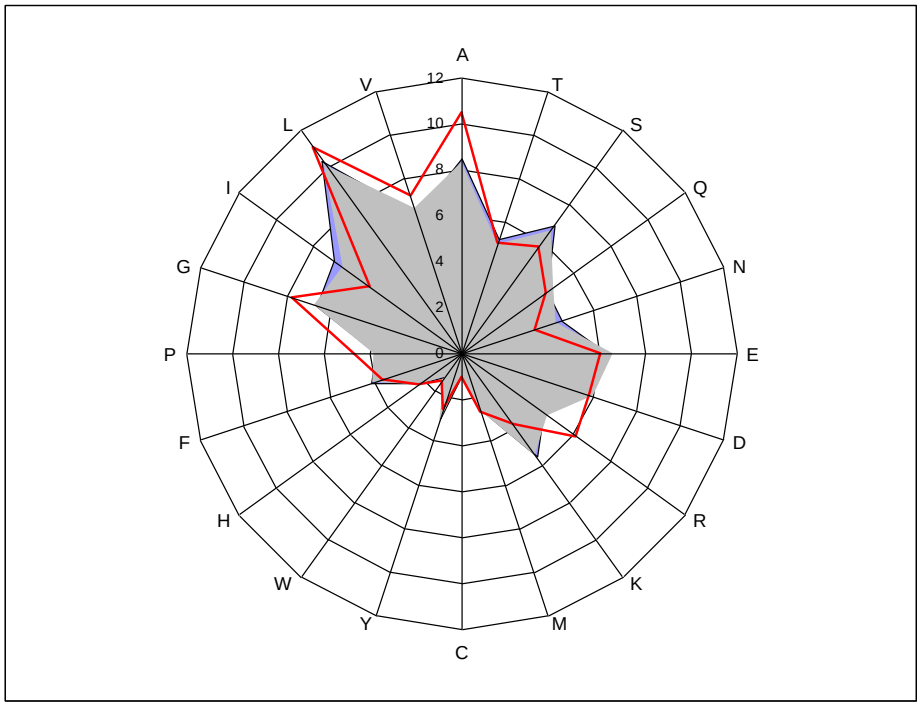
Supplementary Information

Kube et al.

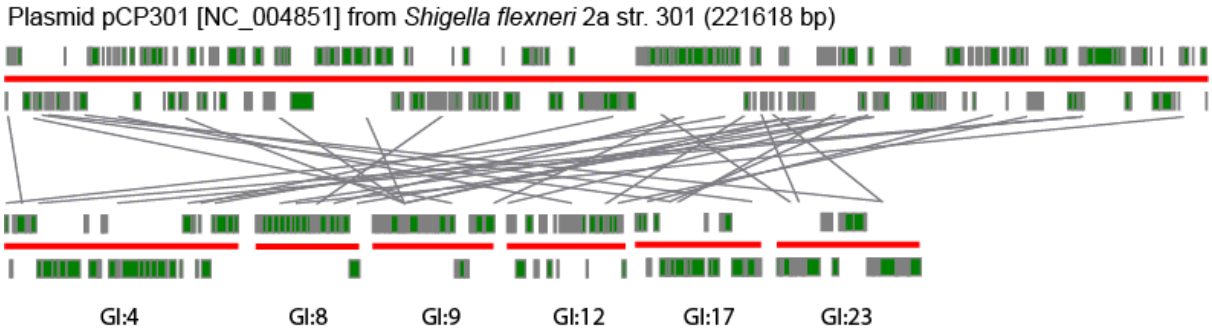
Genome sequence and functional genomic analysis of the oil-degrading bacterium *Oleispira antarctica*

CONTENTS

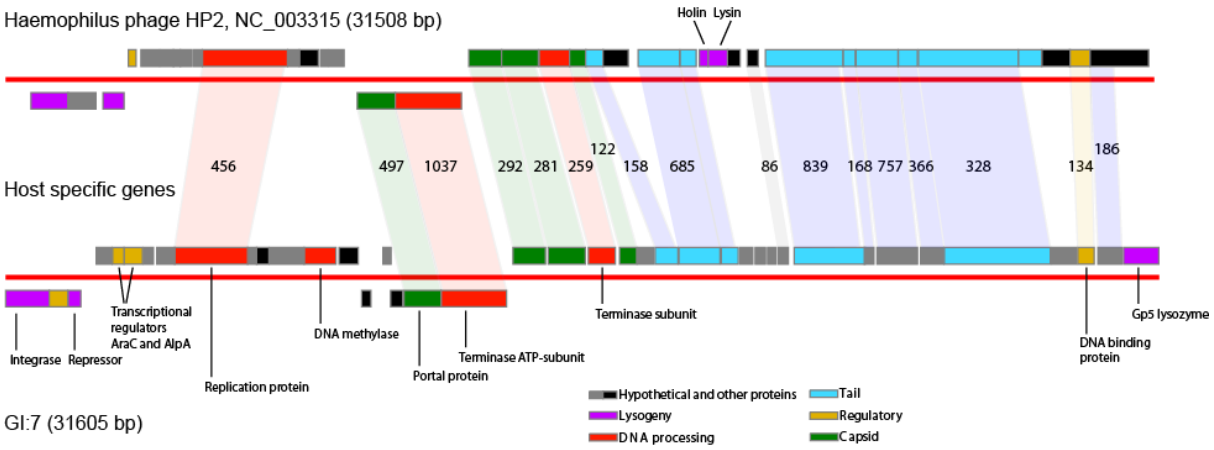
SUPPLEMENTARY FIGURES



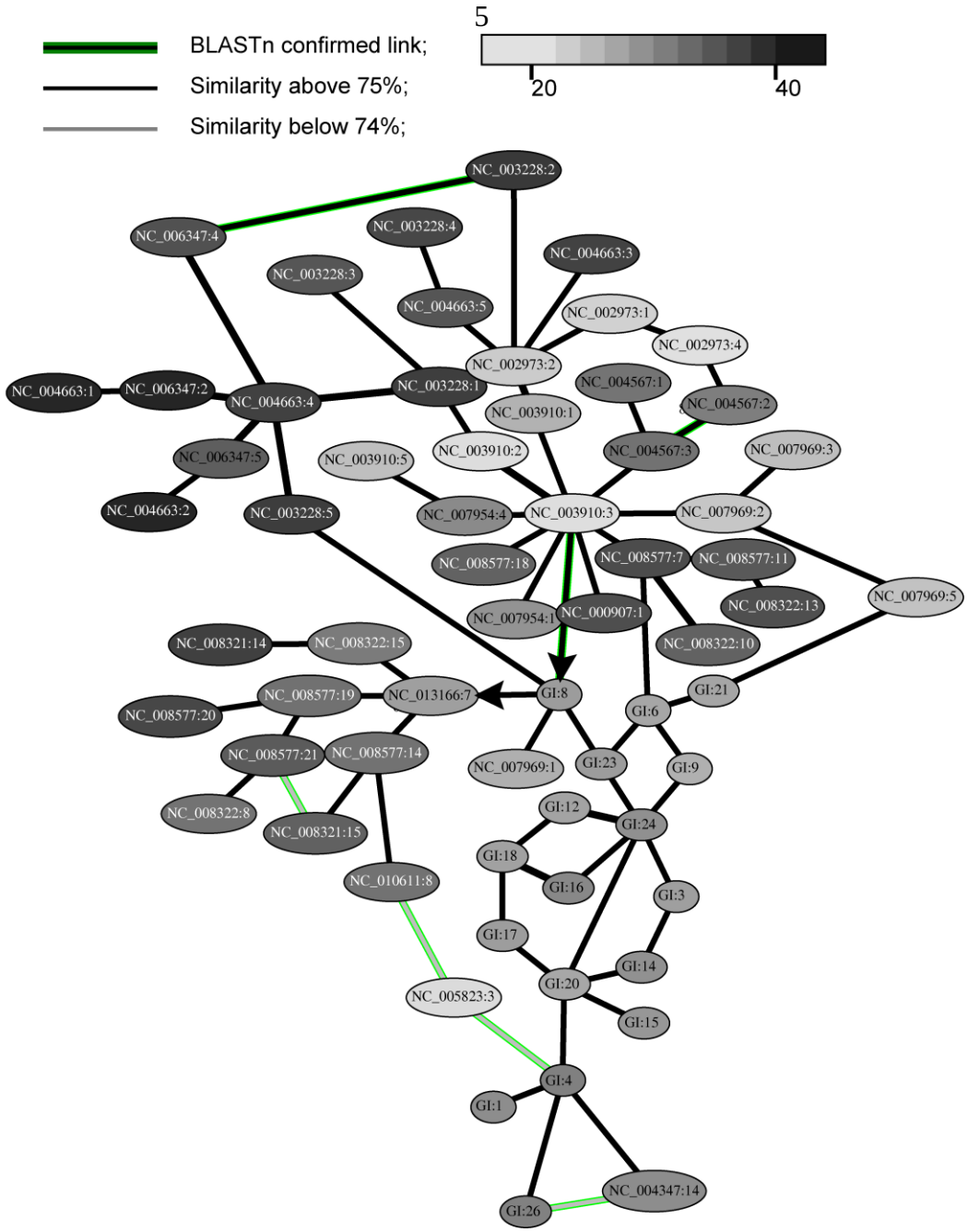
Supplementary Figure S1 Amino acid mol % composition in proteins of *Oleispira antarctica* (blue plot), *Bermanella marisrubri* (grey plot) and *Alcanivorax borkumensis* (red solid line-plot). Two former organisms have been isolated from very distinct environments and exhibit very distinct temperature profiles for growth. However, the AA composition plots of *Bermanella marisrubri,* a Red Sea bacterium, and *O. antarctica*, the Antarctic isolate, are almost identical, exhibiting differences within 0.5 per cent, which suggests the overall AA compositional bias could hardly be a considered as a valid factor for cold resistance of *O. antarctica.*



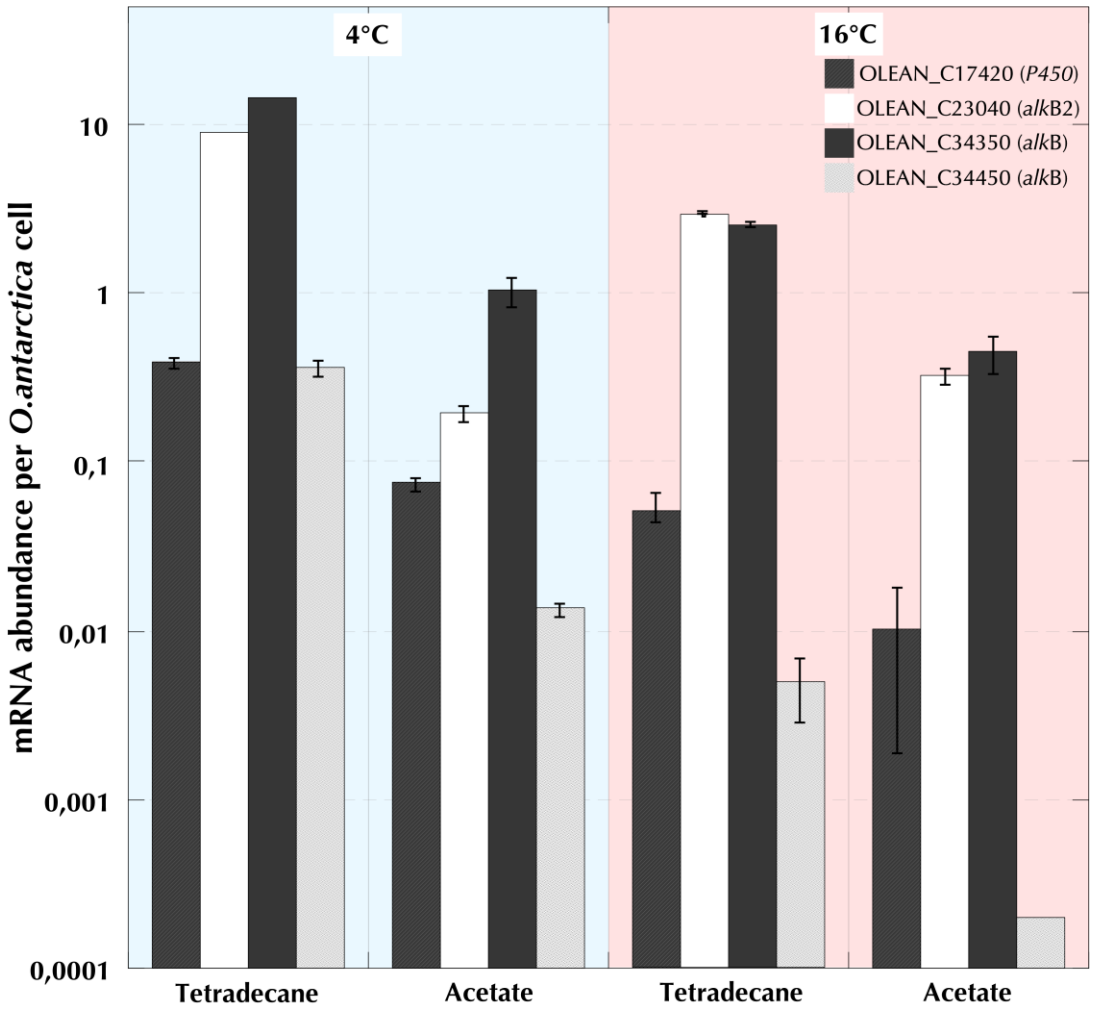
Supplementary Figure S2 Regions of DNA sequence similarity in six genomic islands of *Oleispira antarctica* RB-8 and the plasmid pCP301 from *Shigella flexneri* 2a strain 301



Supplementary Figure S3 Amino acid sequence similarity between proteins encoded by genes of the prophage GI:7 and the Haemophilus phage HP2. Shaded quadrilaterals indicate arrangement of homologous genes in phage and in the genomic island (GI:7) of *O. antarctica*, general protein functions are color-coded.

