	50 7'23E	Sanger
GS001a	4441576.3	15/05/03	Open Ocean	Sargasso Sea, Hydrostation S	32 10'00N	64 30'00W	Sanger
GS001c	4441578.3	25/02/03	Open Ocean	Sargasso Sea, Hydrostation S	32 10'00N	64 30'00W	Sanger

GS121	4441614.3	29/09/06	Open Ocean	Indian Ocean	29 20'56S	43 12'56E	Sanger
GS119	4441568.3	26/09/05	Open Ocean	Indian Ocean	23 12'58S	52 18'22E	Sanger
GS015	4441586.3	8/01/04	Coastal	Caribbean Sea	24 29'18N	83 4'12W	Sanger
GS027	4441595.3	4/02/04	Coastal	Floreana Island, Galapagos Islands	1 12'58S	90 25'22W	Sanger
GS036	4441602.3	2/03/04	Coastal	Isabella Island, Galapagos Islands	0 1'15S	91 11'52W	Sanger
GS108a	4441139.3	3/08/05	Lagoon Reef	Cocos Keeling, Indian Ocean	12 5'33S	96 52'54E	Sanger
GS028	4441596.4	4/02/04	Coastal	Floreana Island	1 13'1S	90 19'11W	Sanger
GS029	4441596.3	8/02/04	Coastal	Santiago Island	0 12'0S	90 50'7W	Sanger
GS116	4441149.3	17/08/05	Open Ocean	Indian Ocean	4 38'6S	56 50'10E	Sanger
GS016	4441660.3	8/01/04	Coastal	Caribbean Sea	24 10'29N	84 20'40W	Sanger
GS111	4441156.3	7/08/05	Open Ocean	Indian Ocean	9 35'49S"	84 11'51E"	Sanger
GS117a	4441613.3	9/09/05	Coastal	St. Anne Island, Seychelles	4 36'49S"	55 30'31E"	Sanger
GS112a	4441609.3	8/08/05	Open Ocean	Indian Ocean	8 30'18S"	80 22'32E"	Sanger
GS030	4442626.3	9/02/04	Warm Seep	Galapagos Island	0 16'20N"	91 38'0W"	Sanger
GS030a	4441662.3	9/02/04	Warm Seep	Galapagos Island	0 16'20N"	91 38'0W"	Sanger
GS148	4441617.3	11/09/05	Fringing Reef	Indian Ocean	6 19'S	39 33E"	Sanger
GS149	4441618.3	12/09/05	Harbour	Zanzibar	6 7'S	39 07'E	Sanger
GS017	4441587.3	9/09/05	Open Ocean	Caribbean Sea	20 31'21N"	85 24'49W"	Sanger

GS110	4441607.3	6/08/05	Open Ocean	Indian Ocean	10 26'46S"	88 18'10E"	Sanger
GS109	4441155.3	5/08/05	Open Ocean	Indian Ocean	10 56'37S"	92 3'32E"	Sanger
GS051	4441604.3	22/05/04	Coral Atoll	Polynesian Archipelago	15 8'37S"	147 26'6W"	Sanger
GS018	4441588.3	10/01/04	Open Ocean	Caribbean Sea	18 2'12N"	83 47'5W"	Sanger
S_35155	4443697.3	25/04/07	Open Ocean	Pacific Ocean	1 40'N	163 40'W	454
S_35171	4443700.3	27/04/07	Open Ocean	Pacific Ocean	16 00'S	170 00'W	454
GS034	4441600.3	19/02/04	Coastal	Galapagos Islands	0 22'59S"	90 16'47W"	Sanger
GS113	4441610.3	9/08/05	Open Ocean	Indian Ocean	7 0'27S"	76 19'53E"	Sanger
Bios-10m	4449104.3	Jun-2010	Open Ocean	Sargasso Sea	32 10'29.4N	64 00'36.6W	454
GS021	4441591.3	19/01/04	Coastal	Gulf of Panama	8 7'45N"	79 41'28W"	Sanger
S_35162	4443698.3	26/04/07	Open Ocean	Pacific Ocean	5 8'N	161 97'W	454
S_35163	4443699.3	26/04/07	Open Ocean	Pacific Ocean	5 5'N	161 58'W	454
GS019	4441589.3	12/01/04	Coastal	Caribbean Sea	10 42'59N"	80 15'16W"	Sanger
GS026	4441594.3	1/02/04	Open Ocean	Galapagos Islands	1 15'51N"	90 17'42W"	Sanger
GS115	4441150.3	16/08/05	Open Ocean	Indian Ocean	4 39'45S"	60 31'23E"	Sanger
GS037	4441145.3	17/03/04	Open Ocean	Pacific Ocean	1 58'26S"	95 0'53W"	Sanger
S_35179	4443701.3	29/04/07	Open Ocean	Pacific Ocean	13 29'N	159 53'W	454
GS114	4441611.3	15/08/05	Open Ocean	Indian Ocean	4 59'25S"	64 58'36E"	Sanger

GS047	4441146.3	28/03/04	Open Ocean	Pacific Ocean	10 7'53S"	135 26'58W"	Sanger
GS023	4441661.3	21/01/04	Open Ocean	Pacific Ocean	5 38'24N"	86 33'55W"	Sanger
GS049	4441605.3	18/05/05	Coastal	Polynesian Archipelago	17 27'11S"	149 47'56W"	Sanger
GS048a	4441603.3	17/05/04	Coral Reef	Polynesian Archipelago	17 28'33S"	149 48'44W"	Sanger
S_35131	4443766.3	21/04/07	Open Ocean	Pacific Ocean	16 00'S	170 00'W	454
GS022	4441592.3	20/01/04	Open Ocean	Pacific Ocean	6 29'34N"	82 54'14W"	Sanger
S_35139	4443701.3	23/04/07	Open Ocean	Pacific Ocean	7 31'S	167 00'W	454

Table S3. Environmental data associated with metagenomes used in this study.

Sample	Temp (°C)	Depth of water column (m)	Salinity	Chlorophyll $l\alpha$
GS399*	-2	700	3.3 ^a	4.52 ^a
GS400*	-2	700	3.3 ^a	4.52 ^a
GS401*	-2	700	3.3 ^a	4.52 ^a
GS402*	-2	700	3.3 ^a	4.52 ^a
GS369*	-2	300	3.35	4.46
GS370*	-2	300	3.6	3.17
GS371*	-2	300	3.3	4.72
GS372*	-2	300	3.2	5.23
GS353*	-1.8	178	3.45	0
GS389*	-1.76	500	3.47 ^a	0.14
GS390*	-1.74	116	3.43	0.32
GS388*	-1.73	2500	3.4 ^a	1.5
GS392*	-1.53	3847	3.36	0.04
GS387*	-1.5	3200	3.45 ^a	0.22
GS391*	-1.4	378	3.42	5
GS349*	-1.3	365	3.4	3.7
GS355*	-0.89	891	3.4	8.4
GS352*	-0.84	169	3.4	1
GS347*	-0.7	443	3.4	3.2
GS360*	-0.7	308	3.41	7.5
GS351*	-0.7	597	3.4	1.6
GS348*	-0.6	649	3.42	12.6
GS235*	-0.5	60	3.39	8.6
GS357*	-0.5	533	3.42	2.7
GS236*	-0.2	2500	3.37	12.1
GS358*	-0.03	3561	3.35	0.3
GS359*	0.05	364	3.41	2.4
GS362*	0.8	1027	3.62	0.2
GS231	1	25	2.2	0.8
GS232	1	25	2.2	1.0
GS012	1	25	0.347	21.0
GS393*	2.0	2246	3.39	0.5
GS386*	2.0	3200	3.38	0.22
GS394*	2.4	100	3.39	0.6
GS346*	2.9	3294	3.37	0.3
GS363*	3.5	4473	3.37	0.1
GS364*	4.2	3693	3.37	0.5
GS366*	7.7	3180	3.38	0.3
PLY3704	9.0	N/A	3.47 ^a	2.8 ^a
GS013	9.3	20	3.06 ^a	3.0

GS008	9.4	12	2.65	2.2
Aprnight	9.6	50	3.5	1.32
PLY3705	9.7	N/A	3.5 ^a	2.73 ^a
PLY3702	9.7	N/A	3.5 ^a	2.73 ^a
PLY3703	9.7	N/A	3.5 ^a	2.73 ^a
Aprday	9.7	50	3.5	0.85
Jannight	10.0	50	3.42	0.85
Jannight	10.1	50	3.42	0.85
GS367	10.9	3490	3.44	0.2
GS009	11	32	3.1	4.0
GS011	11	8	3.01 ^a	4.8
MontBay5	11	1000	3.3	1.83 ^a
GS006	11.2	11	3.26 ^a	2.8
GS003	11.7	119	2.99	1.4
GS010	12	10	3.1	2.0
MontBay1	14	1000	3.3	2.1 ^a
MontBay2	14	1000	3.3	2.1 ^a
MontBay3	14	1000	3.3	2.1 ^a
MontBay4	14	1000	3.3	2.1 ^a
MontBay6	14	1000	3.3	2.1 ^a
GS368	14.2	3201	3.47	1.3
GS005	15	64	3.02	6.0
Aug4am	15.7	50	3.3	9.8
Aug10pm	15.8	50	3.3	8.17
Aug4pm	15.9	50	3.21	9.24
BBAY01	17.1	2	3.36	0.8
GS004	17.3	142	2.83	0.4
GS007	17.9	139	3.17	1.4
GS002	18.2	106	2.92	1.4
BBAY02	18.3	2	3.62	0.82
GS014	18.6	31	3.34 ^a	1.7
GS031	18.6	19.6	2.95 ^a	0.35
GS000c	19.8	4200	3.67	0.17
GS000d	20	4200	3.66	0.17
NASB174 1	20	N/A	3.6	0.1 ^a
NASB174 2	20	N/A	3.6	0.1a
NASB179 1	20	N/A	3.6	0.1 ^a
NASB179 2	20	N/A	3.6	0.1 ^a
GS122a	20.2	4921	3.58	0.151
GS123	20.4	1860	3.58	0.228
GS000b	20.5	4200	3.66	0.17

GS035	21.8	71	3.45	0.28
GS120	22.5	5081	3.56	0.117
GS001a	22.9	4200	3.67	0.1
GS001c	22.9	4200	3.67	0.1
GS121	23.1	4309	3.54	0.14
GS119	23.8	2995	3.54	0.08
GS015	25	47	3.6	0.2
GS027	25.5	2.3	3.49	0.4
GS036	25.8	67	3.46	0.65
GS108a	25.8	7	3.24	0.11
GS028	26	156	3.12 ^a	0.35
GS029	26.2	12	3.45	0.40
GS116	26.2	2150	3.31	0.29
GS016	26.4	3333	3.58	0.16
GS111	26.4	3841	3.23	0.20
GS117a	26.4	4513	3.58	0.13
GS112a	26.6	4573	3.25	0.134
GS030	26.9	19	3.1 ^a	1.04 ^a
GS030a	26.9	19	3.1 ^a	1.04 ^a
GS148	27	1	3.17 ^a	1.19 ^a
GS149	27	5	3.17 ^a	1.18 ^a
GS017	27	4513	3.58	0.13
GS110	27	1220	3.27	0.128
GS109	27.2	4573	3.26	0.138
GS051	27.3	10	3.42	0.8 ^a
GS018	27.4	4470	3.54	0.14
S_35155	27.4	4200	3.53	0.01 ^a
S_35171	27.5	4200	3.47	0.01 ^a
GS034	27.5	35	3.16 ^a	0.36
GS113	27.5	4573	3.33	0.24
GS021	27.6	76	3.07	0.5
Bios-10m	27.58	3300	3.66 ^a	0.37
S_35162	27.6	4200	3.48	0.01 ^a
S_35163	27.6	4200	3.48	0.01 ^a
GS019	27.7	3336	3.54	0.23
GS026	27.8	2386	3.26	0.22
GS115	27.9	3220	3.32	0.14
GS037	28	3334	3.27 ^a	0.21
S_35179	28.2	4200	3.44	0.01 ^a
GS114	28.2	3649	3.31	0.14
GS047	28.6	2400	3.73	0.01 ^a
GS023	28.7	1139	3.26	0.03
GS049	28.8	900	3.26	0.09
GS048a	28.9	34	3.51	0.09
S_35131	29.1	4200	3.52	0.01 ^a

GS022	29.3	2431	3.23	0.33
S_35139	30.1	4200	3.53	0.01 ^a

* sample sets generated during this study.

^a Estimated using expectation maximization algorithm in the Primer V6 + PERMANOVER software (Clarke and Gorley, 2006).

Table S4. Results of DISTLM fitting environmental data to relative abundances of SAR11 phylotypes in 128 surface marine metagenome samples. Marginal tests identify the explanatory power of each variable in isolation. The BEST solution identifies the variables best describing the observed phylotype distribution for any given number of combinations. Analyses were carried out with both the full sample set (top) and also on only the 79 samples for which data on all variables was available. SS = sum of squares; Pseudo-F (multivariate analogue of Fishers F ratio, as the pseudo-F statistic gets larger, the likelihood of the null hypothesis being true [i.e. no relationship between phylotype distribution and environmental variables], diminishes); Prop. = proportion of variation explained by each variable or combination of variables; P = significance by permutation; No. Var = number of the variable; DWC = depth of the water column.

Full set: 128 metagenomes						
Marginal Test						
Variable	SS (trace)	Pseudo-F	Prop.	P		
DWC	30963	13.425	0.097	0.001		
Salinity	5706	2.2748	0.0178	0.089		
Chlorophyll	43655	19.8	0.13674	0.001		
Latitude	1.83E+05	167.86	0.57318	0.001		
Temperature	1.84E+05	169	0.57483	0.001		
BEST solution						
R ²	RSS	No. Var	Selections		Variable	
0.57483	1.36E+05	1	4		1	DWC
0.59675	1.29E+05	2	4,5		2	Salinity
0.61536	1.23E+05	3	1,4,5		3	Chlorophyll
0.62373	1.20E+05	4	1,2,4,5		4	Temperature
0.6262	1.19E+05	5	All		5	Latitude
Fit Last in Sequential Test						
Variable	SS (trace)	Pseudo-F	Prob.	P		
Temperature	6693.7	6.8321	2.0967E-2	0.001		
Latitude	7231.8	7.3813	2.2652E-2	0.001		
DWC	6307.8	6.4382	1.9758E-2	0.001		
Salinity	2054.7	2.0972	6.436E-3	0.107		
Chlorophyll	786.53	0.80279	2.4637	0.501		
Analysis of P1a clade only						
Marginal Test	SS (trace)	Pseudo-F	Prop.	P		
DWC	36557	16.59	0.11885	0.001		
Salinity	3187.7	1.288	1.0363E-2	0.295		
Chlorophyll	50053	23.905	0.16272	0.001		
Latitude	2.0681E5	252.37	0.67232	0.001		
Temperature	2.0259E5	237.29	0.65861	0.001		

Subset: 79 metagenomes with all variables available						
Marginal Test						
Variable	SS (trace)	Pseudo-F	Prop.	P		
DWC	17901	8.7738	0.120229	0.002		
Salinity	2607.5	1.1646	1.49E-02	0.315		
Chlorophyll	15846	7.6658	9.05E-02	0.002		
Temperature	1.09E+05	127.09	0.62271	0.001		
Latitude	1.08E+05	122.78	0.61458	0.001		
BEST solution						
R ²	RSS	No. Var	Selections		Variable	
0.62271	66029	1	4		1	DWC
0.64583	61983	2	4,5		2	Salinity
0.6699	57769	3	1,4,5		3	Chlorophyll
0.68259	55549	4	1,2,4,5		4	Temperature
0.68732	54721	5	All		5	Latitude
Fit Last in Sequential Test						
Variable	SS (trace)	Pseudo-F	Prop.	P		
Temperature	5663.1	7.5547	3.2359E-2	0.001		
Latitude	4374.6	5.8359	2.4997E-2	0.004		
DWC	4179.8	5.576	2.3883E-2	0.002		
Salinity	2116.2	2.8231	1.2092E-2	0.044		
Chlorophyll	827.8	1.1043	4.7301E-3	0.366		
Analysis of P1a clade only						
Marginal Test	SS (trace)	Pseudo-F	Prop.	P		
DWC	22521	10.787	0.12287	0.001		
Salinity	863.05	0.36429	4.7088E-03	0.755		
Chlorophyll	19653	9.2483	0.10723	0.001		
Latitude	1.2419E5	161.83	0.6776	0.001		
Temperature	1.1958E5	144.53	0.65242	0.001		

Table S5. Specific genes, paralogous genes, and genes with signatures of positive selection in ACE_P1a.1, ACE_P3.2, and IMCC9063 genomes. For positive selection comparisons are between ACE_P3.2 vs HIMB114, ACE_P1a.1 vs HTCC7211, and IMCC9063 vs HIMB114. NA, not available. For positive selection, * = maximum dN/dS ratio over a 60 codon window.

Genes specific to ACE_P1a.1			
Gene ID	Annotation	COG	COG Category
feature:167713860	hypothetical protein	NA	NA
feature:167713862	hypothetical protein	NA	NA
feature:167713864	hypothetical protein RHECIAT_CH0003834	NA	NA
feature:167713866	hypothetical protein	NA	NA
feature:167713868	hypothetical protein	NA	NA
feature:167713870	hypothetical protein	NA	NA
feature:167713872	hypothetical protein	NA	NA
feature:167713874	hypothetical protein	NA	NA
feature:167713876	probable integrase protein	COG0582	[L]
feature:167713882	lipoprotein precursor	NA	NA
feature:167713894	cell division protein FtsI	NA	NA
feature:167713902	phospho-N-acetylmuramoyl-pentapeptide-transferase	COG0472	[M]
feature:167713906	hypothetical protein	NA	NA
feature:167713920	cell division protein FtsA	COG0849	[D]
feature:167713922	cell division protein FtsA	NA	NA
feature:167713924	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase	COG0774	[M]
feature:167713930	Xaa-Pro aminopeptidase	COG0006	[E]
feature:167713932	hypothetical protein	NA	NA
feature:167713936	hypothetical protein	NA	NA
feature:167713938	hypothetical protein	NA	NA
feature:167713940	Fumarate hydratase class II (Fumarase C)	COG0114	[C]
feature:167713944	Fumarate hydratase class II (Fumarase C)	COG0114	[C]
feature:167713948	hypothetical protein	NA	NA
feature:167713950	T4-like clamp loader subunit	COG0470	[L]
feature:167713952	hypothetical protein	NA	NA
feature:167713954	hypothetical protein	NA	NA
feature:167713956	hypothetical protein	NA	NA
feature:167713958	beta-lactamase	COG1680	[V]
feature:167713960	hypothetical protein	NA	NA
feature:167713962	hypothetical protein	NA	NA
feature:167713964	hypothetical protein	NA	NA
feature:167713966	hypothetical protein	NA	NA
feature:167713970	hypothetical protein	NA	NA
feature:167713972	hypothetical protein	NA	NA
feature:167713974	hypothetical protein	NA	NA
feature:167713982	probable integral membrane proteinase	COG0330	[O]
feature:167713986	periplasmic serine protease DO-like precursor	COG0265	[O]
feature:167714002	rod shape-determining protein MreD	NA	NA
feature:167714014	putative SAM-dependent methyltransferase	COG0500	[QR]
feature:167714022	LysE type translocator	NA	NA
feature:167714058	hypothetical protein	NA	NA
feature:167714060	pyrroline-5-carboxylate reductase	COG0345	[E]
feature:167714076	hypothetical protein	NA	NA

