

Supplementary data for

Core Metabolic Features and Hot Origin of Bathyarchaeota

Xiaoyuan Feng ^a, Yinzhaoh Wang ^a, Rahul Zubin ^a, Fengping Wang ^{a,b,*}

^a State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

^b State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

* Corresponding author.

E-mail address: fengpingw@sjtu.edu.cn (F. Wang)

Table S1

General characteristics of all bathyarchaeotal MAGs collected up to date.

Lineage	16S	Bin	Gonomic length	GC (%)	Completeness	Contamination	Gene number	Used for metabolic analysis	Genome quality
Bathy-22	Bathy-22	JdFR-11	1 725 700	51.9	100.00	0.93	1864	Yes	High
Bathy-6	—	AD8-1	1 583 813	32.0	95.79	4.21	1730	Yes	High
Bathy-13	—	UBA233	1 574 115	40.9	94.86	2.80	1662	Yes	High
Bathy-21	Bathy-21	B24	1 731 088	51.1	94.86	1.87	1576	Yes	High
Bathy-15	—	NP77	1 996 065	48.4	93.93	3.31	1864	Yes	High
Bathy-8	Bathy-8	BA2	1 455 657	44.2	93.77	3.74	1800	Yes	High
Bathy-3	Bathy-3	BA1	1 930 879	47.1	91.59	2.80	2468	Yes	High
Bathy-15	Bathy-15	M10_bin139	1 527 575	58.4	90.42	1.87	1605	Yes	High
Bathy-23	Bathy-23	B26-2	1 890 021	51.9	90.19	3.81	1630	Yes	High
Bathy-22	Bathy-22	JdFR-10	1 817 739	51.8	97.09	6.47	1972	Yes	Middle
Bathy-17	Bathy-17	13_1_40CM_3_53_5	2 459 104	52.9	90.65	9.42	2419	Yes	Middle
Bathy-22	Bathy-22	M10_bin185	1 310 867	33.4	89.49	0.00	1462	Yes	Middle
Bathy-1	—	M8_bin1642	1 094 533	40.8	89.25	0.93	1171	Yes	Middle
Bathy-23	—	B26-1	1 347 585	52.7	89.25	6.39	1185	Yes	Middle
Bathy-17	—	13_1_40CM_3_52_17	1 775 293	51.8	88.32	0.93	1852	Yes	Middle
Bathy-23	—	M10_bin26	1 331 265	51.9	88.32	2.80	1390	Yes	Middle
Bathy-17	—	13_1_40CM_2_52_14	1 621 432	51.8	86.92	0.93	1728	Yes	Middle
Bathy-8	Bathy-8	UBA589	1 523 390	44.2	86.14	4.67	1855	Yes	Middle
Bathy-23	-	M8_bin162	1 251 171	46.8	85.12	4.67	1223	Yes	Middle
Bathy-17	Bathy-17	RBG_16_48_13	845 981	48.1	84.11	1.44	949	Yes	Middle
Bathy-15	-	RBG_13_52_12	2 028 665	51.8	83.98	2.91	2508	Yes	Middle
Bathy-23	Bathy-23	M10_bin1901	1 202 259	44.8	83.96	4.67	1188	Yes	Middle
Bathy-17	—	UBA185	1 797 576	54.1	81.40	5.83	2256	Yes	Middle
Bathy-17	—	RBG_13_38_9	1 682 179	37.7	80.53	2.80	2014	Yes	Middle
Bathy-23	Bathy-23	M10_bin1904	1 257 418	42.7	77.57	1.87	1224	Yes	Middle
Bathy-13	—	JdFR-04	1 067 774	41.9	77.28	1.25	1297	Yes	Middle
Bathy-17	—	13_1_40CM_3_52_10	1 324 891	51.9	75.40	0.97	1408	Yes	Middle
Bathy-15	—	RBG_13_60_20	1 208 628	61.2	75.15	0.93	1460	Yes	Middle
Bathy-1	—	M10_bin208	907 528	41.1	74.37	1.71	1030	Yes	Middle
Bathy-15	—	RBG_16_57_9	1 492 913	59.5	73.89	4.67	1914	Yes	Middle
Bathy-22	—	B25	818 523	50.5	72.90	1.32	857	Yes	Middle
Bathy-17	—	13_1_20CM_2_53_14	1 541 478	53.3	72.43	2.80	1598	Yes	Middle
Bathy-1	Bathy-1	JdFR-07	907 518	38.7	64.49	0.74	1022	Yes	Middle
Bathy-23	—	ex4484_40	1 136 090	51.6	63.03	1.94	1267	Yes	Middle
Bathy-17	Bathy-17	SMTZ-80	1 510 195	34.5	61.99	7.79	1527	Yes	Middle