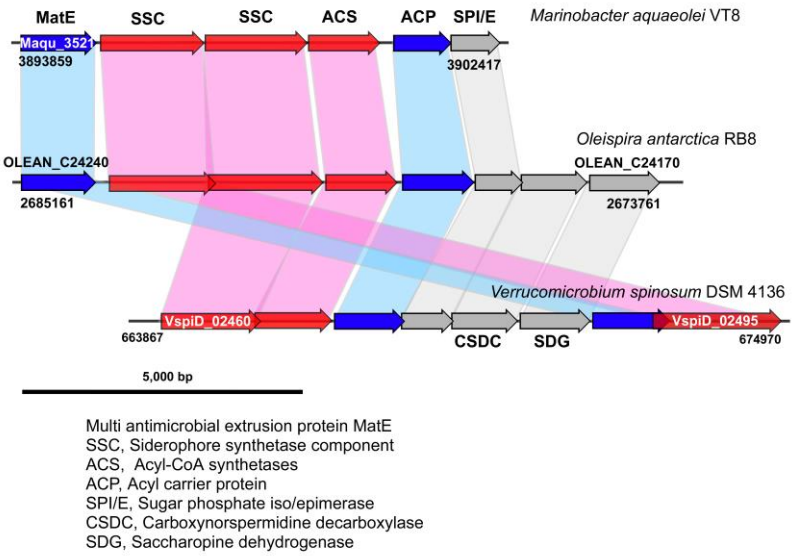


Supplementary Figure S4 Grouping of genomic islands by similarity of tetranucleotide usage patterns. Genomic islands of *Oleispira antarctica* RB-8T are marked as in **Table S1**. Other microorganisms are represented by NC accession numbers: NC_000907 – *Haemophilus influenzae* Rd KW20; NC_002973 – *Listeria monocytogenes* serotype 4b str. F2365; NC_003228 – *Bacteroides fragilis* NCTC 9343; NC_003910 – *Colwellia psychrerythraea* 34H; NC_004347 – *Shewanella oneidensis* MR-1; NC_004567 – *Lactobacillus plantarum* WCFS1; NC_004663 – *Bacteroides thetaiotaomicron* VPI-5482; NC_005823 – *Leptospira interrogans* serovar Copenhageni str. Fiocruz L1-130; NC_006347 – *B. fragilis* YCH46; NC_007954 – *Shewanella denitrificans* OS217; NC_007969 – *Psychrobacter cryohalolentis* K5; NC_008321 – *Shewanella* sp. MR-4; NC_008322 – *Shewanella* sp. MR-7; NC_008577 – *Shewanella* sp. ANA-3; NC_010611 – *Acinetobacter baumannii* ACICU; NC_013166 – *Kangiella koreensis* DSM 16069. Numbers of genomic islands are the same as in www.bi.up.ac.za/SeqWord/sniffer/gidb/index.php. Putative directions of genomic island distribution are depicted by arrows.

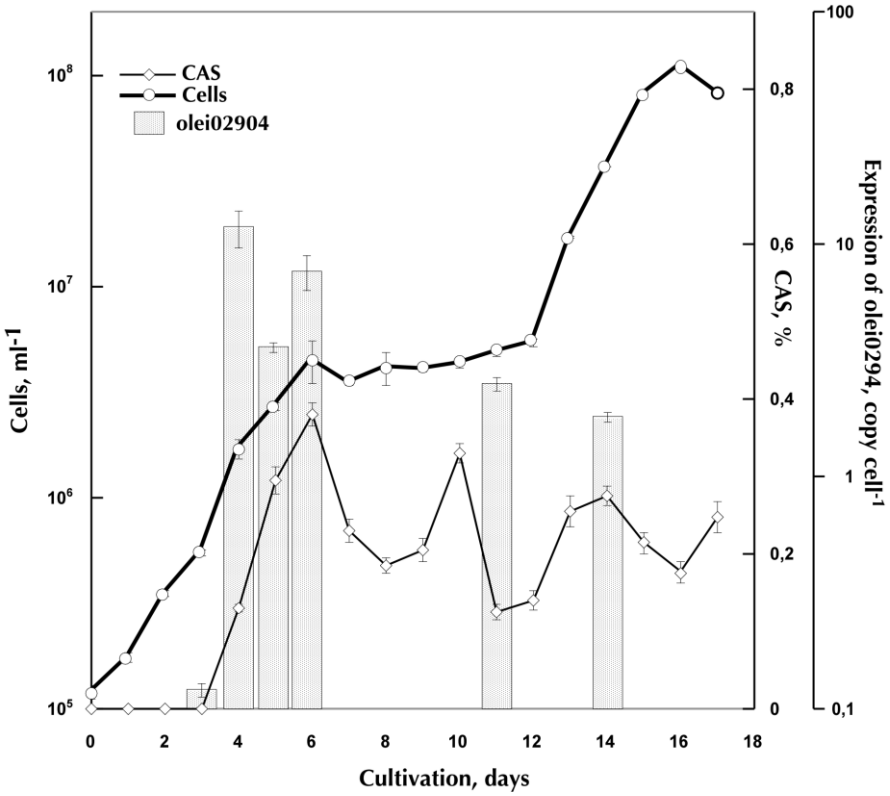


Supplementary Figure S5 Q-RT-PCR analysis of alkane hydrolases and cytochromes P450 expression in *O. antarctica* cells grown on n-tetradecane and acetate at different temperatures. The total number of mRNA copies was determined by Q-RT-PCR and further normalized against cell numbers (assuming one gene copy per cell, after quantification of chromosome copies in DNA yielded after RNA/DNA extraction). All data are mean values of triplicate measurements, standard deviations are shown by vertical bars.

A



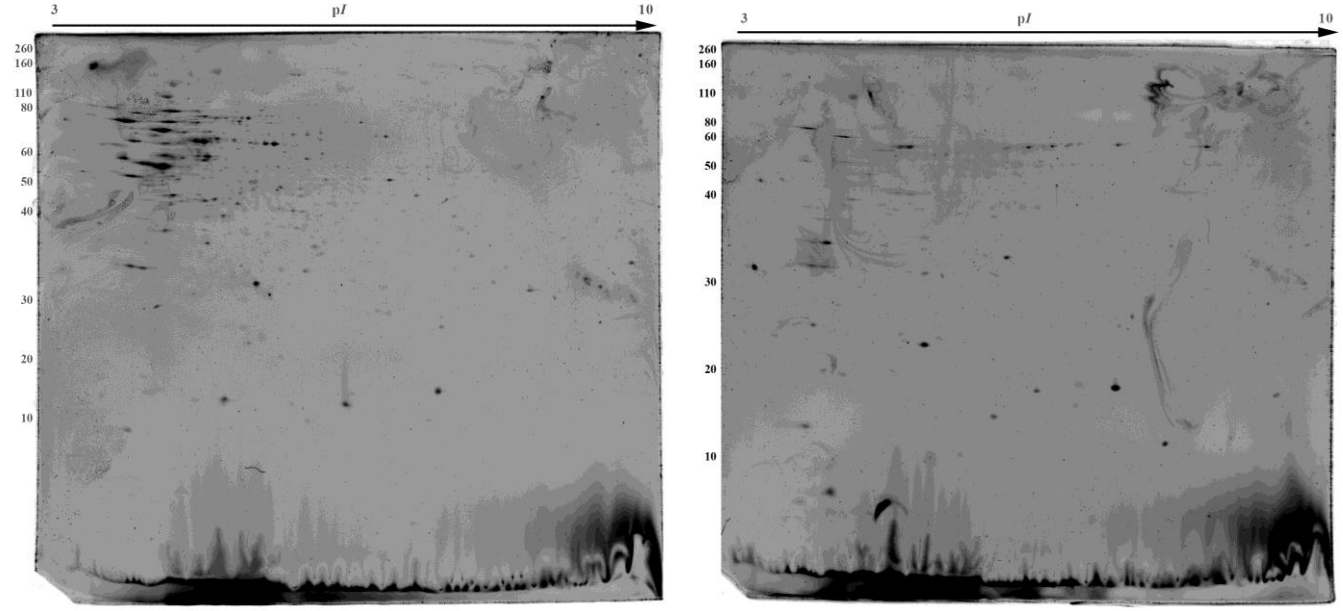
B



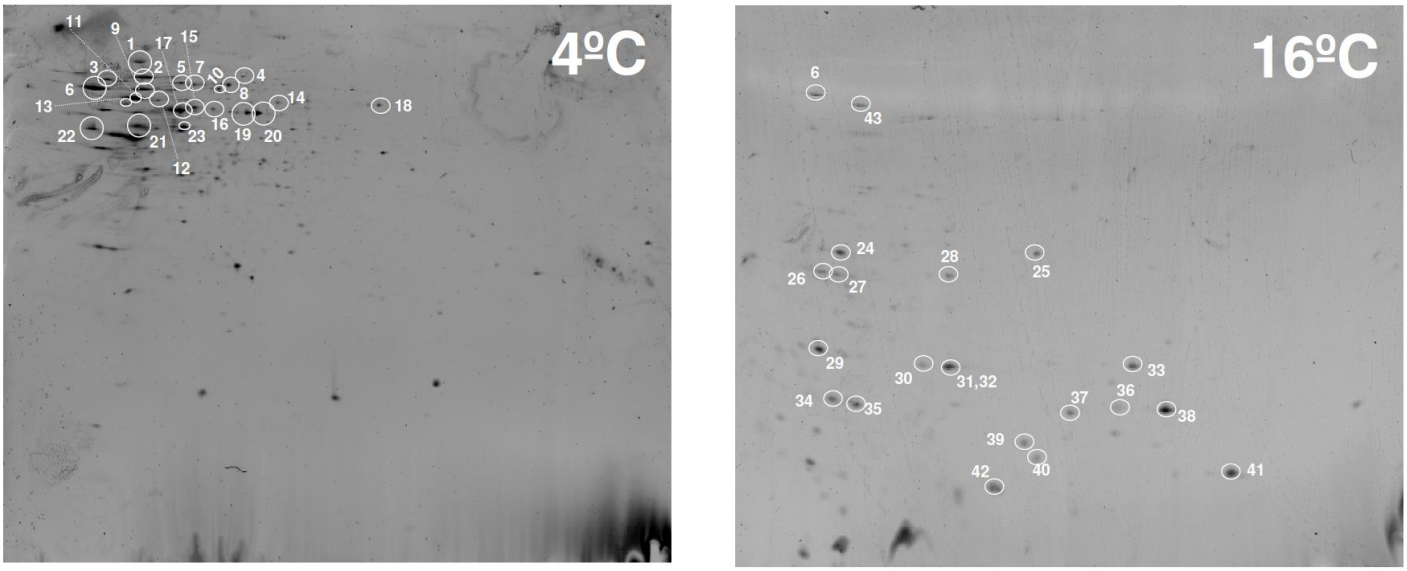
Supplementary Figure S6 Siderophores in *O. antarctica.* (A) Conserved siderophore biosynthesis clusters in genomes of *Oleispira antarctica* RB-8, *Marinobacter aquaolei* VT8 (and a freshwater bacterium *Verrucomicrobium spinosum.* Shaded quadrilaterals indicate gene arrangements in chromosomal fragments. SSC, siderophore synthetase component; ACS, acyl-CoA synthetase; ACP, acyl carrier protein; SPI/E, sugar phosphate iso/epimerase; CSDC, carboxynorspermidine decarboxylase; SDG, saccharopine dehydrogenase. **(B) Siderophore production by *O. antarctica*** growing in the Fe-limited medium ONR7a with *n*-tetradecane at 4ºC and 16ºC as per chrome Azurol S reagent (CAS) assay according to Schwyn & Neilands[68] and enumeration of transcripts of *acs* gene (OLEAN_C24210, grey bars). "CAS, %"is calculated from ratio [CAS-SPL/CAS] x 100 where 'CAS' is the absorbance of iron-chelated CAS reagent at

OD_{630}, ‘SPL’ is that of the medium sample containing siderophores at OD_{630}. Standard deviations (SD) from average values measured from biological triplicates are shown as vertical bars.