feature:167714078	pyrroline-5-carboxylate reductase	COG0345	[E]
feature:167714096	hypothetical protein	NA	NA
feature:167714110	X-Pro dipeptidase	COG0006	[E]
feature:167714112	spermidine/putrescine transport ATP-binding protein	COG3842	[E]
feature:167714134	putative protein of unknown function (DUF989)	COG3748	[S]
feature:167714136	putative cytidine and deoxycytidylate deaminase zinc-binding region	COG0590	[FJ]
feature:167714138	putative aldehyde oxidase and xanthine dehydrogenase molybdopterin binding domain protein	COG4631	[F]
feature:167714140	putative aldehyde oxidase and xanthine dehydrogenase molybdopterin binding domain protein	COG4631	[F]
feature:167714142	putative FAD binding domain in molybdopterin dehydrogenase	COG4630	[F]
feature:167714144	putative transthyretin precursor (formerly prealbumin)	COG2351	[R]
feature:167714146	conserved hypothetical protein	COG3195	[S]
feature:167714148	putative cytochrome c	COG2010	[C]
feature:167714150	hypothetical protein Adeh_3104	COG1975	[O]
feature:167714152	XdhC/CoxI family protein	COG1975	[O]
feature:167714154	hypothetical protein NEISUBOT_02557	COG1816	[F]
feature:167714156	ureidoglycolate hydrolase	COG3194	[F]
feature:167714158	putative allantoicase	COG4266	[F]
feature:167714160	putative polysaccharide deacetylase family protein	COG0726	[G]
feature:167714178	AFG1-like ATPase	NA	NA
feature:167714190	hypothetical protein	NA	NA
feature:167717472	NAD-dependent epimerase/dehydratase	COG0451	[MG]
feature:167717474	hypothetical protein	NA	NA
feature:167717476	hypothetical protein	NA	NA
feature:167717482	HemK	COG2890	[J]
feature:167717514	cytochrome C	COG3474	[C]
feature:167717530	hypothetical protein PU1002_03846	NA	NA
feature:167717574	hypothetical protein PU1002_03956	NA	NA
feature:167717576	hypothetical protein SAR11_0474	NA	NA
feature:167717580	deoxycytidine triphosphate deaminase	COG0717	[F]
feature:167717582	hypothetical protein SAR11_0470	NA	NA
feature:167717584	hypothetical protein	NA	NA
feature:167717586	Glycosyl transferase family 2	COG0463	[M]
feature:167717588	hypothetical protein	NA	NA
feature:167717590	hypothetical protein WH5701_14161	NA	NA
feature:167717592	perosamine synthetase	COG0399	[M]
feature:167717594	deoxycytidine triphosphate deaminase	COG0717	[F]
feature:167717598	glycosyl transferase family protein	COG0463	[M]
feature:167717600	DegT/DnrJ/EryC1/StrS aminotransferase	NA	NA
feature:167717602	hypothetical protein	NA	NA
feature:167717604	hypothetical protein	NA	NA
feature:167717606	hypothetical protein	NA	NA
feature:167717608	hypothetical protein	NA	NA
feature:167717610	Dolichyl-phosphate beta-D-mannosyltransferase	COG0463	[M]
feature:167717614	hypothetical protein	NA	NA
feature:167717616	deoxycytidine triphosphate deaminase	COG0717	[F]
feature:167717622	ribosomal protein L34 (rpmH)	NA	NA
feature:167717640	nuclear protein SET	COG2940	[R]
feature:167717642	tRNA pseudouridine synthase A	COG0101	[J]
feature:167717650	hypothetical protein	NA	NA

feature:167717652	hypothetical protein PU1002_04056	NA	NA
feature:167717658	BolA-like protein	NA	NA
feature:167717662	Cobalamin biosynthesis protein CobS	COG0714	[R]
feature:167717688	translation initiation factor IF-3	COG0290	[J]
feature:167717696	nucleoside-diphosphate-sugar epimerase	COG0451	[MG]
feature:167717766	hypothetical protein PU1002_04341	COG3807	[S]
feature:167717780	polyribonucleotide nucleotidyltransferase	COG1185	[J]
feature:167717782	hypothetical protein	NA	NA
feature:167717784	polyribonucleotide nucleotidyltransferase	COG1185	[J]
feature:167717796	S-adenosylmethionine synthetase	COG0192	[H]
feature:167717798	S-adenosylmethionine synthetase	COG0192	[H]
feature:167717812	polysaccharide deacetylase	COG0726	[G]
feature:167717826	phospholipid-binding domain-containing protein	NA	NA
feature:167717840	Chaperone protein	COG0484	[O]
feature:167717866	hypothetical protein	NA	NA
feature:167717872	ParA family ATPase for plasmid partitioning and ot her plasmid related functions	COG1192	[D]
feature:167717874	ParA family ATPase for plasmid partitioning and ot her plasmid related functions	COG1192	[D]
feature:167717876	Glucose inhibited division protein	COG0357	[M]
feature:167717890	ferrochelatae	COG0276	[H]
feature:167717894	hypothetical protein	NA	NA
feature:167719272	Phage integrase	NA	NA
feature:167719274	hypothetical protein	NA	NA
feature:167719276	hypothetical protein	NA	NA
feature:167719278	hypothetical protein	NA	NA
feature:167719280	hypothetical protein	NA	NA
feature:167719282	hypothetical protein	NA	NA
feature:167719284	hypothetical protein	NA	NA
feature:167719286	N-6 DNA methylase	COG0286	[V]
feature:167719288	type I restriction-modification system specificity subunit	NA	NA
feature:167719290	type I site-specific deoxyribonuclease HsdR famil y protein	COG0610	[V]
feature:167719292	hypothetical protein	NA	NA
feature:167719294	HsdR family type I site-specific deoxyribonuclease	COG0610	[V]
feature:167719296	hypothetical protein	NA	NA
feature:167719310	hypothetical protein	NA	NA
feature:167719314	N6-adenine-specific methylase	COG0742	[L]
feature:167719332	hypothetical protein SAR11_0144	NA	NA
feature:167719338	hypothetical protein MB2181_02345	NA	NA
feature:167719342	transcriptional regulator	COG1733	[K]
feature:167719358	surfeit locus protein 1	COG3346	[S]
feature:167719360	electron transport protein SCO1/SenC	COG1999	[R]
feature:167719362	COG1845: Heme/copper-type cytochrome/quinol oxidase subunit 3	NA	NA
feature:167719370	hypothetical protein GP2143_03903	COG5605	[S]
feature:167719372	hypothetical protein	NA	NA
feature:167719384	ribonuclease H (RNase H)	COG0328	[L]
feature:167719388	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	COG0761	[IM]
feature:167719422	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	NA	NA
feature:167719436	ubiquinol-cytochrome-c reductase	COG2857	[C]
feature:167719484	putative D-ala D-ala ligase	COG1181	[M]
feature:167719486	carbamoyl-phosphate synthase large chain	COG0458	[EF]
feature:167719488	carbamoyl-phosphate synthase large chain	COG0458	[EF]

feature:167719492	Carbamoyl-phosphate synthase large chain	COG0458	[EF]
feature:167719686	sulfatase	COG2194	[R]
feature:167719690	hypothetical protein	NA	NA
feature:167719706	hypothetical protein	NA	NA
feature:167719708	NADPH:quinone reductase	COG0604	[CR]
feature:167719712	hypothetical protein PU1002_00400	COG5467	[S]
feature:167719714	hypothetical protein	NA	NA
feature:167719718	dienelactone hydrolase	COG0412	[Q]
feature:167719720	conserved hypothetical protein	NA	NA
feature:167719724	hypothetical protein SAR11_1729	COG5515	[S]
feature:167719728	hypothetical protein	NA	NA
feature:167719732	hypothetical protein RD1_1866	NA	NA
feature:167719736	hypothetical protein RSK20926_01912	COG2072	[P]
feature:167719744	hypothetical protein Arad_9834	COG2072	[P]
feature:167719746	endoribonuclease L-PSP family protein	COG0251	[J]
feature:167719748	hypothetical protein RB2150_08714	COG3342	[S]
feature:167719752	hypothetical protein	NA	NA
feature:167719754	hypothetical protein PU1002_00600	NA	NA
feature:167719756	hypothetical protein PU1002_00605	NA	NA
feature:167719770	hypothetical protein	NA	NA
feature:167719798	glutaredoxin	COG0278	[O]
feature:167719814	hypothetical protein	NA	NA
feature:167719920	hypothetical protein PU1002_01000	NA	NA
feature:167719924	hypothetical protein	NA	NA
feature:167719952	Novel protein with potential Cupin domain	COG0662	[G]
feature:167719986	recombination protein O	NA	NA
feature:167720012	hypothetical protein SAR11_1039	NA	NA
feature:167720024	ribosomal protein L33	NA	NA
feature:167720028	cytochrome c oxidase assembly protein (isoform 1)	COG1612	[O]
feature:167720036	N-acetyl-gamma-glutamyl-phosphate reductase	COG0002	[E]
feature:167720038	hypothetical protein SAR11_1026	NA	NA
feature:167720044	hypothetical protein	NA	NA
feature:167720052	hypothetical protein	NA	NA
feature:167720054	hypothetical protein	COG0500	[QR]
feature:167720056	hypothetical protein BACUNI_00980	COG0463	[M]
feature:167733984	FeS assembly ATPase SufC	COG0396	[O]
feature:167733990	alpha/beta hydrolase	COG2945	[R]
feature:167734006	ThiS family protein	NA	NA
feature:167734022	ATPase involved in DNA replication initiation	NA	NA
feature:167734054	Guanylate kinase	COG0194	[F]
feature:167734056	hypothetical protein	NA	NA
feature:167734086	hypothetical protein SAR11_0694	COG2081	[R]
feature:167734088	SlyX	NA	NA
feature:167734108	NifU-like domain-containing protein	COG0822	[C]
feature:167734140	glycine cleavage system protein T2	COG0404	[E]
feature:167774839	hypothetical protein	NA	NA
feature:167774841	hypothetical protein	NA	NA
feature:167774843	hypothetical protein	NA	NA
feature:167774845	hypothetical protein	NA	NA
feature:167774849	isocitrate dehydrogenase	COG0538	[C]
feature:167774859	FtsJ-like methyltransferase	COG0293	[J]
feature:167774869	hypothetical protein MBMO_EB80-69G07.0030	COG0861	[P]
feature:167774885	glycine dehydrogenase	COG0403	[E]

feature:16774891	putative glycine cleavage system P-protein	COG1003	[E]
feature:167818776	SurA-like protein	NA	NA
feature:167818778	hypothetical SurA-like protein	NA	NA
feature:167818792	lipid-A-disaccharide synthase (lpxB)	COG0763	[M]
feature:167818804	outer membrane protein	NA	NA
feature:167818808	phosphatidate cytidyltransferase	NA	NA
feature:167818812	Ribosome recycling factor	COG0233	[J]
feature:167818814	Ribosome recycling factor	COG0233	[J]
feature:167818816	protein translation elongation factor Ts (EF-Ts)	COG0264	[J]
feature:167818820	DNA polymerase III alpha subunit	COG0587	[L]
feature:167818834	biotin-protein ligase	NA	NA
feature:167818836	biotin-protein ligase	COG0340	[H]
feature:167818872	hypothetical protein	NA	NA
feature:167818878	hypothetical protein SAR11_0877	COG3313	[R]
feature:167818880	possible transmembrane protein	NA	NA
feature:167818884	hypothetical protein PU1002_02026	NA	NA
feature:167818906	TerC family integral membrane protein	COG0861	[P]
feature:167818908	short chain dehydrogenase	COG1028	[IQR]
feature:167818910	L-xylulose reductase	COG1028	[IQR]
feature:167818918	putative membrane protein	COG0390	[R]
feature:167818920	hypothetical protein mll5080	NA	NA
feature:167818922	hypothetical protein	NA	NA
feature:167818924	hypothetical protein	NA	NA
feature:167818926	integrase family protein	COG0582	[L]
feature:167818928	hypothetical protein SAR11_1220	NA	NA
feature:167818930	sarcosine dehydrogenase	COG0404	[E]
feature:167818932	ATP-binding component of ABC transporter	COG1131	[V]
feature:167818934	ATP-binding component of ABC transporter	COG1131	[V]
feature:167818938	sarcosine dehydrogenase	COG0665	[E]
feature:167818940	conserved hypothetical protein	NA	NA
feature:167818942	cytochrome c'	NA	NA
feature:167818948	ADP-ribosylglycohydrolase	NA	NA
feature:167818952	Enoyl-CoA hydratase/isomerase	COG1024	[I]
feature:167818964	hypothetical protein BAL199_27271	NA	NA
feature:167818966	hypothetical protein CPS_3856	COG0730	[R]
feature:167818968	gamma-glutamyltranspeptidase	COG0405	[E]
feature:167818970	gamma-glutamyltransferase	COG0405	[E]
feature:167818972	Xaa-Pro aminopeptidase	COG0006	[E]
feature:167818974	Xaa-Pro aminopeptidase	COG0006	[E]
feature:167818976	Aldehyde Dehydrogenase	COG4230	[C]
feature:167818982	hypothetical protein MADE_00086	COG0730	[R]
feature:167818984	COG0607: Rhodanese-related sulfurtransferase	COG0607	[P]
feature:167818986	hypothetical protein SAR11_1230	NA	NA
feature:167818994	AzIC protein	NA	NA
feature:167819000	iron(III) ABC transporter	COG1178	[P]
feature:167819006	iron(III) ABC transporter	COG3842	[E]
feature:167819008	iron(III) ABC transporter	COG1178	[P]
feature:167819020	hypothetical protein SAR11_1247	COG3651	[S]
feature:167819024	hypothetical protein SAR11_1248	NA	NA
feature:167819028	hypothetical protein	NA	NA
feature:167819030	hypothetical protein SAR11_1249	NA	NA
feature:167819038	dimethylglycine dehydrogenase	COG0665	[E]
feature:167819076	hypothetical protein RCCS2_14234	COG0697	[GER]

feature:167819082	hypothetical protein PU1002_06946	NA	NA
feature:167819084	hypothetical protein PU1002_06941	COG2391	[R]
feature:167819086	hypothetical protein SAR11_1258	COG2391	[R]
feature:167819090	enoyl-CoA hydratase	COG1024	[I]
feature:167819094	iron(III) ABC transporter	COG1178	[P]
feature:167819098	hypothetical protein	NA	NA
feature:167819108	spe:Spro_4738 outer membrane autotransporter barr el domain	NA	NA
feature:167819110	hypothetical protein	NA	NA
feature:167819112	hypothetical protein	NA	NA
feature:167819114	hypothetical protein	NA	NA
feature:167819116	hypothetical protein	NA	NA
feature:167819120	hypothetical protein PU1002_06886	NA	NA
feature:167819134	acetyltransferase	COG0454	[KR]
feature:167819146	unknown membrane protein	NA	NA
feature:167819154	sarcosine oxidase alpha chain	COG0446	[R]
feature:167819168	Phytanoyl-CoA dioxygenase (PhyH) superfamily	NA	NA
feature:167819170	hypothetical protein	NA	NA
feature:167819176	glycine betaine/l-proline transport ATP-binding protein	COG4175	[E]
feature:167819186	putative aminomethyltransferase protein	COG3665	[S]
feature:167819188	putative aminomethyltransferase protein	COG3665	[S]
feature:167819192	ribosomal protein L6e	NA	NA
feature:167819216	ammonium transporter	COG0004	[P]
feature:167819340	unknown membrane protein	NA	NA
feature:167819346	hypothetical protein	NA	NA
feature:167819362	putative N-acetylglucosaminyltransferase	NA	NA
feature:167819384	possible Rhodanese-related sulfurtransferase	COG0607	[P]
feature:167819394	hypothetical protein	NA	NA
feature:167819402	hypothetical protein	NA	NA
feature:167819404	hypothetical protein ALOHA_HF4000007D16ctg1g30	COG1670	[J]
feature:167819408	6-aminohexanoate-dimer hydrolase	NA	NA
feature:167819410	hypothetical protein OM2255_02972	COG1335	[Q]
feature:167819412	conserved hypothetical protein	COG1670	[J]
feature:167819414	hypothetical protein ALOHA_HF4000007D16ctg1g30	NA	NA
feature:167819416	putative thiol-specific antioxidant related protein (thiol peroxidase Bcp)	COG1225	[O]
feature:167819420	putative site-specific recombinase phage integrase family protein	NA	NA
feature:167819422	hypothetical protein	NA	NA
feature:167866872	hypothetical protein	NA	NA
feature:167866878	hypothetical protein	NA	NA
feature:167866910	tolR protein	COG0848	[U]
feature:167866958	glycyl-tRNA synthetase subunit alpha	COG0752	[J]
feature:167866978	ferredoxin	COG0633	[C]
feature:167866988	Glycosyl transferase	COG0463	[M]
feature:167866994	hypothetical protein SAR11_0461	NA	NA
feature:167867010	isocitrate dehydrogenase	COG0538	[C]
feature:167867018	FtsJ-like methyltransferase	COG0293	[J]
feature:167867028	X-Pro dipeptidase	COG0006	[E]
feature:167867030	X-Pro dipeptidase	COG0006	[E]
feature:167867032	putative putrescine ABC transporter ATP-binding protein	COG3842	[E]
feature:167867042	hypothetical protein SAR11_1332	NA	NA
feature:167888766	hypothetical protein Nham_2527	COG0358	[L]
feature:167888768	primase/helicase	NA	NA

feature:167888770	hypothetical protein	NA	NA
feature:167888772	hypothetical protein	NA	NA
feature:167888774	hypothetical protein	NA	NA
feature:167888798	NADH dehydrogenase subunit G	NA	NA
feature:167888814	biotin-protein ligase	COG0340	[H]
feature:167888824	hypothetical protein	NA	NA
feature:167888830	ABC transporter ATP-binding protein	COG1136	[V]
feature:167888832	DNA polymerase III alpha subunit	COG0587	[L]
feature:167888836	protein translation elongation factor Ts (EF-Ts)	COG0264	[J]
feature:167888838	protein translation elongation factor Ts (EF-Ts)	COG0264	[J]
feature:167888840	Ribosome recycling factor	COG0233	[J]
feature:167888848	outer membrane protein	COG4775	[M]
feature:167888862	lipid-A-disaccharide synthase (lpxB)	COG0763	[M]
feature:167888876	hypothetical SurA-like protein	NA	NA
feature:167888884	hypothetical protein	NA	NA
feature:167888892	septum formation initiator	NA	NA
feature:167907226	hypothetical protein	NA	NA
feature:167907280	YadG-like ABC transporter protein	COG1131	[V]
feature:167907306	prephenate dehydrogenase	COG0287	[E]
feature:167907344	H ⁺ -transporting two-sector ATPase (F0F1-type ATP synthase) delta chain	COG0712	[C]
feature:167907374	membrane portion of succinate dehydrogenase	NA	NA
feature:167907392	3-isopropylmalate dehydratase	COG0066	[E]
feature:167907410	hypothetical protein SAR11_0259	NA	NA
feature:167907428	hypothetical protein PU1002_05026	NA	NA
feature:167907430	membrane protein putative	NA	NA
feature:167907432	hypothetical protein PU1002_05016	NA	NA
feature:167907434	cytochrome c class IC	COG2863	[C]
feature:167907436	hypothetical protein PU1002_05006	NA	NA
feature:167907438	L-sorbose dehydrogenase putative	COG2133	[G]
feature:167907444	(S)-2-hydroxy-acid oxidase	COG0277	[C]
feature:167907446	(S)-2-hydroxy-acid oxidase	COG0277	[C]
feature:167907484	polyketide biosynthesis dithiol-disulfide isomerase	COG2761	[Q]
feature:167907492	marC family protein	COG2095	[U]
feature:167907496	putative transmembrane protein	COG1238	[S]
feature:167907498	possible transmembrane protein	NA	NA
feature:167907510	hypothetical protein SAR11_1722	COG0762	[S]
feature:167907526	hypothetical protein SAR11_1723	NA	NA
feature:167907550	hypothetical protein	NA	NA
feature:167907552	hypothetical protein	NA	NA
feature:167907576	Smr (small MutS related) protein	COG2840	[S]
feature:167907586	Dephospho-CoA kinase	NA	NA
feature:167907596	ferrochelatase	COG0276	[H]
feature:167907604	Glucose inhibited division protein	COG0357	[M]
feature:167907606	ParA family ATPase for plasmid partitioning and other plasmid related functions	COG1192	[D]
feature:167907612	hypothetical protein SAR11_0356	NA	NA
Genes specific to ACE_P3.2			
Gene ID	Annotation	COG	COG Category
feature:167687112	putative short-chain dehydrogenase/reductase	COG1028	[IQR]
feature:167687114	hypothetical protein PROVRETT_00056	COG2079	[R]
feature:167687118	hypothetical protein	NA	NA