Lineage	16S	Bin	Gonomic length	GC (%)	Completeness	Contamination	Gene number	Used for metabolic analysis	Genome quality
Bathy-15	—	B23	717 588	61.8	60.90	0.00	747	Yes	Middle
Bathy-23	—	B63	859 079	53.9	59.81	0.47	823	Yes	Middle
Bathy-22	—	ex4484_218	628 054	33.1	57.61	0.00	737	Yes	Middle
Bathy-13	Bathy-13	ex4484_231	880 221	46.7	56.96	7.12	969	Yes	Middle
Bathy-1	—	JdFR-08	917 816	38.6	52.34	2.88	1002	Yes	Middle
Bathy-6	Bathy-17	SG8-32-3	668 322	44.0	50.54	0.00	731	Yes	Middle
Bathy-8	—	SG8-32-1	1 078 462	36.7	50.34	9.22	1270	Yes	Middle
Bathy-23	—	M8_bin26	503 563	53.6	46.96	2.34	548	Yes	Low
Bathy-17	—	13_1_40CM_3_52_4	756 044	52.1	44.98	0.97	791	Yes	Low
Bathy-13	—	JdFR-05	433 324	42.6	41.12	0.00	551	Yes	Low
Bathy-23	—	M10_bin182	597 004	44.3	37.70	0.00	604	Yes	Low
Bathy-1	Bathy-1	JdFR-06	735 190	38.3	37.38	2.80	874	Yes	Low
Bathy-21	—	M10_bin242	497 722	48.8	33.35	0.97	543	Yes	Low
Bathy-15	—	M10_bin1432	474 510	56.3	32.97	0.00	605	Yes	Low
Bathy-1	—	JdFR-09	497 927	39.0	28.65	0.00	596	Yes	Low
Bathy-15	Bathy-15	AB-539-E09	529 169	43.3	22.33	0.00	670	Yes	Low
unclassified	—	ex4484_205	1 282 076	42.1	91.75	7.99	1448	No	Middle
unclassified	—	M10_bin241	1 132 007	47.1	88.35	2.02	1214	No	Middle
unclassified	—	M8_bin163	816 136	43.4	79.44	0.93	897	No	Middle
unclassified	—	ex4484_135	976 871	58.8	70.09	0.93	1083	No	Middle
unclassified	Bathy-1	M10_bin214	597 499	37.2	65.87	0.00	695	No	Middle
Bathy-23	Bathy-23	M10_bin1903	1 879 918	43.1	92.60	15.26	1833	No	Low
Bathy-5	Bathy-5	RBG_13_46_16b	1 336 882	46.3	89.72	15.80	1628	No	Low
Bathy-15	Bathy-15	DG-45	1 423 247	62.1	73.36	13.01	1527	No	Low
unclassified	—	JdFR-01	200 428	38.2	14.49	0.93	251	No	Low

