

Table S1. Locus tags of putative proteins that may have a direct/indirect role in carotenoid biosynthesis in *Sphingomonadales*

	CrtE	CrtB	LOG	CrtI	CrtY	CrtG	DUF2141*	CrtZ	CrtW
1	Sala_2194 306 aa	Sala_3132 327 aa	Sala_3133 193 aa	Sala_3134 492 aa	Sala_3135 385 aa	Sala_3136 253 aa	Sala_3137 119 aa	Sala_2128 172 aa	None
2	AN936_01555 305 aa	AN936_02675 316 aa	AN936_02670 193 aa	AN936_02665 491 aa	AN936_02660 385 aa	AN936_02655 256 aa	AN936_02650 119 aa	AN936_16185 172 aa	None
3	SKP52_03550 306 aa	SKP52_02265 319 aa	SKP52_02260 193 aa	SKP52_02255 490 aa	SKP52_02250 385 aa	SKP52_02245 255 aa	SKP52_02240 119 aa	SKP52_17370 172 aa	None
4	NX02_22505 293 aa	NX02_00980 312 aa	NX02_00975 192 aa	NX02_00970 494 aa	NX02_00965 393 aa	NX02_00960 265 aa	NX02_00955 119 aa	NX02_21750 172 aa	None
5	AZE99_02430 300 aa	AZE99_09605 311 aa	AZE99_15750 189 aa	AZE99_09600 503 aa	AZE99_09595 388 aa	AZE99_09590 254 aa	AZE99_09585 120 aa	AZE99_11975 171 aa	None
6	LH20_15970 307 aa	LH20_01385 319 aa	LH20_01380 193 aa	LH20_01375 491 aa	LH20_01370 386 aa	LH20_01360 252 aa	LH20_01355 119 aa	LH20_13675 172 aa	None
7	AOA14_02810 299 aa	AOA14_09220 323 aa	AOA14_09215 193 aa	AOA14_09210 490 aa	AOA14_09205 385 aa	AOA14_09195 262 aa	AOA14_09190 119 aa	AOA14_05275 172 aa	None
8	SLG_30760 300 aa	SLG_31750 321 aa	SLG_31740 196 aa	SLG_31730 492 aa	SLG_31720 387 aa	SLG_31700 249 aa	SLG_31690 120 aa	SLG_38860 174 aa	None
9	TS85_16025 299 aa	TS85_20435 311 aa	TS85_20440 193 aa	TS85_20445 491 aa	TS85_20450 393 aa	TS85_20465 262 aa	TS85_20470 119 aa	TS85_19965 166 aa	None
10	ASE13_12915 302 aa	ASE13_12025 318 aa	ASE13_12030 193 aa	ASE13_12035 492 aa	ASE13_12040 386 aa	ASE13_12060 249 aa	None	ASE13_11100 166 aa	None
11	Sphi55_0001 302 aa	Sphi55_0002 307 aa	Sphi55_0003 193 aa	Sphi55_0004 492 aa	Sphi55_0005 385 aa	Sphi55_0007 265 aa	Sphi55_0008 119 aa	Sphi55_0006 166 aa	Sphi55_0009 261 aa
12	SP6_03_00020 299 aa	SP6_42_00220 316 aa	SP6_42_00230 193 aa	SP6_42_00180 491 aa	SP6_42_00170 390 aa	SP6_42_00290 271 aa	SP6_42_00300 119 aa	SP6_29_00510 179 aa	None
13	Sasta_0001 301 aa	Sasta_0002 312 aa	Sasta_0003 193 aa	Sasta_0004 487 aa	Sasta_0005 380 aa	None	Sasta_0008 120 aa	Sasta_0006 176 aa	Sasta_0009 258 aa
14	Sphch_1308 302 aa	Sphch_1941 317 aa	Sphch_1942 193 aa	Sphch_1943 493 aa	Sphch_1944 383 aa	Sphch_2837 241 aa	Sphch_2838 119 aa	Sphch_2465 167 aa	None
15	K663_05560 302 aa	K663_01410 319 aa	K663_01405 193 aa	K663_01400 493 aa	K663_01395 383 aa	K663_13655 242 aa	K663_13650 119 aa	K663_01820 170 aa	None

16	EP837_00635 302 aa	EP837_01648 319 aa	EP837_01647 193 aa	EP837_01646 493 aa	EP837_01645 383 aa	EP837_01002 242 aa	EP837_01001 119 aa	EP837_01682 170 aa	None
17	TZ53_03080 306 aa	TZ53_10675 317 aa	TZ53_10670 193 aa	TZ53_10665 493 aa	TZ53_10660 386 aa	TZ53_01450 237 aa	TZ53_01445 119 aa	TZ53_12060 169 aa	None
18	Saro_2253 303 aa	Saro_1814 323 aa	Saro_1815 193 aa	Saro_1816 492 aa	Saro_1817 398 aa	Saro_0236 236 aa	Saro_0235 121 aa	Saro_1168 172 aa	None
19	WG74_10145 302 aa	WG74_04285 324 aa	WG74_04290 195 aa	WG74_04295 500 aa	WG74_04300 403 aa	WG74_03350 240 aa	WG74_03345 122 aa	WG74_11370 172 aa	None
20	A6F68_02109 304 aa	A6F68_02787 341 aa	A6F68_02786 193 aa	A6F68_02785 497 aa	A6F68_02784 393 aa	A6F68_00557 237 aa	A6F68_00556 121 aa	A6F68_01016 172 aa	None
21	A6F65_01603 303 aa	A6F65_00929 344 aa	A6F65_00930 201 aa	A6F65_00931 508 aa	A6F65_00932 402 aa	A6F65_00501 239 aa	A6F65_00500 122 aa	A6F65_02209 166 aa	A6F65_02210 262 aa
22	AMC99_01057 301 aa	AMC99_00377 341 aa	AMC99_00378 193 aa	AMC99_00379 523 aa	AMC99_00380 394 aa	AMC99_02463 244 aa	AMC99_02464 121 aa	AMC99_01700 173 aa	None
23	WYH_00402 301 aa	WYH_03017 341 aa	WYH_03018 193 aa	WYH_03019 501 aa	WYH_03020 393 aa	WYH_02232 238 aa	WYH_02233 121 aa	WYH_00011 172 aa	None
24	Altish_0001 301 aa	Altish_0002 315 aa	Altish_0003 193 aa	Altish_0004 520 aa	Altish_0005 402 aa	Altish_0007 240 aa	Altish_0008 121 aa	Altish_0006 166 aa	Altish_0009 255 aa
25	ELI_05575 302 aa	ELI_09895 341 aa	ELI_09890 193 aa	ELI_09885 501 aa	ELI_09880 400 aa	ELI_12610 240 aa	ELI_12615 121 aa	ELI_03325 166 aa	ELI_03320 255 aa
26	CP97_10495 301 aa	CP97_13035 348 aa	CP97_13030 193 aa	CP97_13025 498 aa	CP97_13020 398 aa	CP97_01785 241 aa	CP97_01790 121 aa	CP97_11790 172 aa	None
27	HMPREF9718_ 03593 301 aa	HMPREF9718_ 02587 329 aa	HMPREF9718_ 02588 193 aa	HMPREF9718_ 02589 493 aa	HMPREF9718_ 02590 382 aa	HMPREF9718_ 03515 246 aa	HMPREF9718_ 03516 119 aa	HMPREF9718_ 02777 172 aa	None
28	PP1Y_AT24739 301 aa	PP1Y_AT8258 317 aa	PP1Y_AT8268 193 aa	PP1Y_AT8278 492 aa	PP1Y_AT8289 408 aa	PP1Y_AT36759 240 aa	PP1Y_AT36765 120 aa	PP1Y_AT11105 172 aa	None
29	JI59_00550 301 aa	JI59_14775 317 aa	JI59_14780 193 aa	JI59_14790 492 aa	JI59_14795 408 aa	JI59_11990 240 aa	JI59_11985 120 aa	JI59_06385 172 aa	None
30	A9D12_11730 303 aa	A9D12_07620 321 aa	A9D12_07625 193 aa	A9D12_07635 520 aa	A9D12_07640 425 aa	A9D12_05235 257 aa	A9D12_05230 121 aa	A9D12_01070 179 aa	A9D12_01075 267 aa
31	A7E77_04905 300 aa	A7E77_09430 314 aa	A7E77_09425 193 aa	A7E77_09415 491 aa	A7E77_09410 388 aa	A7E77_06785 270 aa	A7E77_06790 119 aa	A7E77_07315 171 aa	None
32	BJP26_00185 300 aa	BJP26_05345 314 aa	BJP26_05340 193 aa	BJP26_05325 491 aa	BJP26_05320 388 aa	BJP26_02355 270 aa	BJP26_02360 119 aa	BJP26_02830 171 aa	None