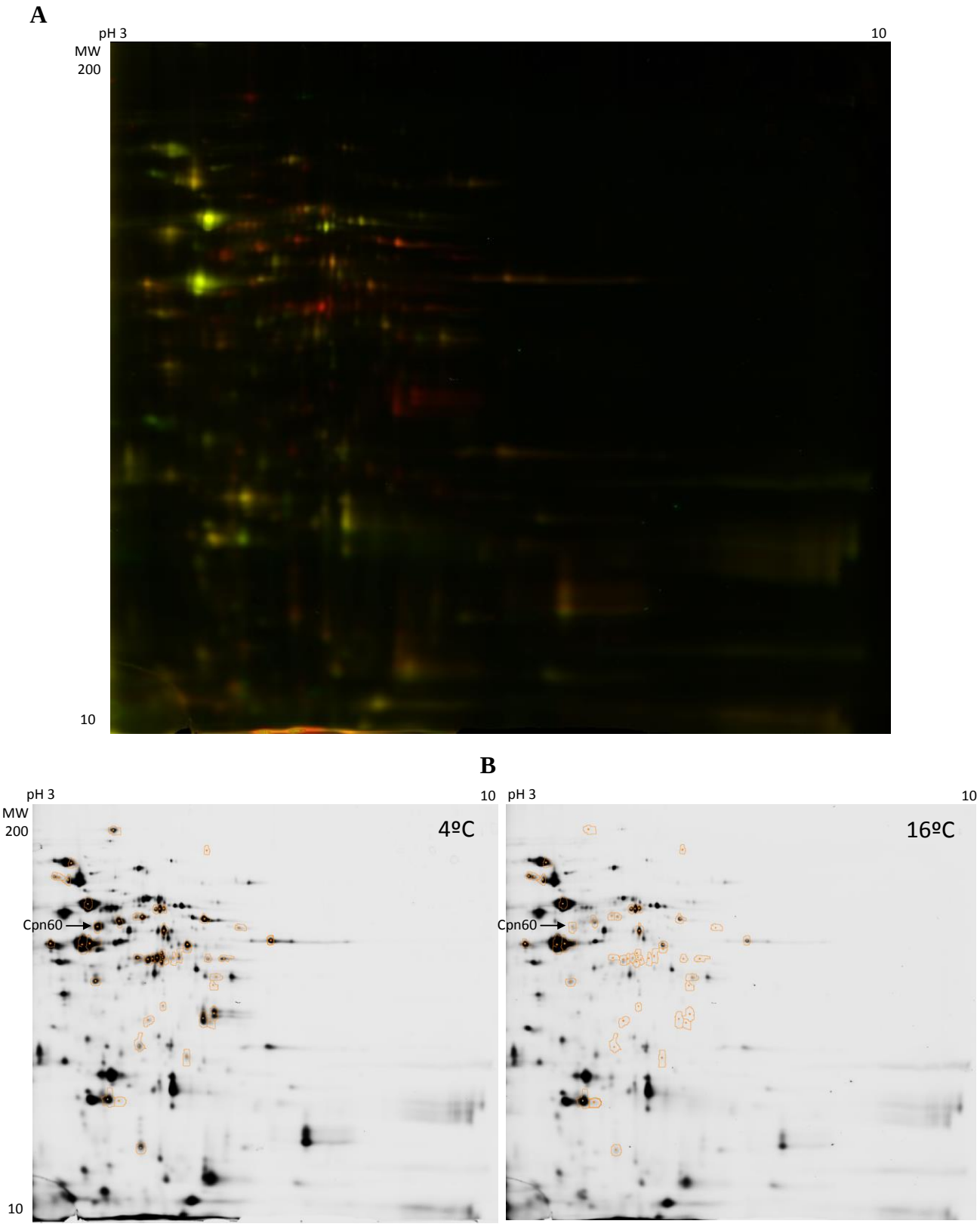
A



B

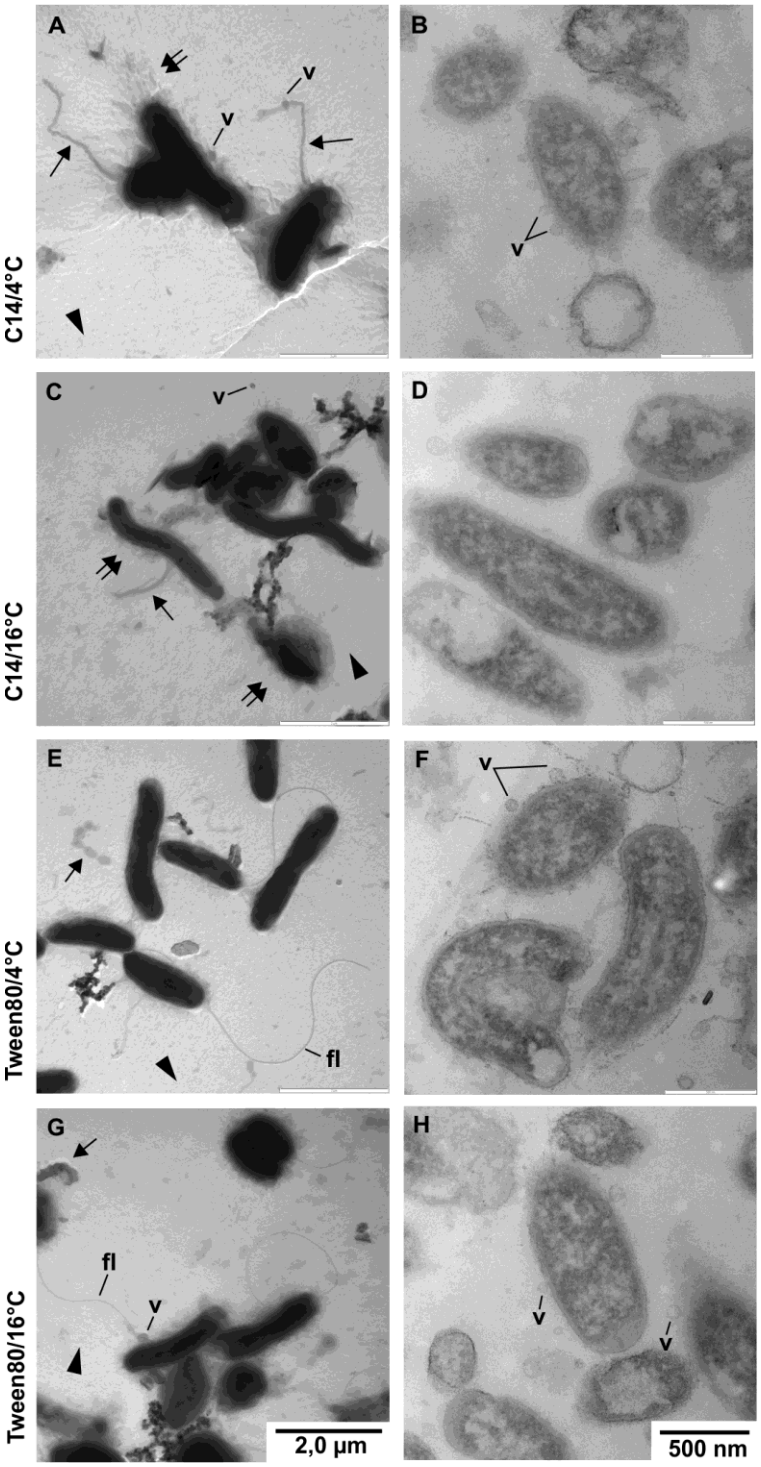


Supplementary Figure S7 (A-B) Cpn60 anti-proteomes of *Oleispira antarctica* grown at two different temperatures on *n*-tetradecane. A. Representative gel of the four anti-proteomes of *O. antarctica* cells grown at 4ºC (left) and 16ºC (right). An average of 209 ± 18 (n=4) spots were detected in 4ºC condition and 72 ± 8 (n=4) in 16ºC condition. B. Representative gel of the four anti-proteomes of *O. antarctica* cells grown at 4 and 16ºC with indication of spots identified.



Supplementary Figure S8 (A, B) Protein profile of *n*-tetradecane-grown *O. antarctica* cells at 4ºC and 16ºC

A: merged DIGE gel image of three gels showing protein extracts from cells grown at 4ºC (Cy5-labelled, red) and grown at 16ºC (Cy3-labelled, green) that have been used for chaperonin interactome study. **B:** Representative gel of the four anti-proteomes with indication of the 48 protein spots showing differential ratios with a 1.5-fold or greater change in abundance and consistent differences between the two conditions. Spot corresponding to the Cpn60 protein is indicated.



Supplementary Figure S9 Ultrastructure of *O. antarctica* from whole-mount samples (A,C,E,G) and ultrathin sections (B,D,F,H). Protrusions of the outer membrane (arrow) and the presence of a slime coat (twin-arrow) are indicative of C14 (*n*-tetradecane) substrate. Flagella can be observed when cells are growing in the presence of the detergent Tween 80 (E,G: fl); outer membrane vesicles = v; twin-arrow = slime coat; arrowheads indicate the shadowing direction. Bar in G is valid for A,C,E; bar in H is valid for B,D,F.

SUPPLEMENTARY TABLES

Supplementary Table S1. Genomic islands in the chromosome of *Oleispira antarctica* RB-8

GI #	Start	Stop	Length (bp)	Identified by	Origin
1	372974	391758	18784	SeqWord Sniffer	Phage and plasmid
2	433342	441012	7670	IslandPath-DIMOB	Plasmid
3	612637	632977	20340	SeqWord Sniffer	Plasmid
4	712000	755099	43099	SeqWord Sniffer and SIGI-HMM	Plasmid and phage
5	838267	843179	4912	IslandPath-DIMOB	Transposon
6	846168	865099	18931	SeqWord Sniffer	Plasmid
7	921136	952739	31603	Annotation data	Prophage
8	1520001	1539146	19145	SeqWord Sniffer	Plasmid
9	1767898	1790275	22377	SeqWord Sniffer	Plasmid and retrotransposon
10	1873853	1880869	7016	IslandPath-DIMOB	Transposon
11	1883942	1889558	5616	IslandPath-DIMOB	Transposon
12	1947441	1969355	21914	SeqWord Sniffer	Plasmid
13	1987888	1997548	9660	IslandPath-DIMOB and SIGI-HMM	Transposon
14	2016477	2036441	19964	SeqWord Sniffer	Plasmid and phage
15	2055256	2074015	18759	SeqWord Sniffer	Plasmid
16	2118996	2143270	24274	SeqWord Sniffer	Plasmid
17	2261552	2284734	23182	SeqWord Sniffer	Plasmid
18	2352305	2370661	18356	SeqWord Sniffer	Plasmid
19	2406863	2413882	7019	SIGI-HMM	Phage or retrotransposon
20	2541406	2569317	27911	SeqWord Sniffer	Plasmid
21	2705877	2731988	26111	SeqWord Sniffer	Plasmid and phage
22	2741361	2784124	42763	IslandPath-DIMOB and SIGI-HMM	Plasmid and phage
23	2955970	2982514	26544	SeqWord Sniffer and SIGI-HMM	Plasmid
24	3102000	3130802	28802	SeqWord Sniffer	Plasmid
25	3143177	3158750	15573	IslandPath-DIMOB	Transposon
26	3490711	3515412	24701	SeqWord Sniffer and IslandPath-DIMOB	Plasmid and phage
27	4215053	4219222	4169	IslandPath-DIMOB	Plasmid

Supplementary Table S2. Plasmids and phages showing amino acid sequence similarities with putative polypeptides in genomic islands of *Oleispira antarctica* RB-8

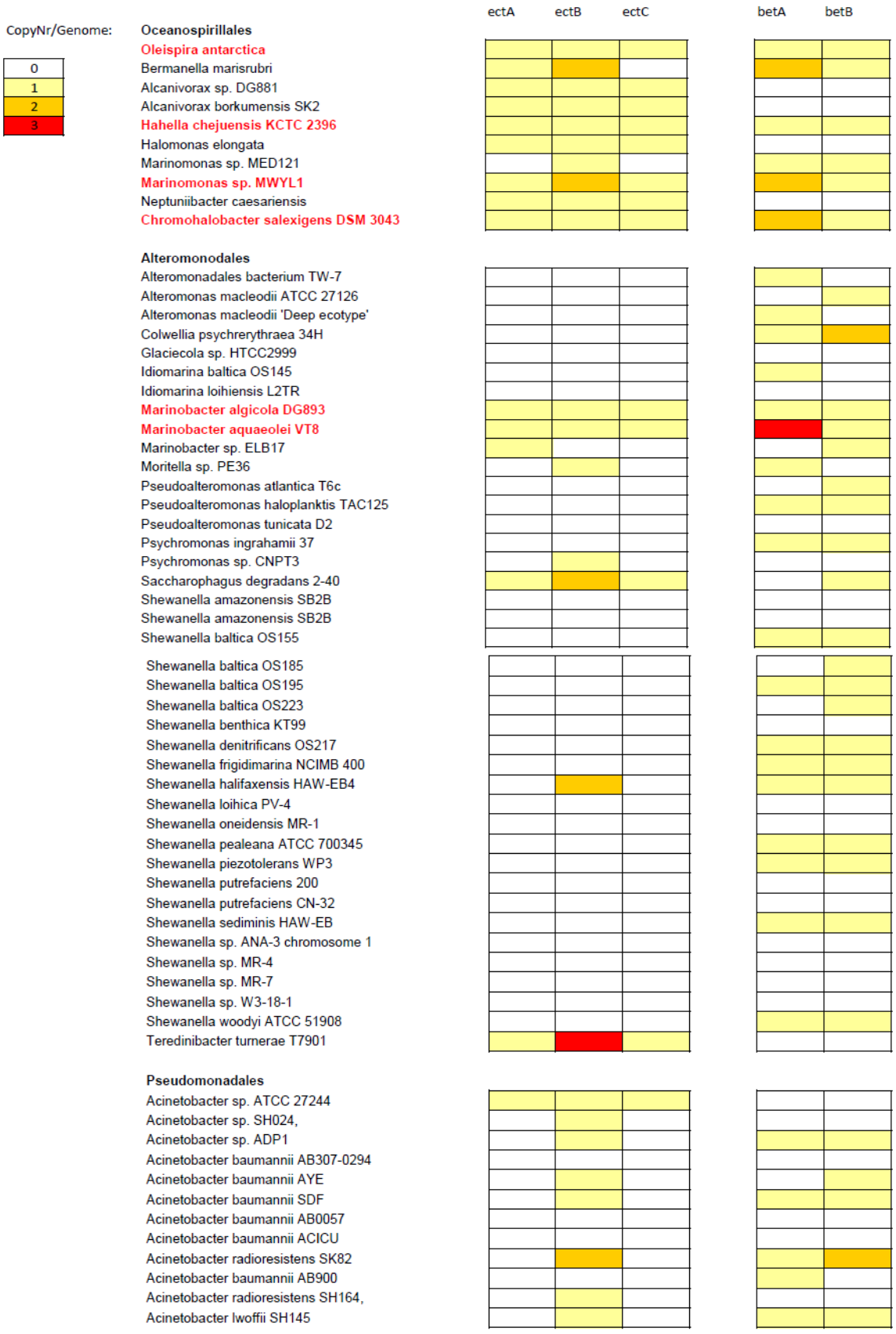
Top 5 plasmids	**Number of shared proteins**
Shigella flexneri 2a str. 301: pCP301 [NC_004851]	130
Xanthobacter autotrophicus Py2: pXAUT01 [NC_009717]	120
Polaromonas sp. JS666: plasmid 2 [NC_007950]	114
Marinobacter aquaeolei V8: pMAQU02 [NC_008739]	108
Pelobacter propionicus DSM 2379: pPRO1 [NC_008607]	102
Top 5 phages	**Number of shared proteins**
Haemophilus phage HP2 [NC_003315]	18
Clostridium phage c-st [NC_007581]	17
Enterobacteria phage RB43 [NC_007023]	12
Thermus thermophilus phage YS40 [NC_008584]	12
Cyanophage S-PM2 [NC_006820]	11

Supplementary Table S3. Census of genetic loci for alkane monooxygenases in Oceanospirillales and Alteromonadales with available genome sequences