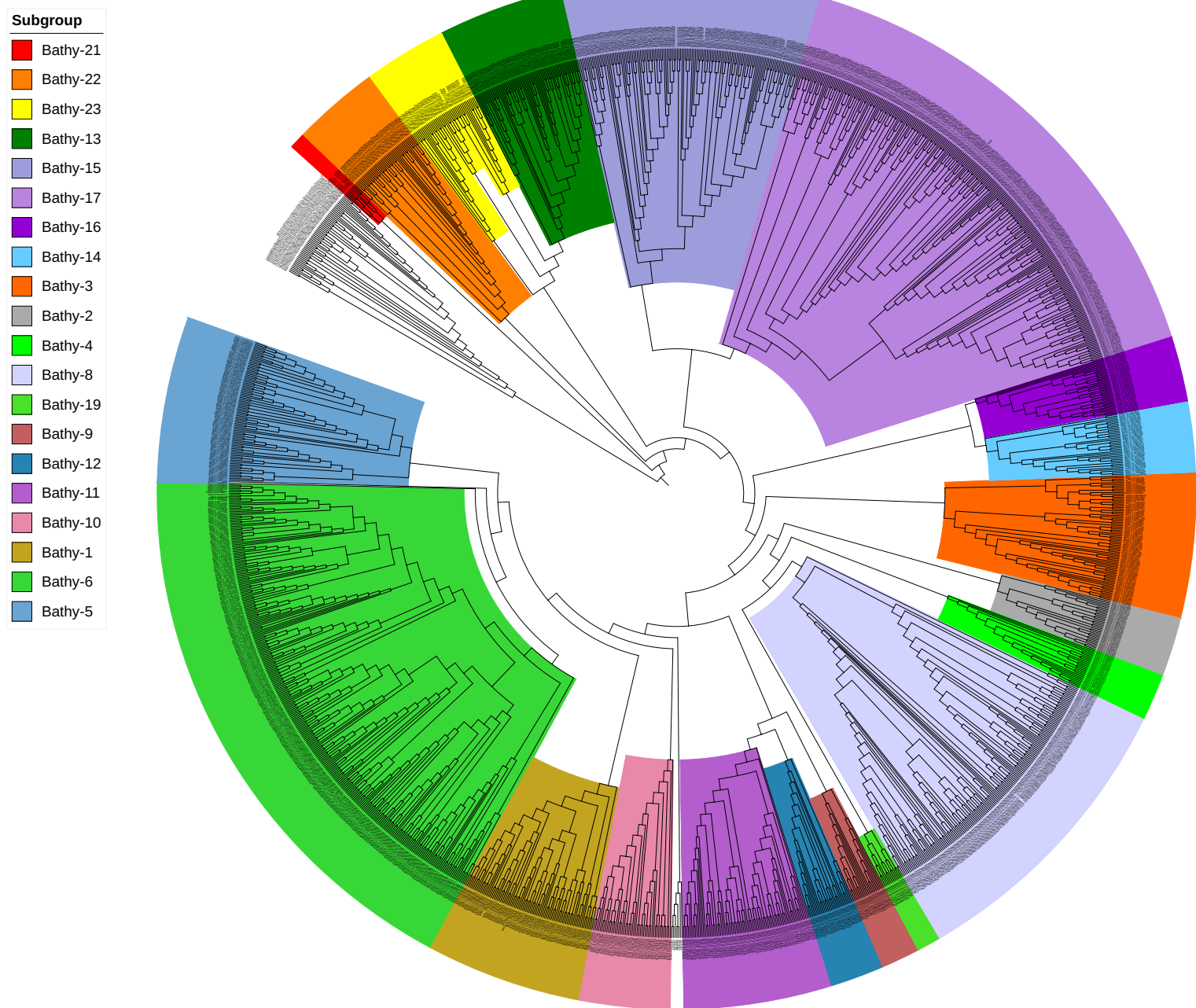


Fig. S1. The full version of the uncollapsed phylogeny tree based on the 16S rRNA gene sequence. Branch length was ignored for conciseness.

Table S2

Sequence information for three novel bathyarchaeotal subgroups (Bathy-21/22/23). Clone ID, sampling country and isolation source information was collected from SILVA database.