33	AM2010_812 302 aa	AM2010_155 346 aa	AM2010_156 193 aa	AM2010_158 520 aa	AM2010_159 393 aa	AM2010_2362 240 aa	AM2010_1236 120 aa	AM2010_1583 172 aa	None
34	AB433_06870 312 aa	AB433_09240 328 aa	AB433_12870 189 aa	AB433_09235 496 aa	AB433_09230 398 aa	AB433_15915 250 aa	AB433_15920 119 aa	AB433_00865 174 aa	None
35	MC45_16265 296 aa	MC45_14215 316 aa	MC45_14210 193 aa	MC45_14760 491 aa	MC45_14765 388 aa	MC45_14775 274 aa	MC45_14780 119 aa	MC45_05470 155 aa	MC45_05465 238 aa
36	BW41_00905 296 aa	BW41_00781 315 aa	BW41_00782 193 aa	BW41_03520 491 aa	BW41_03521 388 aa	BW41_03523 272 aa	BW41_03524 119 aa	BW41_00240 155 aa	BW41_00241 236 aa
37	SJA_C1-33430 305 aa	SJA_C1-01890 317 aa	SJA_C1-01590 193 aa	SJA_C1-01600 493 aa	SJA_C1-01610 386 aa	SJA_C1-28960 241 aa	SJA_C1-28950 119 aa	SJA_C1-06860 169 aa	None
38	G432_12610 305 aa	None	G432_07845 193 aa	G432_07840 492 aa	G432_07835 394 aa	None	None	None	None
39	ATN00_08710 303 aa	None	ATN00_06065 193 aa	ATN00_06060 493 aa	None	None	None	None	None
40	Swit_0022 297 aa	None	Swit_3199 196 aa	None	Swit_3200 380 aa	None	None	None	None
41	SGRAN_0972 307 aa	None	SGRAN_3508 193 aa	None	None	None	None	None	None

Serial numbers in the first column are the same as in Tables 1 and S2.

*Number of aa is counted from the first conserved substitution (V/L/I).

Sequences for locus tags in serial numbers 11, 13, and 24 are provided after Fig. S1.

Table S2. Comparison of the features of *crt* and *log* ORFs identified in strains of *Sphingomonadales*

	Species (strain)	<i>crtY</i> Length (Start. .Stop)	<i>crtI</i> Length (Start. .Stop)	Gap/overlap between <i>crtY</i> and <i>crtI</i>	Gap/overlap between <i>crtI</i> and <i>log</i>	<i>log</i> Length (Start. .Stop)	<i>crtB</i> Length (Start. .Stop)	Gap/overlap between <i>log</i> and <i>crtB</i>	<i>crtG</i> Length (Start. .Stop)	<i>crtZ</i> Length (Start. .Stop)	<i>crtW</i> Length (Start. .Stop)
1	<i>Sphingopyxis</i> <i>alaskensis</i> (RB2256)	1158 (ATG. .TGA)	1479 (ATG. .TAA)	Overlap (4 bp)	Gap (35 bp)	582 (ATG. .TGA)	984 (ATG. .TGA)	Overlap (1 bp)	762 (ATG. .TGA)	519 (ATG. .TAA)	None
2	<i>Sphingopyxis</i> <i>macrogoltabida</i> (EY-1)	1158 (ATG. .TGA)	1476 (ATG. .TAA)	Overlap (4 bp)	Gap (98 bp)	582 (ATG. .TAA)	951 (ATG. .TGA)	Overlap (1 bp)	771 (GTG. .TGA)	519 (ATG. .TAG)	None
3	<i>Sphingopyxis</i> <i>fribergensis</i> (Kp5.2)	1158 (ATG. .TGA)	1473 (ATG. .TGA)	Overlap (4 bp)	Gap (116 bp)	582 (ATG. .TAA)	960 (ATG. .TGA)	Overlap (1 bp)	768 (ATG. .TGA)	519 (ATG. .TAA)	None
4	<i>Sphingomonas</i> <i>sanxanigenens</i> (DSM 19645)	1182 (ATG. .TGA)	1485 (ATG. .TGA)	Overlap (4 bp)	Gap (28 bp)	579 (ATG. .TGA)	939 (GTG. .TGA)	Overlap (4 bp)	798 (GTG. .TGA)	519 (ATG. .TGA)	None
5	<i>Sphingorhabdus</i> sp. (M41)	1167 (ATG. .TGA)	1512 (ATG. .TGA)	Gap (13 bp)	<i>log</i> is off- located	570 (TTG. .TGA)	936 (ATG. .TAG)	<i>crtB</i> is separated from <i>crtI</i> by 75 bp	765 (ATG. .TGA)	516 (ATG. .TAG)	None
6	<i>Sphingopyxis</i> sp. (113P3)	1161 (ATG. .TGA)	1476 (ATG. .TAG)	Overlap (4 bp)	Gap (35 bp)	582 (ATG. .TGA)	960 (ATG. .TGA)	Overlap (1 bp)	759 (ATG. .TGA)	519 (ATG. .TAG)	None
7	<i>Sphingopyxis</i> <i>terrae</i> (NBRC 15098)	1158 (ATG. .TGA)	1473 (ATG. .TGA)	Overlap (4 bp)	Overlap (1 bp)	582 (ATG. .TAA)	972 (ATG. .TGA)	Overlap (1 bp)	789 (ATG. .TGA)	519 (ATG. .TGA)	None
8	<i>Sphingobium</i> sp. (SYK-6)	1164 (ATG. .TGA)	1479 (ATG. .TGA)	Overlap (4 bp)	Overlap (4 bp)	591 (ATG. .TGA)	966 (GTG. .TGA)	Gap (36 bp)	750 (ATG. .TGA)	525 (ATG. .TAG)	None
9	<i>Sphingomonas</i> <i>hengshuiensis</i> (WHSC-8)	1182 (ATG. .TGA)	1476 (ATG. .TAA)	Overlap (4 bp)	Gap (1 bp)	582 (TTG. .TGA)	936 (ATG. .TAG)	Gap (17 bp)	789 (GTG. .TGA)	501 (ATG. .TGA)	None