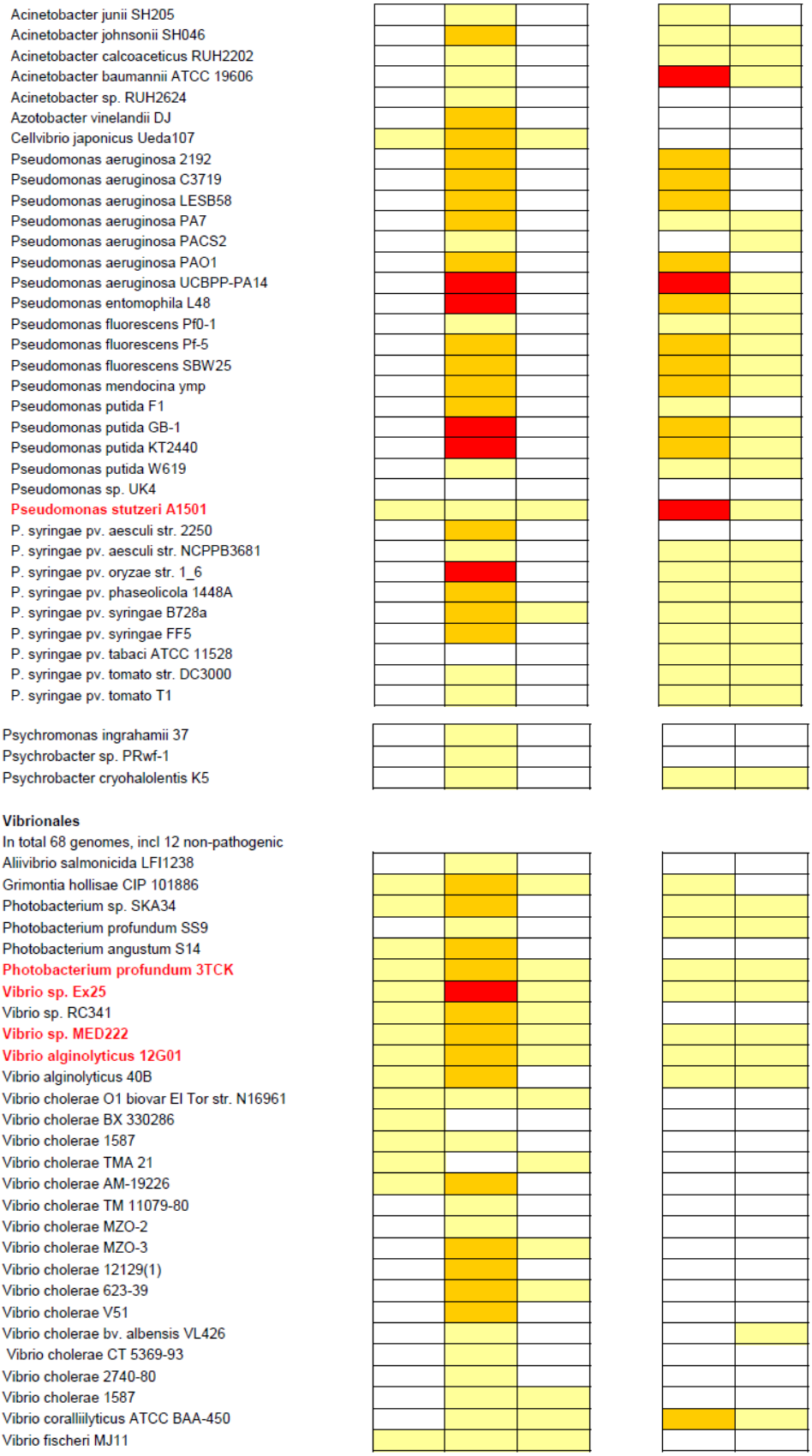
CopyNr/Genome:
0
1
2
3

	AlkB	p450
Oceanospirillales		
Oleispira antarctica RB8	3	1
Alcanivorax borkumensis SK2	2	3
Alcanivorax sp. DG881	2	3
Bermanella marisrubri	1	0
Chromohalobacter salixigens DSM 3043	0	0
Hahella chejuensis	1	2
Marinomonas MED121	0	0
Marinomonas MWYl1	1	0
Neptuniibacter caesariensis	0	0
Alteromonadales		
Alteromonadales bacterium TW-7	0	0
Alteromonas macleodii ATCC 27126	0	0
Alteromonas macleodii 'Deep ecotype'	0	0
Colwellia psychrerythraea 34H	0	0
Glaciecola sp. HTCC2999	0	0
Idiomarina baltica OS145,	0	1
Idiomarina loihiensis L2TR	0	0
Marinobacter algicola DG893	1	0
Marinobacter aquaeolei VT8	3	2
Marinobacter sp. ELB17	1	1
Moritella sp. PE36	0	1
Pseudoalteromonas atlantica T6c	0	1
Pseudoalteromonas haloplanktis TAC125	0	0
Pseudoalteromonas tunicata D2	0	0
Psychromonas ingrahamii 37	0	0
Psychromonas sp. CNPT3	0	0
Saccharophagus degradans 2-40	0	0
Shewanella amazonensis SB2B	0	0
Shewanella baltica OS155	0	0
Shewanella baltica OS185	0	0
Shewanella baltica OS195	0	0
Shewanella baltica OS223	0	0
Shewanella benthica KT99	0	0
Shewanella denitrificans OS217	0	0
Shewanella frigidimarina NCIMB 400	0	0
"Shewanella halifaxensis" HAW-EB4	0	0
Shewanella loihica PV-4	0	0
Shewanella oneidensis MR-1	0	0
"Shewanella pealeana" ATCC 700345	0	0
"Shewanella piezotolerans" WP3	0	0
Shewanella putrefaciens 200	0	0
Shewanella putrefaciens CN-32	0	0
Shewanella sediminis HAW-EB	0	0
Shewanella sp. ANA-3 chromosome 1	0	0
Shewanella sp. MR-4,	0	0
Shewanella sp. MR-7	0	0
Shewanella sp. W3-18-1	0	0
Shewanella woodyi ATCC 51908	0	0
"Teredinibacter turnerae" T7901	0	1

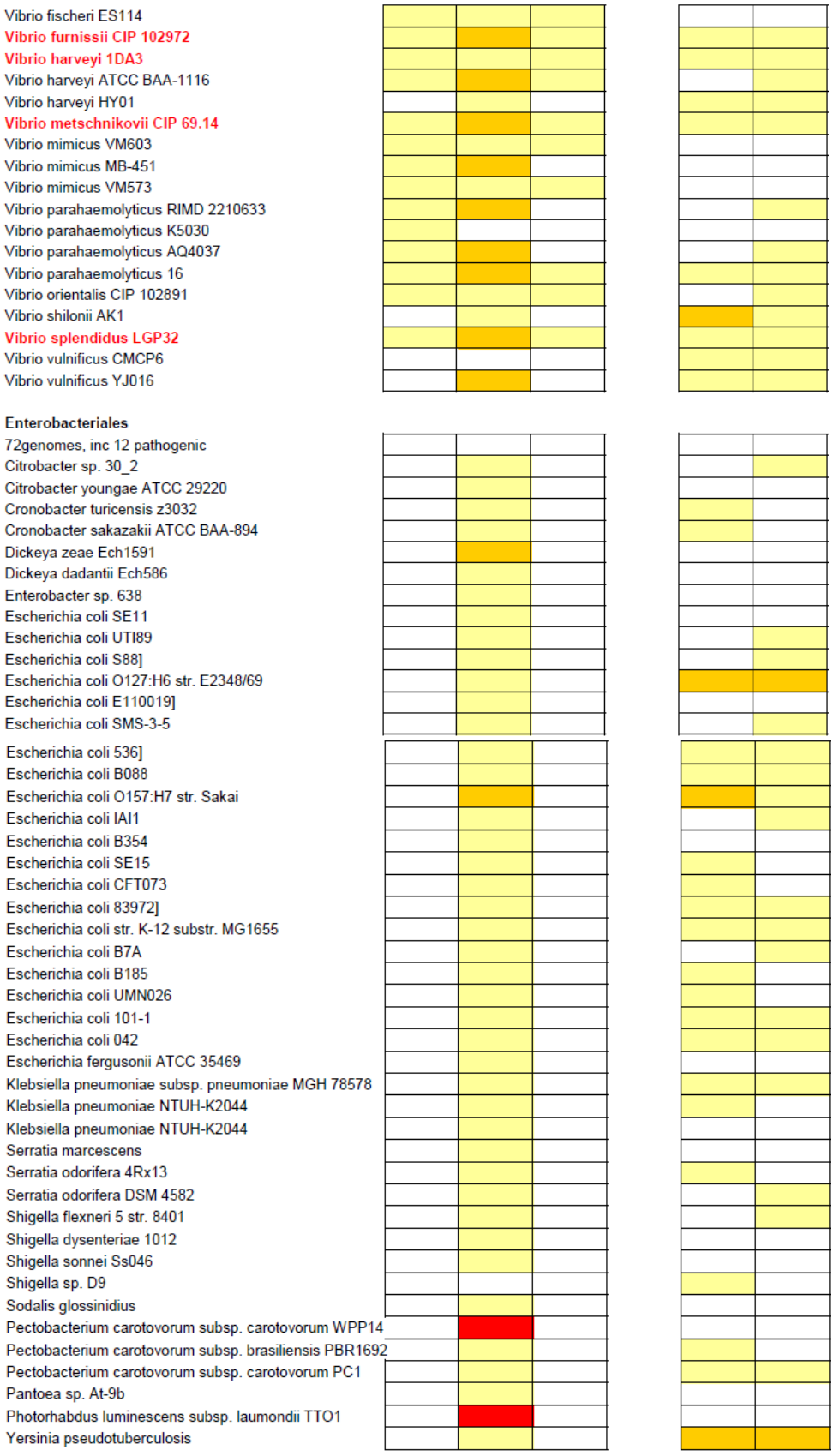
Supplementary Table S4 The gene inventory for ectoine and betaine synthesis pathways in representatives of Gammaproteobacteria with sequenced genomes



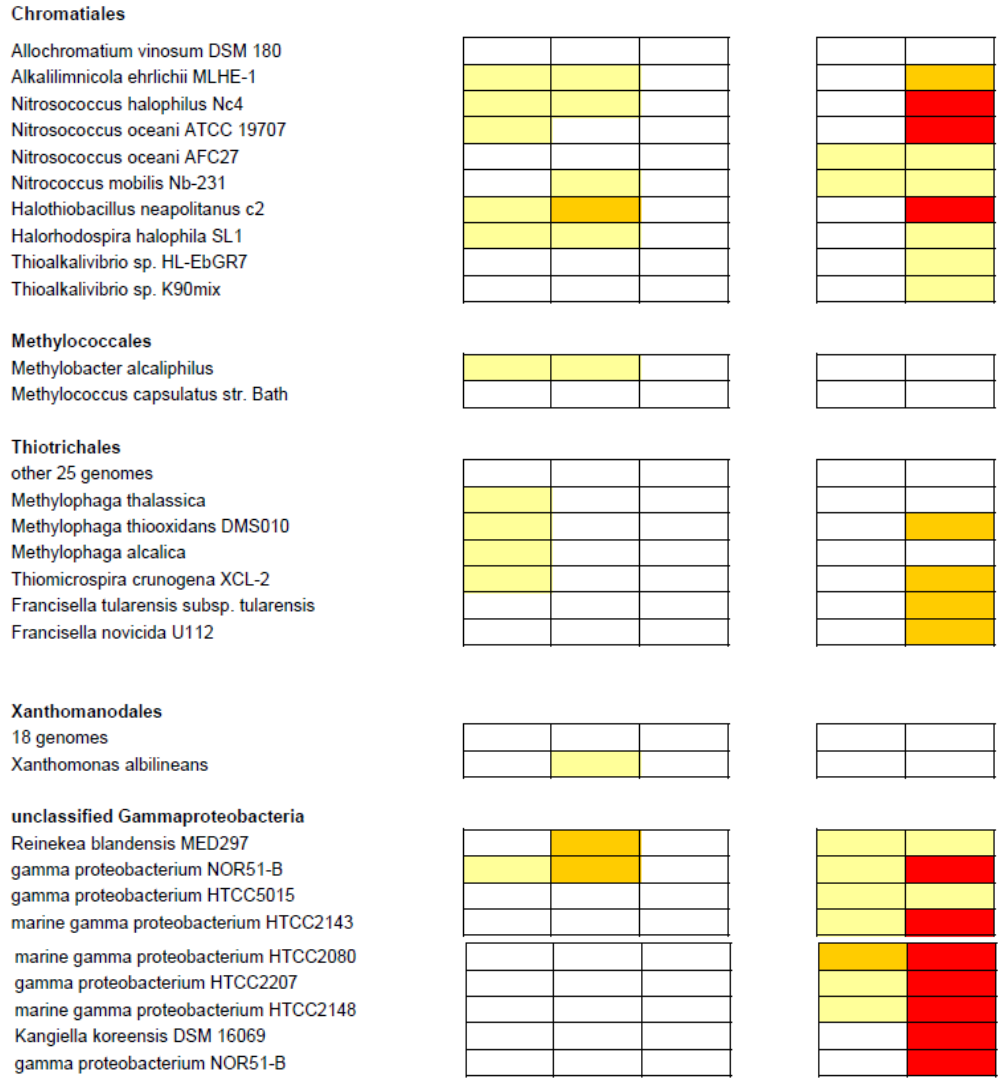
Supplementary Table S4 (continued) **The gene inventory for ectoine and betaine synthesis pathways in representatives of Gammaproteobacteria with sequenced genomes**



Supplementary Table S4 (continued) **The gene inventory for ectoine and betaine synthesis pathways in representatives of Gammaproteobacteria with sequenced genomes**



Supplementary Table S4 (continued) **The gene inventory for ectoine and betaine synthesis pathways in representatives of Gammaproteobacteria with sequenced genomes**



Supplementary Table S5. Census of genes for heme biosynthesis in *Oleispira antarctica* RB-8

ID	Name
OLEAN_C04820	*hemA, Glutamyl-tRNA reductase*
OLEAN_C02150	*hemL, Glutamate-1-semialdehyde 2,1-aminomutase*
OLEAN_C01970	*hemL, glutamate-1-semialdehyde-2,1-aminomutase*
OLEAN_C00640	*hemB, Delta-aminolevulinic acid dehydratase, porphobilinogen synthase*
OLEAN_C37010	*hemC, Porphobilinogen deaminase*
OLEAN_C37000	*hemD, Uroporphyrinogen-III synthase*
OLEAN_C36990	*hemX , Uncharacterized enzyme of heme biosynthesis, polytopic membrane protein which by an unknown mechanism down-regulates the level of HemA*
OLEAN_C36980	*hemY, Uncharacterized enzyme of heme biosynthesis, involved in a late step of protoheme IX synthesis*
OLEAN_C00240	*hemF, Coproporphyrinogen III oxidase, aerobic*
OLEAN_C20820	*hemH, Ferrochelatase 1*
OLEAN_C23270	*hemN, Oxygen-independent coproporphyrinogen III oxidase*
OLEAN_C36290	*hemE, Uroporphyrinogen decarboxylase*

Supplementary Table S6. Fatty acid compositions in *O. antarctica* when grown on different substrates and at two different temperatures.