Accession ID	Subgroup	Clone ID	Sampling country	Isolation source
AB213074	21	FnvA63	—	PCR-derived sequence from hydrothermal fluid
AB611625	21	TKTMmva-A42	Japan: Okinawa, off Taketomi Island	Microbial mat developed at the main vent site
AY861949	21	OPPD013	USA: Yellowstone National Park	Obsidian Pool prime community
EU635903	21	SSW_L4_H01	USA: Nevada, Mud Hot Springs	Sediment from hot spring; temperature 87 °C
EU924232	21	LHC4_L2_A09	USA: California, Little Hot Creek	Hot spring sediment at temperature of 80 °C
LUCB01000007	21	B24	Mexico: Guaymas Basin, Gulf of California	—
U63340	21	pSL22	Obsidian Pool hot spring, Yellowstone National Park	Perspectives on archaeal diversity, thermophily, and monophyly from environmental rRNA sequences
AB177007	22	ODP1227A17.20	—	PCR-derived sequence from seafloor sediment at the Peru margin (ODP Leg 201)
AB213070	22	FnvA51	—	PCR-derived sequence from hydrothermal fluid
AB213075	22	FnvA65	—	PCR-derived sequence from hydrothermal fluid
AB213100	22	Papm3A46	—	PCR-derived sequence from hydrothermal fluid
AB213101	22	Papm3A49	—	PCR-derived sequence from hydrothermal fluid
AB329759	22	YS16Ac22	—	PCR-derived sequence from hydrothermal deposits
AB329777	22	YS16Ac76	—	PCR-derived sequence from hydrothermal deposits
AB644796	22	C9001C_A13_4_C020	Japan	PCR derived sequence from seafloor sediments off Shimokita Peninsula
DQ190083	22	ARCHEUKDBF30	—	Snake River Plain Aquifer
FJ902719	22	LP30MA65	Mexico: Tamaulipas	Biomat 30 m deep in cenote La Palita
FN666140	22	P10a5RS55	Tunisia	Hot spring
GQ410805	22	TWP1-90	China	Marine sediments from the tropical western Pacific Ocean
GQ410893	22	TWP4-42	China	Marine sediments from the tropical western Pacific Ocean
GQ410906	22	TWP5-6	China	Marine sediments from the tropical western Pacific Ocean
GQ410949	22	TWP7-54	China	Marine sediments from the tropical western Pacific Ocean
GQ926390	22	MD3044F-84	China	Deep-sea sediment
GQ927584	22	MD3059P-54	China	Marine sediment from the tropical western Pacific Ocean
GQ927596	22	MD3059P-74	China	Marine sediment from the tropical western Pacific Ocean
GQ927599	22	MD3059P-86	China	Marine sediment from the tropical western Pacific Ocean
GQ927606	22	MD3059K-17	China	Marine sediment from the tropical western Pacific Ocean
GQ927614	22	MD3059K-38	China	Marine sediment from the tropical western Pacific Ocean