10	<i>Sphingomonas</i> sp. (Root241)	1161 (ATG. . TGA)	1479 (ATG. . TGA)	Overlap (4 bp)	Overlap (4 bp)	582 (ATG. . TGA)	957 (GTG. . TAA)	Overlap (4 bp)	750 (ATG. . TAG)	501 (ATG. . TGA)	None
11	<i>Sphingomonas</i> sp. (ATCC 31555)	1158 (ATG. . TGA)	1479 (ATG. . TGA)	Overlap (1 bp)	Gap (171 bp)	582 (ATG. . TAA)	924 (ATG. . TAA)	Overlap (1 bp)	798 (ATG. . TGA)	501 (ATG. . TGA)	786 (ATG. . TGA)
12	<i>Sphingomonas paucimobilis</i> (NBRC 13935)	1173 (ATG. . TGA)	1476 (GTG. . TAA)	Overlap (4 bp)	Gap contains 3 ORFs	582 (ATG. . TGA)	951 (GTG. . TAG)	Overlap (4 bp)	816 (GTG. . TGA)	540 (ATG. . TAA)	None
13	<i>Sphingomonas astaxanthinifaciens</i> (DSM 22298)	1143 (ATG. . TGA)	1464 (GTG. . TGA)	Overlap (4 bp)	<i>log</i> is off- located	582 (GTG. . TAG)	939 (GTG. . TGA)	<i>crtB</i> overlaps with <i>crtI</i> by 4 bp	None	531 (ATG. . TGA)	777 (ATG. . TGA)
14	<i>Sphingobium chlorophenicum</i> (L-1)	1152 (ATG. . TGA)	1482 (ATG. . TAG)	Gap (13 bp)	Gap (1 bp)	582 (ATG. . TGA)	954 (ATG. . TGA)	Gap (363 bp)	726 (ATG. . TGA)	504 (ATG. . TAG)	None
15	<i>Sphingobium</i> sp. (MI1205)	1152 (ATG. . TGA)	1482 (ATG. . TGA)	Overlap (1 bp)	Overlap (4 bp)	582 (GTG. . TGA)	960 (ATG. . TGA)	Gap (40 bp)	729 (ATG. . TGA)	513 (ATG. . TGA)	None
16	<i>Sphingobium</i> sp. (EP60837)	1152 (ATG. . TGA)	1482 (ATG. . TGA)	Overlap (1 bp)	Overlap (4 bp)	582 (TTG. . TAG)	960 (ATG. . TGA)	Gap (37 bp)	729 (ATG. . TGA)	513 (ATG. . TAG)	None
17	<i>Sphingobium</i> sp. (YBL2)	1161 (ATG. . TGA)	1482 (ATG. . TGA)	Gap (13 bp)	Gap (12 bp)	582 (ATG. . TGA)	954 (ATG. . TGA)	Gap (354 bp)	714 (ATG. . TAG)	510 (ATG. . TGA)	None
18	<i>Novosphingobium aromaticivorans</i> (DSM 12444)	1197 (ATG. . TGA)	1479 (ATG. . TGA)	Overlap (4 bp)	Overlap (4 bp)	582 (GTG. . TGA)	972 (ATG. . TGA)	Gap (4 bp)	711 (ATG. . TGA)	519 (ATG. . TAG)	None
19	<i>Citromicrobium</i> sp. (JL477)	1212 (ATG. . TGA)	1503 (ATG. . TGA)	Overlap (8 bp)	Gap (299 bp)	588 (ATG. . TGA)	975 (GTG. . TAG)	Gap (32 bp)	723 (ATG. . TGA)	519 (ATG. . TAG)	None
20	<i>Altererythrobacter dongtanensis</i> (KCTC 22672)	1182 (ATG. . TGA)	1494 (TTG. . TGA)	Overlap (4 bp)	Overlap (4 bp)	582 (GTG. . TGA)	1026 (CTG. . TGA)	Gap (61 bp)	714 (ATG. . TGA)	519 (ATG. . TAG)	None