	Acetate		Tetradecane		Tween 80	
	4°C	**15°C**	**4°C**	**15°C**	**4°C**	**15°C**
12:1 *cis*	0.7	0.3	4.2	4.2	0.8	1.1
14:0	1.4	3.0	1.5	0.6	5.0	4.1
14:1 cis	0.1	0.1	16.1	17.9	0.8	0.5
16:0	25.5	37.5	29.1	36.9	23.7	31.9
16:1 *trans*	0.8	0.8	1.2	1.2	2.5	4.7
16:1 *cis*	61.5	52.8	40.4	34.8	31.8	26.8
18:0	8.7	4.2	7.3	4.1	6.0	8.4
18:1 *cis*Δ9	0.5	0.3	0.2	0.1	28.9	21.6
18:1 *cis*Δ11	0.8	0.9	0.1	0.2	0.4	0.9
Degree of Sat.	**0.55**	**0.81**	**0.61**	**0.72**	**0.53**	**0.78**
16 *trans/cis*	**0.01**	**0.02**	**0.02**	**0.04**	**0.10**	**0.16**

All values are averages from three independent cultures

Supplementary Table S7. Cpn60 client proteins

No.	Protein name	Gene No.	Matches	M	E-value	Score
	ANTI-PROTEOME AT 4ºC					
1	AcnB Aconitate hydratase 2	OLEAN_C22770	34	99448	9.8e-050	526
2	FusA1 Elongation factor G 1 (EF-G 1)	OLEAN_C02300	26	76098	2e-028	313
3	PilQ Type 4 pilus biogenesis protein	OLEAN_C36360	31	76083	9.8e-068	706
4	TonB-dependent receptor	OLEAN_C00450	19	77353	9.8e-025	276
5	GlcB Malate synthase G	OLEAN_C20790	25	78631	2.5e-038	412
6	**DnaK Chaperone protein DnaK***	OLEAN_C06850	48	68143	3.9e-114	1170
7	Peroxidase/catalase	OLEAN_C17540	24	79780	2e-038	413
8	3-hydroxyacyl-CoA dehydrogenase	OLEAN_C14270	35	77612	6.2e-048	508
9	**Pnp Polyribonucleotide nucleotidyltransferase**	OLEAN_C32770	31	74632	1.2e-062	655
10	GTP-binding protein TypA	OLEAN_C01170	24	67042	1.6e-040	434
11	RpsA 30S ribosomal protein S1	OLEAN_C13950	31	61588	3.9e-077	800
12	OadA Oxaloacetate decarboxylase alpha chain	OLEAN_C28140	24	63658	9.8e-040	426
13	LysU Lysine-tRNA ligase	OLEAN_C23770	24	60238	3.1e-041	441
14	Acyl-CoA synthase	OLEAN_C27480	7	62874	7.6e-006	87
15	Acyl-CoA dehydrogenase	OLEAN_C03900	21	64746	1.2e-035	385
16	SdhA Succinate dehydrogenase subunit A	OLEAN_C16820	17	63873	3.1e-032	351
17	**Icl (AceA) Isocitrate lyase**	OLEAN_C12850	27	58058	9.8e-058	606
18	Nitrous-oxide reductase	OLEAN_C03170	25	70828	6.2e-036	388
19	lpdG Dihydrolipoyl dehydrogenase.	OLEAN_C16860	23	51062	1.6e-032	354
20	lpdG Dihydrolipoyl dehydrogenase	OLEAN_C16860	24	51062	6.2e-066	688
21	IlvC Ketol-acid reductoisomerase	OLEAN_C04020	28	54758	6.2e-037	398
22	IlvC Ketol-acid reductoisomerase	OLEAN_C04020	28	54758	6.2e-037	398
23	AtpA2 ATP synthase F1, alpha subunit	OLEAN_C39040	25	55059	2e-054	573
	ANTI-PROTEOME AT 16ºC					
24	Efp Elongation factor P	OLEAN_C05480	6	20842	0.0013	65
25	ProX Glycine betaine/L-proline-binding	OLEAN_C28210	17	36481	2.5e-055	582
26	Peroxiredoxin family protein/glutaredoxin	OLEAN_C10420	12	26475	1.2e-022	255
27	Peroxiredoxin family protein/glutaredoxin	OLEAN_C10420	10	26475	3.9e-022	250
28	Response regulator, CheY-like family	OLEAN_C12610	18	27806	9.8e-029	316
29	SodB Superoxide dismutase	OLEAN_C04950	4	21300	0.0031	61
30	NusG Transcription antitermination	OLEAN_C02320	7	20316	4.9e-010	129
31	NusG Transcription antitermination	OLEAN_C02320	16	20316	2e-072	753
32	Efp Elongation factor P	OLEAN_C05480	10	20842	3.9e-008	110
33	Frr Ribosome recycling factor	OLEAN_C10190	10	20733	1.2e-013	165
34	PpiB Peptidyl-prolyl cis-trans isomerase B	OLEAN_C20570	7	18164	9.8e-024	266
35	GreA Transcription elongation factor GreA	OLEAN_C06920	15	17538	4.9e-037	399
36	conserved hypothetical protein YceI-like, periplasmatic	OLEAN_C38350	5	21450	1.2e-012	155
37	Peptidyl-prolyl cis-trans isomerase (FKPB type)	OLEAN_C10560	10	17487	9.8e-011	136
38	conserved hypothetical protein YceI-like, periplasmatic	OLEAN_C38350	17	21450	1.6e-058	614
39	OsmC family protein (peroxyredoxin)	OLEAN_C04010	8	15198	3.1e-019	221
40	Periplasmatic chaperone Skp family	OLEAN_C10250	4	18660	0.0094	56
41	Anti-sigma factor antagonist	OLEAN_C03930	9	13554	1.6e-026	294
42	PilH Type IV response regulator (pilus retraction)	OLEAN_C00900	4	13347	0.0034	61
43	**DnaK Chaperone protein DnaK**	OLEAN_C06850	13	68143	1.6e-016	194
44	RpsA 30S ribosomal protein S1	OLEAN_C13950	13	61588	9e-007	100

Proteins are revealed through immune-coprecipitation with antibodies raised against Cpn60, 2-D gel electrophoresis and mass-spectrometry of tryptic digests of protein spots

*Proteins homologous to those from, and interacting with, Oleispira-Cpn60 at 4 C in, *E. coli* [43] are highlighted in bold.

Supplementary Table S8. The overview of *Oleispira* proteins with solved crystal structures and their structural homologues

Oleispira protein	PDB code	Tm, ºC	Nr of monomers in calculated oligomeric state	PDB code of closest structural homologue	Origin of closest structural homologue	Sequence similarity ID %	Nr of monomers in calculated oligomeric state
OLEAN_C30460	3I4Q	> 88	6	3D63	*Burkholderia pseudomallei*	56	3
OLEAN_C09750*	3I6Y/3S8Y	45	2	3FCX	*Homo sapiens*	54	2
OLEAN_C33610	3IRU	53	6	1FEZ	*Bacillus cereus*	40	2
OLEAN_C35840	3L53	52	12	1NR9	*E. coli*	50	2
OLEAN_C25130	3LAB	70	3	1FQ0	*E. coli*	51	3
OLEAN_C10530	3LMB	__	2	1T82	*S. oneidensis*	31	2
OLEAN_C13880	3LNP	67	2	1J6P	*T. maritima*	33	2
OLEAN_C07660	3LQY	70	2	2A67	*Enterococcus faecalis*	32	2
OLEAN_C18160	3M16	40	2	1ONR	*E. coli*	63	2
OLEAN_C08020	3QVM	49	2	1WOM	*B. subtilis*	42	1
OLEAN_C20330	3QVQ	56	2	2OTD	*Shigella flexneri*	31	1

*See the reference[46] for details

Supplementary Table S9 Comparison of *Oleispira* structures with mesophilic homologues.

		Hydrophobicity		Interactions				Oligomer interface				Charged amino acid content				Other amino acid content					
#	PDB codes *Oleispira* / mesophile	Decrease in core	Increase in surface	Less H-bonds	Less salt bridges overall / interface	Less aromatic-aromatic interactions	Less disulfides	Increased buried surface area	Decreased hydrophobicity	Less H-bonds	Less salt-bridges	Increased surface negative charge	Less total charged residues	High Glu+Asp / Lys+Arg ratio	Low Arg/Lys ratio	More Gly	More His	Less Ile	More Met	Less Pro (loops)	More Pro (helices)
1	3I4Q / 3SW5	-	+	+	-	-	equal	N/A	N/A	N/A	N/A	+	+	+	-	-	+	-	equal	+	equal
2	3I6Y / 3FCX	+	+	-	-	+	equal	+	+	-	+	+	+	+	-	+	-	+	+	+	equal
3	3IRU / 1SWV	+	+	-	-	+	equal	N/A	N/A	N/A	N/A	-	+	-	-	+	-	+	-	+	equal
4	3V77 / 3S52	-	+	-	equal	+	equal	N/A	N/A	N/A	N/A	+	equal	+	-	-	-	-	-	-	-
5	3VCR / 1FQ0	+	+	-	-	-	equal	+	+	-	equal	+	+	+	+	equal	+	+	+	equal	-
6	3LMB / 1T82	-	-	-	-	-	equal	+	-	+	+	+	-	-	+	+	-	-	equal	equal	equal
7	3LNP / 1J6P	-	-	-	+	+	equal	+	+	+	+	+	-	+	+	-	+	-	+	-	+
8	3L6Y / 2A67	+	+	-	-	-	equal	+	+	-	-	+	-	+	+	+	-	+	+	+	equal
9	3M16 / 1UCW	-	+	-	+	-	equal	N/A	N/A	N/A	N/A	+	+	+	+	equal	+	equal	+	-	-
10	3QVM / 1WOM	+	+	-	+	+	equal	N/A	N/A	N/A	N/A	+	+	+	+	+	-	-	equal	+	-
11	3QVQ / 2OTD	+	+	-	+	+	equal	N/A	N/A	N/A	N/A	+	+	+	-	-	+	-	-	+	equal

Oleispira / mesophilic homolog pairs with similar oligomeric structures were analyzed based in this shared oligomeric structure; those with distinct oligomers were analyzed in the context of an isolated chain. "+" refers to an observed adaptation, as listed across the columns of the table, "-" refers to an absent adaptation, N/A refers to analysis not performed due to distinct oligomer interfaces.

SUPPLEMENTARY DISCUSSION

Supplementary Discussion 1

Mobile genetic elements, genomic islands (GIs) and the horizontal gene transfer (HGT).

Genomic islands. There are several transposition hotspots in the chromosome, the first one in GI:4 (pos. 715,000-743500) affects the gene clusters of the type I restriction-modification system (with phage addiction *doc* (death on curing) homolog[55] which is adjacent with a gene cluster encoding GrpE, DnaK and DnaJ. Type I and type III restriction-modification system and DNA repair proteins are also encoded in GI:23 and GI:26. In GI:10 between positions 1,875,000 and 1,896,000 transposases IS4 and IS66 are neighboured by a number of hypothetical and truncated passenger genes (including the one for RNA-directed DNA polymerase /reverse transcriptase), gene clusters for TCA cycle components (OLEAN_C16790-OLEAN_C16850) and fatty acid biosynthesis (OLEAN_C17390-OLEAN_C17450). In GI:22 an IS66 and a mutator transposase are associated with chitin deacetylase (OLEAN_C24710), D-isomer-specific 2-hydroxyacid dehydrogenase (OLEAN_C24730), taurine catabolism dioxygenase (OLEAN_C24750) and sigma 54-dependent transcriptional activator (OLEAN_C24780). In GI:25 electron transport and NADH:ubiquinone oxidoreductase genes (OLEAN_C30670-OLEAN_C30680) are neighboring IS4 transposase and phage integrase.

Other horizontally-acquired metabolic genes are involved in fatty acid biosynthesis (GI:2, OLEAN_C03900-OLEAN_C03910); GI:24, OLEAN_C25950; GI:3, OLEAN_C05730-OLEAN_C05760); cytochrome/heme biogenesis and transportation (GI:20, OLEAN_C23150-OLEAN_C23230). In GI:14 several metabolic operons of genes involved in folate polyglutamylation pathway, O-succinyl-L-homoserine biosynthesis, CMP-KDO biosynthesis and adenosine *de novo* biosynthesis are adjacent to multiple hypothetical genes, colicin and sporulation genes of likely foreign origin. These metabolic genes have never been reported in mobile genomic elements and it is quite possible that they were falsely assigned to the genomic island because of alternations with horizontally acquired genes

There are efflux proteins (OLEAN_C05600 and OLEAN_C05620) and cobalt-zinc-cadmium resistance protein (OLEAN_C05610) in GI:3; GrpE, DnaKJ and DnaJ (OLEAN_C06840-OLEAN_C06860) chaperones in GI:4; heavy metal efflux pump protein in GI:15 (OLEAN_C18890); metallo-beta-lactamase (OLEAN_C19460) and capsule synthesis proteins (OLEAN_C19480-OLEAN_C19490) in GI:16; beta-lactamase (OLEAN_C21870) in GI:19; efflux transporter (OLEAN_C27790) in GI:24; and mercury resistance operon activator MerR (OLEAN_C13790) and an efflux protein gene (OLEAN_C37450) in GI:27. All these genes may be of importance for *O. antarctica* by allowing surviving of this hydrocarbonoclastic micro-organism in a hostile environment of natural or industrial oil spills that also often are associated with the heavy metal contamination.