Accession ID	Subgroup	Clone ID	Sampling country	Isolation source
GQ927663	22	MD3059T-41	China	Marine sediment from the tropical western Pacific Ocean
GQ927688	22	MD3059S-29	China	Marine sediment from the tropical western Pacific Ocean
HM187495	22	HDBA_SISW527	USA	Saturated zone of the Hanford Site 300 Area subsurface
HM244046	22	HWA1012-3-26	China	Upper sediment from Honghu Lake
HM244065	22	HWA1012-2-14	China	Upper sediment from Honghu Lake
HM244187	22	HWA5257-3-37	China	Lower sediment from Honghu Lake
HM244190	22	HWA5257-3-85	China	Lower sediment from Honghu Lake
HM244235	22	HWA5257-1-28	China	Lower sediment from Honghu Lake
HQ611193	22	HQ350.48	China	Marine sediments from HQ08-48PC cruise HY4-2008-1 at the station in the northern South China Sea, 1474 m
JQ257976	22	K-UA-90	India: Karwar, Arabian Sea coast	Tropical marine sediment, around 20 m water depth
JQ258743	22	R-UA-96	India: Ratnagiri, Arabian Sea coast	Tropical marine sediment, around 20 m water depth
JQ327887	22	MD3047-41	—	Marine sediments from western Pacific Ocean
JQ989677	22	ORIII-1368-04-G_S161-163_043A04	China: Taiwan	Subseafloor sediment at the KP-9
JQ989712	22	ORIII-1368-11-G_S188-190_114A09	China: Taiwan	Subseafloor sediment at the KP-9
JX000679	22	ORI-961-19-1-P_S008-010_19-11A18	China: Taiwan	Subseafloor sediment at the Deformation Front
JX196217	22	IMCUGW13A3	China	Groundwater samples from Inner Mongolia
JX507247	22	4559-4-C2-20	Mexico	Sediment
JX507256	22	4559-4-C2-110	Mexico	Sediment
KC003524	22	H11_4a_2	USA: New Jersey	IODP expedition 313, 8.94 mbsf
KC003804	22	H10_7a_2	USA: New Jersey	Sediment; IODP expedition 313, 12.99 mbsf
KJ881610	22	Ai2_16SA_75	—	Upper convective layer of Atlantis II Deep, Red Sea
KJ881627	22	Ai2_16SA_92	—	Upper convective layer of Atlantis II Deep, Red Sea
KP090994	22	ARC_OTU_2	—	Hydrothermal seep
KP210129	22	SRBZUarch62	—	Thermal soil
AB213061	23	Fapm1aA37	—	PCR-derived sequence from hydrothermal fluid
AB260058	23	CORK.A40	USA: Eastern Flank of Juan de Fuca Ridge	Black rust formation on a borehole observatory, CORK 1026B
AB301882	23	pIta-vmat-1	Japan: Okinawa, off Taketomi Island	Microbial mat at a shallow submarine hot spring
AB329810	23	YS18As31	—	PCR-derived sequence from hydrothermal deposits

Accession ID	Subgroup	Clone ID	Sampling country	Isolation source
AB611658	23	TKTMmvm-A11	Japan: Okinawa, off Taketomi Island	Microbial mat developed at the main vent site
AF419653	23	AT_R021	—	—
DQ300318	23	VulcPIw.10	Italy: Vulcano Island	Geothermal well
DQ300322	23	VulcPIw.115	Italy: Vulcano Island	Geothermal well
DQ300326	23	VulcPIw.156	Italy: Vulcano Island	Geothermal well
DQ300334	23	VulcPIw.36	Italy: Vulcano Island	Geothermal well
DQ300343	23	VulcPIw.88	Italy: Vulcano Island	Geothermal well
EU924228	23	LHC4_L2_G04	USA: California, Little Hot Creek	Hot spring sediment at temperature of 80 °C
FN554034	23	A257-6	Atlantic Ocean: Mid-Atlantic Ridge, Logatchev hydrothermal vent field	Logatchev hydrothermal vent field, site F, watercolumn depth = 3000 m
FN554041	23	A257-14	Atlantic Ocean: Mid-Atlantic Ridge, Logatchev hydrothermal vent field	Logatchev hydrothermal vent field, site F, watercolumn depth = 3000 m
FN554050	23	A257-26	Atlantic Ocean: Mid-Atlantic Ridge, Logatchev hydrothermal vent field	Logatchev hydrothermal vent field, site F, watercolumn depth = 3000 m
FN554068	23	A257-45	Atlantic Ocean: Mid-Atlantic Ridge, Logatchev hydrothermal vent field	Logatchev hydrothermal vent field, site F, watercolumn depth = 3000 m
GQ848425	23	TVG11AR01	—	Deep sea hydrothermal vent sediment
GQ848429	23	TVG11AR06	—	Deep sea hydrothermal vent sediment
GQ848432	23	TVG11AR09	—	Deep sea hydrothermal vent sediment
GQ848441	23	TVG11AR18	—	Deep sea hydrothermal vent sediment
GQ848445	23	TVG11AR22	—	Deep sea hydrothermal vent sediment
HQ317335	23	g-53	Indonesia: Padang Cermin	Hot spring
JF935161	23	PNG_Kap4_A90	Papua New Guinea: Ambitle Island, Kapkai	Hot spring sediment
JX507244	23	4559-4-C2-95	Mexico	Sediment
JX507246	23	4559-4-C2-107	Mexico	Sediment
JX507250	23	4559-4-C2-11	Mexico	Sediment
JX507252	23	4559-4-C2-106	Mexico	Sediment
JX507258	23	4559-4-C2-110	Mexico	Sediment
JX507263	23	4559-4-C2-111	Mexico	Sediment
KM221269	23	W8A-11	—	Hot spring
KM585423	23	630arch34	China: Tengchong National Geothermal Park	Thermal soil
KM585435	23	666arch-21	China: Tengchong National Geothermal Park	Thermal soil