21	<i>Altererythrobacter namhicola</i> (JCM 16345)	1209 (ATG. .TGA)	1527 (ATG. .TGA)	Overlap (4 bp)	Overlap (1 bp)	606 (ATG. .TGA)	1035 (ATG. .TAA)	Gap (44 bp)	720 (ATG. .TGA)	501 (ATG. .TGA)	789 (ATG. .TGA)
22	<i>Altererythrobacter epoxidivorans</i> (CGMCC 1.7731)	1185 (ATG. .TGA)	1572 (ATG. .TGA)	Gap (6 bp)	Gap (8 bp)	582 (ATG. .TAA)	1026 (CTG. .TGA)	Gap (61 bp)	735 (ATG. .TGA)	522 (ATG. .TAA)	None
23	<i>Altererythrobacter atlanticus</i> (26DY36)	1182 (ATG. .TGA)	1506 (ATG. .TGA)	Overlap (4 bp)	Overlap (4 bp)	582 (GTG. .TAA)	1026 (CTG. .TGA)	Gap (43 bp)	717 (ATG. .TGA)	519 (ATG. .TGA)	None
24	<i>Altererythrobacter ishigakiensis</i> (NBRC 107699)	1209 (ATG. .TGA)	1563 (ATG. .TGA)	Overlap (4 bp)	Overlap (4 bp)	582 (CTG. .TGA)	948 (GTG. .TAG)	Overlap (4 bp)	723 (ATG. .TGA)	501 (ATG. .TAA)	768 (ATG. .TGA)
25	<i>Erythrobacter litoralis</i> (HTCC2594)	1203 (ATG. .TGA)	1506 (ATG. .TGA)	Overlap (4 bp)	Overlap (4 bp)	582 (ATG. .TGA)	1026 (CTG. .TAG)	Gap (56 bp)	723 (ATG. .TAG)	501 (ATG. .TGA)	768 (ATG. .TGA)
26	<i>Erythrobacter atlanticus</i> (s21-N3)	1197 (ATG. .TGA)	1497 (ATG. .TAG)	Overlap (4 bp)	Gap (98 bp)	582 (ATG. .TGA)	1047 (GTG. .TGA)	Gap (35 bp)	726 (ATG. .TGA)	519 (ATG. .TGA)	None
27	<i>Sphingobium yanoikuyae</i> (ATCC 51230)	1149 (GTG. .TGA)	1482 (ATG. .TGA)	Overlap (1 bp)	Gap (31 bp)	582 (ATG. .TGA)	990 (ATG. .TGA)	Gap (54 bp)	741 (ATG. .TGA)	519 (ATG. .TAG)	None
28	<i>Novosphingobium</i> sp. (PP1Y)	1227 (ATG. .TGA)	1479 (ATG. .TGA)	Overlap (4 bp)	Gap (435 bp, contains an ORF)	582 (ATG. .TGA)	954 (ATG. .TAG)	Gap (37 bp)	723 (TTG. .TGA)	519 (ATG. .TGA)	None
29	<i>Novosphingobium pentaromativorans</i> (US6-1)	1227 (ATG. .TGA)	1479 (ATG. .TGA)	Overlap (4 bp)	Gap (435 bp, contains JI59_14785)	582 (ATG. .TGA)	954 (ATG. .TAG)	Gap (37 bp)	723 (TTG. .TGA)	519 (ATG. .TGA)	None
30	<i>Porphyrobacter neustonensis</i> (DSM 9434)	1278 (ATG. .TGA)	1563 (ATG. .TGA)	Overlap (4 bp)	Gap (432 bp, contains A9D12_07630)	582 (ATG. .TAG)	966 (ATG. .TGA)	Gap (11 bp)	774 (ATG. .TGA)	540 (ATG. .TAA)	804 (ATG. .TGA)

31	<i>Sphingomonas</i> sp. (NIC1)	1167 (ATG. .TAG)	1476 (ATG. .TGA)	Gap (16 bp)	Gap (832 bp, contains A7E77_09425)	582 (ATG. .TGA)	945 (CTG. .TAG)	Overlap (11 bp)	813 (ATG. .TGA)	516 (ATG. .TAA)	None
32	<i>Sphingomonas melonis</i> (TY)	1167 (ATG. .TAG)	1476 (ATG. .TGA)	Gap (16 bp)	Gap (1035 bp, contains BJP26_05330 BJP26_05335)	582 (ATG. .TGA)	945 (CTG. .TAG)	Overlap (11 bp)	813 (ATG. .TGA)	516 (ATG. .TAA)	None
33	<i>Altererythrobacter marensis</i> (KCTC 22370)	1182 (ATG. .TGA)	1563 (ATG. .TGA)	Overlap (4 bp)	Gap (510 bp, contains AM2010_157)	582 (ATG. .TGA)	1041 (CTG. .TAG)	Gap (61 bp)	723 (ATG. .TAG)	519 (ATG. .TGA)	None
34	<i>Croceicoccus naphthovorans</i> (PQ-2)	1197 (ATG. .TGA)	1491 (ATG. .TGA)	Gap (11 bp)	<i>log</i> is off- located	570 (TTG. .TAA)	987 (GTG. .TGA)	<i>crtB</i> is separated from <i>crtI</i> by 4 bp	753 (ATG. .TGA)	525 (ATG. .TGA)	None
35	<i>Sphingomonas taxi</i> (ATCC 55669)	1167 (ATG. .TAA)	1476 (ATG. .TAG)	Gap (25 bp)	<i>log</i> and <i>crtB</i> are off-located	582 (ATG. .TAG)	951 (CTG. .TAA)	Overlap (11 bp)	825 (ATG. .TGA)	468 (GTG. .TGA)	717 (ATG. .TGA)
36	<i>Sphingomonas</i> sp. (RIT328)	1167 (ATG. .TGA)	1476 (ATG. .TAG)	Gap (23 bp)	<i>log</i> and <i>crtB</i> are off-located	582 (ATG. .TAG)	948 (CTG. .TGA)	Overlap (11 bp)	819 (ATG. .TGA)	468 (ATG. .TGA)	711 (ATG. .TAG)
37	<i>Sphingobium japonicum</i> (UT26S)	1161 (ATG. .TGA)	1482 (ATG. .TAA)	Gap (13 bp)	Gap (9 bp)	582 (ATG. .TGA)	954 (ATG. .TGA)	<i>crtB</i> is off- located	726 (ATG. .TGA)	510 (ATG. .TGA)	None
38	<i>Sphingomonas</i> sp. (MM-1)	1185 (ATG. .TGA)	1479 (ATG. .TGA)	Overlap (4 bp)	Gap (51 bp)	582 (ATG. .TGA)	None		None	None	None
39	<i>Sphingobium baderi</i> (DE-13)	None	1482 (ATG. .TGA)		Overlap (4 bp)	582 (GTG. .TAA)	None		None	None	None

40	<i>Sphingomonas wittichii</i> (RW1)	1143 (ATG. .TGA)	None	<i>crtY</i> and <i>log</i> overlap by 4 bp	591 (ATG. .TAG)	None		None	None	None
41	<i>Sphingopyxis granuli</i> (TFA)	None	None		582 (ATG. .TGA)	None		None	None	None
Serial numbers in the first column are the same as in Tables 1 and S1										