A significant level of DNA compositional similarity was observed between GIs from Gammaproteobacteria, Bacteroides, Firmicutes and Spirochaetes. In the Supplementary Fig. S4 the divergence of GIs from the host chromosomes is depicted by grey colour gradient. GIs of *O. antarctica* very likely are older inserts than similar GIs in many other Gammaproteobacteria and Bacteroidetes except for *Colwellia psychrerythraea* 34H chromosome where a similar GI was an even older insert. Comparison of tetranucleotide usage patterns of GIs and the host chromosomes showed that the majority of GIs including those in *Oleispira* have originated from *Colwellia* spp. Initially these GIs were shared by

micro-organisms inhabiting the same environment – marine cold-adapted *Colwellia*, *Oleispira*, and *Psychrobacter*. Relatively recently these organisms have become donors of mobile vectors for a broad range of bacteria. Particularly, *Oleispira* might have donated mobile elements for *Kangiella* and later for several *Shewanella* and *Acinetobacter* (Supplementary Fig. S4).

Some weak sequence and composition similarity was observed even with a GI in such distant organism as *Leptospira interrogans* (NC_005823), however, in this particular case it was impossible to unambiguously establish the donor-recipient relationships. Oligonucleotide usage patterns of GIs also retained similarity with their previous hosts.

Hydrocarbonoclastic bacteria in general and *Oleispira* spp. in particular, have a good potential to be implemented for tackling oil spills through bioaugmentation or biostimulation. However, the bioaugmentation option (i.e. the large-scale introduction of pre-grown biomass) must be treated with an extreme care, considering aspects of the horizontal gene transfer. The genes that ensure survival of oil-degrading strains in the environment hostile for common terrestrial microorganisms may become virulence factors in the latter bacteria. For example, the recent outbreak of an enterohemorrhagic *Escherichia coli* in Germany in 2011 was caused, to some extent, by acquisition of the drug resistance genes by *E. coli* from marine beta-Proteobacteria[40]. Supplementary Fig. S4 shows intimate relations between GIs of *O. antarctica* and those in notorious nosocomial pathogens of *Acinetobacter* and *Bacteroides*. Those formerly saprophytic and commensal microorganisms have emerged to cause severe infections only after a recent development of an extreme resistance against many antibiotics and disinfectants[56]. Indigenous marine bacteria may play a significant role in the development of new emerging pathogens by providing them with effective efflux pumps and drug-resistance factors. The rise of ocean temperatures may destabilize chromosomes of cold-adapted *Colwellia* and *Oleispira* and activate mobile elements within their genomes. Having become ubiquitous, those mobile genetic elements could be acquired by pathogenic and conditionally pathogenic microorganisms turning them into super-pathogens. GIs of *O. antarctica* comprise plenty of potentially dangerous genes encoding multiple efflux proteins, beta-lactamases, heavy metal resistance proteins and even a virulence-associated protein RhuM (OLEAN_C06710) in GI:4 which also is present in the pathogenicity genomic island SPI-3 of *Salmonella*[57]**.** Taking that into account, a release of great numbers of this microorganism into environment be considered with a great care and the biostimulation of oil-degrading bacteria

Prophage. The genes of lysogeny module are in the opposite orientation from the rest of the phage genes, resembling that in the type A lambdoid prophages. It starts with a putative replication protein A (OLEAN_C08560) a cluster of genes to represent a remnant DNA replication module (including hypothetical DNA binding proteins, DNA methylase and a terminase subunits). Remarkably, terminase subunits are in the opposite orientation and separated genes encoding portal protein (OLEAN_C08650) and major capsid protein (OLEAN_C08680). A cluster of P2-like putative tail assembly and structural genes follows the capsid assembly genes. The complexity of these genes including at least ten putative CDSs involved in tail assembly and the strong identity score for a contractile tail sheath protein (OLEAN_C08730) suggests that the prophage was a member of the *Myoviridae*, i.e., phages possessing a contractile tail. The lateral gene in the prophage-like sequence was similar to a phage Gp5 lysozyme, which helps to release mature phage particles from the cell wall by breaking down the peptidoglycan. This prophage also shares significant sequence similarity with the *Myoviridae* phage HP2 isolated from *Haemophilus* **(**Supplementary Fig. S3**)**. The prophage was immobilized in the *O. antarctica* chromosome most likely because of a disruption of genes encoding holin and lysin, which are important for the release of viral particles from bacterial cells. At the end of the infection by an intact phage, holin permeabilizes the cytoplasmic membrane allowing access of the phage lysin to its murein substrate[58]. Knocking out of the lysis enzymes might suppress the excising activity of the

phage integrase that caused immobilization of the prophage in the chromosome. Prophages with similar genome organisation are frequently occurring in marine proteobacteria, e.g. *Vibrio splendidus* 12B1, *V. alginolyticus* 12G01, *Thiomicrospira crunogena* XCL-2, *Roseobacter* SK209-2-6 and *Silicibacter* TMS1040[11]. As it is quite common for marine prophages, to exhibit extremely low homologies of their DNA and polypeptide sequences with those from phages of the same type. In contrast, genes of the lytic module of the *O. antarctica* prophage had a significant similarity to the corresponding genes of two almost identical prophages found in the chromosome of another marine gamma-proteobacterium, *Hahella chejuensis*[59] (Fig. 2B). Noticeably, these prophages lack a discernable phage repressor protein and belong to type B of marine prophages[11]. Such a mixed heritage is often the result of the modular evolution of phages. The most versatile are lysogeny and replication modules that are probably host-specific. Interestingly, that this obvious prophage was not detected by any genomic island identification tools. The prophages in *O. antarctica* and *H. chejuensis* are indistinguishable by DNA composition from the host chromosomes. Precise adaptation towards specific codon usage and compositional constraints of the host chromosomes is likely very important for these phages. As a common feature for type A marine putative prophages is a high frequency of occurrence of transcriptional regulators and repressors. Coupling these observations with the fact that coliphage λ repressors can actively repress and modulate host metabolic genes, Paul[11] hypothesized that marine prophages serve to repress host growth in times of resource partitioning. Such metabolic economization seems to be very logic bearing in mind the life style and specialization of *O. antarctica*, which metabolic activity strongly depends on aliphatic hydrocarbons, sporadically appearing in polar marine ecosystems.

CRISPR-Cas. The search for CRISPR-Cas-related sequences in *O. antarctica* genome returned only a short 53aa-long peptide fragment (OLEAN_C21530) homologous with C-terminus of Cse1 from *Psychromonas ingrahamii* 37, which suggests the CRISPR-Cas-based phage defense mechanism is absent in *O. antarctica*.

Supplementary Discussion 2

Genome inspection for genetic loci featuring marine lifestyle

Transport of divalent cations. The genome contains a number of genes for systems involved into the efflux of metals: the high-affinity zinc uptake system protein ZnuA (OLEAN_C00390); the possible cobalt-zinc-cadmium resistance protein (cation efflux system protein) (OLEAN_C05610); the cation efflux system protein CzcA (OLEAN_C05600); the outer membrane efflux protein, Co/Zn/Cd efflux system component precursor (OLEAN_C05620); OLEAN_C21330, the arsenate resistance protein ArsH; and OLEAN_C01780, the arsenical pump membrane protein. Few putative copper transporting/resistance systems are present in *O. antarctica* genome. The system Cus with two copies for *cus*A genes (OLEAN_C31490, OLEAN_C37610), cation efflux system protein and one copy for *cus*B gene (OLEAN_C31500) is present in the *O. antarctica* genome. Other system, Cop, with two copies of genes for CopA (OLEAN_C27140, OLEAN_C37680) and one copy for CopB (OLEAN_C37690) was found in the genome as well. Both copper transport systems are known be also involved into the control of reactive oxygen that increase their concentrations at lower temperatures due to the higher solubility of gases.

The gene OLEAN_C00650 coding for MerA, mercuric reductase, could be probably responsible for the mercury resistance in *O. antarctica,* however no other common components for mercury reduction pathway[60], i.e. neither periplasmatic MerP, or transfer protein MerT, or any common regulatory proteins have been found to be encoded in the vicinity of the gene for mercuric reductase. Regarding resistance to arsenite/arsenate, *Oleispira* genome bears the gene OLEAN_C21330, the arsenate resistance protein ArsH, and a gene cluster consisting of OLEAN_C01770-OLEAN_C01780 most likely coding for a probable transcriptional regulator of ArsR family and for a putative membrane arsenical pump protein, correspondingly.

Compatible solutes. *O. antarctica* was isolated from superficial sea-water samples from the inlet Rod Bay (Ross Sea). At the site of isolation, the salinity of the Ross Sea is about 33.7-34.4 ‰ with average temperatures around -1,8° -- 0°C and only during Summer, when significant part of Rod Bay is free of ice, the water temperature slightly rises, increasing its temperature to +2ºC. One could therefore anticipate the presence of biosynthesis systems for compatible solutes to allow *O. antarctica* coping with both, low temperatures and salinity. Indeed, all corresponding genes (OLEAN_C31120, OLEAN_C31110 and OLEAN_C31100) are present in the genome as well. The ABC-type proline/glycine betaine transporters (genes OLEAN_C35510, OLEAN_C35520 and OLEAN_C35540**),** and other numerous glycine betaine transporters (OLEAN_C20610, OLEAN_C27450, OLEAN_C27670, OLEAN_C28210, OLEAN_C28220, OLEAN_C28230, OLEAN_C31310, OLEAN_C35510, OLEAN_C35530), repressor involved in choline regulation of the *bet* genes (OLEAN_C35560)**,** choline phosphate cytidylyltransferase/choline kinase (OLEAN_C35570) and the regulator with the marR-type HTH domain (gene OLEAN_C31130) were further identified with a high probability in the genome of *O. antarctica.* From 343 fully sequenced Gammaproteobacteria analysed only 15 possess both systems: just 1 genome of 49 in the order Pseudomonadales; only 2 genomes from 40 of Alteromonadales; 8 of 68 in Vibrionales and 4 genomes (including *O. antarctica*) among 10 genome sequences available for Oceanospirillales. The system for cryoprotectant choline betaine production is well known in psychrophiles[61] and is also present in *Oleispira.* The pathway for production of choline betaine occurs less frequently than that for ectoine in closest relatives of *O. antarctica* from the family Oceanospirillales (Supplementary Table S4). Thus, the combination of alternative systems for synthesis of osmoprotectants in *O. antarctica* certainly makes this organism well-equipped for the life in the sea-ice environment.

Biosynthesis and degradation of lipophilic macromolecules. The genes related to PHAs synthesis and degradation are not clustered in operons, as it is the case in *Ralstonia eutropha*

H16 or in some pseudomonads. The gene OLEAN_C05780 encodes an enoyl-CoA hydratase/isomerase which could alternatively function as acetoacetyl-CoA-reductase (PhaB). OLEAN_C13230 encodes a protein with mol weight of 14,187 Da related to the phasin (PhaP) superfa OLEAN_C13230 mily and exhibiting 29% AA sequence identity with the phasin of *Ectothiorhodospira* sp. PHS-1. The above gene is co-clustered with OLEAN_C13240 encoding PhaC, a typical poly-3-hydroxyalkanoic acid synthase of 67,350 Da with 60 % AA sequence identity with the poly(3-hydroxyalkanoate) synthetase of *Hahella chejuensis* KCTC 2396. Unlike in *A. borkumensis* SK2 which contains in its genome two *phaC* genes of the class II, *phaC1* and *phaC2*[15], the genome of *O. antarctica* RB-8 *phaC* has only a single copy, and the enzyme size and some conserved reactive motifs resemble those of typical PHA synthases of the class I. Two alternative PhaAs have been identified as acetyl-CoA acetyltransferases of thiolase superfamily. The PhaA1 is encoded by OLEAN_C32530 and has about 74 % protein sequence identity with the enzyme from *Moritella* sp. PE36. The other PhaA2 is encoded by OLEAN_C35320, which composed of 391 amino acids with 71 % identity to the *phb*A gene product of *Halomonas elongata* DSM 2581. The *pha*A1 (OLEAN_C32530) is clustered with, and situated between, the OLEAN_C32520 encoding the 3-hydroxyacyl-CoA-dehydrogenase of type II and the OLEAN_C32540 encoding the large subunit of 3-hydroxyacyl-CoA dehydrogenase. A similar co-expression pattern has also been identified for PhaA of *R. eutropha* H16. Thus, the PhaA1 is likely a crucial enzyme in the PHAs synthesis in *O. antarctica*. The occurrence of a typical *pha*B, which encodes an acetoacetyl-CoA-reductase, has not been confirmed, whereas the OLEAN_C05780-encoded protein, the enoyl-CoA hydratase/isomerase of crotonase superfamily, could possibly serve as an alternative PhaB. The corresponding gene is not clustered with any other genetic locus of PHA biosynthetic components.

The genome analysis further revealed the absence of genes responsible to wax ester synthetase, including wax ester synthase/acyl coenzyme A:diacyl glycerol acyltransferase (WE/DGAT) as well as their alternative acetyltransferases (AtfA1 and AtfA2) in *O. antarctica*. In spite of the presence of the PHA biosynthetic components, no deposition of biopolymers has been observed in ultrathin sections of cells under standard growth conditions tested[4]. The reason for that could possibly be the expression of the typical poly-3-hydroxyalkanoate depolymerase encoded by OLEAN_C28830 and sharing up to 64 % protein sequence identity to the poly-beta-hydroxyalkanoate depolymerase (PhaZ) of *Grimontia hollisae* CIP101886. This enzyme belongs to the one of the intracellular type and is capable of a simultaneous degradation of PHAs deposits in the cells. In many cases the accumulation of PHAs could only occur when the *pha*Z is inactivated or when the suitable carbon source is in an excess over the nitrogen in the growth medium.