feature:167687122	hypothetical protein	NA	NA
feature:167687128	hypothetical protein GP2143_10052	NA	NA
feature:167687130	Bacterial extracellular solute-binding protein, family 5	COG0747	[E]
feature:167687132	extracellular solute-binding protein, family 5	COG0747	[E]
feature:167687134	binding-protein-dependent transport systems inner membrane component	COG0601	[EP]
feature:167687136	ABC transporter, permease protein, putative	COG1173	[EP]
feature:167687140	oligopeptide ATP-binding ABC transporter protein	COG4608	[E]
feature:167687142	oligopeptide/dipeptide ABC transporter, ATP-binding protein, C-terminal domain protein	COG4608	[E]
feature:167687144	Succinylglutamate desuccinylase/aspartoacylase	NA	NA
feature:167687146	putative hydrolase	COG3741	[E]
feature:167687148	amidohydrolase	COG1473	[R]
feature:167687150	peptidase M20D, amidohydrolase	COG1473	[R]
feature:167687152	hypothetical protein SAR11_0694	COG2081	[R]
feature:167687154	hypothetical protein	NA	NA
feature:167687162	penicillin-binding protein 1A	COG5009	[M]
feature:167687164	Penicillin-binding protein 1A	COG5009	[M]
feature:167687200	hypothetical protein	NA	NA
feature:167687204	carboxymuconolactone decarboxylase family protein	COG0599	[S]
feature:167687208	dihydroorotase	COG0044	[F]
feature:167687212	sulfate permease family protein	COG0659	[P]
feature:167687218	radical SAM domain-containing protein	COG0535	[R]
feature:167687224	hypothetical protein	NA	NA
feature:167687226	hypothetical protein	NA	NA
feature:167687232	hypothetical protein	NA	NA
feature:167687236	alternative oxidase 2	NA	NA
feature:167687238	hypothetical protein	NA	NA
feature:167687240	hypothetical protein PB7211_523	NA	NA
feature:167687244	hypothetical protein	NA	NA
feature:167687246	hypothetical protein	NA	NA
feature:167687248	hypothetical protein Pnap_3934	NA	NA
feature:167687250	hypothetical protein P9515_00371	COG4889	[R]
feature:167687252	hypothetical protein	NA	NA
feature:167687254	hypothetical protein	NA	NA
feature:167687256	hypothetical protein	NA	NA
feature:167687258	hypothetical protein	NA	NA
feature:167687260	hypothetical protein PU1002_00600	NA	NA
feature:167687262	hypothetical protein	NA	NA
feature:167687294	hypothetical protein	NA	NA
feature:167687298	immunogenic protein	COG2358	[R]
feature:167687306	hypothetical protein	NA	NA
feature:167687308	hypothetical protein	NA	NA
feature:167687322	NADH dehydrogenase I chain E	COG1905	[C]
feature:167687336	NADH dehydrogenase i chain I	COG1009	[CP]
feature:167687338	NADH Dehydrogenase I Chain L	COG1009	[CP]
feature:167687340	hypothetical protein	NA	NA
feature:167687342	NADH dehydrogenase subunit M	COG1008	[C]
feature:167687344	NADH dehydrogenase subunit N	COG1007	[C]
feature:167687380	1-deoxy-D-xylulose 5-phosphate reductoisomerase	COG0743	[I]
feature:167687384	COG1044: UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	NA	NA
feature:167687392	lipid-A-disaccharide synthase (lpxB)	COG0763	[M]
feature:167730502	uroporphyrinogen III synthase HEM4	NA	NA

feature:167730510	hypothetical protein PMM1242	NA	NA
feature:167730512	methylase involved in ubiquinone/menaquinone biosynthesis	NA	NA
feature:167730514	hypothetical protein PMM1249	NA	NA
feature:167730520	hypothetical protein	NA	NA
feature:167730544	hypothetical protein Mmar10_2939	COG2835	[S]
feature:167730594	hypothetical protein	NA	NA
feature:167730596	hypothetical protein	NA	NA
feature:167730602	COX RNA-associated protein	COG0486	[R]
feature:167730618	leucyl-tRNA synthetase	COG0495	[J]
feature:167730646	hypothetical protein	NA	NA
feature:167730648	hypothetical protein	NA	NA
feature:167730674	hypothetical protein	NA	NA
feature:167730678	hypothetical protein	NA	NA
feature:167730684	hypothetical protein	NA	NA
feature:167730714	hypothetical protein	NA	NA
feature:167730716	hypothetical protein	NA	NA
feature:167730718	hypothetical protein	NA	NA
feature:167730724	molybdopterin biosynthesis protein	COG0476	[H]
feature:167730758	hypothetical protein	COG1131	[V]
feature:167730764	YadG-like ABC transporter protein	COG0258	[L]
feature:167730766	DNA polymerase I	COG0749	[L]
feature:167730768	DNA polymerase I	NA	NA
feature:167730770	DNA polymerase I	COG0749	[L]
feature:167730788	hypothetical protein	NA	NA
feature:167730800	hypothetical protein	NA	NA
feature:167730810	major facilitator superfamily protein	COG0477	[GEPR]
feature:167730812	putative sugar (and other) transporter	COG0477	[GEPR]
feature:167730836	fumarylacetoacetate (FAA) hydrolase family protein	NA	NA
feature:167730840	putative membrane protein	NA	NA
feature:167730864	P0A3K3 YDNAA_RICMO Uncharacterized protein in dnaA 5' region	COG0607	[P]
feature:167730866	hypothetical protein	NA	NA
feature:167730886	conserved hypothetical protein	NA	NA
feature:167730890	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	COG2609	[C]
feature:167730892	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	COG2609	[C]
feature:167730912	GntR family transcriptional regulator	COG1802	[K]
feature:167730922	membrane protein, putative	COG2358	[R]
feature:167730924	O-methyl transferase	NA	NA
feature:167730926	hypothetical protein	NA	NA
feature:167730934	phosphate ABC transporter	COG0226	[P]
feature:167730950	extracellular solute-binding protein, family 1	COG1840	[P]
feature:167730952	Binding-protein-dependent transport systems inner membrane component	COG1178	[P]
feature:167730966	protein of unknown function DUF47	COG1392	[P]
feature:167730984	hypothetical protein ALOHA_HF4000133G03ctg1g2	NA	NA
feature:167730986	hypothetical protein ALOHA_HF4000133G03ctg1g2	COG0665	[E]
feature:167730992	acetylornithine aminotransferase	COG4992	[E]
feature:167731028	hypothetical protein	NA	NA
feature:167731030	hypothetical protein	NA	NA
feature:167731064	hypothetical protein	NA	NA
feature:167731066	hypothetical protein	NA	NA
feature:167731068	putative beta-ketoacyl-acyl-carrier-protein reductase (fabG-like)	COG1028	[IQR]

feature:167731070	Beta-ketoacyl-acyl-carrier-protein synthase I	COG0332	[I]
feature:167731072	hypothetical protein	NA	NA
feature:167731074	hypothetical protein	NA	NA
feature:167731076	aminopeptidase domain protein	COG4310	[R]
feature:167731078	hypothetical protein RS9917_04675	COG0223	[J]
feature:167731080	hypothetical protein	NA	NA
feature:167731082	acylneuraminate cytidyltransferase	COG1861	[M]
feature:167731084	polysaccharide biosynthesis protein CapD	COG1086	[MG]
feature:167731086	N-acetylneuraminate synthase	NA	NA
feature:167731088	hypothetical protein	NA	NA
feature:167731090	hypothetical protein	NA	NA
feature:167731092	N-acetylneuraminate synthase	COG2089	[M]
feature:167731094	hypothetical dTDP-4-dehydrorhamnose reductase	COG1091	[M]
feature:167731096	hypothetical protein	NA	NA
feature:167731098	hypothetical protein Ddes_0146	COG4310	[R]
feature:167731100	hypothetical protein	NA	NA
feature:167731102	dTDP-glucose 4,6-dehydratase	COG0451	[MG]
feature:167731104	hypothetical protein	NA	NA
feature:167731106	hypothetical protein	NA	NA
feature:167731108	hypothetical protein	NA	NA
feature:167731110	glycosyltransferase group 2 family protein	COG0463	[M]
feature:167731112	hypothetical protein	NA	NA
feature:167731114	glycosyl transferase family protein	COG0463	[M]
feature:167731116	hypothetical protein	NA	NA
feature:167731118	hypothetical protein	NA	NA
feature:167731122	hypothetical protein SAR11_0567	COG0472	[M]
feature:167731124	hypothetical protein	NA	NA
feature:167731126	TDP-4-oxo-6-deoxy-D-glucose transaminase	COG0399	[M]
feature:167731128	hypothetical protein P9515_12791	NA	NA
feature:167731132	hypothetical protein PBCV1_A295L	COG0451	[MG]
feature:167731136	Glycosyl transferase, family 2	COG0463	[M]
feature:167731138	hypothetical protein	NA	NA
feature:167731140	glycosyl transferase group 1	NA	NA
feature:167731142	dTDP-glucose 4,6-dehydratase	COG1088	[M]
feature:167731144	hypothetical protein	NA	NA
feature:167731148	Carbamoyltransferase	COG2192	[O]
feature:167731150	hypothetical protein	NA	NA
feature:167731152	hypothetical protein	NA	NA
feature:167731154	hypothetical protein	NA	NA
feature:167731156	lipoprotein	NA	NA
feature:167731158	hypothetical protein VspiD_05775	NA	NA
feature:167731160	NAD-dependent epimerase/dehydratase	COG0451	[MG]
feature:167731162	hypothetical protein	NA	NA
feature:167731164	glycosyl transferase family protein	COG0463	[M]
feature:167731196	hypothetical protein	NA	NA
feature:167731208	probable glycosyltransferase protein	COG0463	[M]
feature:167731210	glycosyl transferase family protein	COG0463	[M]
feature:167731212	hypothetical protein	NA	NA
feature:167731214	methyltransferase type 12	NA	NA
feature:167731216	hypothetical protein	NA	NA
feature:167731218	NAD-dependent epimerase/dehydratase	NA	NA
feature:167731222	syn:SynRCC307_2084 mrp; ATPases involved in chromosome partitioning; K03593 ATP-binding protein involved in c chromosome partitioning	COG0489	[D]

feature:167731248	rod shape-determining protein MreC	NA	NA
feature:167731258	hypothetical protein	NA	NA
feature:167774671	glutamate dehydrogenase	COG0334	[E]
feature:167774673	anti-ECFsigma factor, ChrR	COG3806	[T]
feature:167774675	OsmC family protein	COG1073	[R]
feature:167774677	hypothetical protein	NA	NA
feature:167774679	hypothetical protein	NA	NA
feature:167774681	hypothetical protein	NA	NA
feature:167774683	COG3152: Uncharacterized membrane protein	COG3152	[S]
feature:167774685	hypothetical protein	NA	NA
feature:167774687	intracellular protease/amidase, putative	COG0693	[R]
feature:167774689	hypothetical protein	NA	NA
feature:167774693	thiamine biosynthesis protein ThiS	COG2104	[H]
feature:167774727	hypothetical protein	NA	NA
feature:167774733	hypothetical protein	NA	NA
feature:167774755	hypothetical protein	NA	NA
feature:167774759	amidophosphoribosyltransferase	COG0034	[F]
feature:167774773	hypothetical protein	NA	NA
feature:167774787	ADP-heptose:LPS heptosyltransferase, putative	NA	NA
feature:167774789	Q8RFJ8 SYA_FUSNN Alanyl-tRNA synthetase	COG0013	[J]
feature:167774791	hypothetical protein	NA	NA
feature:167774793	alanyl-tRNA synthetase	COG0013	[J]
feature:167774827	hypothetical protein	COG0365	[I]
feature:167816100	hypothetical protein	NA	NA
feature:167816108	protein-L-isoaspartate(D-aspartate) O-methyltransferase	COG2518	[O]
feature:167816110	hypothetical protein	NA	NA
feature:167816112	valine-tRNA ligase	COG0525	[J]
feature:167816114	valyl-tRNA synthetase	COG0525	[J]
feature:167816116	valine-tRNA ligase	COG0525	[J]
feature:167816130	COG0018: Arginyl-tRNA synthetase	COG0018	[J]
feature:167816134	hypothetical protein	NA	NA
feature:167816146	COG0172: Seryl-tRNA synthetase	COG0172	[J]
feature:167816154	putative SinH-like protein	NA	NA
feature:167816156	hypothetical protein	NA	NA
feature:167816158	hypothetical protein	NA	NA
feature:167816160	membrane-bound lytic transglycolase-related protein	COG2951	[M]
feature:167816174	hypothetical protein	NA	NA
feature:167816196	putative transmembrane transport protein	COG1133	[I]
feature:167816202	hypothetical protein	NA	NA
feature:167816206	protein-export membrane protein SecD	COG0342	[U]
feature:167824446	hypothetical protein	NA	NA
feature:167824462	hypothetical protein	NA	NA
feature:167824466	hypothetical protein	NA	NA
feature:167824468	hypothetical protein ALOHA_HF4000005I08ctg1g18	NA	NA
feature:167824470	hypothetical protein	NA	NA
feature:167824472	ATP-dependent RNA helicase	COG0513	[LKJ]
feature:167824476	hypothetical protein	NA	NA
feature:167824484	hypothetical protein	NA	NA
feature:167824492	conserved hypothetical protein	COG3564	[S]
feature:167824504	hypothetical protein PU1002_02296	NA	NA
feature:167824514	hypothetical protein	NA	NA
feature:167824526	nitrogen regulatory protein P-II	COG0347	[E]
feature:167824540	hypothetical protein	NA	NA

feature:167824548	COG3114: Heme exporter protein D	COG3114	[U]
feature:167824556	hypothetical protein Meso_3181	COG0730	[R]
feature:167824572	hypothetical protein OG2516_00624	COG2079	[R]
feature:167824590	hypothetical protein	NA	NA
feature:167824598	hypothetical protein Arad_2918	COG2120	[S]
feature:167824600	hypothetical protein MED297_02905	COG0684	[H]
feature:167824610	hypothetical protein BAL199_08298	COG2175	[Q]
feature:167824618	Fumarate reductase/succinate dehydrogenase flavoprotein-like protein	COG1053	[C]
feature:167824620	hypothetical protein	NA	NA
feature:167824628	MmgE/PrpD family protein	NA	NA
feature:167824630	bur:Bcep18194_B1382 hypothetical protein	NA	NA
feature:167824632	hypothetical protein	NA	NA
feature:167824662	MarR family transcriptional regulator	NA	NA
feature:167824668	alkyl hydroperoxide reductase/ Thiol specific anti oxidant/ Mal allergen	COG0450	[O]
feature:167824674	hypothetical protein	NA	NA
feature:167824692	hypothetical protein	NA	NA
feature:167824696	probable aminopeptidase	COG0006	[E]
feature:167824718	hypothetical protein	NA	NA
feature:167824722	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	COG0771	[M]
feature:167824724	hypothetical protein	NA	NA
feature:167824740	hypothetical protein	NA	NA
feature:167824760	hypothetical protein	NA	NA
feature:167824764	hypothetical protein	NA	NA
feature:167824766	hypothetical protein	NA	NA
feature:167824770	carbamoyl-phosphate synthase large chain	COG0458	[EF]
feature:167824772	carbamoyl-phosphate synthase large subunit	COG0458	[EF]
feature:167824774	carbamoyl-phosphate synthase, large subunit	COG0458	[EF]
feature:167824786	excinuclease ABC, B subunit	COG0556	[L]
feature:167824788	excinuclease ABC, B subunit	COG0556	[L]
feature:167824790	short chain dehydrogenase	COG1028	[IQR]
feature:167824792	3-oxoacyl-(acyl-carrier protein) reductase	COG1028	[IQR]
feature:167824798	hypothetical protein	NA	NA
feature:167824836	hypothetical protein	NA	NA
feature:167824842	COG0617: tRNA nucleotidyltransferase/poly(A) polymerase	COG0617	[J]
feature:167824874	hypothetical protein	NA	NA
feature:167824876	COG0711: F0F1-type ATP synthase b subunit	NA	NA
feature:167824878	hypothetical protein	NA	NA
feature:167824894	threonine synthase	COG0498	[E]
feature:167824896	surfeit locus protein 1	COG3346	[S]
feature:167824902	hypothetical protein	NA	NA
feature:167824908	putative modulator of DNA gyrase	COG0312	[R]
feature:167824912	putative modulator of DNA gyrase	COG0312	[R]
feature:167824952	RNA pseudouridine synthase family protein	COG1187	[J]
feature:167824962	isoleucyl-tRNA synthetase	COG0060	[J]
feature:167824972	hypothetical protein	NA	NA
feature:167824974	methyltransferase FkbM family	COG0500	[QR]
feature:167824976	hypothetical protein	NA	NA
feature:167824978	hypothetical protein	NA	NA
feature:167824980	hypothetical protein	NA	NA
feature:167824982	hypothetical protein	NA	NA

feature:167824984	hypothetical protein	NA	NA
feature:167824986	hypothetical protein	NA	NA
feature:167824988	hypothetical protein	NA	NA
feature:167824990	hypothetical protein	NA	NA
feature:167824992	hypothetical protein	NA	NA
feature:167824994	hypothetical protein	NA	NA
feature:167824996	hypothetical protein	NA	NA
feature:167824998	hypothetical protein	NA	NA
feature:167825000	truncated transposon gamma-delta resolvase	NA	NA
feature:167825002	hypothetical protein	NA	NA
feature:167825004	hypothetical protein	NA	NA
feature:167825006	hypothetical protein	NA	NA
feature:167825008	hypothetical protein	NA	NA
feature:167825010	hypothetical protein	NA	NA
feature:167825022	nucleotide-binding protein	COG0424	[D]
feature:167825028	hypothetical protein	NA	NA
feature:167825048	dihydrolipoamide dehydrogenase	COG0508	[C]
feature:167825050	putative biotin-requiring enzyme	COG2609	[C]
feature:167825062	hypothetical protein	NA	NA
feature:167825064	GTP-binding protein LepA	COG0481	[M]
feature:167825066	GTP-binding protein LepA	COG0481	[M]
feature:167825068	GTP-binding protein LepA	COG0481	[M]
feature:167825080	COG0290: Translation initiation factor IF3	NA	NA
feature:167825084	threonyl-tRNA synthetase	COG0441	[J]
feature:167825086	hypothetical protein	NA	NA
feature:167825090	conserved hypothetical protein	COG0451	[MG]
feature:167825112	hypothetical protein	NA	NA
feature:167825114	3-phosphoshikimate 1-carboxyvinyltransferase	COG0128	[E]
feature:167825128	hypothetical protein	NA	NA
feature:167825134	acetyl-CoA carboxylase subunit beta	COG0777	[I]
feature:167825144	RmuC domain protein	COG1322	[S]
feature:167825150	hypothetical protein	NA	NA
feature:167825152	hypothetical protein	NA	NA
feature:167825160	hypothetical protein	NA	NA
feature:167825220	muropeptide permease AmpG	COG0477	[GEPR]
feature:167825222	major facilitator transporter	COG0477	[GEPR]
feature:167825226	hypothetical protein	NA	NA
feature:167825228	hypothetical protein	NA	NA
feature:167825232	hypothetical protein	COG0500	[QR]
feature:167825234	hypothetical protein OA307_2585	COG4310	[R]
feature:167825236	acylphosphatase	COG1254	[C]
feature:167825244	hypothetical protein PU1002_04226	NA	NA
feature:167825246	hypothetical protein	NA	NA
feature:167825248	hypothetical protein	NA	NA
feature:167825250	hypothetical protein RAZWK3B_09726	NA	NA
feature:167825252	hypothetical protein	NA	NA
feature:167825254	COG0582: Integrase	COG0582	[L]
feature:167825256	hypothetical protein	NA	NA
feature:167825258	putative site-specific recombinase, phage integrase family protein	NA	NA
feature:167825260	hypothetical protein	NA	NA
feature:167825262	DedA family protein	COG1238	[S]
feature:167825264	long-chain-fatty-acid--CoA ligase, putative	COG1022	[I]
feature:167825266	hypothetical protein	NA	NA