Overlaps among *crtB*, *log*, and *crtI*

In bacteria that contain *crt* loci, *crtB* usually occurs immediately downstream of *crtI*, and in some cases their reading frames overlap. Whereas the two ORFs overlapped by 4 bp in strain DSM 22298, they were separated by 4 and 75 bp in strains PQ2 and M41, respectively (Table S2). Surprisingly, in the genomes of 24 strains of *Sphingomonadales*, an unusual ORF was present between *crtI* and *crtB* (Figs. 1 and S1). In nine of the 24 strains, the reading frames of *crtB* and *log* overlapped by 1–4 bp, with 1 bp being the most common overlap (Table S2). In 15 strains where *crtB* and *log* occur together without an overlap of the reading frames, their separation range was 4–363 bp (Figs. 1 and S1; Table S2), indicating that *crtB* could be transcribed independently in some of them. Furthermore, in 11 of the 24 strains, the reading frames of *crtI* and *log* overlapped by 1–4 bp, with 4 bp being the most common overlap (Table S2). In 13 strains where *crtI* and *log* occur together without an overlap of the reading frames, their separation range was 1–299 bp (Figs. 1 and S1; Table S2), indicating that *log* could be transcribed independently in some of them. Interestingly, in 11 of the 24 strains, the reading frame of *crtY* overlapped with that of *crtI*, and the reading frame of *crtI* overlapped with that of *log*. However, only in 3 (*Sphingopyxis terrae* NBRC 15098, *Sphingomonas* sp. Root241, and *Altererythrobacter ishigakiensis* NBRC

107699) of the 24 strains, the reading frame of *crtY/crtI/log* overlapped with that of *crtI/log/crtB*, respectively (thus forming a putative *crtY–crtI–log–crtB* operon).

In the genomes of 6 strains, the *crtY–crtI* and *log–crtB* pairs were separated by 1–2 ORFs (Figs. 1 and S1; Table S2). In two of these strains (*Sphingomonas* sp NIC1 and *Sphingomonas melonis* TY), the reading frames of *log* and *crtB* overlapped by 11 bp, and in the other four, they were separated by 11–61 bp. In the genomes of *Sm. taxi* ATCC 55669 and *Sphingomonas* sp. RIT328, the *log–crtB* pair was located away from the *crt* locus (Figs. 1 and S1) and their reading frames overlapped by 11 bp (Table S2). Furthermore, in the genomes of *Sb. japonicum* UT26S, *Sphingomonas* sp. MM-1, and *Sb. baderi* DE-13, *log* was found immediately downstream of *crtI* (Figs. 1 and S1). Whereas the reading frames of *crtI* and *log* overlapped by 4 bp in *Sb. baderi* DE-13, they were separated by 9 and 51 bp in the other two strains.

Overlaps among *crtZ* and *crtW*

In four strains (JCM 16345, NBRC 107699, HTCC2594, and DSM 9434, which were members of *Erythrobacteraceae*), *crtW* was located immediately upstream of *crtZ* (Figs. 1 and S1), and their reading frames overlapped by 4 bp. Conversely, in two strains (ATCC 55669 and RIT328, which were members of *Sphingomonas*), *crtW* was located immediately downstream of *crtZ* (Figs. 1 and S1). Whereas the reading frames of *crtW* and *crtZ* overlapped by 4 bp in strain ATCC 55669, they were separated by 7 bp in strain RIT328. Therefore, it is likely that there is transcriptional and/or translational coupling of *crtW* and *crtZ* in these strains.