Accession ID	Subgroup	Clone ID	Sampling country	Isolation source
KP091006	23	ARC_OTU_30	—	Hydrothermal seep
KP091007	23	ARC_OTU_31	—	Hydrothermal seep
KP091048	23	ARC_OTU_74	—	Hydrothermal seep
KP091051	23	ARC_OTU_77	—	Hydrothermal seep
KP091069	23	ARC_OTU_93	—	Hydrothermal seep
KP784731	23	LX3-A101	—	Hot spring
U63339	23	pSL17	Obsidian Pool hot spring, Yellowstone National Park	Perspectives on archaeal diversity, thermophily, and monophyly from environmental rRNA sequences

IODP: International Ocean Discovery Program; mbsf: meter below seafloor; PCR: polymerase chain reaction.

Table S3

Genes annotated as extracellular peptidases in MEROPS database.

	Subgroup									
	1	3	6	8	13	15	17	21	22	23
A22B						+			+	
C01A	+		+	+	+	+	+	+		
C11						+	+			
C14B			+	+			+			
C26	+			+	+	+	+	+		+
C39				+	+		+			+
G01							+			
I02		+		+			+		+	
I08						+	+			
I39							+			
I42	+					+	+	+	+	+
I87						+				+
M04						+	+			
M06			+			+	+		+	
M24A						+	+			
M38						+	+			
M43A	+					+	+			
M48B			+			+	+	+	+	+
M50B						+	+			
N10						+	+			+
P01			+				+	+		+
S01A							+	+		+
S01C			+			+				+
S08A	+	+	+	+	+	+	+	+	+	+
S09X			+	+			+			+
S12	+		+			+				
S26B					+					+
S33				+	+	+	+		+	+
S45	+				+					
S49C		+		+		+	+		+	
S53							+			
S54			+				+			+
S71							+			
T03					+	+				+

+/: pathway near complete.

+/: presence of gene family.

Table S4

Analyzed metabolic potential of different bathyarchaeotal subgroups in Fig. 2 and Table 2. Abbrev was used same as Table2.