feature:167825268	hypothetical protein	NA	NA
feature:167825270	hypothetical protein	NA	NA
feature:167825272	hypothetical protein	NA	NA
feature:167825274	hypothetical protein	NA	NA
feature:167825276	hypothetical protein PU1002_04226	COG1329	[K]
feature:167825278	long-chain-fatty-acid--CoA ligase, putative	COG1022	[I]
feature:167825280	DedA family protein	COG1238	[S]
feature:167825298	diaminopimelate decarboxylase	COG0019	[E]
feature:167825322	hypothetical protein	NA	NA
feature:167825350	ATP synthase F1, delta subunit	COG0712	[C]
feature:167825356	fructose-6-phosphate aldolase 2	COG0176	[G]
feature:167825360	dihydrolipoamide dehydrogenase	COG1249	[C]
feature:167825384	aspartate-semialdehyde dehydrogenase	COG0136	[E]
feature:167825386	aspartate-semialdehyde dehydrogenase	COG0136	[E]
feature:167825392	3-isopropylmalate dehydratase, large subunit	COG0065	[E]
feature:167825410	hypothetical protein	NA	NA
feature:167825412	possible transmembrane protein	NA	NA
feature:167825422	hypothetical protein	NA	NA
feature:167825434	(S)-2-hydroxy-acid oxidase	COG0277	[C]
feature:167905352	conserved hypothetical protein	COG1809	[S]
feature:167905356	COG2855: Uncharacterized membrane protein	NA	NA
feature:167905360	K+-dependent Na+/Ca+ exchanger related-protein	COG0530	[P]
feature:167905362	CaCA family Na(+)/Ca(+) antiporter	COG0530	[P]
feature:167905370	efflux ABC transporter, permease protein	NA	NA
feature:167905372	tricarballoylate dehydrogenase	COG1053	[C]
feature:167905374	putative succinate dehydrogenase/fumarate reductas e flavoprotein subunit	COG1053	[C]
feature:167905382	conserved hypothetical protein	NA	NA
feature:167905392	5'-methylthioadenosinenucleosidase/S- adenosylhomoc ysteinenucleosidase	COG0775	[F]
feature:167905398	Cytochrome c	NA	NA
feature:167905404	conserved hypothetical protein	NA	NA
feature:167905406	NMT1/THI5 like domain protein	NA	NA
feature:167905408	ABC transporter related	COG1116	[P]
feature:167905410	ABC transporter ATP-binding protein (taurine)	COG1116	[P]
feature:167905412	ABC transporter permease protein	COG0600	[P]
feature:167905416	Substrate-binding region of ABC-type glycine betai ne transport system	COG2113	[E]
feature:167905418	hypothetical protein SAR11_0965	NA	NA
feature:167905420	ferric-uptake regulator	NA	NA
feature:167905426	iron(III) ABC transporter	COG3842	[E]
feature:167905428	iron(III) ABC transporter	COG1178	[P]
feature:167905430	hypothetical protein	NA	NA
feature:167905432	hypothetical protein	NA	NA
feature:167905434	iron-uptake ABC transport system periplasmic iron-binding protein	COG1840	[P]
feature:167905436	deoxyribodipyrimidine photolyase-related protein	COG3046	[R]
feature:167905438	deoxyribodipyrimidine photolyase-related protein	COG3046	[R]
feature:167905448	putative cytochrome c	COG2010	[C]
feature:167905472	hypothetical protein	NA	NA
feature:167905482	hypothetical protein	NA	NA
feature:167905490	putative D-ala D-ala ligase	COG1181	[M]
feature:167905504	phosphoglucosmutase/phosphomannosmutase family prote in MrsA	COG1109	[G]

feature:167932898	hypothetical protein	NA	NA
feature:167932910	GTP cyclohydrolase II /3,4-dihydroxy-2-butanone 4-phosphate synthase	COG0108	[H]
feature:167932922	delta-aminolevulinic acid dehydratase	COG0113	[H]
feature:167932966	aspartyl/glutamyl-tRNA amidotransferase subunit B	COG0064	[J]
feature:167932968	glutamyl-tRNA(gln) amidotransferase chain B	COG0064	[J]
feature:167932980	dihydroorotase	COG0044	[F]
feature:167933000	lipoate-protein ligase B	COG0321	[H]
feature:167933016	hypothetical protein	NA	NA
feature:167933018	hypothetical protein	NA	NA
feature:167933074	hypothetical protein	NA	NA
feature:167933082	hypothetical protein	NA	NA
feature:167933094	hypothetical protein Bpro_2642	NA	NA
feature:167933102	trigger factor	COG0544	[O]
feature:167933104	trigger factor	COG0544	[O]
feature:167933108	hypothetical protein	NA	NA
feature:167933116	Quaternary ammonium compound-resistance protein qa cH	COG2076	[P]
feature:167933118	hypothetical protein	NA	NA
feature:167933122	hypothetical protein	NA	NA
feature:167933124	ABC transporter, nucleotide binding/ATPase protein (zinc)	COG1121	[P]
feature:167933132	hypothetical oxidoreductase yiaK	NA	NA
feature:167933134	hypothetical protein	NA	NA
feature:167933136	hypothetical protein	NA	NA
feature:167933138	hypothetical protein	NA	NA
feature:167933140	hypothetical protein	NA	NA
feature:167933142	hypothetical protein	NA	NA
feature:167933144	hypothetical protein PB7211_875	NA	NA
feature:167933146	hypothetical protein	NA	NA
feature:167933148	hypothetical protein Noc_2239	NA	NA
feature:167933150	hypothetical protein	NA	NA
feature:167933152	hypothetical protein	NA	NA
feature:167933154	hypothetical protein	NA	NA
feature:167933156	hypothetical protein	NA	NA
feature:167933158	hypothetical protein	NA	NA
feature:167933160	hypothetical protein S-PM2p152	NA	NA
feature:167933162	hypothetical protein Nham_2527	NA	NA
feature:167933164	hypothetical protein	NA	NA
feature:167933166	hypothetical protein	NA	NA
feature:167933168	hypothetical protein	NA	NA
feature:167933170	hypothetical protein	NA	NA
feature:167933172	hypothetical protein	NA	NA
feature:167933174	hypothetical protein	NA	NA
feature:167933176	hypothetical protein	NA	NA
feature:167933178	hypothetical protein	NA	NA
feature:167933180	hypothetical protein	NA	NA
feature:167933182	Glycosyl transferase	COG0463	[M]
feature:167933184	glycosyl transferase family protein	COG0463	[M]
feature:167933212	RNA methyltransferase	COG0565	[J]
feature:167933214	tRNA/rRNA methyltransferase	COG0565	[J]
feature:167933218	hypothetical protein	NA	NA
feature:167933226	hypothetical protein PU1002_05956	NA	NA
feature:167933228	hypothetical protein PU1002_05956	NA	NA

feature:167933230	protein of unknown function DUF218	COG1434	[S]
feature:167933232	hypothetical protein	NA	NA
feature:167933236	hypothetical protein	COG0449	[M]
feature:167933254	hypothetical protein HPDFL43_18047	NA	NA
feature:167933256	COG0412: Dienelactone hydrolase and related enzyme s	NA	NA
feature:167933264	phosphomannomutase	COG1109	[G]
feature:167933266	hypothetical protein SAR11_0794	NA	NA
feature:167933268	hypothetical protein SAR11_0793	NA	NA
feature:167933270	hypothetical protein PU1002_00295	NA	NA
feature:167933272	conserved hypothetical protein	COG2961	[R]
feature:167933294	D-lactate dehydrogenase (cytochrome)	COG0277	[C]
feature:167933296	hypothetical protein	NA	NA
feature:167933302	DNA photolyase-like protein	COG0415	[L]
feature:167933304	hypothetical protein	NA	NA
feature:167933306	conserved hypothetical protein	NA	NA
feature:167933308	COG2849: Uncharacterized BCR	COG2849	[S]
feature:167933310	hypothetical protein SAR11_0636	NA	NA
feature:167933312	hypothetical protein	NA	NA
feature:167933314	hypothetical protein PU1002_00600	NA	NA
feature:167933316	hypothetical protein	NA	NA
feature:167933318	hypothetical protein	NA	NA
feature:167933320	hypothetical protein	NA	NA
feature:167933322	hypothetical protein BACINT_04027	COG0317	[TK]
feature:167933324	transporter	COG0477	[GEPR]
feature:167933326	hypothetical protein	NA	NA
feature:167933328	hypothetical protein	NA	NA
feature:167933330	hypothetical protein	NA	NA
feature:167933332	hypothetical protein	NA	NA
feature:167935902	Glycosyl transferase group 1	NA	NA
feature:167935904	Glycosyl transferase group 1	COG0438	[M]
feature:167935912	extracellular solute-binding protein	COG0725	[P]
feature:167935924	pantetheine-phosphate adenyltransferase	COG0669	[H]

Genes specific to IMCC9063

Gene ID	Annotation	COG	COG Category
SAR11G3_00011	hypothetical protein	NA	NA
SAR11G3_00012	hypothetical protein	NA	NA
SAR11G3_00030	putative flippase	NA	NA
SAR11G3_00031	glycosyl transferase, family 2	NA	NA
SAR11G3_00032	hypothetical protein	NA	NA
SAR11G3_00033	hypothetical protein	NA	NA
SAR11G3_00034	putative glycosyltransferase	COG0463	[M]
SAR11G3_00035	hypothetical protein	NA	NA
SAR11G3_00036	hypothetical protein	NA	NA
SAR11G3_00037	hypothetical protein	NA	NA
SAR11G3_00038	hypothetical protein	NA	NA
SAR11G3_00041	hypothetical protein	NA	NA
SAR11G3_00042	hypothetical protein	NA	NA
SAR11G3_00043	glycosyltransferase involved in cell wall biogenesis	COG0463	[M]
SAR11G3_00044	nucleotidyl transferase superfamily	NA	NA
SAR11G3_00045	capsular polysaccharide biosynthesis protein CapD	COG1086	[MG]
SAR11G3_00046	hypothetical protein	NA	NA
SAR11G3_00047	hypothetical protein	NA	NA

SAR11G3_00048	hypothetical protein	NA	NA
SAR11G3_00049	wbpG	COG0037	[D]
SAR11G3_00050	putative LPS biosynthesis protein WbpG	COG0037	[D]
SAR11G3_00051	hypothetical protein	NA	NA
SAR11G3_00054	hypothetical protein	COG0118	[E]
SAR11G3_00055	hypothetical protein	COG1207	[M]
SAR11G3_00056	WblQ protein	COG0399	[M]
SAR11G3_00057	oxidoreductase, Gfo/Idh/MocA family/transferase hexapeptide repeat protein	NA	NA
SAR11G3_00058	hypothetical protein	NA	NA
SAR11G3_00059	glycosyltransferase ycbB	COG0463	[M]
SAR11G3_00060	hypothetical protein	NA	NA
SAR11G3_00061	hypothetical protein	NA	NA
SAR11G3_00064	hypothetical protein	COG0438	[M]
SAR11G3_00065	hypothetical protein	NA	NA
SAR11G3_00066	hypothetical protein	NA	NA
SAR11G3_00067	hypothetical protein	NA	NA
SAR11G3_00068	hypothetical protein	NA	NA
SAR11G3_00069	lipoprotein	NA	NA
SAR11G3_00071	cps9E, putative	COG1086	[MG]
SAR11G3_00072	nucleotide-diphosphate-sugar epimerase, membrane associated	COG1086	[MG]
SAR11G3_00073	glycosyltransferase	COG0438	[M]
SAR11G3_00074	hypothetical protein	COG1670	[J]
SAR11G3_00075	hypothetical protein	NA	NA
SAR11G3_00076	hypothetical protein	NA	NA
SAR11G3_00077	ubiquinone/menaquinone biosynthesis methyltransferase UBIE	NA	NA
SAR11G3_00078	protein containing aminopeptidase domain	COG4310	[R]
SAR11G3_00079	UDP-glucose 4-epimerase	COG0451	[MG]
SAR11G3_00080	hypothetical protein	NA	NA
SAR11G3_00083	hypothetical protein	NA	NA
SAR11G3_00084	3-deoxy-manno-octulosonate cytidyltransferase	COG1212	[M]
SAR11G3_00085	4-hydroxy-2-oxovalerate aldolase	COG0119	[E]
SAR11G3_00086	4-hydroxy-2-oxovalerate aldolase	NA	NA
SAR11G3_00087	hypothetical protein	NA	NA
SAR11G3_00088	hypothetical protein	NA	NA
SAR11G3_00089	spore coat polysaccharide biosynthesis protein spsC	NA	NA
SAR11G3_00090	hypothetical protein	NA	NA
SAR11G3_00091	spore coat polysaccharide biosynthesis protein spsC	COG0399	[M]
SAR11G3_00092	putative dehydrogenase and related protein	NA	NA
SAR11G3_00093	hypothetical protein	NA	NA
SAR11G3_00094	hypothetical protein	NA	NA
SAR11G3_00095	hypothetical protein	NA	NA
SAR11G3_00097	LPS biosynthesis protein	COG0037	[D]
SAR11G3_00098	hypothetical protein	NA	NA
SAR11G3_00099	glycosyltransferase	COG0463	[M]
SAR11G3_00100	N-acetylneuraminate synthase	COG2089	[M]
SAR11G3_00101	N-acetylneuraminate synthase	COG2089	[M]
SAR11G3_00102	UDP-N-acetylglucosamine 4,6-dehydratase	COG1083	[M]
SAR11G3_00103	UDP-N-acetylglucosamine 4,6-dehydratase	COG1086	[MG]
SAR11G3_00104	hypothetical protein	NA	NA
SAR11G3_00105	hypothetical protein	NA	NA
SAR11G3_00107	acylneuraminate cytidyltransferase	NA	NA

SAR11G3_00108	ptmF	NA	NA
SAR11G3_00111	hypothetical protein	NA	NA
SAR11G3_00112	legionaminic acid biosynthesis protein PtmG	COG0037	[D]
SAR11G3_00114	D-glycero-D-manno-heptose 1-phosphate guanosyltransferase	NA	NA
SAR11G3_00115	hypothetical protein	COG2120	[S]
SAR11G3_00116	formyltransferase, putative	NA	NA
SAR11G3_00117	UDP-N-acetylglucosamine 2-epimerase	COG0381	[M]
SAR11G3_00118	N-acetylneuraminate synthase	COG2089	[M]
SAR11G3_00119	hypothetical protein	COG0110	[R]
SAR11G3_00120	putative aminotransferase, DegT family	COG0399	[M]
SAR11G3_00123	hypothetical protein	NA	NA
SAR11G3_00124	phosphoenolpyruvate synthase / Pyruvate phosphate dikinase	COG0574	[G]
SAR11G3_00125	phosphoenolpyruvate synthase / Pyruvate phosphate dikinase	COG0574	[G]
SAR11G3_00126	glutamine amidotransferase class-I	COG2071	[R]
SAR11G3_00127	sugar metabolism cluster protein	COG1213	[M]
SAR11G3_00128	adenylylsulfate kinase	COG0529	[P]
SAR11G3_00129	hypothetical protein	NA	NA
SAR11G3_00130	hypothetical protein	NA	NA
SAR11G3_00131	hypothetical protein	NA	NA
SAR11G3_00132	hypothetical protein	NA	NA
SAR11G3_00133	hypothetical protein	NA	NA
SAR11G3_00134	hypothetical protein	NA	NA
SAR11G3_00135	hypothetical protein	NA	NA
SAR11G3_00136	hypothetical protein	NA	NA
SAR11G3_00137	hypothetical protein	NA	NA
SAR11G3_00138	hypothetical protein	COG0500	[QR]
SAR11G3_00154	hypothetical protein	NA	NA
SAR11G3_00164	hypothetical protein	NA	NA
SAR11G3_00169	hypothetical protein	NA	NA
SAR11G3_00173	D-amino acid dehydrogenase small subunit	COG0665	[E]
SAR11G3_00174	permease of the drug/metabolite transporter (DMT) superfamily	COG0697	[GER]
SAR11G3_00178	aldehyde dehydrogenase	NA	NA
SAR11G3_00189	ferric iron ABC transporter, permease protein	COG1178	[P]
SAR11G3_00190	ferric iron ABC transporter, iron-binding protein	COG1840	[P]
SAR11G3_00191	hypothetical protein	NA	NA
SAR11G3_00192	metal-dependent hydrolases of the beta-lactamase superfamily I; PhnP protein	COG1235	[R]
SAR11G3_00204	hypothetical protein	NA	NA
SAR11G3_00205	VcbS	NA	NA
SAR11G3_00206	outer membrane autotransporter	NA	NA
SAR11G3_00207	hypothetical protein	NA	NA
SAR11G3_00209	putative secretory autotransporter	NA	NA
SAR11G3_00210	putative secretory autotransporter	NA	NA
SAR11G3_00212	hypothetical protein	NA	NA
SAR11G3_00213	hypothetical protein	NA	NA
SAR11G3_00214	putative pseudo-pilin PulG	NA	NA
SAR11G3_00215	fimbrial protein pilin	NA	NA
SAR11G3_00216	sulfate transporter/antisigma-factor antagonist STAS	COG1366	[T]
SAR11G3_00219	type II Secretion System PilC	NA	NA
SAR11G3_00222	type 4 fimbrial biogenesis protein PilP	NA	NA

SAR11G3_00223	hypothetical protein	NA	NA
SAR11G3_00227	hypothetical protein	NA	NA
SAR11G3_00230	type IV fimbrial biogenesis protein PilY1	COG3419	[NU]
SAR11G3_00231	hypothetical protein	NA	NA
SAR11G3_00232	type IV pilus assembly protein PilW	NA	NA
SAR11G3_00233	type IV pilus assembly protein PilW	NA	NA
SAR11G3_00234	putative pilV-like pilin protein	NA	NA
SAR11G3_00235	putative pilV-like pilin protein	NA	NA
SAR11G3_00236	putative pilin protein PilU	NA	NA
SAR11G3_00238	outer membrane autotransporter	NA	NA
SAR11G3_00239	hypothetical protein	NA	NA
SAR11G3_00240	hypothetical protein	NA	NA
SAR11G3_00244	transcriptional regulator, GntR family	COG1802	[K]
SAR11G3_00266	hypothetical protein	NA	NA
SAR11G3_00277	putative manganese transporter, 11 TMS	NA	NA
SAR11G3_00305	hypothetical protein	NA	NA
SAR11G3_00323	hypothetical protein	NA	NA
SAR11G3_00348	hypothetical protein	NA	NA
SAR11G3_00352	hypothetical protein	NA	NA
SAR11G3_00354	hypothetical protein	NA	NA
SAR11G3_00372	uroporphyrinogen-III synthase	NA	NA
SAR11G3_00378	hypothetical protein	NA	NA
SAR11G3_00382	YGGT family	NA	NA
SAR11G3_00389	UPF0434 protein YcaR	COG2835	[S]
SAR11G3_00393	uroporphyrinogen decarboxylase	COG0407	[H]
SAR11G3_00414	hypothetical protein	NA	NA
SAR11G3_00428	hypothetical protein	NA	NA
SAR11G3_00431	thiaminase II	COG0819	[K]
SAR11G3_00437	hypothetical protein	COG4338	[S]
SAR11G3_00458	putative transmembrane protein	NA	NA
SAR11G3_00464	hypothetical protein	NA	NA
SAR11G3_00506	hypothetical protein	NA	NA
SAR11G3_00510	hypothetical protein	NA	NA
SAR11G3_00518	hypothetical protein	NA	NA
SAR11G3_00521	hypothetical protein	NA	NA
SAR11G3_00526	multi-domain protein	COG1254	[C]
SAR11G3_00528	hypothetical protein	NA	NA
SAR11G3_00547	hypothetical protein	NA	NA
SAR11G3_00561	hypothetical protein	NA	NA
SAR11G3_00563	hypothetical protein	NA	NA
SAR11G3_00567	hypothetical protein	NA	NA
SAR11G3_00571	hypothetical protein	NA	NA
SAR11G3_00572	hypothetical protein	NA	NA
SAR11G3_00573	hypothetical protein	NA	NA
SAR11G3_00579	hypothetical protein	NA	NA
SAR11G3_00580	hypothetical protein	NA	NA
SAR11G3_00581	glycosyl transferase, family 2	COG0463	[M]
SAR11G3_00582	hypothetical protein	NA	NA
SAR11G3_00594	hypothetical protein	NA	NA
SAR11G3_00596	hypothetical protein	NA	NA
SAR11G3_00604	nucleoside-diphosphate-sugar epimerase	COG0451	[MG]
SAR11G3_00614	hypothetical protein	NA	NA
SAR11G3_00615	hypothetical protein	NA	NA