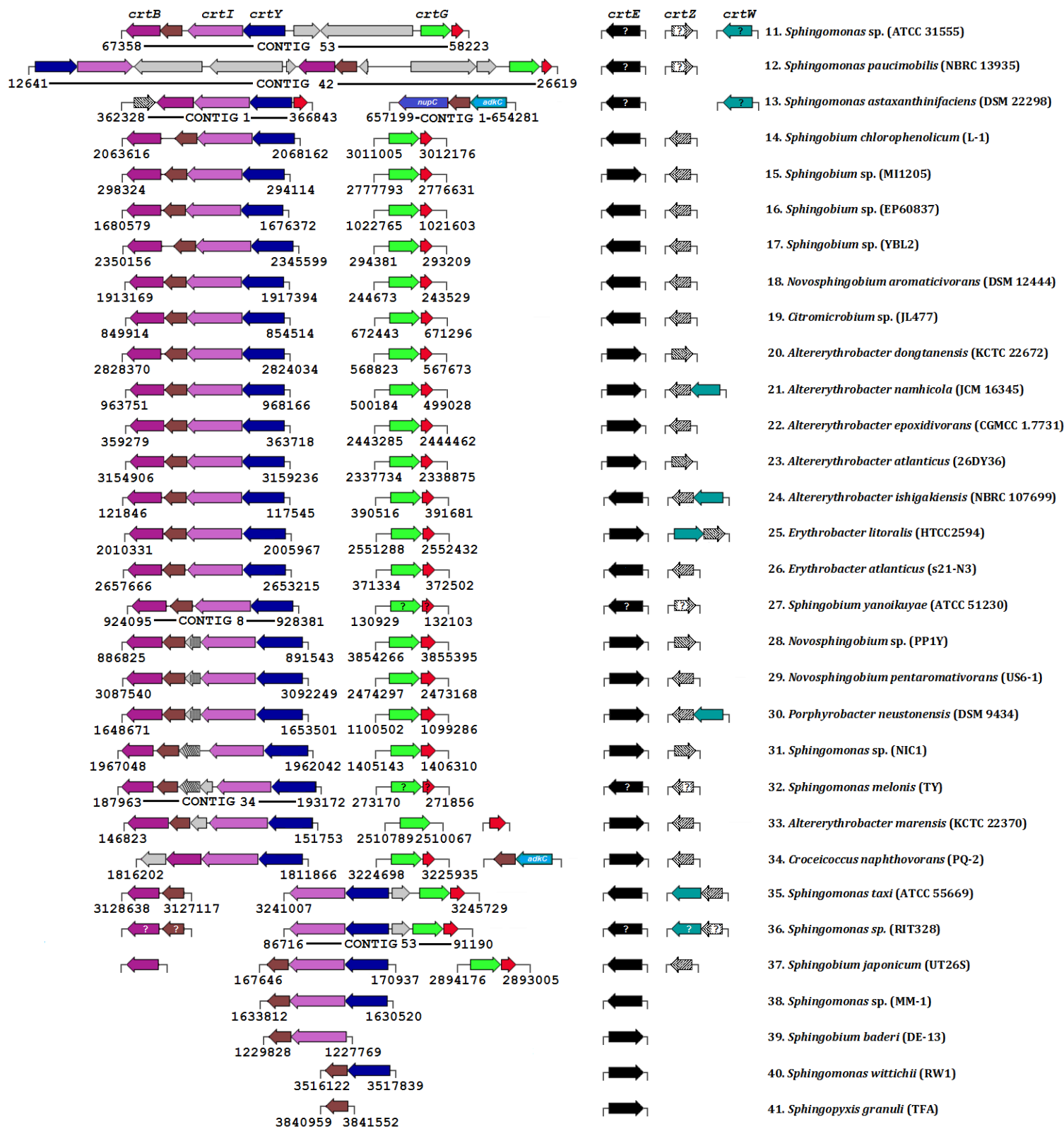


Fig. S1. Comparison of *crt* ORFs/loci in 31 strains of *Sphingomonadales*. Serial numbers of the bacteria are the same as in Table 1. The *crt* ORF that each color-coded or shaded arrow represents is indicated on the top. The brown and red arrows represent ORFs encoding the putative LOG and DUF2141 proteins, respectively. The gray arrows represent the *separating* ORFs that are unlikely to be involved in carotenoid biosynthesis. Numbers at the beginning and end of gene clusters indicate the coordinates within each genome or contig. In the case of *crtE*, *crtZ*, and *crtW*, arrows pointing to the right indicate that they were located on the *plus* strand, while those pointing to the left indicate that they were located on the *minus* strand. In serial numbers 14, 22, 23, and 28, *crtY* was located on the *minus* strand, but *crtG* was located on the *plus* strand. In serial numbers 15, 16, 17, and 20, *crtY* was located on the *plus* strand, but *crtG* was located on the *minus* strand. In serial numbers 18, 19, 21, 29, 30, and 33, *crtY* and *crtG* were located on the *minus* strand. In serial numbers 24, 25, 26, 31, and 34, *crtY* and *crtG* were located on the *plus* strand. In *Sphingomonas taxi* (serial number 35), *crtY* was located on the *minus* strand, but *crtB* was located on the *plus* strand. In *Sphingobium japonicum* (serial number 37), *crtY* and *crtG* were located on the *minus* strand, but *crtB* was located on the *plus* strand. The uncertainty of the location of some of the ORFs on the *plus* or *minus* strands in the draft genomes of serial numbers 11, 12, 13, 27, 32, and 36 is denoted by a '?' within the respective arrows.

Sphingomonas sp. ATCC 31555 locus tags and sequences

Sphi55_0001, 302 aa, WP_019370787

MASVTHASLALLESALRQVAEEIDAQFDRLLLEVPDDPRADLYRAMRHAAIIGGGKRLRPLLVFATAQLFAVD
KSCAARVATALECIHVYSLIHDDLPMDDDDMRGKPTVHKAFFEATAILAGDCLHDLAFELLAHEATHP
DPFVVRVDLISELARASGPSMAGGQMMDLAEKASFDLATVTRLQAMKTGALISCAVECAAAILGRVAPEG
RTGLRGYARDIGLAFQIVDDILD AEGDEALVGKKLGKGDAGKETFLSLLGIDRAREQARMLVDQAITHL
HAYGPEADLLRDIARFTLERDR

Sphi55_0002, 307 aa, WP_019370947

MRFPGRDAIVATAGESIARGSRSF AAASKLFDPKTRERAWLLYAWCRRCDIADGQDHGHGTMVVVEDPQA
RLAEVRERTAAALAGEWVGDAAFDALRIVAQETGMPHRFAWDLIDGFALDADGWQPTNEDDLLRYCYHVA
GVVGCMMAVVMGVQPEDEAVLDRACDLGLAFQLANIARDIGEDAEAGRCYLPAAWLAEVGLTADNPLADR
SALAILGTRLADRAEAYEISARGGAAALSFRSAWAVLAAADIYGGIARKVRARGAGAWESRTTTS GGEKL
AAIGGAWIAAWKRAPTRPAGLWTRPR

Sphi55_0003, 193 aa, WP_019370946

MKRIAVYCGSASPANPVYIEAARTLGRTLAERGIGVVYGGGRLGLMGAVADAALAEAGGEVIGVIP TLLVN
AEVAHRGLTSLEVCETMHERKARFTELADGFVNLPGGTGTMDLWEAMSWAQLGYHADPIGLLNIAGYYD
KLVEFWEHMGTVGFVRSQHQLLLVDSTIDGLLAKMADAQPIVPIAQLRRDQL

Sphi55_0004, 492 aa, WP_019370945

MRRAAVIGAGFGGLALAIRLQSAGVDTTLIEARDKPGGRAYWDRDGFIFDGGPTVITDPDALKELWRLS
GHDISEDVTCLKPVTPFYRLSWPDGTTFDYTNDDAVLAGEIAKLDASDIAGYRKFLYSEVYQEGYVKLG
AKAFDFKSMKKAAPALMQNAWRSVYSIVSSFVKNEKLRREALSFHTLLVGGNPMATSSIYALIHKLERD
GGVWVAEGGTNRLVAGMIRHFERLGGAVRIGDPVDEILTLGDRVTGLRTRKSGFELEVD AVASNADIVHSY
RDLLRNSRSARRTAGRLVNKRFSPSLFVVHFGLRGTFPDI PHHMLFGPRYKGLLDDIYTHGVISEDFSL
YLHHPASDASMAPEGHSTFYVLA PVPHLGKFPADWSEVGP ILEQRILAEVQRRLIPDLADRI VTKFHYA
PTDFRDDLAAHHGSGFSLEPILTQSAWFRVHNRDDAISNLYFVGAGTHPGAGIPGVVGS AKATAELMLGE
MV

Sphi55_0005, 385 aa, WP_019370944

MPSTISCDVAIVGGGLAGGLIALAIRKKKPHCSVRLIEGSARI GGNHLWSFFADDVAPEHRWLAAPLISH
GWSNYDVAFPGHSRTL PARYYSVESEHFDRVVREALGPKGLLLGREAADITPNAVALADGDLVEADGVID
CRGVGDTSALQLGWQKFHGAELALALNTVQRPLIMDATVPQIDGYRFVYCLPFSPTRMFVEDTYYS DTP
SIDLLATGARIEAYADAAGWQVERVLREEVGS LVPVAMGGDHAAYWAAGGRGI AKAGMKAGLFHPMTGYSF
PDAVRTAALIAEEKDYSGEALHQMLYGYSRGLWRKRGFYRMLAAMLFKAAEPEERYVLERFYRLDAGLI
QRFYAGQSTWGDVRVLSGKPPVPIGRALSAINR

Sphi55_0006, 166 aa, WP_019369087

MSLPTGIALFVVTVLGMELFAYAAHRWIMHGPGWFLHASHHRPRTGNWEWNDLYAAIFAVPSFVLLL GGA
HLGWWTGAWIGAGIAAYGMIYFGFHD LIVHKRVRHRYIPRSSYMKRIVQAHRLHHAVHTRKGTVSFGFL
YAPRPEVLKAQLRSSGALTQSSTIPD

Sphi55_0007, 265 aa, WP_019370940

MDIALALLTSAVVMTVIVGVRYLAASAGFALATRLRHPGLYAGLDRQIRKEIGWSLASTAIYGI PAGIVA
WGWAQGWTRIYDDVHAWPLWYLPVSAFLYLFLHDSWFYWTHR AMHPALFRRYHAVHHASRPPTAWAAM
SFHWGEALS GAIVIPLLVFVPIHVAALGLVLTIMTVMGVTNHMGWEIFPAWLWRGPVGAWVITASHHQR
HHERYGCNYGLYFRFWDRLCGTDQGLGQFQGT PRARRAGSVPDTGRRIATGDAGD