Pathway	Gene name	Abbrev.	Subgroup									
			1	3	6	8	13	15	17	21	22	23
Cellulose degradation	Endoglucanase	-	+			+	+	+	+		+	+
Glycogen degradation	1,4-alpha-glucan branching enzyme	<i>glg</i>	+		+				+			
	Glycogen phosphorylase	<i>glg</i>	+		+		+	+	+			+
Galactan degradation (Leloir pathway)	Beta-galactosidase	<i>lac</i>	+							+		+
	Galactokinase	<i>galK</i>	+					+		+		+
	UDP-glucose–hexose-1-phosphate uridylyltransferase	<i>galT</i>	+	+	+	+	+		+	+		+
	UDP-glucose 4-epimerase	<i>galE</i>	+	+	+	+	+	+	+	+	+	+
Mannan degradation	Alpha-mannosidase	<i>amy</i>	+					+		+		+
Lipid degradation	Long-chain acyl-CoA synthetase	<i>fadD</i>	+	+	+	+	+	+	+		+	+
	Acyl-CoA dehydrogenase	<i>acd</i>	+	+			+	+			+	
	Acetyl-CoA C-acetyltransferase	<i>atoB</i>	+	+	+	+	+	+	+	+	+	+
Benzoate degradation	Acylphosphatase	<i>acyP</i>	+		+	+	+	+	+	+	+	+
	Amidase	<i>amiE</i>	+					+	+	+		+
Organic sulfur compound degradation	Steryl-sulfatase	<i>sts</i>										+
	Choline-sulfatase	<i>betC</i>	+							+		+
	Iduronate 2-sulfatase	<i>ids</i>										+
	Extracellular sulfatase	<i>sulF</i>										+
Embden–Meyerhoff glycolysis pathway	Glucose-6-phosphate isomerase	<i>pgi</i>	+	+	+	+	+	+	+	+	+	+
	6-phosphofructokinase	<i>pfk</i>	+		+	+	+		+			+
	Triosephosphate isomerase	<i>tpi</i>	+	+	+	+	+	+	+	+	+	+
	Glyceraldehyde-3-phosphate dehydrogenase	<i>gap</i>	+	+	+		+	+	+	+	+	+
	3-phosphoglycerate kinase	<i>pgk</i>	+	+	+	+	+	+	+	+	+	+
	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	<i>apg</i>	+	+	+	+	+	+	+	+	+	+
	Enolase	<i>eno</i>	+	+	+	+	+	+	+	+	+	+
	Pyruvate kinase	<i>pyk</i>	+				+	+	+		+	+
Pyruvate → acetyl-CoA	Pyruvate ferredoxin oxidoreductase alpha subunit	<i>por</i>	+	+		+	+	+	+	+	+	+

Pathway	Gene name	Abbrev.	Subgroup									
			1	3	6	8	13	15	17	21	22	23
Archaeal WL pathway	2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase subunit alpha	<i>kor</i>	+	+		+	+	+	+		+	+
	Formylmethanofuran dehydrogenase subunit A	<i>fwd</i>	+	+	+	+	+	+	+	+	+	+
	Formylmethanofuran tetrahydromethanopterin <i>N</i> -formyltransferase	<i>frt</i>	+	+	+	+	+	+	+	+	+	+
	Methenyltetrahydromethanopterin cyclohydrolase	<i>mch</i>	+	+	+	+	+	+	+	+	+	+
	Methylenetetrahydromethanopterin dehydrogenase	<i>mtd</i>	+		+		+	+	+	+	+	+
	5,10-methylenetetrahydromethanopterin reductase	<i>mer</i>	+	+		+	+	+	+	+	+	+
Bacterial WL pathway	Acetyl-CoA decarbonylase/synthase alpha subunit	<i>cdh</i>	+	+	+	+	+	+	+	+	+	+
	Anaerobic carbon-monoxide dehydrogenase catalytic subunit	<i>coo</i>	+	+	+	+	+	+	+			
	Formate dehydrogenase alpha subunit	<i>fdh</i>										
	Formate–tetrahydrofolate ligase	<i>fhs</i>	+					+	+			
	Methylenetetrahydrofolate dehydrogenase (NADP ⁺)/methenyltetrahydrofolate cyclohydrolase	<i>folD</i>	+					+	+	+	+	
	Methylenetetrahydrofolate dehydrogenase (NADP ⁺)/methenyltetrahydrofolate cyclohydrolase	<i>folD</i>	+					+	+	+	+	
Methanogenesis	Methylenetetrahydrofolate reductase (NADPH)	<i>metF</i>	+					+	+			+
	Acetyl-CoA decarbonylase/synthase complex subunit gamma	<i>cdh</i>	+	+	+	+	+	+	+	+	+	+
	Tetrahydromethanopterin <i>S</i> -methyltransferase subunit A	<i>mtr</i>	+						+		+	
	Methyl-coenzyme M reductase alpha subunit	<i>mcr</i>		+		+						
	Citrate synthase	<i>cs</i>	+		+	+		+	+		+	
	Aconitate hydratase	<i>acn</i>	+		+	+		+	+			
TCA cycle	Isocitrate dehydrogenase	<i>icd</i>	+		+			+	+	+	+	+
	2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase subunit alpha	<i>kor</i>	+	+		+	+	+	+		+	+
	Succinyl-CoA synthetase	<i>suc</i>	+	+		+	+	+	+		+	
	Succinate dehydrogenase, flavoprotein subunit	<i>sdh</i>	+	+		+	+	+	+	+	+	+
	Malate dehydrogenase	<i>mdh</i>	+		+	+		+	+	+	+	+
	ADP-forming acetyl-CoA synthetase	<i>acd</i>	+					+	+	+	+	+
Acetate	Acetate kinase	<i>ack</i>	+							+	+	+
	Phosphate acetyltransferase	<i>pta</i>	+							+	+	+
	Sulfhydrogenase subunit gamma (sulfur reductase)	<i>hydG</i>	+							+		+
Nitrite reduction	Ferredoxin-nitrite reductase	<i>nirA</i>						+				+