SAR11G3_00616	hypothetical protein	COG0451	[MG]
SAR11G3_00620	hypothetical protein	NA	NA
SAR11G3_00628	hypothetical protein	NA	NA
SAR11G3_00634	septum formation protein Maf	COG0424	[D]
SAR11G3_00641	hypothetical protein	COG3788	[R]
SAR11G3_00672	hypothetical protein	NA	NA
SAR11G3_00673	putative subunit of Alternative cytochrome c oxidase	COG5605	[S]
SAR11G3_00674	alternative cytochrome c oxidase polypeptide CoxP	COG1845	[C]
SAR11G3_00675	alternative cytochrome c oxidase polypeptide CoxN	COG1845	[C]
SAR11G3_00678	hypothetical protein	NA	NA
SAR11G3_00680	hypothetical protein	NA	NA
SAR11G3_00683	cytochrome oxidase biogenesis protein Surf1, facilitates heme A insertion	COG3346	[S]
SAR11G3_00690	phytoene synthase	NA	NA
SAR11G3_00691	phytoene synthase	NA	NA
SAR11G3_00701	hypothetical protein	NA	NA
SAR11G3_00751	hypothetical protein	NA	NA
SAR11G3_00769	hypothetical protein	NA	NA
SAR11G3_00772	hypothetical protein	NA	NA
SAR11G3_00771	hypothetical protein	NA	NA
SAR11G3_00777	hypothetical protein	NA	NA
SAR11G3_00779	hypothetical protein	COG0620	[E]
SAR11G3_00780	putative transporter component	NA	NA
SAR11G3_00784	alkyl hydroperoxide reductase subunit C-like protein	COG0450	[O]
SAR11G3_00787	hypothetical protein	NA	NA
SAR11G3_00802	hypothetical protein	NA	NA
SAR11G3_00805	hypothetical protein	NA	NA
SAR11G3_00810	hypothetical protein	NA	NA
SAR11G3_00828	putative integral membrane protein	COG0730	[R]
SAR11G3_00832	hypothetical protein	NA	NA
SAR11G3_00841	nitrogen regulatory protein P-II	COG0347	[E]
SAR11G3_00852	hypothetical protein	COG3564	[S]
SAR11G3_00856	hypothetical protein	NA	NA
SAR11G3_00860	hypothetical protein	NA	NA
SAR11G3_00861	hypothetical protein	NA	NA
SAR11G3_00863	cold-shock DEAD-box protein A	COG0513	[LKJ]
SAR11G3_00864	snoK-like protein	NA	NA
SAR11G3_00865	hypothetical protein	COG0500	[QR]
SAR11G3_00866	hypothetical protein	COG4301	[S]
SAR11G3_00867	glutaredoxin-like protein NrdH, required for reduction of Ribonucleotide reductase class Ib	NA	NA
SAR11G3_00870	hypothetical protein	NA	NA
SAR11G3_00872	tiTiN family	NA	NA
SAR11G3_00874	hypothetical protein	NA	NA
SAR11G3_00885	hypothetical protein	NA	NA
SAR11G3_00887	hypothetical protein	NA	NA
SAR11G3_00891	hypothetical protein	NA	NA
SAR11G3_00890	hypothetical protein	COG1238	[S]
SAR11G3_00907	ADP-heptose:LPS heptosyltransferase	COG0859	[M]
SAR11G3_00913	hypothetical protein	NA	NA
SAR11G3_00914	hypothetical protein	NA	NA
SAR11G3_00917	hypothetical protein	NA	NA
SAR11G3_00918	hypothetical protein	NA	NA
SAR11G3_00957	hypothetical protein	COG2104	[H]

SAR11G3_00959	hypothetical protein	NA	NA
SAR11G3_00961	DNA repair and recombination protein, putative helicase	COG0507	[L]
SAR11G3_00962	hypothetical protein	NA	NA
SAR11G3_00963	hypothetical protein	NA	NA
SAR11G3_00964	hypothetical protein	NA	NA
SAR11G3_00965	guanosine-3',5'-bis(Diphosphate) 3'-pyrophosphohydrolase	COG0317	[TK]
SAR11G3_00966	hypothetical protein	NA	NA
SAR11G3_00967	hypothetical protein	NA	NA
SAR11G3_00968	hypothetical protein	NA	NA
SAR11G3_00969	hypothetical protein	NA	NA
SAR11G3_00970	hypothetical protein	NA	NA
SAR11G3_00971	hypothetical protein	NA	NA
SAR11G3_00972	hypothetical protein	NA	NA
SAR11G3_00974	alternative oxidase	NA	NA
SAR11G3_00975	hypothetical protein	NA	NA
SAR11G3_00976	hypothetical protein	NA	NA
SAR11G3_00978	hypothetical protein	NA	NA
SAR11G3_00979	hypothetical protein	NA	NA
SAR11G3_00980	hypothetical protein	NA	NA
SAR11G3_00981	hypothetical protein	NA	NA
SAR11G3_00982	hypothetical protein	COG4889	[R]
SAR11G3_00983	hypothetical protein	NA	NA
SAR11G3_00984	hypothetical protein	NA	NA
SAR11G3_00985	hypothetical protein	NA	NA
SAR11G3_00998	hypothetical protein	NA	NA
SAR11G3_01004	hypothetical protein	NA	NA
SAR11G3_01022	hypothetical protein	NA	NA
SAR11G3_01027	hypothetical protein	NA	NA
SAR11G3_01028	hypothetical protein	NA	NA
SAR11G3_01037	hypothetical protein	NA	NA
SAR11G3_01040	hypothetical protein	NA	NA
SAR11G3_01044	hypothetical protein	NA	NA
SAR11G3_01057	hypothetical protein	NA	NA
SAR11G3_01074	hypothetical protein	NA	NA
SAR11G3_01088	protein-L-isoaspartate(D-aspartate) O-methyltransferase	COG2518	[O]
SAR11G3_01105	hypothetical protein	NA	NA
SAR11G3_01106	hypothetical protein	NA	NA
SAR11G3_01107	hypothetical protein	NA	NA
SAR11G3_01128	hypothetical protein	NA	NA
SAR11G3_01138	hypothetical protein	NA	NA
SAR11G3_01141	extracellular solute-binding protein, family 1	COG0725	[P]
SAR11G3_01161	hypothetical protein	NA	NA
SAR11G3_01160	porphobilinogen synthase	COG0113	[H]
SAR11G3_01231	hypothetical protein	NA	NA
SAR11G3_01247	hypothetical protein	NA	NA
SAR11G3_01250	quaternary ammonium compound-resistance protein qacH	COG2076	[P]
SAR11G3_01251	hypothetical protein	NA	NA
SAR11G3_01252	hypothetical protein	NA	NA
SAR11G3_01264	tRNA/rRNA methyltransferase	COG0565	[J]
SAR11G3_01265	tRNA:Cm32/Um32 methyltransferase	COG0565	[J]
SAR11G3_01270	hypothetical protein	NA	NA
SAR11G3_01271	protein of unknown function DUF218	NA	NA
SAR11G3_01274	N-acetylglucosamine-1-phosphate uridyltransferase /	COG1207	[M]

	Glucosamine-1-phosphate N-acetyltransferase		
SAR11G3_01279	hypothetical protein	NA	NA
SAR11G3_01282	hypothetical protein	COG0412	[Q]
SAR11G3_01288	hypothetical protein	NA	NA
SAR11G3_01289	protein involved in catabolism of external DNA	COG2961	[R]
SAR11G3_01291	dolichol-phosphate mannosyltransferase	NA	NA
SAR11G3_01292	hypothetical protein	NA	NA
SAR11G3_01302	hypothetical protein	NA	NA
SAR11G3_01304	deoxyribodipyrimidine photolyase	COG0415	[L]
SAR11G3_01305	hypothetical protein	NA	NA
SAR11G3_01306	hypothetical protein	NA	NA
SAR11G3_01307	sulfate permease	COG0659	[P]
SAR11G3_01308	copper-binding protein	COG4454	[P]
SAR11G3_01309	copper-binding protein	COG4454	[P]
SAR11G3_01310	hypothetical protein	NA	NA
SAR11G3_01315	4-carboxymuconolactone decarboxylase	COG0599	[S]
SAR11G3_01336	HI0933 family protein	COG2081	[R]
SAR11G3_01341	K ⁺ -dependent Na ⁺ /Ca ⁺ exchanger related-protein	COG0530	[P]
SAR11G3_01342	hypothetical protein	NA	NA
SAR11G3_01347	hypothetical protein	NA	NA
SAR11G3_01352	hypothetical protein	NA	NA
SAR11G3_01357	5'-methylthioadenosine nucleosidase / S-adenosylhomocysteine nucleosidase	COG0775	[F]
SAR11G3_01360	cytochrome c	COG2010	[C]
SAR11G3_01364	substrate-binding region of ABC-type glycine betaine transport system	COG2113	[E]
SAR11G3_01365	hypothetical protein	NA	NA
SAR11G3_01373	ferric iron ABC transporter, permease protein	COG1178	[P]
SAR11G3_01374	hypothetical protein	NA	NA
SAR11G3_01375	hypothetical protein	NA	NA
SAR11G3_01376	ferric iron ABC transporter, iron-binding protein	COG1840	[P]
SAR11G3_01380	hypothetical protein	NA	NA
SAR11G3_01383	putative cytochrome c protein	COG2010	[C]
SAR11G3_01396	hypothetical protein	NA	NA
SAR11G3_01397	hypothetical protein	NA	NA
SAR11G3_01399	membrane associated lipoprotein precursor	NA	NA
SAR11G3_01400	hypothetical protein	NA	NA
SAR11G3_01401	hypothetical protein	NA	NA
SAR11G3_01408	D-alanine--D-alanine ligase	COG1181	[M]

Paralogous genes in ACE_P1a.1

Gene ID	Annotation	COG	COG Category
feature:167713900	putative bifunctional UDP-N-acetylmuramoylalanyl-D - glutamate--2,6-diaminopimelate ligase/UDP-N-acetylmuramoyl- tripeptide:D-alanyl-D-alanine ligase	NA	NA
feature:167713946	hypothetical protein PU1002_06296	COG4696	[S]
feature:167713968	FAD-dependent thymidylate synthase	COG1351	[F]
feature:167714018	glutaredoxin	COG0695	[O]
feature:167714028	phosphoglycerate dehydrogenase	COG1052	[CHR]
feature:167714030	SelR domain-containing protein	COG0229	[O]
feature:167714038	hypothetical protein PU1002_06416	NA	NA
feature:167714040	LysE type translocator	NA	NA
feature:167714042	Leu/Ile/Val-binding protein precursor	COG0683	[E]

feature:167714044	ABC-type branched-chain amino acid transport system permease protein I	COG0559	[E]
feature:167714046	high-affinity branched-chain amino acid transport permease protein	COG4177	[E]
feature:167714048	ATP-binding protein of ABC transporter NatA	COG0411	[E]
feature:167714050	branched-chain amino acid ABC transporter, ATP-binding protein	COG0410	[E]
feature:167714056	hypothetical protein SAR11_1354	NA	NA
feature:167714070	hypothetical protein SAR11_1356	NA	NA
feature:167714072	hypothetical protein PU1002_06456	NA	NA
feature:167714104	luciferase-like monooxygenase	NA	NA
feature:167714114	spermidine/putrescine-binding periplasmic protein	COG0687	[E]
feature:167714116	spermidine/putrescine transport system permease protein potb	COG1176	[E]
feature:167714118	spermidine/putrescine transport system permease protein potc	COG1177	[E]
feature:167714120	3-hydroxyacyl-CoA dehydrogenase	COG1250	[I]
feature:167714122	peptidylprolyl cis-trans isomerase precursor	COG0652	[O]
feature:167714124	conserved hypothetical protein	NA	NA
feature:167714130	Permease for cytosine/purines, uracil, thiamine, allantoin	COG1953	[FH]
feature:167714132	Xanthine/uracil/vitamin C permease family protein	COG2252	[R]
feature:167714168	aconitate hydratase	COG1048	[C]
feature:167714200	glutamate synthase large subunit-like protein	COG0069	[E]
feature:167714202	GXGXG motif-containing protein	COG0070	[E]
feature:167714206	GlnT-like protein	COG0174	[E]
feature:167714208	glutamine synthetase catalytic domain	COG0174	[E]
feature:167717506	DNA-directed DNA polymerase gamma/tau subunit	COG2812	[L]
feature:167717562	DNA recombination protein	NA	NA
feature:167717618	hypothetical protein SAR11_0470	NA	NA
feature:167717670	putative inner membrane protein translocase component YidC	COG0706	[U]
feature:167717674	DNA topoisomerase (ATP-hydrolyzing) chain B (DNA gyrase B)	COG0187	[L]
feature:167717678	phenylalanine-tRNA ligase	COG0072	[J]
feature:167717710	pyruvate dehydrogenase (lipoamide) e1 component	COG2609	[C]
feature:167717718	ATP-dependent helicase	COG0513	[LKJ]
feature:167717728	Putative long-chain-fatty-acid--CoA ligase (Long-chain acyl-CoA synthetase)	NA	NA
feature:167717792	protein translation initiation factor 2	COG0532	[J]
feature:167717828	possible methylase or methyltransferase	NA	NA
feature:167717844	dihydrodipicolinate reductase	COG0289	[E]
feature:167717868	DNA-directed DNA polymerase III	NA	NA
feature:167717880	tRNA modification GTPase	COG0486	[R]
feature:167717882	sulfotransferase domain-containing protein	NA	NA
feature:167717884	putative Sulfotransferase domain protein	NA	NA
feature:167717886	transcription termination factor Rho	COG1158	[K]
feature:167717896	hypothetical protein SAR11_0343	COG1806	[S]
feature:167717900	Maf-like protein	COG0424	[D]
feature:167719306	isoleucyl-tRNA synthetase	COG0060	[J]
feature:167719324	3-deoxy-D-manno-octulosonic-acid transferase	COG1519	[M]
feature:167719348	cytochrome c oxidase polypeptide II (cytochrome aa 3 subunit 2)	COG1622	[C]
feature:167719364	cytochrome C oxidase	COG1622	[C]
feature:167719366	cytochrome C oxidase	COG0843	[C]

feature:167719368	cytochrome c oxidase subunit III	COG1845	[C]
feature:167719376	hypothetical protein SAR11_0901	NA	NA
feature:167719378	threonine synthase	COG0498	[E]
feature:167719470	excinuclease ABC subunit B	COG0556	[L]
feature:167719476	thioredoxin reductase	COG0492	[O]
feature:167719494	Carbamoyl-phosphate synthase small chain	COG0505	[EF]
feature:167719496	YqeY-like protein	COG1610	[S]
feature:167719498	DNA primase	NA	NA
feature:167719500	RNA polymerase sigma factor	COG0568	[K]
feature:167719502	carbamoylphosphate synthase small subunit	COG0505	[EF]
feature:167719692	putative dehydrogenase	NA	NA
feature:167719710	hypothetical protein ALOHA_HF4000141F21ctg1g27	NA	NA
feature:167719716	NADPH:quinone reductase	COG0604	[CR]
feature:167719740	Dihydrodipicolinate synthetase	COG0329	[EM]
feature:167719750	acetylornithine deacetylase	COG0624	[E]
feature:167719760	aminomethyl transferase family protein	NA	NA
feature:167719790	2-isopropylmalate synthase	COG0119	[E]
feature:167719832	ribosomal protein L1	COG0081	[J]
feature:167719840	DNA-directed RNA polymerase	COG0085	[K]
feature:167719914	probable UTP-glucose-1-phosphate uridylyltransferase	COG1210	[M]
feature:167719926	DNA topoisomerase I	COG0550	[L]
feature:167719928	hypothetical protein SAR11_1082	NA	NA
feature:167719934	aspartate carbamoyltransferase	COG0540	[F]
feature:167734036	permease YjgP/YjgQ family protein	NA	NA
feature:167734040	organic solvent tolerance-like protein	COG1452	[M]
feature:167734046	4-hydroxythreonine-4-phosphate dehydrogenase	NA	NA
feature:167734050	dimethyladenosine transferase	COG0030	[J]
feature:167734136	glycine dehydrogenase	COG0403	[E]
feature:167734138	glycine cleavage H-protein	COG0509	[E]
feature:167734142	integral membrane protein	COG0697	[GER]
feature:167774847	alanine-tRNA ligase	COG0013	[J]
feature:167774851	phosphatidylserine decarboxylase	COG0688	[I]
feature:167774857	glycosyl transferase family 8	NA	NA
feature:167774861	IMP dehydrogenase-like protein	COG0516	[F]
feature:167774887	glycine dehydrogenase	COG0403	[E]
feature:167774889	glycine cleavage system protein P2 gcvP	COG0403	[E]
feature:167774893	glycine dehydrogenase	COG1003	[E]
feature:167774895	Flavin containing amine oxidoreductase	COG2907	[R]
feature:167774897	hypothetical protein PU1002_03016	COG3496	[S]
feature:167818750	nitrogen regulation protein ntrY	NA	NA
feature:167818752	IspD/ispF bifunctional enzyme	COG1211	[I]
feature:167818756	phosphatidylglycerophosphatase A	COG1267	[I]
feature:167818758	competence/damage-inducible protein CinA	COG1546	[R]
feature:167818760	polyketide cyclase/dehydrase superfamily protein	COG2867	[I]
feature:167818762	lipoyl synthase	COG0320	[H]
feature:167818766	Enolase	COG0148	[G]
feature:167818768	2-dehydro-3-deoxy-phosphooctonate aldolase	COG2877	[M]
feature:167818772	preprotein translocase SecE subunit	COG1314	[U]
feature:167818774	triosephosphate isomerase (tim)	COG0149	[G]
feature:167818780	anthranilate synthase component I-like protein	COG0147	[EH]
feature:167818782	anthranilate synthase component II	COG0512	[EH]
feature:167818786	indole-3-glycerol phosphate synthase/anthranilate isomerase	COG0134	[E]
feature:167818790	glutamyl-tRNA synthetase	COG0008	[J]

feature:167818794	conserved hypothetical protein	COG3494	[S]
feature:167818798	glucosamine N-acyltransferase	COG1044	[M]
feature:167818800	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase	COG0764	[I]
feature:167818802	Outer membrane protein (OmpH-like)	NA	NA
feature:167818818	30S ribosomal protein S2	COG0052	[J]
feature:167818826	prolyl-tRNA synthetase	COG0442	[J]
feature:167818832	Bvg accessory factor	COG1521	[K]
feature:167818840	NADH dehydrogenase subunit M	COG1008	[C]
feature:167818842	NADH Dehydrogenase I Chain L	COG1009	[CP]
feature:167818852	NADH dehydrogenase subunit G	COG1034	[C]
feature:167818854	NADH dehydrogenase I subunit F	COG1894	[C]
feature:167818856	NADH Dehydrogenase I Chain E	COG1905	[C]
feature:167818858	NADH dehydrogenase delta subunit	COG0649	[C]
feature:167818866	ATP-dependent protease La	COG0466	[O]
feature:167818868	ATP-dependent protease La	COG0466	[O]
feature:167818870	ATP-dependent clp proteinase regulatory chain X	COG1219	[O]
feature:167818900	TRAP-type mannitol/chloroaromatic compound transport system protein	COG4665	[Q]
feature:167818950	AMP-dependent synthetase and ligase	NA	NA
feature:167818960	oxidoreductase, FAD-binding	COG0665	[E]
feature:167818962	putative acetolactate synthase large subunit	COG0028	[EH]
feature:167818978	bifunctional proline dehydrogenase/pyrroline-5-carboxylate dehydrogenase	COG4230	[C]
feature:167818990	hypothetical protein PU1002_00095	NA	NA
feature:167818996	AziC protein	COG1296	[E]
feature:167819002	hypothetical protein PU1002_00090	COG4392	[S]
feature:167819010	iron-uptake ABC transport system periplasmic iron-binding protein	COG1840	[P]
feature:167819056	oxidoreductase, short chain dehydrogenase/reductase family	COG1028	[IQR]
feature:167819096	hypothetical protein Avi_3473	COG0678	[O]
feature:167819124	hypothetical protein SAR11_1269	NA	NA
feature:167819138	alpha/beta hydrolase fold protein	COG0596	[R]
feature:167819174	glycine betaine/L-proline transport system permease	NA	NA
feature:167819204	putative monomeric sarcosine oxidase	COG0665	[E]
feature:167819212	glutamate synthase large subunit-like protein	COG0069	[E]
feature:167819336	nucleotidyltransferase family protein	NA	NA
feature:167819344	AcrB/AcrD/AcrF family protein (Acriflavin resistance)	COG0841	[V]
feature:167819356	drug/metabolite transporter	COG0697	[GER]
feature:167819358	DNA-3-methyladenine glycosylase I	COG0122	[L]
feature:167819360	hypothetical protein SAR11_0185	COG5470	[S]
feature:167819368	hypothetical protein PU1002_05456	NA	NA
feature:167819370	oligopeptidase B	COG1770	[E]
feature:167819376	dapA-like protein	NA	NA
feature:167819386	protein that suppresses inhibitory activity of Csr A	COG0312	[R]
feature:167819406	probable ATP-dependent RNA helicase	COG0513	[LKJ]
feature:167866888	fructose-bisphosphate aldolase	COG3588	[G]
feature:167866914	translocation protein TolB	COG0823	[U]
feature:167866920	hypothetical protein PU1002_03336	NA	NA
feature:167866928	ATP-dependent Zn protease	COG0465	[O]
feature:167866970	dihydroorotase	COG0044	[F]
feature:167866972	small multidrug resistance protein	COG2076	[P]
feature:167867006	alanine-tRNA ligase	COG0013	[J]