Sphi55_0008, 128 aa, WP_026009371

MGGGSPLATLEIDFVKLRSSRGLLQVC IAPAPSVFPDCRDGGGAIKRSIPASAGQLRVTGLAPGDYAVAV
IHDANGNGKLDTMMGI PREGFGFSRNP AIGFGPPRF TAAQFPLTGGSDRQEV MRMYLL

Sphi55_0009, 261 aa, WP_037513166

MTRLEAPRDEGRQGRCGIDAQDKRVAQRGRWIAAAIGGAWLAVHVGCIFFWRWSIAAAPLAVFIVLLQAW
LSTGLFIVAHDCMHGSLAPGRPRWNAVVGALCLGAYAGLSYAVLLPKHHAHHAAPGSADDPDFHADAPQR
ALPWFSGSFFRTYYTHGQLARITAAAIYVLLL GASLLNIVVFWAVPALLALGQLFLFGTYLPHRHDGRPFA
DKHHARSNTLGPMLSLITCFHFH GAYHHEHHLSPGTPWWRLPALRRERLARR

Sphingomonas astaxanthinifaciens DSM 22298 locus tags and sequences

Sasta_0001, 301 aa, WP_029942171

MASVAEATDLLSEAKRIAAGVDALFEQALADRCDRDRLCAAMRHAGIGGGKRLRPLITVAAARLFAIDEDRALRAGAAIEAIHVYSLIHDDLPCMDDDDLRRGKPTVHKAFDEATAVLAGDCFHALAFELLADERTHEDPFVRSSELVLELAKASGIDGMGGGQAMDLAEEGETLDLGAITRLQQLKTGALIEYAVEAAVVMGRVAPEGRARYRGYARNVGLAFQIADDLLDHEGDEEAAGKKLQKDAEAGKATFVSLLGADRARQQAQMLVTQAIEHLHGHEEADLLRAIARYAVERDR

Sasta_0002, 312 aa, WP_029940286

MSEADERARLVQAALESISAGSKSFRFASQLFDQOTRERSWLLYSWCRACDDVTDGQTLGHDAERVDDPAARLAFLLKAKTAEAFAGQPTGLVPFDALRVVAAECAIPQAVAGDHLAGFERDAGGWRPTTTDDLLSYCYQVAGAVGVMMAHVMGVPPPEDEDTLNRAADLGIAFQLANIARDIVDDAKVGRVYLPAEWLAAEGLAGADLADPAHRPALARLAHRLADMADAYRRSARVGAARLPFRSRWAVLAASGIYGEIATRAAALGPRAWDERITTSKAEKAALVMEAFWEALWVRPAPRDGLWTRPAHA

Sasta_0003, 193 aa, WP_037504354

MSHNLKRIAVYCGSAPGNPAYADAVALIRAMAGQGIDLVIYGGGKLGMLLIADTMLEAGGEVHGVIPALVDKEVAHLGLTELHQVAGMHERKAKMTELCDAFVCLPGGIGTLDLFEAWTWNALGYHAKPFCLLNNGFWNGLGTFMDHVAGAGFLSQARREQLLLAESPAEAINKLDEAARTATQGMVW

Sasta_0004, 487 aa, WP_029940287

MKSAIVIGAGFGGLALAIRLQASAGVKTIVEARDRPGGRAYVWEKDGHVFDAGPTVITDPDCLQRLWKLSGHDMSDELVELVPVKPFYRLSWPDGVVFDYTNDDAELKAAMDALNPDDWAGYQRFAYSAGVYNEGYVKLGTKAFESLGDMLKAAPALAKYQAWRSVYSIVSSFVKDEHLRQTLFSHTLLVGGNPMTCSSIYALIHKLRRGGVWFAGGTNKLIAGMVRQFERIGGTIRLGDVPTAILAENDRVTVGRTASGWSATADAVASNGDVVHSYGLIEGSDRGQQVRLKRRKRFSPGLFVLHFGLEGTSDLAHHTILFGPRYGGVNDIYKTGRLATDPSLYIHHPTITDPSMAPPGCSTFYALAPVPNAGKADVDWAVEGPKYQEVVLDTIAERLIPDVRQIRITIFHYTPADFSADLAHLGSAFSLPEVWLQSAWFRTHNRDDKLRNLYFVGAGTHPGAGIPGVVGSAAEATAGLMLA

Sasta_0005, 380 aa, WP_029940288

MMKRADLVIVGGGLAGLALALRRRRPDLRLLLVPEGPSIGGNHLWSFFESDVAPADRWLTDPLIRHRWPDYEVRFPAHQRHAEAYQTIIESEALDEAVRKALSAAEIVRAEATDLGPTHVTLATGERIEAKAVLDARGKAEGDLGLGWQKFLGQLLTIPOGHGLTRPIVMDATVDQHDGYRFVYCLPFSPTLFEVDYYSYSDGPELDHDLRLDRIGDYAAAQGWQVADRSREEHGALPVVIGGDFDRLWPAADHVARAGARGGFHPLTSYSLPDAVRFAIWLADKATFDARLGAATRARGRRHWRSGAFYRLLTALLFHAAEPGQRYLVLERFYRLSGPLIGRFYAGMSTGYDKARVLAGKPPVFFRALRVLRDSL

Sasta_0006, 176 aa, WP_029940285

MSWPAGLALFVSTVLLMEGFAYVLHRFVMHSRLGWNWHESHHRARTGWFERNLDYAVVFALPSILLIWWGLNGGWGDWATWMGAGVAFYGVIIYFGFHDVIVHGRLPHRIVPRSTYFKRIVQAHKLHHAVESRDGAVSFGFLYAPPVERLKQALQASREAQLRRARGGSTARHEERA

Sasta_0008, 151 aa, WP_051676451

*MTKRF*TLAATAGLLLSAPVLMQAQAPQAGADLQVHLTGLRSAGVVLHCLSARPKFLDCKADKTALARTVPARAAGRLDLGPVRPGTYALLIVHDENQNGKLDMTLGIPIREGFGFSNNPAMRPRPRWEEIRFTMPATPTTQQIRVRYVL

Sasta_0009, 258 aa, WP_029942176

MAERRRPAYMAPMLSDAQRRRQAMIGLGLAAAITAAFVALHVWSVFFLPLEGAGWWLALPIVAVQTWLSVGLFIVAHDAMHGSLAPGRPATNLFWRGLTLLLYAGFWLDRLSPKHFDHRRHVGTERDPDFSVDPTRFWPWYYAFMRRYFGLREYLVNLALVLAIVLVLKPAPLGNLLLFWALPSILSSIQLFYFGTYLPHRHEDAPFADQHNARSNDFFVWLSLLTCFHFGYHREHLSPGTPWWQLPRRRRELALPA