Pathway	Gene name	Abbrev.	Subgroup									
			1	3	6	8	13	15	17	21	22	23
Eucaryotic chaperons	Nitrite reductase (NADH) large subunit	<i>nirB</i>				+					+	
	Nitrite reductase (NADH) small subunit	<i>nirD</i>							+			
	Heat shock protein HtpX/Zn-dependent protease with chaperone function	<i>Htpx</i>	+	+	+	+	+	+	+	+	+	+
	Small heat shock protein HSP16.5/molecular chaperone IbpA, HSP20 family	<i>sHSP</i>	+	+	+	+	+	+	+	+	+	+
Bacterial DnaK-DnaJ-GrpE chaperones	Molecular chaperone DnaK	<i>DnaK</i>	+	+	+	+		+	+		+	
	Chaperone protein DnaJ	<i>DnaJ</i>	+	+	+	+		+	+		+	
Thermoprotectant formation	Heat shock protein GrpE	<i>GrpE</i>	+	+	+			+	+		+	
	Myo-inositol-1-phosphate synthase	-	+	+	+	+	+	+	+	+	+	+
Reverse gyrase	Reverse gyrase	-								+	+	+
Transporter	Sulfate transport system permease protein	<i>cysU</i>		+		+						+
	Tungstate transport system substrate-binding protein	<i>tup</i>	+		+	+	+	+	+	+	+	+
	Spermidine/putrescine transport system ATP-binding protein	<i>pot</i>	+			+	+	+			+	+
	Phosphate transport system ATP-binding protein	<i>pst</i>	+				+	+	+	+	+	+
	Branched-chain amino acid transport system ATP-binding protein	<i>liv</i>	+	+	+	+	+	+	+	+	+	+
	ABC-type oligopeptide transport system, ATPase component	<i>ABC.PE</i>	+	+	+	+	+	+	+	+		+
	Multiple sugar transport system permease protein	<i>ABC.MS</i>	+						+		+	+
	Iron complex transport system ATP-binding protein	<i>ABC.FEV</i>	+		+	+	+		+	+	+	+
	Ferrous iron transport protein B	<i>feo</i>	+		+		+	+	+	+	+	+
	Zinc transport system substrate-binding protein	<i>znu</i>	+		+	+	+	+	+	+	+	+
	Mg ²⁺ -importing ATPase	<i>mgt</i>	+					+	+	+		+
	Magnesium transporter	<i>cor</i>	+					+	+			
	K ⁺ -stimulated pyrophosphate-energized sodium pump	<i>hpp</i>	+	+	+	+	+	+	+		+	+
	Trk system potassium uptake protein	<i>trk</i>	+						+	+	+	+
	Cu ²⁺ -exporting ATPase	<i>cop</i>	+				+	+		+		+

TCA: tricarboxylic acid; UDP: uridine diphosphate; WL: Wood–Ljungdahl.

+: presence of gene.