feature:167867008	alanine-tRNA ligase	COG0013	[J]
feature:167867014	phosphatidylserine synthase	COG1183	[I]
feature:167867016	Glycosyl transferase family 8	NA	NA
feature:167867022	TPR-like	NA	NA
feature:167867024	bifunctional GMP synthase/glutamine amidotransferase protein	COG0519	[F]
feature:167867026	3-methyl-2-oxobutanoate hydroxymethyltransferase	COG0413	[H]
feature:167867048	agmatinase	COG0010	[E]
feature:167888776	ATP-dependent Clp protease proteolytic subunit	COG0740	[OU]
feature:167888780	ATP-dependent protease La	COG0466	[O]
feature:167888784	NADH dehydrogenase I chain A	COG0838	[C]
feature:167888786	NADH dehydrogenase subunit B	COG0377	[C]
feature:167888788	NADH dehydrogenase I chain C	COG0852	[C]
feature:167888800	NADH Dehydrogenase I Chain H	COG1005	[C]
feature:167888802	NADH dehydrogenase subunit I	COG1143	[C]
feature:167888804	NADH dehydrogenase I chain J	COG0839	[C]
feature:167888812	NADH dehydrogenase subunit N	COG1007	[C]
feature:167888818	metallo-beta-lactamase family protein	COG0595	[R]
feature:167888820	conserved hypothetical protein	NA	NA
feature:167888826	prolyl-tRNA synthetase	COG0442	[J]
feature:167888828	ABC transporter	COG4591	[M]
feature:167888842	putative undecaprenyl diphosphate synthase	COG0020	[I]
feature:167888846	1-deoxy-D-xylulose 5-phosphate reductoisomerase	COG0743	[I]
feature:167888856	acyl-[acyl carrier protein]-UDP-N-acetylglucosamine O-acyltransferase	COG1043	[M]
feature:167888858	hypothetical protein PU1002_01826	NA	NA
feature:167888866	LexA repressor	COG1974	[KT]
feature:167888870	anthranilate phosphoribosyltransferase	COG0547	[E]
feature:167888882	CTP synthetase	COG0504	[F]
feature:167888886	CTP synthase	COG0504	[F]
feature:167888890	Enolase	COG0148	[G]
feature:167888902	IspD/ispF bifunctional enzyme	COG0245	[I]
feature:167907230	heat shock protein	NA	NA
feature:167907232	oligopeptidase B	COG1770	[E]
feature:167907240	Xanthine/uracil/vitamin C permease family protein	COG2252	[R]
feature:167907242	phosphoribosyltransferase	COG2236	[R]
feature:167907250	DMT family permease	COG0697	[GER]
feature:167907252	Peptide methionine sulfoxide reductase	COG0225	[O]
feature:167907254	thiosulfate sulfurtransferase	COG2897	[P]
feature:167907256	HlyD family-like protein	COG0845	[M]
feature:167907258	AcrB/AcrD/AcrF family protein (Acriflavin resistance)	COG0841	[V]
feature:167907260	AcrB/AcrD/AcrF family protein (acriflavin resistance)	COG0841	[V]
feature:167907262	soluble lytic murein transglycosylase	COG4764	[S]
feature:167907264	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	COG0588	[G]
feature:167907322	glutamate 5-kinase	COG0263	[E]
feature:167907346	probable primosomal protein N	NA	NA
feature:167907348	probable primosomal protein N	COG1198	[L]
feature:167907354	2-oxoglutarate dehydrogenase complex E2 component	COG0508	[C]
feature:167907368	succinate dehydrogenase	COG0479	[C]
feature:167907370	succinate dehydrogenase iron-sulfur subunit	COG0479	[C]
feature:167907386	acyl-CoA dehydrogenase	COG1960	[I]
feature:167907406	cell division particle	COG0552	[U]
feature:167907412	Transporter	COG0697	[GER]

feature:167907424	mannitol transporter	COG4665	[Q]
feature:167907442	(S)-2-hydroxy-acid oxidase	COG0247	[C]
feature:167907542	putative Sulfotransferase domain protein	NA	NA
feature:167907592	hypothetical protein SAR11_0343	COG1806	[S]
feature:167907594	uroporphyrinogen decarboxylase	COG0407	[H]
feature:167907602	tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA	COG0445	[D]
feature:167907608	DNA-directed DNA polymerase III	NA	NA
feature:167907614	Leucyl-tRNA synthetase	COG0495	[J]

Paralogous genes in ACE_P3.2

Gene ID	Annotation	COG	COG Category
feature:167687116	IclR family transcriptional regulator	NA	NA
feature:167687124	fumarate lyase	COG0015	[F]
feature:167687138	Oligopeptide/dipeptide transporter domain family protein	COG0444	[EP]
feature:167687158	ribonuclease E/G	COG1530	[J]
feature:167687160	ribonuclease E/G	COG1530	[J]
feature:167687166	peptide chain release factor 2	COG1186	[J]
feature:167687172	anhydro-N-acetylmuramic acid kinase	COG2377	[O]
feature:167687176	alpha/beta hydrolase	COG2945	[R]
feature:167687202	betaine aldehyde dehydrogenase (BADH) protein	COG1012	[C]
feature:167687216	Methionine sulfoxide reductase B	COG0229	[O]
feature:167687234	short chain dehydrogenase	COG1028	[IQR]
feature:167687268	glycine dehydrogenase	COG0403	[E]
feature:167687280	glycyl-tRNA synthetase subunit beta	COG0751	[J]
feature:167687284	glycyl-tRNA synthetase subunit beta	COG0751	[J]
feature:167687290	nicotinate-nucleotide pyrophosphorylase	COG0157	[H]
feature:167687350	metallo-beta-lactamase family protein	COG0595	[R]
feature:167687358	Lipoprotein releasing system, transmembrane protein n, LoIC/E family	COG4591	[M]
feature:167687366	DNA polymerase III alpha subunit	COG0587	[L]
feature:167687372	elongation factor Ts	COG0264	[J]
feature:167687376	putative undecaprenyl diphosphate synthase	COG0020	[I]
feature:167687396	citrate synthase	COG0372	[C]
feature:167687416	3-deoxy-8-phosphooctulonate synthase	COG2877	[M]
feature:167730530	heparinase II/III family protein	NA	NA
feature:167730538	hypothetical protein MELB17_17679	COG0596	[R]
feature:167730560	shikimate 5-dehydrogenase	COG0169	[E]
feature:167730582	putative glutamine amidotransferase class-I	COG0118	[E]
feature:167730634	DNA-(apurinic or apyrimidinic site) lyase / endonuclease III	COG0177	[L]
feature:167730640	dihydrodipicolinate reductase	COG0289	[E]
feature:167730642	Chaperone protein	COG0484	[O]
feature:167730660	possible transmembrane receptor	NA	NA
feature:167730696	hypothetical protein	NA	NA
feature:167730746	DNA gyrase, B subunit	COG0187	[L]
feature:167730754	deoxyhypusine synthase-like protein	COG1899	[O]
feature:167730778	adenosylhomocysteinase	COG0499	[H]
feature:167730794	AcrB/AcrD/AcrF family protein	COG0841	[V]
feature:167730822	oligopeptidase B	COG1770	[E]
feature:167730878	L-aspartate oxidase	COG0029	[H]
feature:167730902	short chain dehydrogenase	COG1028	[IQR]
feature:167730916	mandelate racemase/muconate lactonizing enzyme, N-	COG4948	[MR]

	terminal domain protein		
feature:167730918	mandelate racemase/muconate lactonizing enzyme, N-terminal domain protein	COG4948	[MR]
feature:167730954	Fe(3+) ions import ATP-binding protein FbpC	COG3842	[E]
feature:167730970	Phosphate permease	COG0306	[P]
feature:167730978	betaine aldehyde dehydrogenase (BADH) protein	COG1012	[C]
feature:167730994	acetylornithine aminotransferase	COG4992	[E]
feature:167731052	peptide chain release factor 1	NA	NA
feature:167731054	peptide chain release factor 1	COG0216	[J]
feature:167731146	Carbamoyltransferase	COG2192	[O]
feature:167731170	NAD(p) transhydrogenase subunit beta	COG1282	[C]
feature:167731174	penicillin binding protein transpeptide	COG0768	[M]
feature:167731178	rod shape-determining protein rodA	COG0772	[D]
feature:167731184	transport protein	COG0697	[GER]
feature:167731206	thymidylate synthase-complementing family protein (Pfam)	COG1351	[F]
feature:167731230	putative Trypsin	COG0265	[O]
feature:167731244	putative HMGL-like protein	COG0119	[E]
feature:167731262	thiamine-phosphate pyrophosphorylase	COG0352	[H]
feature:167774801	DSBA oxidoreductase	COG3917	[Q]
feature:167774819	thioredoxin-disulfide reductase	COG0492	[O]
feature:167816076	transcriptional regulator, NtrC family	COG2204	[T]
feature:167816086	amino acid ABC transporter	COG4597	[E]
feature:167816120	dihydroxy-acid dehydratase	COG0129	[EG]
feature:167816152	acetoacetate-CoA ligase	COG0365	[I]
feature:167816200	ATP-dependent DNA helicase RecG	COG1200	[LK]
feature:167816208	excinuclease ABC chain A	COG0178	[L]
feature:167824428	putative LysE type translocator	COG1280	[E]
feature:167824432	fumarylacetoacetate hydrolase family protein	COG0179	[Q]
feature:167824438	malonyl-CoA decarboxylase	COG1593	[G]
feature:167824444	trap dicarboxylate transporter- dctp subunit	COG1638	[G]
feature:167824456	3-hydroxyisobutyrate dehydrogenase	COG2084	[I]
feature:167824494	Acetaldehyde dehydrogenase II (ACDH-II)	COG1012	[C]
feature:167824496	ammonium transporter	COG0004	[P]
feature:167824500	Heptosyltransferase family	NA	NA
feature:167824512	phosphoserine aminotransferase	COG1932	[HE]
feature:167824520	putative peptidase family M50	COG0750	[M]
feature:167824530	putative zinc-binding dehydrogenase	NA	NA
feature:167824574	taurine dioxygenase	COG2175	[Q]
feature:167824580	aconitate hydratase 1	COG1048	[C]
feature:167824582	aconitate hydratase	COG1048	[C]
feature:167824586	aconitate hydratase	COG1048	[C]
feature:167824592	aconitate hydratase 1	COG1048	[C]
feature:167824594	putative short-chain dehydrogenase/reductase SDR	COG1028	[IQR]
feature:167824606	hypothetical protein BAL199_01704	COG0028	[EH]
feature:167824622	Methionine sulfoxide reductase B	COG0229	[O]
feature:167824638	hypothetical protein	NA	NA
feature:167824682	O-succinylhomoserine sulfhydrylase	COG0626	[E]
feature:167824686	Pc22g04140	COG4221	[R]
feature:167824714	UDP-N-acetylmuramate--L-alanine ligase	COG0773	[M]
feature:167824728	UDP-N-acetylmuramoylalanine--D-glutamate ligase	COG0771	[M]
feature:167824732	phospho-N-acetylmuramoyl-pentapeptide-transferase	COG0472	[M]
feature:167824750	lipoprotein	COG2853	[M]
feature:167824782	thioredoxin reductase	COG0492	[O]

feature:167824800	excinuclease ABC subunit C	COG0322	[L]
feature:167824852	site-specific DNA methyltransferase	COG0863	[L]
feature:167824910	TldD protein	COG0312	[R]
feature:167824922	putative modulator of DNA gyrase	COG0312	[R]
feature:167824924	putative modulator of DNA gyrase	COG0312	[R]
feature:167824932	3-deoxy-D-manno-octulosonic-acid transferase	NA	NA
feature:167824942	exodeoxyribonuclease VII, large subunit	COG1570	[L]
feature:167825016	ATP phosphoribosyltransferase regulatory subunit, putative	COG3705	[E]
feature:167825054	dihydrolipoamide S-acetyltransferase	COG0508	[C]
feature:167825074	phenylalanyl-tRNA synthetase subunit alpha	COG0016	[J]
feature:167825092	glutamate synthase, large subunit	COG0070	[E]
feature:167825106	arabinose 5-phosphate isomerase	COG0517	[R]
feature:167825130	tryptophan synthase, alpha subunit	NA	NA
feature:167825168	hypothetical protein	NA	NA
feature:167825170	'Cold-shock' DNA-binding domain, putative	COG1278	[K]
feature:167825216	cytochrome C	NA	NA
feature:167825296	argininosuccinate lyase	COG0165	[E]
feature:167825326	gamma-glutamyl phosphate reductase	COG0014	[E]
feature:167825344	FOF1 ATP synthase subunit gamma	COG0224	[C]
feature:167825348	FOF1 ATP synthase subunit alpha	COG0056	[C]
feature:167825362	dihydrolipoamide dehydrogenase	COG1249	[C]
feature:167825418	mannitol transporter	COG4665	[Q]
feature:167825426	(S)-2-hydroxy-acid oxidase	COG0247	[C]
feature:167825430	(S)-2-hydroxy-acid oxidase	COG0277	[C]
feature:167825438	probable aspartate transaminase	COG0436	[E]
feature:167825442	putative FMN-dependent dehydrogenase	COG1304	[C]
feature:167825450	Cytochrome c-type biogenesis protein	COG1138	[O]
feature:167905348	hypothetical protein MED121_17284	COG3333	[S]
feature:167905350	hypothetical protein SSE37_02345	COG3181	[S]
feature:167905396	hypothetical protein PU1002_00095	COG0697	[GER]
feature:167905452	hypothetical protein Oant_0061	NA	NA
feature:167905458	glucose-methanol-choline oxidoreductase	COG2303	[E]
feature:167905496	RNA polymerase sigma factor	COG0568	[K]
feature:167905500	phosphomethylpyrimidine kinase	COG0351	[H]
feature:167905508	ATP-dependent metalloprotease FtsH	COG0465	[O]
feature:167905542	hypothetical protein Sbal_4504	COG0057	[G]
feature:167932900	V-type H(+)-translocating pyrophosphatase	COG3808	[C]
feature:167932904	membrane-bound proton-translocating pyrophosphatase	COG3808	[C]
feature:167932928	DNA topoisomerase IV subunit A	NA	NA
feature:167932954	putative soluble lytic murein transglycosylase	NA	NA
feature:167932972	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	[J]
feature:167933044	ribosomal protein S3	COG0092	[J]
feature:167933066	DNA-directed RNA polymerase, beta' subunit	COG0086	[K]
feature:167933070	DNA-directed RNA polymerase, beta subunit	COG0085	[K]
feature:167933112	DNA topoisomerase IV subunit B	COG0187	[L]
feature:167933204	phosphoribosylformylglycinamide synthase II (FGA M synthase II)	COG0046	[F]
feature:167933238	glutamine-fructose-6-phosphate transaminase (isomerizing)	COG0449	[M]
feature:167933278	L-carnitine dehydrogenase	COG1804	[C]
feature:167933292	Mandelate racemase / muconate lactonizing enzyme, C-terminal domain protein	COG4948	[MR]
feature:167933300	histone deacetylase family protein	COG0123	[BQ]

feature:167935908	integrase/recombinase xerD-like protein	COG0582	[L]
feature:167935926	DNA gyrase, A subunit	COG0188	[L]
Paralogous genes in IMCC9063			
Gene ID	Annotation	COG	COG Category
SAR11G3_00007	HflK protein	COG0330	[O]
SAR11G3_00022	hypothetical transport protein	COG0697	[GER]
SAR11G3_00052	imidazole glycerol phosphate synthase cyclase subunit	COG0107	[E]
SAR11G3_00053	imidazole glycerol phosphate synthase amidotransferase subunit	COG0118	[E]
SAR11G3_00062	dTDP-glucose 4,6-dehydratase	COG0451	[MG]
SAR11G3_00063	carbamoyltransferase	COG2192	[O]
SAR11G3_00081	CDP-glucose 4,6-dehydratase-like protein	COG0451	[MG]
SAR11G3_00082	glucose-1-phosphate cytidylyltransferase	COG1208	[MJ]
SAR11G3_00096	amidase	COG0154	[J]
SAR11G3_00106	putative short-chain dehydrogenase family protein	COG1028	[IQR]
SAR11G3_00109	imidazole glycerol phosphate synthase cyclase subunit	COG0107	[E]
SAR11G3_00110	hypothetical protein	COG0118	[E]
SAR11G3_00122	phospholipid-lipopolysaccharide ABC transporter	COG1132	[V]
SAR11G3_00188	ferric iron ABC transporter, permease protein	COG3842	[E]
SAR11G3_00195	phosphate regulon transcriptional regulatory protein PhoB (SphR)	COG0745	[TK]
SAR11G3_00211	hypothetical protein	NA	NA
SAR11G3_00229	leader peptidase (Prepilin peptidase) / N-methyltransferase	COG1989	[NOU]
SAR11G3_00241	TRAP transporter solute receptor, TAXI family precursor	COG2358	[R]
SAR11G3_00249	dehydrogenases with different specificities (short-chain alcohol dehydrogenase)	COG1028	[IQR]
SAR11G3_00254	pyruvate dehydrogenase E1 component	COG2609	[C]
SAR11G3_00296	peptide methionine sulfoxide reductase MsrA	COG0225	[O]
SAR11G3_00300	RND multidrug efflux transporter; Acriflavin resistance protein	COG0841	[V]
SAR11G3_00316	agmatinase	NA	NA
SAR11G3_00329	phosphopantothenoylcysteine decarboxylase / Phosphopantothenoylcysteine synthetase	COG0452	[H]
SAR11G3_00336	peptide methionine sulfoxide reductase MsrB	COG0229	[O]
SAR11G3_00375	ygjD/Kae1/Qri7 family, required for threonylcarbamoyladenosine (t(6)A) formation in tRNA	COG0533	[O]
SAR11G3_00484	putative primosomal protein N	COG1198	[L]
SAR11G3_00519	long-chain-fatty-acid--CoA ligase	COG1022	[I]
SAR11G3_00562	cold shock DNA-binding domain protein	COG1278	[K]
SAR11G3_00586	tryptophan synthase alpha chain	COG0159	[E]
SAR11G3_00638	histidine-tRNA ligase	NA	NA
SAR11G3_00639	histidine-tRNA ligase	COG3705	[E]
SAR11G3_00656	exodeoxyribonuclease VII large subunit	COG1570	[L]
SAR11G3_00665	(S)-2-haloacid dehalogenase	COG1011	[R]
SAR11G3_00671	hypothetical protein	COG0517	[R]
SAR11G3_00676	alternative cytochrome c oxidase polypeptide CoxN	COG0843	[C]
SAR11G3_00677	alternative cytochrome c oxidase polypeptide CoxM	COG1622	[C]
SAR11G3_00721	uncharacterized ACR	COG1496	[S]
SAR11G3_00730	excinuclease ABC subunit C	COG0322	[L]
SAR11G3_00755	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase	COG0770	[M]

SAR11G3_00758	cell division protein FtsW	COG0772	[D]
SAR11G3_00782	hypothetical protein	COG2391	[R]
SAR11G3_00801	mmgE/PrpD	NA	NA
SAR11G3_00815	3-oxoacyl-[acyl-carrier protein] reductase	COG1028	[IQR]
SAR11G3_00850	ammonium transporter	COG0004	[P]
SAR11G3_00851	aldehyde dehydrogenase	COG1012	[C]
SAR11G3_00868	fumarate lyase	NA	NA
SAR11G3_00869	putative exported protein	NA	NA
SAR11G3_00871	tricarboxylate transport membrane protein TctA	COG3333	[S]
SAR11G3_00884	fumarylacetoacetate hydrolase family protein	COG0179	[Q]
SAR11G3_00888	transporter, LysE family	COG1280	[E]
SAR11G3_00915	Glutamine-hydrolyzing GMP synthase	COG0518	[F]
SAR11G3_00925	membrane protein	COG1033	[R]
SAR11G3_00927	alanine racemase	COG0787	[M]
SAR11G3_00945	putative phosphoribosylformylglycinamide cyclo-ligase	COG0150	[F]
SAR11G3_00973	short chain dehydrogenase	COG1028	[IQR]
SAR11G3_00994	glycyl-tRNA synthetase beta chain	COG0751	[J]
SAR11G3_01050	indole-3-glycerol phosphate synthase	COG0134	[E]
SAR11G3_01053	anthranilate synthase, amidotransferase component	COG0512	[EH]
SAR11G3_01064	lipoate synthase	COG0320	[H]
SAR11G3_01065	lipoate synthase	COG0320	[H]
SAR11G3_01072	nitrogen assimilation regulatory protein ntrX	COG2204	[T]
SAR11G3_01086	deoxyribodipyrimidine photolyase	NA	NA
SAR11G3_01170	pyridoxine 5'-phosphate synthase	COG0854	[H]
SAR11G3_01286	phosphoglucosamine mutase / Phosphomannomutase	COG1109	[G]
SAR11G3_01300	putative mandelate racemase/muconate lactonizing enzyme	COG4948	[MR]
SAR11G3_01316	aldehyde dehydrogenase	COG1012	[C]
SAR11G3_01320	AFG1-like ATPase	COG1485	[R]
SAR11G3_01343	putative UBX domain	NA	NA
SAR11G3_01348	succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	[C]
SAR11G3_01359	hypothetical protein	COG0697	[GER]
SAR11G3_01372	iron(III) ABC transporter	COG3842	[E]
SAR11G3_01393	hypothetical protein	NA	NA
SAR11G3_01395	major facilitator superfamily (MFS) transporter	NA	NA
SAR11G3_01438	permease of the major facilitator superfamily	NA	NA