Altererythrobacter ishigakiensis NBRC 107699 locus tags and sequences

Altish_0001, 301 aa, WP_067599965

MAAAIDHSFLKDGLSRVQREVDVDFDALLPVPQDSRARLVEAMRYAAIGGGKRVRLPLLLVSTAEELFGVNR
DSALRAACAVEAIHVYSLIHDDLPCMDDDALRHGKPTLHLEFDEATAVLAGDSLHALAFEILSDTATSE
PFTRSELILTLGKASGKDGMMAGGQMMDDMAEEEDYDLHTITRLQQLKTGALLAASVEMGAIMGHVQPEGR
AHLRAYARDIGLAFQIADDLLDVEGDADKAGKALRKDEGQGKQTFVTLMGAGKAREQARALVDQAIGHLS
VHGDEAKMLAAVARYIVERDR

Altish_0002, 315 aa, WP_067596482

MKRDELVEASRLAIKQGSKSFSASRLFDRTTRERVWMLYAWCRRCDDLTDAQDMGGELGDQSDIARVE
HLRQLTANALAGEETGDIADFAGQVAREVGLTMEADAVIAGFQLDAEDWRPRTEADLMRYCWHVAGAV
GVMMQVMGVKREDDTLDLRACDLGLAFQLANIARDVVEDDAAAARCYLPAEWLVEEDIEPGQHTKPHRR
ELADMAARLVRRMEVHAASSRIGAARLPFRCRWAILSAARIYTAIGREVLAKGPAAWDARVYTSKTQKIG
HAMAAFVEALVNRPPKLDDEPRWRRSDLAENAASA

Altish_0003, 193 aa, WP_067602122

MNRLAVYCGSATPDDPRYIELATEVGRALALREIGVVYGGGRLGLMGAVARGALDAGGEVIGVIPEALFN
SEVANHDCTKLYTVSGMHERKQRFDTLSDGFITIPGGVGTMDDELWEAMSWAQLGYHSPVGLLNTMGFFD
HLIAFNEKMAEVGFVRPAHQNILIHDTTVGGLIDKMAAYEPHTPIFRMKAEDL

Altish_0004, 520 aa, WP_067602119

MNAATPVTPPVRAPIQVDPKAYEGKTACVIGSGFGGMALAIRLQASAGITTTVIEARDKPGGRAYFWEKDG
FTFDGGPTVVTDPPCLKELWEITGHDMAEDVELMKVMPFYRLNWPDGTFNFDYSNDEDSLNSEIAKLDPGD
VAGYAEFLKYSEGVRHREGYLKLGTVPFDFKSMKKAAPALVKKQAWRSVYSMVSSYVKSEKLREALSFHT
LLVGGNPMNTSAIYALIHKLEKDGCVWVTRGGTNRLIAGMVRHFERLGGTMRVGDVAVTQVHTIGNKASEV
ETESGWKQRFDAVASNADIMHSYRDLLSGSKRGKDYAKSLSRKSFSPGLFVVHFGLEGTWPGIPHHMILF
GPRYKGLLDDIYKQGVLPQDFSIYLHHPVTVDPSMAPKKGKSTFYALVPVAHMGKLAVDWDEVGPVLEKRI
LDEIGRRLIPDIHDIRVTKFSYAPKNFAHDLNAHLGSAFSLPEVPLWQSAYLRGHNRRDDVIDNFYLVGAGT
HPGAGIPGVVGSAKATAGLMLLEDLAIKTTA

Altish_0005, 402 aa, WP_067596481

MGGRETDAIVGGGLAGGLIALALHRAQPDLRFQLIEAGETLGGHHRWSWFDSDLGHEGRALLSPFTYVA
WDQGYEVAFPPQYRRPMRAGYRSLASADFDAALREILPERTLSLQRVCEL DANGVTLQSGERINARTVID
CRSVLPSQHLLGGWQIFAGRHFYDRPHGITQPMIMDASIDQHAPAGNRSAYRFMYVPLAEDELFFEDT
YYDEARALDEDLLSRRIDQYAADMGFADGVEIGHERGILPVIITGGNFKAYRASIEIPGVAMAGARGGFSH
PLTSYTMPPIAVENALAIASHAHLGPELLAHFVRGRANEHWRKTAIYRALGRMLFKATEPERRVNIQRFY
QLPEELIERFYACHTTLPDQVRILSGKPPVSI PRAIRALASKGKPLLMEKTE

Altish_0006, 166 aa, WP_067596417

MSWWEIIAIVLAALLGMEFFAWYAHKYIMHGWGWEWHRDHHEPHDNTLEKNDLFAVVFGTINAAMYIYGA
LYSDALWWFALGITLYGVIYTLVHDGLVHQRYFRWVPKKGYAKRLVQAHKLHHTIGKEGGVSFGFVFAR
DPVKLKAELKEQRKAGIAKVREADLS

Altish_0007, 240 aa, WP_067596721

MVFGMLFAAVSMTLIVALRYLISSGMFAWITVRLRPGQHDGLKQPQIRREIKWSLASAAIYGVPAGLAFWS
WRHLGWTHMYLDPLDYPVWYMPLSALIYLFADHTWFWYTHRAMHWPVLFRLAHSVHNSRPPTAWAAMSF
HPIEALSGAIVIPALVLLVPIHMGILGLVLAVMTIMGVTNHMGWEMFPRWLVHSRFGKVVITASHHERHH
EEYQCNYGLYFRFWDKVCGTDRGLTTRFAQ

Altish_0008, 129 aa, WP_067602229

*MNEPPTGTVTVTIT*EMRSSEGVVRACMTTNREIFPRCRKDPASHRTVVTASENMTLTFTGVKPGNYAIAL
LHDENDNGKADRVLGMPKEGFGFSRDAKVQMAPPKFDDAVFEFDGSAEELTIRMRYML

Altish_0009, 255 aa, WP_067596418

MPYRASTVSHPMKLPSPNSQRLGLVLAALIAGSWLGIHFYAMFVFELESWQAWPQVLLMATLQCWLSVGV
FIVCHDAMHGSLAPGWQRVNSALGAILLFLYAGFAWRKIRDAHFAHKKHTGKDGDPDFDTANPTHFWAWY
WTFKRYFGWQSLLYVHMVVGIIYLFVFGIPIFMQIFLLYGAPALLSSLQLFYFGTYRPHRHLGESFADGHN
ARSDNFSTLASLASCFFHGYHLEHHRPDPVPPWALPGARRAGVAA