Flagellation and motility. Flagellar protein FlaS; transcriptional flagellar regulator FleQ; sensory box histidine kinase FleS; Sigma-54 dependent DNA-binding response regulator FleR; receiver protein CheY; signal transduction histidine kinase CheA; chemotaxis signal transduction protein CheW; flagellar M ring protein FliF; motor switch protein G FliG; flagellum-specific ATP-synthase/H^+-transporting ATPase FliI; flagellum biosynthesis chaperone FliJ; flagellar hook length-control protein FliK; flagellar basal body-associated protein FliL, flagellar motor switch proteins FliM and FliN, flagellar biogenesis proteins FliO, FliP, FliQ and FliR, FlhA, FlhB FlhR; flagellar biosynthesis regulator SRP54 subunit GTPase FlhF; flagellar number regulator FleN; RNA polymerase sigma factor for flagellar operon FliA; CheY-like receiver protein; chemotaxis phosphatase CheZ; CheA-like chemotaxis protein histidine kinase; chemotaxis response regulator receiver CheY; and finally, two CheW-like proteins are encoded in the same order within an unidirectionally transcribed gene cluster OLEAN_C12070-OLEAN_C12510.

Two potential stator systems have been found in *O. antarctica* – MotAB (OLEAN_C05070, OLEAN_C05080, OLEAN_C35090, OLEAN_C35080) and PomAB (OLEAN_C12460, OLEAN_C12470, OLEAN_ OLEAN_C36090, OLEAN_C36100). Each stator unit is

determined by different ion motive force, H+ - driven motor for Mot and Na+ - driven motor for Pom systems; B subunits of these stators anchor to the peptidoglycan, organize the channel for the driving H+ or Na+ ions and interact with Asp24 and Asp32 residues of PomB and MotB, respectively[17]. The presence of both stator systems could be very beneficial for *O. antarctica* allowing not only swimming, but also swarming abilities on surfaces in the environments with lower sodium concentrations such as the Antarctic sea ice.

Phosphate and sulfur uptake and the nitrogen metabolism.

Phosphate uptake is facilitated by high-affinity ABC transporter systems including PhoB (OLEAN_C31720, OLEAN_C38760) with the regulator PhoU (OLEAN_C31690); PstCAB (OLEAN_C31660- OLEAN_C31680) phosphate binding protein (OLEAN_C31650). Regarding the sulfur uptake, the *Oleispira* genome harbors two copies of genes for CysD for sulfate adenylyltransferase subunit 2 (OLEAN_C10920, OLEAN_C19550), for sulfate adenylyltransferase subunit 1 CysN (OLEAN_C19540), CysZ (OLEAN_C38780), CysH (OLEAN_C20240), sulfate transporter SulP (OLEAN_C30340) and sulfate permease family protein (OLEAN_C30700). In relation to the nitrogen metabolism, the organism is able to grow under anaerobic conditions by nitrate reduction[4]. Accordingly, in the genome of *O. antarctica* we have identified the whole array of genes (*nap*, *nir*, *nor* and *nos*) responsible for the denitrification. The first step of the process is conducted by the periplasmic nitrate reductase (Nap) ubiquitous in Gram-negative bacteria[62,63]; the nitrate reduction genes in the *O. antarctica* are arranged in the cluster *napEFDABC*, in the same manner as in *Pseudomonas* spp. G-179[63]. The first step in denitrification is followed by the reduction of nitrite to nitric oxide through a dissimilatory nitrite reductase (Nir) encoded by a gene cluster at pos. 2443401-2455751. Finally, the last step of denitrification[64] is likely catalyzed by nitrous-oxide reductases encoded by the clusters *nor*CBD (OLEAN_C21730-OLEAN_C21760) and *nos*RZDFYL (OLEAN_C03130- OLEAN_C03180).

Albeit *O. antarctica* genome encodes RnfABCDGE constituting the six-subunit complex of an apparent NADH oxidoreductase responsible for electron transport to nitrogenase, neither molybdenum-, nor vanadium-dependent nitrogenases have been found to be encoded in the genome, suggesting that *Oleispira* is not able to fix the molecular nitrogen.

The *Oleispira* genome encodes two ammonium transporter systems - Amt (AmtB, OLEAN_C31170; OLEAN_C37150 and AmtE - OLEAN_C31860); one ammonia permease (OLEAN_C16050), two-component response regulator NtrC (OLEAN_C03720) and proteins similar to the nitrogen regulatory protein PII (OLEAN_C37160; OLEAN_C37730 and OLEAN_C37820). In eubacteria, the AmtB protein was found to be associated with the P_{II} signal transduction protein (GlnK) and is known as a part of a system for sensing external ammonium in the Ntr regulon[65]. We have found one operon for the uptake and assimilation of urea (*ureDABCEFG*, OLEAN_C37290- OLEAN_C37230) and the adjacent gene cluster with few related transporters (OLEAN_C37220- OLEAN_C37180).

Transport of divalent cations. For import of cations, such as Zn, Mo, Mg, Co and Ni, the genome encodes the array of proteins ZnuC and ZnuB; Zur transcriptional repressor of Zn transport system (genes OLEAN_C00400-OLEAN_C00420 in the same order) the high-affinity zinc uptake system protein ZnuA (OLEAN_C00390); zinc transporter family protein Zip (OLEAN_C23060) and zinc-binding protein (OLEAN_C29960); the molybdate ABC transporter ModCBA, (OLEAN_C31340- OLEAN_C31360) and MobA (OLEAN_C03070); the magnesium and cobalt transport protein CorA (OLEAN_C27680); the magnesium transporter OLEAN_C04380; the Mg/Co/Ni transporter MgtE (OLEAN_C04380); the magnesium chelatase CobN (OLEAN_C15180). The genome inspection has also revealed no specific Mn-transport systems in *O. antarctica*, however some "zinc transporters" mentioned above could be capable of Mn-uptake[66].

Iron and siderophores, The low availability of iron in most environments has been well documented and poses a challenge for virtually all life forms, due to the essential catalytic and structural roles this element plays in proteins. Siderophores, the low-molecular-weight

metabolites possess extremely high-affinity to Fe(III) ions. They solubilise and coordinate iron by formation of Fe(III)-siderophore complex, which is recognized by the siderophore-specific cell surface receptors that transport chelated iron into the bacterium.
The genome inspection for siderophore-producing loci has revealed the following inventory. Within the gene cluster spanning the ORFs OLEAN_C24170 and OLEAN_C24310, the OLEAN_C24210 is coding for the acyl-CoA synthase (ACS); OLEAN_C24220 for siderophore synthetase component, IucA/IucC outer membrane receptor for aerobactin; OLEAN_C24230 for IucA and IucC family protein as well as OLEAN_C24310 for siderophore-interacting protein. Genes OLEAN_C24200 and OLEAN_C24190 code for acyl carrier protein and sugar phosphate iso/epimerase homologous to *Bacillus*-like type acyl-CoA synthase or petrobactin biosynthesis proteins AsbE and AsbF, respectively. The gene product of OLEAN_C24180 is pyridoxal-dependent carboxynorspermidine decarboxylase which could act on ornithine, lysine, arginine and related substrates and saccharopine dehydrogenase encoded by OLEAN_C24170 has been reported to act as a bifunctional polypeptide with lysine ketoglutarate reductase activity. Proposed aerobactin synthesis reaction from lysine and citric acid as substrates requires the products of genes *iuc*D, B, A and C that belong to IucA/IucC family proteins[67]. All above proteins could be involved in synthesis of aerobactin-like, or any other type of, siderophore in *O. antarctica*. Proteins encoded by gene OLEAN_C24310, siderophore-interacting protein and numerous periplasmic binding and inner membrane transport proteins are serving to actively deliver siderophore complexes across cell membrane. Gene clusters with similar arrangements of ORFs for siderophore production could be found in genomes of other bacteria, such as in ubiquitous marine oil-degrader *Marinobacter aquaolei* and in *Verrucomicrobium spinosum* fairly distant taxonomically (Supplementary Fig. S6a). Furthermore, we have obtained the experimental evidence of elevated expression of the gene for ACS in *O. antarctica* culture
Applying the standard chrome Azurol S reagent (CAS) assay, according to Schwyn & Neilands[68], we monitored the siderophore production by *O. antarctica* grown on tetradecane as sole carbon and energy source under iron-limiting conditions (Supplementary Fig. S6b). At the low cell density the available iron is likely sufficient for the growth of *O. antarctica* and the presence of siderophore for the first time was detected in the medium only after 72 hours of cultivation. Further monitoring revealed the undulate mode of siderophore appearance, suggesting the complexities of siderophore recirculation mechanism and/or regulation of expression of corresponding genes. To elucidate this, we performed the q-RT-PCR analysis. As discussed above, the genome of *O. antarctica* contains the gene cluster OLEAN_C17080-OLEAN_C16830 for enzymes potentially involved in the siderophore synthesis comprising putative siderophore synthetase components, acyl-CoA synthetase and acyl carrier protein (Supplementary Fig. S6). Expression profile of OLEAN_C16920 gene for acyl-CoA synthetase correlated with the appearance of siderophore, although its expression levels appeared to be rather uniform during the late exponential and stationary phases of growth. This finding suggests that the dynamic interplay between the uptake and release of siderophores drive its notched appearance in the medium, rather than the (basal) expression of genes for siderophore biosynthesis.
Lipid analysis. FAME data on *Oleispira* cultures grown in three different substrates and at two different temperatures decreased the degree at saturation, which is a common way to increase the membrane fluidity at low temperatures (Supplementary Table S6). Quite interesting were also the differences emerged depending on the growth substrate: when the cells grew on tetradecane and Tween-80, they showed a different fatty acid pattern because of the fact that *Oleispira* was able to either convert the alkanes to the corresponding or similar fatty acids, or to incorporate the oleic acid (18:1 delta 9 *cis*) present in Tween 80 directly in their membrane phospholipids.

Supplementary Discussion 3

General features of *O. antarctica* proteome and anti-proteome. Differential In-Gel Electrophoresis (DIGE) was performed to compare the total proteome from *O. antarctica* cells growing at 4ºC and 16ºC. An average of 549 spots was detected in the 4ºC condition and 498 at 16ºC when the total proteomes were examined. Further, 48 protein spots showed differential ratios with a 1.5-fold or greater change in abundance and consistent differences between the two conditions (Student's t test, $p<0.05$ n=3, applying FDR correction). Thirty four spots were found to be more expressed at 4ºC (from 1.59 to 36.7-fold), while only 14 were at 16ºC (from 1.61 to 2.63-fold). Cpn60 protein is included among the differentially expressed proteins (6.4-fold higher expression at 4ºC). Using a densitometry analysis, a tentative relative proportion of Cpn60-to-protein spots in the detected protein extract, could be obtained: 0.64% (at 4ºC) and 0.11% (at 16ºC) in the total proteomes. These values accounted for 1.97% (at 4ºC) and 0.91% (at 16ºC) in the anti-proteomes.

The Cpn60-interactome study was conducted using immune-precipitation, two-dimensional gel electrophoresis (2D-E) and MALDI-TOF peptide mass fingerprinting, as described earlier[39] (see Supplementary Methods for further details). Under the experimental conditions tested, the 2D anti-proteome of cultures at 4ºC revealed that a significant proportion of Cpn60 protein substrates had molecular masses above 50 kDa, which, most certainly, refers to *Oleispira*-derived Cpn60 acting as a single heptameric ring at low temperatures[25]. At 16ºC the protein substrates for chaperonin were found to have molecular masses mostly below 40 kDa, which is typical for Cpn60-like protein substrates in their classical two-barrel (hexadecameric) conformation. Using MALDI-TOF peptide mass fingerprinting and MS/MS analysis 23 and 21 spots in anti-proteomes from 4 ºC and 16 ºC, respectively, were unambiguously identified (Supplementary Table S7). Theoretical molecular masses of chaperonine partner proteins identified at 4ºC ranged from 51 to 99 kDa whereas those from 16 ºC ranged from 13 to 28 kDa, thus confirming that there is no or a little size limitation for protein folding mediated by Cpn60 at 4ºC. This may be a consequence of a heptadecameric single-barrel conformation of Cpn60 at this thermal condition which is a unique feature for *O. antarctica*[25].

Supplementary Discussion 4

Ultrastructural analysis of *Oleispira* cells

Oleispira antarctica was analyzed by transmission electron microscopy to study the ultrastructural features that could be observed on *in vitro* impact of the temperature (4 ºC versus 16 ºC) and *n*-tetradecane (C14) versus detergent (Tween 80). Both, whole mount samples, air-dried and Pt-C shadow-casted, and ultrathin sections of chemically fixed and resin-embedded cells were analyzed, in order to reveal the physiological/experimental alterations, intrinsic to these growth conditions. Supplementary Fig. S9 demonstrates typical morphological details from the extra- (shadow-casts) and intracellular (thin sections) milieus. At low and high temperature growth conditions and in the presence of C14 substrate bacteria produce an amorphous coat of extracellular polymers (EP; Supplementary Fig. S9: A, C; twin arrows), which apparently is not that way visible with cells grown in the presence of Tween 80. Additionally, growth with C14 substrate supports the formation of tubular protrusions of the outer membrane (Supplementary Fig. S9: A, C; arrow) and vesicles, independently on the growth temperature.

At higher temperature, cells from whole mount preparations (Supplementary Fig. S9: C, G) show a distinct trend in pleomorphic cell shape from nearly spheres to elongated rods. Equally vesicles shedding from the outer membrane appear similar, both in C14 and Tween 80 growth medium. Further under the growth conditions used it appears that cells are flagellated to a higher degree in the Tween 80 growth medium, independent on growth temperature (Supplementary Fig. S9: E, G).

Independent of the growth conditions the cytoplasm, as observed from ultrathin sections, appears similar in matrix density and amount of electron translucent inclusions or voids (fig. X: B,D,F,H), and labeling with cationic colloidal ThO_2 did not reveal the presence of acidic extracellular polymers.

SUPPLEMENTARY METHODS

RNA extraction and real-time PCR.