Genes with signatures of positive selection in ACE_P3.2 vs HIMB114

HIMB114 gene number & annotation	ACE_P3.2 gene identifier & annotation	dN/dS*	COG	COG Category
HIMB114_1149, drug/metabolite transporter	feature:167687126, integral membrane protein, putative	1.4586	COG0697	[GER]
HIMB114_0784, conserved hypothetical protein	feature:167687168, conserved hypothetical protein	1.4274	NA	NA
HIMB114_0782, alpha/beta hydrolase	feature:167687178, alpha/beta superfamily hydrolase	1.018	COG2945	[R]
HIMB114_0778, FeS assembly protein SufD	feature:167687184, FeS assembly protein SufD	1.5523	COG0719	[O]
HIMB114_0855, putative glycine cleavage T protein	feature:167687206, predicted aminomethyltransferase	1.1647	COG0354	[R]
HIMB114_0303, pyridine nucleotide-disulphide oxidoreductase dimerization	feature:167687220, membrane protein, putative	1.2797	COG0398	[S]

region				
HIMB114_0114, conserved hypothetical protein	feature:167687242, hypothetical protein SAR11_1254	1.7505	NA	NA
HIMB114_0859, pyruvate, phosphate dikinase	feature:167687286, pyruvate, phosphate dikinase	35.2411	COG0574	[G]
HIMB114_0714, NADH dehydrogenase (quinone), g subunit	feature:167687326, NADH dehydrogenase gamma subunit	1.3748	COG1034	[C]
HIMB114_0697, translation elongation factor Ts	feature:167687370, elongation factor Ts	1.0298	COG0264	[J]
HIMB114_0692, outer membrane protein assembly complex, YaeT protein	feature:167687382, outer membrane protein	1.3497	COG4775	[M]
HIMB114_0680, anthranilate synthase component I	feature:167687406, anthranilate synthase component I-like protein	1.5901	COG0147	[EH]
HIMB114_0679, hypothetical protein	feature:167687408, peptidyl-prolyl cis-trans isomerase domain-containing protein	1.5899	NA	NA
HIMB114_0669, phosphatidylglycerophosphatase A	feature:167687430, ott:OTT_1074 pgpA; phosphatidylglycerophosphatase A; K01095 phosphatidylglycerophosphatase A	1.5501	COG1267	[I]
HIMB114_1291, hypothetical protein	feature:167730482, hypothetical protein	1.2739	NA	NA
HIMB114_1272, potential esterase	feature:167730522, potential esterase	1.2854	NA	NA
HIMB114_0558, uroporphyrinogen decarboxylase (URO-D) (UPD)	feature:167730552, uroporphyrinogen decarboxylase	1.2913	COG0407	[H]
HIMB114_1246, probable Smr	feature:167730572, Smr (small MutS related) protein	1.1418	NA	NA
HIMB114_1243, imidazole glycerol phosphate synthase, glutamine amidotransferase subunit	feature:167730580, putative glutamine amidotransferase class-I	1.4974	COG0118	[E]
HIMB114_1224, conserved hypothetical protein	feature:167730624, hypothetical protein SAR11_0362	1.577	COG3786	[S]
HIMB114_1220, conserved hypothetical protein	feature:167730630, hypothetical protein RB2654_17506	1.7518	COG3152	[S]
HIMB114_1150, conserved hypothetical protein	feature:167730798, 3-mercaptopyruvate sulfurtransferase	1.0854	COG2897	[P]
HIMB114_0187, omega-amino acid--pyruvate aminotransferase (Omega-APT)(Beta-alanine--pyruvate aminotransferase)	feature:167730814, putative aminotransferase class-III	1.2202	COG0161	[H]
HIMB114_1141, protease 2 (Protease II) (Oligopeptidase B)	feature:167730820, oligopeptidase B	3.9625	COG1770	[E]
HIMB114_1138, putative membrane-bound pili assembly protein Fimc	feature:167730826, hypothetical protein	1.7379	NA	NA
HIMB114_1135, conserved hypothetical protein	feature:167730834, hypothetical protein ALOHA_HF4000133G03ctg1g26	1.2948	NA	NA
HIMB114_0605, MatE efflux family protein	feature:167730858, MatE efflux family protein	1.4433	COG0534	[V]
HIMB114_1096, hypothetical protein	feature:167730928, hypothetical protein SKA53_04713	1.6863	NA	NA

HIMB114_1054, molecular chaperone	feature:167731016, molecular chaperone	1.449	COG2214	[O]
HIMB114_1046, acetyl-CoA carboxylase, carboxyl transferase, alpha subunit	feature:167731034, putative acetyl co-enzyme A carboxylase carboxyltransferase alpha subunit	1.0383	COG0825	[I]
HIMB114_1047, preprotein translocase, SecA subunit	feature:167731036, putative SecA protein, amino terminal region	1.0218	COG0653	[U]
HIMB114_0156, glucose-1-phosphate thymidyltransferase (dTDP-glucosyltransferase) (dTDP-glucose pyrophosphorylase)	feature:167731120, glucose-1-phosphate thymidyltransferase	3.1787	COG1209	[M]
HIMB114_0179, GDP-mannose 4,6-dehydratase (GDP-D-mannose dehydratase)	feature:167731130, gdpmannose 4,6-dehydratase	2.2966	COG1089	[M]
HIMB114_0998, NAD dependent epimerase/dehydratase family protein	feature:167731134, NAD-dependent epimerase/dehydratase	1.2891	COG0451	[MG]
HIMB114_0960, cytochrome c-type biogenesis protein CcdA	feature:167731192, probable holocytochrome-c synthase ccdA	1.118	COG0785	[O]
HIMB114_0956, conserved hypothetical protein	feature:167731198, hypothetical protein PU1002_00130	4.6627	COG1262	[S]
HIMB114_0938, hypothetical protein	feature:167731250, putative rod shape-determining protein MreD	1.4186	NA	NA
HIMB114_0797, probable ATPase involved in DNA replication initiation	feature:167774709, probable ATPase involved in DNA replication initiation	1.0344	COG0593	[L]
HIMB114_0799, hypothetical protein	feature:167774711, possible membrane protein	1.6918	NA	NA
HIMB114_0803, probable cytosol aminopeptidase (Leucine aminopeptidase)(LAP) (Leucyl aminopeptidase)	feature:167774719, leucyl aminopeptidase	1.0534	COG0260	[E]
HIMB114_0806, hypothetical protein	feature:167774725, organic solvent tolerance-like protein	1.487	COG1452	[M]
HIMB114_0808, pyridoxal phosphate biosynthetic protein PdxA	feature:167774729, 4-hydroxythreonine-4-phosphate dehydrogenase	1.1864	COG1995	[H]
HIMB114_0810, dimethyladenosine transferase	feature:167774731, dimethyladenosine transferase	1.5821	COG0030	[J]
HIMB114_0886, bacteriorhodopsin	feature:167774825, proteorhodopsin	5.0057	NA	NA
HIMB114_0659, conserved hypothetical protein	feature:167816098, hypothetical protein OA2633_02196	1.5937	COG3496	[S]
HIMB114_0602, single-stranded-DNA-specific exonuclease RecJ	feature:167816188, single-stranded-DNA-specific exonuclease RecJ	1.0656	COG0608	[L]
HIMB114_0374, protein YhdH	feature:167824532, zinc-binding alcohol dehydrogenase	1.999	COG0604	[CR]
HIMB114_0332, taurine dioxygenase	feature:167824576, taurine dioxygenase	3.498	COG2175	[Q]
HIMB114_0326, uptake gene product superfamily protein	feature:167824604, hypothetical protein MED121_17279	1.1876	NA	NA
HIMB114_0287, cell division protein FtsA	feature:167824706, cell division protein FtsA	1.1314	NA	NA
HIMB114_0286, conserved hypothetical protein	feature:167824708, putative cell division protein FtsQ	1.0794	COG1589	[M]

HIMB114_0284, UDP-N-acetylenolpyruvoylglucosamine reductase	feature:167824712, UDP-N-acetylenolpyruvoylglucosamine reductase	1.0368	COG0812	[M]
HIMB114_0279, UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate-D-alanyl-D-alanyl ligase	feature:167824734, UDP-N-acetylmuramoylalanyl-D-glutamate	1.4254	COG0770	[M]
HIMB114_0278, UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2, 6-diaminopimelate ligase	feature:167824736, UDP-N-acetylmuramoyl-tripeptide synthetase	1.2665	COG0769	[M]
HIMB114_0274, cell division protein FtsZ	feature:167824744, cell division protein FtsZ	1.23	COG0206	[D]
HIMB114_0233, putative tetratricopeptide repeat protein	feature:167824846, TPR repeat-containing protein	1.0868	COG0457	[R]
HIMB114_0232, uracil DNA glycosylase superfamily	feature:167824848, phage SPO1 DNA polymerase-related protein	1.3824	COG1573	[L]
HIMB114_0230, putative salicylate 1-monooxygenase	feature:167824856, salicylate 1-monooxygenase	1.3615	COG0654	[HC]
HIMB114_0228, A/G-specific adenine glycosylase	feature:167824858, A/G-specific adenine glycosylase	1.4595	COG1194	[L]
HIMB114_0225, chromosome segregation protein SMC	feature:167824864, smc protein	3.2042	COG1196	[D]
HIMB114_0221, H ⁺ -transporting two-sector ATPase	feature:167824872, H ⁺ -transporting two-sector ATPase (subunit b')	2.91	COG0711	[C]
HIMB114_0888, phytoene synthase	feature:167824882, phytoene synthase	1.0418	COG1562	[I]
HIMB114_0890, putative transmembrane protein	feature:167824886, transmembrane protein	2.247	NA	NA
HIMB114_0218, homoserine kinase	feature:167824890, homoserine kinase	1.0382	COG2334	[R]
HIMB114_0206, 4-hydroxybenzoate polyprenyl transferase	feature:167824920, 4-hydroxybenzoate octaprenyltransferase	1.1922	COG0382	[H]
HIMB114_0200, lipid A biosynthesis lauroyl acyltransferase	feature:167824936, lipid A biosynthesis lauroyl acyltransferase	1.0948	COG1560	[M]
HIMB114_0193, RNA methyltransferase, RsmD family	feature:167824954, N6-adenine-specific methylase	1.5678	NA	NA
HIMB114_0172, UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1 (Enolpyruvate transferase 1) (UDP-N-acetylglucosamine enolpyruvyltransferase 1) (EPT 1)	feature:167825032, UDP-N-acetylglucosamine 1-carboxyvinyltransferase	1.0406	COG0766	[M]
HIMB114_0162, dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	feature:167825052, pyruvate dehydrogenase, E2 component, dihydrolipoamide acetyltransferase	1.7956	COG0508	[C]
HIMB114_0184, methyltransferase FkbM	feature:167825060, methyltransferase FkbM	1.8875	NA	NA
HIMB114_0133, abc transporter ATP-binding protein	feature:167825110, ABC transporter, ATP-binding protein	1.0384	COG1137	[R]
HIMB114_0125, acetyl-coenzyme A carboxylase carboxyl transferase subunit beta (ACCase beta chain)	feature:167825132, acetyl-CoA carboxylase, carboxyl transferase, beta subunit	1.7998	COG0777	[I]
HIMB114_0124, folylpolyglutamate synthase	feature:167825136, folylpolyglutamate synthetase	1.1331	COG0285	[H]

HIMB114_1389, peptide deformylase	feature:167825204, polypeptide deformylase	4.0255	COG0242	[J]
HIMB114_1380, conserved hypothetical protein	feature:167825224, hypothetical protein Bamb_4034	1.5937	COG0730	[R]
HIMB114_1216, sodium:solute transporter family protein	feature:167825230, sodium:solute transporter family protein	3.1746	NA	NA
HIMB114_1376, ATP-dependent helicase	feature:167825238, helicase, putative	1.0371	COG0513	[LKJ]
HIMB114_1346, putative nicotinate-nucleotide adenyltransferase	feature:167825328, nicotinic acid mononucleotide adenyltransferase	1.3789	COG1057	[H]
HIMB114_1324, succinate dehydrogenase, cytochrome b556 subunit	feature:167825382, succinate dehydrogenase cytochrome B560 subunit (s dhC)	1.0625	COG2009	[C]
HIMB114_1299, lipid A 3-O-deacylase (PagL) superfamily	feature:167825444, conserved hypothetical protein	1.4157	COG2855	[S]
HIMB114_0335, putative integral membrane protein	feature:167905354, putative integral membrane protein	1.0439	NA	NA
HIMB114_0899, glycine betaine/L-proline transport system permease protein ProW	feature:167905414, Binding-protein-dependent transport systems inner membrane component	1.5376	NA	NA
HIMB114_0874, pyrroline-5-carboxylate reductase	feature:167905444, delta 1-pyrroline-5-carboxylate reductase	1.0775	NA	NA
HIMB114_1423, twin-arginine translocation pathway signal	feature:167905454, conserved hypothetical secreted protein	1.8724	NA	NA
HIMB114_1422, gluconate 2-dehydrogenase	feature:167905456, putative 2-keto-gluconate dehydrogenase	5.3099	COG2303	[E]
HIMB114_0881, hypothetical protein	feature:167905474, hypothetical protein	1.6622	NA	NA
HIMB114_0894, Sua5/YciO/YrdC family protein	feature:167905484, Sua5/YciO/YrdC family protein	7.2431	COG0009	[J]
HIMB114_0911, cell cycle protein mesj	feature:167905512, cell cycle protein mesj	1.1273	COG0037	[D]
HIMB114_0582, hypothetical protein	feature:167932914, 5-amino-6-(5-phosphoribosylamino)uracil reductase	1.5513	COG0117	[H]
HIMB114_0557, delta-aminolevulinic acid dehydratase	feature:167932924, delta-aminolevulinic acid dehydratase	1.1726	COG0113	[H]
HIMB114_0568, 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase	feature:167932948, 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase	1.0203	COG0801	[H]
HIMB114_0530, CrcB protein	feature:167933002, hypothetical protein Hac_1601	1.1735	COG0239	[D]
HIMB114_0529, ribosomal protein L17	feature:167933004, ribosomal protein L17	1.4755	COG0203	[J]
HIMB114_0487, putative porin	feature:167933090, putative porin	1.3362	NA	NA
HIMB114_0452, probable methyltransferase	feature:167933258, probable methyltransferase	4.664	NA	NA
HIMB114_0447, L-carnitine dehydratase/bile acid-inducible protein F	feature:167933280, L-carnitine dehydrogenase	1.88	COG1804	[C]

Genes with signatures of positive selection in ACE_P1a.1 vs HTCC7211

HTCC7211 gene number & annotation	ACE_P1a.1 gene identifier & annotation	dN/dS*	COG	COG Category
PB7211_759, probable N-acetylmuramoyl-L-alanine amidase YbjR	feature:167713884, N-acetylmuramoyl-L-alanine amidase	1.5912	COG3023	[V]
PB7211_420, cell division protein FtsZ	feature:167713888, cell division protein FtsZ	1.0831	COG0206	[D]
PB7211_1331, UDP-N-acetylmuramoylalanine--D-glutamate ligase	feature:167713904, UDP-N-acetylmuramoylalanine--D-glutamate ligase	1.3265	COG0771	[M]
PB7211_637, DNA uptake lipoprotein	feature:167713926, DNA uptake lipoprotein	1.8721	COG4105	[R]
PB7211_63, penicillin-binding protein 2	feature:167714004, penicillin binding protein transpeptide	1.1065	COG0768	[M]
PB7211_386, branched-chain amino acid ABC transporter, ATP-binding protein	feature:167714068, branched-chain amino acid ABC transporter, ATP-binding protein	1.2909	COG0410	[E]
PB7211_616, Integral membrane protein DUF6	feature:167714094, hypothetical protein SAR11_1344	2.2957	NA	NA
PB7211_366, PQQ enzyme repeat family protein	feature:167714174, PQQ repeat-containing protein	1.443	NA	NA
PB7211_1005, permease	feature:167717520, YadH-like ABC-2-type transport protein	1.6095	COG0842	[V]
PB7211_723, GDP-mannose 4,6-dehydratase	feature:167717612, NAD-dependent epimerase/dehydratase	2.353	NA	NA
PB7211_342, 3-ketoacyl-CoA thiolase (Acetyl-CoA acyltransferase)	feature:167717712, Acetyl-CoA acetyltransferase	1.1539	COG0183	[I]
PB7211_557, formamidopyrimidine-DNA glycosylase	feature:167717754, Formamidopyrimidine-DNA glycosylase	1.67	COG0266	[L]
PB7211_1337, catalase/peroxidase HPI	feature:167717770, catalase/peroxidase HPI	1.0808	COG0376	[P]
PB7211_1321, polysaccharide deacetylase	feature:167717814, polysaccharide deacetylase	2.0428	COG0726	[G]
PB7211_1264, Sulfotransferase domain family protein	feature:167717858, Sulfotransferase domain family protein	1.2428	NA	NA
PB7211_469, putative porin	feature:167717860, putative porin	1.5338	NA	NA
PB7211_1272, riboflavin biosynthesis protein RibF	feature:167719302, Riboflavin biosynthesis protein	2.0297	COG0196	[H]
PB7211_1446, lipid A biosynthesis lauroyl acyltransferase	feature:167719320, lipid A biosynthesis lauroyl acyltransferase	1.0693	COG1560	[M]
PB7211_1330, tetraacyldisaccharide 4'-kinase	feature:167719322, tetraacyldisaccharide 4'-kinase	1.0645	COG1663	[M]
PB7211_868, salicylate 1-monooxygenase	feature:167719412, salicylate 1-monooxygenase	1.2038	NA	NA
PB7211_1100, prolipoprotein diacylglycerol transferase	feature:167719450, prolipoprotein diacylglycerol transferase	2.8813	COG0682	[M]
PB7211_1203, O-Antigen Polymerase family protein	feature:167719510, O-antigen polymerase	1.4019	NA	NA

PB7211_811, Diacylglycerol kinase catalytic domain protein, putative	feature:167719684, Diacylglycerol kinase catalytic domain protein, putative	2.2579	NA	NA
PB7211_970, TctA protein	feature:167719730, integral membrane protein, putative	1.0108	COG3333	[S]
PB7211_965, putative tricarboxylic transport TctC	feature:167719734, exported protein, putative	1.7508	COG3181	[S]
PB7211_638, glycine/D-amino acid oxidase deaminating	feature:167719742, hypothetical protein MED121_20906	2.0999	COG0665	[E]
PB7211_1215, sugar fermentation stimulation protein	feature:167719816, sugar fermentation stimulation protein	3.8764	COG1489	[R]
PB7211_580, pantoate--beta-alanine ligase	feature:167719910, pantoate-beta-alanine ligase	1.1087	COG0414	[H]
PB7211_440, conserved hypothetical protein	feature:167720014, hypothetical protein SAR11_1038	1.0161	COG5452	[S]
PB7211_1401, glycosyl transferase, group 2 family	feature:167720050, glycosyl transferase, group 2 family	2.2803	COG1216	[R]
PB7211_49, glycosyl transferase family 8	feature:167774855, glycosyl transferase family 8	1.1375	NA	NA
PB7211_262, methyltransferase	feature:167818944, methyltransferase	1.0035	NA	NA
PB7211_986, ADP-ribosylglycohydrolase	feature:167818946, ADP-ribosylglycohydrolase	2.5882	NA	NA
PB7211_972, bacterial extracellular solute-binding protein, family 7	feature:167818958, TRAP dicarboxylate transporter- DctP subunit	1.2802	COG1638	[G]
PB7211_1440, probable transcription regulatory protein	feature:167819026, hypothetical protein SAR11_1248	1.3527	NA	NA
PB7211_644, ABC proline/glycine betaine transporter, periplasmic substrate-binding protein	feature:167819050, glycine betaine/L-proline ABC transporter,substrate-binding periplasmic protein	1.2334	COG2113	[E]
PB7211_88, ABC transporter, membrane spanning protein (glycine betaine)	feature:167819052, PA5095	2.5	COG4176	[E]
PB7211_552, phosphoenolpyruvate carboxylase	feature:167819194, putative phosphoenolpyruvate carboxylase	1.1113	COG2352	[C]
PB7211_123, hypothetical protein	feature:167819374, Glycosyltransferase family 9 (heptosyltransferase)	1.6466	COG0859	[M]
PB7211_1161, membrane protein	feature:167819396, membrane protein	1.4594	COG2364	[S]
PB7211_132, phosphatidate cytidyltransferase	feature:167888844, phosphatidate cytidyltransferase	1.2456	COG0575	[I]
PB7211_625, UDP-3-O-(3-hydroxymyristoyl) glucosamine N-acyltransferase	feature:167888854, UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	1.0432	COG1044	[M]
PB7211_158, drug/metabolite transporter	feature:167907248, DMT family permease	1.0154	COG0697	[GER]
PB7211_244, phosphotransferase enzyme family protein	feature:167907268, phosphotransferase enzyme family protein	1.0189	COG3178	[R]
PB7211_1136, two-component sensor histidine kinase	feature:167907274, two-component sensor histidine kinase	1.0954	COG0642	[T]

PB7211_401, acid tolerance regulatory protein actr	feature:167907284, response regulator	1.2681	COG4567	[TK]
PB7211_1162, D-lactate dehydrogenase (cytochrome)	feature:167907448, (S)-2-hydroxy-acid oxidase	1.3169	COG0277	[C]
PB7211_1296, homocysteine S-methyltransferase	feature:167907494, betaine-homocysteine methyltransferase	1.0481	COG0646	[E]
PB7211_114, bifunctional purine biosynthesis protein	feature:167907536, bifunctional purine biosynthesis protein	1.5277	COG0138	[F]
PB7211_999, putative Sulfotransferase domain protein	feature:167907540, putative Sulfotransferase domain protein	1.2712	NA	NA

Genes with signatures of positive selection in IMCC9063 vs HIMB114

HIMB114 gene number & annotation	IMCC9063 gene identifier & annotation	dN/dS*	COG	COG Category
HIMB114_0965, penicillin-binding protein 2	SAR11G3_00024, penicillin-binding protein 2 (PBP-2)	1.0483	COG0768	[M]
HIMB114_0990, putative uNdecaprenyl-phosphate-alpha-n-acetylglucosaminyltransferase	SAR11G3_00039, undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase	1.3074	NA	NA
HIMB114_0991, putative O-Antigen Polymerase	SAR11G3_00040, O-Antigen Polymerase	1.3288	NA	NA
HIMB114_0993, putative glycosyltransferase	SAR11G3_00070, glycosyl transferase family protein, putative	3.0041	COG0463	[M]
HIMB114_0989, nucleotidyl transferase	SAR11G3_00113, D-glycero-D-mannoheptose 1-phosphate guanosyltransferase	1.6271	COG1208	[MJ]
HIMB114_1044, mutator MutT protein	SAR11G3_00148, 5-methyl-dCTP pyrophosphohydrolase	1.1454	COG0494	[LR]
HIMB114_1046, acetyl-CoA carboxylase, carboxyl transferase, alpha subunit	SAR11G3_00151, acetyl-coenzyme A carboxyl transferase alpha chain	1.0629	COG0825	[I]
HIMB114_1048, hypothetical protein	SAR11G3_00153, hypothetical protein	1.5092	NA	NA
HIMB114_1096, hypothetical protein	SAR11G3_00203, hypothetical protein	1.6464	NA	NA
HIMB114_0761, type 4 fimbrial assembly protein PilB	SAR11G3_00220, type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB	1.6743	COG2804	[NU]
HIMB114_0762, putative outer membrane protein	SAR11G3_00221, hypothetical protein	1.8769	COG4796	[U]
HIMB114_0883, conserved hypothetical protein	SAR11G3_00226, hypothetical protein	1.2802	NA	NA
HIMB114_0765, twitching motility protein	SAR11G3_00237, twitching motility protein PilT	1.242	COG2805	[NU]
HIMB114_1138, putative membrane-bound pili assembly protein Fimc	SAR11G3_00284, putative membrane-bound pili assembly protein Fimc	1.3821	NA	NA
HIMB114_0187, omega-amino acid--pyruvate aminotransferase (Omega-APT)(Beta-alanine--pyruvate aminotransferase)	SAR11G3_00289, omega-amino acid--pyruvate aminotransferase	1.3953	COG0161	[H]
HIMB114_1415, major facilitator superfamily protein	SAR11G3_00290, putative MFS transporter (Major Facilitator Superfamily)	1.0301	COG0477	[GEPR]

HIMB114_1150, conserved hypothetical protein	SAR11G3_00297, thiosulfate sulfurtransferase, rhodanese	1.0576	COG2897	[P]
HIMB114_1163, ABC transporter, ATP-binding protein	SAR11G3_00311, ABC transporter, ATP-binding protein	1.7968	COG1131	[V]
HIMB114_1171, putative DNA replication and repair protein RecF	SAR11G3_00320, DNA recombination and repair protein RecF	1.1476	NA	NA
HIMB114_1178, phosphopantothencysteine decarboxylase/phosphopantothencysteine ligase	SAR11G3_00328, phosphopantothencysteine decarboxylase / Phosphopantothencysteine synthetase	1.811	COG0452	[H]
HIMB114_0722, apolipoprotein N-acyltransferase	SAR11G3_00345, apolipoprotein N-acyltransferase	1.0882	NA	NA
HIMB114_1194, conserved hypothetical protein	SAR11G3_00347, metal-dependent hydrolase	1.4376	COG0319	[R]
HIMB114_0558, uroporphyrinogen decarboxylase (URO-D) (UPD)	SAR11G3_00394, uroporphyrinogen III decarboxylase	2.4278	COG0407	[H]
HIMB114_1246, probable Smr	SAR11G3_00403, putative Smr (Small MutS Related) protein	1.0497	NA	NA
HIMB114_1238, putative outer membrane lipoprotein carrier protein	SAR11G3_00413, hypothetical protein	3.2924	NA	NA
HIMB114_1291, hypothetical protein	SAR11G3_00439, hypothetical protein	1.358	NA	NA
HIMB114_1295, cytochrome c-type biogenesis protein CcmF	SAR11G3_00444, cytochrome c heme lyase subunit CcmF	1.0139	COG1138	[O]
HIMB114_1299, lipid A 3-O-deacylase (PagL) superfamily	SAR11G3_00447, hypothetical protein	1.4959	NA	NA
HIMB114_1303, FAD linked oxidase domain protein	SAR11G3_00451, glycolate dehydrogenase, FAD-binding subunit GlcE	1.295	COG0277	[C]
HIMB114_1329, aspartate-semialdehyde dehydrogenase	SAR11G3_00471, aspartate-semialdehyde dehydrogenase	1.6153	COG0136	[E]
HIMB114_1333, oxoglutarate dehydrogenase (succinyl-transferring), E1 component	SAR11G3_00480, 2-oxoglutarate dehydrogenase E1 component	1.0675	COG0567	[C]
HIMB114_0986, transaldolase	SAR11G3_00483, transaldolase	1.0517	COG0176	[G]
HIMB114_1346, putative nicotinate-nucleotide adenyltransferase	SAR11G3_00495, putative nicotinate-nucleotide adenyltransferase	1.0708	COG1057	[H]
HIMB114_1356, histidinol-phosphate aminotransferase (Imidazole acetol-phosphate transaminase)	SAR11G3_00505, histidinol-phosphate aminotransferase	3.8665	COG0079	[E]
HIMB114_1373, transcriptional regulator, CarD family	SAR11G3_00522, carD-like transcriptional regulator	1.0069	COG1329	[K]
HIMB114_1376, ATP-dependent helicase	SAR11G3_00525, ATP-dependent DNA helicase	1.0874	COG0513	[LKJ]
HIMB114_1383, 3-deoxy-manno-octulosonate cytidyltransferase (CMP-KDOsynthetase) (CMP-2-keto-3-deoxyoctulosonic acid synthetase) (CKS)	SAR11G3_00534, 3-deoxy-manno-octulosonate cytidyltransferase	1.0777	COG1212	[M]
HIMB114_1398, 60 kDa inner membrane insertion protein	SAR11G3_00546, inner membrane protein translocase component YidC, long form	1.0934	COG0706	[U]

HIMB114_0124, folylpolyglutamate synthase	SAR11G3_00584, dihydrofolate synthase / Folylpolyglutamate synthase	1.1114	COG0285	[H]
HIMB114_0125, acetyl-coenzyme A carboxylase carboxyl transferase subunit beta (ACCcase beta chain)	SAR11G3_00585, acetyl-coenzyme A carboxyl transferase beta chain	1.0386	COG0777	[I]
HIMB114_0158, putative glycosyl transferase	SAR11G3_00617, glycosyl transferase, group 2 family protein	1.5765	NA	NA
HIMB114_0161, beta-lactamase	SAR11G3_00619, esterase	1.0828	COG1680	[V]
HIMB114_0163, pyruvate dehydrogenase (acetyl-transferring), homodimeric type	SAR11G3_00622, pyruvate dehydrogenase E1 component	1.1217	COG2609	[C]
HIMB114_0176, septum formation protein Maf	SAR11G3_00635, septum formation protein Maf	1.3608	COG0424	[D]
HIMB114_0193, RNA methyltransferase, RsmD family	SAR11G3_00651, ribosomal RNA small subunit methyltransferase D	1.5166	NA	NA
HIMB114_0200, lipid A biosynthesis lauroyl acyltransferase	SAR11G3_00659, lipid A biosynthesis lauroyl acyltransferase	1.1711	COG1560	[M]
HIMB114_0202, tetraacyldisaccharide 4'-kinase	SAR11G3_00660, 3-deoxy-D-manno-octulosonic-acid transferase	1.2561	COG1663	[M]
HIMB114_0218, homoserine kinase	SAR11G3_00686, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase	1.051	COG2334	[R]
HIMB114_0890, putative transmembrane protein	SAR11G3_00688, putative transmembrane protein	1.4711	NA	NA
HIMB114_0225, chromosome segregation protein SMC	SAR11G3_00699, chromosome partition protein smc	3.2042	COG1196	[D]
HIMB114_0228, A/G-specific adenine glycosylase	SAR11G3_00703, A/G-specific adenine glycosylase	1.4059	COG1194	[L]
HIMB114_0230, putative salicylate 1-monooxygenase	SAR11G3_00704, salicylate hydroxylase	1.3615	COG0654	[HC]
HIMB114_0232, uracil DNA glycosylase superfamily	SAR11G3_00707, uracil-DNA glycosylase, family 4	1.0421	COG1573	[L]
HIMB114_0233, putative tetratricopeptide repeat protein	SAR11G3_00708, conserved hypothetical protein; putative TPR repeat domain	1.2818	COG0457	[R]
HIMB114_0234, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK)(4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase)	SAR11G3_00709, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	1.1503	NA	NA
HIMB114_0247, putative cyclopropane-fatty-acyl-phospholipid synthase	SAR11G3_00723, conserved hypothetical protein	2.1235	NA	NA
HIMB114_0260, conserved hypothetical protein	SAR11G3_00737, DUF1022 domain-containing protein	1.1488	COG3660	[M]
HIMB114_0265, DNA primase	SAR11G3_00742, DNA primase	1.1256	COG0358	[L]
HIMB114_0274, cell division protein FtsZ	SAR11G3_00749, cell division protein FtsZ	1.1549	COG0206	[D]
HIMB114_0278, UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2, 6-diaminopimelate ligase	SAR11G3_00753, UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase	1.284	COG0769	[M]
HIMB114_0286, conserved hypothetical protein	SAR11G3_00764, hypothetical protein	1.2103	NA	NA

HIMB114_0287, cell division protein FtsA	SAR11G3_00765, cell division protein FtsA	1.1601	NA	NA
HIMB114_0293, hypothetical protein	SAR11G3_00774, hypothetical protein	1.5466	NA	NA
HIMB114_0317, putative transcriptional regulator, IclR family	SAR11G3_00798, transcriptional regulator, IclR family	3.4672	NA	NA
HIMB114_0384, putative phosphoserine phosphatase	SAR11G3_00848, phosphoserine phosphatase	1.1134	COG0560	[E]
HIMB114_0416, transporter major facilitator superfamily MFS_1	SAR11G3_00876, putative membrane permease of unknown function	1.0186	NA	NA
HIMB114_0833, alanyl-tRNA synthetase	SAR11G3_00906, alanyl-tRNA synthetase	1.1745	COG0013	[J]
HIMB114_0806, hypothetical protein	SAR11G3_00939, outer membrane protein Imp, required for envelope biogenesis / Organic solvent tolerance protein precursor	1.3512	COG1452	[M]
HIMB114_0805, putative permease, YjgP/YjgQ family	SAR11G3_00940, putative permease YjgP/YjgQ family protein	1.1595	NA	NA
HIMB114_0799, hypothetical protein	SAR11G3_00947, putative membrane protein	1.5301	NA	NA
HIMB114_0797, probable ATPase involved in DNA replication initiation	SAR11G3_00948, chromosomal replication initiator protein DnaA	1.1275	NA	NA
HIMB114_0114, conserved hypothetical protein	SAR11G3_00977, hypothetical protein	1.5491	NA	NA
HIMB114_0858, probable nicotinate-nucleotide pyrophosphorylase [carboxylating]	SAR11G3_00996, quinolinate phosphoribosyltransferase/decarboxylase	1.0814	COG0157	[H]
HIMB114_1100, immunogenic protein	SAR11G3_01000, TRAP transporter solute receptor, TAXI family precursor	1.0421	COG2358	[R]
HIMB114_0714, NADH dehydrogenase (quinone), g subunit	SAR11G3_01013, NADH-ubiquinone oxidoreductase chain G	1.6723	COG1034	[C]
HIMB114_0701, ABC transporter	SAR11G3_01029, lipoprotein releasing system transmembrane protein LolC	2.0175	COG4591	[M]
HIMB114_0699, DNA polymerase III subunit alpha	SAR11G3_01031, DNA polymerase III alpha subunit	1.3333	COG0587	[L]
HIMB114_0694, phosphatidate cytidyltransferase	SAR11G3_01036, phosphatidate cytidyltransferase	1.0445	COG0575	[I]
HIMB114_0693, 1-deoxy-D-xylulose 5-phosphate reductoisomerase	SAR11G3_01038, 1-deoxy-D-xylulose 5-phosphate reductoisomerase	2.1952	COG0743	[I]
HIMB114_0692, outer membrane protein assembly complex, YaeT protein	SAR11G3_01039, outer membrane protein assembly factor YaeT precursor	1.4305	NA	NA
HIMB114_0688, conserved hypothetical protein	SAR11G3_01045, hypothetical protein	6.5561	COG3494	[S]
HIMB114_0687, lipid-A-disaccharide synthase	SAR11G3_01046, lipid-A-disaccharide synthase	1.2279	COG0763	[M]
HIMB114_0680, anthranilate synthase component I	SAR11G3_01055, hypothetical SurA-like protein	1.0698	COG0147	[EH]
HIMB114_0679, hypothetical protein	SAR11G3_01056, hypothetical SurA-like protein	1.3423	NA	NA
HIMB114_0678, triose-phosphate isomerase	SAR11G3_01058, triosephosphate isomerase	3.1289	COG0149	[G]

HIMB114_0664, putative membrane protein	SAR11G3_01075, trkH-like cation transport protein	1.7126	NA	NA
HIMB114_0659, conserved hypothetical protein	SAR11G3_01083, conserved hypothetical protein	1.177	NA	NA
HIMB114_0653, valyl-tRNA synthetase	SAR11G3_01089, valyl-tRNA synthetase	1.0367	COG0525	[J]
HIMB114_0636, putative lipoprotein	SAR11G3_01110, hypothetical protein	1.5106		
HIMB114_0616, protein-export membrane protein SecD	SAR11G3_01130, protein-export membrane protein SecD	1.3136	COG0342	[U]
HIMB114_0630, glycosyl transferase group 1	SAR11G3_01144, putative hexosyltransferase	1.2026	COG0438	[M]
HIMB114_0633, putative DNA polymerase III	SAR11G3_01147, DNA polymerase III delta prime subunit	1.0547	COG0470	[L]
HIMB114_0583, riboflavin synthase, alpha subunit	SAR11G3_01155, riboflavin synthase alpha chain	1.0005	COG0307	[H]
HIMB114_0577, DNA gyrase subunit A	SAR11G3_01164, topoisomerase IV subunit A	1.2291	COG0188	[L]
HIMB114_0571, pyridoxal phosphate biosynthesis protein PdxJ	SAR11G3_01171, pyridoxine 5'-phosphate synthase	1.2895	COG0854	[H]
HIMB114_0564, putative soluble lytic murein transglycosylase	SAR11G3_01177, soluble lytic murein transglycosylase precursor	1.6949	COG0741	[M]
HIMB114_0533, pantoate--beta-alanine ligase	SAR11G3_01195, pantoate--beta-alanine ligase	1.0448	COG0414	[H]
HIMB114_0532, ribosomal large subunit pseudouridine synthase C	SAR11G3_01198, crcB-like protein	1.4439	COG0564	[J]
HIMB114_0487, putative porin	SAR11G3_01240, putative porin	1.252	NA	NA
HIMB114_1132, histone deacetylase family protein	SAR11G3_01303, histone deacetylase family protein	1.4163	COG0123	[BQ]
HIMB114_1467, conserved hypothetical protein	SAR11G3_01317, hypothetical protein	1.0859	NA	NA
HIMB114_0776, SUF system FeS assembly protein, NifU family	SAR11G3_01324, putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2	1.4051	COG0822	[C]
HIMB114_0778, FeS assembly protein SufD	SAR11G3_01326, iron-sulfur cluster assembly protein SufD	1.2291	COG0719	[O]
HIMB114_0781, anhydro-N-acetylmuramic acid kinase	SAR11G3_01330, anhydro-N-acetylmuramic acid kinase	1.0901	COG2377	[O]
HIMB114_0784, conserved hypothetical protein	SAR11G3_01332, hypothetical protein	1.5917	NA	NA
HIMB114_0785, peptide chain release factor 2	SAR11G3_01333, peptide chain release factor 2	1.2872	COG1186	[J]
HIMB114_0786, penicillin-binding protein 1A	SAR11G3_01334, multimodular transpeptidase-transglycosylase	1.3044	COG5009	[M]
HIMB114_0787, ribonuclease E	SAR11G3_01335, cytoplasmic axial filament protein CafA and Ribonuclease G	1.1753	COG1530	[J]
HIMB114_0899, glycine betaine/L-proline transport system permease protein ProW	SAR11G3_01363, L-proline glycine betaine ABC transport system permease protein ProW	1.5242	COG4176	[E]
HIMB114_0356, ferric uptake regulation protein	SAR11G3_01369, zinc uptake regulation protein ZUR	1.0975	COG0735	[P]
HIMB114_0874, pyrroline-5-carboxylate reductase	SAR11G3_01381, pyrroline-5-carboxylate reductase	2.2476	COG0345	[E]

HIMB114_0881, hypothetical protein	SAR11G3_01394, hypothetical protein	1.4347	NA	NA
HIMB114_0929, thiamine-phosphate pyrophosphorylase (TMPpyrophosphorylase) (TMP-PPase) (Thiamine-phosphate synthase)	SAR11G3_01434, thiamin-phosphate pyrophosphorylase	1.0194	COG0352	[H]
HIMB114_0935, AIR carboxylase	SAR11G3_01441, phosphoribosylaminoimidazole carboxylase catalytic subunit	1.0747	COG0041	[F]
HIMB114_0938, hypothetical protein	SAR11G3_01444, hypothetical protein	1.0427	NA	NA
HIMB114_0939, putative rod shape-determining protein MreC	SAR11G3_01445, rod shape-determining protein MreC	1.4622	NA	NA