Nucleic acids were extracted from 100 ml of batch cultures using RNA/DNA mini extraction kit (QIAGEN) according to the manufacturer's protocol. The total RNA was eluted in a final volume of 300 µl, followed by a treatment with DNAse I (Invitrogen, Carlsbad, CA, USA) and stored at -80 °C. The quality of RNA samples was examined by 0.8%-agarose-gel electrophoresis and concentrations were determined using the NanoDrop® ND-1000 Spectrophotometer (Celbio). About 4 µg RNA was reversely transcribed by using 20 ng random primer and the SuperScript II RNase H-free Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. Briefly, total RNA was initially denatured by incubating at 70°C for 10 min. The reverse transcription reaction mix was further incubated for 5 min at 65°C and placed on ice for 2 min, followed by the addition of first-strand buffer [50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$] and 75 units of RNase inhibitor, and then incubated at 37°C for 2 min. A 200-unit aliquot of SuperScript II RNaseH-free Reverse Transcriptase was added prior to a 50 min incubation at 42°C that resulted in the transcription of RNA into complementary DNA (cDNA). The reverse transcriptase reaction was then stopped by 5 min heating the solution to 80°C. Control reactions without reverse transcriptase were conducted to verify the absence of genomic DNA. No contaminating DNA was detected in any of these reactions. Primers were designed using Primer Express 2.0 software package (Primer Express software, version 2.0 (Applied Biosystems, Foster City, Calif.). Specificity of the primer sets against *Oleispira antarctica* genes was checked using the BLAST search function http://www.ncbi.nlm.nih.gov). Designed primers were synthesized by Invitrogen (Carlsbad, CA, USA) and their specificity was experimentally confirmed against the genomic DNA of *Alcanivorax borkumensis* and *Thalassolituus oleivorans.* The abundance of transcripts was evaluated by quantitative real-time RT-PCR analysis using TaqMan® gene expression assay. The reaction was performed in an ABI 7500 Fast Real-Time PCR System thermocycler. The primers and the TaqMan® probe used are listed below.

List of oligonucleotides used for real time PCR-quantification of transcripts for alkane-hydroxylases in *O. antarctica* RB-8

Primer / Probe	Sequence (5′–3′)	PCR product size
P450_OleiF	CGCCACCTCTACGCCAGTT	150
P450_Taq	CACGCCCTTGGCCTTTAGT	
P450_OleiR	GGCCCCAAGTGCATTTTAAA	
AlkB2-1F	TAGCCATTAACGTCGGACATGAG	150
AlKB2-1_Taq	ATTCACAAAGATCCGCTTAT	
AlKB2-1R	GCATCTTCAGGTGTGGAAACG	
AlkB2-2F	ACGGGTGCAGTCTTCCTGAT	150
AlkB2-2_Taq	ACTTCATGGAGCACTACG	
AlkB2-2R	ACGAGCTTAGGCGCTTG	
AlkB2-3F	CCGGCAACAGCACCTAGAG	150
AlkB2-3_Taq	AGCCTCAACAATTAAAGGT	
AlkB2-3R	GGCCACGAATCACACCATTT	
OLEAN_C24210F	GGCGACTTGGTTAGAGCAAG	159
OLEAN_C24210R	GCCGAAGATCCAGACTTAAC	

5'-6-FAM and 3'-TAMRA-labeled TaqMan probe was obtained from PE Applied Biosystems. The experiment was performed using 10^{-1} and 10^{-2} and undiluted cDNA templates originating

from RT reaction. Mixtures for Q-PCRs and the reaction conditions were as follows: initial denaturation for 5 min at 95°C, followed by 45 cycles of 95°C for 30 s and 60°C for 30 s; each 25 µl of reaction contained 1 µl of template cDNA, 12.5 µl of 2x TaqMan® Master Mix (Applied Biosystems) and 100 nM of each primer and 6mM of TaqMan® probe. Q-PCR amplification was analyzed using an automatic setting of the baseline and threshold values. In all experiments, appropriate negative controls were subjected to the same procedure to exclude or detect any possible contamination. The standard curve method was applied to generate the absolute quantification of gene copies using a serial dilution ranging from 10^7 until 10 copies of *O. antarctica* genomic DNA quantified using a NanodropND-1000 spectrophotometer (Wilmington, DE, USA). Expression of the house keeping gene, *gyrB*, was used as reference gene to normalize tested genes in *O. antarctica* RB-8. The relative fold change in mRNA quantity was calculated for the gene of interest in each sample using the ΔΔCt method. For each RNA preparation, at least three independent real-time PCR.

Cpn60 anti-proteome isolation and analysis

Anti-proteome preparation for identification of Cpn60 client proteins

Anti-proteomes were prepared as described previously[43], with small modifications. Briefly, for immune-precipitation experiments, 50 mL *O. antarctica* cultures (in triplicates) were grown at 4 or 16ºC to mid-log phase (OD600: 0.6), harvested, and the cells resuspended in 1000 µl BugBuster solution containing 4 µl lysonase solution and then incubated for 5 min at the temperatures used for the culture incubations. Cell debris were subsequently removed by centrifugation (12000 *g*, 20 min), and 2000 µL of cleared cell lysate diluted to a final protein concentration of 1.0 mg/ml was mixed with 200 µL of protein G-sepharose 4 Fast Flow bead slurry (Amersham Pharmacia Biotech) and 40 µL of anti-Cpn60 antiserum (raised against the N-terminal peptide AAKDVLFGDSARAK, of Cpn60, and provided by SEQLAB, Göttingen, Germany). The suspension was then incubated overnight at 4ºC with gentle agitation. The beads were washed by ultrafiltration through low-adsorption hydrophilic 30000 NMWL cutoff membranes (regenerated cellulose, Amicon) three times with 50 M HEPES buffer, pH 7.0 to eliminated not bound proteins after which the beads were collected, and bound proteins were eluted after washing the beads three times with 0.1 M glycine buffer, pH 2.5. The protein solution was then neutralized with 0.5 M HEPES buffer, pH 7.0, to get a final pH close to 7.0. The protein solution thus obtained was used directly for proteome analysis as described below.

Electrophoresis and staining

Each sample aliquot (5 µg) were mixed with Rehydratation buffer (7 M urea, 2 M thiourea, 4% CHAPS, 0.5% carrier ampholites pH 3-10, 10 mM DTT) to obtain a final volume of 100 µl and applied by Cup Loading to 18 cm IPG strips, 3-10 L (General Electric, Healthcare) previously rehydrated with 340 µl of Rehydration Buffer (7 M urea, 2 M thiourea, 4% CHAPS, 0,5% carrier ampholites pH 3-10 L, 1,2% DeStreak). Four replicates were made for each sample (4ºC and 16ºC). The first dimension was run at 0.05 mA/IPG strip in the IPGphor IEF System (GE Healthcare) following a voltage increase in 5 steps: 300 V/h for 3h, linear gradient to 1000V in 6h, linear gradient to 8000 V in 3h, 8000 V/h until 43000 Vhrs were reached. After the first dimension, strips were equilibrated with SDS Equilibration Buffer (50 mM Tris pH 8.8, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, traces of bromophenol blue) containing 1% (w/v) DTT for 10 min and thereafter in SDS Equilibration Buffer containing 4% (w/v) iodoacetamide for 10 additional min. Second dimension (SDS-PAGE) was performed on 12.5% polyacrylamide gels (1 mm, 16x15 cm) using Hoefer SE 600 series electrophoresis unit (General Electric, Healthcare)**.** 1.2 µl of Novex Sharp pre-stained protein standards (Invitrogen) were pipetted into 2mm 3MM Whatman filter sheet and was positioned over the gel near de acid end of each IPG strip. Gels were run at approximately 7 mA per gel

overnight maintaining buffer temperature at 4ºC. Gels were stained with SYPRO Ruby Protein Gel Stain from Invitrogen: fixed in 10% MeOH/7% acetic acid for 30 min, incubated in SYPRO Ruby staining solution overnight, washed in 10% MeOH/7% acetic acid twice for 30 min/each, and finally, two washing steps with water for 10 min/each.

Image Analysis

Gels stained with SYPRO Ruby were scanned in a Typhoon™ 9400 Variable Mode Imager (General Electric, Healthcare) equipped with a 532 nm excitation laser (green) with the emission filter 610 nm BP 30 nm (SYPRO Ruby, ROX, EtBr) and 100 μm resolution. The photomultiplier tube setting was altered to 700 V to optimize sensitivity to background ratios. The images were analyzed with DeCyder v7.0 (General Electric, Healthcare) to enable gel-to-gel staining comparison and the reproducibility among replicates. After automatic spot detection the background was removed from each gel and the images were edited manually removing spots if the program did not define the spots properly. Spots differentially expressed were automatically excised with Spot Picker (General Electric, Healthcare).

In-gel protein digestion and sample preparation

Proteins of interest from Sypro stained 2D SDS-PAGE gels were excised automatically with an Ettan Spot Picker (GE), deposited in 96-well plates and processed automatically in a Proteineer DP (Bruker Daltonics, Bremen, Germany). The digestion protocol used was exactly the same in the three cases and was based on established protocol[38] with minor variations: gel plugs were washed with 50 mM ammonium bicarbonate and then with acetonitrile prior with further reduction with 10 mM DTT in 25 mM ammonium bicarbonate solution. Alkylation was carried out with 55 mM IAA in 50 mM ammonium bicarbonate solution. Gel pieces were then rinsed firstly with 50 mM ammonium bicarbonate and secondly with ACN, and then were dried under a stream of nitrogen. Trypsin, proteomics grade (Sigma, CA, USA) at a final concentration of 16 ng/μl in 25% ACN/50 mM ammonium bicarbonate solution was added and the digestion took place at 37 ºC for 4 h. The reaction was stopped by adding 0.5% TFA for peptide extraction. The tryptic eluted peptides were dried by speed-vacuum centrifugation and were resuspended in 4 ul of MALDI solution (30% ACN/15% isopropanol/0.1% TFA). A 0.8 μl aliquot of each peptide mixture was deposited onto a 386-well OptiTOF™ Plate (Applied Biosystems, Framingham, MA, USA) and allowed to dry at room temperature. A 0.8 μl aliquot of matrix solution (3 mg/mL α-cyano-4-hydro-cinnamic acid in MALDI solution) was then deposited onto dried digest and allowed to dry at room temperature.

MALDI peptide mass fingerprinting, MS/MS analysis and database searching

For MALDI-TOF/TOF analysis, samples were automatically acquired in an ABi 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA, USA) in positive ion reflector mode (the ion acceleration voltage was 25 kV to MS acquisition and 1 kV to MSMS) and the obtained spectra were stored into the ABi 4000 Series Explorer Spot Set Manager. PMF and MSMS fragment ion spectra were smoothed and corrected to zero baseline using routines embedded in ABi 4000 Series Explorer Software v3.6. Each PMF spectrum was internally calibrated with the mass signals of trypsin autolysis ions to reach a typical mass measurement accuracy of <25 ppm. Known trypsin and keratin mass signals, as well as potential sodium and potassium adducts (+21 Da and +39 Da) were removed from the peak list. To submit the combined PMF and MS/MS data to MASCOT software v.2.2.04 (Matrix Science, London, UK), GPS Explorer v4.9 was used, searching in a custom protein database cp188907 -20110630- (3912 sequences; 1302226 residues) that contains all possible *Oleispira antarctica* protein sequences. The following search parameters were used: enzyme, trypsin; allowed missed cleavages, 1; carbamidomethyl cystein as fixed modification by the treatment with iodoacetamide; variable modifications, oxidation of methionine; mass tolerance for precursors was set to ± 50 ppm and for MS/MS fragment ions to ± 0.3 Da. The confidence interval for protein identification was set to ≥ 95% ($p < 0.05$) and only peptides with an individual ion score above the identity threshold were considered correctly identified.

SUPPLEMENTARY REFERENCES

55 Lehnherr, H., Maguin, E., Jafri, S., & Yarmolinsky, M.B. Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J. Mol. Biol.* **233**, 414-428 (1993).

56 Durante-Mangoni, E. & Zarrilli, R. Global spread of drug-resistant Acinetobacter baumannii: molecular epidemiology and management of antimicrobial resistance. *Future Microbiol.* **6**, 407-422 (2011).

57 Blanc-Potard, A.B., Solomon, F., Kayser, J. & Groisman, E.A. The SPI-3 pathogenicity island of *Salmonella enterica*. *J. Bacteriol.* **181,** 998-1004 (1999).

58 Vasala, A., Välkkilä, M., Caldentey, J. & Alatossava, T. Genetic and biochemical characterization of the *Lactobacillus delbrueckii* subsp. *lactis* bacteriophage LL-H lysin. *Appl. Environ. Microbiol*. **61,** 4004-4011 (1995).

59 Jeong, H. *et al.* Genomic blueprint of Hahella chejuensis, a marine microbe producing an algicidal agent *Nucleic Acids Res*. **33**, 7066-7073 (2005).

60 Osborn, A.M., Bruce, K.D., Strike, P. & Ritchie, D.A. Distribution, diversity and evolution of the bacterial mercury resistance (*mer*) operon. *FEMS Microbiol. Rev.* **19**, 239-262 **(**1997).

61 Methé, B.A. *et al.* The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proc Natl Acad Sci U.S.A.* 102, 10913-10918 (2005).

62 Sabirova, J.S. *et al.*. Mutation in a "tesB-like" hydroxyacyl-coenzyme A-specific thioesterase gene causes hyperproduction of extracellular polyhydroxyalkanoates by *Alcanivorax borkumensis* SK2. *J. Bacteriol*. **188**, 8452-8459 (2006).

63 Bedzyk, L., Wang, T. & Ye., R.W. The Periplasmic Nitrate Reductase in *Pseudomonas* sp. Strain G-179 catalyzes the First Step of Denitrification *J. Bacteriol,* **181**, 2802-2806 (1999).

64 Philippot, L. Denitrification in pathogenic bacteria: for better or worse? *Trends Microbiol.* 13, 191-192 (2005).

65 Thomas, G., Coutts, G., & Merrick, M. The *glnKamtB* operon: a conserved gene pair in prokaryotes. *Trends Genet.* **16**, 11-14 (2000).

66 Canovas, D., Cases, I. & de Lorenzo, V. Heavy metal tolerance and metal homeostasis in *Pseudomonas putida* as revealed by complete genome analysis. *Environ. Microbiol.* **5**, 1242-1256 (2003).

67 De Lorenzo, V. & Neiland, J.B. Characterization of *iucA* and *iucC* genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli. J. Bacteriol* **167**, 350-355 (1986).

68 Schwyn, B. & Neilands, J.B. Universal chemical assay for the detection and determination of siderophores. *Anal. Biochem.* **160**:47-56 (1987).